Skills, Technologies, and Development

Juan I. Vizcaino *

The University of Nottingham

Abstract

I study how the productivity of skilled and unskilled labor varies with development. Using harmonized, occupational labor market outcomes for a broad set of countries across the development spectrum, I document that employment in high-skill occupations, or jobs that are relatively more intensive in non-routine cognitive tasks, grows with development. In addition, the income of workers in high-skill occupations falls relative to earnings in low-skill occupations as countries grow richer. To understand the forces driving these findings, I develop a stylized model of the labor market across development. In the model, labor productivity is determined endogenously as a result of the selection of heterogeneous workers into occupations and education. I use a quantitative version of the model to decompose the observed decline in relative labor income between less-developed countries and the US into a component embedded in technologies, or relative skilled labor efficiency, and a fraction due to workers' characteristics, or relative skilled labor quality. I find that relative quality explains 25 percent of the decline in relative labor income, with the remaining fraction due to relative efficiency. In less-developed countries, the relatively few skilled workers are the most productive in performing high-skill jobs, which reduces the magnitude of skill-biased technological progress needed to rationalize the cross-country data by one half when compared to a world where labor quality is purely determined by educational attainment.

Keywords: development, human capital, skilled labor, skill-biased technical change.

JEL Codes: 011, 015, 040

*I thank Rodolfo Manuelli, Francisco Buera, and Ping Wang for their support and guidance. For helpful comments and suggestions, I would also like to thank Yongseok Shin, Limor Golan, and Joseph Kaboski.

1 Introduction

Cross-country differences in standards of living and labor force attributes are noticeable. Understanding what shapes the observed disparities in workers' qualifications and how they translate into different prosperity paths is a central element in the analysis of economic development. The traditional view is that there are sizable gaps in factor-neutral productivity across countries. More recent studies, based on newly available data and improved measuring techniques find that technological progress is biased towards skilled labor, reflecting a shift in demand in favor of workers with higher levels of educational attainment along the development path (Caselli and Coleman (2006), Jones (2014), Caselli (2005), Rossi (2017), Malmberg (2018)). This relative demand shift is needed in order to reconcile the cross-country empirical fact that large improvements in educational attainment are associated with a relatively small decline in the educational skill premium as countries develop.

Strikingly, measured skill bias in cross-country technological efficiency is of such magnitude that models where workers' labor productivity is solely determined by their educational attainment predict that higher-income countries use unskilled labor not only relatively but also absolutely less efficiently.

In this paper I revisit this issue and study how the relative efficiency of the technologies that use skilled and unskilled labor and the quality of skilled and unskilled labor vary with development. Consistent with new data on occupational labor income across the development spectrum, I find that allowing for heterogeneity in workers' abilities reduces the magnitude of measured skill-bias technological progress by one half when compared to a world where differences in the quality of the labor force are purely determined by educational attainment. As a consequence, unlike the most-recent strand of the literature, more-developed countries use both skilled and unskilled labor more efficiently.

The key difference between my work and previous studies is in the modeling and measurement of the relative quality of skilled and unskilled labor. The literature typically assumes that differences in the quality of workers are mostly explained by their educational attainment, which implicitly considers that individuals are homogeneous in their characteristics once differentials in schooling levels are accounted for. In my case, instead, I allow for workers to have heterogeneous attributes, which together with the state of technology shape their occupational and educational choices and determine the quality of skilled and unskilled labor in equilibrium.

My paper also introduces a novel empirical approach. Earlier studies often rely on cross-country data on educational attainment and Mincerian returns to education to disentangle the behavior of relative skilled labor productivity and efficiency along the development path. Thus, splitting countries' labor forces into skilled and unskilled employees requires choosing a minimum level of education that differentiates workers qualitatively into different production factors. In my case, I group workers according to the main occupation they perform, which more closely reflects the task component of their jobs and more clearly reflects qualitative differences between them.

Following this alternative classification criterion, I use harmonized, occupational labor market outcomes for a broad set of countries across the development spectrum and document that employment in high-skill occupations, or jobs that are relatively more intensive in non-routine cognitive tasks, grows with development. In addition, workers earnings in high-skill occupations falls with respect to those in low-skill occupations as countries grow richer, with elasticities in line with those found by studies based on educational attainment and Mincerian returns to education.

To shed light on these findings and disentangle the mechanisms that determine the relative quality and efficiency of skilled labor, I build a general equilibrium model of occupational choice and human capital accumulation through education. The labor demand side is characterized by a representative, cost-minimizing firm that operates the single aggregate technology available in the economy. As in previous studies, the technology features a labor aggregator where skilled and unskilled labor are imperfect substitutes in production. Given this structure, exogenous skill-biased shifts in technological efficiency attract more workers towards high-skill occupations. A novel feature of the model is that the effective supply of skilled and unskilled labor is determined endogenously by workers' occupational choice and their decision to accumulate human capital through education. As a result, the equilibrium productivities of skilled and unskilled labor crucially depend on the properties of the joint distribution of skills in the population.

I perform a mixture of calibration and estimation of the model's deep parameters, including those governing the joint distribution of skills, to match some labor market moments in the US. I use the quantitative version of the model to conduct two exercises that highlight the importance of workers' attributes when making their occupational and educational decisions, and how they are related to skilled-biased technological change.

In the first exercise, I fix the parameters of the joint skill distribution and compute the levels of relative skilled labor efficiency and two model parameters that capture fixed costs of schooling that are required to rationalize the observed shares of skilled workers and educational attainment across countries. As a by-product, I obtain the non-targeted levels of relative labor income in high- and low-skill occupations. Qualitatively, the model reproduces the decline and the pattern of relative labor income in high- and low-skill occupations we observe in the data as countries develop. Quantitatively, my framework explains 70 percent of the total decline, which I decompose into a relative efficiency and a relative quality component. I find that between 25 percent of the observed relative labor income differentials are explained by the relative quality of skilled labor, while the remaining fraction is due to skill-biased technological progress.

A natural outcome of this exercise is to compare the model predictions for technological efficiency gaps between countries. When I do so, I find a sizeable skill-bias in technological progress as countries develop, with relative labor efficiency being more than a hundred times higher in the US than in the set of least developed countries in my sample, and about fifty times higher than in the average country in the second development quartile. Since my sample encompasses a broader set of countries at the lower end of the development spectrum, the latter group is more relevant for comparison with earlier studies. The fact that the relative quality of skilled labor is higher in poor countries dampens the measured differences in skill-bias between rich and poor countries when compared to a model that measures labor quality based exclusively on educational attainment. In a world where labor quality is purely determined by educational attainment, the measured gap in relative efficiency between rich and poor countries would be two times larger, or 120 times higher in the US than in the average country in the second development quartile.

In the second main quantitative exercise, I investigate the role of an educational expansion. To asses the effects of increased access to education on development outcomes, I perform a reduction in the costs of acquiring education for the average country in the first development quartile. The engineered expansion is such that, after the thought policy is in place, the educational attainment levels in least-developed countries are the same as those observed for the average country in the most-developed group. In line with the results of a major educational expansion that took place in Brazil between 1995 and 2014 (see Jaume (2019)), I find that the occupational structure remains fairly unchanged compared to the educational attainment structure, with workers of all educational groups increasingly employed in lower-wage occupations.

In addition, the model predicts a growth in output per worker in the order of ten percent, a relative small number in comparison with the observed income gaps between most- and least-developed nations. The effects on GDP per-worker are at least one-third smaller than what a model that classifies workers into high- and low-skill according to their educational attainment would anticipate. The reason is that, in my model, the expansion benefits workers in high-skill occupations proportionately more, which through general equilibrium effects reduces the fraction of workers in high-skill occupations after the policy is implemented, compensating the initial increase in the effective supply of high-skill labor. On the other hand, in the model based on educational attainment, the expansion leads to a direct increase in the effective supply of high-skill labor, without any reduction in quantities coming through the workers' selection channel.

The quantitative exercises all together highlight the importance the selection of workers into occupations and education together with skill-biased technological progress in shaping differences in standards of living across countries and improving labor market outcomes for both high- and low-skill workers.

Related Literature. This paper contributes to the strand of the economic development literature that studies skill-biased technological differences between countries. An early contribution is by Caselli and Coleman (2006), who propose an aggregate technology framework to unveil cross-country skilled-biased gaps in technological efficiency. Caselli (2016) updates and expands this study to a broader set of countries and other factors besides labor, but still based on an aggregate technology approach. More recently, Malmberg (2018) proposes a novel approach to estimate the relative efficiency of skilled and unskilled labor based on disaggregated trade and industry data. Another recent contribution is given by Rossi (2017), who compares labor market outcomes of immigrants with different levels of educational attainment to identify differences in the relative efficiency and the relative quality of skilled and unskilled labor. The main difference between my work and these studies is that I propose a model-based method to estimate the relative quality of skilled and unskilled labor. In addition, and except for the case of Malmberg, I use occupational attainment instead of educational attainment data to identify qualitative differences between workers.

My paper is also related to a large body of literature that finds evidence of skilled-biased technical change across time and within countries. Katz and Autor (1999) provide a comprehensive survey of this literature, that includes Katz and Murphy (1992), Acemoglu (1998), Autor et al. (1998), Acemoglu (2002).

From a methodological perspective, a paper that is closely related to mine is Lagakos and Waugh (2013). However, in their case, the self-selection of workers of heterogeneous abilities into different sectors is used to explain differences agricultural productivities between rich and poor countries. Moreover, in their work countries only differ in an economy-wide efficiency parameter and there is no role for skill-biased technological progress.

From a broader point of view, this paper is also related to the literature that measures the contribution of human capital to development. Earlier contributors to this literature are Mankiw et al. (1992), Klenow and Rodriguez-Clare (1997), and Hall and Jones (1999). More recently, Erosa et al. (2010), Manuelli and Seshadri (2014), Jones (2014) find human capital to be an important factor in explaining disparities in wealth levels between countries. In my case, instead, I focus on productivity differentials between two groups of workers that are different in their nature due to the tasks they more commonly perform, rather than studying the role of labor quality as a whole.

The paper is organized as follows. Sections 2 and 3 present the empirical analysis, including sources, a detailed description of the data, and robustness checks. Section 4 presents the model. In order to build up some intuition, I show in Subsection 4.2 the model's basic structure, given by the workers' selection problem into occupations according to their ***unobservable characteristics***. Subsection 4.5 shows the properties of the joint skill distribution that are key to understand what mechanisms in the model generate the patterns we observe in the data, while Subsection 4.6 builds endogenous human capital accumulation into the model. Section 5 describes the strategy followed to estimate the model parameters, while in Section 6.1 I present my main quantitative exercises. Section 7 concludes.

2 Empirical Analysis

2.1 Data Description

My main data source is the International Labor Organization (ILO) [1]. In particular, I use cross-country, harmonized, occupational level data on average nominal labor income of employees [2], average weekly hours worked per employee, and number of people employed.

The ILO provides occupational statistics at the one-digit level of aggregation following the International Standard Classification of Occupations (ISCO). I focus on countries

[1]See http://www.ilo.org/ilostat. Appendix A presents a description of the ILO data, issues and treatment.

[2]More precisely, the ILO has data on labor earnings. The concept of earnings, as applied in wages statistics, relates to gross remuneration in cash and in kind paid to employees, as a rule at regular intervals, for time worked or work done together with remuneration for time not worked, such as annual vacation, other type of paid leave or holidays. Earnings exclude employers' contributions in respect of their employees paid to social security and pension schemes and also the benefits received by employees under these schemes.

that have data for the latest version of the ISCO classification, ISCO-08 [3]. ILO's aggregated statistics are based on micro-data that is representative of the labor force of each country. The micro-data comes from surveys or studies that vary between countries, depending on availability. Data sources include Labor Force Surveys, Employment Surveys, Establishment Surveys, Household Surveys, Insurance Records, and Administrative data. For countries with multiple data sources available, I prioritize labor force, employment and household surveys [4].

My main analysis takes into account workers of all ages [5] and both sexes [6]. Regarding the time frame, I focus on countries with information available between 2000 and 2018 and average data across time when information for multiple years is available for a preferred data source in any given country.

My cross-country empirical analysis and calibration uses data on GDP, capital, number of employees, and average hours worked per worker, which I obtain from Penn World Table 9.0 [7]. From this broad sample of GDP per worker for 182 countries I calculate the 25th, 50th, and 75th percentiles of GDP per worker, which I later on use to classify the 81 countries in my ILO sample into four development quartiles.

I exclude from my analysis countries intensive in natural resources, as defined in Appendix B, and those with a population smaller than one million. My sample has data for 81 countries and covers one-third of the world's population, including 20 countries in the first, 18 in the second, 23 in the third, and 20 in the fourth development quar-

[3] I discard data from ISCO-88 to avoid issues arising from methodological changes in occupational aggregation between the two ISCO releases. Even though the titles of the ten major occupational groups are the same under ISCO-08 and ISCO-88, some minor occupational groups were moved between major groups when ISCO-88 was updated to ISCO-08 to reflect the effects of technology on professional, technical, and clerical work. My sample is reduced in eleven observations by discarding countries with data available for ISCO-88 only.

[4] In these studies income is reported by workers rather than by employers, as in Administrative Data or Insurance Records. Thus, these type of studies provide a more accurate description of labor market outcomes for less-developed countries where the informal economy represents a large share of employment and it is more common for employees to receive payments under the table.

[5] It is customary in the empirical labor literature to restrict the analysis to either prime-aged workers or to workers with high labor force attachment. That is not possible in my case, as the ILO only provides summary statistics from for workers of all ages. My unit of analysis is employees instead of employed people, since the latter classification includes not only employees receiving remuneration but also working proprietors and unpaid family workers.

[6] It might be a concern that the inclusion of women might affect the analysis. For example, the level of attachment to the labor force could potentially differ for women who work in high- and low-skill occupations, especially at low development levels. Another concern could arise from the a gender wage gap that varies across development, especially in contexts of weak labor institutions. To tackle these type of issues I perform a robustness analysis in Section 3.2.

[7] The variables used are Output-side real GDP at chained PPPs in millions of 2011 US$ *(rgdpo)*, people engaged *(emp)*, average annual hours worked by persons engaged *(avh)*, and capital stock at 2011 national prices in millions of 2011 US$ *(rnna)*. See Feenstra et al..

tiles. Regarding regional coverage, my sample includes 17 Advanced Economies, 10 from East Asia and Pacific, 10 from Europe and Central Asia, 15 from Latin America and the Caribbean, 4 from Middle East and North Africa, 4 from South Asia, and 15 from Sub-Saharan Africa.

2.2 Occupational Aggregation.

The ILO provides labor market data by occupation at the one-digit level, ISCO-08's highest degree of aggregation. At that level occupations are collected into ten major groups [8]. I discard workers in Armed Forces Occupations, since their labor market outcomes might not necessarily be determined by market forces.

To simplify the analysis, I put together the nine major occupational groups left into two broader categories, which I label as high-skill and low-skill occupations. Behind this aggregation procedure lies the assumption that occupations are qualitatively different between groups. One way of capturing differences in nature between groups is to separate jobs according to the tasks that are more commonly performed in them.

Acemoglu and Autor (2011) provide a clear characterization of the task intensity performed by each occupation. While professional, managerial and technical jobs are relatively more intensive in abstract, non-routine cognitive tasks, clerical and sales occupations, production workers and operators, and service workers more commonly perform routine cognitive, routine manual, and non-routine manual tasks. In turn, they document that for the US, job abstract intensity is positively and monotonically related to the skill level of an occupation, as approximated by the average wage of the workers in them. In addition, workers with higher educational attainment levels tend to concentrate proportionately more in occupations that are abstract task intensive.

Non-routine cognitive task intensity of occupations seems to provide a good criterion to group them into high- and low-skill groups. As an additional robustness check to this criterion, I compare occupational wages at different development stages to see if high-skill occupations are indeed the ones with higher wages, as Acemoglu and Autor document to be the case for the US. To do so, I use ILO data to compute hourly labor income at purchasing power parity by country and occupation. I group the countries in my sample into four development quartiles, using the GDP per capita thresholds described in Subsection 2.1 above, and calculate the median hourly labor income at PPP for each quartile.

[8]These major categories are: 1. Managers, 2. Professionals, 3. Technicians and Associate Professionals, 4. Clerical Support Workers, 5. Service and Sales Workers, 6. Skilled Agricultural, Forestry and Fishery Workers, 7. Craft and Related Trades Workers, 8. Plant and Machine Operators, and Assemblers, 9. Elementary Occupations, and 10. Armed Forces Occupations.

Table 1: Occupational Skill-Intensity
(median hourly labor income at PPP in US$ of 2011 by development quartile)

Broad Group	High Skill			Low Skill					
ISCO Occupation Title	*1. Managers*	*2. Professionals*	*3. Technicians*	*4. Clerical*	*5. Machine Operators*	*6. Craft Workers*	*7. Service Workers*	*8. Skilled Agricultural Workers*	*9. Elementary Occupations*
Bottom Quartile	4.8	3.8	3.0	2.7	1.9	1.9	2.7	2.3	1.7
Second Quartile	7.2	6.1	5.0	3.7	2.5	2.4	2.7	3.0	2.2
Third Quartile	21.4	14.5	10.0	6.9	4.8	4.3	5.8	5.9	4.1
Top Quartile	39.1	32.3	25.3	22.2	18.2	15.8	20.6	19.9	18.7

As Table 1 above shows, managers, professionals and technicians are indeed the highest earning occupations at all development levels. Thus, based both on abstract task intensity and the fact that they exhibit a higher wage level at all development quartiles I choose managers, professionals and technicians to be the occupations in my high-skill group [9].

2.3 Skill-Premium Definition.

I here explain how I construct my skill-premium measure, which I call occupational skill-premium. I start by computing hourly labor income for each of the nine occupational categories by diving monthly labor income, expressed at local currency units, by average monthly hours worked. The latter is obtained using data on average weekly hours worked and assuming that individuals work four weeks per month.

I define my skill-premium measure to be the ratio of the employment-weighted average hourly labor income for the ISCO categories in high-skill occupations with respect to the corresponding employment-weighted average labor income for the occupations in low-skill occupations.

I consider this to be an improvement with respect to previous work for several reasons. First, previous studies rely on information on cross-country Mincerian returns to education and define their skill-premium to be the return to completing certain level of schooling, which could either be elementary school, secondary school, or university. As discussed by the recent empirical labor literature, this approach has the disadvantage that in the last three decades, a period often characterized by an acceleration in technological progress, educational attainment has lost explanatory power in wage regressions. At the same time, the explanatory power of occupations in accounting for wage differences across workers has significantly increased in this period of time, not only in developed, but also in developing countries [10].

[9]To be precise, high-skill occupations include ISCO-08's major groups 1,2, and 3. low-skill occupations are those in groups 4,5,6,7,8, and 9.

[10]For example,Acemoglu and Autor (2011) document that even though the university/ secondary

Second, there are other characteristics that affect worker's skills besides their formal education, like their own innate ability to perform different tasks, their general health status, and the learning that might be acquired on the job, among others. This set of characteristics are better captured by observed labor income rather than by expected average returns to education. Third, the development literature has not reached an agreement on what level of educational attainment should be considered to be the minimum in order to classify workers as skilled. The main implication of choosing different educational attainment levels is that the higher the educational attainment level required to classify workers as skilled, the larger the variability of the resulting aggregate human capital stock across countries, increasing its explanatory power in development accounting exercises.

