

Distributional Growth Accounting: Education and the Reduction of Global Poverty, 1980-2019

Amory Gethin*

September 8, 2024

Abstract

This article quantifies the role played by education in the decline of global poverty. I propose tools for identifying the contribution of schooling to economic growth by income group, integrating imperfect substitution between skill groups into macroeconomic growth decomposition. I bring this “distributional growth accounting” framework to the data by exploiting a new microdatabase representative of nearly all of the world’s population, new estimates of the private returns to schooling, and historical income distribution statistics. Under conservative assumptions, education accounts for 50% of global economic growth, 70% of income gains among the world’s poorest 20%, and 40% of extreme poverty reduction since 1980. It also explains over 50% of the reduction in gender labor income inequality. Standard growth accounting methods substantially underestimate economic benefits of schooling, especially for women and the global poor.

*World Bank Development Research Group. I am particularly grateful to Thomas Piketty, Vincent Pons, Emmanuel Saez, and Gabriel Zucman for their continuous guidance. I thank Daron Acemoglu, Pierre Bachas, Denis Cogneau, Eva Davoine, Jérémy Do Nascimento Miguel, Sebastian Hager, Eric Hanushek, Hilary Hoynes, Anders Jensen, Pete Klenow, Karen Macours, Giulia La Mattina, Ségal Le Guern Herry, Wouter Leenders, David Levine, Antoine Levy, Hongbin Li, Rosanne Logeart, Nicolas Longuet Marx, Clément Mazet-Sonilhac, Edward Miguel, Mathilde Muñoz, Emi Nakamura, Jon Piqueras, Tom Raster, Anne-Sophie Robilliard, Benjamin Schoefer, Andrei Shleifer, Jón Steinsson, Christopher Walters, Duncan Webb, Danny Yagan, and Céline Zipfel for useful comments. This paper benefited from fruitful discussions with seminar participants at UC Berkeley, Paris School of Economics, and Bordeaux School of Economics. I also thank the International Labour Organization for providing me with data tabulations from their collection of harmonized household surveys; I am especially thankful to David Bescond for his help and assistance throughout. Juan Ferrer provided excellent research assistance. This research benefited from funding from the Paris School of Economics (Grant ANR-17-EURE-001).

1. Introduction

The past decades witnessed dramatic improvements in access to basic public services among the global poor. These improvements were evident in progress made on indicators as diverse as school enrollment, healthcare coverage, and access to drinkable water ([United Nations, 2023](#)).

Despite such progress, how useful these policies have been at generating income growth for the global poor remains an open question. Education, in particular, has expanded massively, yet its contribution to global poverty reduction is anything but clear. The quality of education is low in many developing countries, casting doubt on the ability of schooling to increase productivity. Many competing factors could also have played a more important role than schooling, from globalization and technical progress to other government policies. The economic effects of education depend on who benefits from educational expansion, the associated returns to schooling, and general equilibrium effects. Because of difficulties in quantifying these channels, we lack estimates of how large the benefits from schooling have been for the global poor. Addressing this gap is of fundamental importance for policy, in a world where the vast majority of children from low-income households are enrolled in public schools.

This article constructs the first estimates of the aggregate and distributional effects of worldwide educational expansion since 1980. Leveraging a unique microdatabase on individual incomes, a “distributional growth accounting” model of education and the wage structure, and complementary quasi-experimental evidence, I quantify the contribution of schooling to growth for different groups of the world distribution of income. I find that education explains a considerable fraction of global economic growth, extreme poverty reduction, and gender inequality reduction since 1980, putting education policies at the forefront of economic progress. My analysis also contributes to our broader understanding of the general equilibrium effects of schooling and their role in the reduction of poverty and inequality, which I find to be substantial.

The starting point is a new microdatabase representative of 95% of the world’s population. It is based on household surveys fielded in 150 countries around 2019, which I collect from data repositories and country-specific sources. Each survey provides comprehensive information on individual incomes, education, age, gender, and other socioeconomic variables. I harmonize each survey to construct a single dataset covering these variables for a total of 10 million surveyed individuals. To the best of my knowledge, this represents the first microdatabase on the contemporary structure of global poverty and inequality.

Starting from this new dataset, I develop methods to estimate what the world distribution of income would look like if there had been no improvement in schooling since 1980. I combine standard growth accounting tools with a 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. It 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 three steps. First, I downgrade education levels in each of my 150 surveys until matching the distribution of educational attainment that prevailed in 1980. Second, I reduce individual incomes accordingly, using new country-specific measures of returns to primary, secondary, and tertiary education. Third, I estimate general equilibrium effects: the supply of skilled workers would be lower, and hence their relative wage higher, if education had not improved. Finally, I compare the resulting counterfactual to the actual evolution of incomes, yielding an estimate of the contribution of education to growth for different income groups within each country.

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% since 1980 (see [Figure 1](#)). They also explain 40% of the reduction in the share of the world’s population living in extreme poverty. On the whole, education has been a major source of global inequality reduction, driving a large share of income gains in low-income countries and among the poorest individuals in all world regions. This progressive nature of global educational expansion has intensified over time and has been most pronounced among recent cohorts. Given the predominant role that governments have had in providing education and other basic services to low-income households, this puts public policies at the center of the fall of global poverty. Combining measures of direct government redistribution from a companion paper ([Gethin, 2024](#)) with indirect investment benefits from education estimated in this paper brings the total contribution of public policies to global extreme poverty reduction to at least 50%.

A natural concern with this analysis is that the private return to schooling may differ from the aggregate effect of education on growth. This framework may also miss important channels through which education affects the income distribution. To address these potential limitations, I validate my method with new quasi-experimental evidence, extending existing work on three large-scale schooling initiatives in India

([Khanna, 2023](#)), Indonesia ([Duflo, 2001](#)), and the United States ([Acemoglu and Angrist, 2000](#)). Combining rich microdata on the distribution of income with differential exposure to each program across subnational regions, I document two main facts. First, educational expansion induced by these policies had large causal effects on aggregate regional incomes, comparable to individual returns found in the same contexts. Second, the three policies disproportionately benefited low-income earners, strongly reducing inequality. This represents some of the first causal evidence that expanding education delivers particularly high returns for the very poor. Crucially, the distributional growth accounting framework reproduces these two findings with a remarkable degree of accuracy—if anything, it underestimates the benefits of schooling—, indicating that it provides a good methodological foundation to study the role of education in shaping income growth. More generally, my estimates should be considered conservative. The estimation relies on standard Mincerian returns, which are typically lower than causal estimates derived from natural experiments, as evidenced by the labor economics literature ([Card, 2001](#); [Deming, 2022](#)) and my own meta-analysis of 15 papers. It assumes that education only affects labor income, abstracting from potential positive effects on capital income. It also ignores human capital externalities, on which there is now significant empirical evidence.¹ All in all, my findings are governed by two 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%.

Methodologically, accounting for distributional effects within countries appears to be essential to quantify the role of schooling in global poverty reduction. Standard growth accounting, as in [Barro and Lee \(2015\)](#), typically relies on combining cross-country data on average years of schooling with a uniform 10% return to derive the same counterfactual as in this paper. This approach could, after all, provide a good approximation. I find that it does not. Education is found to account for 23% of growth among the world’s poorest 20%, compared to 71% in my benchmark specification. In other words, the canonical exercise underestimates the contribution of education to poverty reduction by a factor of three.² One reason is that cross-country data cannot accurately measure poverty: the poorest individuals in the world do not all live in the poorest countries. More importantly, the classic approach fails to account for key dimensions

¹See in particular [Acemoglu and Angrist \(2000\)](#), [Moretti \(2004\)](#), [Ciccone and Peri \(2006\)](#), [Iranzo and Peri \(2009\)](#), [Gennaioli et al. \(2013\)](#), [Wantchekon, Klačnja, and Novta \(2015\)](#), [Chauvin et al. \(2017\)](#), [Glaeser and Lu \(2018\)](#), [Guo, Roys, and Seshadri \(2018\)](#), and [Queiró \(2022\)](#).

²The standard approach also underestimates the contribution of education to aggregate growth by about 40%, for two main reasons. First, it relies on labor income shares that exclude mixed income entirely, while I provide evidence that mixed income is affected by schooling just as much as wages. Second, I account for imperfect substitution, which increases the contribution of schooling by magnifying losses from not expanding education.

of educational expansion, such as labor income shares being greater at the bottom of the distribution and general equilibrium effects redistributing schooling gains from high-skilled to low-skilled workers. The contribution of this paper lies in the use of new microdata and tools allowing me to identify and quantify each of these channels, with major implications for the role of education in reducing extreme poverty.

One should also stress that the large contribution of schooling does not preclude that other factors may have played a significant role too. In the model, the return to schooling increases with the skill bias of technology. Had technology not improved, this return would be lower than the one observed today. This interaction effect between schooling and technology can be estimated. Indeed, my benchmark estimate relies on reducing incomes observed in 2019, fixing technology at its 2019 value. An alternative exercise would be to increase incomes observed in 1980, fixing technology at its 1980 value. The difference between these two measures turns out to identify the role of skill-biased technical change in *amplifying* the economic benefits of schooling. Quantifying this difference would require microdata on the world distribution of income in 1980, which unfortunately are not available. Nonetheless, I was able to systematically compile and harmonize additional surveys covering 33 countries around 2000, allowing me to isolate this channel for this particular sample. I find that skill-biased technical change accounts for 20-30% of the benefits of schooling in the average country. Technology has thus played an important role in making schooling profitable, although education would still have yielded large income gains even in the absence of technological progress.

I also 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. I account 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.

A large literature in labor economics uses the canonical labor supply-and-demand framework to relate changes in the wage distribution to educational expansion.³ Concurrently, a considerable literature in

³This framework has been extensively applied to the study of wage inequality in the United States (Katz and Murphy, 1992; Goldin and Katz, 2007; 2008; Acemoglu and Autor, 2011; Autor, Goldin, and Katz, 2020; Hershbein, Kearney, and Pardue, 2020; Deming, 2023). A growing literature extends this analysis to low- and middle-income countries (Fernández and Messina, 2018; Vu and Vu-Thanh, 2022; Khanna, 2023).

macroeconomics investigates the contribution of human capital to economic development .⁴ 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 integrate them into a unified “distributional growth accounting” framework, enabling me to derive the first estimate of the role played by education in the reduction of global poverty and gender inequality.

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 and cross-country income convergence ([Sala-i-Martin, 2006](#); [Chen and Ravallion, 2010](#); [Pinkovskiy and Sala-i-Martin, 2016](#); [Hammar and Waldenström, 2020](#)), the emergence of a new “global median class” ([Lakner and Milanovic, 2016](#)), skyrocketing top income inequality ([Chancel and Piketty, 2021](#)), and decreasing gender inequality ([Kleven and Landais, 2017](#); [Neef and Robilliard, 2021](#)). Amongst the numerous factors shaping these dynamics, I isolate the contribution of one of them: education. I 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 extensive empirical evidence on the economic effects of education. A vast literature documents positive impacts of education on individual earnings ([Card, 2001](#); [Psacharopoulos and Patrinos, 2004](#); [Deming, 2022](#)). A more limited number of studies examine the general equilibrium effects of education policies (e.g., [Duflo, 2004](#); [Che and Zhang, 2018](#); [Porzio, Rossi, and Santangelo, 2022](#); [Fujimoto, Lagakos, and VanVuren, 2023](#); [Hsiao, 2023](#); [Khanna, 2023](#)). I contribute to this literature by extending previous work on India ([Khanna, 2023](#)), Indonesia ([Duflo, 2001](#)), and the United States ([Acemoglu and Angrist, 2000](#)). Unlike these studies, which were primarily focused on individual returns to schooling or human capital spillovers, I provide new evidence on the effects of these policies on growth and inequality. I also draw heavily on existing studies to calibrate the parameters guiding my results, such as elasticities of substitution between skill groups and the differential economic effects of education by gender. In doing so, this article is an attempt at leveraging the best microeconomic evidence to draw conclusions on the aggregate effects of education on global poverty and gender inequality. This approach aligns with the growing literature highlighting the need to bridge the micro-macro gap in the study of

⁴The recent literature has focused more heavily on explaining cross-country differences in development: see for instance [Hall and Jones \(1999\)](#), [Hanushek and Kimko \(2000\)](#), [Caselli and Coleman \(2006\)](#), [Hsieh and Klenow \(2010\)](#), [Gennaioli et al. \(2013\)](#), [Jones \(2014\)](#), [Hendricks and Schoellman \(2018\)](#), and [Rossi \(2020\)](#). Growth accounting decompositions have been used considerably since [Solow \(1957\)](#), although worldwide perspectives are more recent (e.g., [Mankiw, Romer, and Weil, 1992](#); [Barro and Lee, 2015](#); [Collin and Weil, 2020](#)).

economic development (Buera, Kaboski, and Townsend, 2023).

The rest of the paper is organized as follows. Section 2 outlines the conceptual framework. Section 3 presents the data and methodology. Section 4 describes the main results on the role of education in shaping the world distribution of income. 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 skill levels:

$$Y = F(K, L) = F(K, L_0, L_1, \dots, L_m) \quad (1)$$

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:

$$w_s = w_s(L) \quad (2)$$

Consider a group p (e.g., the poorest 20%) receiving income from labor and capital. Its average income is:

$$y^p = y_K^p + \sum_s w_s(L) L_s^p \quad (3)$$

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 skills $\tilde{L} = (\tilde{L}_1, \dots, \tilde{L}_m)$:

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

This requires characterizing four sets of parameters: the initial distributions of labor and capital incomes, initial wages $w_s(L)$, counterfactual skill supplies \tilde{L}_s^p , and counterfactual wages $w_s(\tilde{L})$.

2.1.2. Model Specification

Throughout the paper, following the labor economics literature (e.g., Katz and Murphy, 1992; Goldin and Katz, 2007; Khanna, 2023), I consider variants of the CES production function with two skill groups:

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

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:

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

The relative wage of skilled workers depends on their relative supply, as well as on the skill bias of technology. A higher elasticity of substitution implies greater substitutability between skill groups and a lower sensitivity of wages to relative supplies.

2.2. Aggregate Effects of Schooling

I now characterize the aggregate effects of schooling 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 this change $\Delta L_H = \tilde{L}_H - L_H$ and wage

changes as Δw_L and Δw_H . The effect of skill upgrading on output is:

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

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 schooling on output 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 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).

How do these four effects play out in practice, and where does the true aggregate effect of schooling lie? One option would be to assume that each additional skilled worker increases output by the initial return $r(L)$. Alternatively, one could use the final return observed after educational expansion $r(\tilde{L})$. The true return turns out to lie in-between (Caselli and Ciccone, 2013). Newly skilled workers end up with lower benefits than they might have hoped, but part of this loss benefits, on net, the unskilled:

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

A graphical illustration providing the main intuition can be found 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 $r(L^{2019})$, corresponding to the derivative of the production function after educational expansion. Alternatively, one could use marginal gains from schooling observed if there had been no educational progress, $r(L^{1980})$, corresponding to counterfactual (initial) returns. With imperfect substitution, the log of output is concave in schooling: the true effect of education lies in-between these two estimates.

Appendix B.3 provides a longer theoretical discussion, as well as results from a simple simulation. For

parametrizations similar to those found in the data, the true return 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:

$$\Delta Y^R - \Delta Y^P = \underbrace{\left(L_H^R - L_H^P \right) \Delta w_H}_{\substack{\text{Differential Supply} \\ \text{Effects on Always Skilled} \\ \leq 0}} + \underbrace{\left(L_L^R - L_L^P \right) \Delta w_L}_{\substack{\text{Differential Supply} \\ \text{Effects on Unskilled} \\ \leq 0}} + \underbrace{\left(\Delta L_H^R - \Delta L_H^P \right) r(\tilde{L})}_{\substack{\text{Differential Selection} \\ \text{Into Education}}} \quad (9)$$

The first two terms reveal that supply effects reduce inequality. The intuition is simple. Increasing the share of skilled workers puts downward pressure on their wage. Because R has a greater fraction of skilled workers than P , this negative effect will be greater for them (differential supply effects on always skilled). Conversely, supply effects benefit unskilled workers, who are more concentrated at the bottom of the distribution (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 , this could increase inequality. 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 bring this framework to the data. The objective is to estimate the contribution of education to the real earnings growth of groups p , who differ in their relative supply of workers belonging to skill groups s . This problem can be reformulated as the following counterfactual question: what would have been the distribution of income in 2019, had there been no progress in education since 1980? I estimate this counterfactual in five steps, starting from microdata reporting information on the joint distribution of labor income and education. Appendix B.1 provides more details.

1) Downgrade Education Levels The first step is to reduce educational attainment to match its distribution in 1980: absent educational progress, the distribution of skill groups would be $L^{1980} = (L_1^{1980}, \dots, L_m^{1980})$ instead of $L^{2019} = (L_1^{2019}, \dots, L_m^{2019})$. I implement this “first stage” in the microdata by sampling individuals and downgrading their education levels until reaching the counterfactual (as in, e.g., [Hershbein, Kearney, and Pardue, 2020](#)). For instance, if the share of tertiary-educated workers expanded from 10% to 15%, I sample 5% of tertiary-educated workers and downgrade their education level from tertiary to secondary.

2) Reduce Wages Using Returns to Schooling The second step is to reduce the earnings of workers by the return to schooling: absent educational progress, incomes would be lower. As discussed in section 2.2, this return lies in-between initial and final returns. Assuming that the final return $r(L^{2019})$ and the elasticity of substitution σ are known, the true return can be directly recovered from the model (see Appendix B.1). I then reduce the earnings of treated workers by this return. For instance, workers downgraded from tertiary education to secondary education see their earnings reduced by the return to tertiary education.

3) Adjust Relative Wages to Account for Supply Effects The third step is to account for supply effects: absent educational progress, the wage of skilled workers would be higher, while the wage of unskilled workers would be lower. Changes in relative wages are given by:

$$\Delta \log \left(\frac{w_H}{w_L} \right) = -\frac{1}{\sigma} \Delta \log \left(\frac{L_H}{L_L} \right) \quad (10)$$

Assuming that σ is known, one can directly adjust wages in the microdata. For instance, for an increase in the relative supply of skilled workers of 1 log point and an elasticity of 6, the wage gap between skilled and unskilled is reduced by 0.17 log points.

4) Derivation of Total Income Steps 1 to 3 allow us to estimate how the distribution of labor income would look like today if there had been no educational expansion since 1980. 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.

5) Growth Accounting The final step is to calculate the share of growth accounted for by education. It is equal to the gap between counterfactual and actual growth rates, expressed as a fraction of actual growth from 1980 to 2019:

$$\text{Success}^P = \frac{g^P - \tilde{g}^P}{g^P} \quad (11)$$

With $g^P = \frac{y_{2019}^P - y_{1980}^P}{y_{1980}^P}$ and $\tilde{g}^P = \frac{\tilde{y}_{2019}^P - y_{1980}^P}{y_{1980}^P}$. If growth had been exactly the same absent educational progress, then $\tilde{g}^P = g^P$ and hence $\text{Success}^P = 0\%$. On the contrary, if there would have been no growth absent educational progress, then $\tilde{g}^P = 0$ and $\text{Success}^P = 100\%$: education can explain all of economic growth.

3. Data and Methodology

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 how much lower wages would 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 using new evidence from 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 (see appendix E for more details).

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 statistical institutes, ILOSTAT have harmonized over 1,300 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. Surveys are nationally representative. 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 download additional household surveys for 55 countries. 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.

Figure 2 maps the 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 GDP⁵ To the best of my knowledge, this represents the first microdatabase 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. It records the share of individuals with no schooling, primary education, secondary education, and higher education in 146 countries since 1950. A number of countries are covered by the survey microdata but missing from the Barro-Lee database. For these, I construct my own estimates, using census data, cohort-level trends observed in the labor force surveys, or other sources (see appendix F).

Figure 3 plots the distribution of educational attainment from 1980 to 2019 in the world as a whole. There has been a dramatic expansion of secondary education, from about 25% to 60%. This 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

⁵In nearly all countries, the survey was fielded after 2015, ensuring that the database is broadly representative of global inequality in 2015-2019. Appendix Figure E1 provides a map of the corresponding survey years in each country.

expanded significantly, from less than 5% in 1980 to over 10% in 2019.⁶

3.3. Returns to Schooling

The second step is to reduce individual incomes by the returns to schooling. To do so, one needs estimates of returns to schooling observed in 2019 in each country—corresponding, as discussed in section 2.4, to final returns $r(L^{2019})$. There are two options, each 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:

$$\ln y_{ic} = \alpha_c + \beta_c^{pri} D_{ic}^{pri} + \beta_c^{sec} D_{ic}^{sec} + \beta_c^{ter} D_{ic}^{ter} + X_{ic} \beta_c + \varepsilon_{ic} \quad (12)$$

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. Appendix G shows that my results are largely insensitive to alternative methodological choices.

The main advantage of OLS returns lies in their coverage and comparability: they can be estimated for all 150 countries using a unified methodology. The potential disadvantages are twofold. First, OLS returns may suffer from omitted variable bias, although this bias has been found to be typically small (Card, 1999; Gunderson and Oreopoulos, 2020; Deming, 2022). Second, our parameter of interest is not the *average* return. It is the return to schooling for those who benefited from education since 1980, which can differ significantly from the average return in the presence of heterogeneity (e.g., Heckman, Humphries, and Veramendi, 2018). In particular, improved access to schooling has predominantly benefited children from

⁶Improvements in schooling have coincided with significant declines in overall educational attainment inequalities, both between and within countries: see Appendix Figure F1, which provides a Theil decomposition of global educational attainment inequality following the methodology first proposed by Morrisson and Murtin (2013). Appendix Figure F2 charts average years of schooling of the working-age population by world region since 1980.

lower socioeconomic backgrounds (Gethin, 2024), a population that is often found to have higher returns. Using OLS returns might then lead to underestimating benefits of schooling during our period of interest. Figure 4 plots returns to schooling by world region. Returns 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 return to tertiary education is 22%. Returns to primary education are extremely low in India, barely reaching 3%. Overall, there are large variations in returns both across countries and across education levels, in line with previous evidence covering a more limited number of countries (Psacharopoulos, 1994; Psacharopoulos and Patrinos, 2004; Banerjee and Duflo, 2005; Caselli, Ponticelli, and Rossi, 2014).⁷

3.3.2. IV Returns

An alternative option is to use instrumental variable estimates of returns to schooling resulting from large-scale education programs. The estimation of these returns generally relies on comparing cohorts and 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).

For the purpose of this paper, I have assembled a collection of IV estimates of the returns to schooling, selected 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. Fourth, I select one estimate per country and education level.

Figure 5 compares the resulting OLS and IV estimates. OLS and IV returns are highly correlated, and IV returns are almost always higher than OLS returns or not significantly different (in line with previous estimates compiled by Card, 1999). The gap ranges from close to 0% in Nigeria to almost 250% in China. The average gap is in the order of 80% (see Appendix Figure G7). This provides reassuring evidence

⁷Appendix 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 pooled regressions by level and decomposes secondary education into upper and lower secondary.

that OLS estimates provide a good approximation, and if anything are likely to underestimate the true contribution of schooling to economic growth during my period of interest.

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).

Production Function Specification Until now, I worked with a CES production function with two skill groups for illustrative purposes, yet the data cover 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. My benchmark specification directly follows the labor economics literature (Goldin and Katz, 2007; Fernández and Messina, 2018). Output is produced by combining tertiary-educated and non-tertiary-educated workers:

$$L = \left(A_{t\bar{e}r} L_{t\bar{e}r}^{\frac{\sigma_1-1}{\sigma_1}} + A_{ter} L_{ter}^{\frac{\sigma_1-1}{\sigma_1}} \right)^{\frac{\sigma_1}{\sigma_1-1}} \quad (13)$$

With $L_{t\bar{e}r} = 1 - L_{ter}$. Non-tertiary-educated workers are subdivided into workers with secondary education L_{sec} and below-secondary education $L_{s\bar{e}c}$:

$$L_{t\bar{e}r} = \left(A_{s\bar{e}c} L_{s\bar{e}c}^{\frac{\sigma_2-1}{\sigma_2}} + A_{sec} L_{sec}^{\frac{\sigma_2-1}{\sigma_2}} \right)^{\frac{\sigma_2}{\sigma_2-1}} \quad (14)$$

With $L_{s\bar{e}c} = 1 - L_{sec} - L_{ter}$. Finally, workers with below-secondary education include those with primary education L_{pri} and those with no schooling L_{non} :

$$L_{s\bar{e}c} = \left(A_{non} L_{non}^{\frac{\sigma_3-1}{\sigma_3}} + A_{pri} L_{pri}^{\frac{\sigma_3-1}{\sigma_3}} \right)^{\frac{\sigma_3}{\sigma_3-1}} \quad (15)$$

These equations yield three formulas for the returns to primary, secondary, and tertiary education, which directly map onto the categories reported in the Barro-Lee database and the returns plotted in Figure 4.

True returns and supply effects can then be calculated separately for each of the three nests.⁸

Elasticities of Substitution To close the model, the only parameters that need to be calibrated are the elasticities of substitution between skill groups. I specify a benchmark elasticity of $\sigma_1 = \sigma_2 = \sigma_3 = 6$, following the recent macroeconomics literature (Bils, Kaymak, and Wu, 2022; Hendricks and Schoellman, 2023). This corresponds to a long-run elasticity, accounting for the fact that firms tend to adjust their technological mix as a response to educational progress. Appendix B.2 provides a longer discussion.

3.5. Global Labor and Capital Income Inequality Data

The fourth and fifth steps of the methodology consist in moving from labor to total income, and comparing counterfactual to actual growth rates. This requires data on the world distribution of income. I use the World Inequality Database (Blanchet et al., 2021), which reports average per-capita labor and capital income by percentile in all countries in the world from 1980 to 2019 (see Appendix H). All components of net national income (GDP, minus consumption of fixed capital, plus net foreign income) are allocated to individuals, so that income distributions are consistent with macroeconomic growth rates.

3.6. Validation from Three Natural Experiments

Despite the popularity that the canonical labor-demand-and-supply framework has encountered in the literature, the ability of my estimates to capture the true effect of education on growth and inequality may naturally be questioned. To assess the validity of my framework, I turn to causal evidence from three natural experiments. I outline the main results and relegate further details to appendix C.

3.6.1. Contexts, Data, and Methodology

I investigate the aggregate and distributional effects of three large-scale policies expanding access to education: 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

⁸Appendix 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.

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 overall effects on aggregate growth and inequality.

Combining data from existing studies and additional sources, I exploit these natural experiments to estimate the causal effect of educational expansion on aggregate economic growth and its distribution across subnational regions. I run variants of the following specification:

$$\ln y_{rt}^i = \gamma_0^i + \gamma_1^i S_{rt} + X_{rt} \beta + \delta_r + \delta_t + \varepsilon_{rt} \quad (16)$$

$$S_{rt} = \alpha_0 + \alpha_1 Z_{rt} + \eta_{rt} \quad (17)$$

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} , average years of schooling of the working-age population in region r . X_{rt} 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 real average 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 ([Acemoglu and Angrist, 2000](#); [Guo, Roys, and Seshadri, 2018](#)).

After estimating the aggregate and distributional effects of each program, I compare them to those predicted by the model. In India, for instance, I simulate the effect of increasing average 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.

3.6.2. Main Results

Figure 6 plots the main results, comparing estimated and simulated effects of education on the average income of each quintile. All three policies led to large reductions in 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 in this district, compared to a null effect on that 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 comparable to individual returns found by articles studying the same policies (Duflo, 2001; Clay, Lingwall, and Stephens, 2021; Khanna, 2023).

The model performs remarkably well when it comes to reproducing results from these natural experiments. In all three cases, simulations predict significantly higher returns to schooling expansion at the bottom of the distribution. If anything, the model slightly underestimates benefits for low-income earners, in particular in the United States. Together, these results provide reassuring evidence that the methodology developed in this paper delivers a good approximation of 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 Since 1980

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% of Global 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 per-capita income growth since 1980.

2) Education Explains 70% of Growth for the Global Poorest 20% The contribution of schooling to growth varies significantly by global income group. For the world's poorest 50%, growth has been higher, exceeding 150%. Yet, benefits 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 55% to 75% of growth for the global bottom 90%. This share is highest for the bottom 20% (71%) and middle 40% of the income distribution (74%), two groups that have witnessed lower growth and large gains from 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).⁹ 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%.

⁹The interested reader will find additional figures describing these trends in more detail in the appendix. In particular, Appendix Figure H1 provides a Theil decomposition of global income inequality since 1980, while Figure H2 charts the average share of pretax income received by the richest 10% by world region. Figures H3, H4, H5, and H6 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.

Taking the ratio of the contribution of education to actual growth rates yields Figure 7, which represents the share of growth explained by education by global income percentile. This share ranges from 55% to 90% 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% of Extreme Poverty Reduction Another indicator that has received considerable attention is the share of the world's individuals living in extreme poverty. A difficulty 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 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 linear in growth rates).

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 to measure absolute poverty. 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. Private returns to schooling can thus account for 42% of global extreme poverty reduction. The corresponding figures are 32% at \$3.65 per day and 44% at \$6.85 per day. As a robustness check, I reproduce this analysis using World Bank income distribution data (see Appendix Table A3). Education explains 40% of global poverty reduction at \$2.15 per day, 56% at \$3.65 per day, and 73% at \$6.85 per day, similar to or higher than the figures obtained with the WID data.

4.2. Decomposing Global Schooling Gains

Why does education explain a higher share of growth at the bottom and upper-middle of the world distribution of income? Are the results primarily driven by differences in growth within countries or between countries? This section addresses these questions by decomposing schooling gains along several dimensions.

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 and bottom 20% real income growth. Two results stand out.

1) Education Explains More Growth in Low- and High-Income Countries Education explains the totality of growth in low-income countries and about 60% in high-income countries. Sub-Saharan Africa and Latin America are the two world regions where education explains the highest share of growth, mainly because these are the two regions that have witnessed the lowest growth. Indeed, education explains by construction over 100% of growth in countries where growth has been zero or negative, as in any growth accounting exercise.¹⁰ The interpretation is simply that other factors than education—such as civil wars and economic crises—have affected growth negatively, leading to a negative counterfactual growth rate.

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 growth has been about two times slower in India than in China, education can explain about the same share of income gains in the two countries. 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 50% of Growth For Low-Income Groups in All Regions In most regions, education explains more growth at the bottom of the distribution for two main reasons. First, because of rising inequality, actual growth has been significantly lower at the bottom of the distribution in many countries (in particular in Europe, Northern America, China, and India). Education thus explains more growth at the bottom of the distribution, simply because there is less growth to be explained. Second, 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.¹¹ Even in India, which displays extraordinarily low returns to basic education, private returns to schooling can explain about half of real income gains for the poorest 20%.

¹⁰See for instance Barro and Lee (2015), pages 139-140, who also find that human capital explains close to 100% of growth in Sub-Saharan Africa since 1960.

¹¹Again, it is important to mention that figures exceeding 100% should not be seen as extraordinary. In the United States, for instance, the real average income of the bottom 20% *declined* over the 1980-2019 period, which implies that education necessarily explains over 100% of growth for that group.

Combining these two facts enables a better understanding of the patterns presented in Figure 1. Economic growth has been 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 U.S. 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 or even negative economic growth in Sub-Saharan Africa, but also because of rapidly rising inequality in middle-income countries such as 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.

From 1980 to 2019, global inequality more or less stagnated (first row of Table 5). Absent educational progress, it would have instead risen from 1.06 to 1.33 (second row). Educational progress has thus contributed to strongly reducing global inequality in the past decades.

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 both low- and high-income countries. The effect of education on the overall dispersion of cross-country average incomes has been close to zero.

While schooling has not reduced between-country inequality, it has unambiguously decreased inequality within countries. The Theil index of within-country inequality would have risen about two times faster absent educational expansion. In other words, education has sufficiently mitigated the rise of within-country

inequality to keep overall global inequality constant.¹²

4.2.3. Distributional Growth Accounting by Time Period

The data also allow estimating results by time period. Figure 8 plots the contribution of education to the annual income growth of each global percentile in 1980-2019 and 2000-2019. Consistently with previous results, schooling gains have been greatest at the middle and bottom of the global income distribution since 1980, generating a 1.2-1.4% annual increase in earnings for these groups. In the recent period, growth effects of education have declined overall, but not at the bottom of the world distribution of income. The benefits of worldwide improvements in schooling have thus increasingly accrued to the global poor.

Appendix Table A1 presents a distributional growth accounting decomposition of the world distribution of income for the 2000-2019 period. Education explains 27% of global average economic growth since 2000, about two times lower than the corresponding figure since 1980. This is mainly due to the acceleration of growth 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. For the global bottom 20%, however, education still accounts for about 50% of income gains, because actual growth for this group has not been particularly high, while gains from schooling have remained about the same. The same pattern holds over the 2010-2019 period.¹³

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 between cohorts born in the 1910s-1960s and 1960s-1990s. The effect of education on economic growth thus results from two separate mechanisms: the replacement of old cohorts by new 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:

¹²Appendix 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.

¹³Appendix 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.

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 general equilibrium effects as in the rest of the paper.

Worldwide educational progress appears to have been particularly progressive when narrowing the focus to post-1980 generations.¹⁴ It explains about 12% of global economic growth since 1980, but 61% of income gains among the world's poorest 20% individuals and 37% among the global bottom 50%. 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; Gethin, 2024), 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 than 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. Most importantly, it allows comparing my results to those of 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.

¹⁴Appendix Table A7 presents the main results. Appendix Figures A4 and A5 plot gains from schooling and the share of growth explained by global percentile.

4.3.1. The Standard Growth Accounting Decomposition

I start by presenting results from a standard growth accounting decomposition. I follow the same methodology as [Barro and Lee \(2015\)](#), who exploit GDP and educational attainment data to estimate the fraction of global economic growth explained by human capital from 1960 to 2010. This decomposition only requires three ingredients: 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 (set at 10% per year). Counterfactual income absent educational progress is:

$$\tilde{y}^{2019} = v_L \frac{y^{2019}}{(1+r)^{\Delta S}} + v_K y^{2019} \quad (18)$$

With y^{2019} GDP per capita in 2019, v_L and v_K 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, 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 explains 33% of global average economic growth. The second column shows results for the global bottom 20%. Because this growth accounting decomposition relies only on cross-country data, the poorest 20% have to be defined as the poorest 20% countries (population-weighted). Schooling gains have been relatively small for these countries. As a result, education explains 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 in the Penn World Tables treat all mixed income as capital income. The implicit assumption is that only wages are affected by schooling; there is no return 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.¹⁵ 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.

¹⁵See [Appendix Table G2](#), which compares returns to schooling by level estimated with and without including mixed income in total personal income.

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 increases the contribution of schooling to average growth to 43%, and its contribution to bottom 20% growth to 35%.

4.3.3. Incorporating Within-Country Inequality

In a third step, I account for within-country inequality: the global poorest 20% correspond no more to the poorest 20% of countries. All other methodological ingredients stick to the standard growth accounting exercise. 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 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 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 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. 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 aggregate labor income share. 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 microdatabase covering education and incomes in 150 countries collected for the purpose of this paper. The main difference is that returns to schooling are now allowed to vary by country and education level. This 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 more strongly the contribution of education to global bottom 20% growth, which falls from 51% to 41%. The 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% (see Figure 4). Heterogeneity in returns to schooling thus has important effects on growth accounting estimates, implying lower gains from schooling for the global poor than what a homogeneous 10% return would suggest.

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). 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 Table 6, accounting for changes in relative wages raises the contribution of education among the world's poorest 20% from 41% to 52%. This large impact is consistent with the empirical literature documenting large effects of changes in the supply of skilled workers on inequality (e.g., Moretti, 2004; Goldin and Katz, 2007; Vu and Vu-Thanh, 2022; Khanna, 2023).

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 what returns to schooling observed in 2019 suggest. The true returns to schooling that should be used are thus higher than the returns observed in 2019 (see section 2.2).

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 52% to 71%, 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 75%.

4.3.8. 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 much lower schooling gains for the global poor than a usual 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

I now briefly discuss the sensitivity of my main findings to alternative specifications.

Alternative Elasticities of Substitution My benchmark specification relies on an elasticity of substitution of 6. Appendix Table A4 presents results of two alternative specifications, corresponding to long-run elasticities at the lower and upper ends of those found in the literature. The results are similar: education explains 53-61% of average growth, and 66-85% among the world's poorest 20%.

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 described in section 3.4 is standard and has been validated empirically (Goldin and Katz, 2007; Fernández and Messina, 2018). Yet, one may still be concerned that alternative specifications may yield different results. In Appendix Table A5, I consider a variant in which firms first pick between workers with below- and above-secondary education, and then choose between subcategories of workers within these two nests. The results are comparable: education

explains 53% of global growth and 65% among the world's 20%.

Alternative Patterns of Schooling Expansion Another step of the estimation has to do with identifying who benefits from schooling. In the main specification, I randomly sample individuals within age-gender cells and downgrade their education levels until matching those observed in 1980. While section 3.6 suggests that it provides a good approximation for reproducing results from three natural experiments, one might still be concerned about unobserved heterogeneity. For instance, if educational expansion mostly benefited poor children, one might underestimate benefits of schooling for low-income earners.

Extending the estimation beyond age-gender cells is not possible for the world as a whole given the lack of data. However, I investigate the implications of using more refined categories for India (1983-2019), South Africa (2002-2019), and the United States (1980-2019). Using historical survey data, I expand the analysis to cover heterogeneity in educational expansion by subnational region (India) and race (South Africa, United States). The results are presented in Appendix Figures A6, A7, and A8. Using more refined categories has limited effects. It marginally reduces gains for top-income earners in India and the U.S., while it slightly raises them for all groups in South Africa. I take this as reassuring evidence that my methodology provides a good approximation of actual distributional effects of schooling.

Capital Income Affected by Schooling In my benchmark specification, I assume that education has no effect on capital income at all. This is a very conservative assumption. 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 has large effects on entrepreneurial income (Gennaioli et al., 2013; Queiró, 2022) and innovation (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. In this scenario, education explains about two-thirds of global economic growth, 80% among the world's poorest 20% individuals, and 54% among the top 10%.

Education Quality A last source of concern is that the quality of education might have changed. Changes in educational attainment since 1980 might then not be comparable. 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 when there have been none. I investigate this concern at length in Appendix

D. Although data are scarce, available evidence suggests stagnating quality in most countries (Altinok, Angrist, and Patrinos, 2018; Angrist et al., 2021; Le Nestour, Moscoviz, and Sandefur, 2022). Even under conservative assumptions on a potential decline in quality, my main results 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 to the results, it is useful to distinguish the channels through which educational expansion can affect gender inequality. Consider an individual i with 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:

$$y_i = e_i(s_i) \cdot r_i^{s_i} \quad (19)$$

The gender gap is:

$$\Delta y = \ln y_f - \ln y_m \quad (20)$$

$$= \ln e_f(s_f) - \ln e_m(s_m) + (s_f \ln r_f - s_m \ln r_m) \quad (21)$$

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

$$\Delta y = \underbrace{\underbrace{(s_f - s_m) \ln r_f}_{\text{Differential Educational Expansion}} - \underbrace{s_m \ln \alpha}_{\text{Differential Returns to Schooling}}}_{\text{Intensive Margin}} + \underbrace{\ln e_f(s_f) - \ln e_m(s_m)}_{\text{Extensive Margin}} \quad (22)$$

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 relative propensity of women to be 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 used until now. 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 9 shows that with the exception of Sub-Saharan Africa, there has been a significant decline in gender schooling inequalities 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: working-age women are now slightly more educated than men. Convergence in educational attainment by gender can thus be expected to have acted as a significant driver of the reduction in gender income 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 and education level, as in equation 12. I then reduce earnings by gender-level-specific returns.

Figure 10 reproduces a well-known fact: returns to schooling are higher for women than for men (e.g.,

Montenegro and Patrinos, 2021).¹⁶ This gap can be large: 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:

$$e_{ic} = \alpha_c + \vartheta_c \text{Educ}_{ic} \cdot \text{Gender}_{ic} + \xi_c \text{Educ}_{ic} + \varphi_c \text{Gender}_{ic} + X_{ic} \beta_c + \varepsilon_{ic} \quad (23)$$

With e_{ic} a dummy for being economically active, Educ_{ic} years of schooling, Gender_{ic} a dummy taking one for women, 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. It is equal to 0.7 in the average country: increasing the average education of women by one year relative to that of men increases female labor force participation by 0.7 percentage points.

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 construct the first database on gender labor income inequality in nearly all countries in the world since 1991. The indicator of interest is the female labor income share, defined as:

$$y_L^f(s_f, s_m) = \frac{N_f \cdot e_f(s_f) \cdot r_f^{s_f}}{\sum_{i \in \{f, m\}} N_i \cdot e_i(s_i) \cdot r_i^{s_i}} \quad (24)$$

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.

¹⁶Appendix 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.

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 worldwide female labor income share since 1991 to its counterfactual evolution absent educational expansion.

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 2019.¹⁷ The second row reports the counterfactual female labor income share, 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 last two rows account for heterogeneous returns to schooling by gender and differential effects of education on labor force participation. This brings the contribution of schooling to gender inequality reduction to almost 80%.

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 the particular role that education has played in reducing income gaps within countries. The last column of Table 7 presents the share of gender inequality reduction explained in the average country (population-weighted).¹⁸ From this angle, differential educational expansion alone accounts for 50% of gender inequality reduction in the average country. Incorporating heterogeneous returns brings this share to 58%, while accounting for employment effects raises it to almost 60%.

Robustness I investigate the sensitivity of the results to alternative specifications in the appendix. First, one might be worried that results are driven by countries where gender inequality increased. In these countries, education explains over 100% of gender inequality reduction by construction. China stands out

¹⁷Appendix 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.

¹⁸To construct this indicator, I calculate the share of gender inequality reduction explained in each country. I then take the population-weighted average of this indicator over all 150 countries. 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%.

as an important source of concern, given its large population. Appendix Table [A13](#) reproduces Table 7 after excluding China: the results are similar. Appendix Table [A14](#) considers a more narrow restriction of the sample, excluding all 28 countries in which gender inequality increased, while still keeping countries in which education explains none of gender inequality reduction. The contribution of education in the average country declines to about 40%.

Second, it is useful to investigate how sensitive these results are to assumptions on the effect of education on female labor force participation. Appendix Table [A15](#) compares the benchmark estimates to alternative specifications of employment effects commensurable to causal estimates found in the existing literature (see Appendix Table [G5](#)). 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). In all specifications, the share of gender inequality reduction explained exceeds 90% in the world as a whole, and 65% in the average country.

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 Appendix Figure [A11](#) extends this analysis to the evolution of gender inequality in different regions of the world.¹⁹ 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 70% in the MENA region.

2) The Effect of Educational Expansion on Gender Inequality Has Increased Finally, I present results for different time periods in Appendix Table [A16](#). Income gains from schooling have been about 2 times greater for women since 1991 in the average country, compared to 2.5 since 2010. This rising convergence is visible in most regions, although it has been most pronounced in China and India. 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 pace as schooling gains.

