# The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation*

Federico Rossi†

August 2017

**Abstract**

This paper studies how the relative productivity of skilled and unskilled labor varies across countries. I use both micro-data and other sources for countries at different stages of development to document that the skill premium varies little between rich and poor countries, in spite of large differences in the relative skill supply. This pattern is consistent with the view that the relative productivity of skilled workers is higher in rich countries. I propose a methodology based on the comparison of labor market outcomes of immigrants with different levels of educational attainment to discriminate between technology and unobserved human capital as drivers of these patterns. I find that human capital quality plays a minor role in explaining cross-country differences in relative skill efficiency.

**JEL Classification**: O11, O47, I25, E24

*I am grateful to Francesco Caselli, Pete Klenow and David Lagakos for useful comments and discussions.

†London School of Economics and CFM. Corresponding address: 32 Lincoln's Inn Fields, WC2A 3PH, London, UK; email: f.rossi2@lse.ac.uk.

# 1 Introduction

A question of major interest in macroeconomics is how the structure of production varies across countries. The traditional view is that rich and poor countries are set apart by large differences in a factor-neutral productivity shifter, while gaps in the relative amount and productivity of various factors of production are of more limited importance (Hall and Jones, 1999). Recently, this view has been challenged, thanks both to improved measurements of production inputs (Schoellman, 2012; Lagakos et al., 2016) and richer characterizations of the production technology (Jones, 2014; Caselli, 2016).

An emerging view in this line of research is that the relative efficiency of skilled and unskilled workers varies substantially across countries (Caselli and Coleman, 2006; Caselli, 2016; Malmberg, 2017). This conclusion typically follows from the analysis of quantities and prices. In a world with imperfect substitutability, a higher relative supply of skilled labor should be reflected in a lower relative price. However, existing estimates for the skill premium display limited variability across countries, in spite of large gaps in enrollment rates and educational achievements. This suggests that high-skilled workers are much more productive in rich (and skill-abundant) countries, attenuating the downward pressure on the skill premium stemming from their high supply. Cross-country gaps in the productivity of unskilled labor are instead moderate in size.

Different interpretations have been proposed to explain these patterns. One possibility, first advanced by Caselli and Coleman (2006) and Caselli (2016), is that technological differences across countries are factor-biased, and firms in rich countries adopt technologies more suitable for skilled workers. A natural alternative is that the human capital gap between high and low skilled workers is larger in rich countries, because of differences in educational quality, training or workers' intrinsic characteristics (Jones, 2014; Malmberg, 2017). In a cross-country setting, distinguishing between the two interpretations has important implications for various open questions in macro-development, such as the degree of transferability of technology across space and the role of human capital in accounting for cross-country gaps in economic performance.

In this paper I re-examine the measurement and intepretation of cross-country differences in relative skill efficiency. Using both aggregate and micro-level data, I confirm that gaps in the relative productivity of skilled and unskilled labor are large and not driven by the limited comparability or reliability of some of the sources used in previous studies. Building on this finding, I propose an approach based on the analysis of US immigrants to separately identify the role of technology and human capital in explaining the cross-country variation in relative skill efficiency.

The main data contribution of the paper consists in the construction of highly comparable estimates for the skill premium across countries. The lack of such information has represented a major drag on the existing literature, which has relied either on imputations based on related quantities, or on the use of sources not fully consistent with the underlying modelling strategy. To improve on this, I use micro-data from IPUMS International on 12 countries at different stages of development, ranging from the United States to India. I estimate the skill premium using the same specifications and similar sample restrictions for all countries. While the magnitude of some of the estimates is quite different from existing sources, I confirm the finding that the skill premium varies little across countries.

Through the lens of a simple production function setting, I back out the implied relative efficiency of skilled labor for each country, using both micro-data from IPUMS and more traditional sources to estimate the relevant parameters. I embed in this framework differences in both relative human capital and technology bias, and show that the estimated relative skill efficiency is a composite of the two. I confirm that relative skill efficiency varies substantially across countries. The productivity gap between skilled workers in the United States and in the average country of the sample is 4 times as large as the corresponding gap for unskilled workers.