3 Empirical Results.

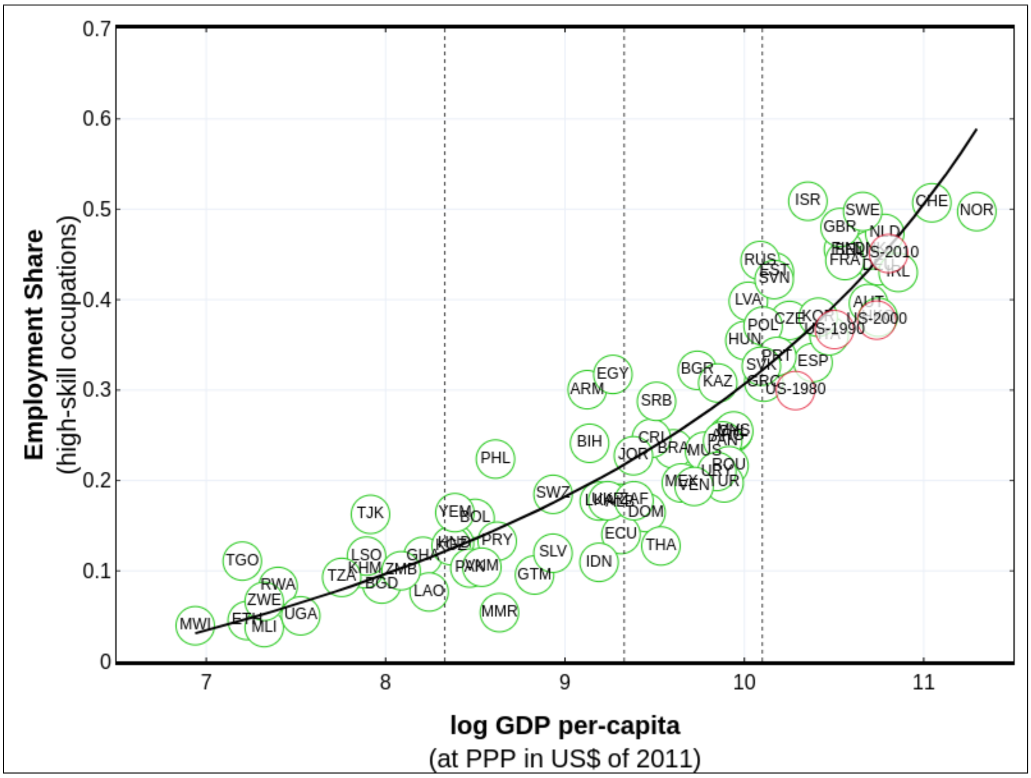
I begin by studying the evolution of occupational employment at different development levels. To that end, I calculate the fraction of total employment in high-skill occupations for the countries in my sample and plot them against their corresponding GDP per capita. In addition, I fit a regression for the employment share in high-skill occupations on a fractional polynomial model in GDP per capita, which is represented by the solid black line in the figure [11]. The vertical dashed lines separate the sample into development quartiles. The results are presented in Figure 1 below.

The main message from Figure 1 is that employment in high-skill occupations rises as countries grow richer. Quantitatively, the average share of employment in high-skill occupations rises from 8.5 percent to 18.1, 27.5, and 42.8 respectively as we move from the first to the fourth development quartile. When compared to educational attainment data, the share of employment in high-skill occupations is higher than the average fraction of workers with university complete (1.98 percent) and lower than the proportion of individuals with secondary education complete (15.8) at the first development quartile. However, it grows at a faster speed that both measures with development, reaching an average of 43.5 percent for countries in the highest quartile of development, while the average fractions of individuals with secondary and university complete total 38.8 and

school wage premium has monotonically increased since the 1970s in the US, these changes in wage levels and the distribution of wages have been accompanied by systematic, non-monotone shifts in the composition of employment across occupations, with rapid simultaneous growth of both high education high wage occupations and low education, low wage occupations in the United States and the European Union.

[11] The best fitting model is a one-dimensional polynomial of the form $\frac{L_{hs,c}}{L_c} \quad \beta_0 \quad \frac{\beta_1}{log(y_c) \quad y_0} \quad \epsilon_c$. Estimated coefficients are both jointly and individually statistically significant at the one percent level and given by: β_0 0.36, β_1 2.90, and y_0 14.35.

Figure 1: Employment Share in High-Skill Occupations vs GDP per capita.



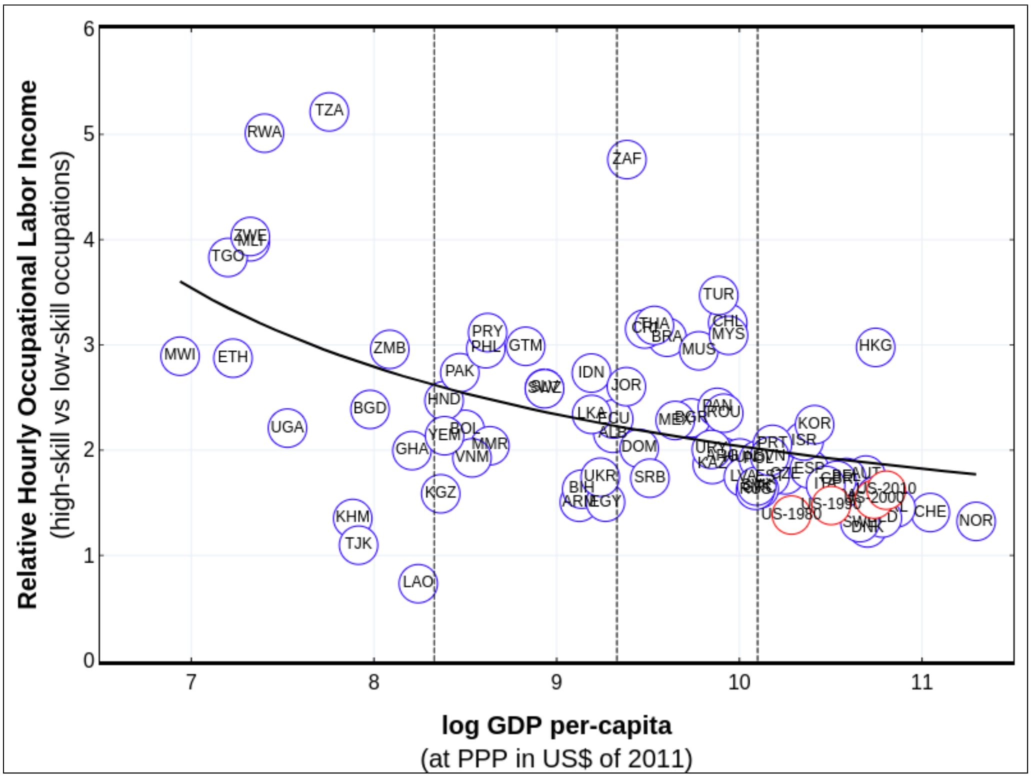
13.7 percent, respectively.

I proceed to study the behavior of the occupational skill-premium across the development spectrum. Therefore, I plot each country's occupational skill-premium and GDP per capita in Figure 2 below. As in the case of the employment shares, the solid black line in the figure represents the best fitting fractional polynomial model [12] and the vertical dashed lines separate the sample into development quartiles.

The main message from Figure 2 is that there is a negative relationship between the occupational skill-premium and GDP per capita. As the regression line shows, this relationship is fairly non-linear, exhibiting a steeper decline at lower development levels. On average, the occupational skill-premium falls from 2.5, to 2.3, 2.1 and 1.6 as we move from the first to the fourth quartile of my sample. Compared to the skill-premium measures based on Mincerian returns to education used by Caselli and Coleman, these numbers lie in between the estimates that takes as high-skill workers those with primary complete and the one that defines as high-skill workers those with university complete, respectively. The occupational skill-premium lies below the educational skill-premium

[12]The best fitting model is a one-dimensional polynomial of the form $\frac{w_{hs,c}}{w_{ls,c}} \quad \beta_0 \quad \frac{\beta_1}{log\text{p}y_c\text{q} \quad y_0} \quad \epsilon_c$. Estimated coefficients are both jointly and individually statistically significant at the one percent level and given by: β_0 0.47, β_1 9.61, and y_0 3.86.

Figure 2: Relative Hourly Labor Income vs GDP per capita.



metric that takes workers with university or above as the high-skill group [13]. Figure 3 below presents a comparison between the occupational skill-premium and the measures of skill-premia based on educational attainment, assuming that skill premium elasticities are constant across development for all measures. I expand on the non-linearities of the skill-premium elasticity in the following section.

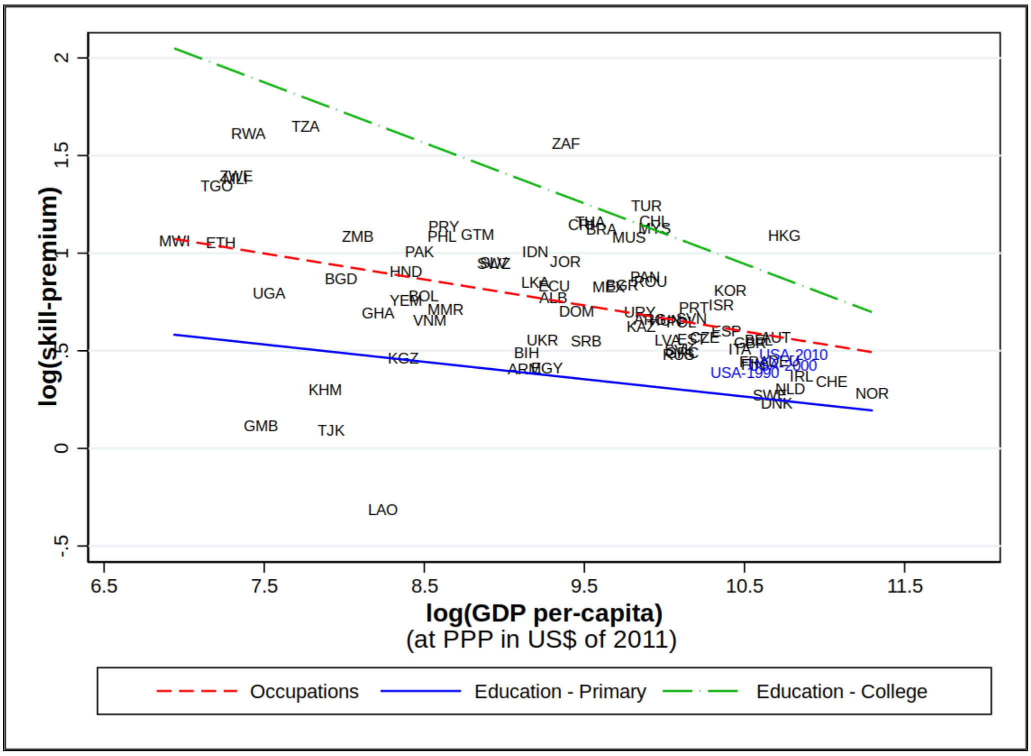
3.1 Occupational Skill-Premium Non-Linearities Across Development.

I here study in further detail the non-linear relationship between the occupational skill-premium and GDP per capita. In order to quantitatively asses these non-linearities in a more tractable framework than the one suggested by the fractional polynomial fit, I estimate the skill-premium development elasticity by running the following linear regression

$$log \left(\frac{w_{hs,c}}{w_{ls,c}} \right) \quad \beta_0 \quad \beta_1 \ log(y_c), \qquad (1)$$

[13] When workers with primary education or higher are defined as high-skill Caselli and Coleman's measure falls from 1.7 to 1.5, 1.4 and 1.3 as we move from the first to the fourth quarter of income per capita in their sample. When secondary complete is taken as the minimum level of school attainment to be considered high-skill, there measure falls from 3.7 to 2.7, 2.4 and 2.0, respectively. If the educational attainment threshold is university complete instead, the corresponding numbers are 9.8, 4.7, 4.0, and 2.9

Figure 3: Development Elasticity for Different Skill-Premium Measures



and compare it with a model where I let the elasticity vary across development quartiles, by fitting the following OLS regression to the data

$$log\left(\frac{w_{hs,c}}{w_{ls,c}}\right) = \beta_0 + \beta_1 \, log(y_c) + \sum_{q=2}^{4} \mathbb{1}_{\{c \in q\}} \beta_q \, log(y_c) \,. \tag{2}$$

In Equations 1 and 2 , $log\left(\frac{w_{hs,c}}{w_{ls,c}}\right)$ is the natural logarithm of the occupational skill-premium in country c and $log(y_c)$ is the natural logarithm of real GDP per capita at PPP in constant US$ of 2011 for country c. In Equation 2, q is quartile indicator, the higher the more developed a country is, and the corresponding quartile indicator functions take the value of one $\mathbb{1}_{\{c \in q\}}$ if country c belongs to development quartile q.

Equation 1 is standard and does not require further discussion. Equation 2, β_1 measures the occupational skill-premium elasticity with respect to GDP per capita for countries in the first development quartile, which I choose to be the base. In this model β_2, β_3, and β_4 measure the change in the base elasticity as countries move to higher development levels, represented by the second, third, and fourth quartiles, respectively. For these coefficients, classical statistical tests of significance (i.e. t-tests) show if the estimated elasticity for the corresponding quartile is statistically different from the base one.

The estimation results for Equations 1 and 2 are presented under the names of Model

(1) and Model (2) in Table 2 below.

Table 2: Development Elasticity of the Occupational Skill-Premium
(quantitative assessment of non-linearities)

Variables	**Model (1)** $log\ \frac{w_{hs}}{w_{ls}}$	**Model (2)** $log\ \frac{w_{hs}}{w_{ls}}$	**Model (3)** $log\ \frac{w_{hs}}{w_{ls}}$
$log(y_c)$	-0.127***	-0.391***	-0.416***
	(0.035)	(0.118)	(0.097)
$\mathbb{1}_{\{c \in 2\}} \times log(y_c)$	-	0.051**	0.054***
	-	(0.019)	(0.018)
$\mathbb{1}_{\{c \in 3\}} \times log(y_c)$	-	0.076***	-
	-	(0.006)	-
$\mathbb{1}_{\{c \in 4\}} \times log(y_c)$	-	0.072*	-
	-	(0.033)	-
$\mathbb{1}_{\{c \in 3 \vert c \in 4\}} \times log(y_c)$	-	-	0.080***
	-	-	(0.025)
Constant	1.949***	3.906***	4.101***
	(0.329)	(0.923)	(0.754)
Observations	80	80	80
R-squared	0.146	0.254	0.253
Adjusted R-squared	0.135	0.214	0.223
Prob >F	0.0005	0.0002	0.0001

[*] Standard errors in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

Model (1) confirms that there is a statistically significant negative relationship between GDP per capita and the skill-premium. The estimated elasticity is -0.12 and is statistically significant at the one percent level. The model implies that the skill-premium falls from 2.4 to 1.5 when countries move from the median GDP per capita in the first quartile to the corresponding one in the fourth quartile.

Model (2) shows the skill-premium elasticity with respect to GDP per capita is not only negative, but also that it declines in absolute terms as countries move from the first to the fourth development quartile. Measured elasticities are -0.339, -0.294, -0.275, and -0.279 for the first, second, third, and fourth quartile, respectively, and are all statistically significant, at least at the ten percent level. The model implies that the skill-premium falls from 2.20 to 1.65 when countries move from the average GDP per capita in the first quartile to the average income per capita in the fourth quartile.

A statistical test of joint linear restrictions rejects the nulls that $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, and $\beta_1 = \beta_2 = \beta_3 = \beta_4$, both at the 1 percent level, and that $\beta_2 = \beta_3 = \beta_4 = 0$ at the 5 percent level. However, the null that $\beta_2 = \beta_3 = \beta_4$ can not be rejected at ten percent level, which favors the alternative that at least one pair of these coefficients are equal.

As a consequence, I estimate four additional regressions that capture all possible combinations of equality between these three coefficients. The best specification, as measured by the highest adjusted R^2 attained, is given by the model where $\beta_1 \quad \beta_2 \quad (\beta_3 \quad \beta_4)$, which I present in third column of Table 2 and name Model (3).

Under the specification given by Model (3), the skill-premium elasticity falls from -0.362 in the bottom quartile of development to -0.314 in the second quartile, stabilizing at -0.293 in the third and fourth quartiles. These elasticities are all statistically significant at the one percent level and the model predicts a decline in the skill-premium from 2.21 to 1.65 as we move from the median GDP per capita in the poorest to quartile to the median GDP per capita in the richest quartile.

As a summary, I find that there is a negative, statistically significant relationship between the occupational skill-premium and development. This relationship is fairly non-linear, exhibiting a steeper decline at the first development quartile. The best model specification to measure these non-linearities implies an predicted cccupational skill-premium that falls from 2.21 to 1.65 as we move from the poorest to the richest development group [14].

3.2 Robustness.

In this section I perform sensitivity exercises to asses the robustness of the quantitative results presented in Section 3.1 above. In the first set of robustness checks I explore to what extent my results are driven by the criterion used to group occupations into broad categories and the role of extreme occupational skill-premium values. The second group of exercises analyze the quantitative relevance of institutions, how the results change if I exclude female workers from my sample, and the role of hours worked across development.

The first set of checks is to guarantee that my results are not driven by the criterion used to define high-skill occupations or by countries with relatively high or low occupational skill-premium values. The spirit of the second group of exercises is to disentangle the underlying mechanisms that drive the decline in the occupational skill-premium across development so that at the time of writing a model one can focus on frameworks that shed light on those forces.

[14] I estimate the same regressions using the data on GDP per-worker and educational skill-premium provided by Caselli and Coleman and find a statistically significant negative elasticity for their three measures. In this case, the best specifications throw skill-premium elasticities that do not vary with development. Estimated elasticities are -0.32, -0.20, and -0.08, depending is the schooling threshold for high-skill workers is university, secondary, or primary complete, respectively.

Appendix D presents a detailed description on how this robustness checks are performed. As a summary, neither changing the criterion to classify occupations into high- and low-skill, nor excluding extreme values or adding institutional controls produce any major qualitative changes to the results, with the best model that fits the data still being the one in which the occupational skill-premium falls with development until it reaches the third GDP per capita quartile. Quantitatively, the estimated elasticities are in the same range as those reported in Table 2 above, increasing modestly in absolute terms if institutional controls are added or if the lowest median wage occupation in the high-skill group is included in the low-skill group. Shifting occupations in the margin of the low-skill group to the high-skill group reduces the predicted skill-premium and the corresponding elasticities, but again, only modestly.

Moving to the second set of exercises, the exclusion of women does not produce any major qualitative or quantitative changes in the results. The best model is still given by the one were the occupational skill-premium elasticity declines with development until countries reach the third quartile of GDP per capita. The estimated elasticities are fairly similar as those presented in Table 2. This robustness check suggests that, for example, one can safely discard frameworks that exploit the role of women in the labor force and their attachment to the labor market when the object of study is the decline in occupational skill-premium across development.

Computing the occupational skill-premium without controlling for hours worked leads to similar quantitative and qualitative results. The best statistical fit is still given by the model where elasticities decline with development. When compared to the estimated that control for hours worked, measured elasticities decline marginally in absolute terms, but are still in line with those presented in Table 2. This is mainly due to the fact that hours worked exhibit a higher decline in low- than in high-skill occupations across development. As in the previous case, this suggests that the main driver of the decline of the occupational skill-premium is not hours worked and one can abstract from models whose main focus is on the intensive margin of labor.