¹⁹Appendix Figure [A12](#) plots income gains for women relative to men by region, while Appendix Figure [A13](#) plots annualized income gains from schooling for men and women in each country. Appendix Figure [A14](#) 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.

6. Discussion

This section presents a general discussion in three directions. Section 6.1 discusses complementarities between education and technology and provides an empirical analysis of the role of skill-biased technical change in magnifying the benefits of education. Section 6.2 investigates what might have been the various factors explaining educational progress since 1980 and implications for the results presented in this paper. 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. 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 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? The difference between these two estimates turns out to identify the role of other production factors in amplifying the growth effects of schooling.

6.1.1. 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 .²⁰ Output is $Y^t = Y(A^t, L^t)$. The main objective of this paper was to construct:

$$\tilde{Y}_{\text{backward}}^{2019} = Y(A^{2019}, L^{1980}) \quad (25)$$

²⁰I present a technological interpretation of A^t here, but it should be thought of as containing all production factors, such as skill-complementary capital, affecting the relative demand for skilled labor.

The share of growth explained by education is:

$$\text{Success}_{\text{backward}} = 1 - \frac{\frac{Y(A^{2019}, L^{1980})}{Y(A^{1980}, L^{1980})} - 1}{\frac{Y(A^{2019}, L^{2019})}{Y(A^{1980}, L^{1980})} - 1} \quad (26)$$

The estimation thus amounts to comparing growth rates with and without educational expansion, *given the actual evolution of the relative efficiency of skilled workers*, 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 income loss from not expanding education would have been greater than what 2019 returns suggest.

6.1.2. 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:

$$\tilde{Y}_{\text{forward}}^{2019} = Y(A^{1980}, L^{2019}) \quad (27)$$

The share of growth explained by education is:

$$\text{Success}_{\text{forward}} = \frac{\frac{Y(A^{1980}, L^{2019})}{Y(A^{1980}, L^{1980})} - 1}{\frac{Y(A^{2019}, L^{2019})}{Y(A^{1980}, L^{1980})} - 1} \quad (28)$$

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, backward accounting will 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 is a much more profitable investment than in a world where it remains unchanged. Comparing estimates of backward and forward growth accounting across countries then identifies the role of skill-biased technical change in enhancing the growth effects of education.

6.1.3. An Empirical Investigation of the Interaction Between Education and Technology

Unfortunately, 1980s survey data are available for almost none of the countries covered by this paper. However, I was able to find and harmonize surveys covering personal income and education in the early 2000s in 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 are not always perfectly comparable, so the results should be interpreted with caution, 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 2000s 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.

6.1.4. Results

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

Gains from schooling are almost always lower under forward 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 is the region where the gap between the two estimates is the largest. This is consistent with pronounced skill-biased technical change in high-income countries in recent decades (the United States does not stand out mainly because educational progress was small). 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, including about 60% in Europe, 50% in Indonesia, and over 100% in Brazil and Mexico. Even absent skill-biased technical change, educational progress would still have had large 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 generates large increases in inequality with little growth for low-income earners. A perfect example is the United States, where pretax incomes have stagnated for the bottom 50% of earners since 1980 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 imperfect substitution, the classic separation between labor and total factor productivity becomes complex and non-additive.²¹ 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 done in this paper.

6.2. Where Does Schooling Come From?

In developing distributional growth *accounting*, I have done nothing else than to estimate how much lower incomes would be had education not improved, holding constant all other characteristics of the economic environment. An objection 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

²¹Caselli and Ciccone (2019) and Jones (2019) provide an interesting discussion along these lines.

determinant of schooling, separating the contribution of education loses some 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. From the 1970s to the 2000s, corresponding to the period during which individuals studied in this paper were still in school, 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 public contribution of schooling for the global poor at over 95%. In a world where governments are the primary providers of education for poor children, the idea that market forces are the main force behind human capital accumulation 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 technology and schooling to closely follow each other. Yet, there are clear examples of disconnections between the two. Perhaps the most illuminating 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 ([Goldin and Katz, 2008](#); [Autor, Goldin, and Katz, 2020](#)). Another example is Latin America, where labor demand has stagnated despite considerable improvements in educational attainment ([Fernández and Messina, 2018](#)). This is not to say that market incentives do not play any role at all: there is plenty of evidence that they do.²² The main argument is that the counterfactual world with no educational progress studied in this paper is not an impossible world. There are concrete instances of disconnections between education and technology that justify growth accounting as a useful tool.

²²A large literature investigates the consequences of trade or technological shocks on schooling: see for instance [Foster and Rosenzweig \(1996\)](#), [Edmonds, Pavcnik, and Topalova \(2010\)](#), [Oster and Steinberg \(2013\)](#), [Atkin \(2016\)](#), [Blanchard and Olneyb \(2017\)](#), [Li \(2018\)](#), and [Cascio and Narayan \(2022\)](#).

3) An Empirical Investigation on the Cross-Country Determinants of Schooling 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).

Appendix Table [A8](#) runs this regression across countries, for the last year available. Appendix Table [A9](#) extends this analysis to the 1980-2019 period with country and year fixed effects. The skill bias of technology is not significantly related to schooling, and if anything enters the regression with the wrong sign. There is also no clear evidence that countries with greater economic growth or faster adoption of new technologies have had larger increases in schooling (if anything, the opposite seems to be true). In contrast, public education spending is strongly associated with improvements in access to schooling. These results do not provide definitive causal evidence, but they suggest that technological progress is unlikely to have been behind the major schooling expansions of the past decades studied in this paper.

6.3. 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 reduce 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 paper, I show that such direct government redistribution can account for 30% of global poverty reduction since 1980 ([Gethin, 2024](#)). Second, public policies can increase future pretax earnings; studying this indirect effect was the objective of this paper, focusing on education. Putting these two estimates together gives us an approximate estimate of the total contribution of public policies to global poverty reduction. This estimate is partial, given that education is not the only policy contributing to pretax earnings growth, but it can be viewed as a lower bound.

Table [8](#) compares the evolution of global poverty and the average income of the world's poorest 20% and 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 54% ($1 - \frac{32}{70}$) of global poverty reduction. Using a similar reasoning, public policies account for about 80% of real income gains among the world's poorest 20%, and about two-thirds of gains for the poorest 50%. Appendix Table A12 reproduces this analysis using World Bank income distribution data, with similar conclusions. The takeaway is that public policies can account for 50-80% of global poverty reduction since 1980, and potentially more if we were to account for the indirect growth effects of other public policies.

7. Conclusion

This article represented a first attempt at estimating the role played by education in the historical reduction of global poverty. Unifying macroeconomic growth decomposition with the canonical model of education and the wage structure, 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 labor 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 illustrate how combining cross-country microdata with simple conceptual frameworks can help uncover new insights on the long-run drivers of the world distribution of income that would otherwise be missed by more aggregate analyses. Similar approaches

could be used to study other key transformations of the past decades, such as trade globalization, structural change, financialization, and even democratization and changing gender norms. The microeconomics literature provides ample and growing empirical evidence on the economic effects of these factors in specific contexts. Combined with the microdata collected in this paper, additional 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 and inequality.

References

- Solow, Robert M. (1957). "Technical Change and the Aggregate Production Function". In: *Review of Economics and Statistics* 39.1, pp. 312–320.
- 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.
- 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.
- Psacharopoulos, George (1994). "Returns to Investment in Education: A Global Update". In: *World Development* 22.9, pp. 1325–1343.
- 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.
- 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.
- 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.
- 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.
- Murphy, Kevin M., W. Craig Riddell, and Paul M. Romer (1998). "Wages, Skills, and Technology in the United States and Canada". NBER Working Paper.
- 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.
- 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.
- 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.
- 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.

- Card, David (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems". In: *Econometrica* 69.5, pp. 1127–1160.
- 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.
- Lemieux, Thomas and David Card (2001). "Education, Earnings, and the 'Canadian G.I. Bill'". In: *The Canadian Journal of Economics / Revue canadienne d'Economie* 34.2, pp. 313–344.
- 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.
- Duflo, Esther (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.
- 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.
- Psacharopoulos, George and Harry Anthony Patrinos (2004). "Returns to Investment in Education: A Further Update". In: *Education Economics* 12.2, pp. 111–134.
- Banerjee, Abhijit and Esther Duflo (2005). "Growth Theory through the Lens of Development Economics". In: pp. 473–552.
- 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.
- Caselli, Francesco and Wilbur John II Coleman (2006). "The World Technology Frontier". In: *American Economic Review* 96.3, pp. 499–522.
- Ciccone, Antonio and Giovanni Peri (2006). "Identifying Human-Capital Externalities: Theory with Applications". In: *Review of Economic Studies* 73.1, pp. 381–412.
- 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.
- 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.
- 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.
- 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.
- 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.

- Edmonds, Eric V, Nina Pavcnik, and Petia Topalova (2010). "Trade Adjustment and Human Capital Investments: Evidence from Indian Tariff Reform". In: *American Economic Journal: Applied Economics* 2.4, pp. 42–75.
- Hsieh, Chang-Tai and Peter J. Klenow (2010). "Development Accounting". In: *American Economic Journal: Macroeconomics* 2.1, pp. 207–223.
- Oyelere, Ruth Uwaifo (2010). "Africa's education enigma? The Nigerian story". In: *Journal of Development Economics* 91.1, pp. 128–139.
- 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.
- 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.
- 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.
- Schoellman, Todd (2012). "Education Quality and Development Accounting". In: *The Review of Economic Studies* 79.1, 388–417.
- 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.
- 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.
- 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.
- 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.
- 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.
- Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". In: *Econometrica* 82.6, 2295–2326.
- Caselli, Francesco, Jacopo Ponticelli, and Federico Rossi (2014). "A New Dataset on Mincerian Returns". Working Paper.
- Jones, Benjamin F. (2014). "The Human Capital Stock: A Generalized Approach". In: *American Economic Review* 104.11, 3752–3777.
- Stephens, Melvin Jr. and Dou-Yan Yang (2014). "Compulsory Education and the Benefits of Schooling". In: *American Economic Review* 104.6, pp. 1777–1792.
- Zimmerman, Seth D. (2014). "The Returns to College Admission for Academically Marginal Students". In: *Journal of Labor Economics* 32.4, pp. 711–754.
- Barro, Robert J. and Jong-Wha Lee (2015). *Education Matters: Global Schooling Gains from the 19th Century to the 21st Century*. Oxford University Press.
- 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.

- Grépin, Karen A. and Bharadwaj Prashant (2015). “Maternal education and child mortality in Zimbabwe”. In: *Journal of Health Economics* 44.1, 97–117.
- 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.
- Atkin, David (2016). “Endogenous Skill Acquisition and Export Manufacturing in Mexico”. In: *American Economic Review* 106.8, 2046–2085.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Kleven, Henrik and Camille Landais (2017). “Gender Inequality and Economic Development: Fertility, Education, and Norms”. In: *Economica* 84.334, pp. 180–209.
- 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.
- 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.
- 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.
- 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.
- Glaeser, Edward and Ming Lu (2018). “Human-Capital Externalities in China”. NBER Working Paper.
- Guo, Junjie, Nicolas Roys, and Ananth Seshadri (2018). “Estimating Aggregate Human Capital Externalities”. Working Paper.
- 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.

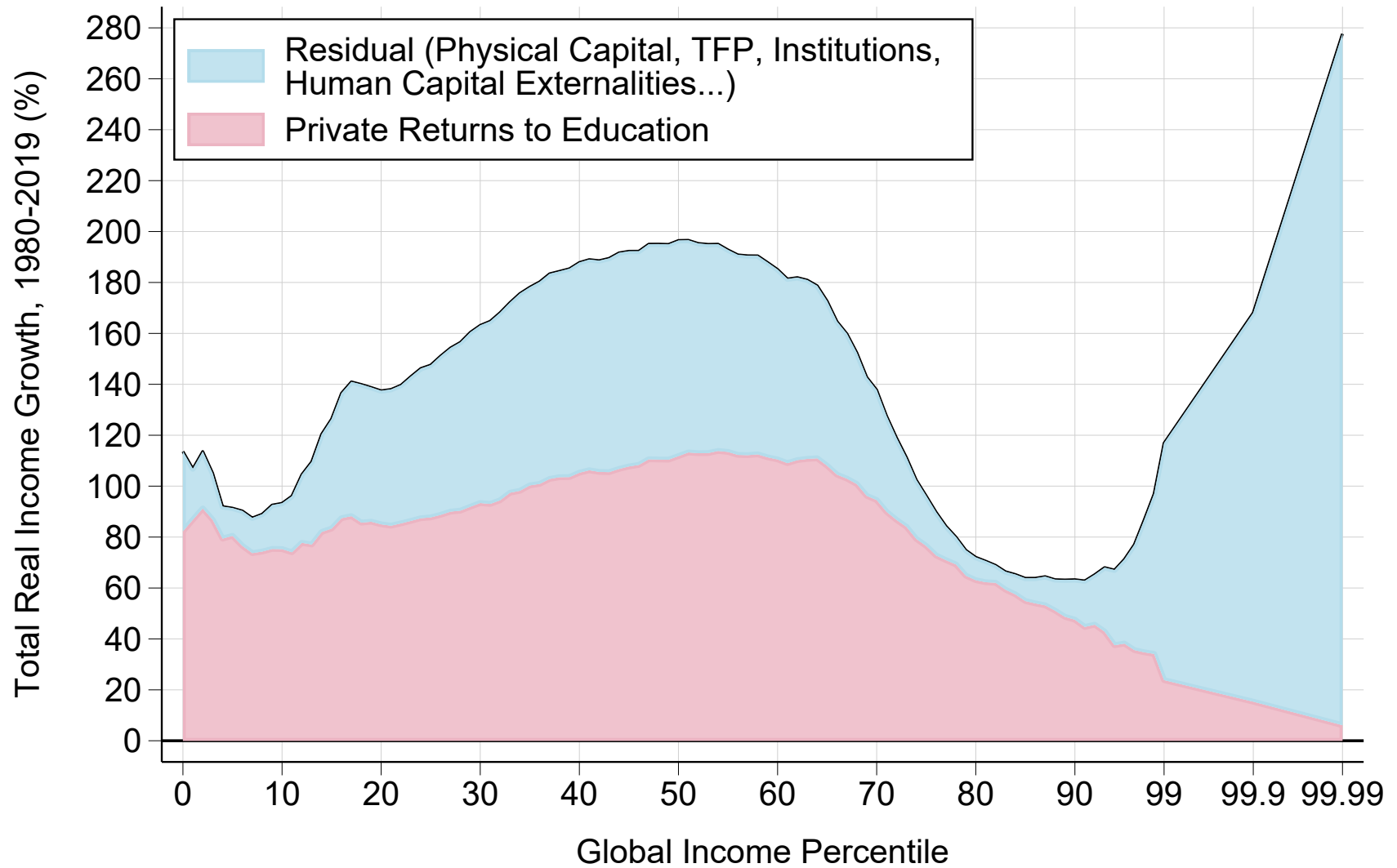
- Hendricks, Lutz and Todd Schoellman (2018). “Human Capital and Development Accounting: New Evidence from Wage Gains at Migration”. In: *The Quarterly Journal of Economics* 133.2, 665–700.
- Keats, Anthony (2018). “Women’s schooling, fertility, and child health outcomes: Evidence from Uganda’s free primary education program”. In: *Journal of Development Economics* 135.1, pp. 142–159.
- Li, Bingjing (2018). “Export expansion, skill acquisition and industry specialization: evidence from China”. In: *Journal of International Economics* 114.1, pp. 346–361.
- 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.
- Caselli, Francesco and Antonio Ciccone (2019). “The Human Capital Stock: A Generalized Approach: Comment”. In: *American Economic Review* 109.3, pp. 1155–1174.
- Chancel, Lucas and Thomas Piketty (2019). “Indian Income Inequality, 1922-2015: From British Raj to Billionaire Raj?” In: *Review of Income and Wealth* 65.S1, S33–S62.
- 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.
- Jones, Benjamin F. (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.
- Piketty, Thomas, Li Yang, and Gabriel Zucman (2019). “Capital Accumulation, Private Property, and Rising Inequality in China, 1978-2015”. In: *American Economic Review* 109.7, pp. 2469–2496.
- 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.
- 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.
- Gunderson, Morley and Philip Oreopoulos (2020). “Returns to education in developed countries”. In: *The Economics of Education*. Ed. by Steve Bradley and Colin Green. Academic Press, pp. 39–51.
- Hammar, Olle and Daniel Waldenström (2020). “Global Earnings Inequality, 1970-2018”. In: *The Economic Journal* 130.632, pp. 2526–2545.
- 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.
- Rossi, Federico (2020). “Human Capital and Macroeconomic Development: A Review of the Evidence”. In: *The World Bank Research Observer* 35.2, pp. 227–262.
- 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.
- Blanchet, Thomas, Lucas Chancel, Ignacio Flores, and Marc Morgan (2021). *Distributional National Accounts Guidelines: Methods and Concepts Used in the World Inequality Database*. Tech. rep.
- Bowlus, Audra, Lance Lochner, Chris Robinson, and Eda Suleymanoglu (2021). “Wages, Skills, and Skill-Biased Technical Change: The Canonical Model Revisited”. Working Paper.
- 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.

- Chatterjee, Aroop, Léo Czajka, and Amory Gethin (2021). “Can Redistribution Keep Up with Inequality? Evidence from South Africa, 1993-2019”. Working Paper.
- 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.
- 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.
- 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.
- Khan, Salman (2021). “Unmet Expectations: The Impacts of School Construction on Female Outcomes in Rural Punjab, Pakistan”. Working Paper.
- 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.
- 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.
- Bachas, Pierre, Matthew H. Fisher-Post, Anders Jensen, and Gabriel Zucman (2022). “Globalization and Factor Income Taxation”. NBER Working Paper 29819.
- Bils, Mark, Baris Kaymak, and Kai-Jie Wu (2022). “Labor Substitutability Among Schooling Groups”. NBER Working Paper.
- 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.
- Cascio, Elizabeth U. and Ayushi Narayan (2022). “Who Needs a Fracking Education? The Educational Response to Low-Skill-Biased Technological Change”. In: *ILR Review* 75.1, pp. 56–89.
- Chancel, Lucas, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2022). *World Inequality Report 2022*. Harvard University Press.
- Chatterjee, Aroop, Léo Czajka, and Amory Gethin (2022). “Wealth Inequality in South Africa, 1993-2017”. In: *World Bank Economic Review* 36.1, pp. 19–36.
- De Rosa, Mauricio, Ignacio Flores, and Marc Morgan (2022). “More Unequal or Not as Rich? Revisiting the Latin American Exception”. Stone Center for Inequality Studies Working Paper Series.
- Deming, David J. (2022). “Four Facts about Human Capital”. In: *Journal of Economic Perspectives* 36.3, pp. 75–102.
- Ekhtator-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.
- Gethin, Amory (2022). “Who Benefits from Public Goods? Evidence from South Africa”. Working Paper.
- 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.

- 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.
- Le Nestour, Alexis, Laura Moscoviz, and Justin Sandefur (2022). “The Long-Run Decline of Education Quality in the Developing World”. Working Paper.
- Oliobi, Ifeatu (2022). “Access to Education and Family Formation: Evidence from University Expansion in Nigeria”. Working Paper.
- Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo (2022). “The Human Side of Structural Transformation”. In: *American Economic Review* 112.8, pp. 2774–2814.
- 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 (2022). “The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation”. In: *American Economic Review* 112.1, pp. 235–266.
- Vu, Khoa and Tu-Anh Vu-Thanh (2022). “Higher Education Expansion, Labor Market, and Firm Productivity in Vietnam”. Working Paper.
- 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.
- 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.
- Buera, Francisco J., Joseph P. Kaboski, and Robert M. Townsend (2023). “From Micro to Macro Development”. In: *Journal of Economic Literature* 61.2, pp. 471–503.
- Chancel, Lucas, Denis Cogneau, Alix Myczkowski, and Anne-Sophie Robilliard (2023). “Income Inequality in Africa, 1990-2019: Measurement, Patterns, Determinants”. In: *World Development* 163.1, pp. 106–162.
- Deming, David J. (2023). “Multidimensional Human Capital and the Wage Structure”. NBER Working Paper.
- Elsayed, Ahmed and Alina Shirshikova (2023). “The women-empowering effect of higher education”. In: *Journal of Development Economics* 163.1, pp. 1031–1041.
- Fisher-Post, Matthew and Amory Gethin (2023). “Government Redistribution and Development: Global Estimates of Tax-and-Transfer Progressivity, 1980-2019”. Working Paper.
- Fujimoto, Junichi, David Lagakos, and Mitchell VanVuren (2023). “Macroeconomic Effects of Free Secondary Schooling in the Developing World”. NBER Working Paper Series.
- Hendricks, Lutz and Todd Schoellman (2023). “Skilled Labor Productivity and Cross-Country Income Differences”. In: *American Economic Journal: Macroeconomics* 15.1, pp. 240–268.
- 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.

- 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.
- Nimier-David, Elio (2023). “Local Human Capital and Firm Creation: Evidence from the Massification of Higher Education in France”. Working Paper.
- United Nations (2023). *The Sustainable Development Goals Report*. United Nations.
- World Bank (2023). *World Development Indicators*. World Bank.
- Gethin, Amory (2024). “Revisiting Global Poverty Reduction: Public Goods and the World Distribution of Income, 1980-2022”. Working Paper.

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. From 1980 to 2019, the average income of the 20th percentile of the world distribution of income grew by 140%, 80 percentage points of which can be rationalized by private returns to education. Education thus accounts for about 60% of growth among this group since 1980.

Figure 2 – Survey Microdata Coverage

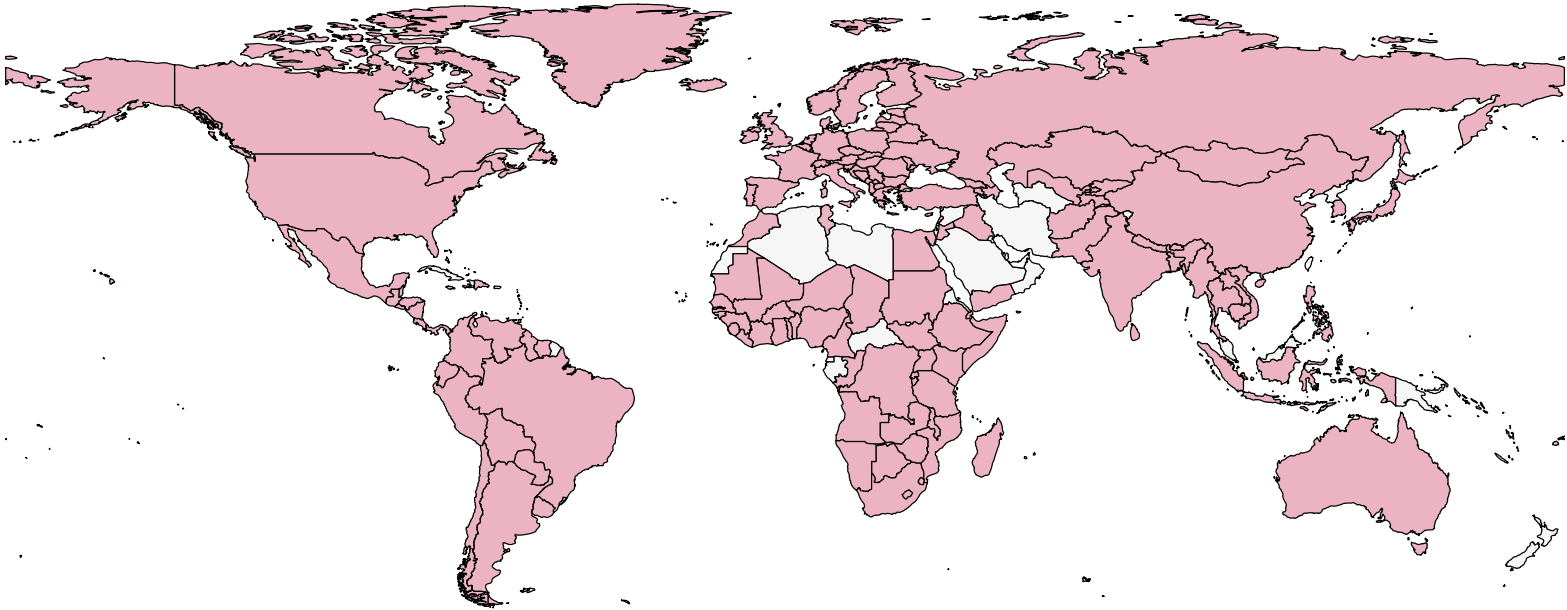
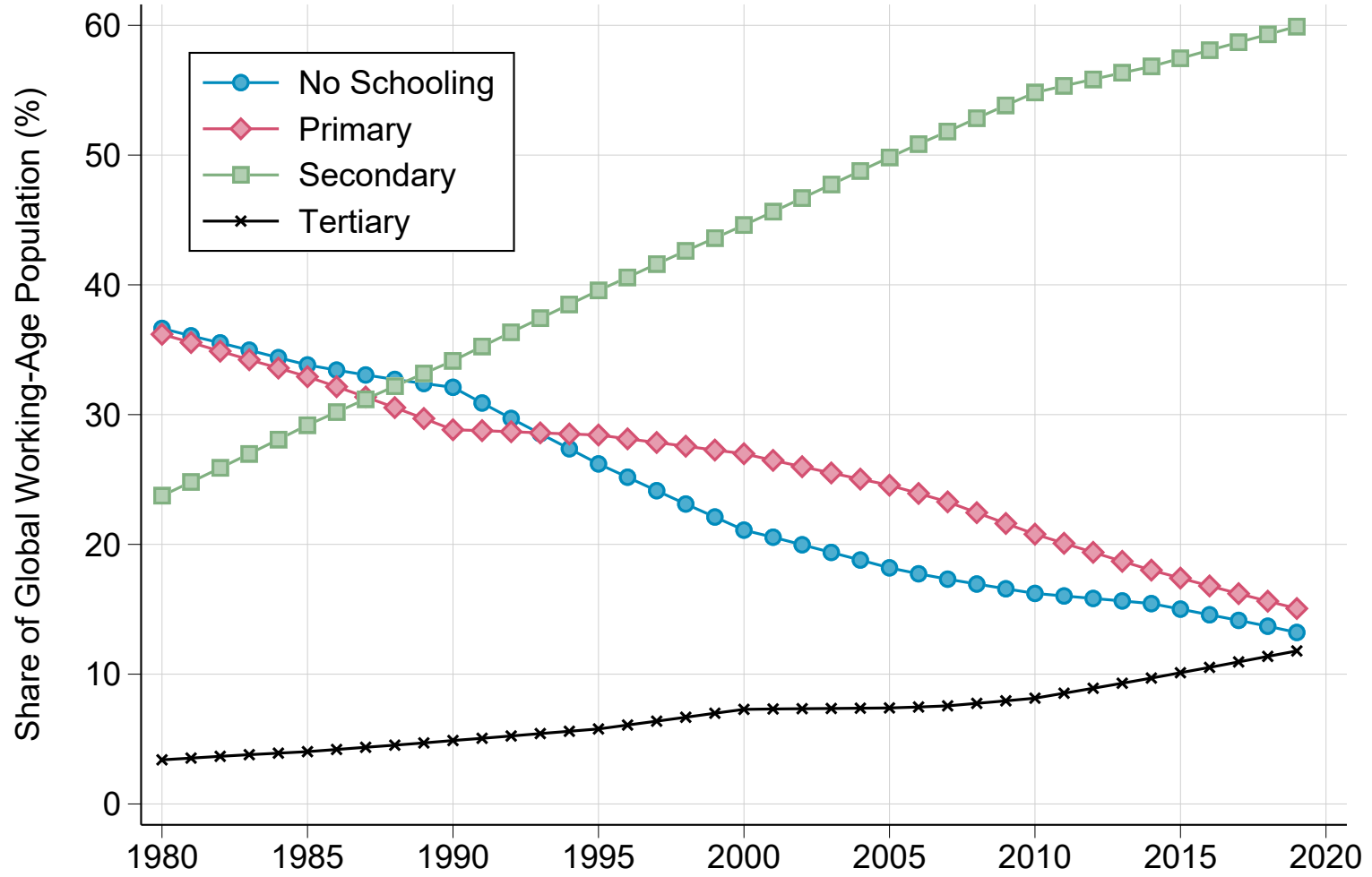
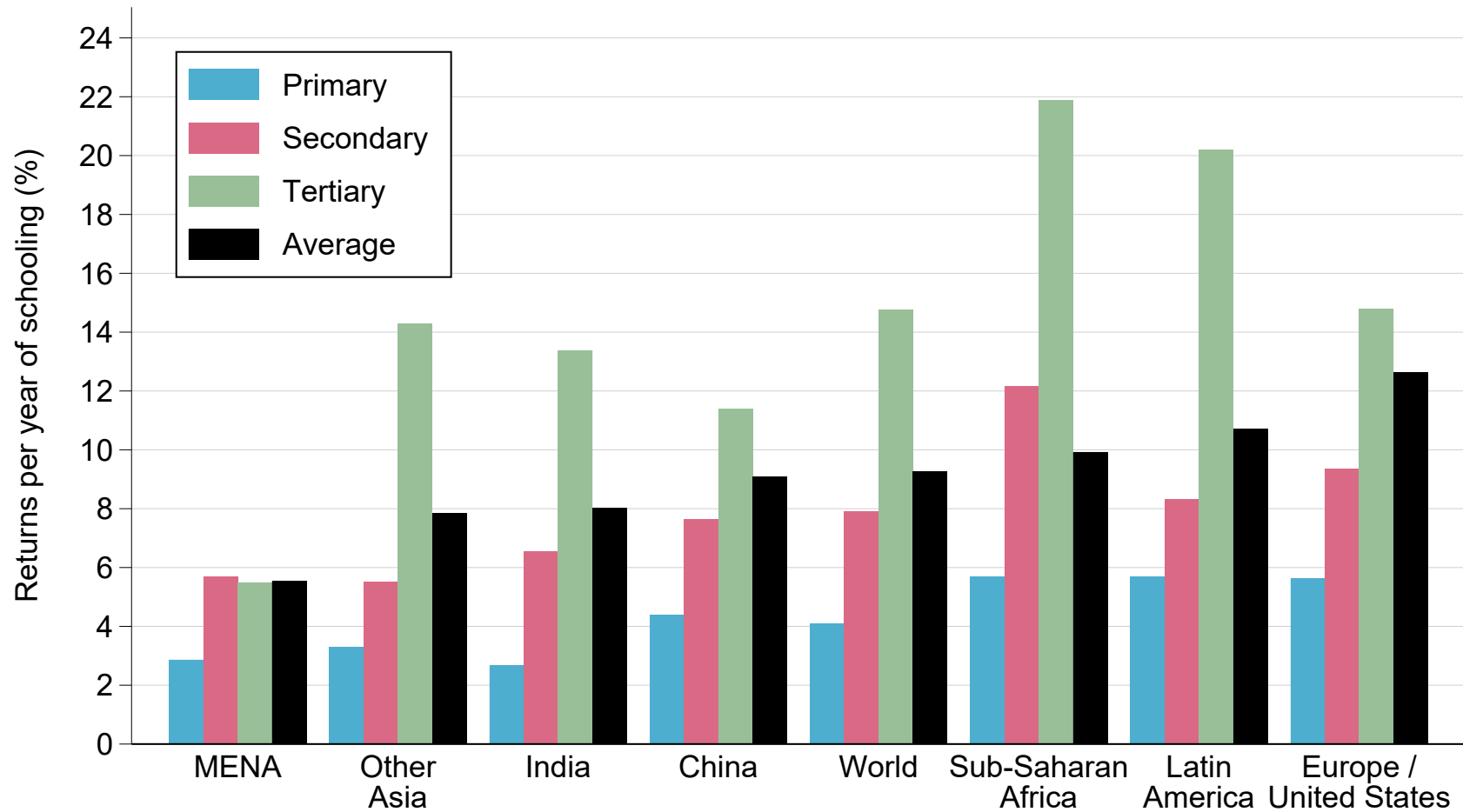


Figure 3 – Educational Attainment of the World’s Working-Age Population, 1980-2019



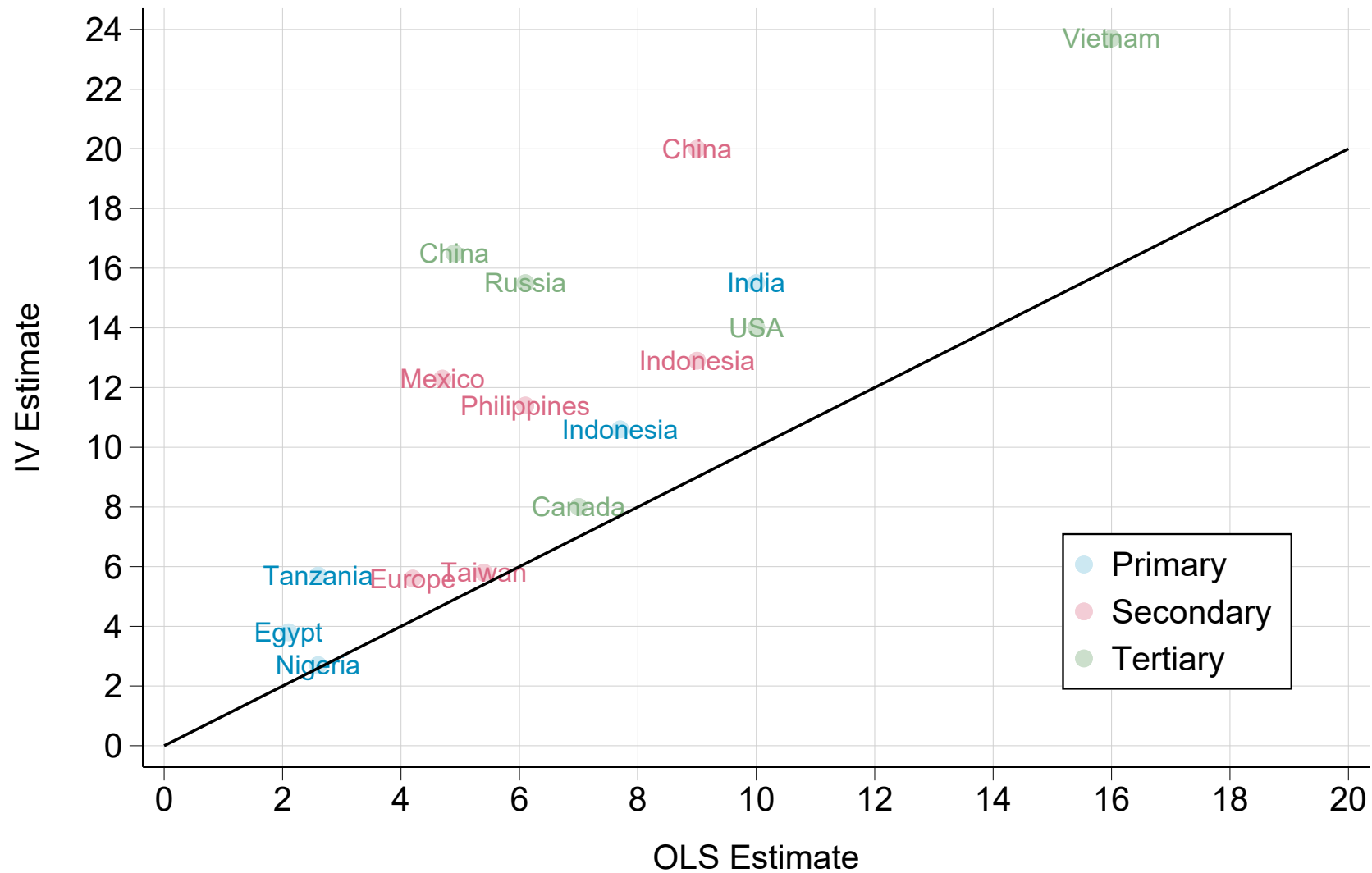
Notes. The figure plots the distribution of educational attainment of the working-age population in the world as a whole. From 1980 to 2019, the share of the world’s working-age population having reached secondary education grew from about 25% to 60%.

Figure 4 – Returns to Schooling by World Region



Notes. The figure plots returns to a year of schooling by education level and world region. Estimates correspond to the effect of an additional year of schooling on the log of personal income, estimated using modified Mincerian equations that control for an experience quartic, gender, and interactions between the experience quartic and gender. Primary: return to an additional year of primary education. Secondary: return to an additional year of secondary education. Tertiary: return to an additional year of tertiary education. Population-weighted averages of coefficients estimated in each country.

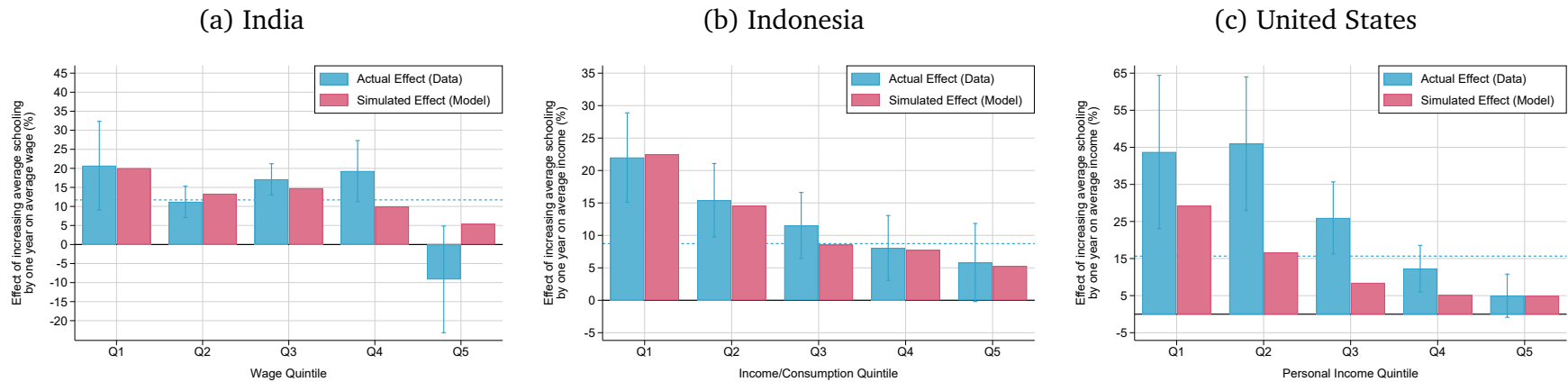
Figure 5 – Returns to Schooling: OLS versus IV Estimates



55

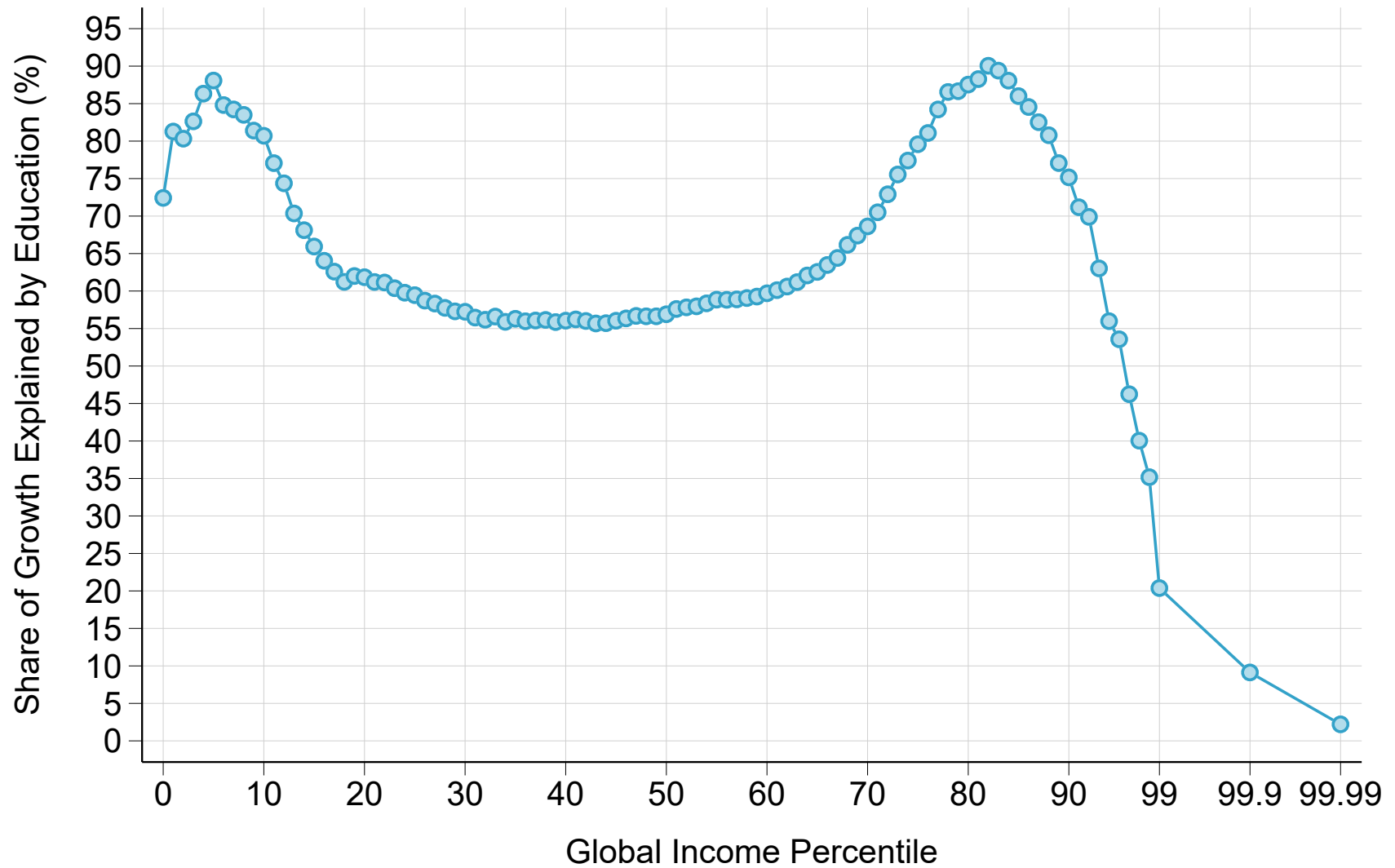
Notes. The figure compares ordinary least squares (x-axis) and instrumental variable (y-axis) estimates of the return to an additional year of schooling. OLS estimates generally correspond to coefficients obtained from a Mincerian equation of the log of earnings on years of schooling, estimated over 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). Author's elaboration compiling estimates from a number of empirical studies: see Appendix Table G4.