I then move to the analysis of US immigrants. The basic idea of my approach is that comparing the within-group skill premia across immigrant groups, educated in their countries of origin but observed in the same labor market provides a way to isolate cross-country differences in relative human capital quality, keeping constant the local technological environment and other institutional characteristics. Gaps in the relative productivity of skilled labor might reflect differences in educational quality, as emphasized in Schoellman (2012), or differential of sorting into higher education across countries. I extend the baseline framework to incorporate immigrant workers with different levels of human capital, and then bring it to the data using the 2000 US Census. Conditional on the relevant controls, the difference between country of origin-specific skill premia identifies the corresponding gap in terms of relative skill quality.

I find that the cross-country variation in relative skill quality is of limited magnitude. While the productivity gap between skilled and unskilled workers is higher in the United States compared to most countries, the differences are much smaller than what it would be expected in a world where human capital quality explained the cross-country gaps in skill efficiency. Indeed, I conclude that differences in the skill bias of technology accounts for more than 90% of the cross-country variance in skill efficiency. While in principle patterns of differential selection into migration as a function of skills and country of origin might contribute to shape these results, I argue that this concern is unlikely to majorly affect the basic conclusion of the paper.

My work fits in the literature on cross-country differences in the structure of production.

The basic approach to isolate skill-biased differences in productivity is introduced by Caselli and Coleman (2006), and subsequently updated by Caselli (2016). Recent work by Malmberg (2017) proposes an alternative methodology, based on trade data, to infer cross-country differences in the efficiency of skilled labor, and discusses the implications for development accounting. Compared to these papers, my main contributions are an improved measurement of skill premia and the development of a methodology to discriminate between relative skill quality and technology bias as sources of differences in skill efficiency. This distinction mirrors, on a cross-country dimension, a related debate on the relative roles of technology, human capital and sorting in explaining the rise of the skill premium over time (Acemoglu, 1998, 2002; Bowlus and Robinson, 2012; Hendricks and Schoellman, 2014).

This paper is also closely related to a growing literature studying the labor market experience of immigrants to learn about cross-country differences in human capital (Schoellman, 2012, 2016; Lagakos et al., 2016). In particular, Schoellman (2012) uses estimated Mincerian returns to schooling across immigrants' nationalities to quantify the role of educational quality for development accounting. While his focus is the aggregate human capital stock (in a model with perfect substitutability across skill levels), the main object of interest of my analysis is the relative quality of high skill and low skill workers. Immigrants from rich countries have higher returns both within and between skill levels, but the variation in returns between skill groups (which drive my estimates of relative skill quality) is more limited.

The paper is structured as follows. Section 2 introduces the basic framework and describes the measurement of relative skill efficiency. Section 3 shows evidence on immigrants, while Section 4 discusses the issue of selection. Finally, Section 5 concludes by discussing some implications and possibilities for future work.

# 2 Measuring Relative Skill Efficiency

In this section I document how the relative efficiency of skilled labor varies across countries. I introduce a simple framework, discuss how I bring it to the data and summarize the main patterns.

## 2.1 Framework

Suppose that country $c$ is endowed with the following production technology,

$$Y_c = A_c K_c^{\alpha} \left[ (A_{Hc} H_c)^{\rho} + (A_{Lc} L_c)^{\rho} \right]^{\frac{1-\alpha}{\rho}}$$

where $K_c$ is physical capital and $H_c$ and $L_c$ are total high-skilled and low-skilled labor services. The production function involves three different technological paramenters, potentially varying across countries: $A_c$ is total factor productivity, while $A_{Hc}$ and $A_{Lc}$ are factor biased technological terms, augmenting high- and low-skill labor. To simplify the notation, in what follows I omit the subscript $c$ where this does not generate confusion.

The total amounts of high- and low-skill services used for production are

$$H = Q_H \tilde{H} \tag{1}$$

$$L = Q_L \tilde{L} \tag{2}$$

where $\tilde{H}$ and $\tilde{L}$ are the quantities of skilled and unskilled workers, while $Q_H$ and $Q_L$ represent their quality, or the amount of labor services provided by a given worker. While $A_H$ and $A_L$ proxy for factors external to individuals, such as the available technologies and the features of the working environment, I think of $Q_H$ and $Q_L$ as capturing workers' embodied productivity, which is possibly the result of both accumulated knowledge and innate characteristics.