4 Model

In what follows I build a stylized model that is useful to understand the labor market structure and evolution through development and allows me to understand the main mechanisms that account for the observations presented in Section 3 above. The ultimate goal is to use the model to understand the determinants of relative labor productivity in high- and low-skill occupations as countries grow richer.

4.1 Environment

The economy is populated by a continuum of individuals of measure one. They are endowed with a unit of time and a pair of occupational-specific labor productivities $z = \{z_l, z_h\}$, where z_h (z_l) represents a realization of a worker's labor productivity in high-skill (low-skill) occupations. Labor productivity is jointly distributed with cumulative distribution function $G(z_h, z_l)$, and $g(z_h, z_l)$ represents the corresponding probability density function. Considering that in my empirical analysis I find that hours worked are not the main driving force of the decline in the occupational skill-premium we observe in the data, I assume that workers provide their time to the labor market fully and inelastically.

Output is produced by combining occupational labor services, according to the technology

$$y = \left[(A_l L_l)^{\left(\frac{\sigma-1}{\sigma}\right)} + (A_h L_h)^{\left(\frac{\sigma-1}{\sigma}\right)}\right]^{\left(\frac{\sigma}{\sigma-1}\right)} \tag{3}$$

where $L_l = \int_{\mathcal{L}} z_l \left(\int_0^{\infty} g(z_h, z_l) dz_h\right) dz_l$ and $L_h = \int_{\mathcal{H}} z_h \left(\int_0^{\infty} g(z_h, z_l) dz_l\right) dz_h$ are total labor output in low- and high-skill occupations, σ represents the elasticity of substitution between skill types, A_l and A_h are occupation-specific productivity parameters that transform labor outputs into labor services, and $\mathcal{L}$ and $\mathcal{H}$ denote the set of individuals who work in low- and high-skill occupations, respectively.

Assuming that the aggregate technology is operated by a cost minimizing firm that acts competitively in both labor markets, relative wages are given by

$$\frac{w_h}{w_l} = \left(\frac{L_h}{L_l}\right)^{\left(-\frac{1}{\sigma}\right)} \left(\frac{A_h}{A_l}\right)^{\left(\frac{\sigma-1}{\sigma}\right)}. \tag{4}$$

Fixing the relative supply of skilled labor, if labor types are are substitutes in production ($\sigma > 1$), economic development processes characterized by skilled-biased technological change ($\uparrow A_h > \uparrow A_l$) lead to an increase in relative efficiency wages. On the other hand, if labor services are complements in production ($\sigma < 1$), relative efficiency wages grow with development if technological progress is unskilled labor biased ($\uparrow A_l > \uparrow A_h$). In the Cobb-Douglas case ($\sigma = 1$), irrespective of its nature, technological progress has a neutral effect on relative wages.

4.2 A Basic Roy Model for The Labor Market.

Assume markets are competitive and workers can freely select their occupation. Let w_h and w_l represent wages per efficiency unit of labor in high- and low-skill occupations. As is standard in Roy models, individuals choose to work in high-skill occupations if their labor income W_h^i is higher than in low-skill occupations W_l^i, or

$$W_h^i > W_l^i$$

$$w_h z_h^i > w_l z_l^i.$$

Re-arranging terms

$$\frac{z_h^i}{z_l^i} > \frac{w_l}{w_h},$$

and agents choose to work in high-skill occupations if their comparative advantage in them is higher than the inverse of relative wages in high- and low-skill occupations. For very low high-to-low skill occupations relative wage levels, only workers with a very high comparative advantage in high-skill occupations choose them. As wages in high-skill occupations grow with respect to those in low-skill occupations workers with lower comparative advantage in high-skill occupations choose to work in them.

In this context, the set of workers who choose high-skill occupations is

$$H = \left\{ i \in [0,1] : w_h z_h^i > w_l z_l^i \right\},$$

and their corresponding effective labor supply is

$$\begin{aligned} L_h &= \int_H z_h \left(\int_0^\infty g(z_h, z_l) dz_h \right) dz_l \\ &= \int_0^\infty z_h \int_0^\infty g(z_h, z_l \wedge w_h z_h > w_l z_l) dz_l dz_h \\ &= \pi_h \int_0^\infty z_h \int_0^\infty g(z_h, z_l \mid w_h z_h > w_l z_l) dz_l dz_h \\ &= \pi_h \int_0^\infty z_h \int_0^{z_h \frac{w_h}{w_l}} g(z_h, z_l) dz_l dz_h = \pi_h \, E\left[Z_h \mid w_h z_h > w_l z_l \right]. \end{aligned} \tag{5}$$

Similarly, the effective labor supply of workers in low-skill occupations is

$$L_l = \pi_l \, E\left[Z_l \mid w_h z_h \le w_l z_l \right]. \tag{6}$$

where π_h denotes the fraction of workers in high-skill occupations, given by

$$\pi_h = Prob\left(w_h Z_h > w_l Z_l \right) = \int_{-\infty}^{\infty} \int_0^{z_h \frac{w_h}{w_l}} g(z_h, z_l) dz_l dz_h, \tag{7}$$

and $\pi_l = (1 - \pi_h)$.

To ease notation, in what follows I call $\bar{z}_h = \mathbb{E}\left[Z_h | w_h z_h > w_l z_l\right]$ and $\bar{z}_l = \mathbb{E}\left[Z_l | w_h z_h \leq w_l z_l\right]$. Thus, from now on $L_h = \pi_h \bar{z}_h$ and $L_l = \pi_l \bar{z}_l$.

As we can see from Equation above, a rise in the high-to-low skill wage ratio $\frac{w_h}{w_l}$ leads to a rise in the faction of workers who choose high-skill occupations. This is as a consequence of the standard selection mechanism described above, and is reflected in an increase in the upper limit of integration in the inner integral of Equation . More importantly, this occurs independently of the properties of the joint skill distribution $g(z_h, z_l)$.

Continuing with the characterization of the main model objects of interest, let $M_h(z_h)$ denote the cumulative labor productivity distribution function conditional on workers selecting high-skill occupations and $m_h(z_h)$ its corresponding pdf. The former is given by

$$
\begin{aligned}
M_h(z) &= Prob\left(Z_h \leq z | w_h z_h > w_l z_l\right) \\
&= \frac{Prob\left(Z_h \leq z, w_h z_h > w_l z_l\right)}{Prob\left(w_h z_h > w_l z_l\right)} \\
&= \frac{1}{\pi_h} \int_0^{z} \int_0^{z_h \frac{w_h}{w_l}} g(z_h, z_l) dz_l dz_h \\
&= \frac{1}{\pi_h} \int_0^{z} g\left(z_h, z_h \frac{w_h}{w_l}\right) dz_h.
\end{aligned}
$$

Similarly, the corresponding labor productivity cdf conditional on choosing low-skill occupations is

$$
M_l(z) = \frac{1}{\pi_l} \int_0^{z} g\left(z_l \frac{w_l}{w_h}, z_l\right) dz_l.
$$

These two distributions are the relevant empirical objects of interest, in consideration of the fact that the researcher observes certain characteristics of workers conditional on their occupational decision. Finally, average labor income in high- and low-skill occupations are represented by

$$
\bar{W}_h = w_h \mathbb{E}\left[Z_h | w_h z_h > w_l z_l\right] = w_h \bar{z}_h,
$$

and

$$
\bar{W}_l = w_l \mathbb{E}\left[Z_l | w_h z_h \leq w_l z_l\right] = w_l \bar{z}_l.
$$

As a consequence, the model's counterpart for the occupational skill-premium, or the ratio of average labor income in high- and low-skill occupations presented in Figure 2 is

$$\frac{\bar{W}_h}{\bar{W}_l} = \frac{w_h}{w_l} \frac{\mathbb{E}\left[Z_h | w_h z_h > w_l z_l\right]}{\mathbb{E}\left[Z_l | w_l z_l \geq w_h z_h\right]} = \frac{w_h}{w_l} \frac{\bar{z}_h}{\bar{z}_l} \tag{8}$$

Through the lens of this framework, and assuming that $\sigma > 1$, economic development processes characterized by skilled-biased technological change ($\uparrow \frac{A_h}{A_l}$) lead to a rise in the efficiency wage of high-skill occupations relative to its low-skill occupations counterpart $\frac{w_h}{w_l}$, which lowers the comparative advantage that workers require to choose high-skill occupations $\frac{z_h}{z_l}$ and drives the fraction of workers in high-skill occupations π_h up.

At this point, one can not give a conclusive answer about the behavior of the relative labor income ratio $\frac{W_h}{W_l}$, since it might be possible that the rise in efficiency wages $\frac{w_h}{w_l}$ lead to either an increase or a decline in relative mean labor productivities $\frac{z_h}{z_l}$, which can potentially augment, offset, or more than offset the initial effect of relative wages.

Simply stated, in order to know how average labor productivity in high- $(\bar{z}_h)$ and low-skill $(\bar{z}_l)$ occupations react to an increase in the share of individuals choosing high-skill occupations, one needs to know the marginal worker's absolute productivity levels in high- and low-skill occupations and how they are related to the population means. This can not be fully described without further knowledge on the properties of the joint skill distribution in the population, denoted in the model by $g(z_h, z_l)$.

4.3 Absolute and Relative Occupational Labor Productivity and The Properties of The Joint Skill Distribution.

Having described the forces in the model that lead to an increase in the fraction of workers in high-skill occupations as a response to skilled-biased technical change, or higher relative skilled labor efficiency, I proceed to study in further detail the model mechanisms that allow me to interpret the decline in the occupational skill-premium I observe in the data.

To illustrate the role that the joint distribution of skills plays in occupational choice and in determining the levels of occupational labor productivity, consider the case were employment in high-skill occupations is low. In such scenario, the relative wage in high-skill occupations is low enough that only workers with a relatively high comparative advantage in these jobs choose them. This might be either because they have high absolute advantage in High-Skill occupations and low absolute advantage in low-skill occupations, or because they have low absolute advantage in high- and low-skill occupations but they are amongst the individuals with lowest absolute advantage in low-skill occupations in the population, or because they are endowed with high absolute advantage in high- and low-skill occupations but they are amongst the individuals with highest absolute advantage in high-skill occupations in the population [15].

In the first case there is positive selection due to unobservable ability both in high- and in low-skill occupations. In this situation, the joint skill distribution is such that, for workers in both occupational groups, their absolute skill level is higher than the corresponding population average. As a consequence, as the fraction of workers in high-skill occupations rises and the required comparative advantage to select them declines, workers of lower absolute advantage in high-skill jobs now enter them, which drives down average labor productivity in these occupations. For low-skill occupations instead, it is the workers with lower absolute advantage who leave this group, increasing average labor productivity.

The second case features positive selection in high-skill occupations and negative selection in low-skill occupations. In this situation, the individuals who move from low-skill to high-skill occupations have an absolute advantage that is smaller than the average for those working in high-skill jobs and are the highest ability individuals in low-skill jobs before switching occupations. Thus, average labor productivity falls both in High- and low-skill occupations.

The third case is when there is negative selection in high-skill occupations and positive selection in low-skill occupations. In this context, an increase in the fraction of workers in high-skill occupations leads to a rise in average labor productivity both in high- and low-skill occupations. The former is straightforward, as the few workers in high-skill occupations before the increase in wages were those with lowest absolute advantage in the population. The fact that the average skill of workers in low-skill occupations rises is as a result of the lowest ability individuals in low-skill jobs switching occupations [16].

It should be clear at this point that identifying the joint distribution of skills in the population is fundamental in order to understand what are the driving forces behind the decline in the ratio of average occupational labor income as countries develop.

In order to identify the underlying properties of the joint distribution of skills in the

[15] Heckman and Honoré (1990) explore these three alternatives and characterize the properties of the Roy model and its identification for the case in which the distribution of skill differences are log-concave.

[16] Heckman and Honoré rule out the possibility for negative selection in both occupational groups since it requires the covariance of the joint skill distribution to be larger than the variances for both marginals, which leads to a variance-covariance matrix that is not positive semi-definite.

population, I proceed as follows. For tractability purposes, I restrict my attention to parametric distributions. This simplifies the identification problem to finding a family of distributions that fit the data reasonably well in the first place, to in turn proceed with the corresponding parameter estimation.

I discipline my distributional choice by looking at the empirical labor productivity distributions conditional on workers occupational choice ($M(z_h)$ and $M(z_l)$). To do so, I use micro-data on labor income, hours worked, workers' main occupation and a set of demographic characteristics from IPUMS International for the United States in 2010 to compute hourly labor income by occupation at the individual level [17]. Separating the component of workers' labor income that corresponds to their productivity or human capital from the wage per efficiency unit of labor requires an identifying assumption [18]. I choose to normalize the labor productivity of white male workers with no experience and no formal education in High and low-skill occupations to unity. As a result, the average labor income for the individuals in these groups represent the wage per efficiency unit of labor in High and low-skill Occupations, respectively. Dividing the labor income of workers in High and low-skill occupations by their corresponding efficiency wages gives me a measure of the productivity of workers in these two occupations. The Kernel density estimates of the labor productivity distributions by broad occupational group are presented in Figure 4 below.

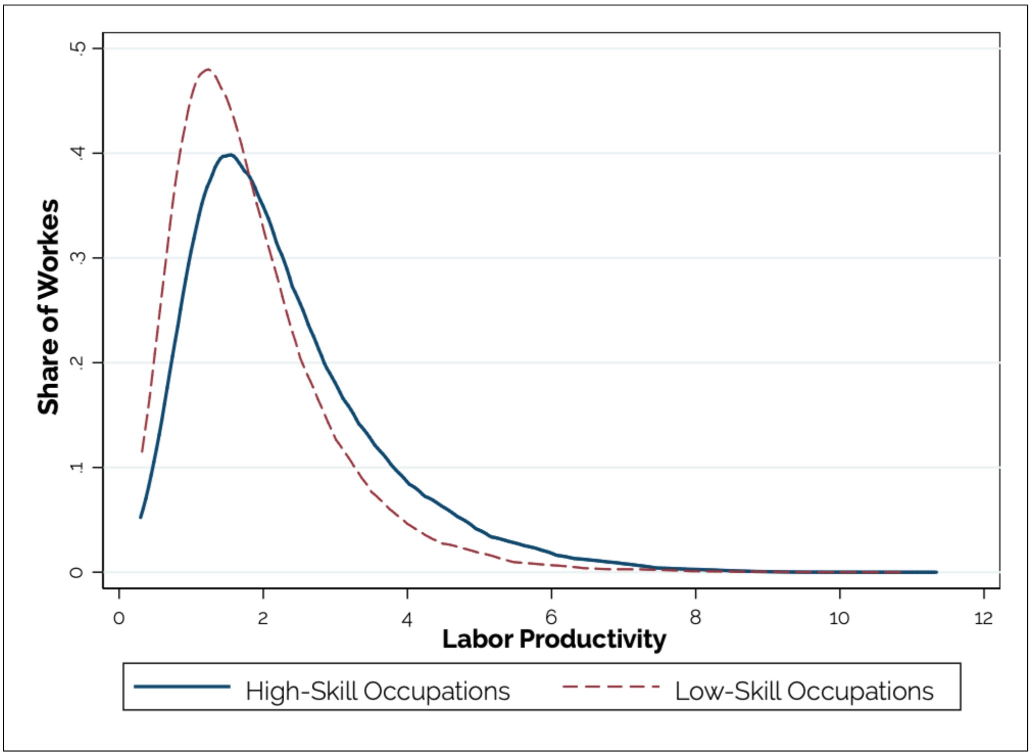
The shape of the empirical densities suggest that the Log-Normal and Frechet cases are good candidates [19]. To decide between these two parametric families, I impose the additional restriction that, for the distribution chosen, the model should be able to reproduce a decline in relative average labor income $\frac{W_h}{W_l}$, like I document in Section 2 above. Two comments are in order. First, notice that this restriction does not entail quantitatively targeting any data moment. It only asks the model to qualitatively reproduce the behavior of relative occupational labor income across development. Second, by imposing this restriction I am not ruling out any of the types of selection discussed above. For example, under bivariate Log-Normally distributed skills one can potentially obtain the decline in relative labor income in the three cases discussed above. The selection type that the model features ultimately depend on the estimated values of the variances

[17] IPUMS International microdata for the US is based on Census data. I use IPUMS International over IPUMS USA because in the former workers occupations are presented using the ISCO classification, as in the ILO databse. Data handling and issues are explained in further detail in Appendix F.

[18] This is standard in the literature. See, for example, Buera et al. (2018).

[19] A Kolmogorov-Smirnov test does not reject the hypotheses that the occupational labor productivity distributions are both distributed Log-Normally or both distributed Frechet at the one percent level. On the other hand, other long tailed distributions like the Beta and Pareto cases are rejected at the ten percent level.

Figure 4: Kernel Density Estimation of the Occupational Labor Productivity Distributions



of the marginal distributions and the covariance between them [20]. However, as I show in Appendix E, in the Log-Normal case one can not obtain a labor income ratio that monotonically declines with development.

As it will become more clear in Section 4.5 below, a monotonically declining ratio of labor income between high- and low-skill occupations can be obtained in the Frechet case under correlated marginal distributions. Before exploring the role of dependence, I characterize the properties of the main model objects under independently distributed Frechet skills below, which are qualitatively preserved once dependence is considered.

4.4 The Roy Model Under Independent Frechet Skill Distributions.

I here analyze some basic properties of the model under the assumption that occupational abilities are drawn from independent Frechet distributions. In order to obtain analytical results I present the common shape parameter case. The properties presented below are preserved under shape parameters that differ across marginals.

[20]Case 1: σ_h ¡ σ_{hl} and σ_l ¡ σ_{hl}; more likely under σ_{hl} 0. Case 2: σ_h ¡ σ_{hl} and σ_l σhl; σhl ¡ 0. Case 3: σh σhl and σl ¡ σhl; σhl ¡ 0

Proposition 1. *Assume skills are drawn from two independent $Frechet(S_j, \theta) = e^{-S_j z_j^{-\theta}}$ distributions, where S_j is an occupation-specific scale parameter governing the overall level of the productivity draws, and θ is a shape parameter that controls the degree of variation in the distribution.*

1. *The share of workers in high-skill occupations is given by*

$$\pi_h = Prob\left(w_h Z_h > w_l Z_l\right) = \frac{S_h}{S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}},$$

2. *The skill distribution conditional on workers selecting high-skill occupations is given by*

$$M_h(z) = e^{-\left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) z^{-\theta}}.$$

3. *The skill distribution conditional on workers selecting low-skill occupations is*

$$M_l(z) = e^{-\left(S_l + S_h \left(\frac{w_h}{w_l}\right)^{\theta}\right) z^{-\theta}}.$$

4. *Average labor productivity in high- and low-skill occupations are given by*

$$\bar{z}_h = \mathsf{E}\left[Z_h | w_h z_h > w_l z_l\right] = \left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right),$$

and

$$\bar{z}_l = \mathsf{E}\left[Z_l | w_l z_l \geq w_h z_h\right] = \left(S_l + S_h \left(\frac{w_h}{w_l}\right)^{\theta}\right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right),$$

where $\Gamma\left(1 - \frac{1}{\theta}\right)$ is the Gamma function evaluated at $\left(1 - \frac{1}{\theta}\right)$.