Figure 6 – Validation: Actual Versus Simulated Distributional Effects of Education Expansion Policies in India, Indonesia, and the United States



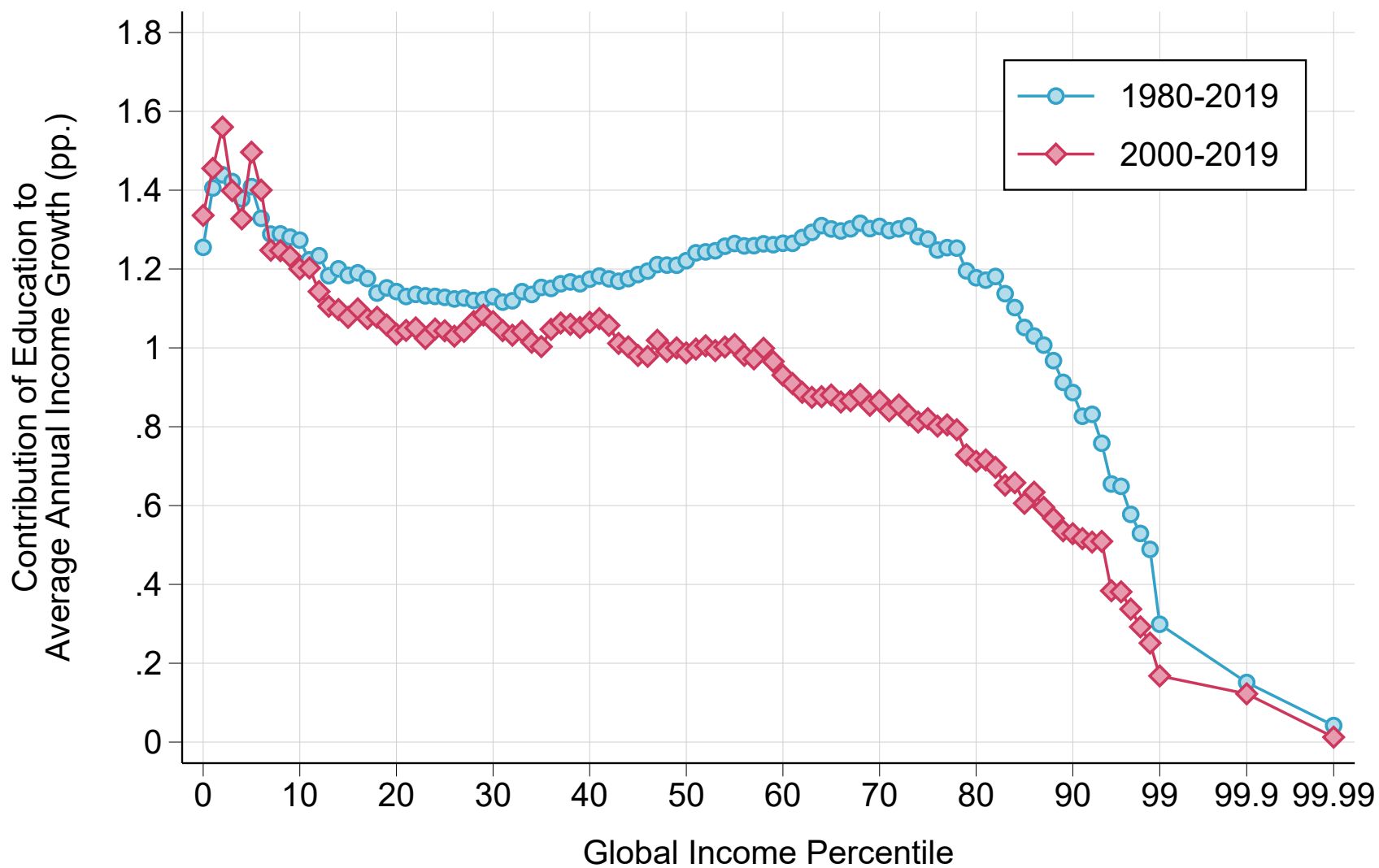
Notes. The three figures compare actual causal effects of educational expansion on the average income of each quintile with simulated effects predicted by the model. Capped spikes correspond to 95% confidence intervals. The dashed line shows the estimated effect of average regional years of schooling on the average income. India: effect of increasing average district schooling by one year on individual wages, instrumenting schooling with exposure to the District Primary Education Program. Estimates combine NES microdata with exposure to the policy from [Khanna \(2023\)](#). Indonesia: effect of increasing average district schooling by one year on per-adult consumption, instrumenting schooling with exposure to the INPRES program. Estimates combine SUSENAS 1993-2019 microdata with INPRES program intensity from [Duflo \(2001\)](#). United States: effect of increasing average state schooling by one year on personal income, instrumenting schooling with state compulsory schooling laws. Estimates combine 1940-2000 census microdata with compulsory schooling laws from [Acemoglu and Angrist \(2000\)](#) and [Clay, Lingwall, and Stephens \(2021\)](#).

Figure 7 – Share of Growth Explained by Education by Global Income Percentile, 1980-2019



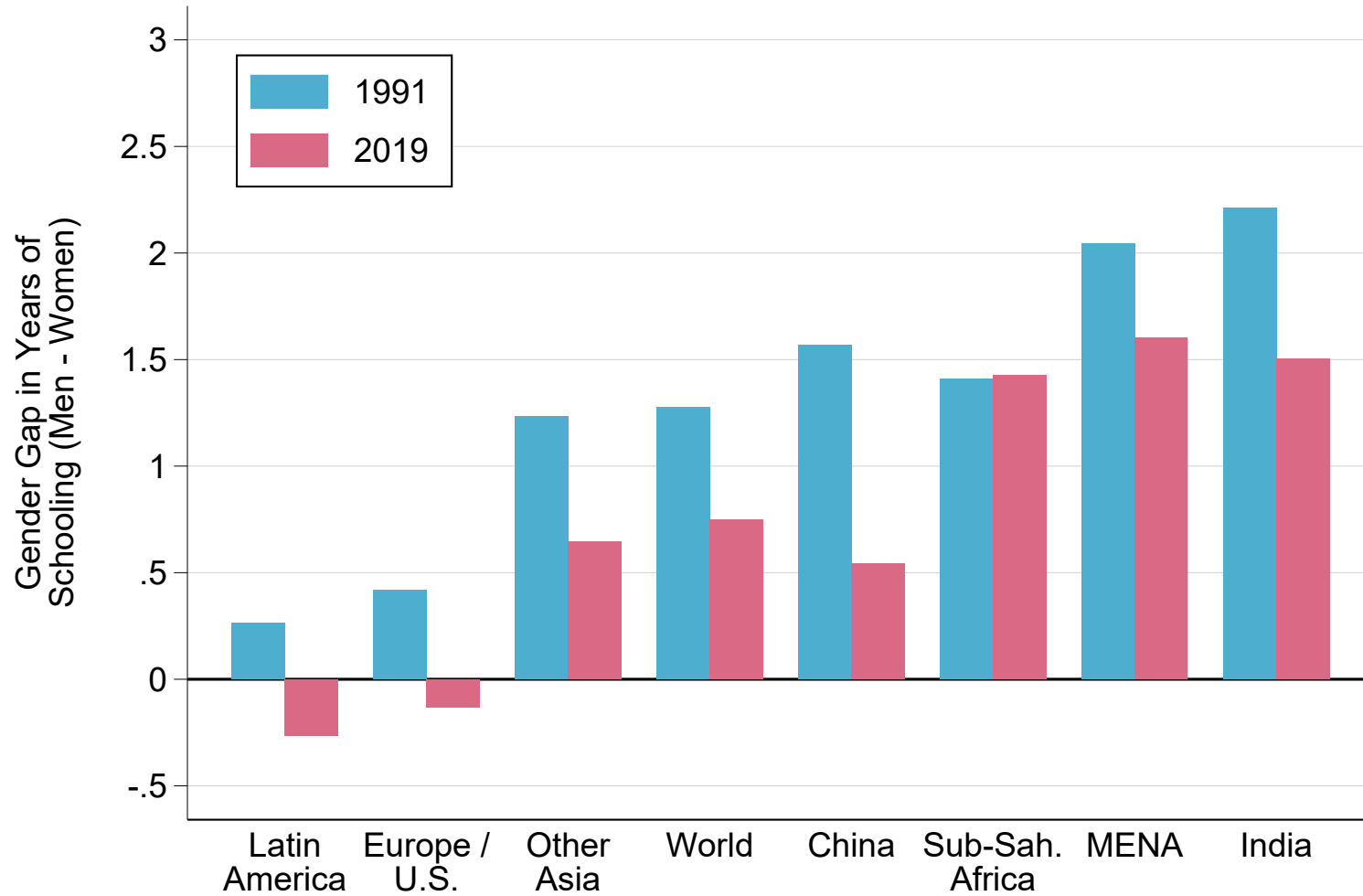
Notes. The figure plots the share of real pretax income growth explained by education by global income percentile. Each point corresponds to the ratio of gains from schooling—equal to actual minus counterfactual income growth absent educational expansion—over actual income growth, calculated for each percentile of the world distribution of income. Education explains about 60% of growth for the 20th percentile of the world distribution of income.

Figure 8 – Annualized Real Income Gains from Schooling by Global Income Group, 1980-2019



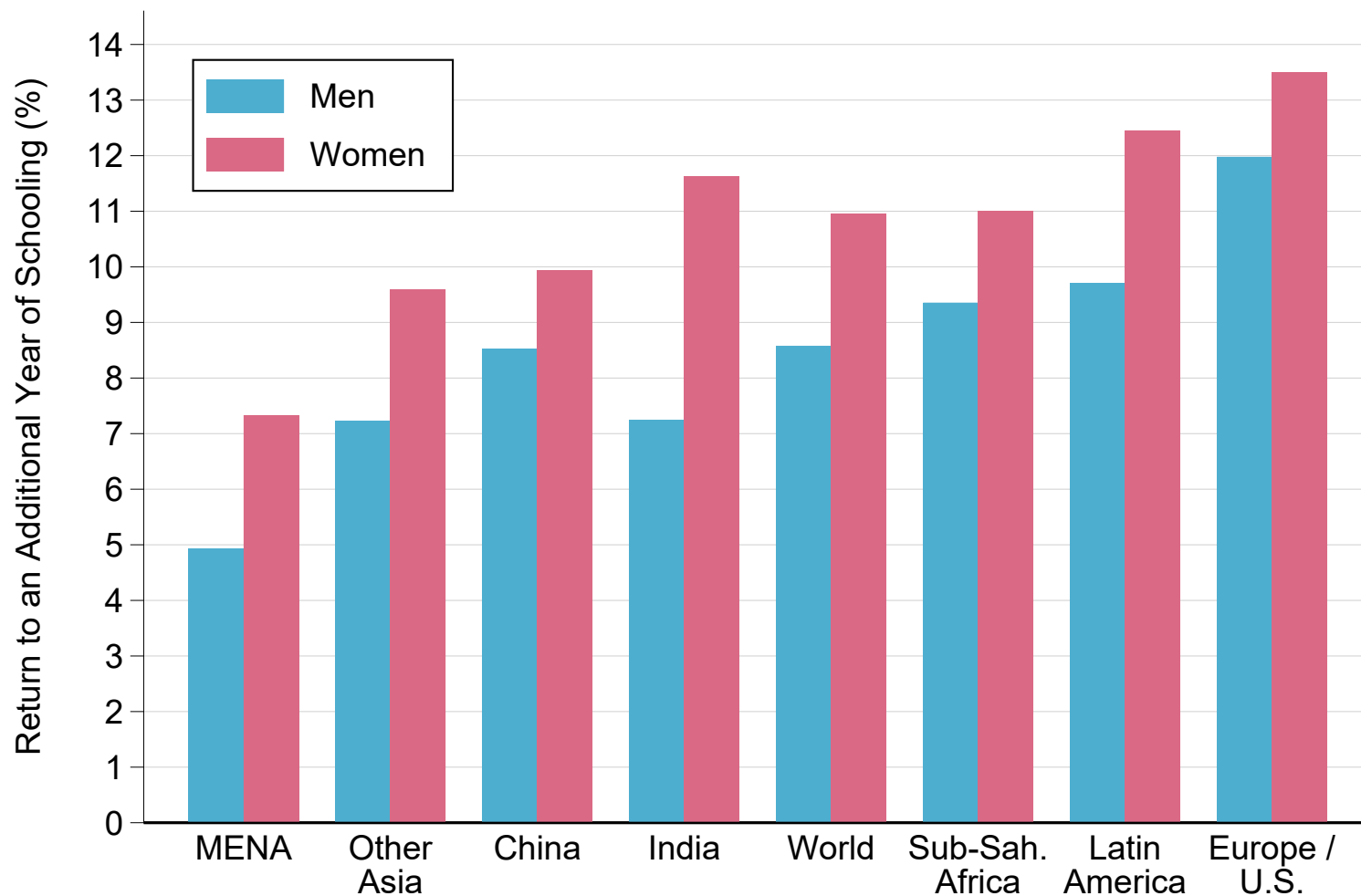
Notes. The figure plots annualized gains from schooling by global income percentile, calculated by taking the percent difference between actual and counterfactual income growth absent educational expansion, and annualizing the resulting figure over the period considered. From 1980 to 2019, education contributed to increasing the average income of the 50th percentile of the world distribution of income by about 1.2% per year.

Figure 9 – Global Gender Schooling Inequality, 1991-2019
 Gender Gap in Years of Schooling (Men - Women)



Notes. The figure shows the gap in years of schooling between working-age men and women by world region and in the world as a whole. Men were more educated than women by over 2 years in India in 1991, compared to 1.5 years in 2019.

Figure 10 – Returns to Schooling by Gender and World Region



Notes. The figure plots returns to a year of schooling by gender and world region in 2019. 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 that control for an experience quartic. In all world regions, the return to a year of schooling is higher for women than for men.

Table 1 – A New Survey Microdatabase on Education and Inequality

	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.

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%	+34%	81	71%
Next 30%	+176%	+76%	100	57%
Middle 40%	+94%	+24%	70	74%
Top 10%	+91%	+57%	34	38%
Top 1%	+131%	+109%	22	17%
Top 0.1%	+173%	+158%	15	9%
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. The average income of the world's poorest 20% grew by 115% from 1980 to 2019. Absent educational expansion, it would have grown by 34%. Education thus accounts for 71% of real income growth among the world's poorest 20% since 1980.

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%	42%
Global Poverty: \$3.65 / Day				
Actual	40%	14%	-65%	
Counterfactual	40%	22%	-44%	32%
Global Poverty: \$6.85 / Day				
Actual	58%	27%	-53%	
Counterfactual	58%	41%	-30%	44%

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. Global poverty at \$2.15 per day declined by 55% from 1980 to 2019. Absent educational expansion, it would have declined by 32%. Education thus accounts for 42% of global poverty reduction 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%	50	62%	+30%	69	>100%
Latin America	+38%	33	86%	+39%	60	>100%
China	+988%	289	29%	+377%	205	54%
India	+410%	124	30%	+208%	107	52%
Other Asia-Pacific	+117%	56	48%	+263%	125	48%
Middle East and North Africa	+118%	46	39%	+35%	46	>100%
Sub-Saharan Africa	+15%	31	>100%	+51%	65	>100%
Low-income	+16%	35	>100%	+50%	70	>100%
Low-middle-income	+182%	68	37%	+198%	100	51%
High-middle-income	+210%	83	40%	+218%	143	65%
High-income	+89%	50	57%	+77%	94	>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. Average pretax income grew by 81% in Europe and Northern America from 1980 to 2019. Absent educational expansion, it would have grown by 50 percentage points less. Education thus accounts for 62% of aggregate economic growth in this region since 1980.

Table 5 – Education and Inequality Between and Within Countries

	1980	2019	Difference
Theil Index of Global Inequality			
Actual	1.06	1.07	0.02
Counterfactual	1.06	1.33	0.27
Between-Country Component			
Actual	0.60	0.34	-0.26
Counterfactual	0.60	0.34	-0.26
Within-Country Component			
Actual	0.46	0.74	0.28
Counterfactual	0.46	0.99	0.53
Within-Country Share (%)			
Actual	43%	69%	25
Counterfactual	43%	74%	31

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. The Theil index of global pretax income inequality was 1.07 in 2019. Absent educational expansion since 1980, it would have reached 1.33. 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	40%	41%
+ Distributional Effects, 2019 Returns	40%	52%
+ Distributional Effects, Adjusted Returns	54%	71%
+ Distributional Effects, IV Returns	54%	75%

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. Standard growth accounting: only rely on cross-country data, assuming a uniform 10% return to schooling and that education only affects wages in each country. Adjusted labor share: account for the fact that education affects not only wages but also mixed income. Within-country inequality: account for within-country inequality—the world’s poorest 20% do not all live in the poorest 20% countries. Within-country labor shares: account for the fact that labor income is concentrated at the bottom of the distribution in each country. Microdata, 2019 returns: bring in the microdata, allowing returns to schooling to vary by country and education level. Distributional effects: account for the distributional implications of supply effects—absent educational expansion, the relative wage of skilled workers would be higher (see Section 2.3). Adjusted returns: account for the aggregate implications of supply effects—absent educational expansion, the return to schooling would be higher (see Section 2.2). IV returns: rely on instrumental variable estimates of returns to schooling instead of adjusted returns.

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	55%	50%
Counterfactual: + Heterogeneous Returns	29.3%	30.1%	0.8	71%	58%
Counterfactual: + Extensive Margin	29.3%	29.9%	0.6	78%	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. No educational progress: only account for differential trends in schooling by gender, applying the same returns to schooling to 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. The share of labor income received by women in the world as a whole increased by 2.8 percentage points from 1991 to 2019. Absent educational expansion, it would have grown by 0.6 points (after accounting for all effects of education). Education thus accounts for 78% of global gender labor income inequality reduction since 1991.

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%	54%
Global Bottom 20% Average Income (\$/Day)				
Pretax Income Absent Educational Expansion	1.3	1.7	+32%	
Pretax Income	1.3	2.8	+115%	
Posttax Income	1.5	4.0	+164%	80%
Global Bottom 50% Average Income (\$/Day)				
Pretax Income Absent Educational Expansion	2.7	4.5	+67%	
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 to individual incomes (see [Gethin, 2024](#)). 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). Global poverty would have declined by 32% absent educational expansion and government redistribution, while it declined by 70% in terms of posttax income. Public policies thus explain 54% of global poverty reduction since 1980. 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.

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

Amory Gethin

August 2024

SUPPLEMENTARY ONLINE APPENDIX

This appendix supplements my article “Distributional Growth Accounting: Education and the Reduction of Global Poverty, 1980-2022.” It contains additional methodological details, as well as supplementary figures and tables.

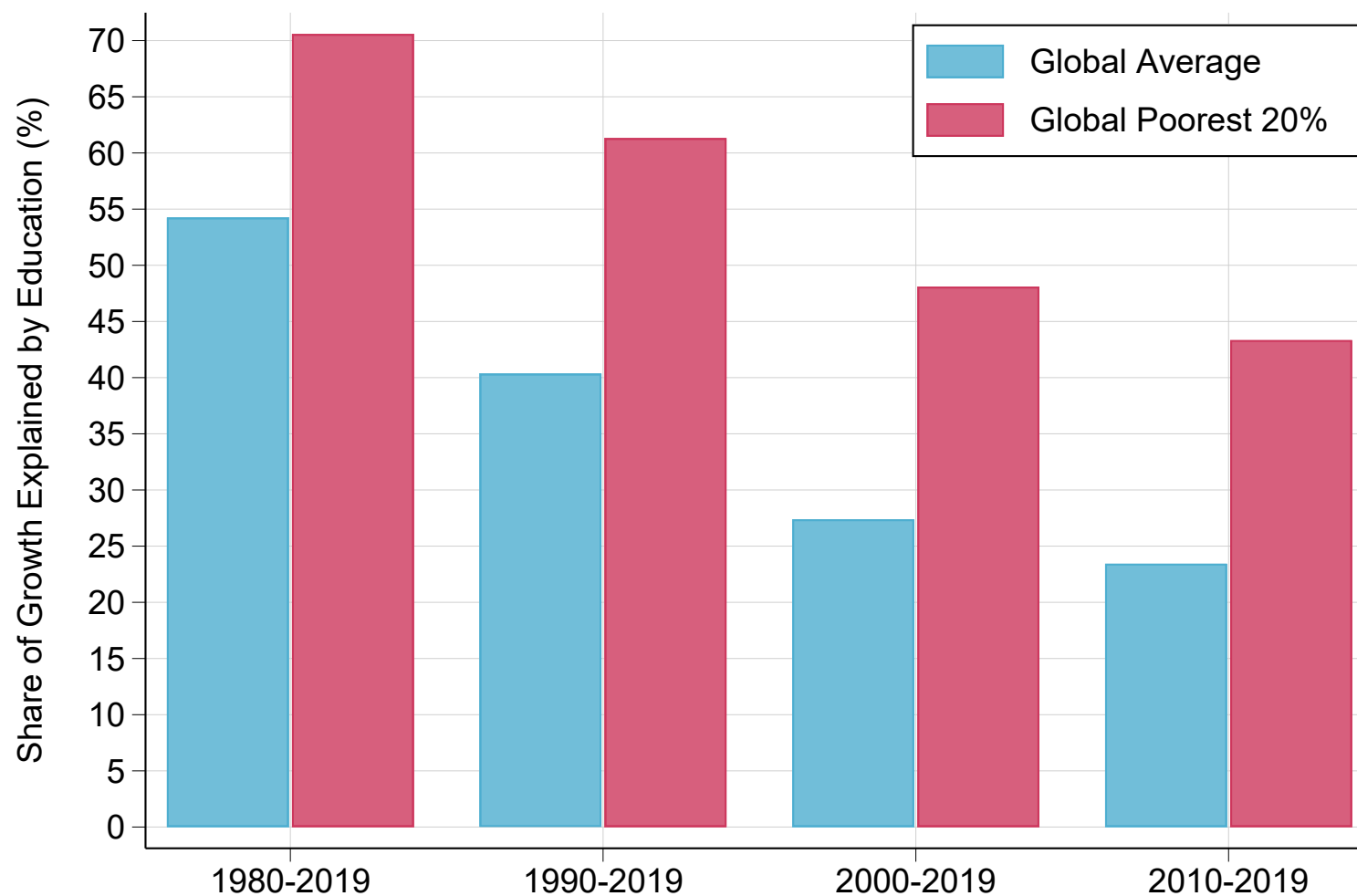
Table of Contents

A	Additional Key Figures and Tables	3
B	Estimation Details	33
C	Natural Experiments	52
D	Education Quality	72
E	Data Appendix: Survey Microdata	90
F	Data Appendix: Educational Attainment Data	101
G	Data Appendix: Returns to Schooling	109
H	Data Appendix: Income Distribution Data	124

A. Additional Key Figures and Tables

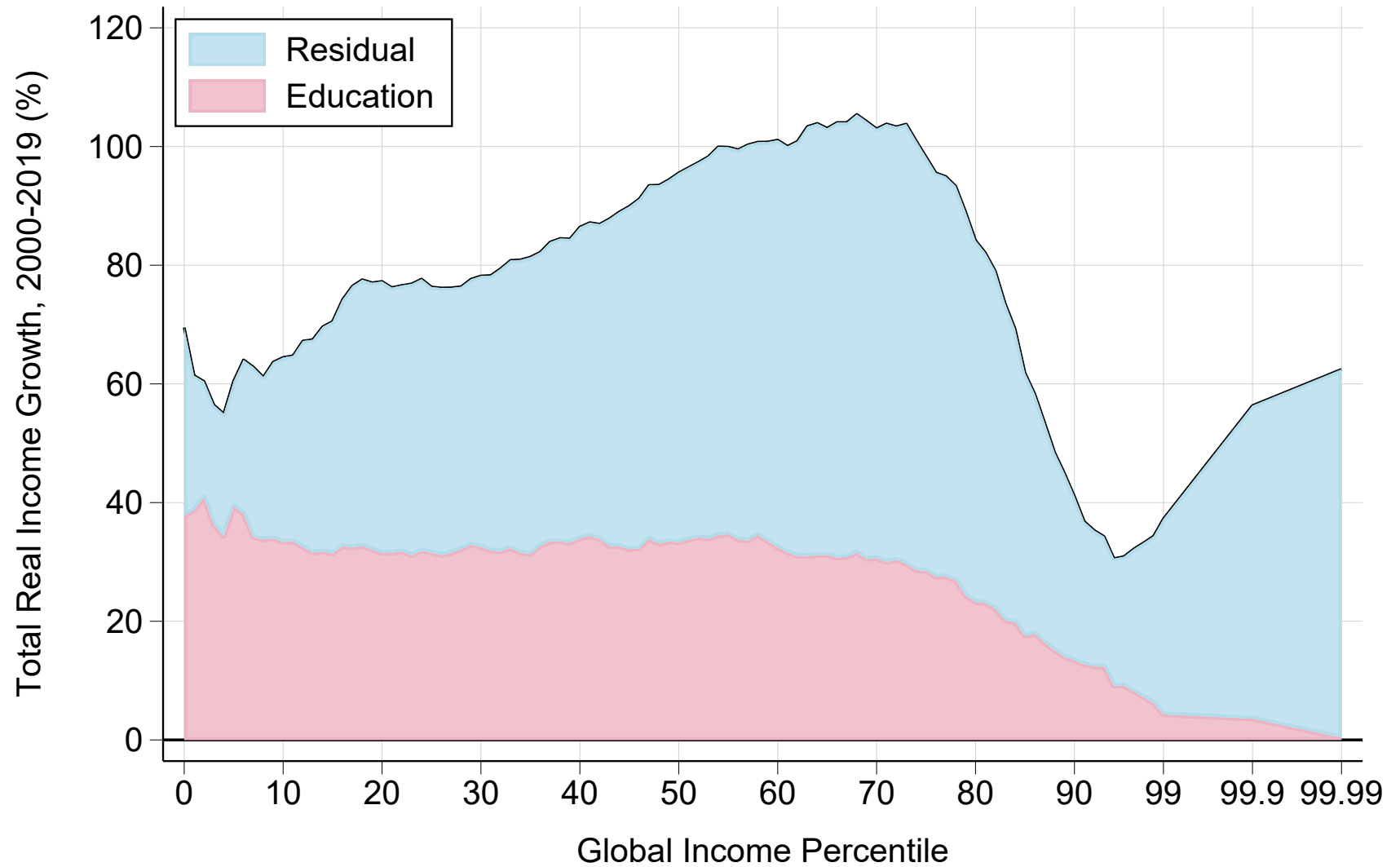
A.1. Additional Results: World Distribution of Income

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



Notes. The figure plots the share of global economic growth and real income growth among the world's poorest 20% that can be explained by education for different time periods.

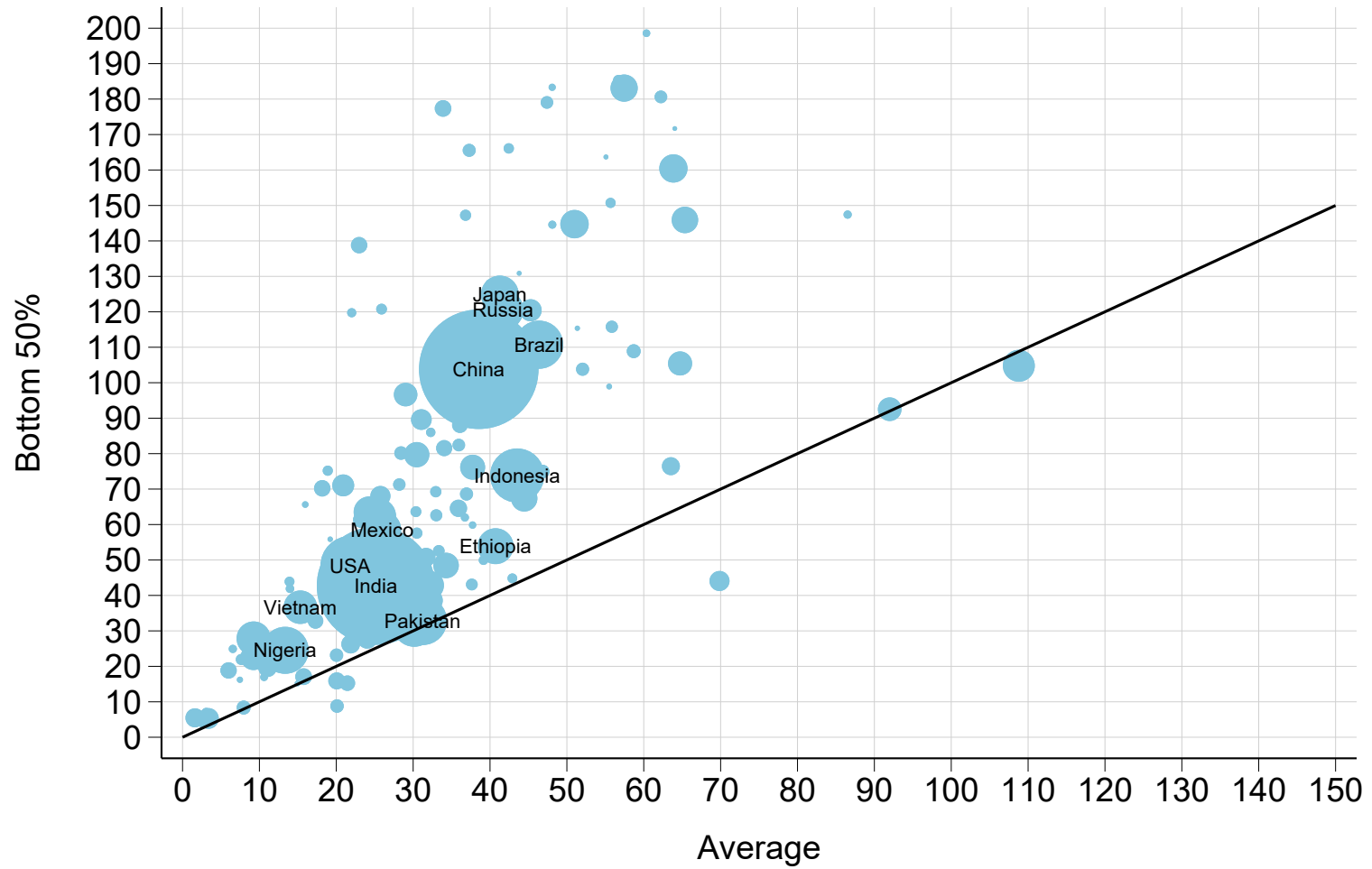
Figure A2 – The Distribution of Global Economic Growth, 2000-2019



4

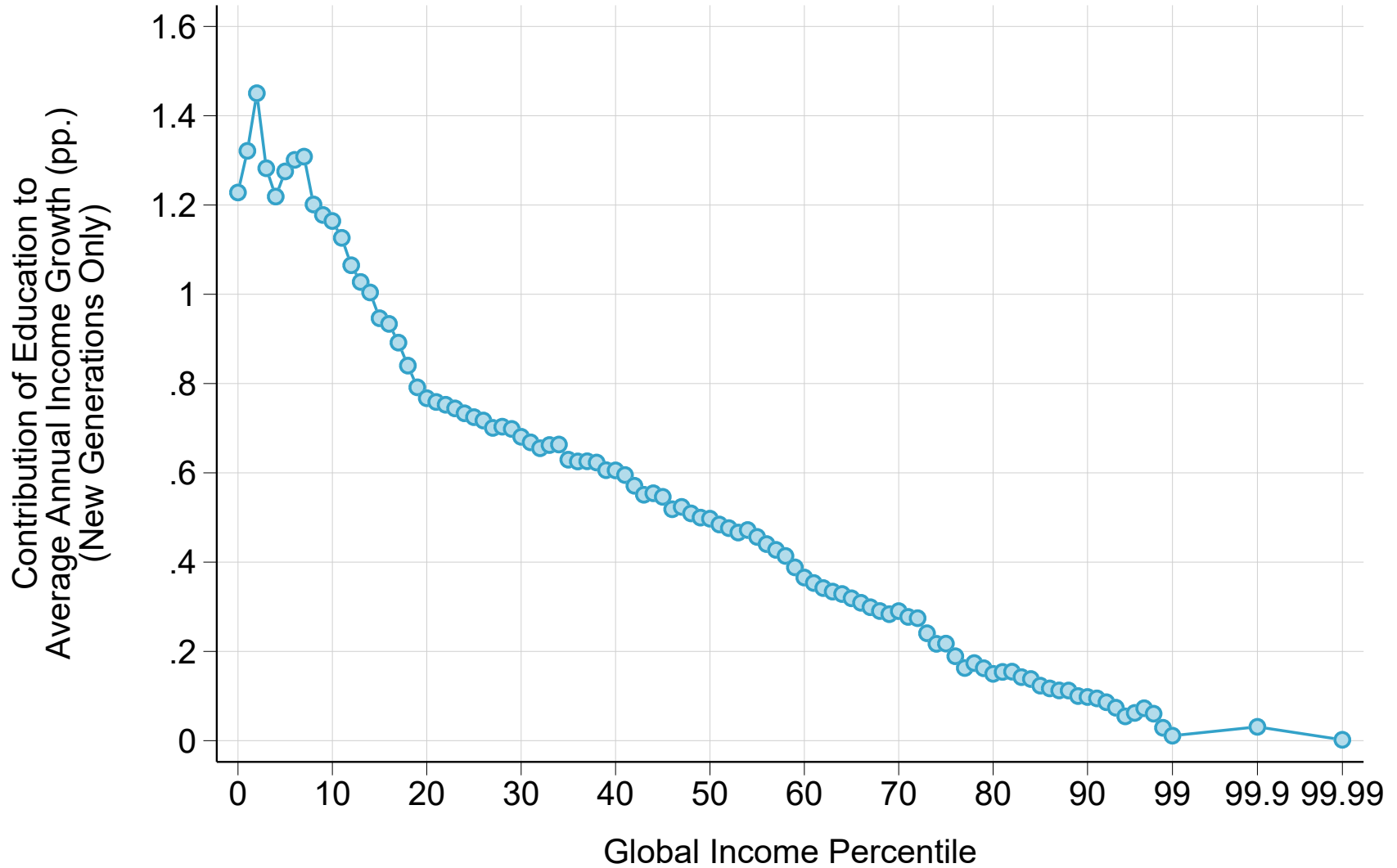
Notes. The figure plots total real income growth by global income percentile from 2000 to 2019, decomposing it into a part that can be explained by private returns to schooling and an unexplained component. The upper shaded area represents the growth rates that would have prevailed absent any improvement in the education of the world's working-age population since 2000. The lower shaded area represents the corresponding contribution of education to economic growth.

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



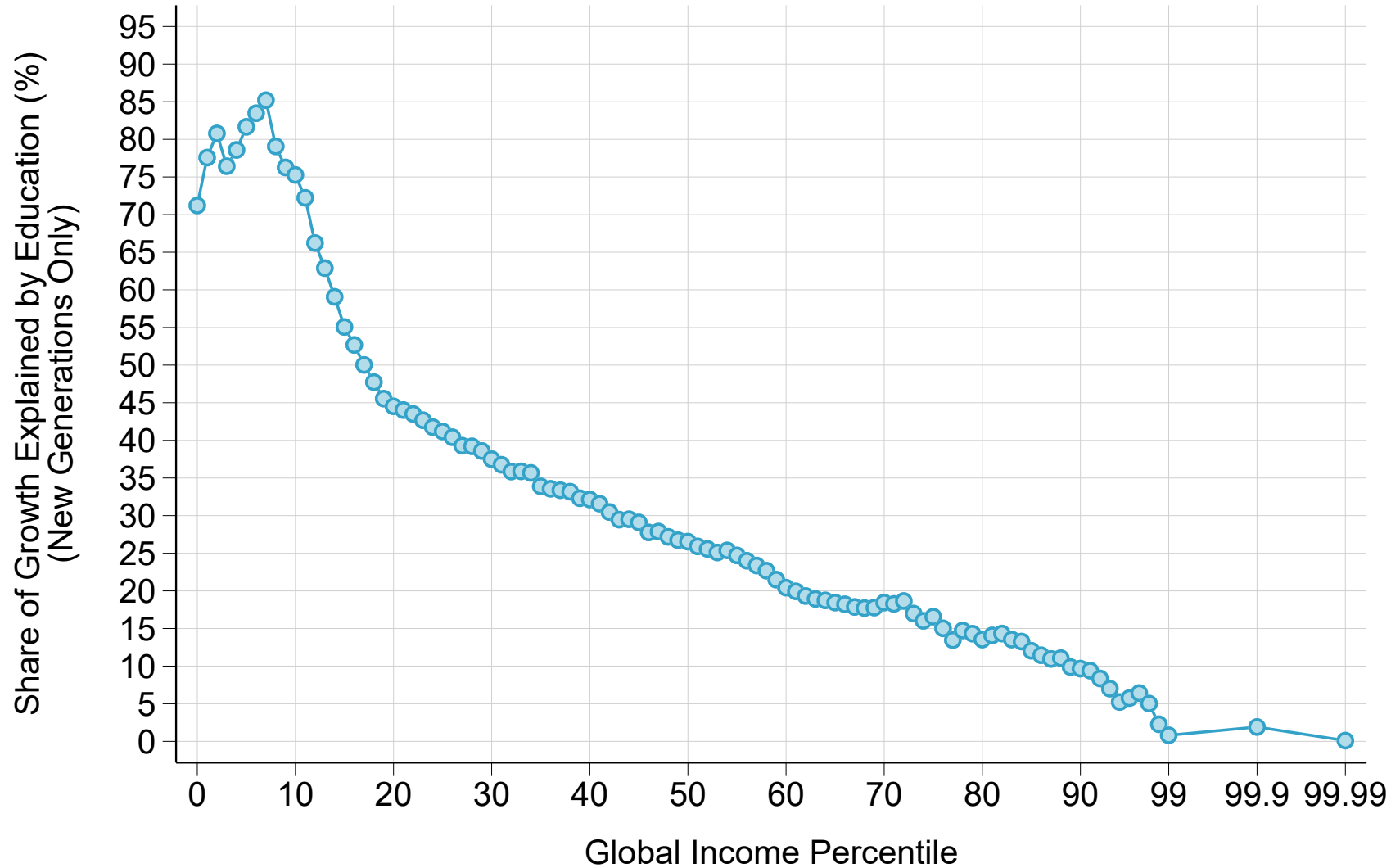
Notes. The figure compares gains from schooling for the population as a whole (average) versus the bottom 50% in each country over the 1980-2019 period. Gains from schooling correspond to the percent increase in income generated by educational expansion since 1980.

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



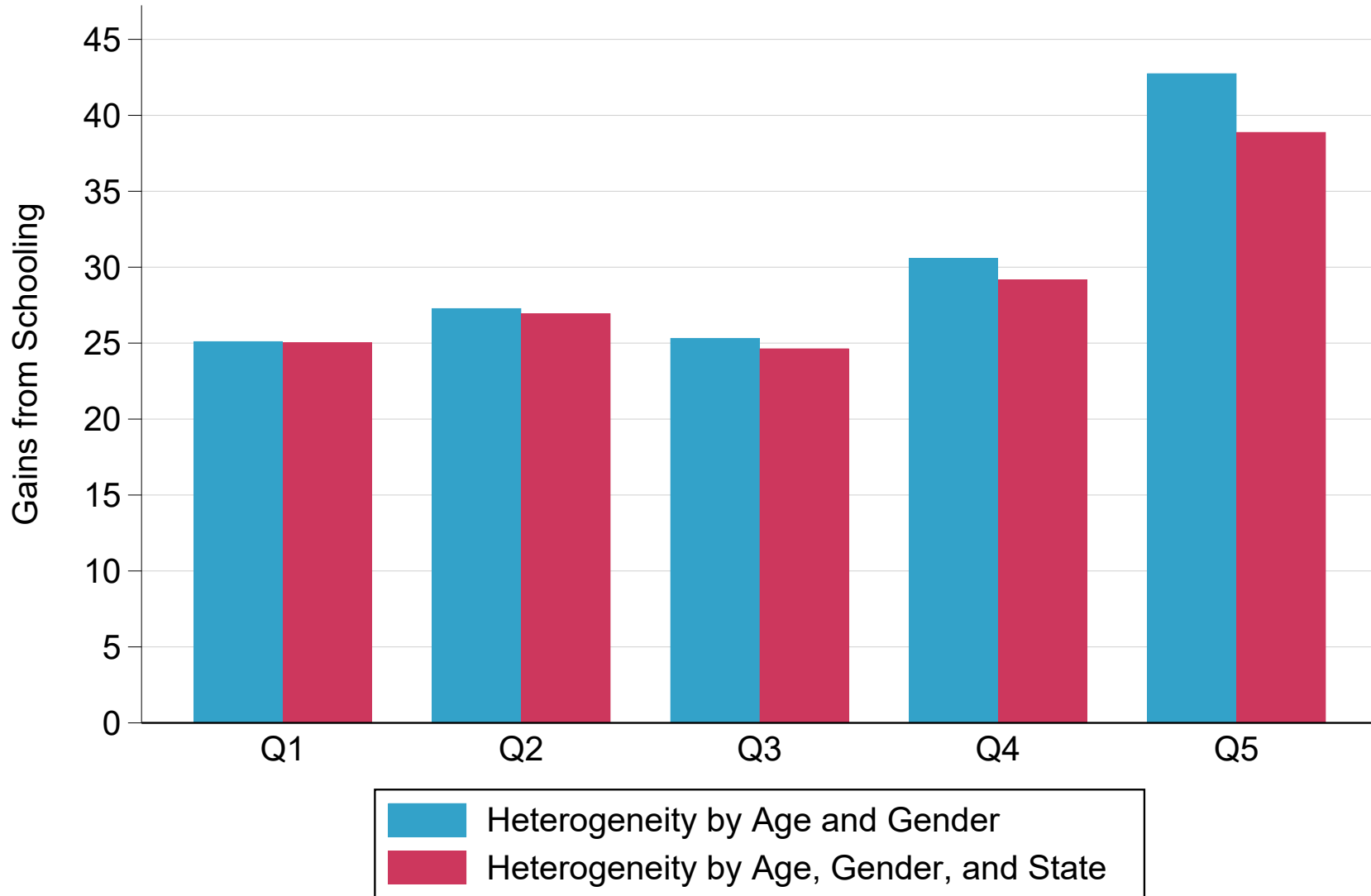
Notes. The figure plots annualized gains from schooling by global income percentile, calculated by taking the percent difference between actual and counterfactual income growth absent educational expansion, and annualizing the resulting figure over the period considered. The analysis focuses on educational expansion among post-1980 generations only, that is, comparing the actual evolution of incomes to a counterfactual in which all post-1980 generations would have reached the education level of the 1980 generation.