Assuming perfectly competitive labor markets, the wage ratio between skilled and unskilled workers is

$$\frac{w_H}{w_L} = \left(\frac{A_H Q_H}{A_L Q_L}\right)^{\rho} \left(\frac{\tilde{H}}{\tilde{L}}\right)^{\rho-1} \tag{3}$$

I refer to $\frac{A_H Q_H}{A_L Q_L}$ as the relative "efficiency" of skilled and unskilled workers. If $\rho > 0$, which is the empirically relevant case given the existing estimates of the elasticity of substitution (Ciccone and Peri, 2005), a higher efficiency of skilled labor raises the skill premium, conditional on factor supplies. There are two reasons why this relative efficiency might vary across countries: differences in the skill bias of technology, $\frac{A_H}{A_L}$, and differences in the relative quality of skilled labor, $\frac{Q_H}{Q_L}$.

## 2.2 Measurement

In this section I describe how I map this framework to the data. The key object of interest is the relative efficiency of skilled labor $\frac{A_H Q_H}{A_L Q_L}$, which I normalize so that it is 1 for the United States. I take 2000 as my baseline year, and consider data sources relative to (or as close as possible to) this date.

A key choice to make is the split between high- and low-skill labor. Following most of the literature, I adopt a criterion based on workers' level of educational attainment. In particular, I consider skilled workers all high-school graduates and above, while individuals with at most some secondary schooling are unskilled. I chose this threshold mainly for two reasons: (i) comparability, as this is a definition of skilled labor commonly used in the literature on skill-bias technology differences (Caselli, 2016), and more broadly on human capital and economic development (Jones, 2014); (ii) the fact that, using the same split, Ciccone and Peri (2005) provide a credibly identified estimate for the elasticity of substitution between skilled and unskilled labor. Indeed, I use their estimated elasticity of 1.5 throughout, which implies a value for $\rho$ of 1/3.

In practice, workers are obviously heterogeneous within these broad skill categories. I allow their productivity to depend on their level of educational attainment (indexed by $j$), gender (indexed by $g$) and experience (indexed by $exp$). For educational attainment, I split the unskilled in two groups (primary or less, some secondary) and the skilled in three groups (secondary completed, some tertiary, tertiary completed).[1] I define (potential) experience as the difference between current age and age at the end of education, and I consider nine groups based on 5-year intervals (0 to 4, 5 to 9, 10 to 14, 15 to 19, 20 to 24, 25 to 29, 30 to 34, 35 to 39, 40 or more). Within skill groups, human capital services provided by workers that differ in these dimensions are perfect subsitutes. I assume the aggregators $\tilde{H}$ and $\tilde{L}$ take the form[2]

$$\tilde{H} = \sum_{j \in H} \sum_{g} \sum_{exp} e^{\beta_j} e^{\lambda_g} e^{\mu_{exp}} n_{j,g,exp} \tag{4}$$

$$\tilde{L} = \sum_{j \in L} \sum_{g} \sum_{exp} e^{\beta_j} e^{\lambda_g} e^{\mu_{exp}} n_{j,g,exp} \tag{5}$$

where $n_{j,g,exp}$ is the number of workers belonging to the ($j$, $g$, $exp$) group, which I calculate from Barro and Lee (2013).[3] The parameters $\beta_j$, $\lambda_g$ and $e^{\mu_{exp}}$ capture the productivity differentials (compared to a baseline category) associated to each educational achievement, gender and experience level. I normalize $\beta_{prim} = \beta_{sec} = \lambda_{male} = \mu_{0to4} = 0$, so that $\tilde{H}$ ($\tilde{L}$) is expressed in terms of equivalents of primary (secondary) educated, male and less than 5 years experienced workers (I refer to these groups as "baseline" skilled and unskilled workers). I start from a specification where $\beta_j$, $\lambda_g$ and $e^{\mu_{exp}}$ do not vary across countries, but I also consider extensions where they do.

[1] While the cross-country data in Barro and Lee (2013) allows me to distinguish also between workers with no education, some primary and primary completed, I chose this broader aggregation for consistency with the other data sources of the paper. In particular, the US Census does not fully distinguish between some primary and primary completed. Moreover, sample sizes are small at these levels of educational attainment, especially by immigrants' countries of origin.

[2] With a slight abuse of notation, I denote by $j \in H$ ($j \in L$) the educational attainment levels assumed to be high- (low-) skill.

[3] In order to measure (potential) experience, I use data from the World Development Indicators on the country-specific statutory duration of each educational stage, as well as on the schooling starting age. For countries lacking this information, I use (World Bank defined) region-specific averages. Incomplete education spells are assigned half of the statutory duration, while for tertiary education (not covered in the WDI data) I use a duration of 4 years for all countries.