5. *The ratio of occupational average labor incomes is*

$$\frac{\bar{W}_h}{\bar{W}_l} = \frac{w_h}{w_l} \frac{\mathsf{E}\left[Z_h | w_h z_h > w_l z_l\right]}{\mathsf{E}\left[Z_l | w_l z_l \geq w_h z_h\right]} = \frac{w_h}{w_l} \frac{\left(S_h + S_l \frac{w_h}{w_l}^{-\theta}\right)^{\frac{1}{\theta}}}{\left(S_h \frac{w_h}{w_l}^{\theta} + S_l\right)^{\frac{1}{\theta}}} = 1.$$

Proof. See Appendix E. □

Item 1 fully characterizes the fraction of workers in high-skill occupations. An increase in the scale parameter of the high-skill innate ability distribution (S_h) rises the share of workers in high-skill occupations. Intuitively, fixing S_l and $(\frac{w_h}{w_l})$, the higher S_h is, the higher worker's comparative advantage in high-skill occupations is, and the more likely they are to select them. The opposite holds for the scale parameter of the low-skill innate ability distribution, S_l. A rise in the skill-premium $\frac{w_h}{w_l}$ leads to an increase in the fraction of workers in high-skill occupations, since now it is more likely that workers with a lower comparative advantage in high-skill occupations will choose to work in them.

Items 2 and 3 show how the selection of workers into occupations endogenously affects the scale parameter of the conditional skill distributions, given by $S_h \quad S_l \; \frac{w_h}{w_l}^{\theta}$ and $S_l \quad S_h \; \frac{w_h}{w_l}^{\theta}$, in the high- and low-skill case, respectively. Focus on the high-skill case for a moment. An increase in the marginal effective wage ratio $(\frac{w_h}{w_l})$ lowers the scale parameter of the high-skill ability distribution. Thus, after an increase in relative wages we are more likely to observe workers with lower ability levels in high-skill occupations. This is due to the fact that the smaller required comparative advantage to work in high-skill occupations after a rise in relative wages is associated with a smaller absolute advantage in these type of jobs, which lowers their average ability level. This effect moves the conditional skill distribution for workers in high-skill occupations up and to the left. The opposite holds for low-skill occupations.

The intuition above is the key to understand the results in item 4, which show that average labor productivity in high-skill (low-skill) occupations is decreasing (increasing) in the efficiency wage ratio.

Item 5 in Proposition 4.5 shows that in the case of independent Frechet distributions with common shape parameter the ratio of average occupational labor income is constant and independent of the wage ratio $(\frac{w_h}{w_l})$. This is due to the fact that the increase in average labor income in high-skill occupations after a rise in the skill-premium is exactly compensated by the decline in average labor productivity in high-skill occupations and the increase in average labor productivity in low-skill occupations [21]. In the following section I show how an increase in the wage premium can lead to a decline in the ratio of average occupational labor incomes if the independence assumption is relaxed.

4.5 The Roy Model Under Joint Frechet Skill Distributions.

In this section I show how the observed increase in the fraction of workers in high-skill occupations and decline in the high-to-low-skill average labor income ratios as countries

[21] This can be shown by computing the wage premium elasticity of the ratio of average occupational labor income.

develop can be rationalized through the lens of the stylized model described above. In particular, I focus on the effects of relaxing the assumption that labor productivity realizations are drawn from independent skill distributions.

Specifically, I assume that the marginal occupational ability distributions are still Frechet, as in Subsection 4.4, but now their joint distribution is given by $G_{\phi}(z_l, z_h) = C_{\phi}\left(G_h(z_h), G_l(z_l)\right)$, where $C_{\phi}\left(G_h(z_h), G_l(z_l)\right) = C_{\phi}(u_h, u_l) = -\frac{1}{\phi} \ln\left(1 + \frac{(e^{-\phi u_h} - 1)(e^{-\phi u_l} - 1)}{e^{-\phi} - 1}\right)$ is a bi-variate Archimedean Copula.[22] with parameter ϕ[23][24][25]

Under these assumptions, average labor income in high- and low-skill occupations are

$$\bar{z}_h = \int_0^{\infty} z_h \left(\int_0^{z_h \frac{w_h}{w_l}} G_{\phi}(z_l, z_h) dz_l \right) dz_h,$$

$$\bar{z}_l = \int_0^{\infty} z_l \left(\int_0^{z_l \frac{w_l}{w_h}} G_{\phi}(z_l, z_h) dz_h \right) dz_l.$$

where $G_{\phi}\left(G_h(z_h), G_l(z_l)\right)$ is joint cumulative distribution function of (Z_h, Z_l), modeled via a Copula with parameter ϕ. Even though Archimedean Copulas admit explicit formulas, in general these expectations can not be computed in closed form.

To make clear what the consequences of relaxing the Independence assumption are, I simulate the four main model objects under interest: expected labor productivity in high- ($\mathbb{E}\left[Z_h | w_h z_h > w_l z_l\right]$) and low-skill ($\mathbb{E}\left[Z_l | w_l z_l \geq w_h z_h\right]$) occupations, the fraction of workers in high-skill occupations (π_h), and the ratio of average labor incomes in high-skill with respect to low-skill occupations $\left(\frac{W_h}{W_l}\right)$. I experiment with different dependence parameter values and compare the results with the independence case. The results are presented in Figure 5 below

The main message from Figure 5 is that, once we allow for dependence, the ratio of labor average labor incomes in high- and low-skill occupations is no longer invariant to changes

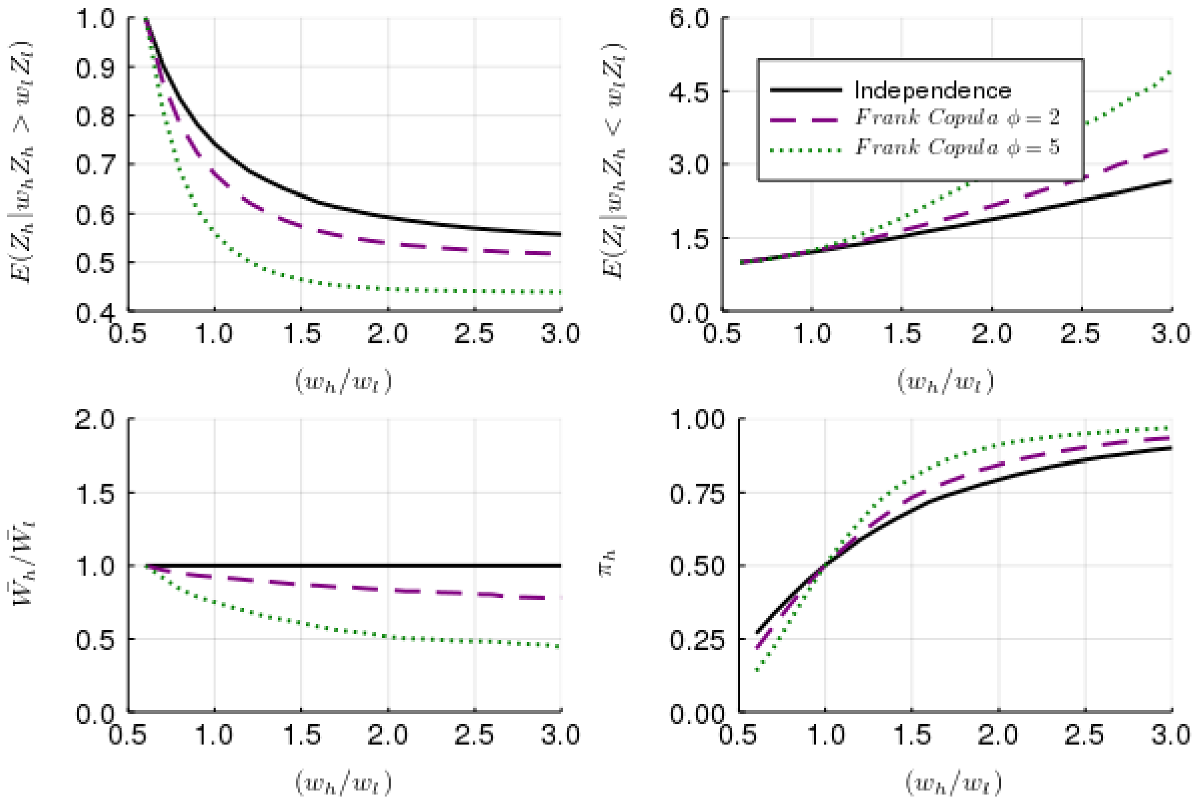
[22] See Nelsen (2006).

[23] Informally, Copulas are functions that join or *couple* multivariate distribution functions to their one-dimensional marginal distribution functions. More formally, according to Sklar's Theorem (1959), for every d-dimensional joint distribution function F, with marginals $F_1, \ldots, F_d$, there exists a Copula C such that $F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))$ $\forall x_i \in [-\infty, \infty]$ and $i = 1, ..., d$. Moreover, if the marginals F_i are continuous, the copula $C(\cdot)$ is unique.

[24] Copulas have the advantage that they contain all the information on the dependence structure between variables, whereas the marginal CDFs contain all the information about the marginal distributions. Thus, the properties of the marginal distributions presented in Section 4.4 are preserved.

[25] Archimedean copulas have the advantages that they admit explicit formulas and that they allow to model dependence in arbitrarily high dimensions with only one parameter. In what follows I use ϕ to denote the parameter that governs the strength of dependence.

Figure 5: Simulated Model Objects Under Different Degrees of Dependence



in the skill-premium [26]. The higher the dependence is, the higher the observed decline in average labor productivity in high-skill occupations, the higher the increase we see in average labor income in low-skill occupations, and the higher the corresponding decline in the average high-to-Low labor income ratio.

The intuition is as follows. Under positive correlation between ability draws, the observed comparative advantage for a given worker is now smaller than what it would be in the independent case. Thus, for a comparable increase in the skill premium more workers are likely to switch from low- to high-skill occupations. In the Frechet case descried in Subsection , since comparative advantage is positively correlated with absolute advantage, workers in high-skill occupations are now expected to have a smaller ability, which leads to a faster decline in average labor productivity in high-skill occupations. The opposite holds for low-skill occupations. As a consequence, the ratio of average labor productivities falls at a higher rate than in the independence case, more than compensating for the increase in the skill-premium, which leads to a decline in the average labor income ratio.

[26] I experimented with negative dependence and it has the opposite results as desired. Similar results can be obtained through lower-tail dependence, using a Clayton Copula. On the other hand, upper-tail dependence modeled trough a Gumbel Copula leads to an increasing average labor income ratio.

4.6 Endogenous Human Capital Accumulation.

I now proceed to relax the assumption that skills are exogenously determined. In particular, I assume that workers can make a discrete and costly decision to build up skills through education. For the sake of tractability, I let individuals choose between no schooling and two different levels of education only ($e \in \{s, u\}$), which can be thought of completing Secondary (s) and University (u), respectively. To fix ideas, denote by $\beta_j^e > 0$ the logarithm of the return to completing educational level e in occupation j and c_j^e the corresponding educational fixed cost. For simplicity, assume that the fixed cost of education is common across occupations and only differs across schooling levels ($c_h^s = c_l^s = c^s$ and $c_h^u = c_l^u = c^u$). Additionally, I suppose that the returns to acquiring the lowest educational level are common across occupations ($\beta_h^s = \beta_l^s = \beta^s$), while the yield of completing the highest educational level is higher in high-skill than in low-skill occupations ($\beta_h^u > \beta_l^u$).

Proposition 2 (Educational Sorting Conditional on Occupational Choice). *Suppose that the fixed cost of attaining the highest level of education c^u is high enough.*

1. *Conditional on working in high-skill occupations, workers' educational decision is characterized by thresholds $\underline{z}_h = \frac{c^s}{w_h(e^{\beta^s}-1)}$ and $\bar{z}_h = \frac{c^u - c^s}{w_h(e^{\beta_h^u}-e^{\beta^s})}$ such that individuals with $z < \underline{z}_h$ acquire no education, workers $\underline{z}_h \leq z_h < \bar{z}_h$ complete the first level of education only, while individuals with $z_h \geq \bar{z}_h$ attain the highest level of education possible.*

2. *Conditional on working in low-skill occupations, workers' educational choice is defined by thresholds $\underline{z}_l = \frac{c^s}{w_l(e^{\beta^s}-1)}$ and $\bar{z}_l = \frac{c^u - c^s}{w_l(e^{\beta_l^u}-e^{\beta^s})}$ such that individuals with $z_l < \underline{z}_l$ opt out of education, and workers with $\underline{z}_l \leq z_l < \bar{z}_l$ and $z_l \geq \bar{z}_l$ complete the first and second occupational levels, respectively.*

Proof. See Appendix E. □

Having fully characterized workers educational decisions conditional on their occupational choices, I proceed to describe the rules that pin down occupational choice.

Proposition 3 (Occupational Choice Under Endogenous Education). *Denote by $\tilde{z}_l = \bar{z}_h \frac{w_h}{w_l} = \frac{c^u - c^s}{w_l(e^{\beta_h^u}-e^{\beta^s})}$. Under endogenous human capital accumulation through schooling, the occupational selection decisions are as follows:*

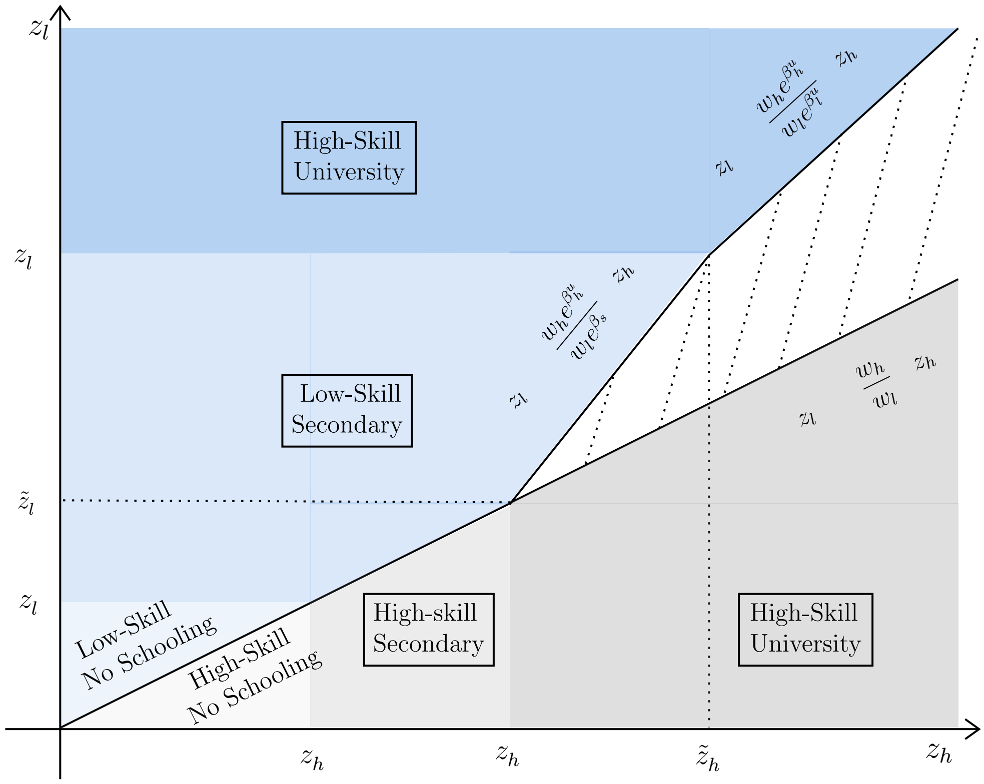
1. *If $z_l \leq \tilde{z}_l$ workers choose high-skill occupations if $z_h > \frac{w_l}{w_h} z_l$.*

2. *If $z_l \in (\tilde{z}_l, \bar{z}_l)$ workers choose high-skill occupations if $z_h > \frac{w_l}{w_h} \frac{z_l e^{\beta^s}}{e^{\beta_h^u}}$.*

3. If z_l ¡ z_l *workers choose high-skill occupations if* z_h ¡ $\frac{w_l}{w_h}$ $\frac{z_l e^{\beta_l^u}}{e^{\beta_h^u}}$.

Proof. See Appendix E. □

Figure 6: Occupational Sorting Under Endogenous Schooling



The intuition for Proposition 3 is summarized in Figure 6 above, which provides a full description of the occupational and educational choices of individuals, for given efficiency wages w_h and w_l.

Item 1 encompasses several possible cases. In all of them, the occupational decision is based on workers' *raw* comparative advantage. This is because workers either find it optimal not to get education (z_h z_h and z_l ¤ z_l), or to get Secondary education in both occupations (z_h ¤ z_h ¤ z_h and z_l z_l ¤ $\tilde{z}_l$), which under common returns does not change their comparative advantage. Additionally, Item 1 includes four more cases where workers educational choice modifies their raw comparative advantage (z_h z_h ¤ z_h and z_l ¤ z_l , z_h ¤ z_h and z_l z_l ¤ $\tilde{z}_l$, z_h ¤ z_h and z_l ¡ $\tilde{z}_l$, z_h z_h ¤ z_h and z_l ¡ $\tilde{z}_l$). However, in all of these cases individuals' comparative advantage before education is already high enough to choose either high- or low-skill occupations, and schooling only augments that comparative advantage in their occupation of choice under the *raw* comparative advantage rule. Hence, comparing workers comparative advantage with relative wages provides a sufficient occupational choice rule for all the cases included in Item 1.

In Item 2 of Proposition 3 the occupational choice decision is modified to take into account that in that region workers go to University if they choose high-skill occupations and, at most, they complete Secondary school if they select low-skill occupations. Thus, since the returns to University are higher than those to going to Secondary school, in this region the schooling decision augments workers' comparative advantage more than proportionately in high-skill occupations. As a consequence, individuals who would not have chosen high-skill occupations according to their *raw* comparative advantage might choose to do so now. This includes the set of individuals between the solid and the dashed line in this region.

Finally, the intuition of Item 3 in Proposition 3 is very close to the one explained above, except for the fact that in this region, if workers were to choose low-skill occupations they would also acquire University education. Since the returns to going to University are higher in high-skill occupations, the comparative advantage of workers grows proportionately more in high-skill occupations, and individuals with who would not have chosen high-skill occupations according to their *raw* comparative advantage might choose to do so now. Again, these workers are the ones whose comparative advantage after education is between the solid and the dashed line in this region.

Proposition 4 (**Endogenous Variables Under Education**). *Denote by* $\tilde{z}_h \equiv z_l \frac{w_l}{w_h} \frac{e^{\beta^s}}{e^{\beta^u_h}} + \frac{c^u - c^s}{w_h \left(e^{\beta^u_l} - e^{\beta^s}\right)} \frac{e^{\beta^s}}{e^{\beta^u_h}}$.