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



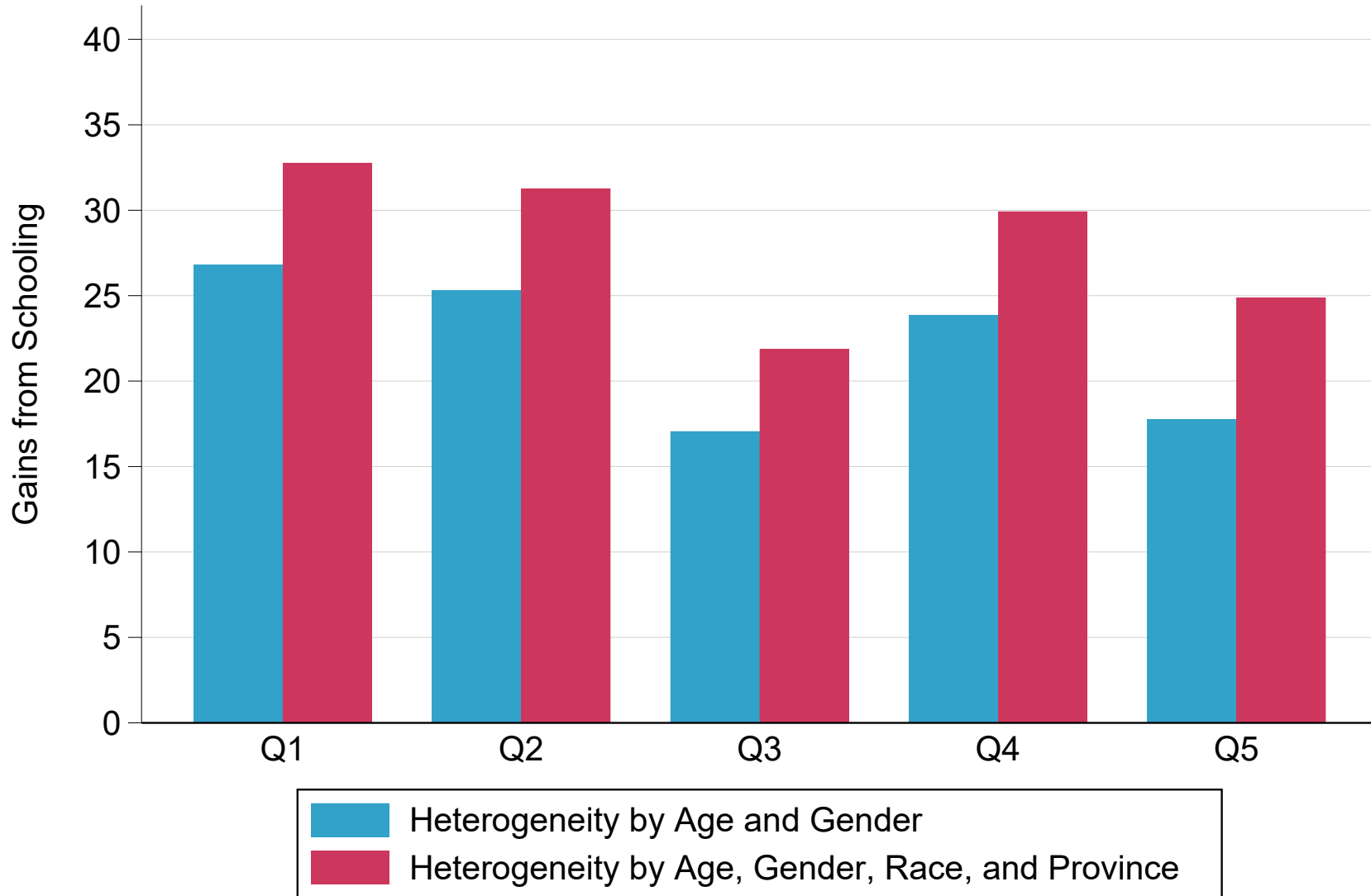
Notes. The figure plots the share of real pretax income growth explained by education by global income percentile. Each point corresponds to the ratio of gains from schooling—equal to actual minus counterfactual income growth absent educational expansion—over actual income growth, calculated for each percentile of the world distribution of income. The analysis focuses on educational expansion among post-1980 generations only, that is, comparing the actual evolution of incomes to a counterfactual in which all post-1980 generations would have reached the education level of the 1980 generation.

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



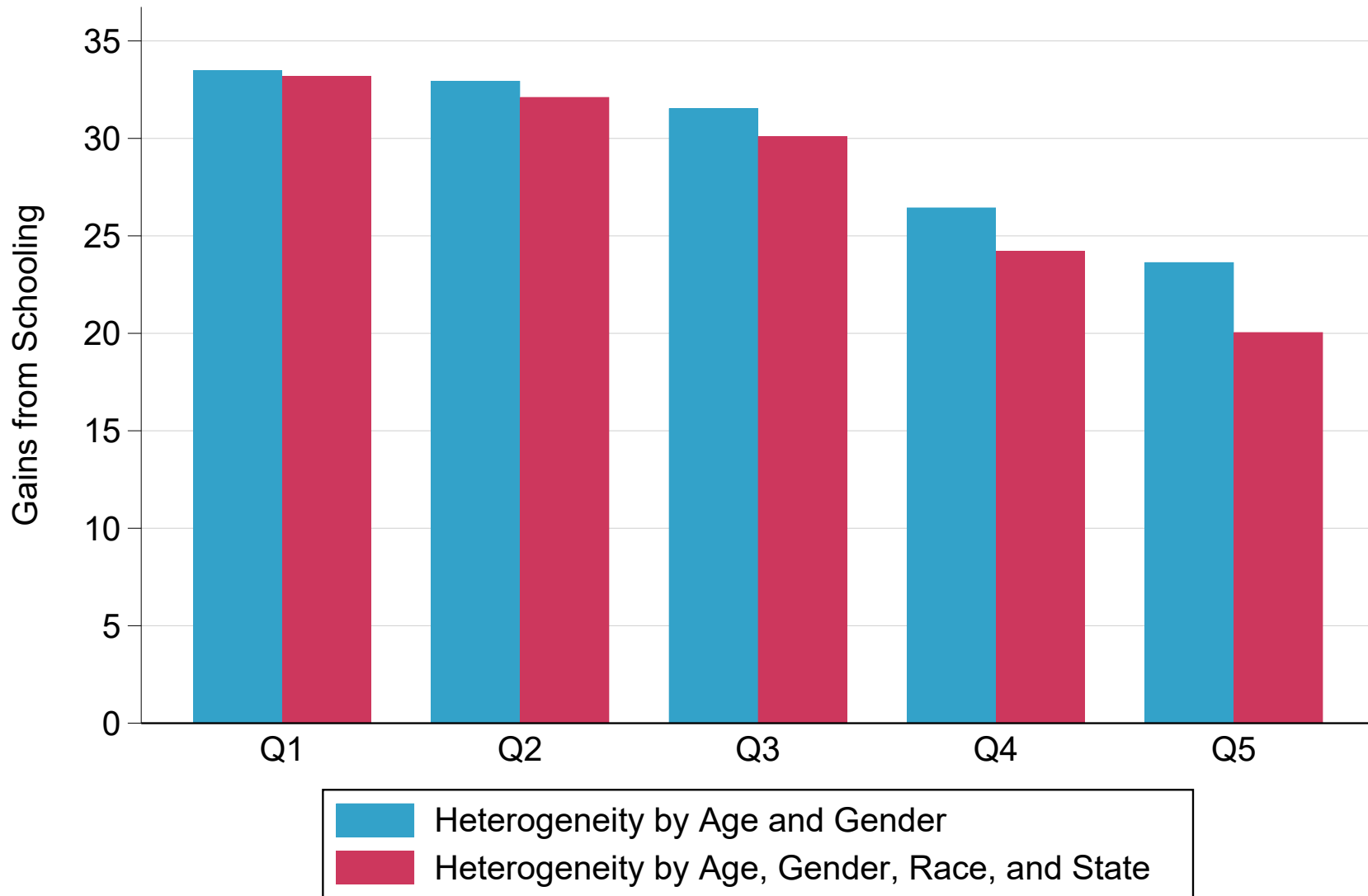
Notes. Educational attainment by age, gender, and state of residence in 1983 is estimated using the 1983 National Sample Survey. Education levels of individuals in the 2019 Periodic Labor Force Survey are then downgraded by age-gender cell (specification 1) or age-gender-state cell (specification 2) until reaching 1983 levels. Their earnings then 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 then 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 then 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%	+40%	15	27%
Bottom 50%	+82%	+49%	33	40%
Bottom 20%	+69%	+36%	33	48%
Next 30%	+84%	+52%	33	39%
Middle 40%	+78%	+55%	23	30%
Top 10%	+37%	+29%	8	21%
Top 1%	+42%	+38%	4	10%
Top 0.1%	+57%	+53%	3	6%
Top 0.01%	+62%	+62%	0.3	0.6%

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%	+17%	77%	+48%	+30%	39%
Bottom 50%	+206%	+79%	61%	+97%	+57%	41%
Bottom 20%	+196%	+74%	62%	+83%	+43%	48%
Next 30%	+208%	+81%	61%	+102%	+62%	39%
Middle 40%	+84%	+7%	91%	+78%	+50%	36%
Top 10%	+61%	+17%	72%	+27%	+16%	42%
Top 1%	+77%	+40%	47%	+26%	+17%	35%
Top 0.1%	+102%	+66%	36%	+20%	+11%	49%
Top 0.01%	+155%	+118%	24%	+5%	+2%	66%

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%	-80%	
Counterfactual	46%	24%	-48%	40%
Global Poverty: \$3.65 / Day				
Actual	60%	25%	-59%	
Counterfactual	60%	45%	-26%	56%
Global Poverty: \$6.85 / Day				
Actual	70%	47%	-32%	
Counterfactual	70%	64%	-9%	73%

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	61%	54%	53%	30%	27%	27%
Bottom 50%	72%	59%	56%	51%	40%	39%
Bottom 20%	85%	71%	66%	61%	48%	46%
Next 30%	70%	57%	54%	49%	39%	38%
Middle 40%	92%	74%	73%	35%	30%	30%
Top 10%	35%	38%	38%	17%	21%	21%
Top 1%	10%	17%	16%	5%	10%	9%
Top 0.1%	7%	9%	8%	5%	6%	6%
Top 0.01%	2%	2%	2%	0.4%	0.6%	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: $\sigma_1 = \sigma_2 = \sigma_3 = 4$. Benchmark: $\sigma_1 = \sigma_2 = \sigma_3 = 6$. 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%	+45%	52	53%
Bottom 50%	+164%	+79%	85	52%
Bottom 20%	+115%	+40%	74	65%
Next 30%	+176%	+89%	87	50%
Middle 40%	+94%	+28%	66	70%
Top 10%	+91%	+54%	37	41%
Top 1%	+131%	+107%	24	18%
Top 0.1%	+173%	+157%	17	10%
Top 0.01%	+278%	+270%	8	3%

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. Firms first pick between below-secondary and above-secondary education, then between workers within each of these two

nests. The production function is thus: $L = \left(A_L L_L^{\frac{\sigma_1 - 1}{\sigma_1}} + A_H L_H^{\frac{\sigma_1 - 1}{\sigma_1}} \right)^{\frac{\sigma_1}{\sigma_1 - 1}}$, 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 $L_H = \left(A_{sec} L_{sec}^{\frac{\sigma_3 - 1}{\sigma_3}} + A_{ter} L_{ter}^{\frac{\sigma_3 - 1}{\sigma_3}} \right)^{\frac{\sigma_3}{\sigma_3 - 1}}$.

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%	+54%	110%	67%
Bottom 20%	+115%	+23%	92%	80%
Next 30%	+176%	+61%	114%	65%
Middle 40%	+94%	+14%	80%	85%
Top 10%	+91%	+42%	49%	54%
Top 1%	+131%	+81%	50%	38%
Top 0.1%	+173%	+111%	62%	36%
Top 0.01%	+278%	+194%	84%	30%

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. The analysis focuses on educational expansion among post-1980 generations only, that is, comparing the actual evolution of incomes to a counterfactual in which all post-1980 generations would have reached the education level of the 1980 generation.

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* (0.699)	10.916*** (1.391)	4.115** (1.949)
Government Effectiveness	0.554* (0.317)	1.961 (1.638)	-2.093 (2.288)
Trade-to-GDP Ratio	-0.547*** (0.150)	-0.624 (0.785)	-3.220*** (1.136)
Internet Usage	0.292 (0.372)	-1.899 (1.961)	1.969 (2.778)
Mobile cellular subscriptions (per 100 people)	-0.045 (0.294)	3.584** (1.535)	1.987 (2.152)
Skill Bias of Technology	-0.124 (0.196)	-1.207 (1.025)	-0.151 (1.425)
Log GDP Per Capita	0.600 (0.539)	-4.813** (2.368)	7.720** (3.392)
Child Population (% Total)	-0.820** (0.347)	-1.457 (1.699)	-13.819*** (2.426)
Constant	12.062*** (0.510)	84.302*** (2.632)	54.134*** (3.714)
Treatment			
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 \(2024\)](#). 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*** (0.246)	6.110*** (1.321)	8.032*** (1.527)
Government Effectiveness	-0.170* (0.097)	0.328 (0.625)	-0.734 (0.669)
Trade-to-GDP Ratio	0.157** (0.061)	0.398 (0.378)	-0.446 (0.353)
Internet Usage	-0.542*** (0.053)	-5.183*** (0.319)	-4.089*** (0.370)
Mobile cellular subscriptions (per 100 people)	0.146*** (0.052)	0.228 (0.307)	1.369*** (0.327)
Log GDP Per Capita	0.137 (0.241)	-2.125 (1.381)	2.472 (1.639)
Child Population (% Total)	-0.527*** (0.147)	3.996*** (0.910)	-8.557*** (1.054)
Constant	12.617*** (0.110)	91.593*** (0.506)	64.970*** (0.801)
Treatment			
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 \(2024\)](#). 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	With Efficiency	Ratio	Without Efficiency	With Efficiency	Ratio
	Gains	Gains		Gains	Gains	
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 With and Without
Relative Efficiency Gains, 2000-2019

	Average			Bottom 50%		
	Without Efficiency	With Efficiency	Ratio	Without Efficiency	With Efficiency	Ratio
	Gains	Gains		Gains	Gains	
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.

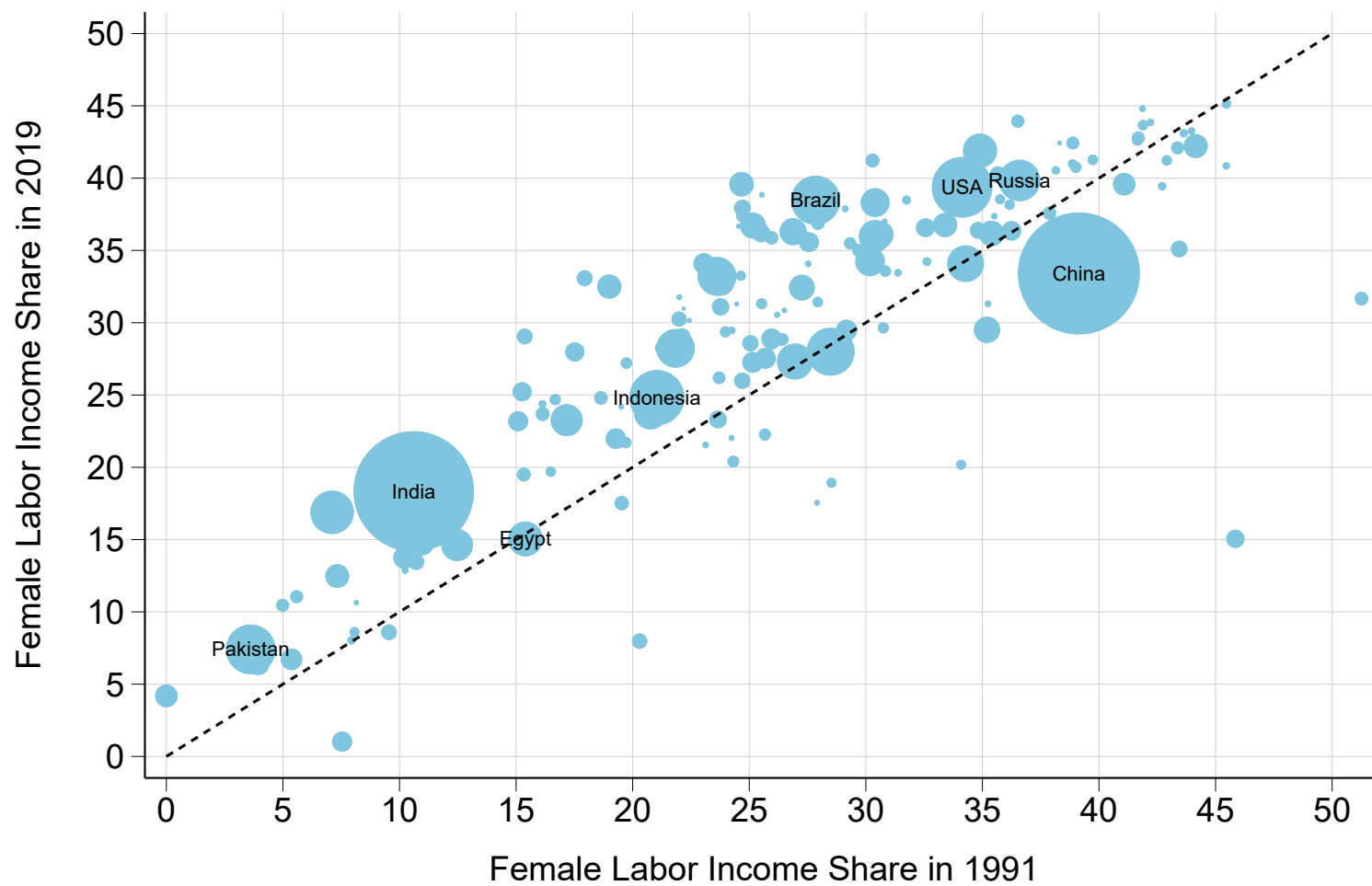
Table A12 – Public Policies and Global Poverty Reduction:
Combining Direct Redistribution and Indirect Investment Benefits from Education (World Bank Data)

	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%	9.5%	-80%	
Posttax Income	44%	4.9%	-89%	46%
Global Bottom 20% Average Income (\$/Day)				
Pretax Income Absent Educational Expansion	0.7	1.2	+73%	
Pretax Income	0.7	2.1	+194%	
Posttax Income	0.8	2.8	+240%	70%
Global Bottom 50% Average Income (\$/Day)				
Pretax Income Absent Educational Expansion	1.3	2.2	+79%	
Pretax Income	1.3	3.8	+205%	
Posttax Income	1.4	5.0	+264%	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, 2024](#)). 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



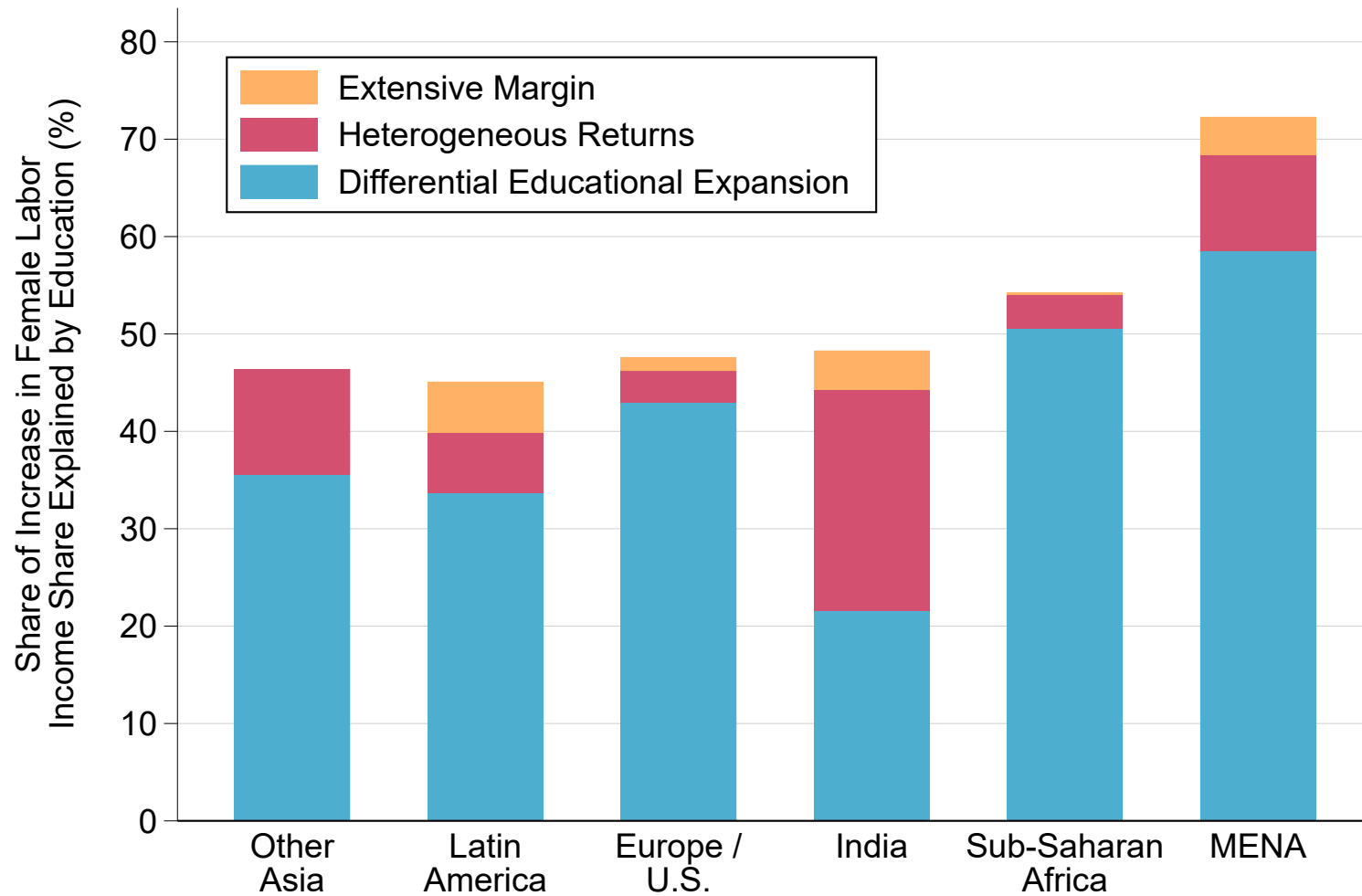
Notes. The figure compares the share of labor income received by women in 1991 (x-axis) and 2019 (y-axis). Data from [Neef and Robilliard \(2021\)](#).

Figure A10 – Returns to Schooling: Men vs. Women



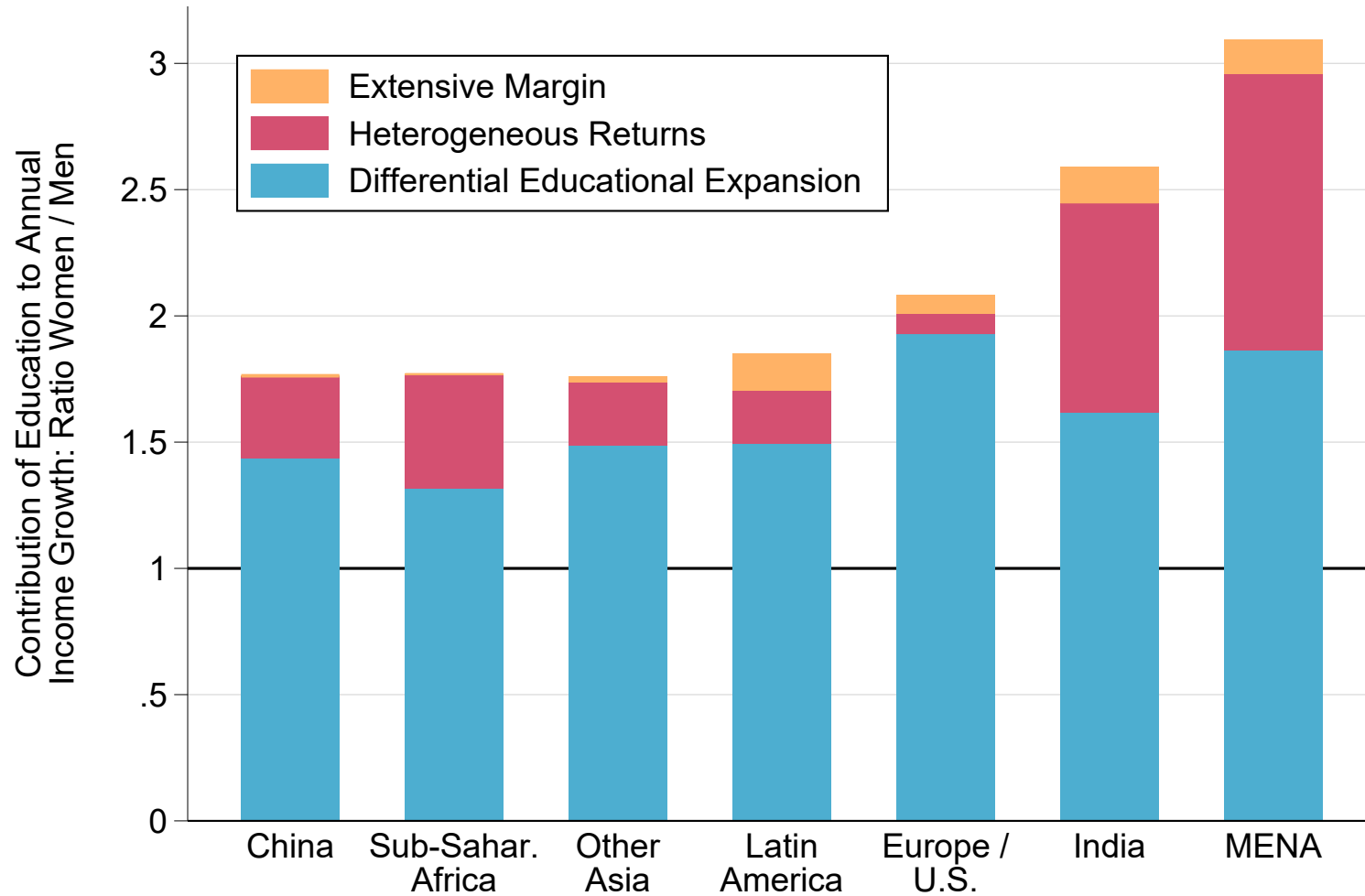
Notes. The figure compares returns to a year of schooling for men (x-axis) and women (y-axis). Estimates correspond to the effect of an additional year of schooling on the log of personal income, estimated using modified Mincerian equations that control for an experience quartic.

Figure A11 – Share of Gender Inequality Reduction Explained by Education by World Region, 1991-2019



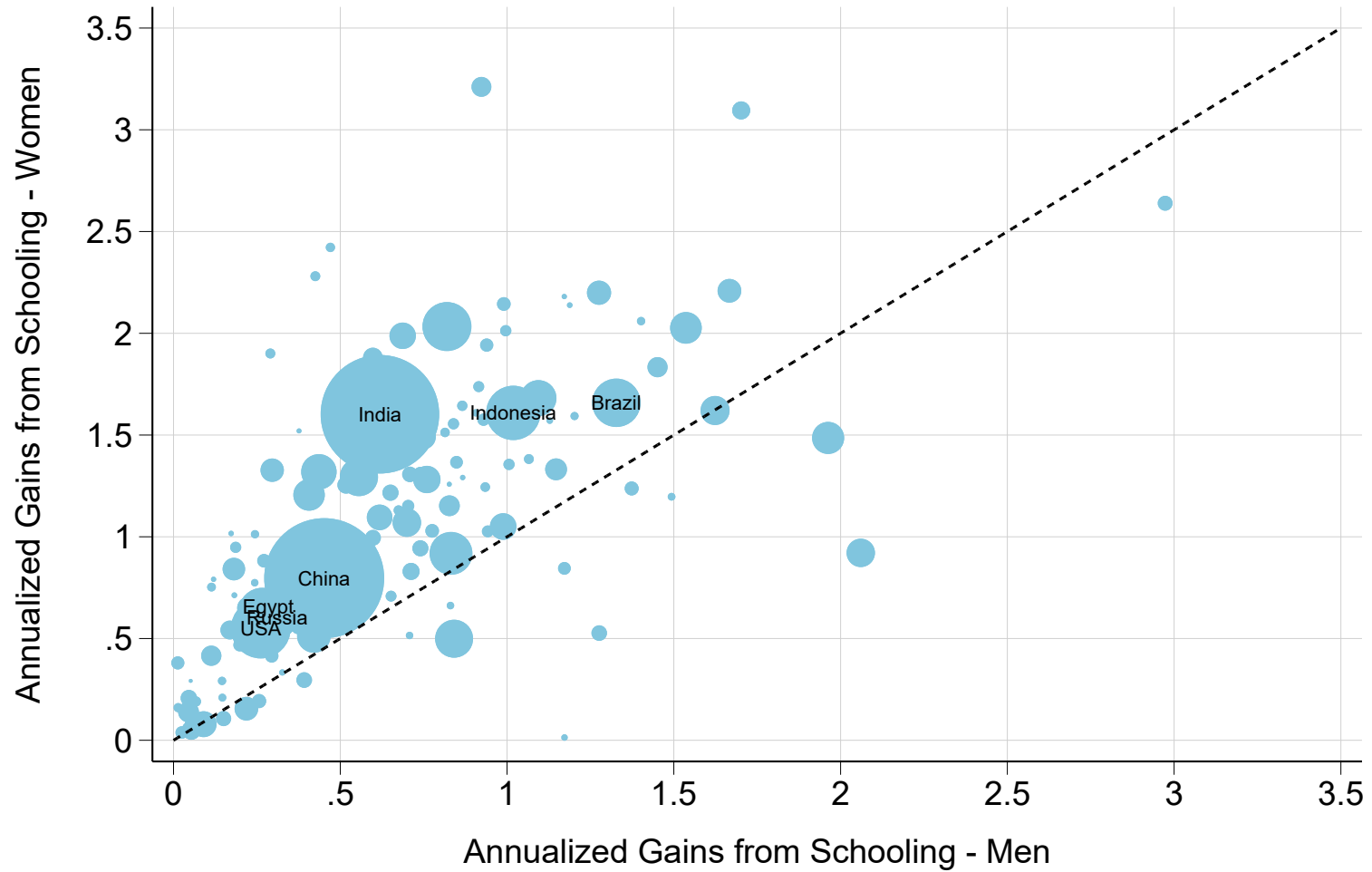
Notes. The figure plots the share of gender labor income inequality reduction explained by education by world region, decomposing it into three components. No educational progress: only account for differential trends in schooling by gender, applying the same returns to schooling to men and women to build the counterfactual. Heterogeneous returns: account for differential returns by gender. Extensive margin: account for differential effects of schooling on employment by gender. Education explains about 50% of improvements in the female labor income share in India since 1991. Most of this effect can be explained by the fact that education increased faster for women than for men (differential educational expansion) and that returns to schooling are higher for women (heterogeneous returns), while differential effects of schooling on employment played a more minor role (extensive margin).

Figure A12 – Contribution of Schooling to Gender Inequality Reduction by World Region
 Gender Ratio of Income Gains from Schooling (Women / Men), 1991-2019



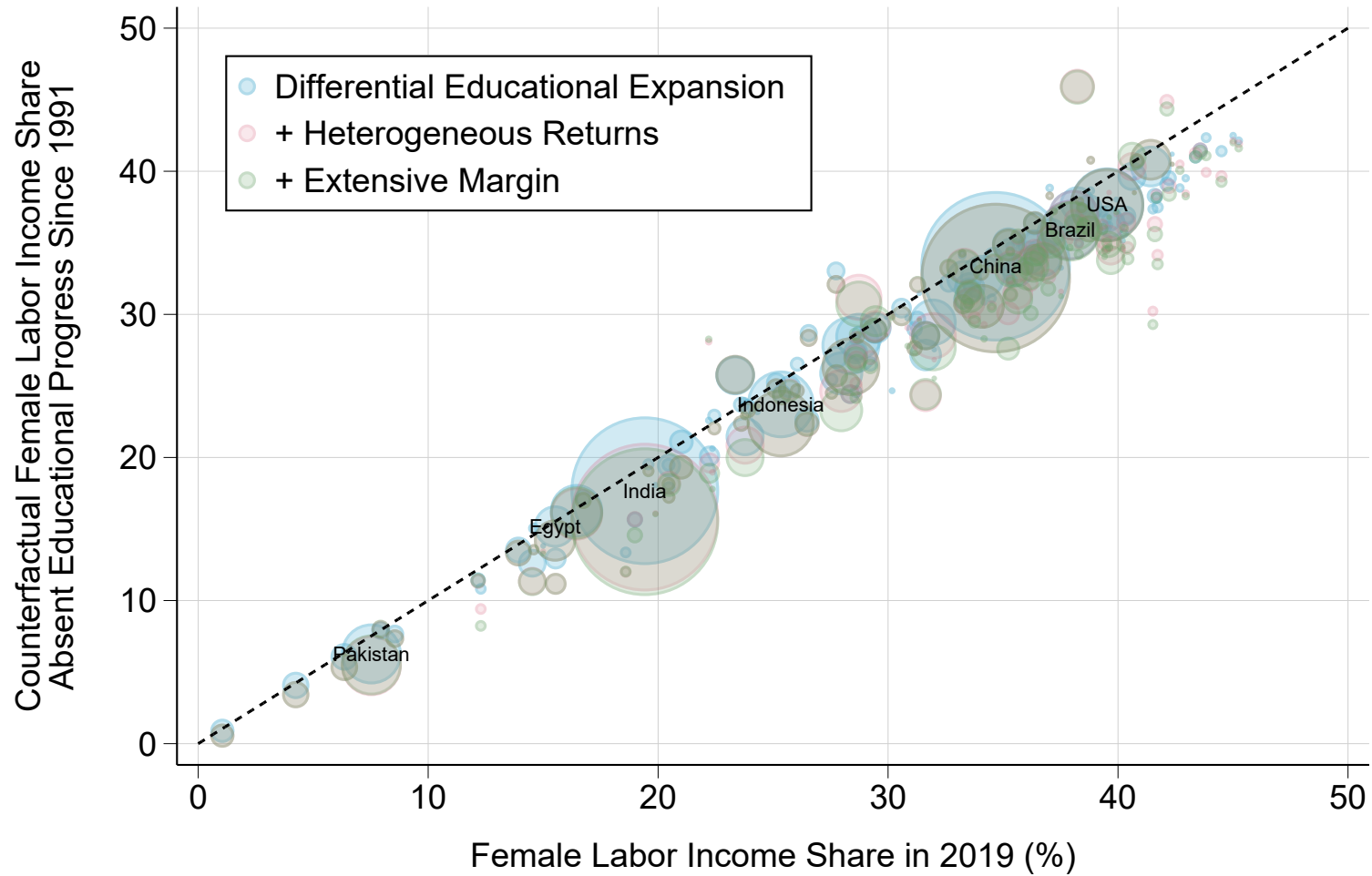
Notes. Each bar corresponds to the ratio of gains from schooling among women over gains from schooling among men. Gains from schooling correspond to the percent difference between actual income and counterfactual income absent educational expansion. No educational progress: only account for differential trends in schooling by gender, applying the same returns to schooling to 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. Population-weighted averages of gains from schooling estimated in each country.

Figure A13 – Annualized Income Gains from Schooling, 1991-2019: Men vs. Women



Notes. The figure plots annualized gains from schooling since 1991 for men (x-axis) and women (y-axis), 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 A14 – Education and Gender Inequality: Actual vs. Counterfactual Female Labor Income Share in 2019



Notes. The figure compares actual female labor income shares in 2019 (x-axis) to the female labor income shares that would have prevailed if education had not improved since 1991 (y-axis). No educational progress: only account for differential trends in schooling by gender, applying the same returns to schooling to 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 A13 – 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%	30.1%	1.5	52%	38%
Counterfactual: + Heterogeneous Returns	28.7%	29.8%	1.1	63%	48%
Counterfactual: + Extensive Margin	28.7%	29.5%	0.9	71%	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. No educational progress: only account for differential trends in schooling by gender, applying the same returns to schooling to 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 A14 – 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%	25.8%	4.9	24%
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. No educational progress: only account for differential trends in schooling by gender, applying the same returns to schooling to 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 A15 – Education and Global Gender Inequality, 1991-2019:
By Specification of Employment Effects

	Share Explained (World)	Share Explained (Cross-Country Average)
No Employment Effect	71%	58%
Benchmark: OLS Employment Effects	78%	59%
Alternative: +4pp. per Year	97%	68%
Alternative: +6pp. per Year	111%	73%
Alternative: +8pp. per Year	124%	76%
Alternative: +15% per Year	117%	70%
Alternative: +20% per Year	130%	73%
Alternative: +25% per Year	143%	75%

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 G5.

Table A16 – 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	1.8	2.2	2.0	>100%	>100%	>100%
Europe / U.S.	1.9	1.5	2.5	48%	33%	50%
India	2.6	4.8	4.3	48%	50%	68%
Latin America	1.8	1.9	2.1	45%	44%	44%
MENA	3.1	2.9	2.6	72%	73%	66%
Other Asia-Pacific	1.8	1.7	1.8	42%	59%	48%
Sub-Saharan Africa	1.8	1.8	1.8	54%	47%	44%
World Average	2.0	2.4	2.5	59%	59%	62%

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.

B. Estimation Details

B.1. Estimation Steps

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 return to moving from s_1 to s_2 . Then, the counterfactual income of individual i is:

$$\tilde{y}_i = \exp[\ln(y_i) - r_{s_1, s_2}] \quad (1)$$

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:

$$\ln y_i = \alpha + \beta^{pri} D_i^{pri} + \beta^{sec} D_i^{sec} + \beta^{ter} D_i^{ter} + X_i \beta + \varepsilon_i \quad (2)$$

Which implies that:

$$r_{non,pri} = \beta^{pri} \quad (3)$$

$$r_{pri,sec} = \beta^{sec} - \beta^{pri} \quad (4)$$

$$r_{sec,ter} = \beta^{ter} - \beta^{sec} \quad (5)$$

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 2.2, 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:

$$\underline{r}_{s_1,s_2} = r_{s_1,s_2}(L) \quad (6)$$

With $L = (L_1, \dots, L_m)$ 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:

$$\bar{r}_{s_1,s_2} = r_{s_1,s_2}(\tilde{L}) \quad (7)$$

With $\tilde{L} = (\tilde{L}_1, \dots, \tilde{L}_m)$ 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:

$$\bar{r}_{s_1,s_2} = r_{s_1,s_2}(L) - \frac{1}{\sigma_{s_1,s_2}} \Delta \ln \left(\frac{L_2}{L_1} \right) \quad (8)$$

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: $\bar{r}_{s_1,s_2} > r_{s_1,s_2}$.

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

$$\bar{r}_{non,pri} = r_{non,pri}(L) - \frac{1}{\sigma_3} \Delta \ln \left(\frac{L_{pri}}{L_{non}} \right) \quad (9)$$

$$\bar{r}_{pri,sec} = r_{pri,sec}(L) - \frac{1}{\sigma_2} \Delta \ln \left(\frac{L_{sec}}{L_{pri}} \right) \quad (10)$$

$$\bar{r}_{sec,ter} = r_{sec,ter}(L) - \frac{1}{\sigma_1} \Delta \ln \left(\frac{L_{sec}}{L_{ter}} \right) \quad (11)$$

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:

$$\Delta \log \left(\frac{w_{pri}}{w_{non}} \right) = -\frac{1}{\sigma_3} \Delta \log \left(\frac{L_{pri}}{L_{non}} \right) \quad (12)$$

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

$$\Delta \log \left(\frac{w_{sec}}{w_{\bar{sec}}} \right) = -\frac{1}{\sigma_2} \Delta \log \left(\frac{L_{sec}}{L_{\bar{sec}}} \right) \quad (13)$$

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

$$\Delta \log \left(\frac{w_{ter}}{w_{\bar{ter}}} \right) = -\frac{1}{\sigma_1} \Delta \log \left(\frac{L_{ter}}{L_{\bar{ter}}} \right) \quad (14)$$

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.¹ I then calculate

¹To 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.

counterfactual total pretax income of generalized percentile p as:

$$\tilde{Y}^p = Y_K^p + \psi^p Y_L^p \quad (15)$$

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.

B.2. Elasticities of Substitution

A key parameter of the model is the elasticity of substitution between skill groups, which governs the magnitude of supply effects. 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, 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 $\sigma_1 = \sigma_2 = \sigma_3 = 6$, in line with 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.

B.3. Aggregate and Individual Returns to Schooling: CES Simulations

B.3.1. 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.3.2. 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:

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

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:

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

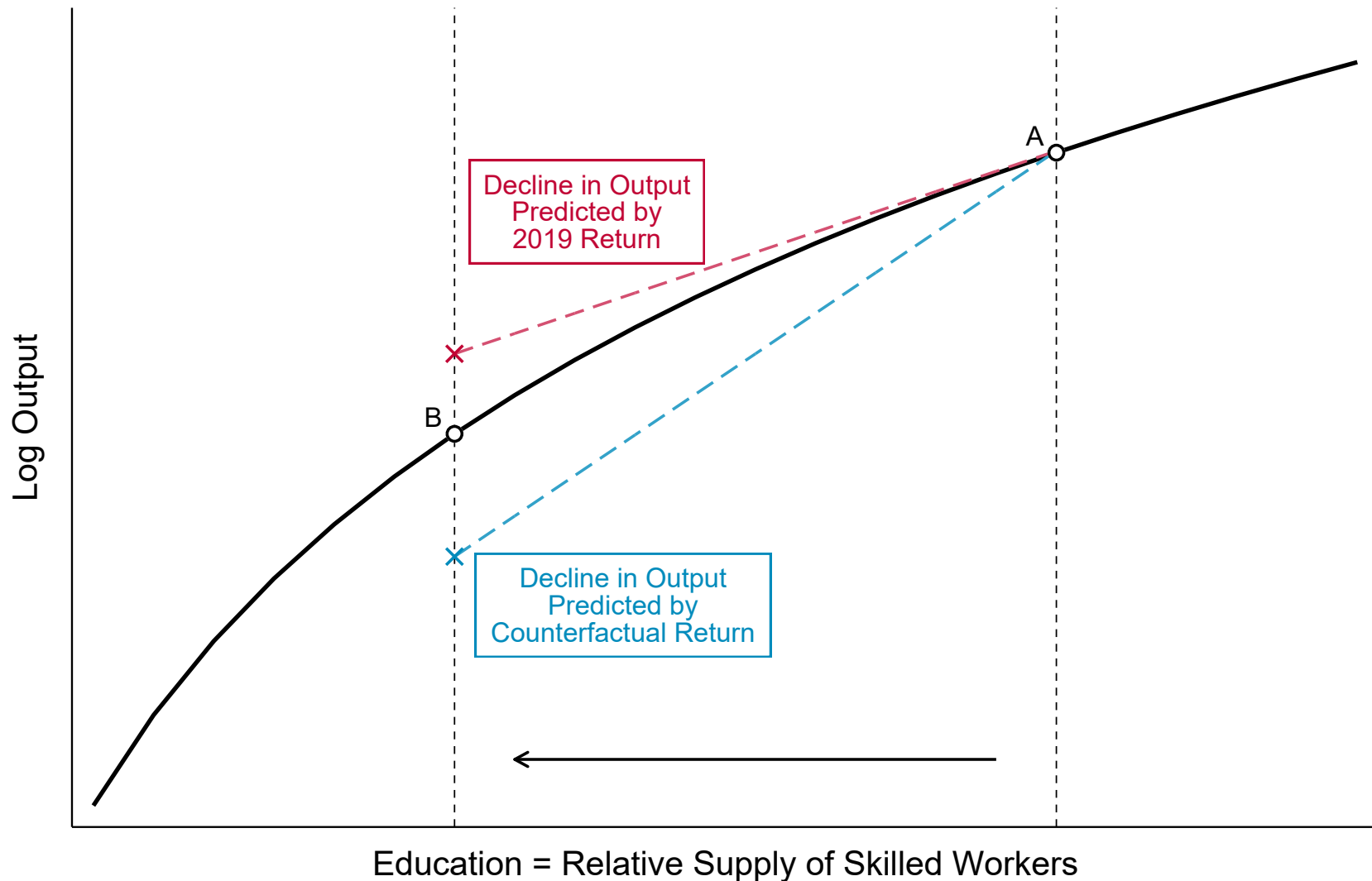
A weight of 0.5 means that the average of initial and final returns captures the true effect of educational expansion on output. A weight greater than 0.5 means that we should give more importance to initial returns in the estimation.