I then use the assumption of perfectly competitive labor markets to estimate $\beta_j$, $\lambda_g$ and $\mu_{exp}$. In particular, perfect competition implies that the average log wage of a worker of skill $S \in \{H, L\}$, with educational attainment $j$, gender $g$ and experience $exp$ is:

$$\log w_{S(j,g,exp)} = \alpha + \gamma_S + \beta_j + \lambda_g + \mu_{exp} \tag{6}$$

where $\alpha$ is a constant and $\gamma_S = \log (A_S Q_S)^{\rho} (\tilde{S})^{\rho-1}$. The parameters $\beta_j$, $\lambda_g$ and $\mu_{exp}$ can be therefore identified from a regression of log wages on skill group, educational attainment, gender and experience fixed effects.[4] Moreover, the coefficient on $\gamma_H$ (with low-skilled workers being the omitted cateogory) identifies the log skill premium, i.e. the log wage differential between baseline skilled and unskilled workers. I run this specification using data from the US Census, focusing on a sample of native individuals between 15 and 64 years old with a relatively high degree of labor market attachment.[5] Table I shows the estimated coefficients. Baseline skilled workers earn approximately 31% more than baseline unskilled workers. Within skill groups, wages rise steeply with educational attainment and experience, while females face a conditional wage gap of about 25%.

With the estimates of $\beta_j$, $\lambda_g$ and $\mu_{exp}$ at hand I can compute $\tilde{H}$ and $\tilde{L}$ for all countries, so that the only missing part to back out $\frac{A_H Q_H}{A_L Q_L}$ from (3) is the skill premium $\frac{w_H}{w_L}$. To the best of my knowledge, no existing dataset provides a measure of the skill premium which is comparable across countries, nationally representative and consistent with the skill categorization used in this paper and the rest of the literature. Sources like ILOSTAT, compiled by the International Labor Organization, allow to contruct, for a limited number of countries, wage gaps between workers in different occupations or economic activities (as opposed to different educational attainments). This is problematic, as occupations and their skill content are difficult to compare across countries at different stages of development. Moreover, these data do not allow to condition in a systematic way on experience, age and other observable characteristics.

To overcome this problem, I adopt two complementary approaches. First, I follow much of the literature (Caselli and Coleman, 2006; Caselli, 2016) in inferring the skill premium from existing data on educational attainment and Mincerian returns to education. Second, I estimate directly the skill premium using micro-data for a smaller sample of countries.

For the first approach, I rely on a dataset I constructed in previous work with Francesco Caselli and Jacopo Ponticelli (Caselli et al., 2016), which includes a collection of estimated

[4] Individual-level heterogeneity, resulting in an error in term in 6, can be easily added to the model; see Section 4. Of course, this specification might fail to capture causal effects, as several relevant unobservables are likely to be correlated with the regressors. The literature on returns to schooling, however, finds that OLS and IV estimates are often close in magnitude (Card, 2001).

[5] I restrict the sample to individuals that report working for wages, for at least 30 weeks and 30 hours per week in the previous year.

Mincerian returns for a large set of countries.[6] As shown in Caselli (2016), there is a direct mapping between the return estimated in a regression of low wages on years of schooling and the skill premium. In particular, if $b$ is the Mincerian return, then

$$\log \frac{w_H}{w_L} = \frac{b \sum_j (s_j - \bar{s})^2 n_j + \sum_j \sum_g \sum_{exp} (\beta_j + \lambda_g + \mu_{exp})(s_j - \bar{s}) n_{j,g,exp}}{\sum_{j \in H} (s_j - \bar{s}) n_j} \tag{7}$$

where $n_j = \sum_g \sum_{exp} n_{j,g,exp}$ and $s_j$ are the number and the completed years of schooling of workers with educational attainment $j$, while $\bar{s}$ is the average of attained years of schooling in the population. I compute $\log \frac{w_H}{w_L}$ from this expression using the estimates for $\beta_j$, $\lambda_g$ and $\mu_{exp}$ illustrated above, data from Barro and Lee (2013) on educational attainment by gender and experience and from the World Bank's World Development Indicators on the duration of each education stage (which I use to infer $s_j$ and $\bar{s}$). This approach has obvious limitations, as it imposes assumptions on the human capital aggregators and it is based on estimates (the Mincerian coefficients) not in all cases coming from nationally representative samples. On the other hand, this methodology has the advantage of wide applicability: it allows me to construct the skill premium for 80 countries, spanning the whole income distribution. In what follows, I refer to this group of countries as the "broad" sample.