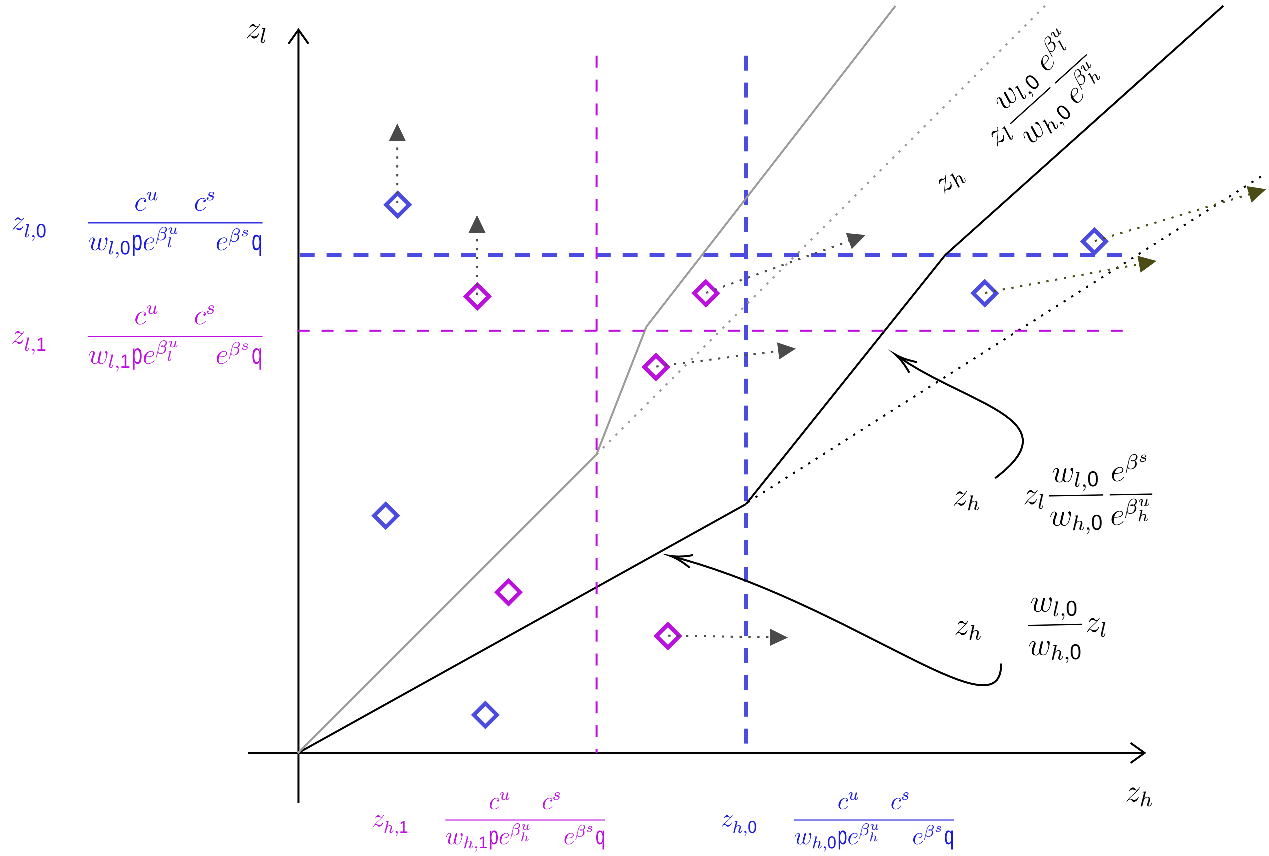
1. *The share of workers in high-skill occupations is given by:*

$$\pi_h = \int_0^{z_h} \int_0^{\frac{w_h}{w_l} z_h} g_\phi(z_h, z_l)\, dz_h dz_l + \int_{z_h}^{z_h} \int_0^{\frac{w_h}{w_l} \frac{e^{\beta^u_h}}{e^{\beta^s}} z_h} g_\phi(z_h, z_l)\, dz_h dz_l + \int_{z_h}^{\infty} \int_0^{\frac{w_h}{w_l} \frac{e^{\beta^u_h}}{e^{\beta^u_l}} z_h} g_\phi(z_h, z_l)\, dz_h dz_l \tag{9}$$

Proof. See Appendix E. □

Proposition 4 is useful to understand how the share of workers in high-skill occupations respond the changes in absolute and relative wages under endogenous education. Assuming that the development process is characterized by skilled-biased changes in productivity that raise both absolute (w_h and w_l) and relative efficiency ($\frac{w_h}{w_l}$) wages, we can use this equation to understand what forces drive the increase in the share of skilled labor as countries develop. These are mainly two channels. The first mechanism is standard and is captured by an increase in the upper limits of integration in the three inner integrals on the right hand side of Equation 9: a higher high-to-low skill efficiency ratio attracts lowers the comparative advantage required by workers to choose high-skill occupations, which attracts more individuals to them.

Figure 7: Sorting Into Occupations and Education After an Increase in Relative Wages



The second mechanism is the educational channel, which is particularly relevant when increased access to education modifies workers' comparative advantage. This might happen because the relative increase in efficiency wages make it proportionately more profitable to go to university in high-skill occupations, as captured by an increase in the distance of the integration limits of the outer integral (Ò z_h z_h) on the second term on the right hand side of the Equation 9, or because more workers now acquire University education in whatever occupation they choose, but the higher returns to University in high-skill occupations lead to an increase in their after education comparative advantage. The latter is captured a reduction in the lower limit of integration on the outer integral in the third term of the right hand side of Equation 9 (Ó z_h).

5 Parameter Estimation.

In this section I describe the strategy I follow to discipline the model parameters. I study a quantitative version of the model where the joint skill distribution is fully flexible, allowing for different scale and shape parameters in the marginal distributions and for dependence between marginals. Additionally, I let workers accumulate human capital through education, as in Subsection 4.6 above.

The goal is to obtain values for thirteen model parameters. The first three objects are technology parameters: the elasticity of substitution between occupational labor services σ, and the technological efficiency parameters in high- (A_h) and low-skill occupations (A_l). To calibrate the elasticity of substitution between labor types I follow Caselli and Coleman and set σ to 1.42. The relative technological efficiency parameters $\frac{A_h}{A_l}$ are calibrated to match the relative wage per efficiency unit of labor in high-skill occupations with respect to low-skill occupations for the US in 2015, which is obtained as explained in Section 4.3, and the fraction of output per hour worked that corresponds to labor for the US in 2015 [27].

The remaining ten parameters are related to the joint distribution of skills: the shape and scale parameters of the high-skill innate ability marginal distribution θ_h and S_h, the shape and scale parameters of the low-skill innate ability marginal distribution θ_l and S_l, the parameter that governs the dependence between marginals in the Frank Copula, ϕ, the parameters that represent the fixed cost of acquiring University c^u and Secondary education c^s, and the log-returns to Secondary β_l^u and University β_h^u, β_l^u education.

Leaving aside the log-returns to education for now, the skill distribution and educational cost parameters do not have direct data counterparts and depend on how workers' select themselves into occupations and education, so I estimate them via a Simulated Method of Moments. This procedure requires at least seven data moments that I match to seven model-simulated counterparts, and themselves are functions of the deep models parameters under interest. The results of the estimation algorithm is a set of parameters that minimize the weighted sum of the squared distances between the data targets and their model simulated analogs.

The data moments under interest come from the occupational labor productivity dis-

[27] Assuming an aggregate technology of the form $y = k^{\alpha}\left((A_l L_l)^{\frac{\sigma-1}{\sigma}} + (A_h L_h)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma(1-\alpha)}{\sigma-1}}$ this requires information on GDP, number of people employed, average hours worked per employee and the capital stock for the US in 2010, which I obtain from Penn World Table 9.2. To calibrate the capital share parameter I follow Caselli and Coleman and set $\alpha = 1/3$.

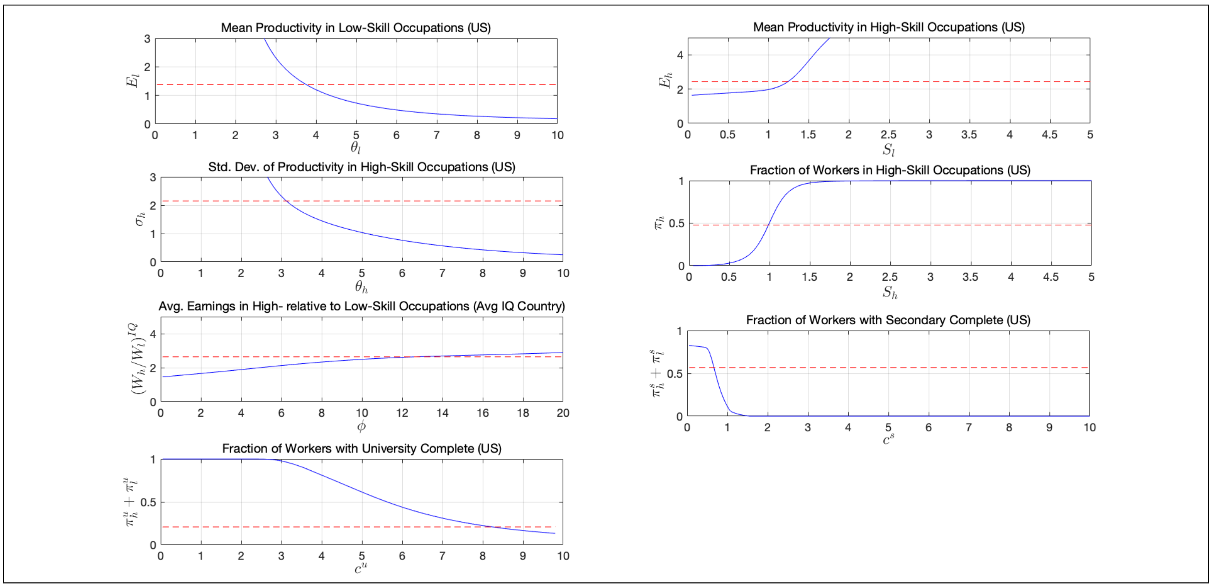
tributions I infer from the data, following the procedure described in Section 4.3. The data moment description and their corresponding values are presented in Table 3 below.

Table 3: Description of Data Moments Targeted

Target	Description	Target Value
$\mathbb{E}[Z_l \mid w_lZ_l \geq w_hZ_h]$	Mean of the labor productivity distribution in low-skill occupations (US)	1.85
$\sigma(Z_h \mid w_hZ_h > w_lZ_l)$	Standard deviation of the labor productivity distribution in high-skill occupations (US)	2.15
$\mathbb{E}[Z_h \mid w_hZ_h > w_lZ_l]$	Mean of the labor productivity distribution in high-skill occupations (US)	2.44
$(W_h/W_l)^{IQ}$	Avg. Earnings in high- relative to low-skill occupations for the (Avg. IQ country)	2.64
π_h	Fraction of workers in high-skill occupations (US)	0.478
$\pi_h^s + \pi_l^s$	Fraction of workers with Secondary education (US)	0.568
$\pi_h^u + \pi_l^u$	Fraction of workers with University education (US)	0.207

Figure 8, on the other hand, shows the model-simulated analogs for these moments and how they react to changes in the underlying parameters. It matches a particular parameter with the model-simulated moment that is more sensitive to it, and plots in the blue lines how the simulated moment varies with changes in the parameter value, fixing the values for the other parameters under interest, and in the red line the targeted value.

Figure 8: Parameter Identification



Note: Model-simulated moments are calculated by running, for a given set of parameter values, 1,000 model simulations. The blue line shows how the value of the simulated moment changes as I change the value of one parameter at a time while leaving the remaining parameters fixed at the values that minimize the weighted sum of squared errors in the Simulated Method of Moments. In each figure, the X axis represents a model parameter and the Y axis a moment.

Finally, to estimate the log returns to education I use, once again, U.S. Census data on worker's wage and salary income, main occupation, hours worked, labor force attachment and demographic characteristics from IPUMS International. To be precise, the returns are estimated by regressing log hourly wages for full-time, full year workers by occupational group on three education dummies (less than Secondary, Secondary complete and University dropouts, University graduates and above), a quartic in experience, interactions of the education dummies and the experience quartic, and two race categories (white, other).

Table 4 below presents the values for the calibrated and estimated model parameters. It

Table 4: Calibrated/Estimated Values for the Model Parameters

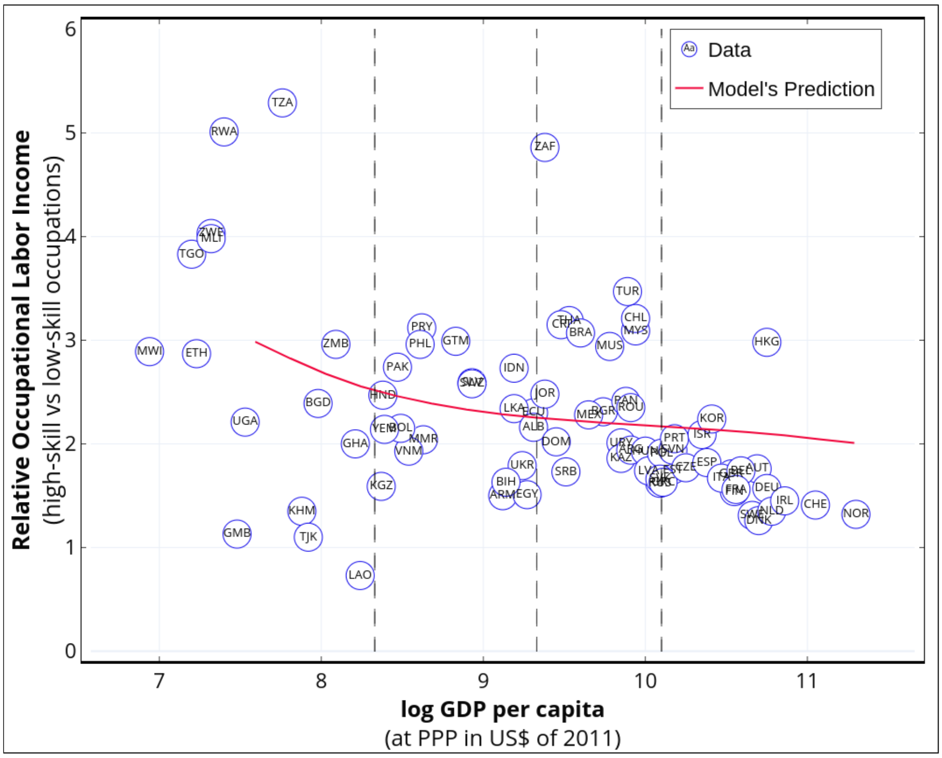
Parameter	Parameter Description	Parameter Value
	Calibrated Parameters	
σ	Elasticity of substitution between labor types	1.50
A_h	Technological Efficiency of high-skill occupations	3.38
A_l	Technological Efficiency of low-skill occupations	0.87
	Parameters Estimated Through SMM	
θ_h	Shape parameter of high-skill labor productivity distribution	2.83
		(0.5084)
S_h	Scale parameter of high-skill labor productivity distribution	0.99
		(0.0010)
θ_l	Shape parameter of low-skill labor productivity distribution	3.72
		(0.0028)
S_l	Scale parameter of low-skill labor productivity distribution	1.31
		(0.0081)
ϕ	Dependence parameter in the Frank Copula	12.11
		(0.5220)
c^s	Fixed cost of Secondary education	0.67
		(0.0041)
c^u	Fixed cost of University education	8.19
		(0.1483)
	Log-returns to education	
β^s	Average log return to Secondary school	0.12
		(0.0010)
β^u_h	Log return to University in high-skill occupations	0.59
		(0.0027)
β^u_l	Log return to University in low-skill occupations	0.44
		(0.0011)

6 Quantitative Analysis.

I begin assessing the model's prediction for relative labor income in high-skill occupations with respect to low-skill occupations. To that end, for each level of GDP per capita I target the fraction of workers in high-skill occupations (π_h) and the share of individuals that acquire Secondary and University education ($\pi^s \quad \pi^s_h \quad \pi^s_l$ and $\pi^u \quad \pi^u_h \quad \pi^u_l$). This three targets are enough to pin down relative efficiency wages ($\frac{w_h}{w_l}$) and the fixed costs of education (c^s and c^u), which suffice to fully characterize the occupational and educational selection decisions of workers' at each level of GDP per capita. Average labor productivity in high- and low-skill occupations can be pinned down as a result.

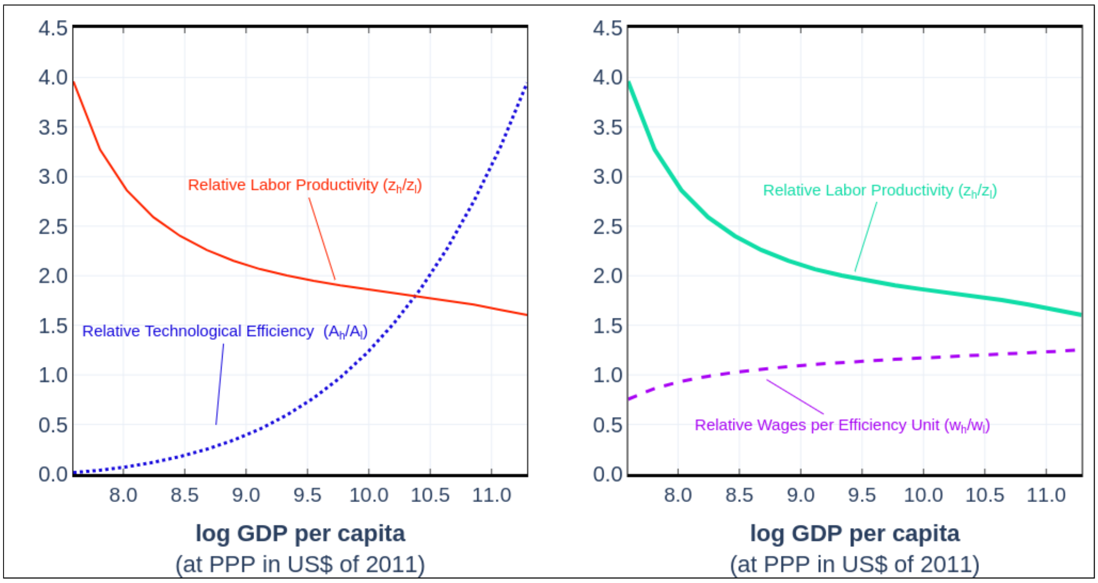
One more comment is in order. The procedure described above requires associating different levels of GDP per capita with their corresponding share of workers in high-skill occupations and the share of workers that acquire each level of education. To obtain a smooth relationship, I run, for the countries in my sample, an OLS regression on each of these two variables on a quadratic for GDP per capita. The results are summarized in Figure 9 below.

Figure 9: Relative Hourly Labor Income in the Model vs the Data



As below above, the model reproduces the empirical pattern of the relative labor income ratio in high-skill versus low-skill occupations. In particular, it mimics the negative relationship documented in Section 1. In absolute terms, the model's relative labor income prediction is below the actual values we observe in the data, particularly for the 4th

Figure 10: Relative Labor Income Decomposition.



development quartile. To be precise the actual average labor income ratios in high-skill versus low-skill occupations are 2.6 in the first development quartile, declining to 2.4, 1.8, and 1.6 in the second, third, and fourth quartiles, respectively. In the case of the model's prediction, these values are 2.8,2.3,2.1, and 2.0. Overall, the model does fairly well in terms of reproducing the empirical behavior of labor income. I take this as evidence that workers' selection into jobs due to observable and unobservable skills is an important feature to understand the relationship between skills and technological progress across development.

The left panel of 10 displays how average abilities in high-skill relative to low-skill occupations and the relative technological efficiency of high-skill labor evolve with development, according to the model. As shown in figure, the average ability of workers in high-skill occupations with respect to those in low-skill jobs falls as countries grow richer. This is particularly more pronounced in the first development quartile. At the same time, relative technological efficiency in high-skill occupations with respect to its low-skill occupations counterpart grows monotonically with GDP per-capita, driving the development process.