B.3.3. 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.

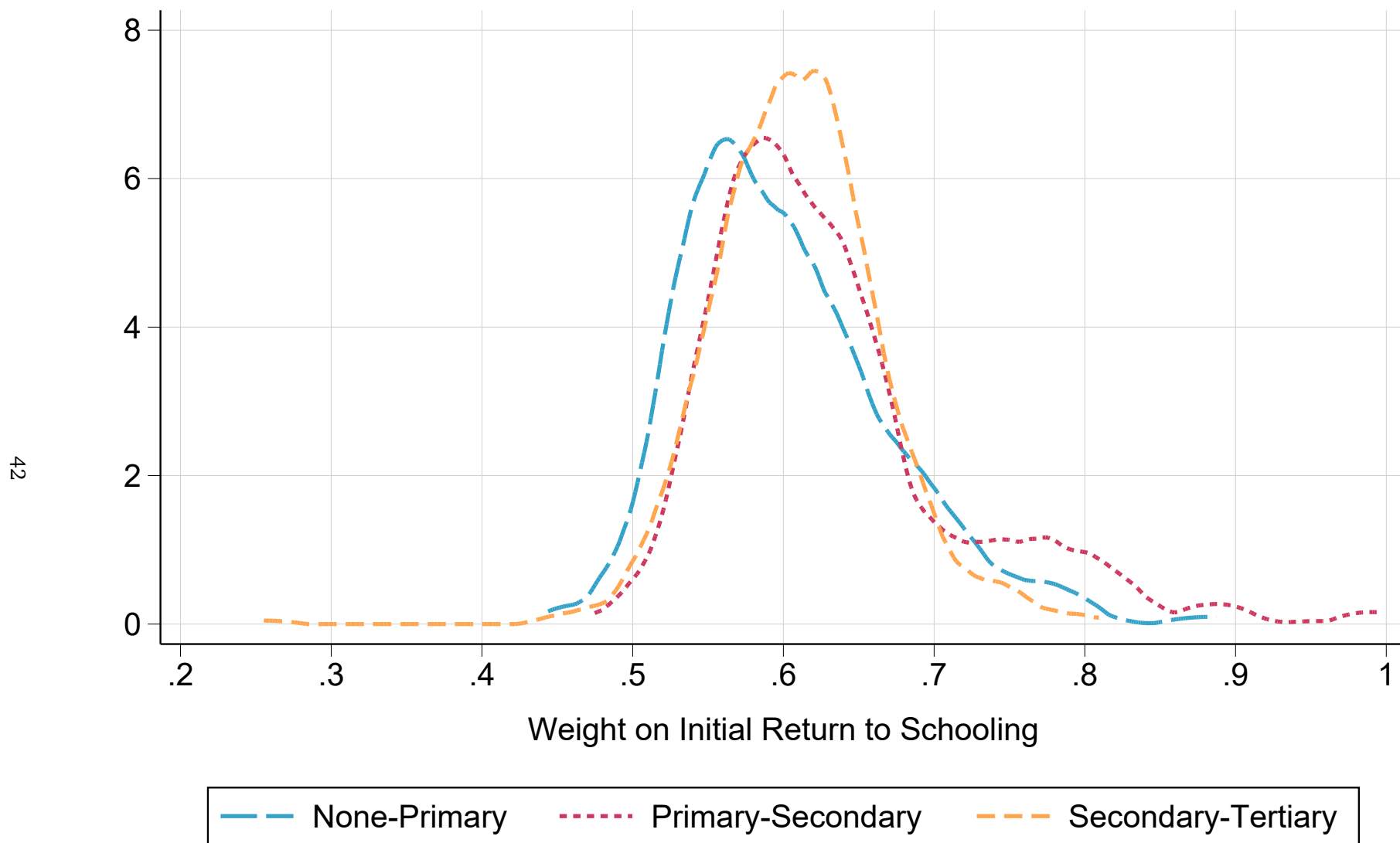
Figure B10 shows how the weight on initial returns varies across specific combinations of relative skill efficiencies and the elasticity of substitution. The weight is higher for lower elasticities of substitution and for higher levels of relative skill efficiencies. For a long-run elasticity of substitution 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 Return to Schooling: Graphical Illustration



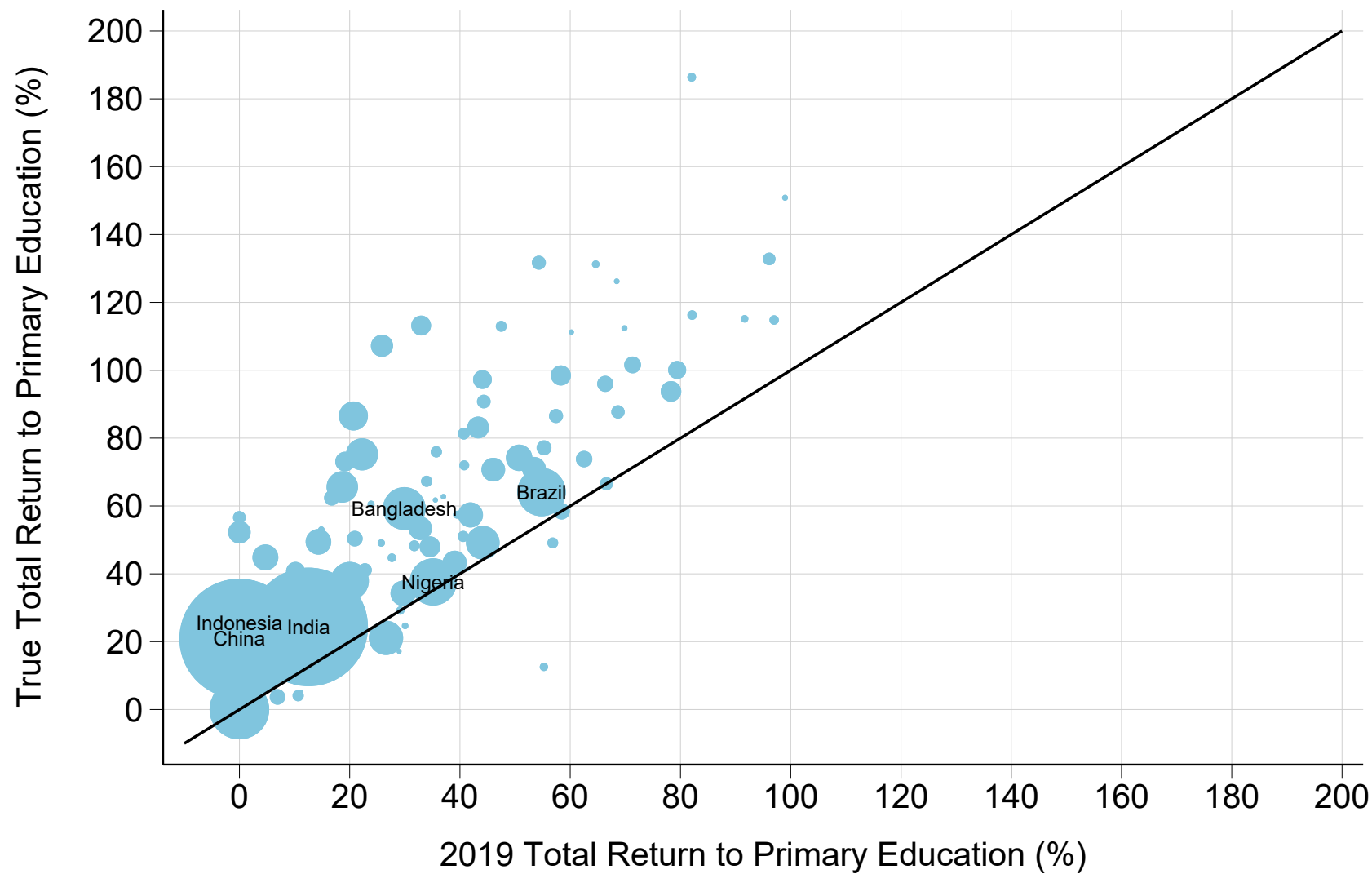
Notes. The figure provides a graphical illustration of why true returns to schooling lie in-between initial and final returns. The objective is to estimate the effect on output of reducing education—moving from point A to point B. One option is to reduce output by the return to schooling observed before reducing education—the 2019 return—, corresponding to the derivative of the production function at point A. The result obtained is materialized by the upper dashed line (“decline in output predicted by 2019 return”). Another option is to reduce output by the return to schooling observed after reducing education—the counterfactual return—, corresponding to the derivative of the production function at point B. The result obtained is materialized by the lower dashed line (“decline in output predicted by counterfactual return”). As shown in the figure, the true return that should be used, materialized by the line going from A to B, falls in-between the 2019 and counterfactual returns.

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



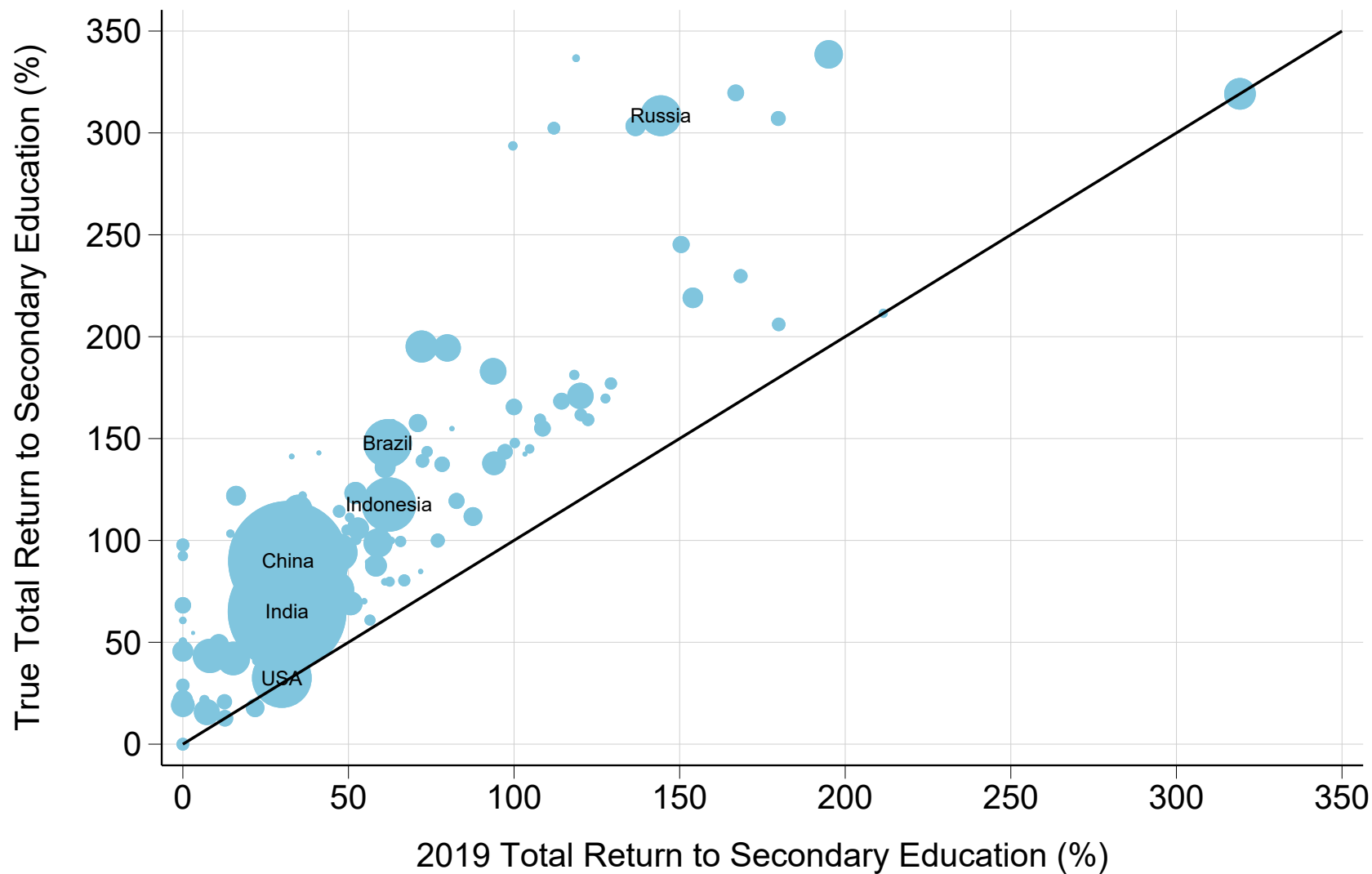
Notes. The figure plots the empirical distribution of optimal weights on initial returns to schooling required to estimate the true effect of educational expansion on output, for each of the three nests of the production function. Estimates assume an elasticity of substitution of 6.

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



Notes. The figure compares the return to primary education observed in 2019 (the final return) to the true return that should be used to predict the decline in output resulting from bringing back education to its 1980 levels in each country. True returns are estimated assuming an elasticity of substitution of 6.

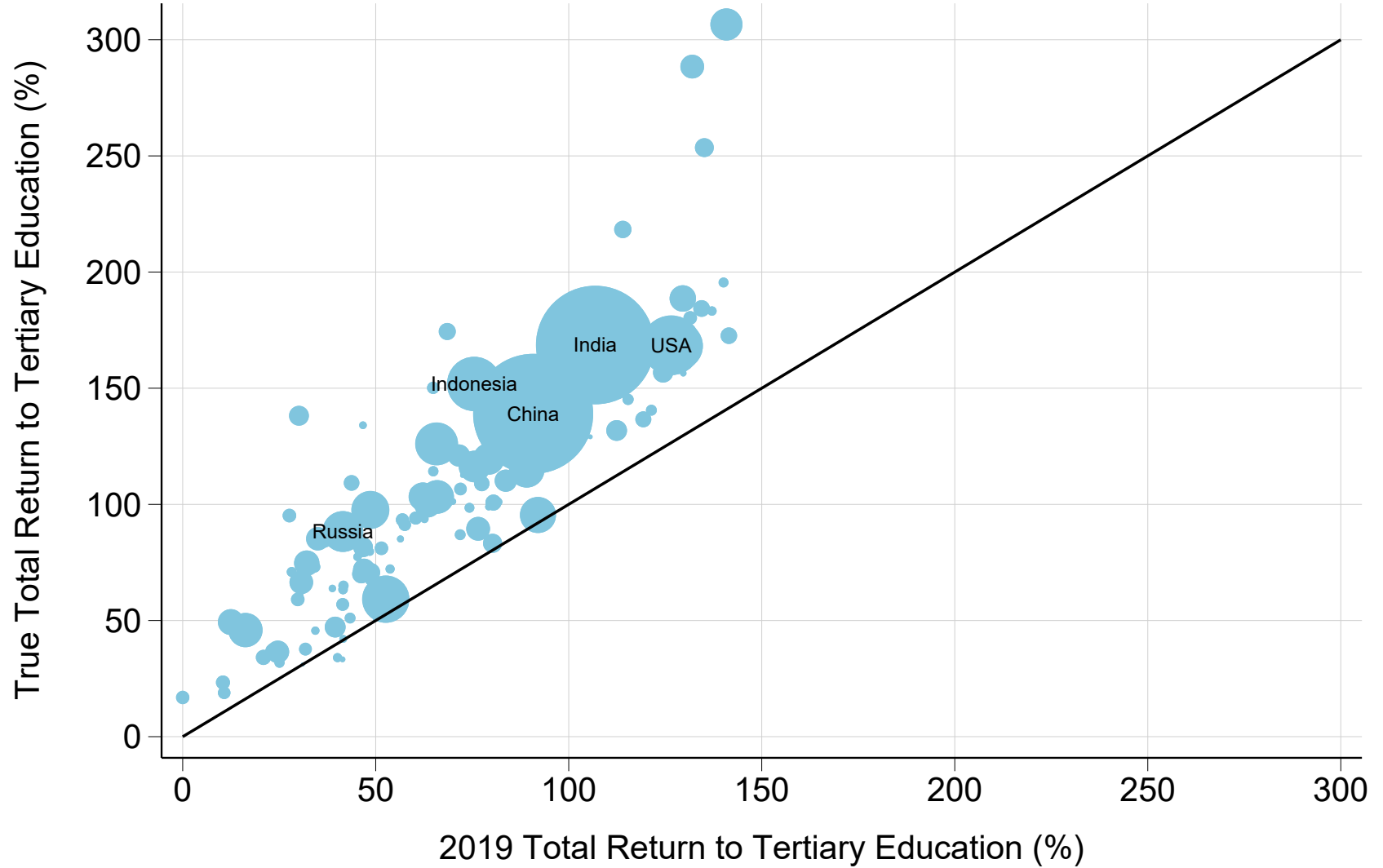
Figure B4 – Returns to Schooling: 2019 vs. True Total Return to Secondary Education



44

Notes. The figure compares the return to secondary education observed in 2019 (the final return) to the true return that should be used to predict the decline in output resulting from bringing back education to its 1980 levels in each country. True returns are estimated assuming an elasticity of substitution of 6.

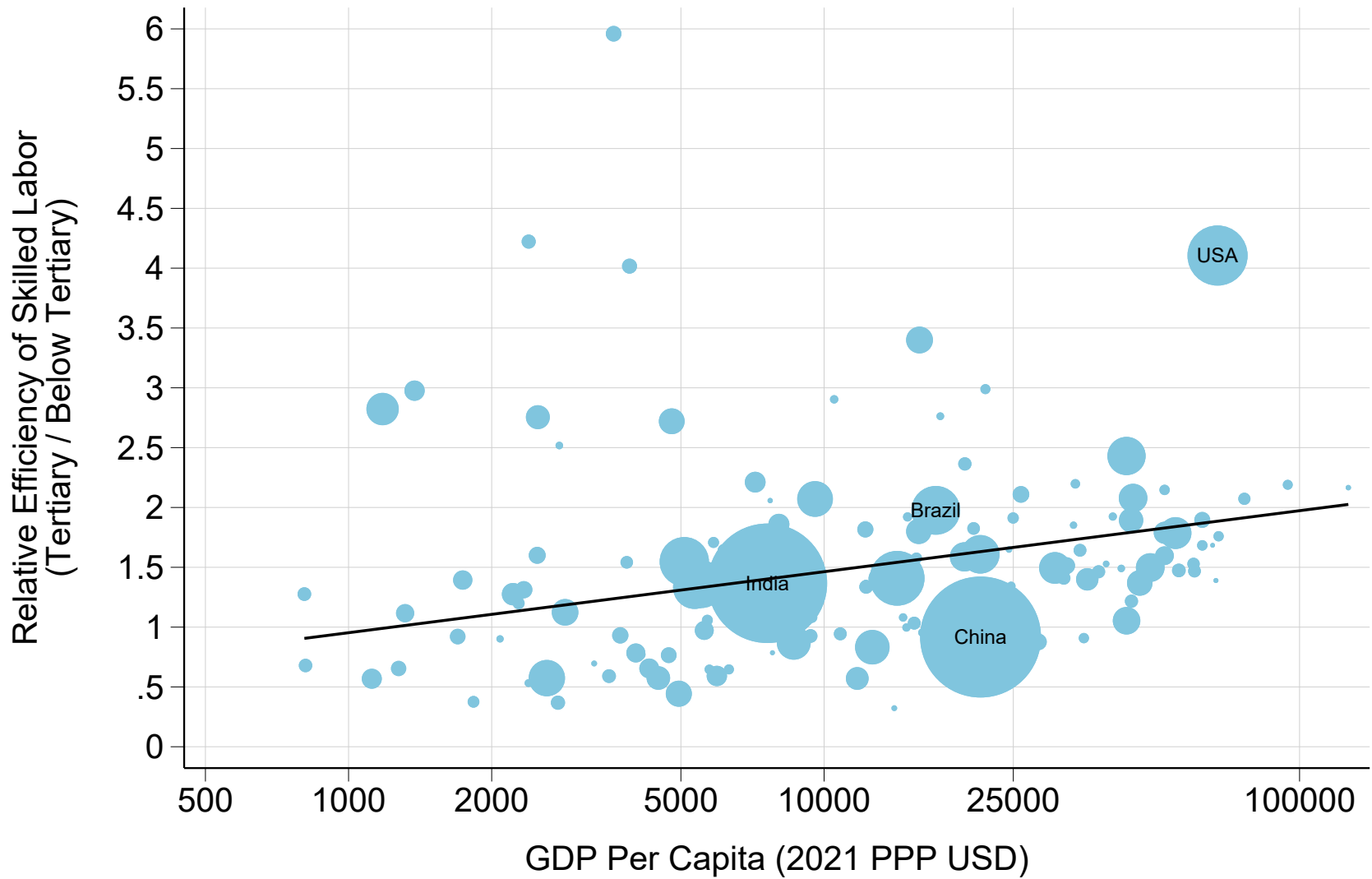
Figure B5 – Returns to Schooling: 2019 vs. True Total Return to Tertiary Education



45

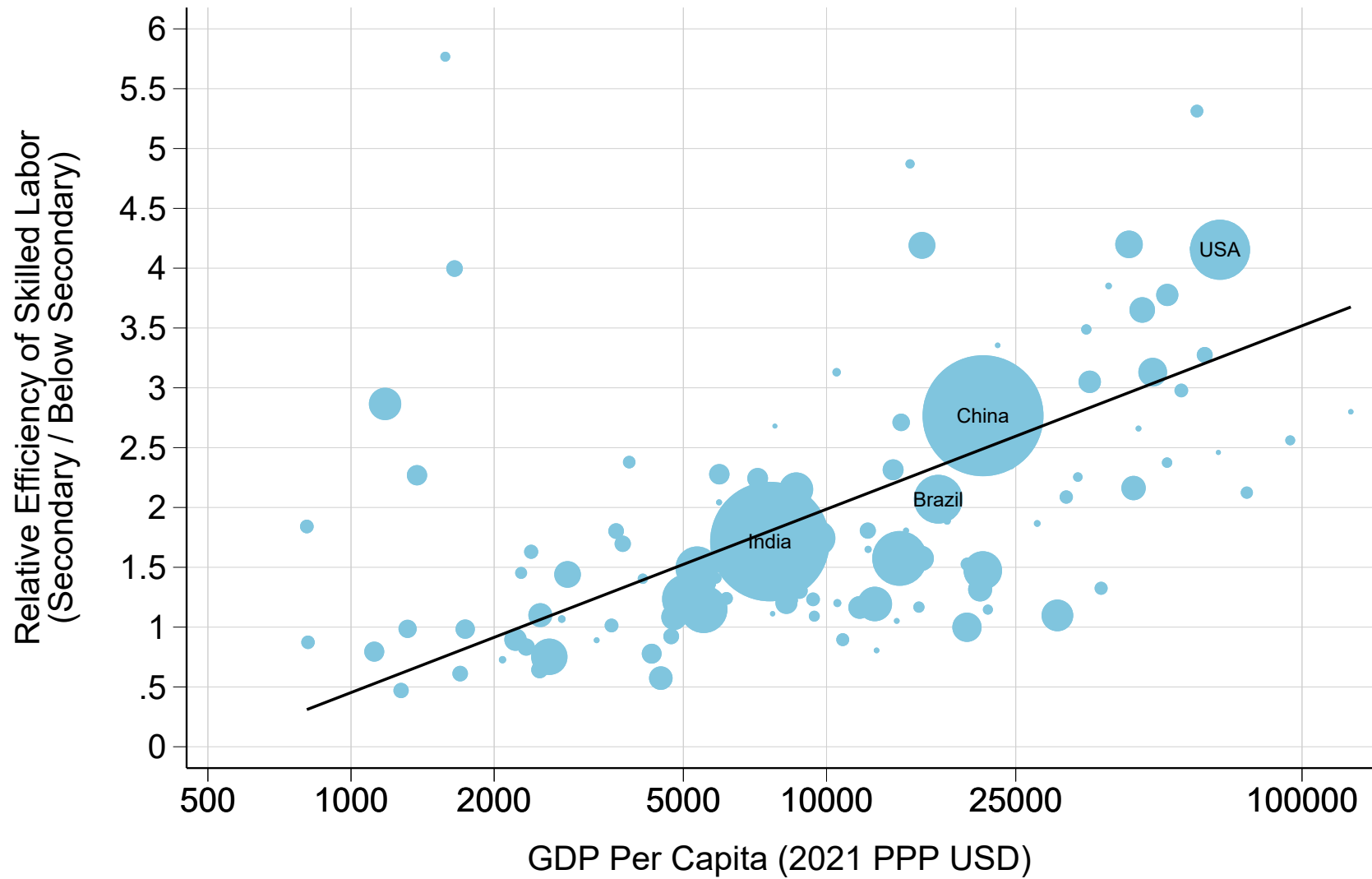
Notes. The figure compares the return to tertiary education observed in 2019 (the final return) to the true return that should be used to predict the decline in output resulting from bringing back education to its 1980 levels in each country. True returns are estimated assuming an elasticity of substitution of 6.

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



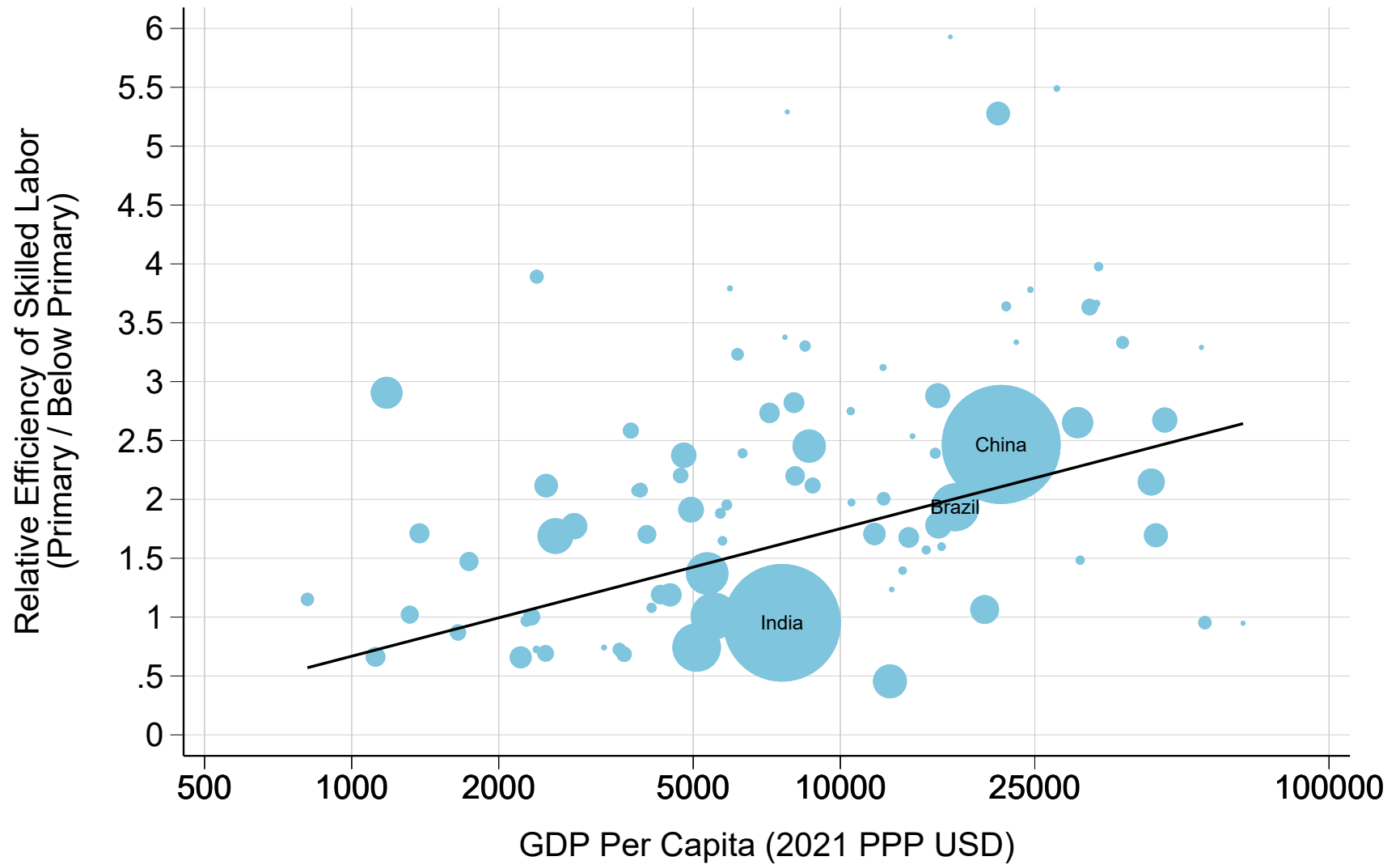
Notes. The figure plots the relative efficiency of tertiary-educated workers relative to non-tertiary-educated workers versus GDP per capita in 2019. The relative efficiency of tertiary-educated workers is recovered from the production function outlined in 3.4, combining data on educational attainment, the return to tertiary education in each country, and an elasticity of substitution of 6. See Rossi (2022) for comparable evidence on a subset of countries.

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



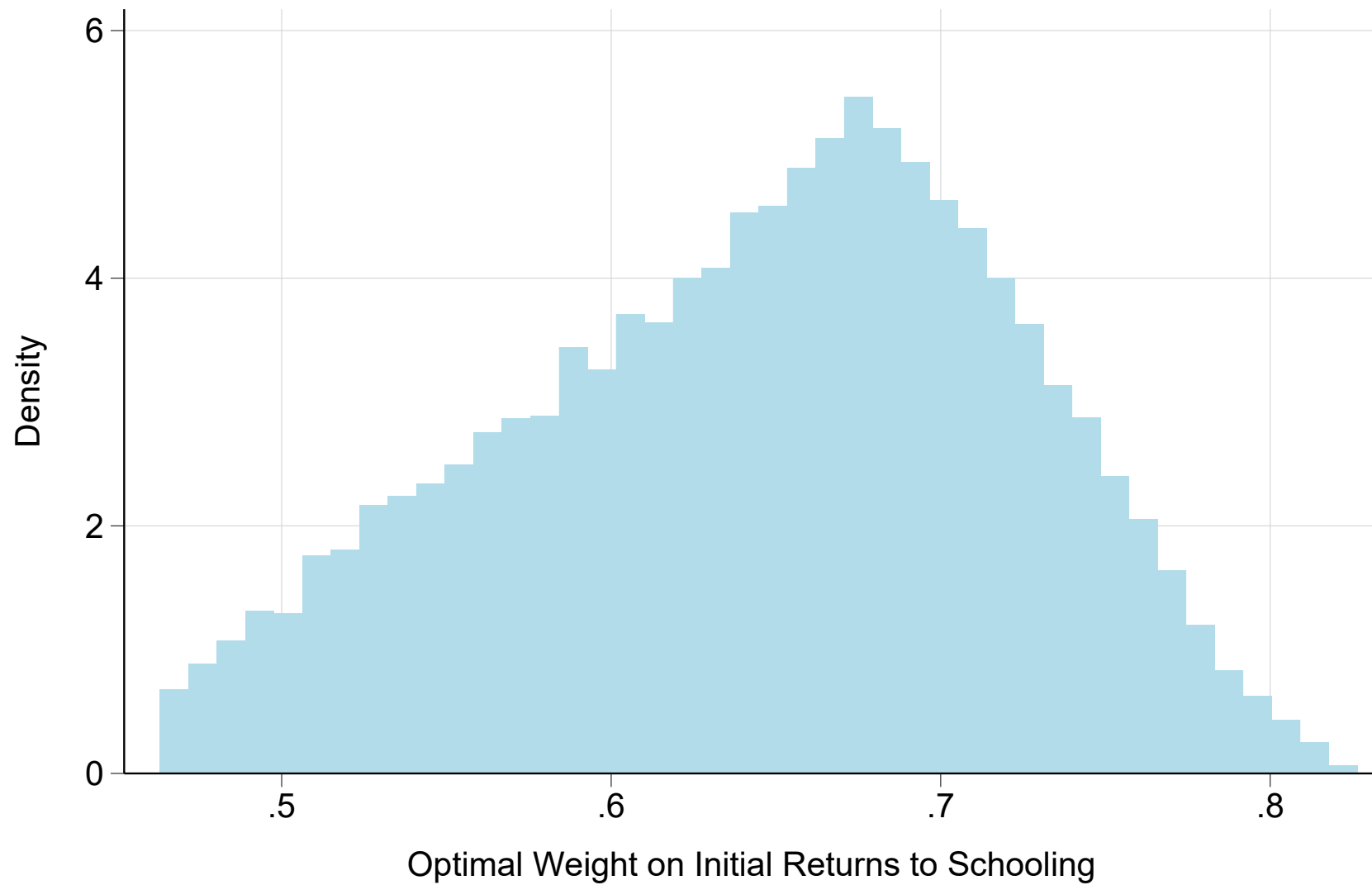
Notes. The figure plots the relative efficiency of secondary-educated workers relative to below-secondary-educated workers versus GDP per capita in 2019. The relative efficiency of secondary-educated workers is recovered from the production function outlined in 3.4, combining data on educational attainment, the return to secondary education in each country, and an elasticity of substitution of 6. See Rossi (2022) for comparable evidence on a subset of countries.

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



Notes. The figure plots the relative efficiency of primary-educated workers relative to those with no schooling versus GDP per capita in 2019. The relative efficiency of primary-educated workers is recovered from the production function outlined in 3.4, combining data on educational attainment, the return to primary education in each country, and an elasticity of substitution of 6. See Rossi (2022) for comparable evidence on a subset of countries.

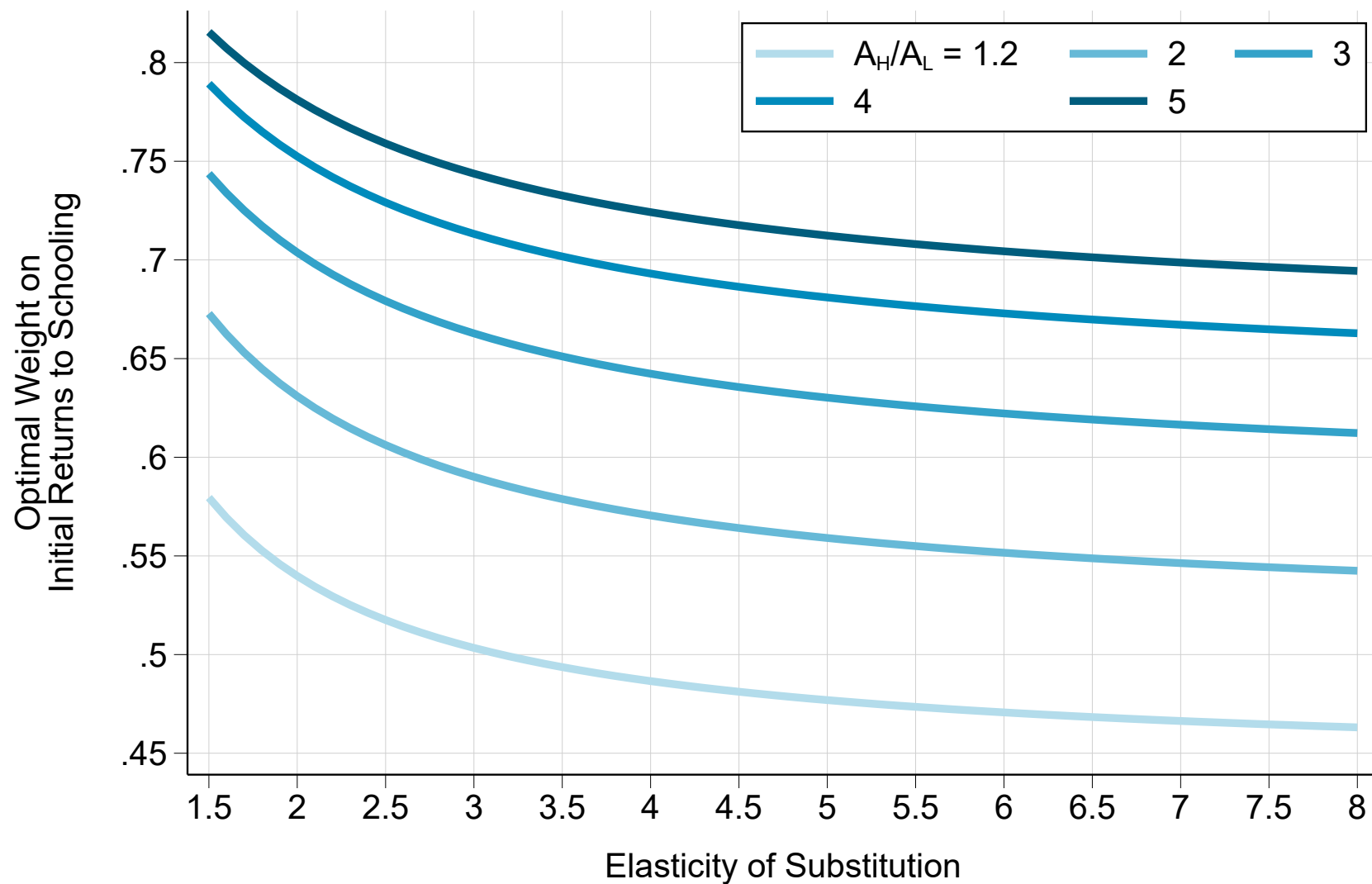
Figure B9 – Returns to Schooling: Simulated Distribution of Optimal Weights on Initial Returns



49

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



50

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.

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:

$$\ln y_{rt}^i = \gamma_0^i + \gamma_1^i S_{rt} + X_{rt} \beta + \delta_r + \delta_t + \varepsilon_{rt} \quad (16)$$

$$S_{rt} = \alpha_0 + \alpha_1 Z_{rt} + \eta_{rt} \quad (17)$$

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} 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 16 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 to estimate human capital externalities, in particular in the case of U.S. compulsory schooling laws ([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 [6a](#) 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 is 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.² 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 between all members of the household.

C.3.3. Empirical Specification

The empirical specification corresponds to the one in equation 16, 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 1978.³ 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

²Some 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.

³Given 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.

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.

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 [6b](#) plots the coefficients by quintile when the return to schooling

is set to 12% and the elasticity of substitution to 4.⁴ 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.

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 ([Stephens and Yang, 2014](#); [Clay, Lingwall, and Stephens, 2021](#)), elasticities of substitution between skill groups ([Ciccone and Peri, 2006](#)), and human capital externalities ([Ciccone and Peri, 2006](#); [Iranzo and Peri, 2009](#); [Guo, Roys, and Seshadri, 2018](#)).

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

⁴[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 high 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.

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 \(2000\)](#).⁵ To estimate the impact of schooling on the distribution of income, I rely on the 1940 to 2000 census microdata samples 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 :

$$S_{st} = \pi_0 + \pi_1 \sum_c \sum_{s'} N_{css't} RS_{cs'} + \theta_s + \theta_t + u_{st} \quad (18)$$

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 [Stephens and Yang, 2014](#); [Clay, Lingwall, and Stephens, 2021](#)). 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

⁵Consider 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.

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 \times 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.