The second approach, which is a novel contribution of this paper, improves the reliability and comparability dimensions of the estimates at the cost of a smaller sample size. Using IPUMS, I estimate the skill premium using micro data for all countries with available information on wages or earnings, education, labor market status, gender and experience. This is possible for 12 countries in 2000 or a close year, including (according to the World Bank classification) high-income (United States, Canada, Israel, Trininidad and Tobago), upper middle-income (Mexico, Panama, Uruguay, Venezuela, Brazil, Jamaica) and lower middle-income (Indonesia, India) countries. I refer to this set of countries as the "narrow sample". For all of them, I run a specification equivalent to (6), imposing the same restriction as in the US sample whenever possible.[7] This also allows me to drop the assumption that $\beta_j$, $\lambda_g$ and $\mu_{exp}$ do not vary across countries.[8]

[6]The dataset contains Mincerian returns estimated on a large set of published and unpublished academic works. It contains up to two estimates for each country, one relative to the 1990s and one to the 2000s. Here I use the observation from the 2000s when possible, and the one from the 1990s otherwise.

[7]The information on weeks worked in the previous year is available only in the US, and therefore I cannot restrict the sample to workers with more than 30 weeks worked for the other countries. With the exception of India, Panama and Uruguay, which provide no information on labor supply, I can still restrict attention to individuals working more than 30 hours per week in the other countries. For Brazil, Mexico and Venezuela I use total earnings as wages are unavailable.

[8]In two countries, Israel and Jamaica, the information in IPUMS does not allow to identify individuals with some (not complete) secondary education, making it impossible to estimate the return to this level of educational attainment. As in Barro and Lee (2013)'s data the share of individuals with incomplete secondary education is positive, I impute their return interpolating the returns to primary and secondary education, using the returns to primary, some secondary and secondary education in the other 10 countries to construct the

## 2.3 Results

In this section I show how the relative efficiency of skilled labor varies across countries. I discuss and compare results on both the broad sample, where wage premia are inferred from Mincerian coefficients, and the narrow sample, where all quantities are directly estimated from micro-level data.

### 2.3.1 Broad Sample

Table II summarizes the variation of the skill premium, the relative supply of skilled and unskilled workers and the relative efficiency of skilled labor across the 80 countries in the broad sample. The wage ratio between skilled and unskilled workers varies substantially across countries, ranging from values close to or (in one case, Ukraine) below 1 to 11.6 in Jamaica. As expected, the correlation between the skill premium and relative skill supply (as well as GDP) is negative.

On the other hand, the relative supply of skilled labor is higher in rich countries. In the United States there are 12.71 as many baseline equivalent high-skilled workers as unskilled ones, while the corresponding number for Cambodia, the poorest country in the sample, is 0.04. This is partially because within the high-skill group US workers are relatively more concentrated in higher levels of educational attainment, blowing up the corresponding amount of baseline worker equivalents, and to a large extent because there are many more skilled workers in the US (the ratio between the numbers of skilled and unskilled workers, not converted to baseline equivalents, is 9.89 in the US and 0.03 in Cambodia).

The last row of Table II shows that there are large cross-country differences in the relative efficiency of skilled labor. The efficiency bias is 4 times larger in the US compared to the average country in the sample. Figure I displays in log scale the strong positive relationship between relative efficiencies and relative supplies: skilled labor is relatively more efficient where it is relatively more abundant.

The result is driven by the fact that the relationship between the skill premium and the relative supply of skilled workers is not steep enough, so that a high efficiency of skilled labor in skill abundant countries is needed to fit the data. Figure II illustrates this point by plotting the log skill premium against the log relative supply. The dashed line has a slope of $\rho - 1 = -0.67$, which is the predicted slope of this relationship in a world where $\log \frac{A_H Q_H}{A_L Q_L}$ was constant across countries (or, more generally, uncorrelated with $\log \frac{\tilde{H}}{\tilde{L}}$). The best linear fit (solid line) has instead an estimated slope of -0.17, with a standard error of 0.04. This implies that $\log \frac{A_H Q_H}{A_L Q_L}$ must increase with $\log \frac{\tilde{H}}{\tilde{L}}$.

weights.