The right panel of Figure 10 decomposes the model's relative average labor income prediction into a relative average ability $\frac{z_h}{z_l}$ and a relative wages per efficiency unit component $\frac{w_h}{w_l}$. The main message of this figure is that the predicted decline in relative earnings in high- versus low-skill occupations is as a consequence of a decline in the ratio of average ability in high-to-low-skill occupations that more than compensates the rise in relative wages per efficiency unit of labor as countries grow richer.

6.1 Effective Skills versus Technologies.

In here use the quantitative version of the model to decompose the decline in labor income in high-skill versus low-skill occupations across development. In particular, I focus on the observed decline between the average country in the first development quartile and the US. To do so, I fix the skill distribution parameters at the values estimated for the US and target, for each country, the fraction of workers in high-skill occupations and the share of educated individuals. In order to match these targets I adjust three model parameters: the ratio of technological efficiency in high-skill with respect to low-skill occupations ($\frac{A_h}{A_l}$), and the fixed costs that workers need to incur to complete Secondary and University schooling (c^s and c^u).

To calibrate these parameters and pin down the equilibrium corresponding to them in each country, I use the results presented in Proposition 4. To be precise, I use bisection methods to first find the relative efficiency wage ($\frac{w_h}{w_l}$) and ability thresholds (z_h and z_l) that rationalize workers' occupational and higher-level educational decisions. Since the decision to get Secondary education does not affect occupational choices, with these three endogenous variables in hand I can compute the Secondary schooling ability thresholds (z_h and z_l) that reconcile the observed educational attainment levels.

These three endogenous variables are sufficient to fully characterize the occupational and educational choices of workers. Thus, with them one can obtain the effective labor supply in high- and low-skill occupations. To find the values of the calibrated parameters of interest, I back out $\frac{A_h}{A_l}$ from

$$\frac{w_h}{w_l} = \left(\frac{\pi_h \bar{z}_h}{\pi_l \bar{z}_l}\right)^{\frac{1}{\sigma}} \left(\frac{A_h}{A_l}\right)^{\frac{\sigma-1}{\sigma}}.$$

Finally, the absolute values of c^u and c^s are obtained by solving the system of two equations in two unknowns defined by $\frac{z_h}{z_h}$ and $\frac{z_l}{z_l}$ define.

The calibrated parameters and equilibrium model variables needed to perform the labor income decomposition are presented in Table 5 below.

Table 5: Relative Labor Income Decomposition in the Model

	$\frac{w_h}{w_l}$	c^u	c^s	$\frac{A_h}{A_l}$	A_h	A_l	$\frac{z_h}{z_l}$	$\frac{\pi_h}{\pi_l}$	$\frac{L_h}{L_l}$	$\frac{\pi_h z_h}{\pi_l z_l}$
US	1.26	8.19	0.67	3.89	3.38	0.87	1.50	0.90		1.35
Avg IQ Country	0.92	3.74	0.41	0.055	0.018	0.33	2.85	0.09		0.26

The counterfactual values for the labor income ratio are exhibited in Table 6 below.

The left panel of presents the labor income ratio for the US in the first row. The second and third rows show the counterfactual relative labor income for a hypothetical country with the US's relative technological efficiency and the average first development quartile's country raw (l.ii) and effective relative labor supply (l.iii). In the fourth row (l.iv) I let the relative technological efficiency adjust to the level calibrated for the average country in the first development quartile. As a consequence, the fourth row presents the level of relative occupational earnings corresponding to the average country corresponding to the poorest 25 per cent of my sample.

By virtue of the calibration procedure, the model matches the actual decline in earnings in high-skill occupations with respect to low-skill occupations as countries grow from the first development quartile to the level of development observed in the US. In what follows, I decompose the fraction of the decline in labor income explained by the model into a relative technological efficiency in high-skill occupations and a relative quality of labor in high- versus low-skill occupations.

As row (l.ii) shows, in a counterfactual country with the US's technological efficiency and the average first development quartile's country raw supply of skills, the labor income ratio rises to 8.76. Moreover, letting relative labor productivity adjust to the average first development quartile's value increases the relative labor income ratio to even further to 10.86. Thus, the goal is to explain a total change in the relative labor income ratio of 8.22 (10.86-2.64). Since the increase in the relative quality of skilled labor leads to a change in relative labor income of 2.10 (10.86-8.76), it follows that 25% of the change in relative labor income ratio is due to relative skilled labor productivity, while the remaining 75% is due to differences in the relative efficiency of the technology that uses skilled labor in production.

Table 6: Relative Labor Income and Counterfactuals

$\bar{W_h}\text{z}\bar{W_l}$			$\bar{W_h}\text{z}\bar{W_l}$		
(l.i)	US	1.91	(r.i)	US	1.91
(l.ii)	US $\frac{A_h}{A_l}$ and $\frac{z_h}{z_l}$ - IQ $\frac{\pi_h}{\pi_l}$	8.76	(r.ii)	US $\frac{z_h}{z_l}$ and $\frac{\pi_h}{\pi_l}$ - IQ $\frac{A_h}{A_l}$	0.46
(l.iii)	US $\frac{A_h}{A_l}$ - IQ $\frac{\pi_h}{\pi_l}$ and IQ $\frac{z_h}{z_l}$	10.86	(r.iii)	US $\frac{z_h}{z_l}$ - IQ $\frac{A_h}{A_l}$ and IQ $\frac{\pi_h}{\pi_l}$	2.11
(l.iv)	Avg IQ Country	2.64	(r.iv)	Avg IQ Country	2.64

A similar exercise can be performed by adjusting relative technological efficiency first and computing the effect of relative labor productivity residually. In the context of a linear model, these two exercises would give the same answer. It turns out that in this case the two answers differ by a small magnitude. Row two in the right panel of Table

6 shows that fixing the effective supply of skills at US's level but adjusting the relative technological efficiency to the average first development quartile's country would have led to a decline in relative labor income to 0.46. Thus, the goal is to decompose a total change of 2.18 (2.64-0.46) in the relative income ratio. Letting the relative raw supply of high-skilled labor adjust to the level in less-developed countries leads to an increase in the income ratio to 2.11. It follows that the residual increase from 2.11 to 2.64 (0.53) is due to differences in the relative high-skill labor productivity. This is roughly a 25% (0.53/2.11).

6.2 Measured Gaps in Cross-Country Skilled Labor Efficiency.

Table 5 above presents the values of the relative ($\frac{A_h}{A_l}$) and absolute (A_h and A_l) technological efficiency parameters for the US and the average country in the less-developed group. As we can see, the model implies that there exist sizeable gaps in high-skill labor technological efficiency. In fact, the skill-bias in technologies is about 70 times higher in rich than in poor countries, and mostly driven by vast gaps in the absolute efficiency of the technologies that employ high-skill labor. For instance, A_h is 190 times higher in the US than in the average country in the first development quartile, while A_l is only about 2.6 times higher in rich countries.

Based on data on educational attainment and returns to education, Rossi finds that relative efficiency is, at least, 100 times higher in rich than in poor countries. However, the poorest country in his sample is Indonesia, which in my case, belongs to the second development quartile. To contrast my model predictions with those in 5 I redo my exercise for Indonesia. The results are presented in Table 7 below.

Table 7: Relative Labor Income Decomposition for the US and Indonesia

	$\frac{w_h}{w_l}$	c^u	c^s	$\frac{A_h}{A_l}$	A_h	A_l	$\frac{z_h}{z_l}$	$\frac{\pi_h}{\pi_l}$	$\frac{L_h}{L_l}$	$\frac{\pi_h z_h}{\pi_l z_l}$
US	1.26	8.19	0.67	3.89	3.38	0.87	1.50	0.90		1.35
Indonesia	0.97	4.72	0.48	0.09	0.063	0.59	2.63	0.12		0.32

As we can see, measured relative technological efficiency is, roughly speaking, 54 times higher in my case, which is around half of what Rossi finds.

In another recent study, using a method that that relies on disaggregated trade and industry data, Malmberg finds relative skilled labor efficiency to be between 3 and 28 higher rich and poor countries. As in the case of Rossi, the poorest countries in his sample belong to the second development quartile in my dataset. Under his preferred calibration, the relative efficiency of skilled labor is 6.5 times higher in rich countries.

However, one major difference with Malmberg is that in his work he focuses on the manufacturing sector, while I study aggregate gaps in skilled labor efficiency. This can translate into major quantitative differences, as gaps in technological efficiency are the lowest in Manufacturing and Services and the highest in Agriculture [28].

6.3 Discussion: Absolute and Relative Skill Bias in Technology.

In addition to relative technological efficiency ($\frac{A_h}{A_l}$), Table 5 presents the absolute values for the technology parameters in high- and low-skill occupations (A_h and A_l). As we can see, the absolute technological efficiency in both high- and low-skill occupations improves with development, with a much more pronounced in the case of the high-occupations. In fact the absolute efficiency of the high-skill technology is roughly 60 times higher in the US than in the average country in the less-developed group. On the other hand, the absolute efficiency parameter in the low-skill labor technology is only about 2.5 times higher.

These results differ from those found by Caselli and Coleman (CC), who find that the absolute efficiency of unskilled labor falls with development. Notice that the difference with CC is due to my finding that average labor quality of skilled labor is higher in poor countries. In an alternative hypothetical country with the same characteristics as the average nation in the first development quartile of my sample and the same relative quality of skilled labor as in the US, the skill-bias in technology for the poor country would be 0.011, or about five times smaller that in the benchmark case (0.055).

6.4 The Effects of an Educational Expansion

I continue to explore the model's quantitative implications of an educational expansion. Within the model, this counterfactual exercise can be performed by analyzing the effects of a reduction in the fixed costs of education c^s and c^u. The most interesting case from policy perspective is studying if increased access to education can push forward less-developed countries towards higher standards of living. Hence, I analyze the effects of a reduction in both educational costs for the average country in the first development quartile. The engineered reduction in educational costs is such that least-developed countries reach the same level of educational attainment as those observed in the group of wealthiest nations after the thought policy is in place.

To do so, I first calibrate the model to match the employment rate in high-skill occupations and the percentage of workers with Secondary and University education in the first development quartile, as explained in Section 6.1 above. With the model calibrated to match the most salient labor market and aggregate outcomes for the first development

[28] See Rossi and Gollin et al.

quartile, I proceed to perform a reduction in the costs of acquiring Secondary and University education (c^s and c^u) until the educational attainment levels in the average poor country in my sample are the same as those observed in the US. Since to perform this exercise I keep relative and absolute technological efficiency fixed at the pre-expansion levels, finding the new equilibrium requires pinning down the level of wages per efficiency unit of labor (w_h,w_l) that clear the labor market, which together with the educational costs are enough to fully characterize workers' occupational and educational decision. The results of the exercise are summarized in Table 8 below.

Table 8: Effects of an Educational Expansion

Variable	Baseline	Educational Expansion pÓ c^s, c^uq	Abs/% Change
π_h	8.5%	5.2%	-3.3 p.p.
$\pi_h^s \quad \pi_l^s$	15.8%	56.9%	41.1 p.p.
π_h^szπ_h	28.8%	2.4%	-26.4 p.p.
π_l^szπ_l	14.6%	59.9%	45.5 p.p.
$\pi_h^u \quad \pi_l^u$	3.6%	20.7%	17.1 p.p.
π_h^uzπ_h	33.4%	79.6%	46.2 p.p.
π_l^uzπ_l	0.8%	16.5%	15.7 p.p.
w_h	1.18	0.80	-32.2%
w_l	1.28	1.17	-8.6%
$\bar{z}_h$	4.74	7.04	48.5%
$\bar{z}_l$	1.66	1.97	18.7%
$\bar{W}_h$z$\bar{W}_l$	2.63	2.43	-7.6%
Output	1.000	1.15	15.0%
c^u	3.74	1.40	-58.8%
c^s	0.41	0.18	-56.1%
c^uzAvg. Earnings	1.5	0.6	-
c^szAvg. Earnings	0.3	0.1	-

A relatively cheaper access to schooling more than doubles the percentage of workers that complete Secondary education (56.9% vs 15.8%). Roughly speaking, this is entirely driven by increased Secondary school completion by individuals in low-skill occupations, which rises from 15.8% to 56.9%. The fraction of workers with University education rises more than proportionately (20.7% vs 3.6%), but in this case, driven by a rise in completion by individuals in high-skill occupations. Since the educational attainment rules are monotonically increasing in ability, the 46.2 percentage point increase in the fraction of workers with University complete in high-skill occupations is given in part by workers who previously would have only completed Secondary in 26.4 percentage points and by individuals who would not have acquired education otherwise in 19.8 percentage points. Something similar occurs in low-skill jobs, but in this case 15.7 percent of workers who were only finished Secondary school before the policy now go to University, while the increased access to Secondary school comes from individuals that would not have accessed education before the expansion (45.3 percent).

Moving on to the occupational structure, more affordable access to education reduces the fraction of the population who works in high-skill occupations. However, when compared with the effects on the educational structure, this reduction is marginal. The slight fall in the equilibrium share of workers in high-skill occupations (-3.3 p.p.) is driven by general equilibrium effects, which come from the fact that the effective supply of labor increases more than proportionately in high-skill occupations, reducing relative efficiency wages in high-skill versus low-skill occupations.

Continuing with the results, even though the occupational structure is almost unchanged, the educational expansion has a positive effect on output per-worker, which grows 15 percent, as a consequence of increased average labor productivity, mainly in high-skill jobs.

Compared to a model where workers are classified into high- and low-skill according to their educational attainment, the growth effects of the educational expansion are between one-third and one-tenth model, depending on if the educational attainment threshold to be considered high-skill is Primary school complete or University complete. In these types of frameworks, since labor quality is purely determined by educational attainment, as larger fractions of the population access education the effective labor supply increases. In the case of skilled labor, both the raw and the effective quantity of skilled labor increases. On the other hand, the effective supply of unskilled labor depends on if the quality effect coming from higher educational attainment within a group compensates or not the fall in the quantity of workers below any given threshold. For the Primary complete case, the quantity effect dominates, while in the University case, the quality effect offsets the quantity effect. Thus, the growth effects are larger in the latter case.

The results are in line with the effects of a major occupational expansion that happened in Brazil between 1990 and 2010. As Jaume (2019) documents, Brazil implemented several important educational reforms to increase the educational level of the population. These reforms included, among others: an increase in public expenditure on education from 2.0 percent of GDP in 1995 to over 5.0 percent in 2010 and reducing the direct and indirect cost of schooling by creating more schools and universities together with conditional cash transfers, which I consider to be empirical counterpart of a reduction of the fixed cost of education (c_e) in my model.

Interestingly, Jaume finds that the share of workers with secondary education doubled from 20.5 to 40.0 percent, while the share with university grew from 11.3 to 23.6 percent, and the share of workers with only primary education or less halved from 68.1 to 36.4 percent. In addition, the occupational structure of employment remained relatively fixed, with workers of all educational groups increasingly employed in lower wage occupations. For example, there was an increase in employment of only 1.9 percentage points in the one third of occupations with the highest wages in the economy, despite the expansion of 13 percentage points in the share of high educated workers. These results are in line with the predictions of my model.

However, he finds that inequality measures improved in Brazil, mostly driven by the fact that wages for primarily educated workers soared as a consequence of a reduced supply and an increase in demand for these type of workers. My model predictions in terms of inequality are ambiguous and depend on the group under consideration.

7 Conclusion.

I study how relative efficiency and the relative quality of skilled and unskilled labor vary with development. Using harmonized, occupational labor market outcomes for a broad set of countries across the development spectrum, I document that employment in high-skill occupations, or jobs that are relatively more intensive in non-routine cognitive tasks, grows with development. In addition, workers earnings in high-skill occupations falls with respect to those in low-skill occupations as countries grow richer, with elasticities in line with those found by studies based on educational attainment and Mincerian returns to education.

To shed light on these findings and disentangle the mechanisms that determine the relative quality and efficiency of skilled labor, I build a general equilibrium model of occupational choice and human capital accumulation through education. In the model, exogenous skill-biased shifts in productivity attract workers with a lower comparative advantage to occupations that are more skill intensive, in the sense that their abstract task component is relatively higher. The resulting average labor productivity of workers in high- and low-skill occupations depends on how their comparative and absolute advantage is correlated, which depends on the properties of the joint distribution of skills in the population.

I discipline the joint skill distribution and other model features using US labor market data and use the quantitative version of the model to study how the relative efficiency and the relative quality of skilled and unskilled labor vary with development. I find that in poor countries, the relatively small share of workers in high-skill occupations is composed by individuals that are both relatively and absolutely more productive in performing them. In addition, there is positive selection of workers into education, with a higher fraction of those in high-skill occupations achieving higher levels of educational attainment. As a consequence, the relative quality of skilled labor is higher in poor than in rich countries.

When used to decompose the decline in relative labor income of high-skill and low-skill occupations between poor and rich countries, I find that relative quality explains between a quarter and a third of this fall, while relative efficiency explains the remaining 70-75 percent. Additionally, the fact that the relative productivity of workers in high-skill occupations is higher in less develop countries doubles the measured gap in relative efficiency between rich and poor countries.

Appendices

A ILO Data Description, Treatment, and Issues.

The ILO's database is constructed from multiple data sources, including establishment surveys, household surveys, insurance records, and administrative data sources. Data sources vary across countries, and when multiple sources are available, ILO presents all the options available. In such case, I pick data from the source I consider most reliable [29]. I discard data from sources that the ILO flags as unreliable, even if that is the only one available for a country.

The ILO data is harmonized, both at the sectoral and occupational level, allowing for comparability across countries. Occupational data is harmonized based on the International Standard Classification of Occupations (ISCO). Statistics on employment by occupation are presented in ILOSTAT according to both the categories of the latest version of the ISCO available (ISCO-08 and ISCO-88). When both versions are available, I take the latest revision (ISCO-08). I take the earlier version (ISCO-88) when it is the only one available and bridge it into the newer (ISCO-08) using the correspondence table provided by ILO [30].

[29] Each source has it's own advantages and disadvantages, depending on the country under study. For example, establishment data tend to be very accurate, but it has limitations in countries where firms routinely pay wages outside their normal book-keeping in order to avoid taxes. Household surveys cover all employees regardless of where they work, but their reliability depends heavily on the accuracy of the respondent.

[30] Link to ILO's occupations documentation.

B Countries Intensive in Natural Resources

Natural resources rents are measured by the World Bank's World Development Indicators variable Total natural resources rents (% of GDP)(NY.GDP.TOTL.RT.ZS). Total natural resources rents are the sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents, and forest rents.

To decide which countries to exclude I take the 217 countries with data available in the WDI database and rank them according to their average natural resources rents as percentage of GDP for the period 2008-2012. The criterion for exclusion is to discard the countries in the top decile in terms of natural resources rents. I end up excluding 21 countries with natural resources rents above 25.1% of GDP.

The countries excluded are Libya (53.2%), Kuwait (52.4%), Republic of Congo (51.4%), Saudi Arabia (46.3%), Iraq (45.2%), Mauritania (44.1%), Angola (42.3%), Oman (40.6%), Papua New Guinea (39.9%), Liberia (39.8%), South Sudan (38.4%), Gabon (37.4%), Equatorial Guinea (35.3%), Mongolia (34.1%), Turkmenistan (33.1%), Azerbaijan (32.5%), Chad (28.4%), Guinea (26.7%), Burundi (25.9%), and Brunei Darussalam (25.1%).

C Hours Worked Across Development.