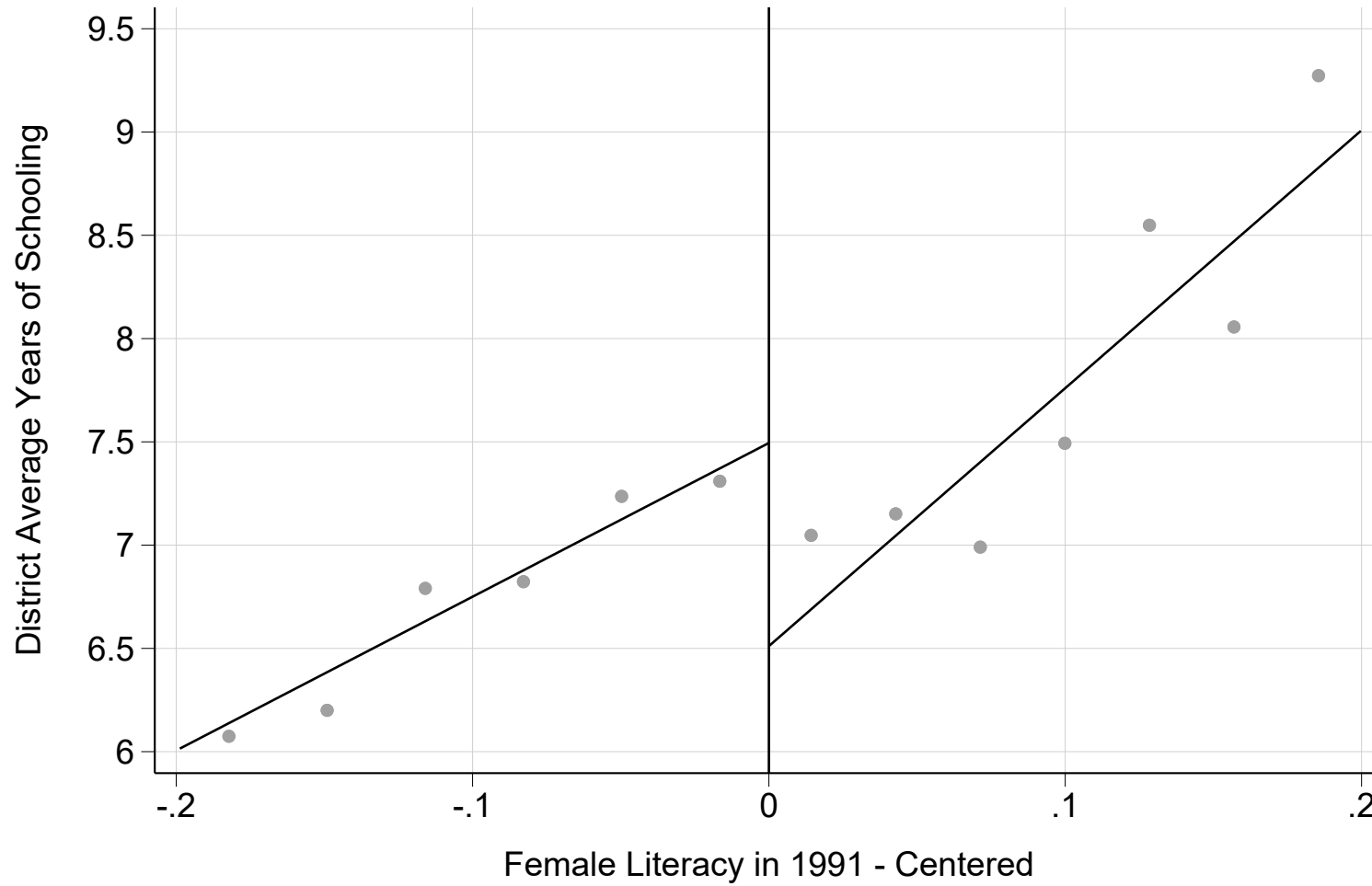
Table C7 compares observed and simulated effects of the policy. Figure 6c plots the coefficients by quintile when the return to schooling is set to 12% and the elasticity of substitution to 4, as in the previous case studies. The simulation is done by upgrading the education of randomly sampled individuals 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

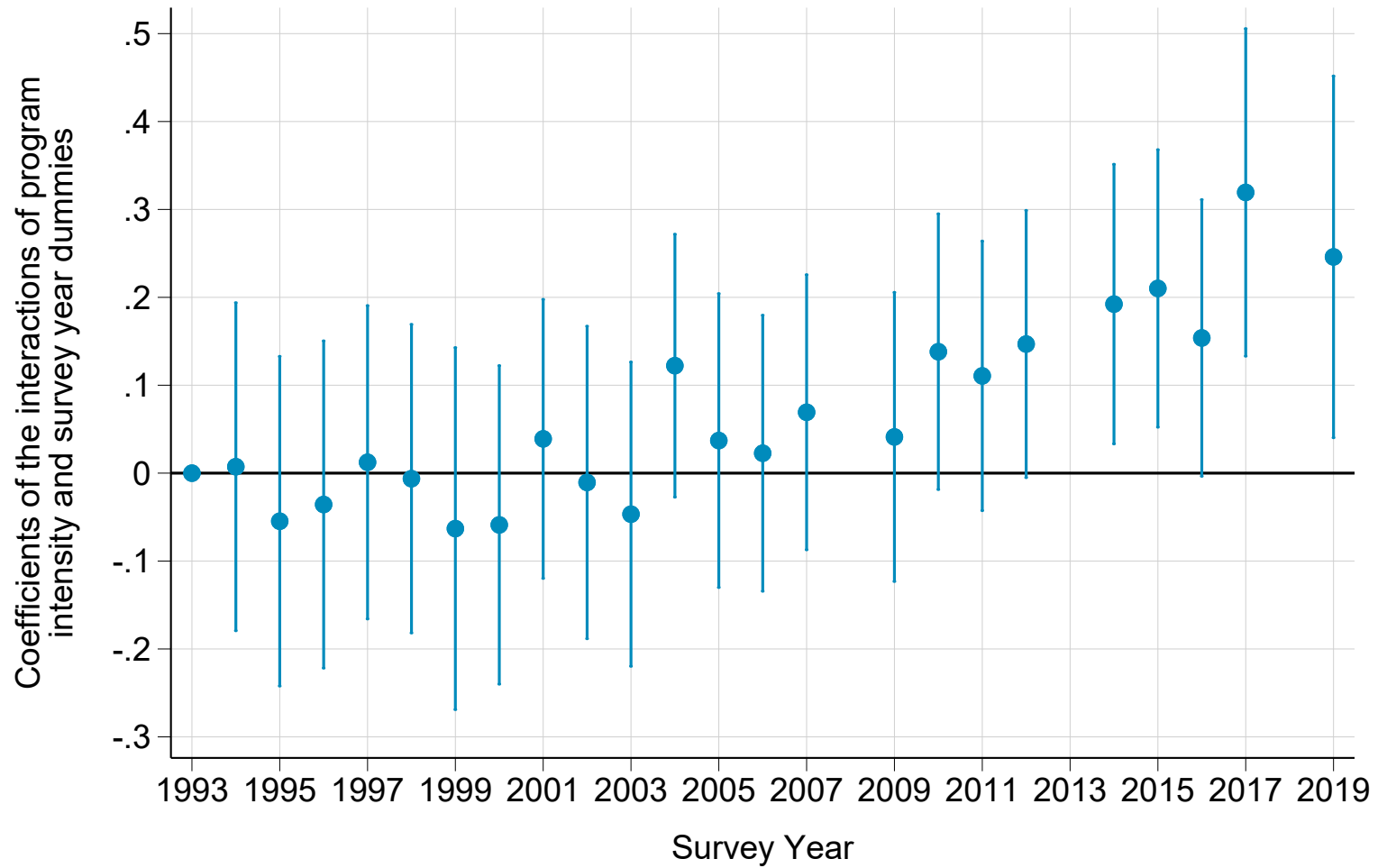
Notes. The table reports the effect of district average years of schooling on district average income, the average of the bottom 20%, and the average income of the top 20%. Bandwidths: “CCT” indicates the [Calonico, Cattaneo, and Titiunik \(2014\)](#) method, “I and K” the [Imbens and Kalyanaraman \(2012\)](#) 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*** (0.026)	0.220*** (0.035)	0.058* (0.031)	0.133 (0.083)	0.505*** (0.145)	-0.002 (0.097)	0.084 (0.108)	0.445** (0.179)	-0.029 (0.131)
N	5520	5520	5520	5352	5352	5352	5304	5304	5304

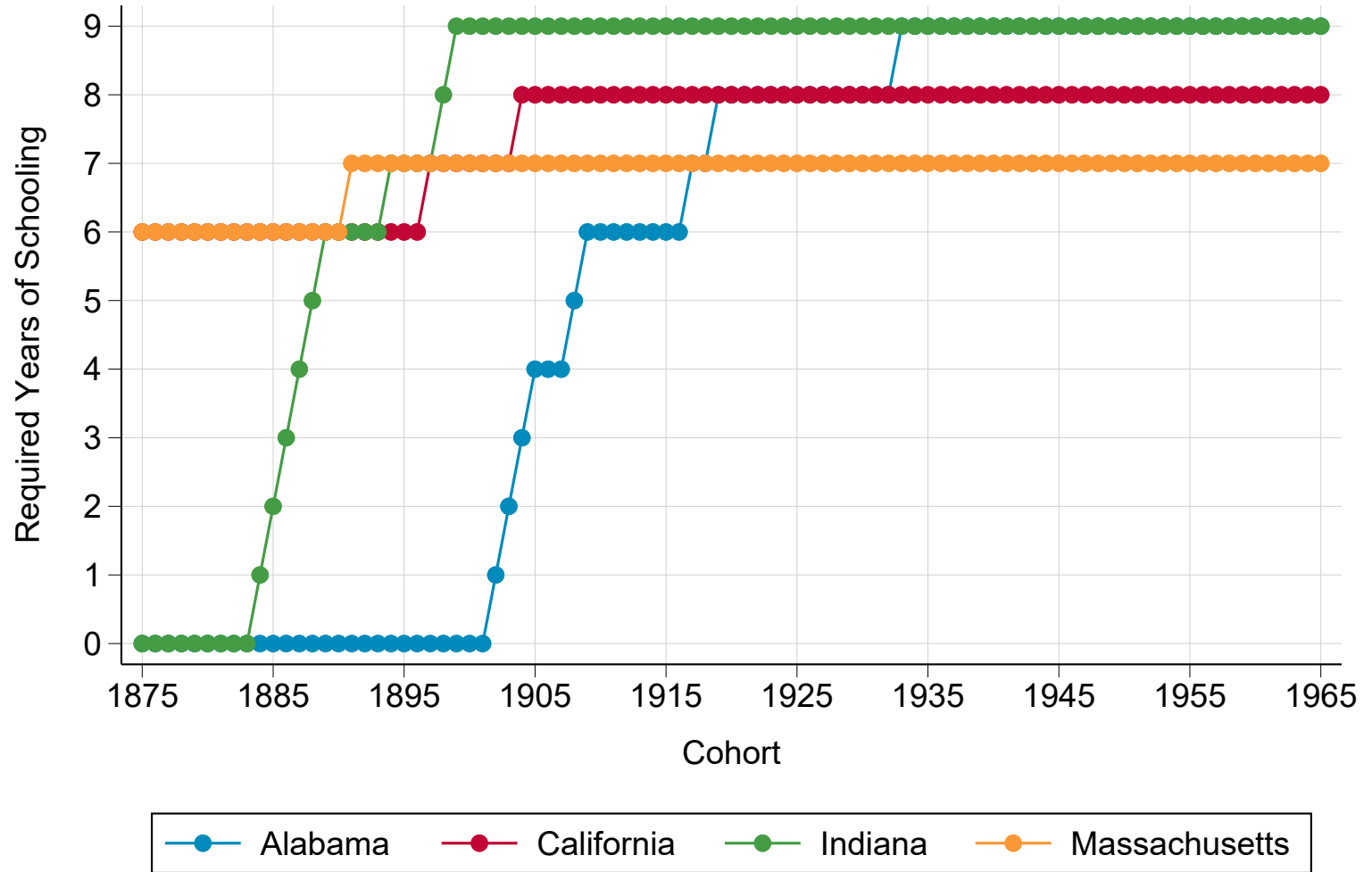
Notes. The table reports the effect of regency average years of schooling on regency average income, the average income of the bottom 20%, and the average income of the top 20%. Columns 1 to 3 control for the demographic composition of the regency, the share of women, and the share of workers with tertiary 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*** (0.032)	0.141*** (0.038)	0.115*** (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*** (0.032)	0.437*** (0.095)	0.050* (0.027)	0.147*** (0.045)	0.431*** (0.110)	0.079* (0.044)	0.082 (0.051)	0.272** (0.114)	0.063 (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 concern relates to changes in education quality, which could affect the results presented in this paper. If education quality has increased or decreased, then educational attainment becomes a biased measure of actual changes in the education of the labor force: 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. 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 a partial and noisy measure of quality. For instance, [Herme 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, 2024). 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:

$$y_{icyt} = \zeta_{cy}s_{icyt} + X_{icyt}\beta_{cy} + \mu_t \quad (19)$$

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 \times 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.⁶ This suggests again

⁶It 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 the two sources tell a broadly similar story on which countries have seen education quality decline or improve most.

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 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](#)):

$$h = \exp(r_L L + r_Q Q) \quad (20)$$

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:

$$\exp(r_L L + r_Q \tilde{Q}) = \exp(r_L \tilde{L} + r_Q Q) \quad (21)$$

Rearranging:

$$\tilde{L} = L - \frac{r_Q}{r_L} (Q - \tilde{Q}) \quad (22)$$

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 return 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

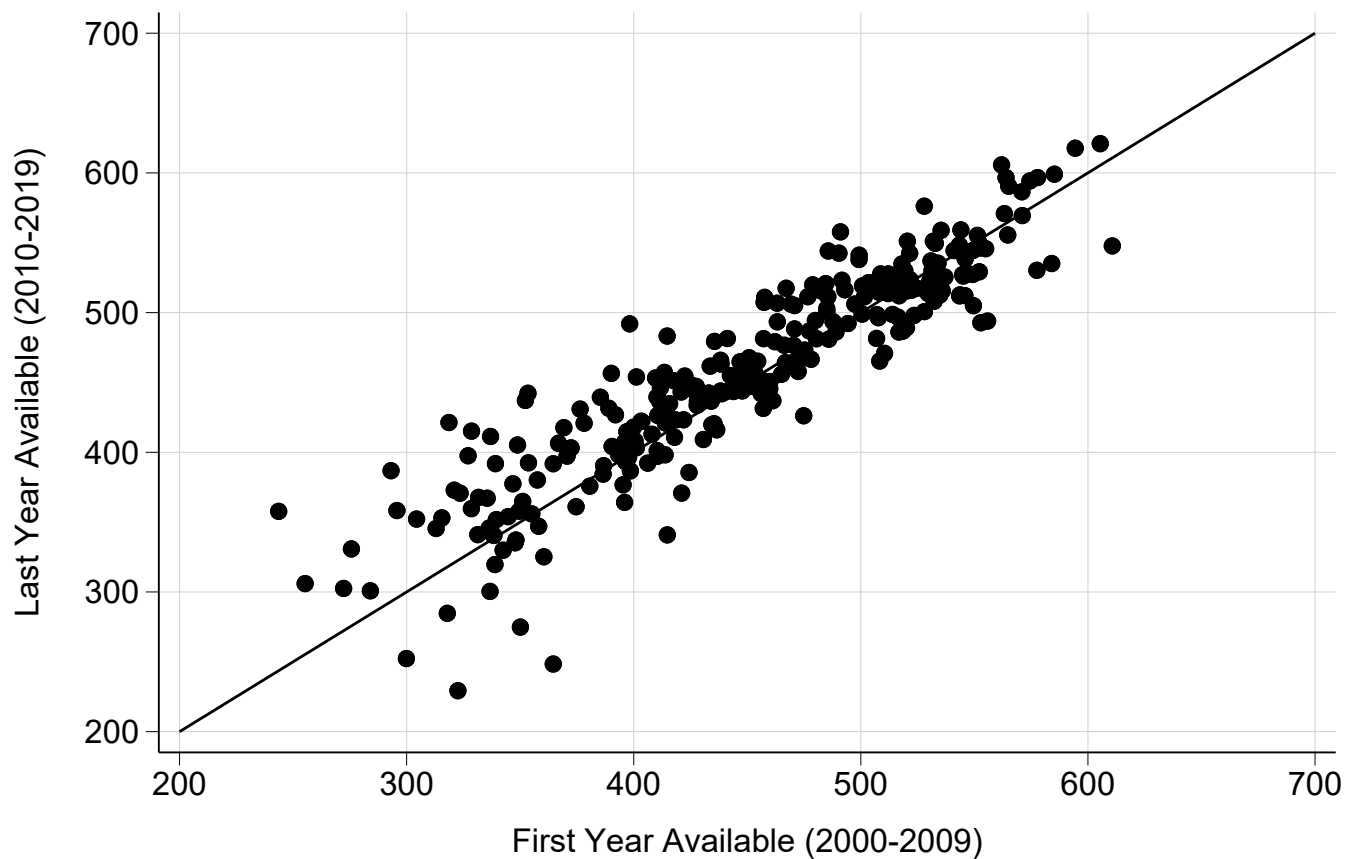
Having estimated changes in education quality and its price relative to education quantity, one can now construct measures of quality-adjusted years of schooling. In practice, I set 1980 as the benchmark year, and adjust estimates of average years of schooling in all other years from 1981 to 2019 so 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%).

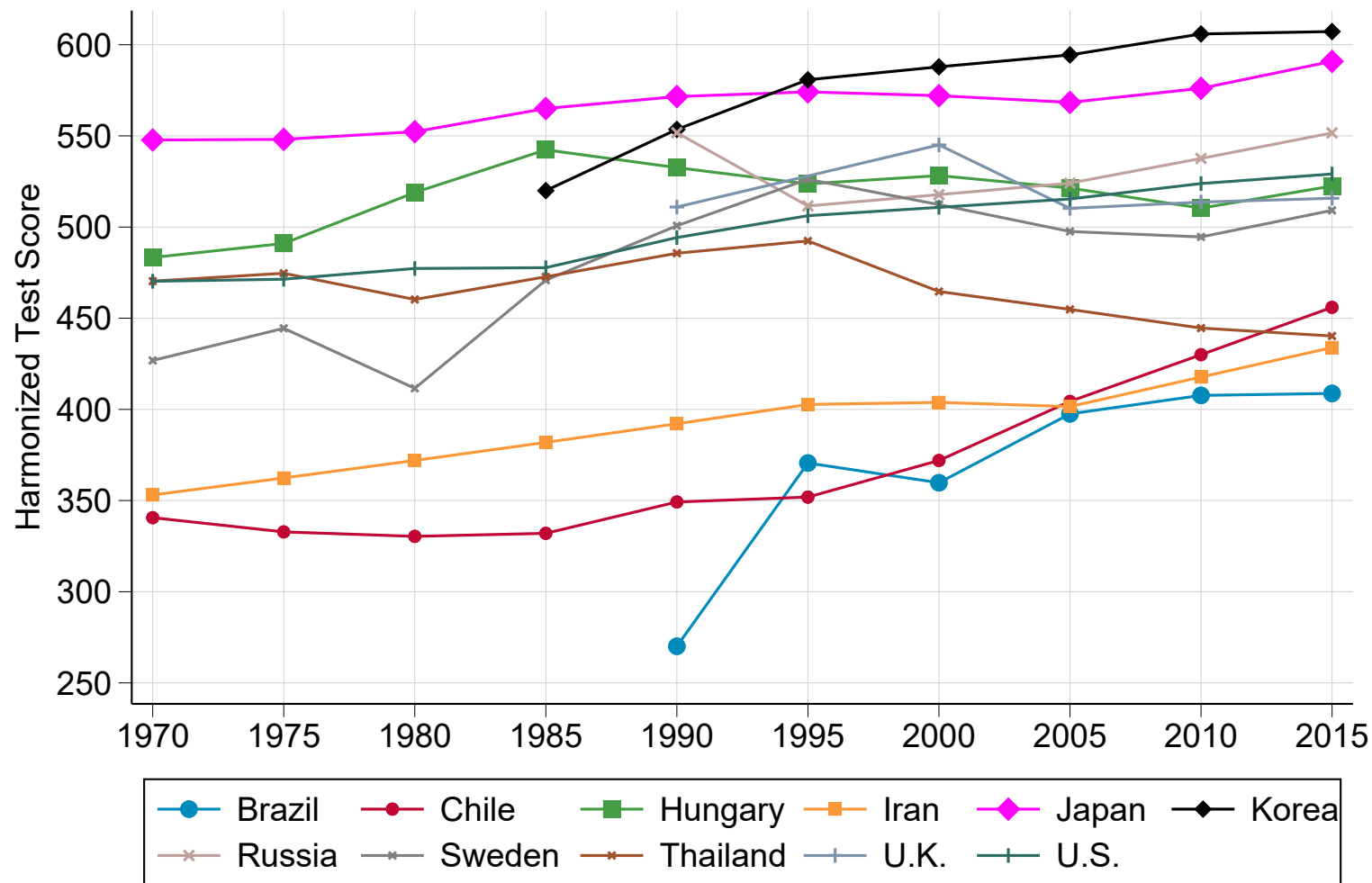
Figure D8 compares the share of growth explained by global income percentile before and after making the lower bound adjustment (assuming that education quality declined as fast as in India for all countries with missing data). The two lines are barely distinguishable: even under strong assumptions on 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



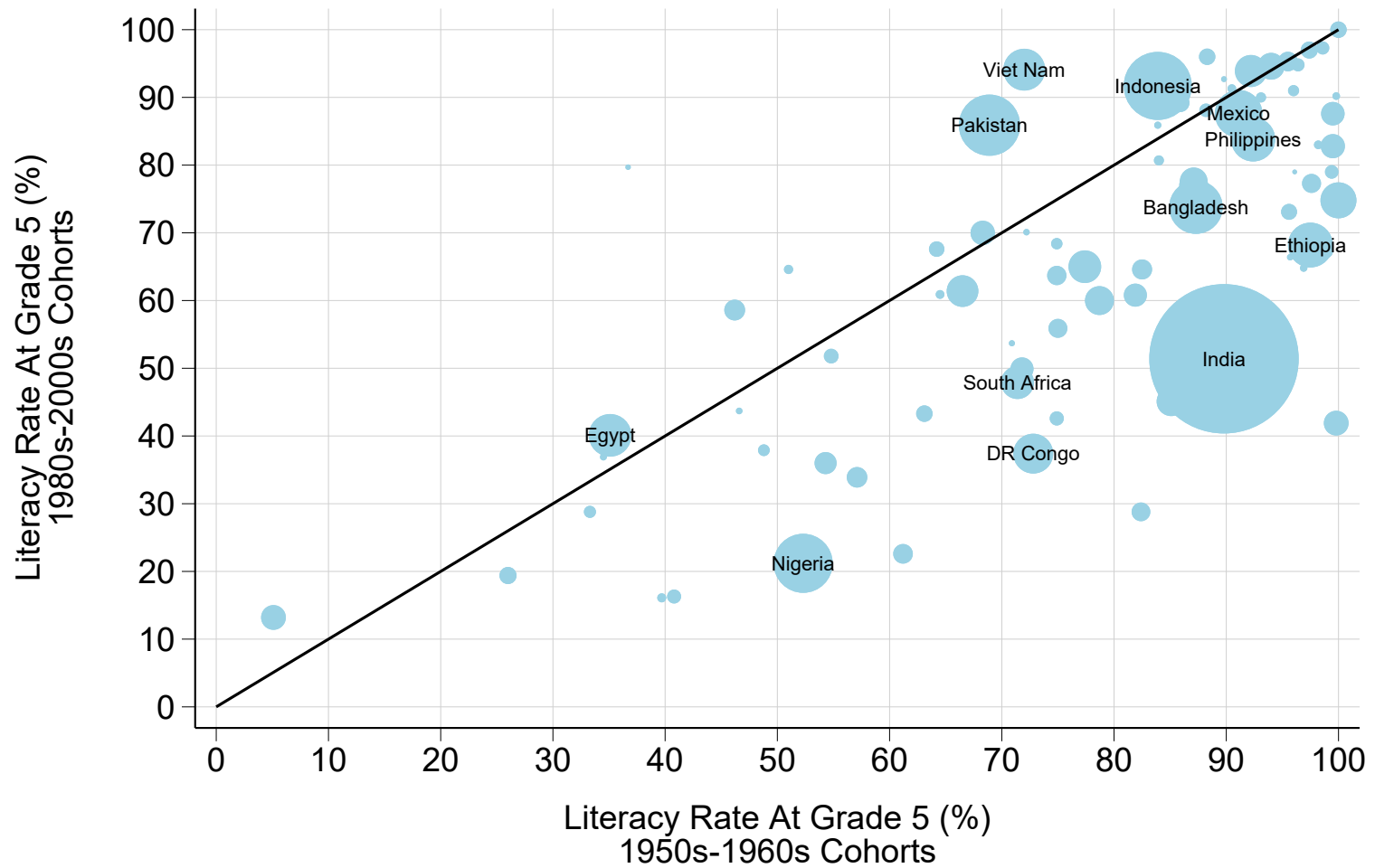
Notes. The figure compares average test scores in 2000-2009 and 2010-2019. Data from [Angrist et al. \(2021\)](#). Each point corresponds to a test score reported for a given country \times education level (primary/secondary) \times subject (maths/science/reading).

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



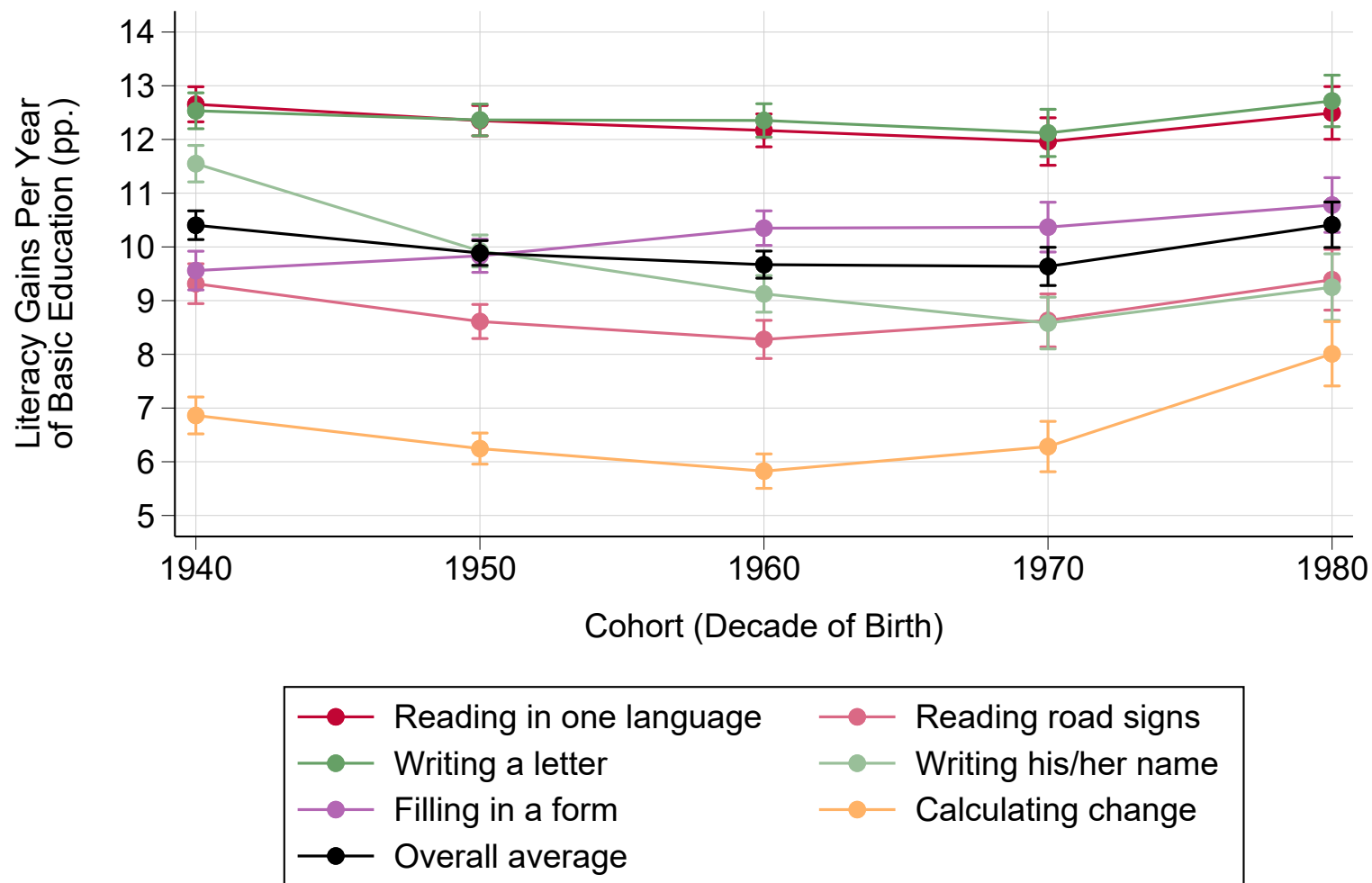
Notes. The figure plots the evolution of harmonized test scores in selected countries from 1970 to 2015, drawing on the database of [Altinok, Angrist, and Patrinos \(2018\)](#).

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



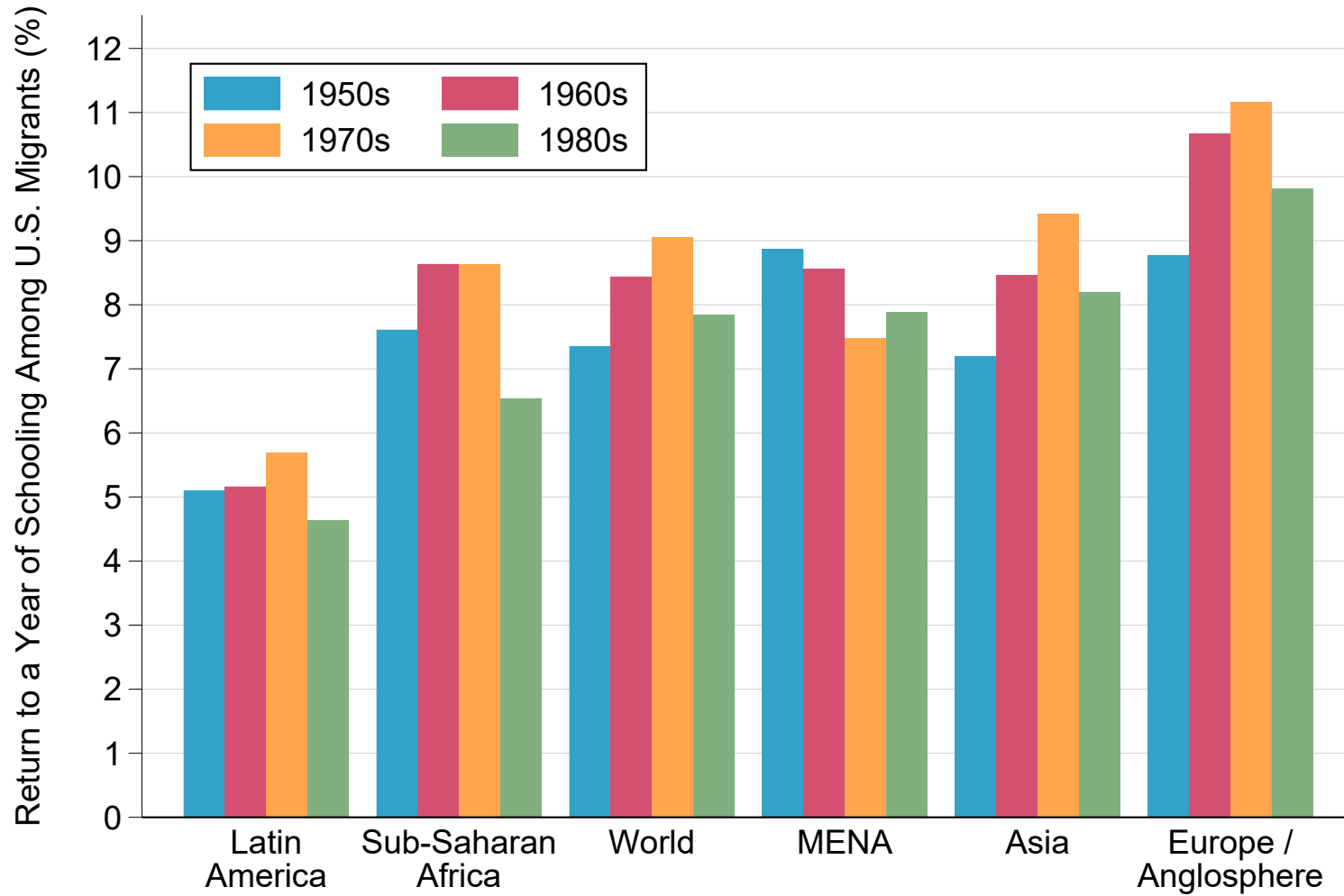
Notes. The figure compares expected literacy at age 5 among 1950s-1960s cohorts (x-axis) and 1980s-2000s cohorts (y-axis), drawing on estimates from [Le Nestour, Moscoviz, and Sandefur \(2022\)](#).

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



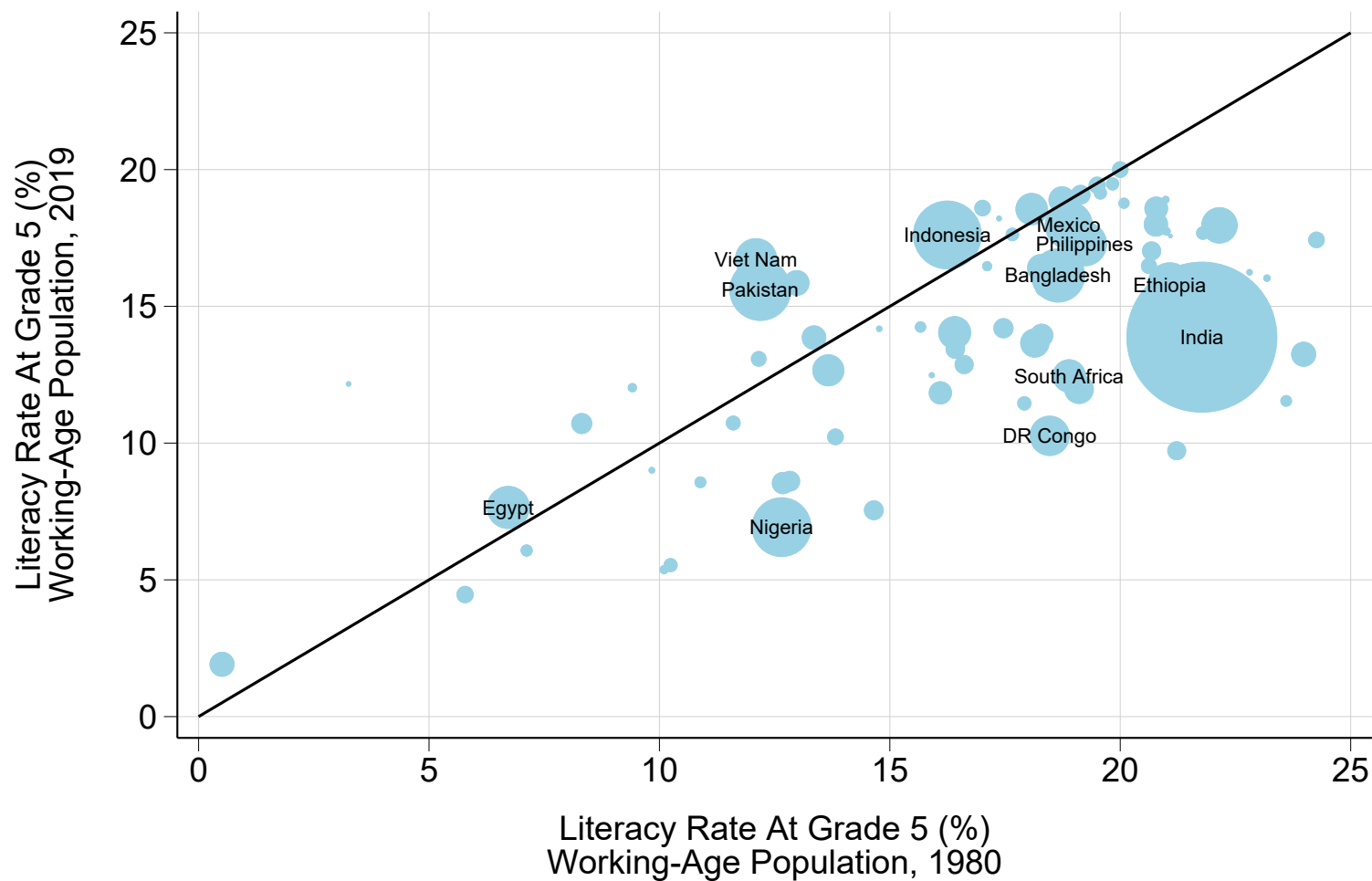
Notes. The figure plots the evolution of education quality in South Africa by cohort for different measures of quality. Author's calculations using South African General Household Surveys (2009-2013). Each data points corresponds to the coefficient of a regression of corresponding literacy scores (0-1) 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



Notes. The figure plots the evolution of education quality by cohort and world region, calculated as the return to a year of schooling among U.S. migrants coming from each world region. 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



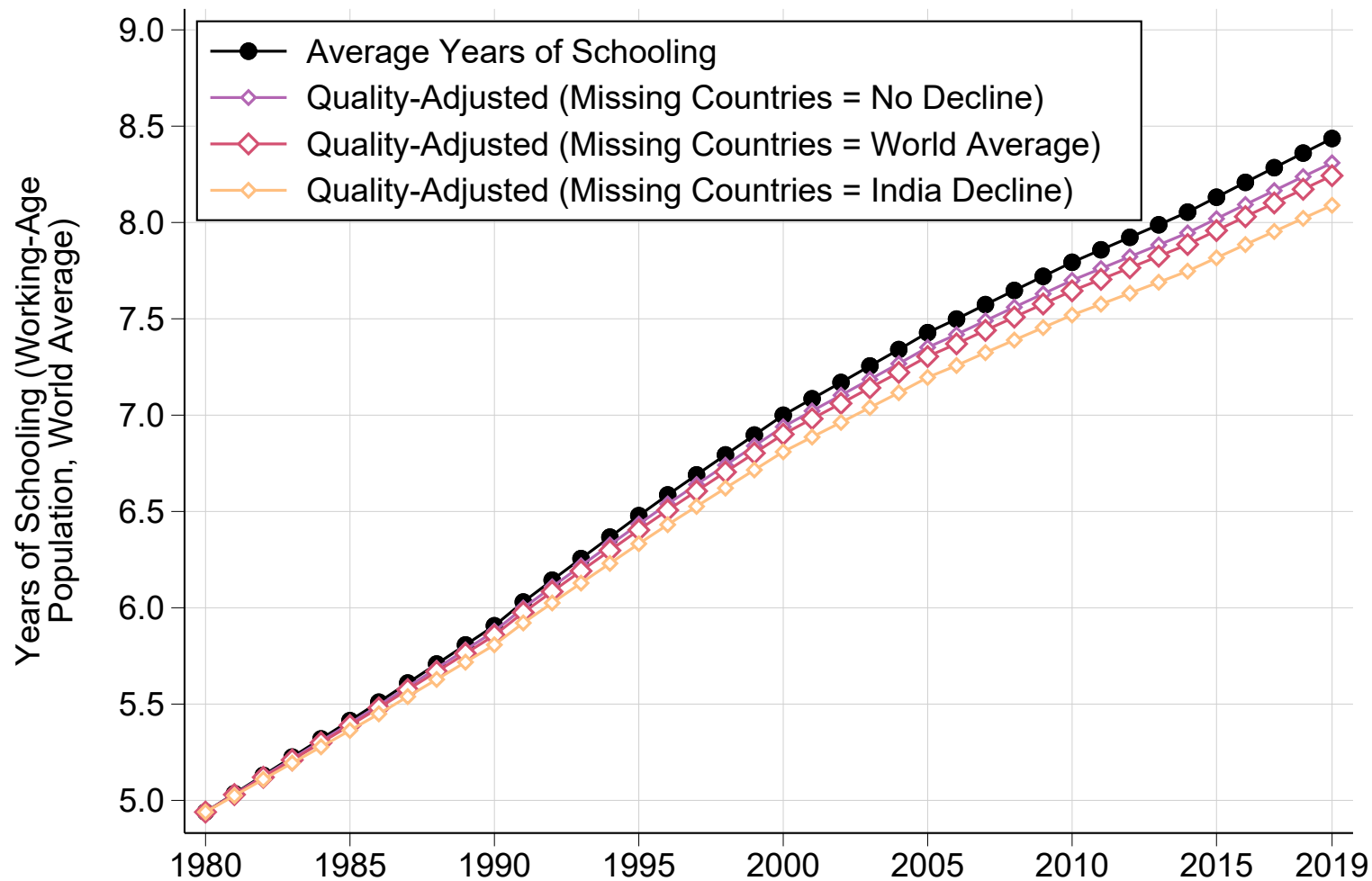
Notes. The figure compares expected literacy at age 5 among the 1980 and 2019 working-age populations. Author's computations combining data from [Le Nestour, Moscoviz, and Sandefur \(2022\)](#) on conditional literacy and population estimates from the United Nations World Population Prospects.

Table D1 – Returns to Literacy

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

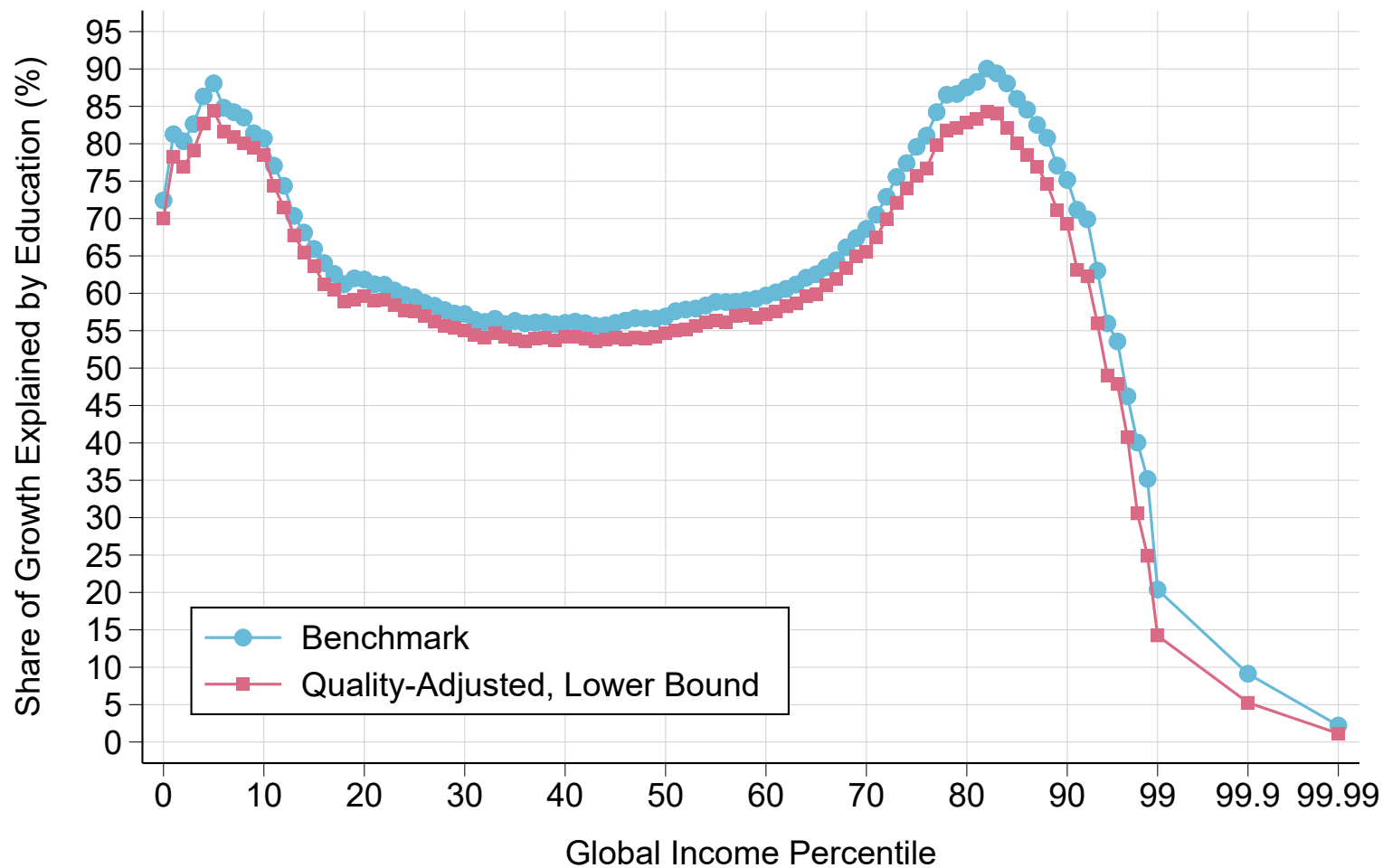
Notes. The table reports estimates of returns to literacy, returns to schooling, and the ratio between the two. The coefficient on literacy corresponds to a regression of the log of personal income on literacy; the coefficient on years of schooling corresponds to a separate regression of the log of personal 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



Notes. The figure plots the evolution of average years of schooling in the world as a whole, before and after adjusting for a potential decline in education quality. Unadjusted estimates correspond to average years of schooling of the world's 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. 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



Notes. The figure plots the share of growth explained by education by global income percentile, before and after adjusting educational attainment for potential changes in education quality. 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 6.1. 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 (ISCO-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

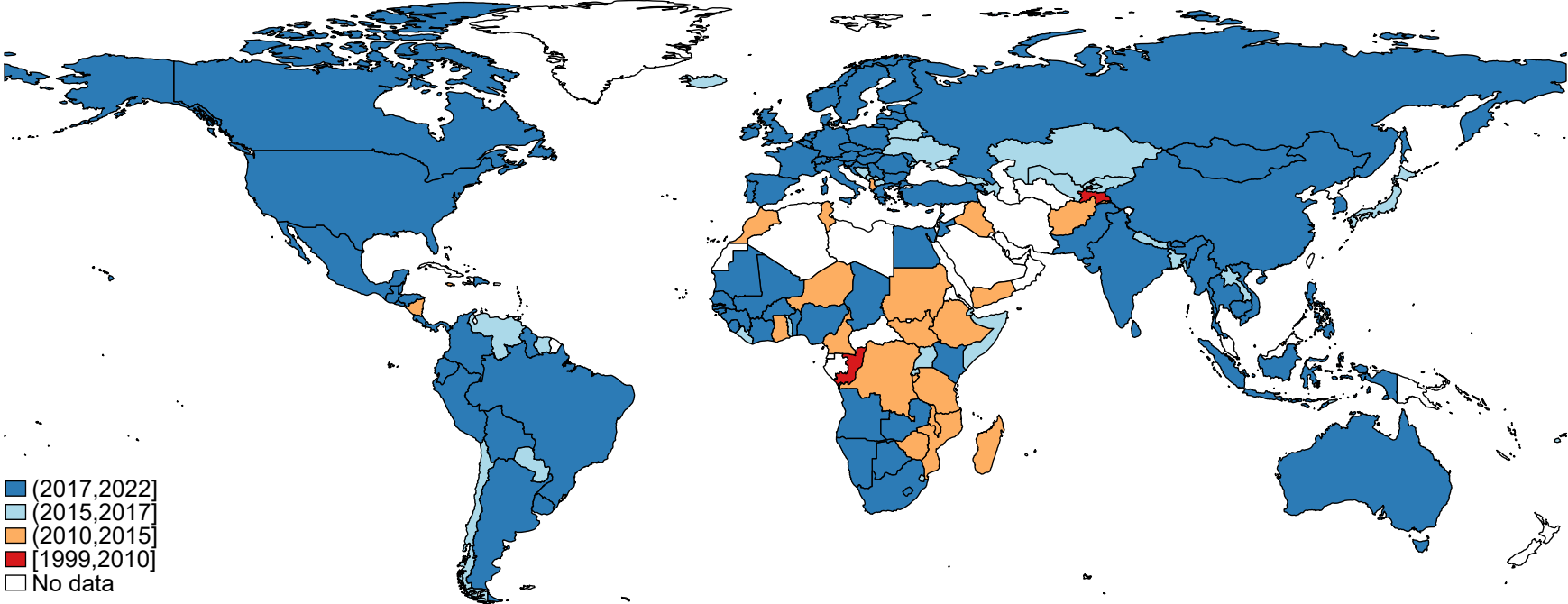
being ideal, with sample sizes of only 3,000-5,000 in each country, yet it is to the best of my knowledge the only data source available to measure individual incomes and education. I use the last wave of the LITS, fielded in 2016, which I harmonize in the same way as the ILO.