### 2.3.2 Narrow Sample

Table III displays the skill premia, skill relative supplies and relative efficiencies for the countries in the narrow sample. As before, the skill premium is on average lower in countries with higher supply of skilled labor, but the range of its variation is relatively modest. Coupled with the large gaps in relative human capital displayed in the second column, this implies once again large cross-country differences in the relative efficiency of skilled labor. The magnitudes are similar to the ones for the broad sample: the average country has about one quarter of the skill efficiency bias compared to the US. Figure III shows that the relative efficiency skilled labor is strongly positively related to its relative supply.

Figure IV illustrates the mechanics behind the result: the log skill premium is remarkably flat across countries (the slope is -0.10), in spite of the important variation in the log relative supply. To give an example, in a world where all countries had the US level of efficiency bias, the model would predict for Indonesia a wage ratio of 26, while the actual ratio is 2.12.

Taking stock, the analysis of micro data for a number of countries at different levels of development supports the existence of large gaps in skill efficiency, with richer and more skill abundant countries having relatively more efficient skilled labor. This pattern, both qualitatively and quantitatively, does not appear to be an artifact of differences in the data sources or in the measurement of the skill premium. This leads naturally to the next question: what explains the variation in relative skill efficiency across countries?

# 3 Sources of Differences in Relative Skill Efficiency

In this section I investigate the role of the skill bias of technology and the quality of skilled labor in explaining the cross-country variation in skill efficiency documented above. My strategy is based on the analysis of immigrants educated in different countries and observed in the same labor market. I first modify the baseline framework to include a specific role for workers' country of origin. I then map the new framework to the data and discuss the emerging patterns.

## 3.1 A Modified Framework

I introduce a new dimension of workers' heterogeneity to the framework in Section 2.1: the fact that some of them are educated in different countries. For clarity, I abstract from educational careers spanning more than one country, and I consider only natives and migrants entirely educated in their own country of origin.

I assume that skilled and unskilled workers' embodied human capital depends on the country where their education was acquired (indexed by $a$). This might reflect the combined impact of several characteristics of the educational environment, but also the mechanisms according to which individuals with different baseline characteristics sort into different levels of educational attainment. I do not wish (or need) to take a stand on the source of embodied productivity differences between skilled and unkilled labor, which might also be different across countries. I take as given their (possible) existence, and attempt to measure them in the data.

Within skill groups, services provided by different immigrant groups are perfect substitutes. The total quantities of high- and low-skill services used for production in country $c$ are

$$H_c = \sum_a Q^a_{Hc} \tilde{H}^a_c \tag{8}$$

$$L_c = \sum_a Q^a_{Lc} \tilde{L}^a_c \tag{9}$$

where $\tilde{H}^a_c$ and $\tilde{L}^a_c$ are the number of (baseline equivalent) skilled and unskilled workers educated in country $a$ and working in $c$, and $Q^a_{Hc}$ and $Q^a_{Lc}$ represent their average quality.

In a competitive labor market, the wage ratio between skilled and unskilled workers educated in a generic country $b$ is

$$\frac{w^b_{Hc}}{w^b_{Lc}} = \left(\frac{A_{Hc}Q^b_{Hc}}{A_{Lc}Q^b_{Lc}}\right)^{\rho} \left(\frac{\sum_a (Q^a_{Hc}/Q^b_{Hc})\tilde{H}^a_c}{\sum_a (Q^a_{Lc}/Q^b_{Lc})\tilde{L}^a_c}\right)^{\rho-1} \tag{10}$$

This expression summarizes the key source of variation for my empirical strategy. Immigrant groups educated in their home countries face similar labor market conditions, both in terms of the degree of technological skill bias ($\frac{A_{Hc}}{A_{Lc}}$) and of the general equilibrium effect of aggregate skill supply (the second term of (10)), but are endowed with different $Q$'s depending on their country of origin. Under some additional assumptions, by comparing skill premia across origin countries one can isolate cross-country differences in the relative quality of skilled and unskilled labor.

## 3.2 Measurement

In this section I describe how I map this framework to the data. The objective is to separately identify $\frac{A_{Hc}}{A_{Lc}}$ and $\frac{Q^c_{Hc}}{Q^c_{Lc}}$, in order to study the variability of both across countries. I normalize $\frac{A_{Hc}}{A_{Lc}}$ and $\frac{Q^c_{Hc}}{Q^c_{Lc}}$ so that they are 1 for the US.