In this section I briefly analyze the evolution of hours worked as countries develop. It is of special interest to see how my data compares to that in Bick, Fuchs-Schundeln, and Lagakos Bick et al. (2018), the paper that, up to my knowledge, more carefully studies this issue. Data limitations allow me to compare hours per worker by sector only. Even though the authors do not study the evolution of hours worked by occupation, I show how hours worked differ for my broad occupational categories across development as well.

The Table above shows that although average hours worked follow the same pattern

Average Hours Worked by Sector Across Development

-ILO data vs B,F-S,L-

Income Group	Agriculture		Industry		Services	
	ILO data	*B,F-S,L*	*ILO data*	*B,F-S,L*	*ILO data*	*B,F-S,L*
Low	35.0	36.0	45.7	44.9	46.0	47.7
Middle	38.5	38.3	42.8	42.5	41.8	41.8
High	41.5	39.7	39.6	37.0	37.2	34.7

Note I : B,F-S,L corresponds to the data in Bick, Fuchs-Schundeln, and Lagakos. In their paper Industry is called Manufacturing. The sectors included in it are roughly the same except for Construction, which I include in Industry and it is not clear to me if they consider it to be in Manufacturing or Services.

Note II : To compute GDP per capita and development percentiles I use as GDP measure PWT's *rgdpo*. I take the average GDP per capita for the period 2005-2014. B,F-S,L use as GDP measure PWT's *rgdpe*, and compute terciles for 2005. I follow their procedure to construct this table.

Note III : B,F-S,L focus on 49 core countries while I have data for 88 countries.

in both cases, magnitudes differ. For instance,in Agriculture there is a 6.5 hour increase between the top and the bottom development tercile, while this number falls to 3.9 in B,F-S,L. At the same time, the decline in hours worked in Industry (6.1 vs 7.9) and Services (8.8 vs 13.0) as countries move from the bottom to the top tercile is smaller in my sample as compared to B,F-S,L. The second Table in this section shows that the pattern described in B,F-S, L is robust to using different measures of GDP per capita as a proxy of development and to splitting to sample into quartiles instead of terciles. Interestingly, the decline in average hours worked is more pronounced in low-skill than in high-skill services (-11.1 vs -7.2 if one compares the Top and the Bottom quartiles).

Average Hours Worked by Sector Across Development

-using ILO data and four broad sectors-

Income Quartile	**Countries**	**Agriculture**	**Industry**	**L-S Services**	**H-S Services**
Bottom	22	35.6	46.1	46.8	46.0
2nd	18	36.6	44.0	44.6	42.3
3rd	25	38.4	41.8	40.5	40.6
Top	23	42.2	39.6	35.7	38.8

Note I : To compute income quartiles I consider all countries in PWT 9.0, use as GDP measure PWT's *rgdpo* and take the average GDP per capita for the period 2005-2014. This criterion differs from B,F-S,L, as described in Note II of the Table above.

Average Hours Worked by Occupation Across Development

-using ILO data-

Income Quartile	**high-skill**	**Low-Skill**
Bottom	40.9	46.0
2nd	40.8	47.4
3rd	40.0	43.1
Top	37.6	37.1

D Robustness Checks.

I here explain in further detail the robustness checks I perform to my quantitative empirical analysis in Section 3.

Different Criteria for Constructing Broad Occupational Groups.

I proceed to study if the results presented in Section 3.1 depend on the criterion used to group one-digit ISCO-08 categories into high- and low-skill occupations. I perform a sensitivity analysis under three different grouping criteria.

The first exercise moves the occupational group with lowest median wages at all development levels in the high-skill group, namely Clerical, from the high-skill to the low-skill group. The second exercise, moves one of the categories with highest median wages in low-skill occupations, Service Workers, into the high-skill group. It is worth pointing out that now the original grouping criterion is no longer respected, as median wages for Service Workers are, in this case, lower than those for Skilled Agricultural Workers in the second and third development quartiles. However, I consider that switching one occupational group at a time presents a more thorough and transparent way to asses how my results are affected by different grouping criteria. In my third exercise I include both Service and Skilled Agricultural Workers in the high-skill group. In this case it holds that median wages in for all the occupations in the high-skill group are higher than those in the low-skill group, at all development levels [31]

The results are presented in Table 9 below. Compared to the baseline case presented in Column 1 of Table 9, switching occupational group four from high- to low-skill occupations has the effect of increasing both the constant and the skill-premium elasticities for all quartiles. Estimated coefficients are both individually and jointly statistically significant at the one percent level.

Switching either occupation seven or seven and eight together from the Low- to the high-skill groups have similar effects. The regression constant falls, more in the latter case, while the skill-premium elasticities are in both instances smaller compared to the baseline, but roughly speaking, very similar. In these two counterfactuals the estimated coefficients are jointly significant at the one percent level, while individual coefficients are now significant at the ten percent level, at least.

[31] The same order holds when I compute average wages by development quartiles. I prefer classify my occupations using median instead of average wages by quartile because the former measure does not depend on extreme values, a feature that is particularly present in my data at low development levels.

Table 9: Development Elasticity of the Occupational Skill-Premium
Sensitivity to Different Grouping Criteria

	(1)	(2)	(3)	(4)
	Baseline	-	-	-
	High-Skill	High-Skill	High-Skill	High-Skill
	1,2,3,4	1,2,3	1,2,3,4,7	1,2,3,4,7,8
Variables	$log \frac{w_{hs}}{w_{ls}}$	$log \frac{w_{hs}}{w_{ls}}$	$log \frac{w_{hs}}{w_{ls}}$	$log \frac{w_{hs}}{w_{ls}}$
$log(y_c)$	-0.360***	-0.415***	-0.251***	-0.250***
	(0.093)	(0.097)	(0.072)	(0.072)
$\mathbb{1}_{\{c\in 2\}}$ $log(y_c)$	0.047***	0.054***	0.022*	0.025*
	(0.017)	(0.018)	(0.013)	(0.013)
$\mathbb{1}_{\{c\in 3 \cup c\in 4\}}$ $log(y_c)$	0.067***	0.080***	0.045**	0.050***
	(0.024)	(0.025)	(0.018)	(0.019)
Constant	3.605***	4.096***	2.644***	2.576***
	(0.724)	(0.755)	(0.562)	(0.565)
Observations	80	80	80	80
R^2	0.229	0.252	0.187	0.160
Adjusted R^2	0.198	0.222	0.155	0.127
Prob > F	0.000181	5.94e-05	0.00123	0.00395

* Standard errors in parentheses (*** p < 0.01, ** p < 0.05, * p < 0.1).

Extreme Values.

In addition to the robustness checks discussed above, I here study if my results are driven by the presence of extreme values in the occupational skill-premium. To that end, I re-estimate Model (3) in Table 2 taking out of my estimation sample the following countries, one at a time: Norway, Tanzania, Rwanda, Hong Kong, Laos, South Africa, Gambia, Tajikistan. I also explore how the results change if I drop all these countries together from my sample. In all cases, the results are robust to the exclusion of these countries. The estimated elasticities are both jointly and statistically significant at the one percent level [32]. The goodness of fit, as measured by the regression's R^2 improves for all cases, with the exception of the estimations that leave Norway and Laos out.

The Role of the Economic Environment and Institutions.

It is still an open discussion in the economic literature to what extent skill-premia is determined by worker's skills or attributes and how much of it depends on variables that affect the economic environment of countries, like the quality of institutions, openness to trade, their economic structure, and other cultural, organizational, or social norms in

[32] The only exception is when I exclude Laos, where the coefficient for the second quartile is significant at the five percent level.

place [33].

A major concern related to this discussion is that the process of economic development is often characterized by significant improvements in the economic environment of countries. These improvements are sometimes driven by enhanced institutions, higher openness to trade, or other organizational. If those changes are neutral, in the sense that they do not affect workers of unlike skill types differently, they should not have an impact on the skill-premium. On the other hand, if these types of changes affect workers of different skill types in heterogeneous ways, one should be cautious before ignoring their effects on skill-premium evolution across the development spectrum.

To attend these concerns, I study if the quantitative results presented in Subsection 3.1 remain robust after controlling for different sets of institutional, organizational, economic policy quality, and economic structure variables that might have an effect on the occupational skill-premium behavior.

To that purpose, I explore specifications that control for different set of institutional quality variables used in the literature. To be precise, I study the effects of three groups of institutional quality controls: the components of the Index Economic Freedom, the set of variables in the Worldwide Governance Indicators, and the institutional controls used by Acemoglu et al. (2014). The estimation results are presented in Table D below, under labels Model (2.1)-(2.3).

[33] For example, in Caselli and Ciccone (2019)'s words: *"it seems extremely implausible that attributes of workers are the sole determinant of skill-premia not accounted for by skill supply. Instead, it seems very likely that skill-premia are also shaped by institutions, technology, organizational structures, infrastructure, the structural composition of the economy, openness to trade, social norms, and other attributes of the environment."*

Table 10: Development Elasticity of the Occupational Skill-Premium
(robustness of non-linearities under institutional controls)

	Model (2.1)	**Model (2.2)**	**Model (2.3)**	**Model (2.4)**
Variables	$log\ \frac{w_{hs}}{w_{ls}}$	$log\ \frac{w_{hs}}{w_{ls}}$	$log\ \frac{w_{hs}}{w_{ls}}$	$log\ \frac{w_{hs}}{w_{ls}}$
$log(y_c)$	-0.430***	-0.454***	-0.373***	-0.461***
	(0.128)	(0.107)	(0.118)	(0.099)
$\mathbb{1}_{\{c\in 2\}}\ log(y_c)$	0.059***	0.050***	0.055***	0.043**
	(0.020)	(0.018)	(0.018)	(0.017)
$\mathbb{1}_{\{c\in 3\cup c\in 4\}}\ log(y_c)$	0.076***	0.070***	0.072***	0.064***
	(0.027)	(0.024)	(0.024)	(0.023)
X_{EFI}	[1]	-	-	-
	-	-	-	-
X_{WGI}	-	[2]	-	-
	-	-	-	-
Rule of Law	-	-	0.057	-
	-	-	(0.063)	-
Years of School.	-	-	-0.024	-
	-	-	(0.022)	-
Ind. VA Share	-	-	-	0.011
	-	-	-	(0.007)
Serv. VA Share	-	-	-	0.016**
	-	-	-	(0.006)
Constant	4.174***	4.424***	3.886***	3.418***
	(1.035)	(0.870)	(0.912)	(0.722)
Observations	79	80	77	80
R-squared	0.399	0.326	0.258	0.290
Adjusted R-squared	0.256	0.239	0.206	0.242
Prob > F	0.00229	0.000670	0.000621	9.55e-05

* Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

[1] X_{EFI} is a vector of controls composed by the sub-indeces in the Index of Economic Freedom Index, including: Property Rights, Government Integrity, Tax Burden, Government Spending, Fiscal Health, Business Freedom, Labor Freedom, Trade Freedom, Investment Freedom, and Financial Freedom. The Government Freedom component is statistically significant at the 1% level. All other components are not statistically significant at the 10% level.

[2] X_{WGI} is a vector of controls composed by the variables in the World Governance Indicators, including: Government Effectiveness, Political Stability and Absence of Violence, Regulatory Quality, Rule of Law,Voice and Accountability, and Control of Corruption. The Political Stability and Absence of Violence component is statistically significant at the 5% level.All other components are not statistically significant at the 10% level.

The Table shows that the results are robust after controlling for the three sets of institutional variables under consideration. The elasticity coefficients are all individually statistically significant at the one percent level and the Adjusted R^2 improves in the three cases, being Model (1) the best specification under this criterion. Quantitatively, the biggest change in the estimated elasticities compared to Model 3 in Table 2 is in Model 2.2 (-0.454,-0.404,-0.384), followed by Model 2.1 (-0.430,-0.371,-0.354), and Model 2.3 (-0.373,-0.318,-0.302).

Model (2.4) shows how the results change after controlling for variables that capture the economic structure of countries, namely, the share of Total Value Added in Industry and the share of Total Value Added in Services [34]. The results are, again, similar to the ones in Model (3), with the individual coefficients being all statistically significant at the one percent level. Quantitatively, the elasticities show the biggest change for all the Models presented in Table increasing in absolute value to 0.461,0.418, and 0.397, respectively [35].

Relative Total Labor Income versus Relative Hourly Labor Income.

I here study if my results are driven by different trends in hours worked across development between my major occupational groups. To that end, instead of studying the behavior of relative hourly labor income by broad occupational groups I focus on the behavior of relative total labor income in high- and low-skill occupations. Table 11 below presents the same regressions as those in Table 2 in Subsection 3.1.

[34] I explored with different combinations of the share of Value Added in Agriculture, Industry, and Services. The elasticities are statistically significant at the one percent level in all cases. Model (2.4) presents the best specification I found.

[35] The results are also robust after jointly controlling for all the variables in the three institutional control groups and the economic structure variables.

Table 11: Development Elasticity of the Occupational Skill-Premium
(relative occupational total labor income)

Variables	Model (1) $log\ \frac{w_{hs}}{w_{ls}}$	Model (2) $log\ \frac{w_{hs}}{w_{ls}}$	Model (3) $log\ \frac{w_{hs}}{w_{ls}}$
$log(y_c)$	-0.102**	-0.317*	-0.371***
	(0.046)	(0.162)	(0.132)
$\mathbb{1}_{\{c \in P2\}}\ log(y_c)$	-	0.040	0.047 *
	-	(0.027)	(0.024)
$\mathbb{1}_{\{c \in P3\}}\ log(y_c)$	-	0.066*	-
	-	(0.036)	-
$\mathbb{1}_{\{c \in P4\}}\ log(y_c)$	-	0.057	-
	-	(0.045)	-
$\mathbb{1}_{\{c \in P3,4\}}\ log(y_c)$	-	-	0.074**
	-	-	(0.034)
Constant	1.631***	3.219**	3.643***
	(0.435)	(1.262)	(1.033)
Observations	80	80	80
R-squared	0.059	0.119	0.115
Adjusted R-squared	0.047	0.072	0.083
Prob >F	0.0294	0.0467	0.0249

* Standard errors in parentheses (*** p $\le$ 0.01, ** p $\le$ 0.05, * p $\le$ 0.1).

The main message from Table 11 is that, qualitatively, the results are similar to the ones obtained in the relative hourly labor income. The best specification is still given by the model where elasticities vary with development until countries reach the third quartile, being all the estimated coefficients statistically significant at least at the ten percent level and jointly significant at the five percent level. Quantitatively, both the initial predicted premium and the estimated elasticities are, in absolute value, smaller. To be precise, the constant falls from 3.63 to 3.15, while the estimated elasticities are -0.315, -0.276, -0.253, compared to -0.362, -0.314, -0.295 in the case I use relative hourly labor income as my skill-premium measure.

The intuition behind this result becomes more clear after looking at the evolution of average weekly hours worked by major occupational groups across development, which I summarize in Table ?? in Appendix C [36]. As we can see, hours are initially higher in low-skill occupations and exhibit a higher decline as we move from the bottom to the top development quartile. These two effects together amplify the decline in the skill-premium for the following reasons: first, the fact that hours worked are higher in Low- than in high-skill occupations at low development levels drives the initial skill-premium up for

[36] I also show in Appendix C how ILO data on hours worked compares to others in the literature.

the relative hourly labor income measure; second, the fact that hours worked decline at a faster pace in low-skill occupations as countries develop increases (in absolute terms) the GDP per capita elasticities the relative hourly labor income measure.

Workers of Both Sexes vs Males Only.

I here analyze to what extent my results depend on the inclusion or not of women when computing relative hourly labor income. I perform this robustness check since, for example, it might be a concern that women's attachment to the labor force or the gender wage gap vary with development.

I thus repeat my empirical analysis but now taking into account employment, hours worked and labor income data for male workers only. The regression results are presented in Table 12 below.

Table 12: Development Elasticity of the Occupational skill-premium
(including males workers only)

Variables	**Model (1)** $log \frac{w_{hs}}{w_{ls}}$	**Model (2)** $log \frac{w_{hs}}{w_{ls}}$	**Model (3)** $log \frac{w_{hs}}{w_{ls}}$
$log(y_c)$	-0.061*	-0.300**	-0.321***
	(0.033)	(0.117)	(0.097)
$\mathbb{1}_{\{c \in 2\}} \; log(y_c)$	-	0.046**	0.049***
	-	(0.020)	(0.018)
$\mathbb{1}_{\{c \in 3\}} \; log(y_c)$	-	0.070**	-
	-	(0.027)	-
$\mathbb{1}_{\{c \in 4\}} \; log(y_c)$	-	0.066**	-
	-	(0.033)	-
$\mathbb{1}_{\{c \in 3,4\}} \; log(y_c)$	-	-	0.073***
	-	-	(0.025)
Constant	1.296***	3.076***	3.243***
	(0.314)	(0.922)	(0.757)
Observations	79	79	79
R-squared	0.042	0.143	0.142
Adjusted R-squared	0.030	0.097	0.107
Prob >F	0.0005	0.021	0.009

* Standard errors in parentheses (*** p 0.01, ** p 0.05, * p 0.1).

As we can see, the results are in line with those obtained when considering workers of both sexes. The best specification is still given by the model where the elasticity varies with development and stabilizes in the third quartile, with the estimated coefficients being all significant at the one percent level.