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 orçamentos 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	Continuous 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	Continuous 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	Continuous 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	Continuous 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.⁷ 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.⁸ 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

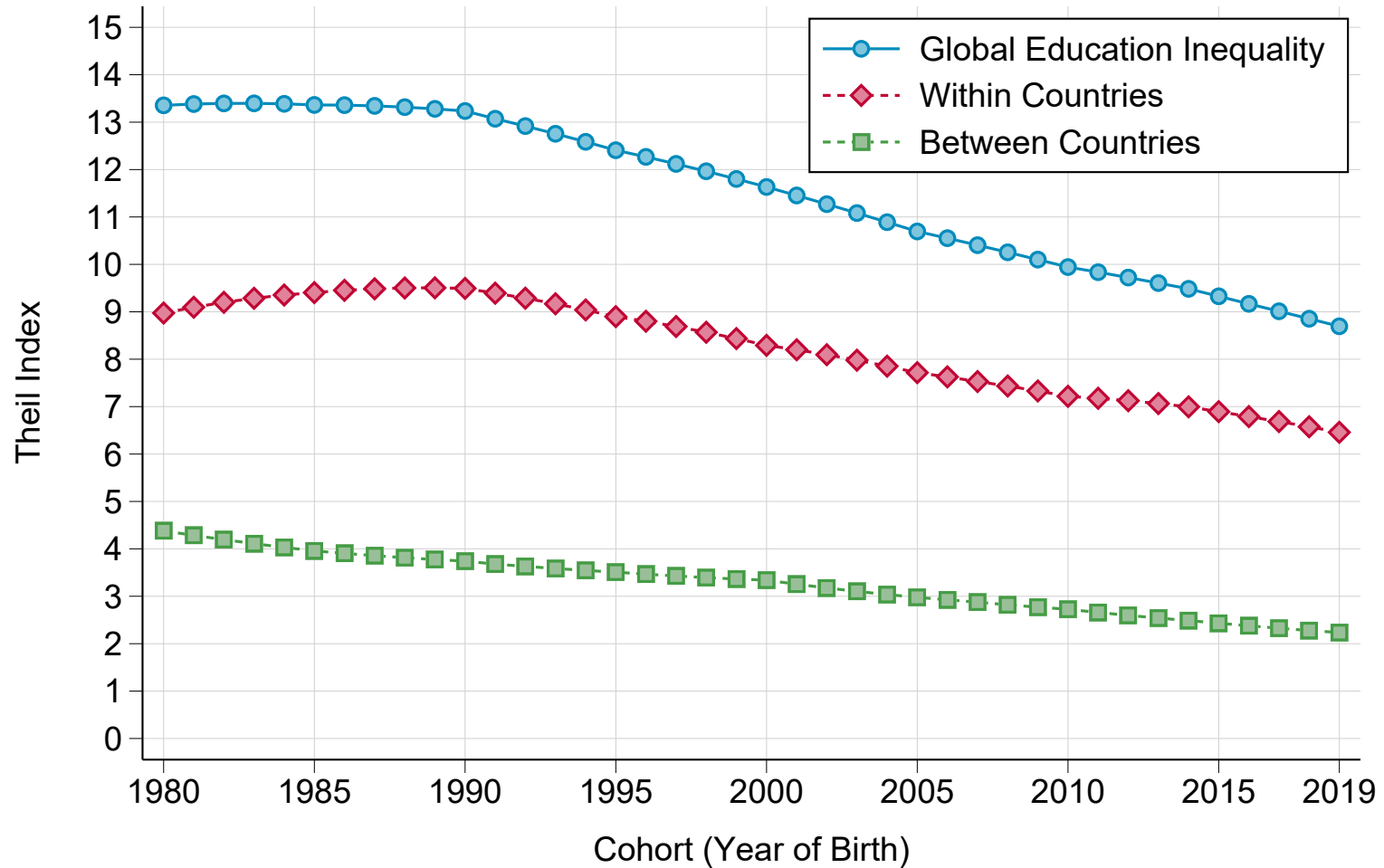
⁷See <http://www.barrolee.com/>.

⁸The 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.

remain. For instance, aggregate and survey 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 [F3](#), [F4](#), [F5](#), and [F6](#), 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.

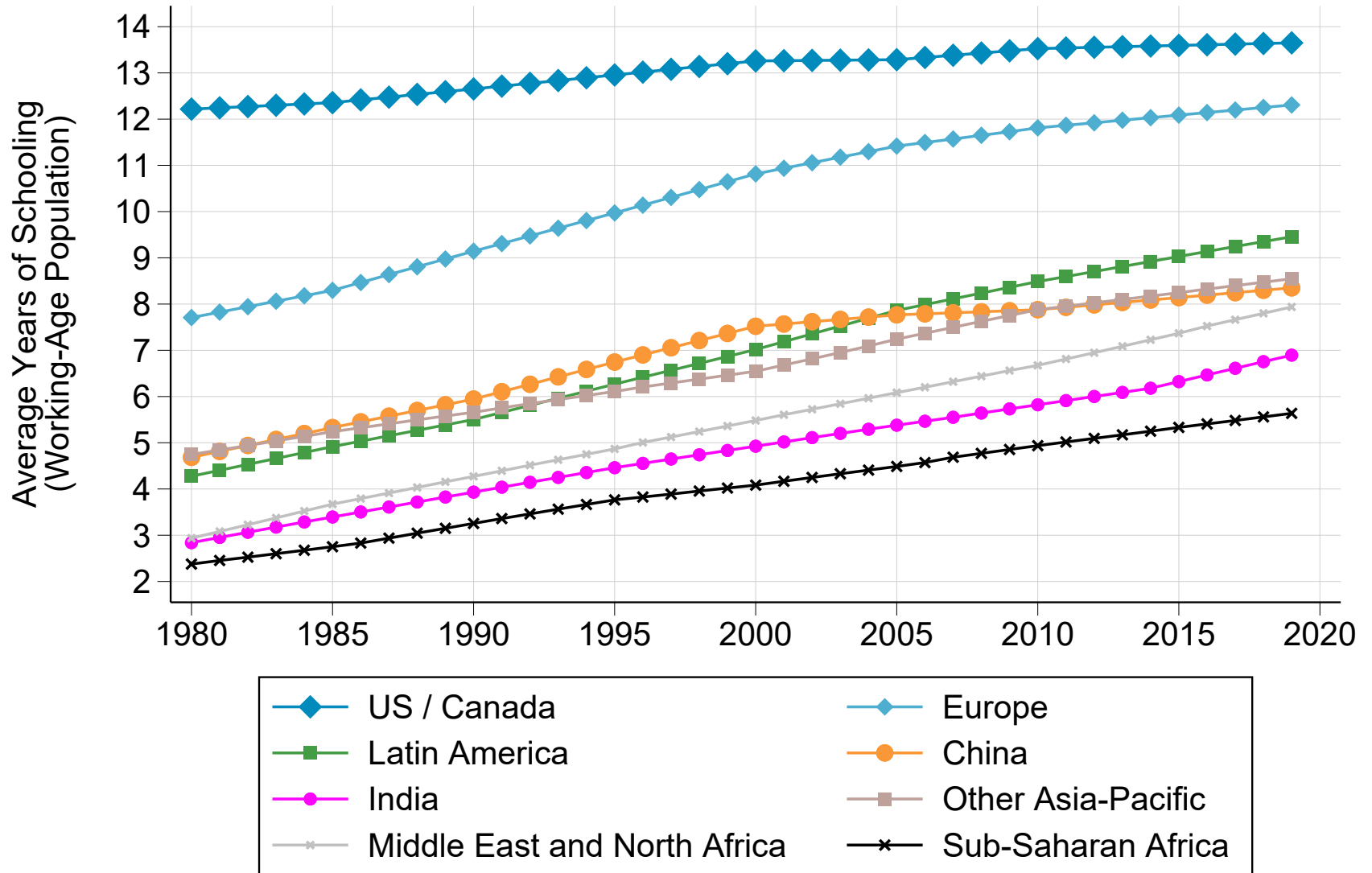
Second, I perform a final small adjustment to the sample weights of each survey to make sure that education levels by age and gender match perfectly in the two sources. I combine aggregate data on the distribution of attainment with data on total population by age and gender from the UN to derive estimates 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 – The Decline of Global Educational Attainment Inequality
Theil Decomposition of Global Human Capital Inequality, 1980-2019



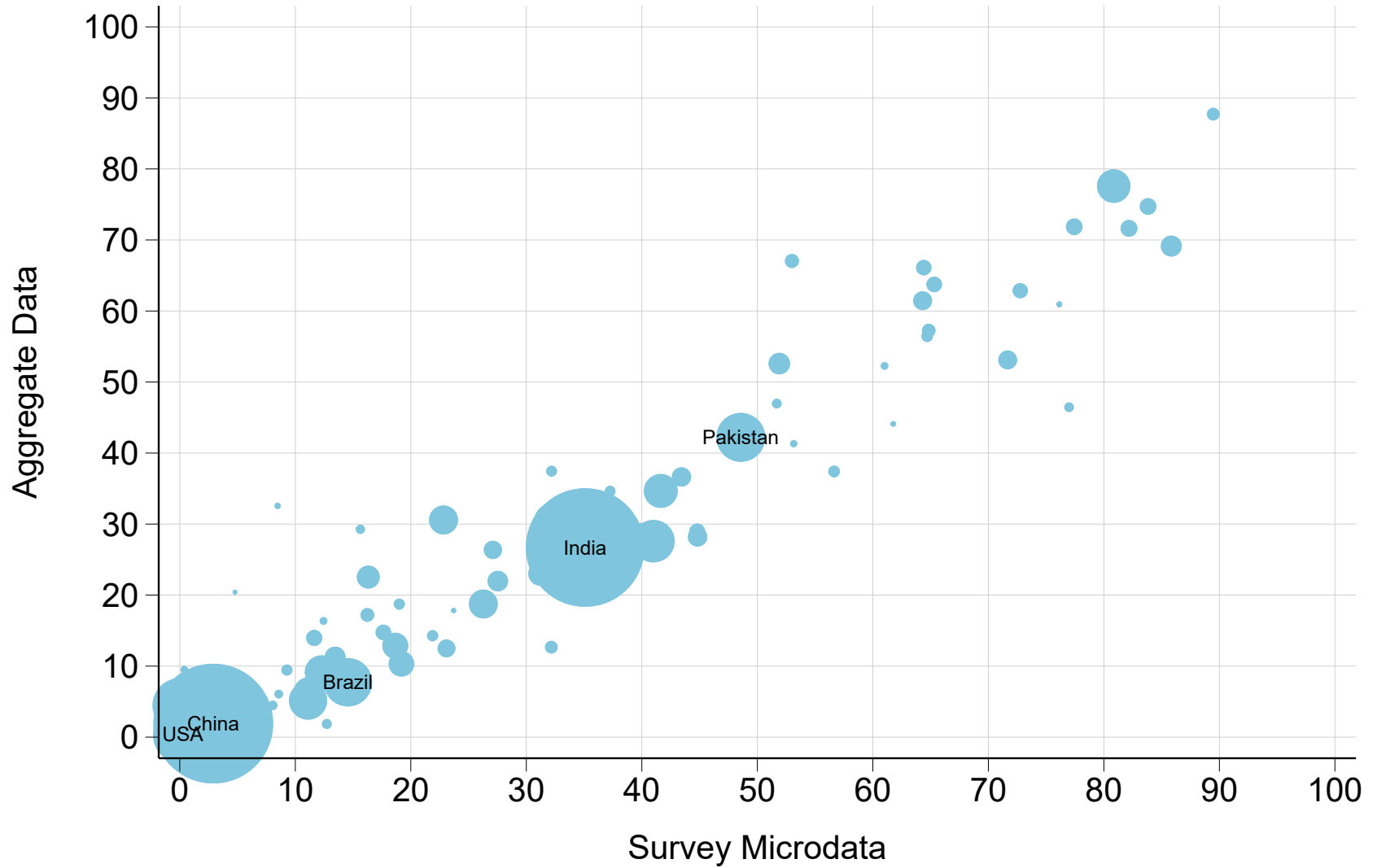
Notes. Author's computations combining data from [Barro and Lee \(2013\)](#), updates, and other sources. The figure plots the evolution of the Theil index of global human capital inequality, as well as its decomposition into a between-country component and a within-country component. Human capital at 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 F2 – Average Years of Schooling by World Region



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

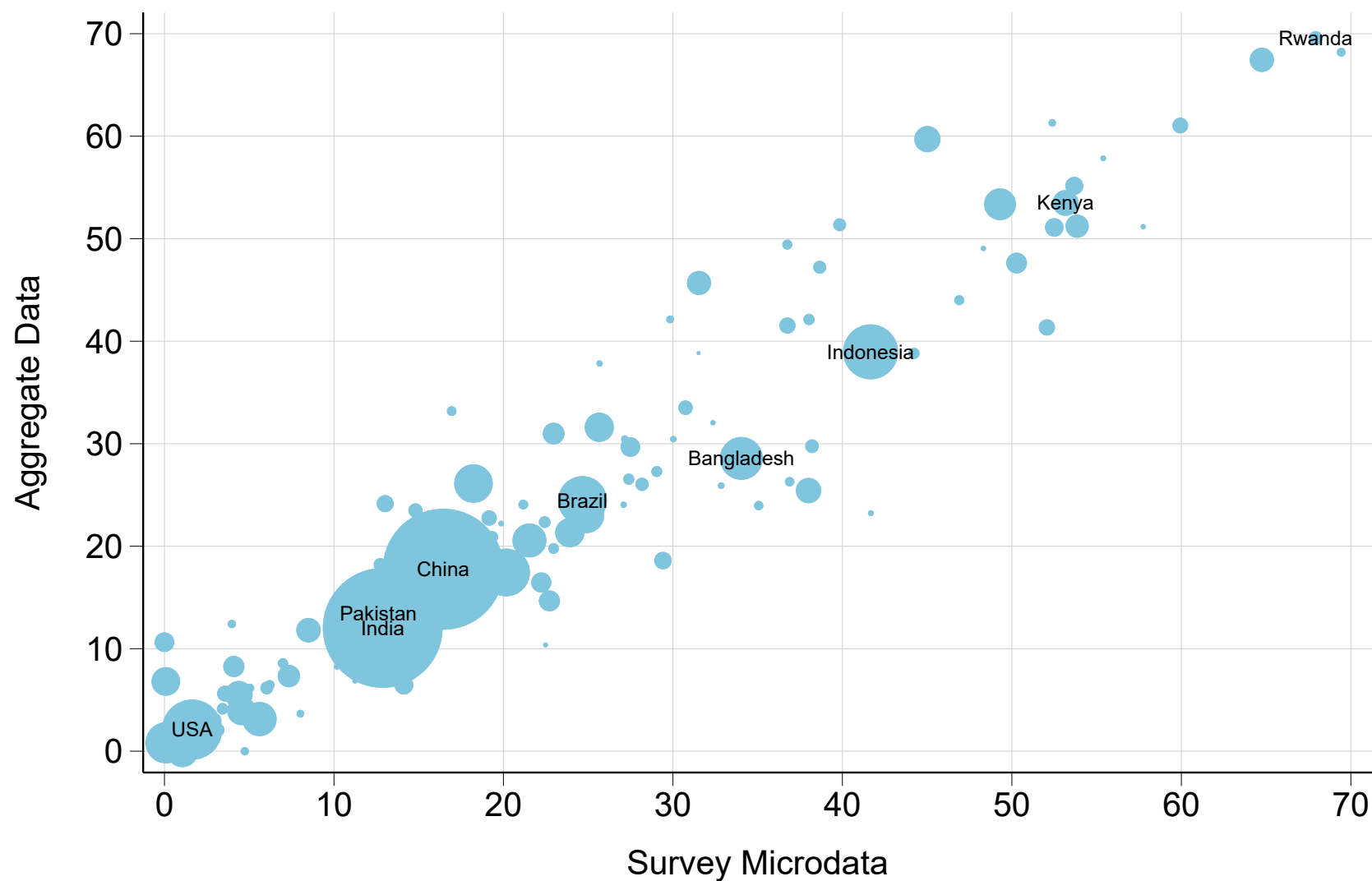
Figure F3 – Barro-Lee Versus Survey Data: Share of Working-Age Population With No Schooling



105

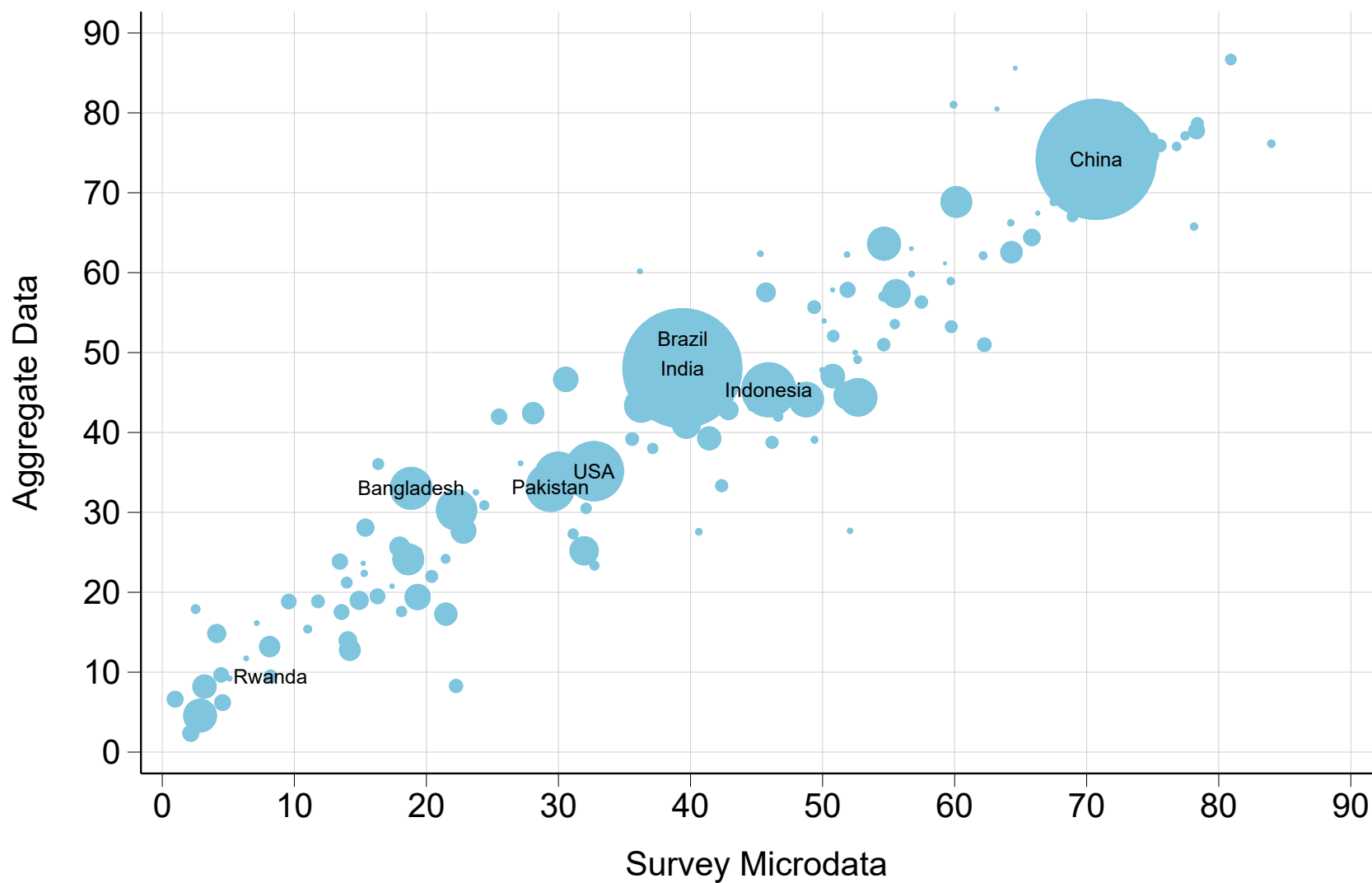
Notes. The figure compares estimates of the share of the working-age population with no schooling in the survey microdata (x-axis) and aggregate data from Barro and Lee (2013) and other sources (y-axis), after manual reclassification of educational categories in each country.

Figure F4 – 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.

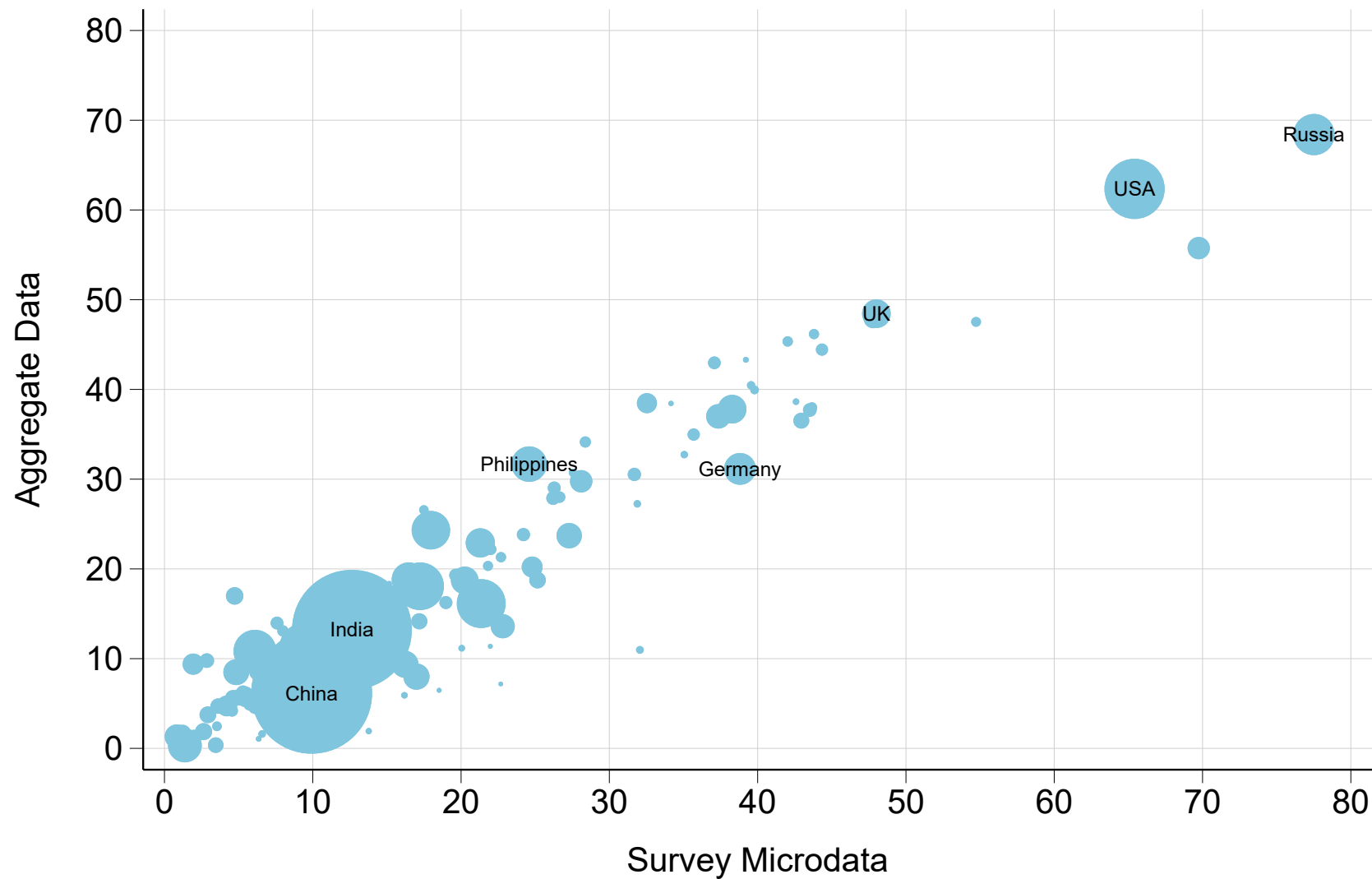
Figure F5 – Barro-Lee Versus Survey Data: Share of Working-Age Population With Secondary Education



107

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 F6 – 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:

$$\ln y_{ict} = \alpha_t + \beta_{ct}^{pri} D_{ict}^{pri} + \beta_{ct}^{sec} D_{ict}^{sec} + \beta_{ct}^{ter} D_{ict}^{ter} + X_{ict} \beta + \varepsilon_{ict} \quad (23)$$

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 positive 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 G1 plots the resulting estimates of returns to schooling obtained from a pooled regression by education level. Figure G2 plots the cross-country 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:

$$r(s_1, s_2) = \ln(w_2) - \ln(w_1) \quad (24)$$

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:

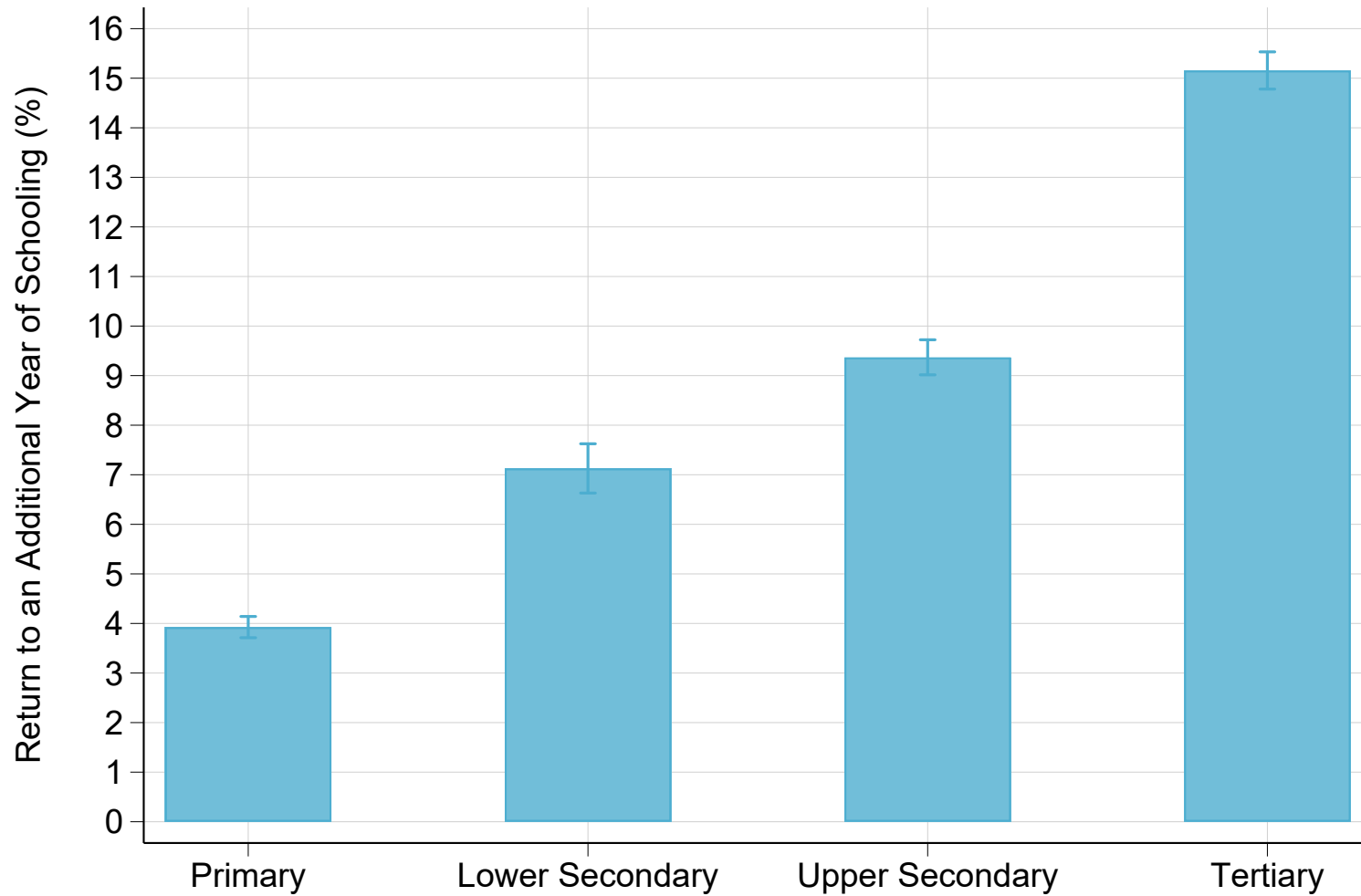
$$\beta(s_1, s_2) = \exp\left(r(s_1, s_2)\right)^{1/T(s_1, s_2)} - 1 \quad (25)$$

With $T(s_1, s_2)$ 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:

$$r^{IV}(s_1, s_2) = T(s_1, s_2) \times \ln\left(1 + \gamma\beta(s_1, s_2)\right) \quad (26)$$

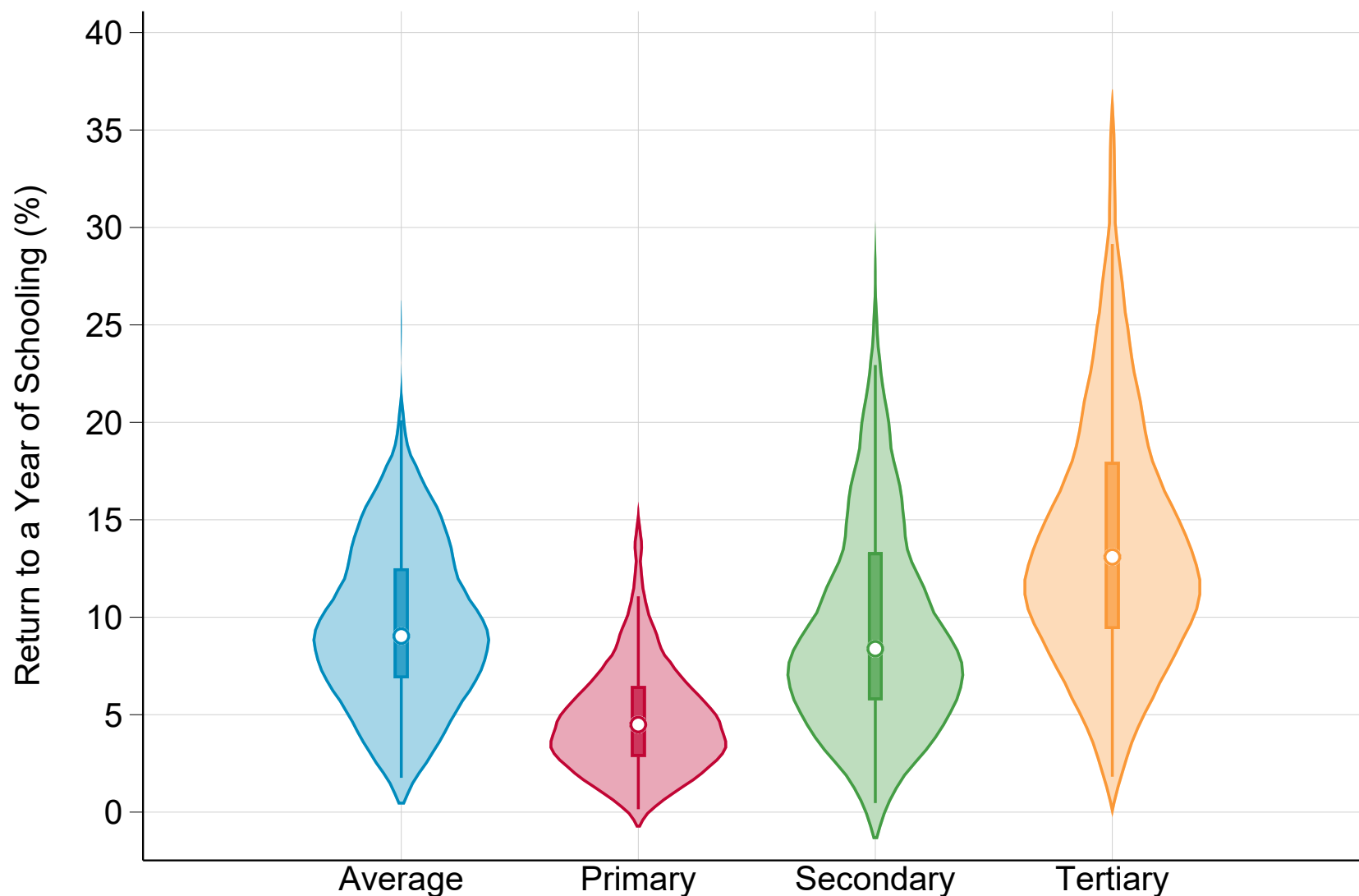
$$= T(s_1, s_2) \times \ln\left[1 + \gamma\left(\exp\left(r(s_1, s_2)\right)^{1/T(s_1, s_2)} - 1\right)\right] \quad (27)$$

Figure G1 – Returns to Schooling: Pooled Estimates by Level



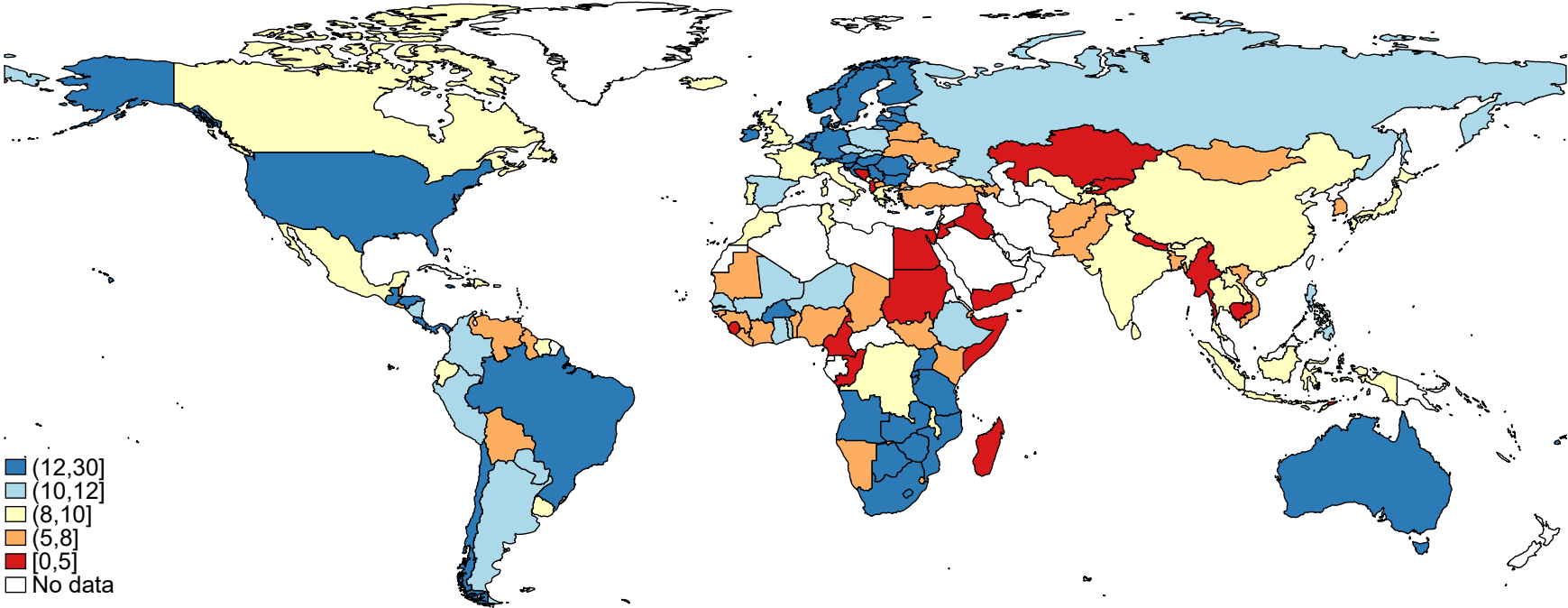
Notes. The figure reports estimates of an additional year of schooling by education level, based on a pooled regression on the full micro dataset. Primary: returns to a year of primary education. Lower secondary: return to a year of lower secondary education, restricting the sample to individuals with 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.

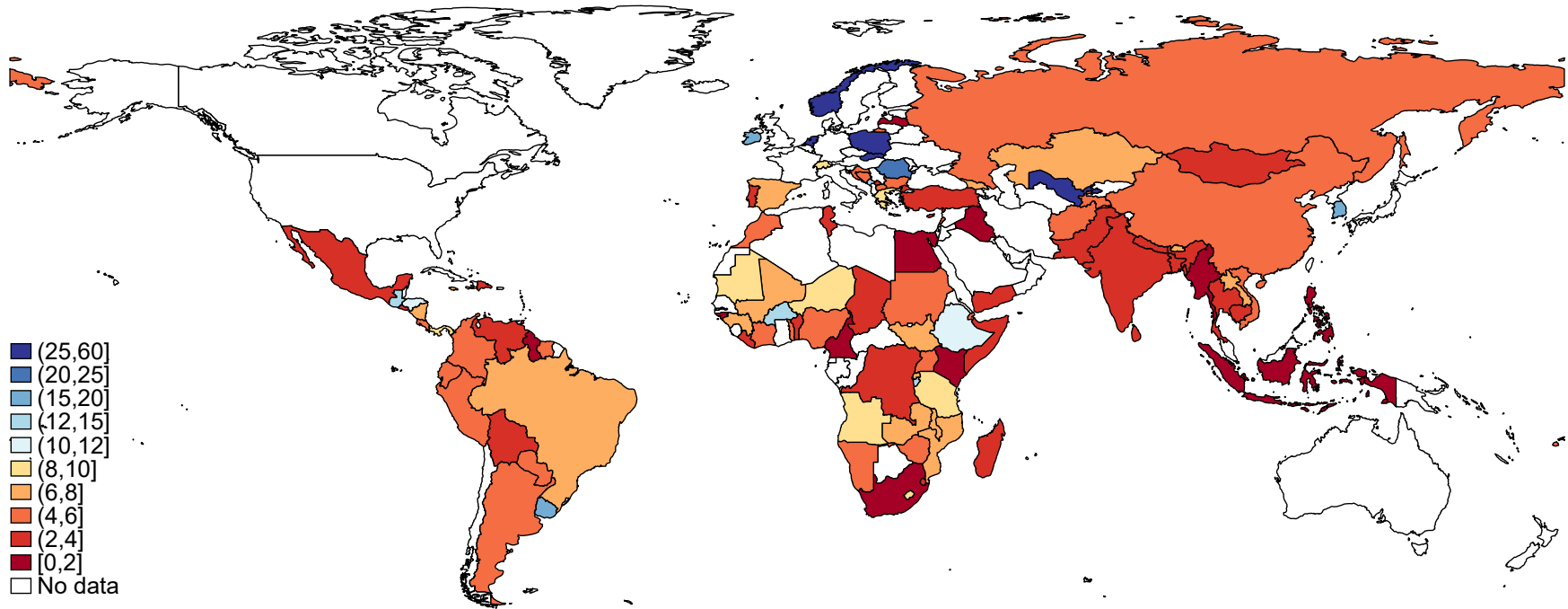
Figure G3 – Return to an Additional Year of Schooling



114

Notes. 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.