E Proofs

Proposition 1

Under independent marginals and common shape parameter the share of workers in high-skill occupations is given by

$$
\begin{aligned}
\pi_h &= P(w_h Z_h \geq w_l Z_l) = P\left(Z_h \geq \frac{w_l}{w_h} Z_l\right) \\
&= \int_0^{\infty}\int_0^{\frac{w_h}{w_l} z_h} g_l(z_l)\, g_h(z_h)\, dz_l dz_h \\
&= \int_0^{\infty} e^{-S_l \left(z_h \frac{w_h}{w_l}\right)^{-\theta}} g_h(z_h)\, dz_h \\
&= \int_0^{\infty} e^{-S_l \left(z_h \frac{w_h}{w_l}\right)^{-\theta}} S_h z_h^{-\theta-1} e^{-S_h (z_h)^{-\theta}} dz_h \\
&= S_h \int_0^{\infty} \theta z_h^{-\theta-1} e^{-\left(S_l \left(\frac{w_h}{w_l}\right)^{-\theta} + S_h\right) z_h^{-\theta}} dz_h \\
&= \frac{S_h}{S_l \left(\frac{w_h}{w_l}\right)^{-\theta} + S_h} \int_0^{\infty} \theta \left(S_l \left(\frac{w_h}{w_l}\right)^{-\theta} + S_h\right) z_h^{-\theta-1} e^{-\left(S_l \left(\frac{w_h}{w_l}\right)^{-\theta} + S_h\right) z_h^{-\theta}} dz_h \\
&= \frac{S_h}{S_l \left(\frac{w_h}{w_l}\right)^{-\theta} + S_h} \left[e^{-\left(S_l \left(\frac{w_h}{w_l}\right)^{-\theta} + S_h\right) z_h^{-\theta}} \right]_0^{\infty} = \frac{S_h}{S_l \left(\frac{w_h}{w_l}\right)^{-\theta} + S_h},
\end{aligned}
$$

and the distribution of labor productivity conditional on workers choosing high-skill occupations is

$$
\begin{aligned}
M_h(z) &= G_h(Z_h \leq z \mid w_h Z_h \geq w_l Z_l) \\
&= \int_0^{z}\int_0^{z_h \frac{w_h}{w_l}} g_h(z_h) g_l(z_l) dz_l dz_h \\
&= \frac{S_h}{S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}} \int_0^{z} \theta z_h^{-\theta-1} \left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) e^{-\left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) z_h^{-\theta}} dz_h \\
&= \pi_h e^{-\left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) z^{-\theta}}.
\end{aligned}
$$

Average labor productivity in high-skill occupations is

$$
\begin{aligned}
\mathbb{E}\left[Z_h|w_hz_h - w_lz_l\right] &= \int_0^\infty z_h M(z_h|w_hz_h - w_lz_l)dz_h,\\
&= \int_0^\infty z_h \frac{g(z_h, w_hz_h - w_lz_l)}{g(w_hz_h - w_lz_l)}dz_h,\\
&= \frac{1}{\pi_h}\int_0^\infty \int_0^{\frac{w_h}{w_l}z_h} z_h g(z_h, z_l)dz_l dz_h,
\end{aligned}
$$

in the independent case

$$
\mathbb{E}\left[Z_h|w_hz_h - w_lz_l\right] = \frac{1}{\pi_h}\int_0^\infty \int_0^{\frac{w_h}{w_l}z_h} z_h g_h(z_h) g_l(z_l)dz_l dz_h.
$$

Under common shape parameter

$$
\begin{aligned}
\mathbb{E}\left[Z_h|w_hz_h - w_lz_l\right] &= \frac{1}{\pi_h}\int_0^\infty z_h g_h(z_h) e^{-\left(S_l \frac{w_h}{w_l} z_h\right)^{-\theta}} dz_l dz_h,\\
&= \frac{1}{\pi_h}\int_0^\infty z_h \theta S_h z_h^{-(\theta+1)} e^{-S_h(z_h)^{-\theta}} e^{-\left(S_l \frac{w_h}{w_l} z_h\right)^{-\theta}} dz_h,\\
&= \int_0^\infty z_h \theta z_h^{-(\theta+1)} \left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) e^{-\left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) z_h^{-\theta}} dz_h,\\
&= \int_0^\infty z_h \theta z_h^{-(\theta+1)} \left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) e^{-\left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) z_h^{-\theta}} dz_h,
\end{aligned}
$$

Let $y = \left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right) z_h^{-\theta}$. It follows that

$$
\mathbb{E}\left[Z_h|w_hz_h - w_lz_l\right] = \int_0^\infty \left(\frac{y}{S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}}\right)^{-(\frac{1}{\theta})} e^{-y}dy,
$$

$$
\mathbb{E}\left[Z_h|w_hz_h - w_lz_l\right] = \left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right)^{(\frac{1}{\theta})} \Gamma\left(1 - \frac{1}{\theta}\right)
$$

□

Proposition 2 (Educational Sorting Conditional on Occupational Choice).

Let c^s and c^u denote the fixed costs of acquiring Secondary and University education, β^s the log-return of completing Secondary education and β_h^u and β_l^u the log-returns to completing University education in high- and low-skill occupations, respectively. Assume $c^u > c^s > 0$ and $\beta_h^u > \beta_l^u > \beta^s > 0$.

Conditional on working in high-skill occupations, workers' incomes after educational costs are:

1. **High-skill occupations, No-Schooling:** $W_h^{ns} = w_h z_h$
2. **High-skill occupations, Secondary Complete:** $W_h^s - c^s = w_h z_h e^{\beta^s} - c^s$
3. **High-skill occupations, University Complete:** $W_h^u - c^u = w_h z_h e^{\beta_h^u} - c^u$

Workers choose no-schooling if:

$$\begin{aligned} W_h^{ns} &\geq W_h^s - c^s \\ w_h z_h &\geq w_h z_h e^{\beta^s} - c^s \\ z_h &\leq \frac{c^s}{w_h (e^{\beta^s} - 1)} = z_h, \end{aligned}$$

and

$$\begin{aligned} W_h^{ns} &\geq W_h^u - c^u \\ w_h z_h &\geq w_h z_h e^{\beta_h^u} - c^u \\ z_h &\leq \frac{c^u}{w_h \left(e^{\beta_h^u} - 1\right)} = \tilde{z}_h. \end{aligned}$$

Secondary Education is chosen if

$$\begin{aligned} W_h^s - c^s &> W_h^{ns} \\ w_h z_h e^{\beta^s} - c^s &> w_h z_h \\ z_h &> \frac{c^s}{w_h (e^{\beta^s} - 1)} = z_h, \end{aligned}$$

and

$$\begin{aligned} W_h^s - c^s &\geq W_h^u - c^u \\ w_h z_h e^{\beta^s} - c^s &\geq w_h z_h e^{\beta_h^u} - c^u \\ z_h &\leq \frac{c^u - c^s}{w_h \left(e^{\beta_h^u} - e^{\beta^s}\right)} = z_h . \end{aligned}$$

Finally, still conditional on working in high-skill occupations, individuals choose to acquire University schooling if

$$
\begin{aligned}
W_h^u - c^u &\geq W_h^{ns} \\
w_h z_h e^{\beta_h^u} - c^u &\geq w_h z_h \\
z_h &\geq \frac{c^u}{w_h\left(e^{\beta_h^u} - 1\right)} \equiv \tilde{z}_h,
\end{aligned}
$$

and

$$
\begin{aligned}
W_h^u - c^u &\geq W_h^s - c^s \\
w_h z_h e^{\beta_h^u} - c^u &\geq w_h z_h e^{\beta^s} - c^s \\
z_h &\geq \frac{c^u - c^s}{w_h\left(e^{\beta_h^u} - e^{\beta^s}\right)} \equiv z_h
\end{aligned}
$$

To guarantee that the educational attainment rule is monotonically increasing in worker's ability, one needs $z_h < \tilde{z}_h < z_h$. Otherwise, if $\tilde{z}_h < z_h$, workers with ability $z_h \in [0, \tilde{z}_h]$ choose no schooling over Secondary school, workers with ability $z_h \in (\tilde{z}_h, z_h]$ choose University over Secondary, individuals with $z_h \in (z_h, z_h]$ choose Secondary over University and no schooling, and workers with $z_h \geq z_h$ choose University education. Following the same logic, if $\tilde{z}_h \geq z_h$, the educational attainment rule is monotonically increasing in ability until z_h, where a region of workers with $z_h \in (z_h, \tilde{z}_h)$ emerges and where workers prefer Secondary Schooling to no schooling, University schooling to Secondary schooling and no schooling over University schooling could emerge for workers with abilities $z_h \in (z_h, \tilde{z}_h)$.

After some algebra, a sufficient condition for $z_h < \tilde{z}_h < z_h$ to hold is

$$
\frac{c^u}{c^s} \geq \frac{\beta_h^u - 1}{\beta^s - 1}
$$

Conditional on working in low-skill occupations, workers' incomes after educational costs are:

1. **No-Schooling:** $W_l^{ns} = w_l z_l$
2. **Secondary Complete:** $W_l^s - c^s = w_l z_l e^{\beta^s} - c^s$
3. **University Complete:** $W_l^u - c^u = w_l z_l e^{\beta_l^u} - c^u$

Following the same logic, the ability thresholds to choose Secondary education over no schooling z_l and University over Secondary education z_l conditional on working in low-skill occupations are given by

$$z_l = \frac{c^s}{w_l(e^{\beta^s} - 1)}$$
$$z_l = \frac{c^u - c^s}{w_l\left(e^{\beta_l^u} - e^{\beta^s}\right)}.$$

A sufficient condition for the educational decision rule to be monotonically increasing in workers' ability in low-skill occupations is

$$\frac{c^u}{c^s} > \frac{\beta_l^u - 1}{\beta^s - 1},$$

which, given that $\beta_h^u > \beta_l^u$ is guaranteed by $\frac{c^u}{c^s} > \frac{\beta_h^u - 1}{\beta^s - 1}$.

□

Proposition 3

Define $\tilde{z}_l = z_h \frac{w_h}{w_l} - \frac{c^u - c^s}{w_l\left(e^{\beta_h^u} - e^{\beta^s}\right)}$. Under endogenous human capital accumulation through schooling, the occupational selection decisions are:

1. If $z_l \leq z_l$ workers choose high-skill occupations if $z_h > \frac{w_l}{w_h} z_l$.

2. If $z_l \in (z_l, \tilde{z}_l]$ workers choose high-skill occupations if $z_h > \frac{w_l}{w_h} z_l$.

3. If $z_l \in (\tilde{z}_l, z_l]$ workers choose high-skill occupations if $z_h > \frac{w_l}{w_h} \frac{z_l e^{\beta^s}}{e^{\beta_h^u}}$.

4. If $z_l > z_l$ workers choose high-skill occupations if $z_h > \frac{w_l}{w_h} \frac{z_l e^{\beta_l^u}}{e^{\beta_h^u}}$.

The proof goes by contraposition.

1. Suppose not. Then, there exists a $z_l \in [0, z_l]$ such that $w_h z_h \leq w_l z_l$ and workers choose to work in high-skill occupations.

Then there exists a $z_l \in [0, z_l]$ such that

$$w_h z_h \leq w_l z_l$$

and

$$\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} > \max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\}.$$

Since $z_l \in [0, z_l]$, $\max_{\{W_l^{ns}, W_l^s, W_l^u\}} \{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\} = W_l^{ns} = w_l z_l$.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\} = W_h^{ns} \Longrightarrow w_h z_h > w_l z_l$, which is a contradiction.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\} = W_h^s = w_h z_h e^{\beta^s}$, $z_h \in (z_h, z_h]$, which together with $w_l z_l \geq w_h z_h \implies z_l > \frac{w_h}{w_l} \underline{z}_h = \underline{z}_l$, which is a contradiction.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left(W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right) = W_h^u = w_h z_h e^{\beta_h^u}$, $z_h \in (\underline{z}_h, \infty)$, which together with $w_l z_l \geq w_h z_h \implies z_l > \frac{w_h}{w_l} \underline{z}_h = \tilde{z}_l > z_l$, which is a contradiction.

2. Suppose not. Then, there exists a $z_l \in (\underline{z}_l, \tilde{z}_l)$ such that $w_h z_h \leq w_l z_l$ and workers choose to work in high-skill occupations.

Then there exists a $z_l \in (\underline{z}_l, \tilde{z}_l)$ such that

$$w_h z_h \leq w_l z_l$$

and

$$\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} > \max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\}.$$

Since $z_l \in (\underline{z}_l, \tilde{z}_l)$, $\max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\} = W_l^s = w_l z_l e^{\beta^s}$.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^{ns} \implies w_h z_h > w_l z_l e^{\beta^s} > w_l z_l$, which is a contradiction.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^s = w_h z_h e^{\beta^s} \implies w_h z_h e^{\beta^s} > w_l z_l e^{\beta^s} \implies w_h z_h > w_l z_l$, which is a contradiction.

3. Suppose not. Then, there exists a $z_l \in (\tilde{z}_l, \overline{z}_l)$ such that $w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta^s}$ and workers choose to work in high-skill occupations.

Then there exists a $z_l \in (\tilde{z}_l, \overline{z}_l)$ such that

$$w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta^s}$$

and

$$\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} > \max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\}.$$

Since $z_l \in (\tilde{z}_l, \overline{z}_l)$, $\max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\} = W_l^s = w_l z_l e^{\beta^s}$. Also, $w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta^s} \implies w_h z_h \leq w_l z_l$.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^{ns}$,

$$w_h z_h > w_l z_l e^{\beta^s}$$
$$w_h z_h > w_l z_l, \text{ which is a contradiction.}$$

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^s$,

$$w_h z_h e^{\beta^s} > w_l z_l e^{\beta^s}$$
$$w_h z_h > w_l z_l, \text{ which is a contradiction.}$$

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^u$,

$$w_h z_h e^{\beta_h^u} > w_l z_l e^{\beta^s} \text{ which is a contradiction.}$$

4. Suppose not. Then, there exists a $z_l \in (z_l, \infty)$ such that $w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta_l^u}$ and workers choose to work in high-skill occupations.
Then there exists a $z_l \in (z_l, \tilde{z}_l)$ such that

$$w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta_l^u}$$
and
$$\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} > \max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{ W_l^{ns}, W_l^s - c^s, W_l^u - c^u \right\}.$$

Since $z_l \in (z_l, \infty)$, $\max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{ W_l^{ns}, W_l^s - c^s, W_l^u - c^u \right\} = W_l^u = w_l z_l e^{\beta_l^u}$. Also, $w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta_h^l} \implies w_h z_h \leq w_l z_l$.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^{ns}$,

$$w_h z_h > w_l z_l e^{\beta_l^u}$$
$$w_h z_h > w_l z_l, \text{ which is a contradiction.}$$

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left[W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right] = W_h^s$,

$$w_h z_h e^{\beta^s} > w_l z_l e^{\beta_l^u}$$

$$w_h z_h > w_l z_l, \text{ which is a contradiction.}$$

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left[W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right] = W_h^u$,

$$w_h z_h e^{\beta_h^u} > w_l z_l e^{\beta^s} \text{ which is a contradiction.}$$

Proposition 4

The share of workers in high-skill occupations is given by:

$$\pi_h = \int_0^{z_h} \int_0^{\frac{w_h}{w_l} z_h} g_\phi(z_h, z_l)\, dz_h dz_l + \int_{z_h}^{z_h} \int_0^{\frac{w_h}{w_l} \frac{e^{\beta_h^u}}{e^{\beta^s}} z_h} g_\phi(z_h, z_l)\, dz_h dz_l + \int_{z_h}^{\infty} \int_0^{\frac{w_h}{w_l} \frac{e^{\beta_h^u}}{e^{\beta_l^u}} z_h} g_\phi(z_h, z_l)\, dz_h dz_l$$

The proof follows from Proposition 3. The first term on the right hand side captures all the workers who choose high-skill occupations if $z_h > \frac{w_l}{w_h} z_l$. The limits of the outer integral capture the abilities in high-skill occupations that make workers indifferent between choosing high- and low-skill occupations in the limits of the interval, namely $[0, z_h]$ (and $[0, \tilde{z}_l]$). The limits of the inner integral captures all the individuals who have ability below the indifference level in low-skill occupations for a given ability level in high-skill occupations, namely, the $z_l \in \left[0, \frac{w_h}{w_l} z_h\right]$.

The second and third terms on the right hand side of the equation are analogous, with the only difference that the second term captures the ability region were workers get Secondary education in low-skill occupations and University education in high-skill occupations, and the third term captures the ability region were workers get University education in high- and low-skill occupations.

F Census/American Community Survey data issues and handling.

To discipline the innate ability distribution I use IPUMS International data for the US in 2010, which contains Household Survey micro-data from the American Community survey, harmonized to allow for international comparisons. The sample includes 1% of the United State's population.

In order to minimize noise in the calculations, wages are computed only for workers with a considerable attachment to the labor force, following Acemoglu and Autor (2011) criteria. Thus, I consider only full-time (i.e. at least 35 hours per week), full year (40 weeks per year or more) workers, aged 16-64, and exclude those who are in the military, institutionalized, or self employed. I construct hourly earnings as the ratio of annual earnings and total hours worked, being the latter the product of average weeks worked per year and average hours worked per week. Calculations are weighted by ACS sampling weights and are converted in real terms using the personal consumer expenditure (PCE) deflator. Eearnings below US$ 1.675 per hour (US$ 67 per week in Acemoglu and Author over 40 hours per week) in 1982 dollars are dropped. I replace income for top-coded earners with 1.45 times the value assigned to the corresponding top-level income, which in 2010 requires identifying the 99.5th percentile of income by state.

Separating occupational labor income from occupational wages that are common across workers and occupations requires and identifying assumption. To that end, I first classify occupations into two broad groups, high-, and low-skill, by following the procedure described in Section 2. I then assume that the average labor productivity of workers with no experience and no educational attainment in high- and low-skill occupations equals unity. Thus, the average labor income of the individuals in these two groups give me efficiency wages in high- and low-skill occupations. Labor productivity for the individuals in high- and low-skill occupations that do not belong to the base groups is obtained by dividing their labor income by the corresponding occupational efficiency wages.

References

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. Quarterly Journal of Economics, 113(4):1055–89.

Acemoglu, D. (2002). Directed technical change. Review of Economic Studies, 69(4):781–809.

Acemoglu, D. and Autor, D. (2011). Skills, Tasks, and Technologies: Implications for Employment and Earnings. Elsevier, North-Holland.

Acemoglu, D., Gallego, F., and Robinson, J. A. (2014). Institutions, human capital, and development. The Annual Review of Economics, 6:875–912.

Autor, D. H., Katz, L. F., , and Krueger, A. B. (1998). Computing inequality: Have computers changed the labor market? Quarterly Journal of Economics, 113(4):1169–1213.

Bick, A., Fuchs-Schundeln, N., and Lagakos, D. (2018). How do hours worked vary with income? cross-country evidence and implications. American Economic Review, 108(1):170–199.

Buera, F., Rogerson, R., Kaboski, J., and Vizcaino, J. I. (2018). Skill-biased technical change. Manuscript.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. Elsevier, North-Holland.

Caselli, F. (2016). Technology differences over space and time. Manuscript, -(-):–.

Caselli, F. and Ciccone, A. (2019). The human capital stock: A generalized approach: Comment. American Economic Review, 109(3):1155–74.

Caselli, F. and Coleman, W. J. I. (2006). The world technology frontier. American Economic Review, 96(3):500–522.

Erosa, A., Koreshkova, T. A., and Restuccia, D. (2010). How important is human capital: A quantitative theory assesment of world income inequality. Review of Economic Studies, 52(01):1–32.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the penn world table. American Economic Review, 105(10):3150–3182.

Gollin, D., Lagakos, D., and Waugh, M. E. (May 2014). The agricultural productivity gap. The Quarterly Journal of Economics, 129(2):939–993.

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? The Quarterly Journal of Economics, 114(1):83–166.

Heckman, J. T. and Honoré, B. E. (1990). The empirical content of the roy model. Econometrica, 58(5):1121–49.

Jaume, D. (2019). The labor market effects of an educational expansion. a theoretical model with applications to brazil. Manuscript, -(-):–.

Jones, B. F. (November 2014). The human capital stock: A generalized approach. American Economic Review, 104(11):3752–77.

Katz, L. F. and Autor, D. H. (1999). Changes in the wage structure and earnings inequality. Elsevier, North-Holland.

Katz, L. F. and Murphy, K. M. (1992). Changes in relative wages, 1963-1987: Supply and demand factors. Quarterly Journal of Economics, 107(1):35–78.

Klenow, P. J. and Rodriguez-Clare, A. (1997). The neoclassical revival in growth economics: Has it gone too far? in: B.S. Bernanke, and J.J. Rotemberg, eds., NBER 66 macroeconomics annual.

Lagakos, D. and Waugh, M. E. (April 2013). Selection, agriculture, and cross-country productivity differences. American Economic Review, 103(2):948–980.

Malmberg, H. (2018). How does the eciency of skilled labor vary across rich and poor countries? an analysis using trade and industry data,. Manuscript, -(-):–.

Mankiw, G. M., Romer, D., and Weil, D. N. (1992). A contribution to the empirics of economic growth. The Quarterly Journal of Economics, 107(2):407–437.

Manuelli, R. E. and Seshadri, A. (September 2014). Human capital and the wealth of nations. American Economic Review, 104(2736-2762).

Nelsen, R. B. (2006). An Introduction to Copulas. Springer.

Rossi, F. (2017). The relative efficiency of skilled labor across countries: Measurement and interpretation. Manuscript, -(-):–.