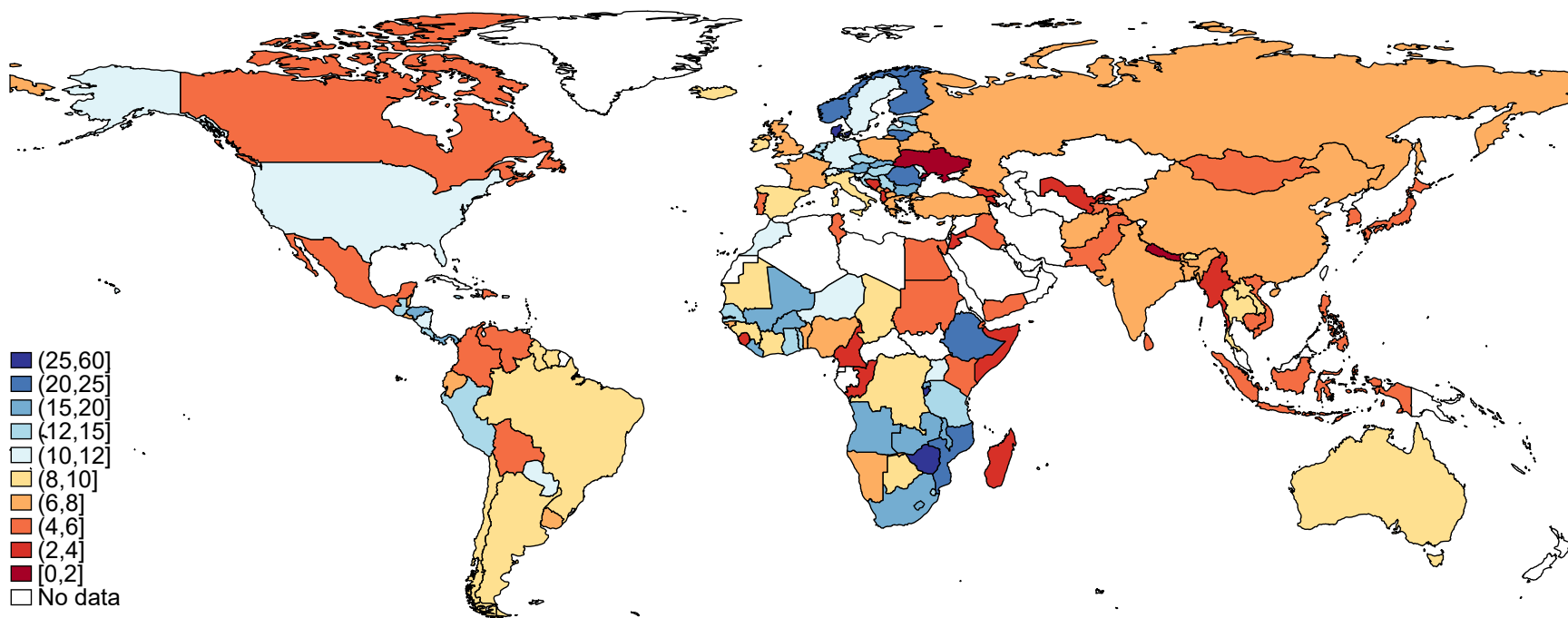
Figure G4 – Returns to an Additional Year of Primary Education



115

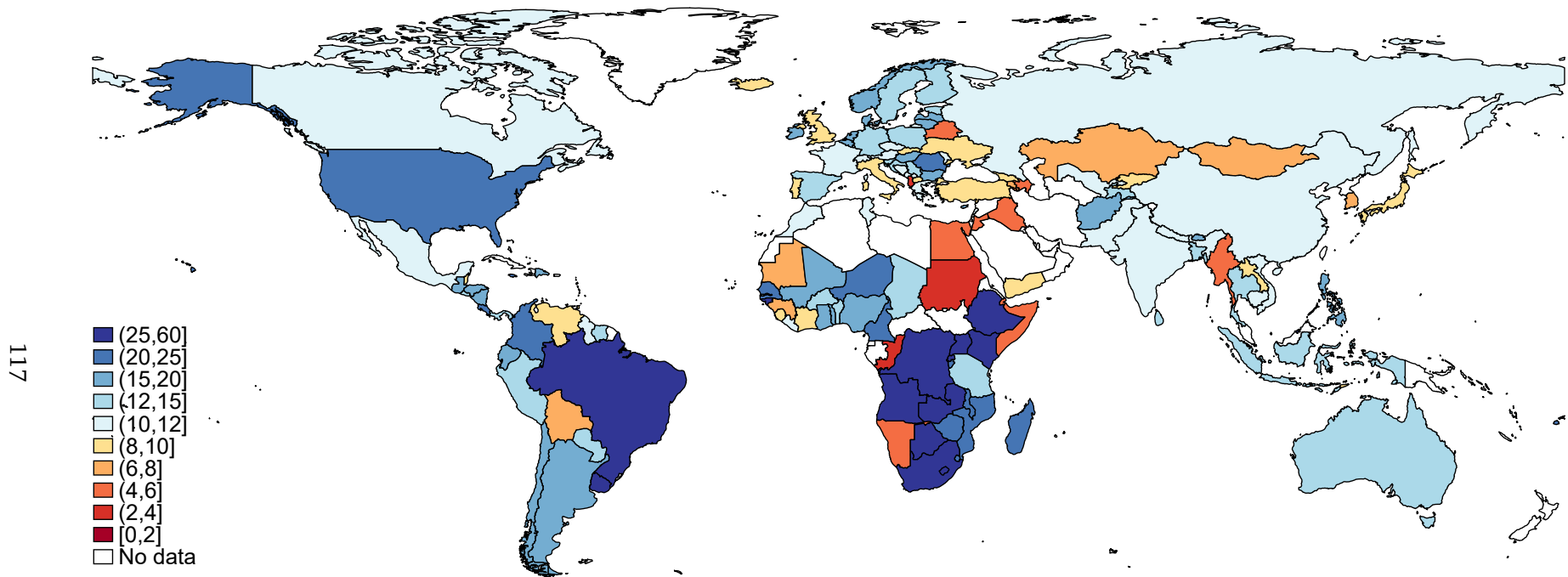
Notes. Estimates correspond to the effect of one additional year of schooling on the log of personal income, estimated using modified Mincerian equations controlling for an experience quartic, gender, and interactions between the experience quartic and gender. The return to a year of schooling is estimated among individuals with either no schooling, some primary education, or completed primary education.

Figure G5 – Returns to an Additional Year of Secondary Education



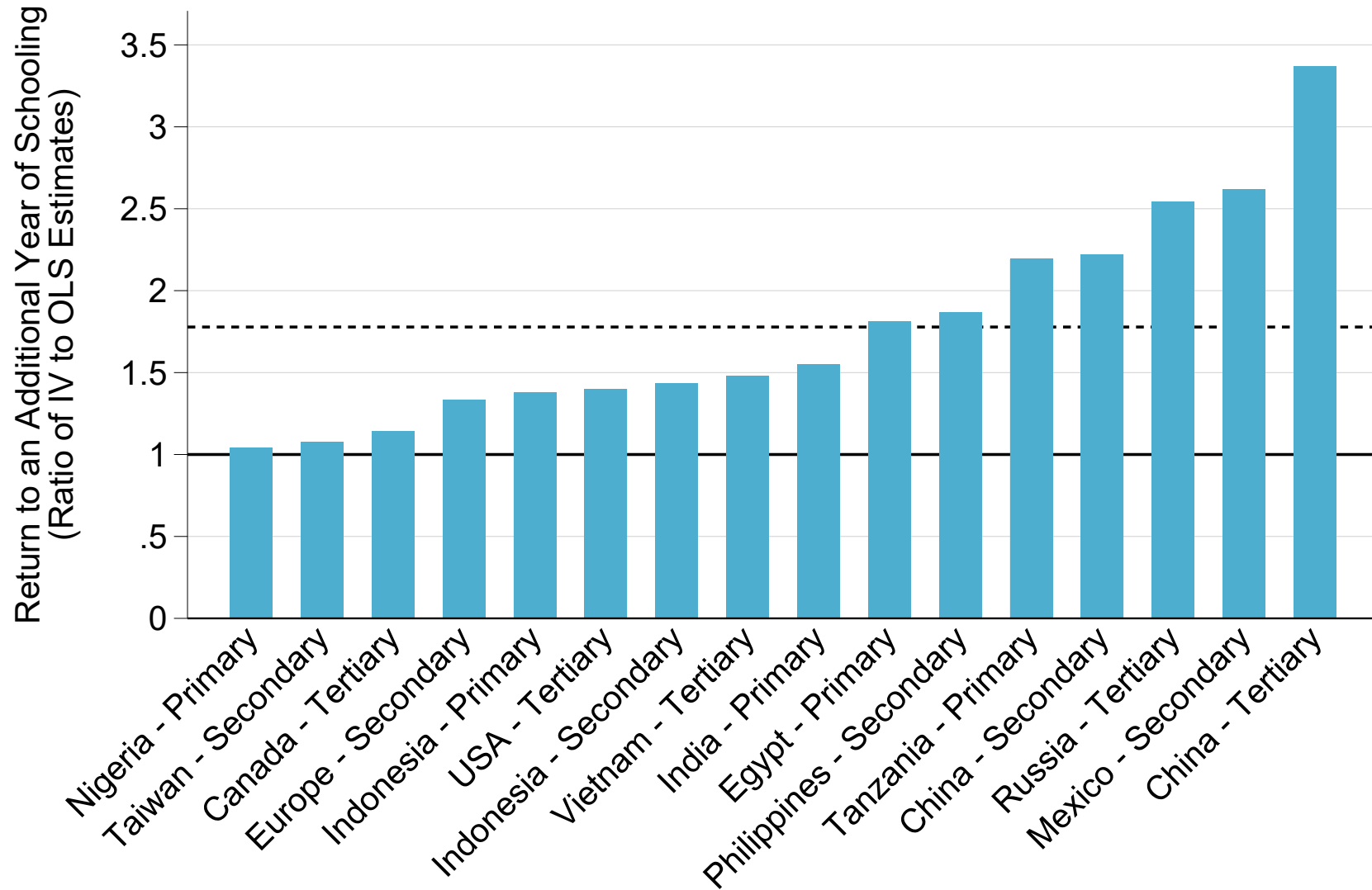
Notes. Estimates correspond to the effect of one additional year of schooling on the log of personal income, estimated using modified Mincerian equations controlling for an experience quartic, gender, and interactions between the experience quartic and gender. The return to a year of schooling is estimated among individuals with either some primary education, completed primary education, some lower or upper secondary education, or completed upper secondary education.

Figure G6 – Returns to an Additional Year of Tertiary Education



Notes. 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. The return to a year of schooling is estimated among individuals with some upper secondary education, completed upper secondary education, some tertiary education, or completed tertiary education.

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



Notes. The figure plots the ratio between instrumental variable and OLS estimates of the returns to schooling, compiled from a number of published studies. Pri/Sec/Ter: returns to a year of primary/secondary/tertiary education. Dashed line: average ratio observed across studies.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years of Schooling	0.089*** (0.001)	0.090*** (0.001)	0.045*** (0.001)	0.046*** (0.001)	0.084*** (0.001)	0.085*** (0.001)	0.150*** (0.001)	0.150*** (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 model: controls for gender, potential experience, and potential experience squared. Extended model: 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*** (0.001)	0.097*** (0.001)	0.046*** (0.001)	0.047*** (0.001)	0.085*** (0.001)	0.088*** (0.001)	0.150*** (0.001)	0.144*** (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*** (0.001)	0.063*** (0.001)	0.063*** (0.001)
Pakistan	0.074*** (0.001)	0.071*** (0.001)	0.071*** (0.001)
DR Congo	0.075*** (0.002)	0.071*** (0.002)	0.070*** (0.002)
Burkina Faso	0.089*** (0.006)	0.106*** (0.003)	0.086*** (0.005)
Benin	0.078*** (0.003)	0.068*** (0.002)	0.071*** (0.003)
Côte d’Ivoire	0.073*** (0.003)	0.058*** (0.002)	0.052*** (0.002)
Guinea-Bissau	0.041*** (0.004)	0.061*** (0.002)	0.052*** (0.003)
Mali	0.072*** (0.005)	0.080*** (0.002)	0.061*** (0.003)
Niger	0.108*** (0.004)	0.103*** (0.003)	0.101*** (0.003)
Sénégal	0.078*** (0.003)	0.074*** (0.002)	0.075*** (0.002)
Togo	0.100*** (0.006)	0.078*** (0.002)	0.076*** (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 OLS and IV estimates of returns to schooling from selected articles. Level: education level at which educational expansion happened, used to derive the instrumental variable estimate of the return to schooling. OLS: return to schooling estimated by ordinary least squares over the entire working-age population in the corresponding article. β : return to a year of schooling. SE: standard error associated with the estimate.

Table G5 – 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.

H. Data Appendix: Income Distribution Data

The final step of the distributional growth accounting exercise 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.

H.1. Pretax Income Inequality Data

Data on the world distribution of income come from the World Inequality Database (WID). It reports average per-capita pretax income by percentile in all countries in the world 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 [Piketty, Saez, and Zucman, 2018](#); [Chancel et al., 2022](#)). 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).

H.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. Alternative results with capital income affected by schooling are presented in section 4.4.

H.3. Capital Income Concentration

The last step is to estimate how the capital income share varies alongside the income distribution in each country. Data on this decomposition are scarce. Drawing on selected high-quality studies, I was able to derive such profiles 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 H7. 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.

H.4. World Bank Income Distribution Data

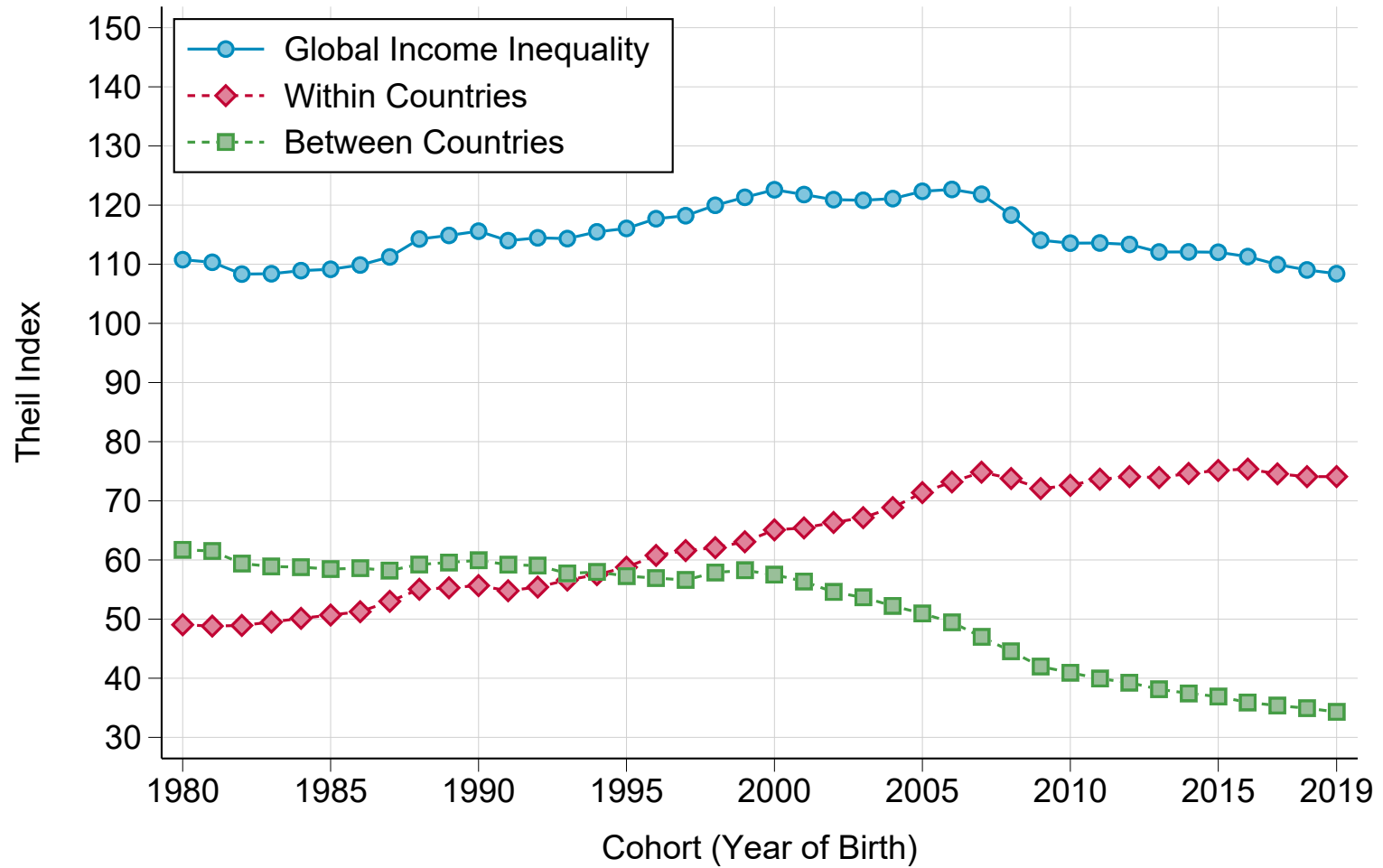
As a robustness check, I investigate the sensitivity of my results to using World Bank income distribution data instead of the World Inequality Database. The data is available from the World Bank's Poverty and Inequality Platform (PIP), and presents itself in the form of distributions available for selected countries and years. The income concept is either consumption or posttax disposable income per capita, depending on the country. All values are reported in 2017 PPP USD per day.

Unfortunately, although the World Bank regularly reports indicators of global poverty, it does not publish underlying estimates of the world distribution of income. I thus attempt to reconstruct measures of pretax and posttax income myself. Starting from available data, I first extrapolate the average income of each country-percentile to missing years using real GDP per capita growth rates. For the 17 countries entirely missing, I use estimates from the World Inequality Database. The resulting database yields trends in global poverty almost identical to those officially reported by the World Bank. I then reconstruct measures of

pretax income. In the absence of any information on savings, I define pretax income in each country as consumption or disposable income, minus social assistance transfers, plus direct taxes. The distribution of social assistance transfers and direct taxes come from two companion papers ([Fisher-Post and Gethin, 2023](#); [Gethin, 2024](#)).

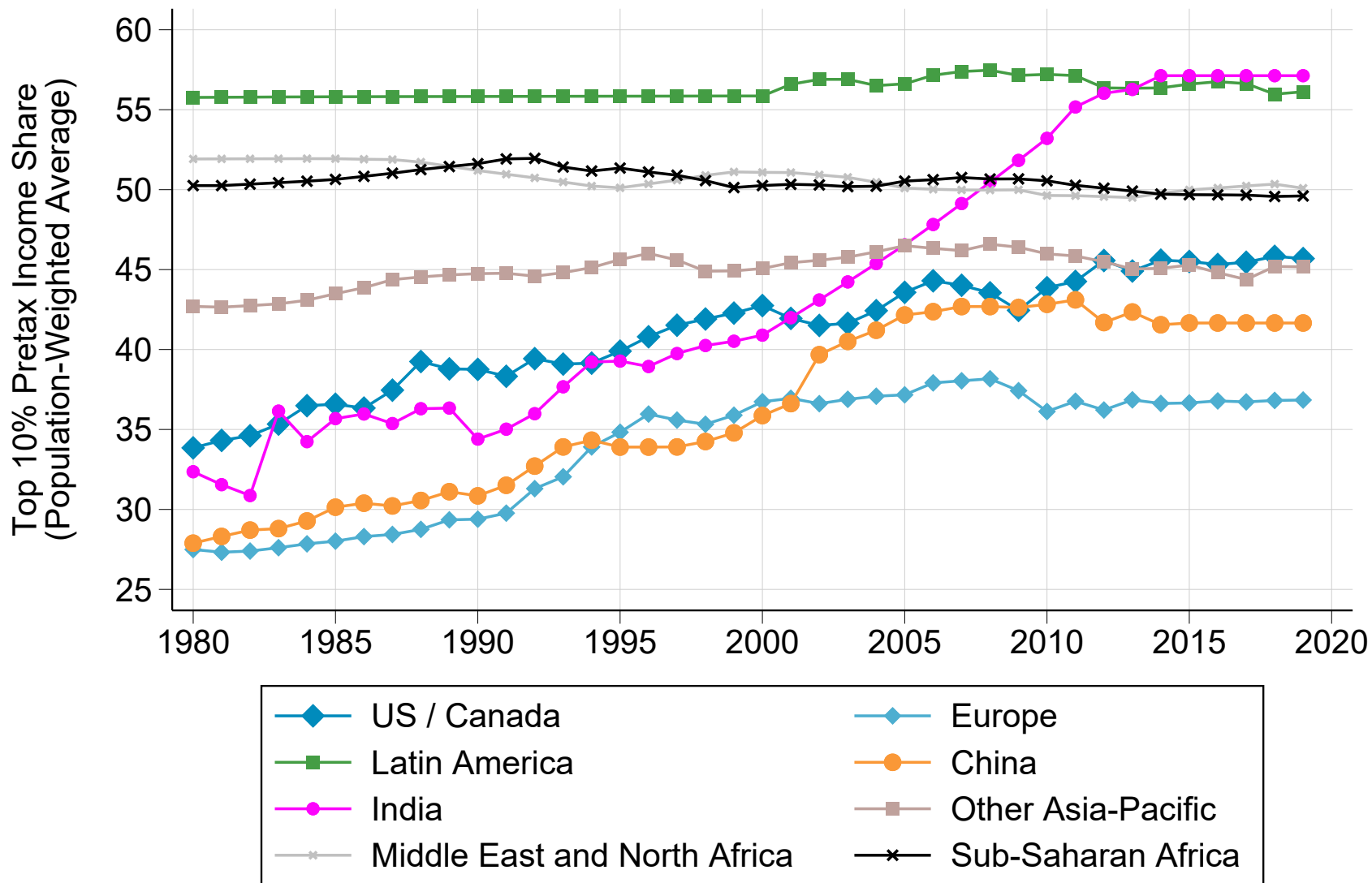
Both the levels and trends in global poverty in the WID data differ from those of the World Bank for at least four main reasons. First, World Bank estimates mostly focus on consumption (posttax disposable income minus net household saving), while my focus here is on income. The difference between consumption and income can be large, with major implications for levels and trends in inequality in some cases (see for instance [Chancel et al., 2023](#)). Second, the estimates presented here are consistent with national income growth rates, while World Bank estimates are based on surveys and do not attempt to bridge gaps between survey and national accounts aggregates. Third, some of the estimates used in this paper are based on detailed country-specific studies that rely on data sources that may differ from those of the World Bank in a number of countries, including China ([Piketty, Yang, and Zucman, 2019](#)), India ([Chancel and Piketty, 2019](#)), and Latin America ([De Rosa, Flores, and Morgan, 2022](#)). See [Chancel and Piketty \(2021\)](#). Fourth, I use GDP purchasing power parity conversion factors, while the World Bank only corrects for price differences in household final consumption expenditure. For all these reasons, I view the World Inequality Database as a more adequate source for conducting the particular analysis developed in this paper. A more systematic comparison of these two datasets is left to future research.

Figure H1 – The Growing Importance of Income Inequality Within Countries
 Their Decomposition of Global Income Inequality, 1980-2019



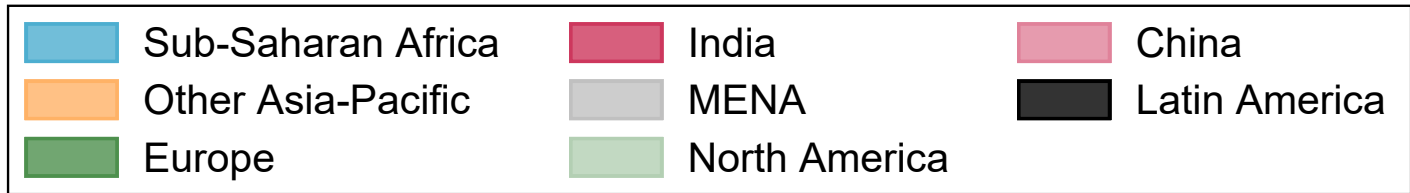
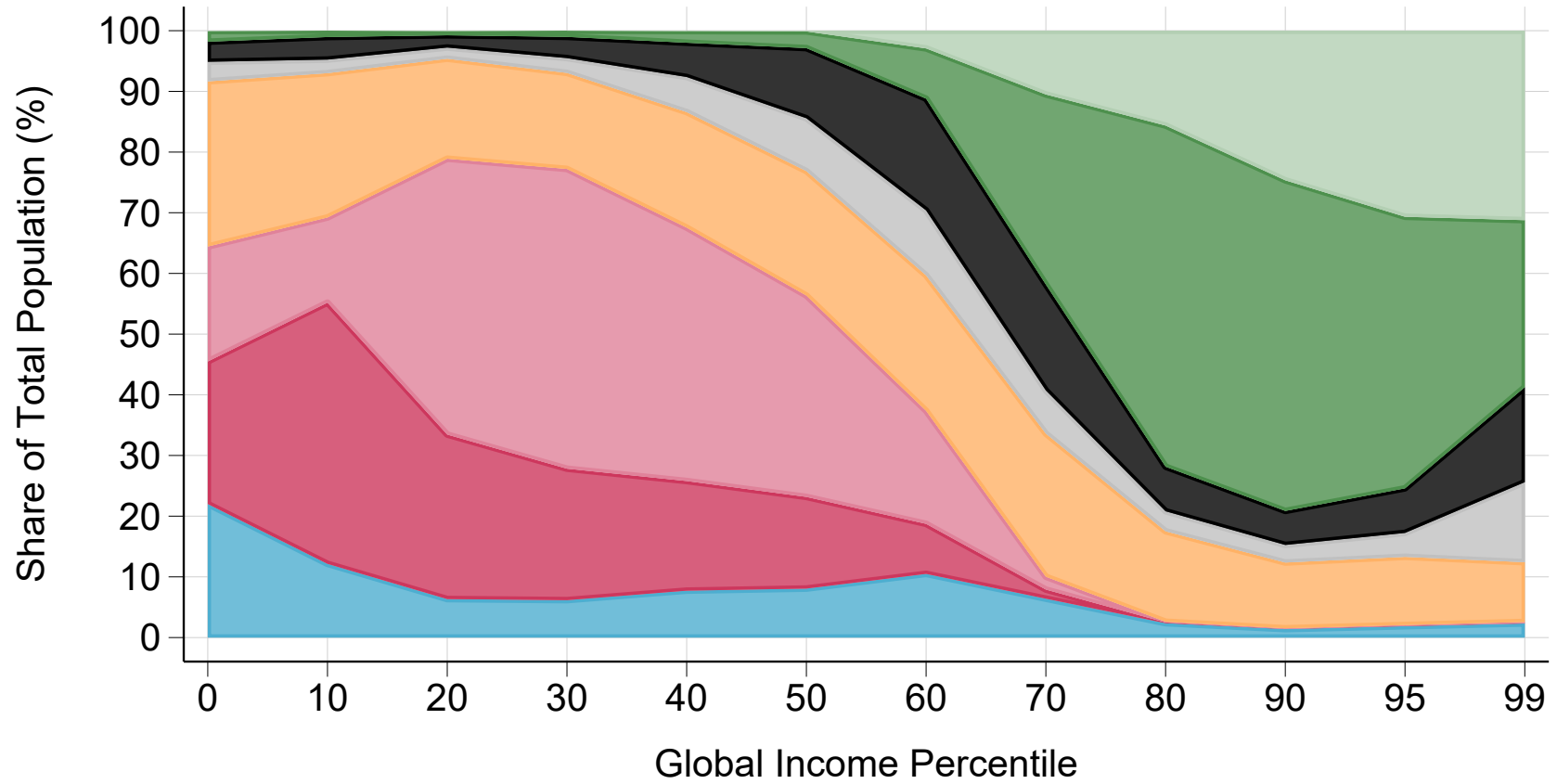
Notes. The figure plots the evolution of the Theil index of global income inequality from 1980 to 2019, as well as its decomposition into a between-country component and a within-country component. Author's computations using data from the World Inequality Database.

Figure H2 – The Rise of Within-Country Inequality: Average Top 10% Pretax Income Share by World Region, 1980-2019



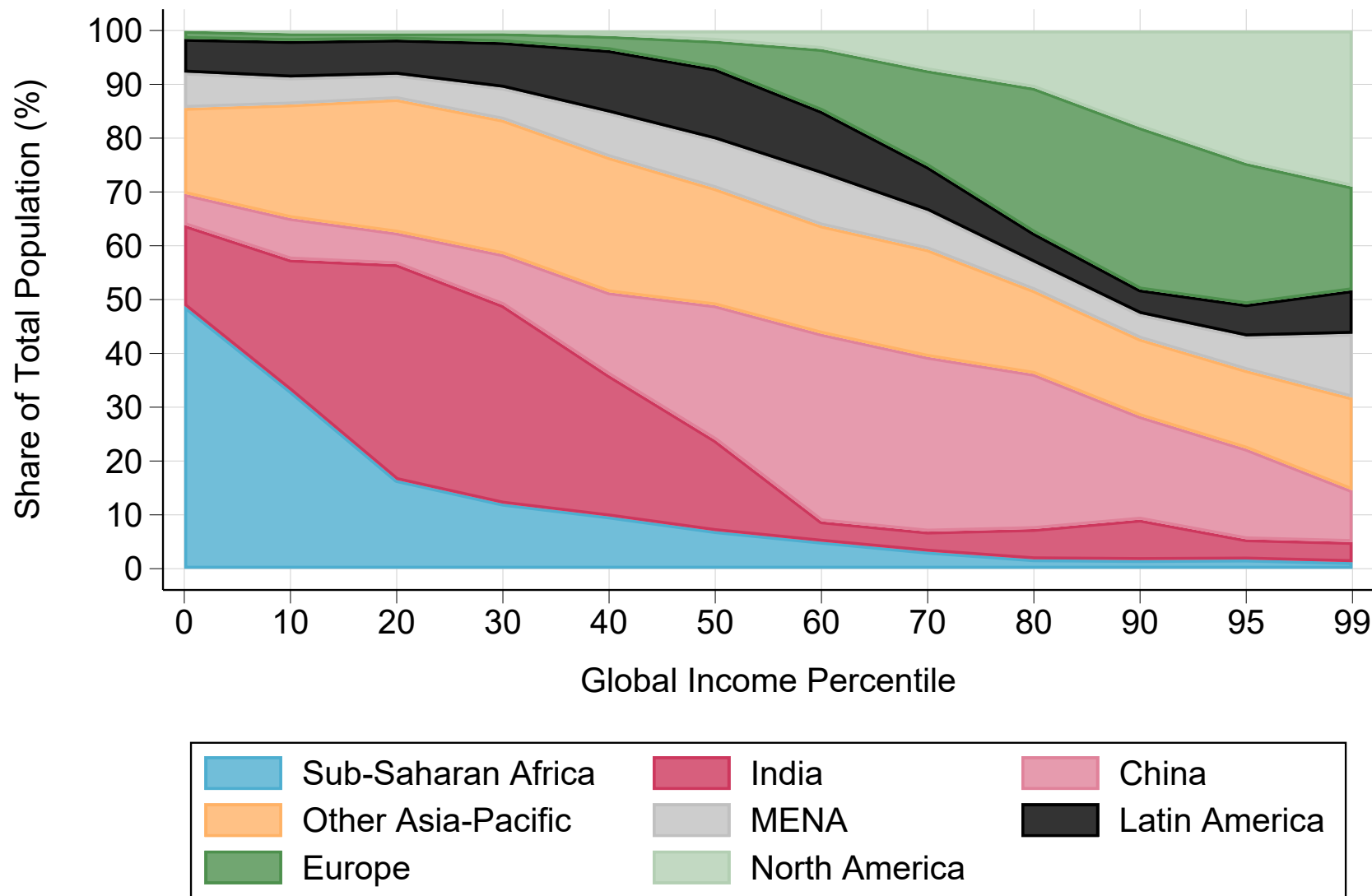
Notes. The figure plots the evolution of the share of pretax income received by the top 10% in selected countries and world regions. Author's computations using data from the World Inequality Database.

Figure H3 – Geographical Breakdown of Global Income Groups, 1980



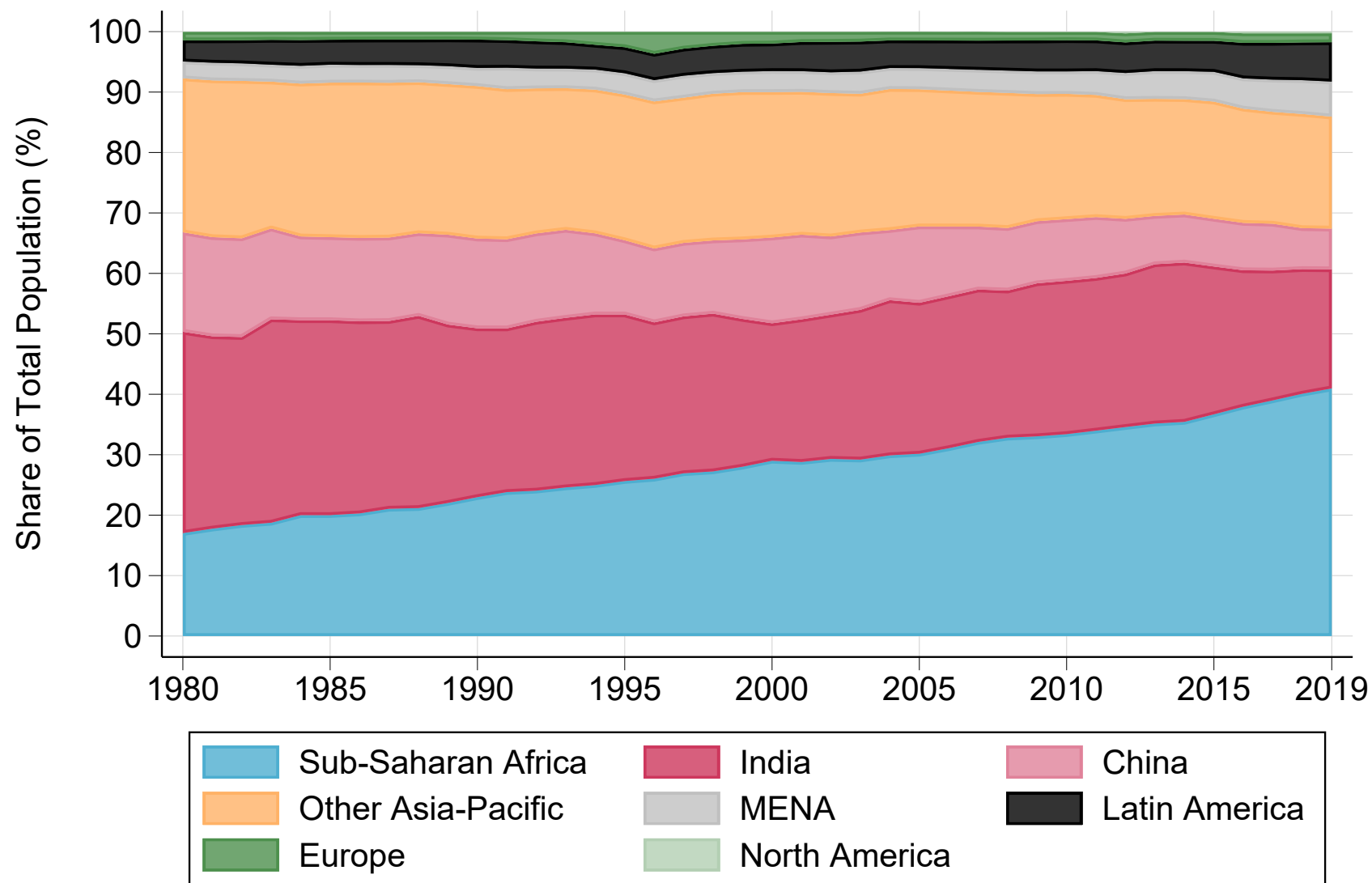
Notes. The figure plots the geographical composition of percentiles of the world distribution of income in 1980. Author's computations using data from the World Inequality Database.

Figure H4 – Geographical Breakdown of Global Income Groups, 2019



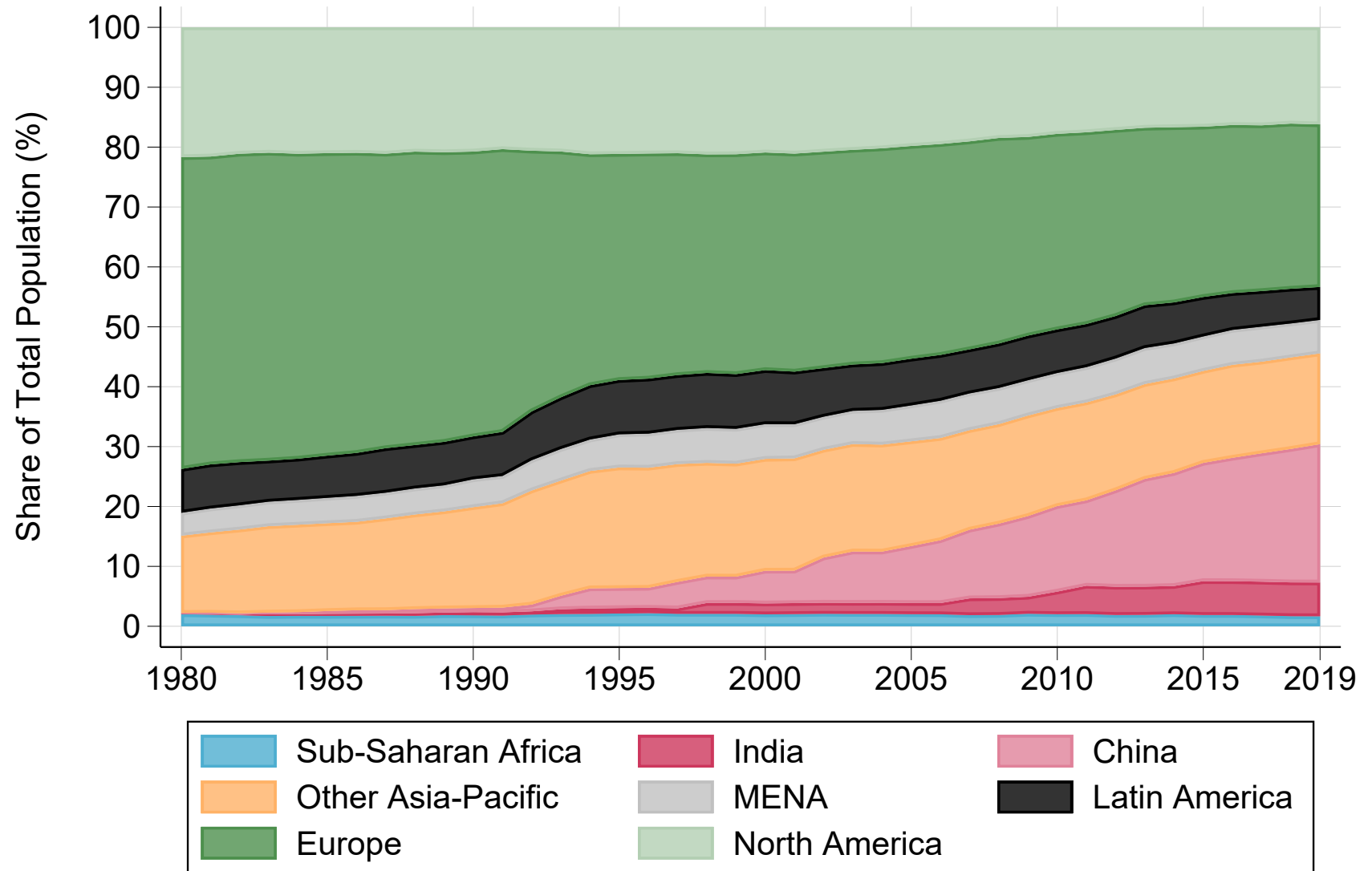
Notes. The figure plots the geographical composition of percentiles of the world distribution of income in 2019. Author's computations using data from the World Inequality Database.

Figure H5 – Geographical Breakdown of the Global Bottom 20%, 1980-2019



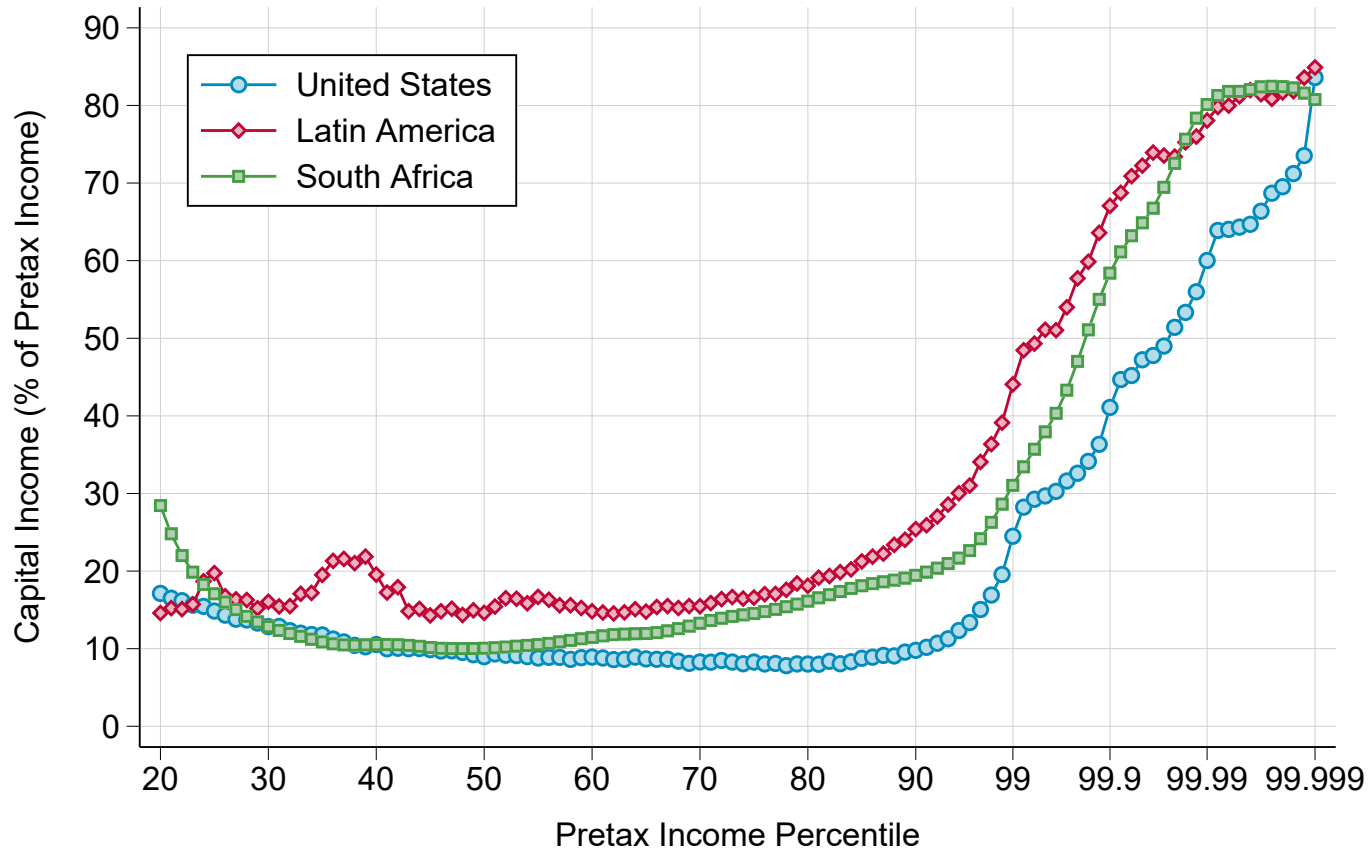
Notes. The figure plots the geographical composition of the world's poorest 20% individuals from 1980 to 2019. Author's computations using data from the World Inequality Database.

Figure H6 – Geographical Breakdown of the Global Top 20%, 1980-2019



Notes. The figure plots the geographical composition of the world's richest 20% individuals from 1980 to 2019. Author's computations using data from the World Inequality Database.

Figure H7 – The Concentration of Capital Income in the United States, Latin America, and South Africa



Notes. The figure plots the capital income share by pretax income percentile in the United States, South Africa, and Latin America in 2019. Capital income is defined as all income other than compensation of employees and mixed income. 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.