I focus on the native and foreign-born workers living in the United States, observed in the 2000 Census. I once again restrict attention to workers between 15 and 64 years old, who have been working for wages for at least 30 weeks and 30 hours per week in the previous year. To isolate the role of education in the origin country, I only consider immigrants which are likely to have completed their education before relocating to the US: as in Schoellman (2012), I restrict the sample to those who migrated at least six years after the age at which they should have ended their studies, given their level of educational attainment. Moreover, when working with the broad sample, I focus my cross-country analysis on the 41 countries for which I observe at least 100 immigrants (satisfying the sample restrictions mentioned above) in each skill group, plus the United States.

As before, I consider human capital aggregators that take into account the heterogeneity in education, gender and experience,

$$\begin{aligned} \tilde{H}_c^a &= \sum_{j \in H} \sum_{g} \sum_{exp} e^{\beta_j} e^{\lambda_g} e^{\mu_{exp}} n^a_{c(j,g,exp)} \\ \tilde{L}_c^a &= \sum_{j \in L} \sum_{g} \sum_{exp} e^{\beta_j} e^{\lambda_g} e^{\mu_{exp}} n^a_{c(j,g,exp)} \end{aligned} \tag{11}$$

where $n^a_{c(j,g,exp)}$ is the number of workers in group $(j, g, exp)$ educated in country $a$. The average log wage of a worker educated in $a$, of skill $S \in \{H, L\}$, with educational attainment $j$, gender $g$ and experience $exp$ is:

$$\log w^a_{Sc(j,g,exp)} = \alpha_c + \gamma_{Sc} + \log Q^a_{Sc} + \beta_j + \lambda_g + \mu_{exp} \tag{12}$$

where $\alpha_c$ is a constant and $\gamma_{Sc} = \log (A_{Sc})^{\rho} (\sum_a Q^a_{Sc}\tilde{S}_c)^{\rho-1}$. In a specification including skill group fixed effects, the interaction terms between skill group and country of origin fixed effects (with US natives as omitted category) identify $\log Q^a_{S,US} - \log Q^{US}_{S,US}$ for $S \in \{H, L\}$, from which $\log \frac{Q^a_{H,US}}{Q^a_{L,US}}$ can be calculated (recall that $\log \frac{Q^{US}_{H,US}}{Q^{US}_{L,US}}$ is normalized to 1). Moreover, $\beta_j$, $\lambda_g$ and $\mu_{exp}$ are identified from the coefficients on educational attainment, gender and experience fixed effects.

For robustness, I also consider alternatives where the returns to (within skill groups) education and experience is country of origin-specific.[9] Given that sample sizes are small for some education-experience-country of origin cells, I use for this purpose a less flexible specification with linear (within skill groups) returns to years of schooling and quadratic

[9] Bratsberg (2002) and Schoellman (2012) document differences in country of origin-specific Mincerian returns for US immigrants, while Lagakos et al. (2016) argue for country-specific returns to experience. Note that the heterogeneity of the relative quality of skilled and unskilled labor already implies heterogeneous Mincerian returns. In future work, I plan to examine more systematically the extent to which heterogeneous returns to schooling are driven by differences within as opposed to between skill groups.

returns to experience,

$$\begin{aligned}\tilde{H}_c^a &= \sum_{j\in H}\sum_g\sum_{exp} e^{\beta_H^a(s_j^a - s_{\text{sec}}^a)} e^{\lambda_g} e^{(\mu_1^a \eta_{exp} + \mu_2^a \eta_{\text{exp}}^2)} n_{c(j,g,exp)}^a \\ \tilde{L}_c^a &= \sum_{j\in L}\sum_g\sum_{exp} e^{\beta_L^a(s_j^a - s_{\text{pri}}^a)} e^{\lambda_g} e^{(\mu_1^a \eta_{exp} + \mu_2^a \eta_{exp}^2)} n_{c(j,g,exp)}^a\end{aligned} \tag{13}$$

where $\beta_H^a$, $\beta_L^a$, $\mu_1^a$, $\mu_2^a$ are the (country of origin-specific) returns to years of schooling (potentially different for skilled and unskilled workers), experience and experience squared, $s_j^a$ is the number of years of schooling of workers with educational attainment $j$ (achieved in country $a$) and $\eta_{exp}$ is years of experience for group $exp$.[10] The units are still baseline category equivalents, and the baseline categories are males, primary educated or less ($s_{\text{pri}}^a$ years of schooling) with no experience for unskilled workers and males, secondary educated ($s_{\text{sec}}^a$ years of schooling) with no experience for skilled workers. Similarly to (12), a regression of log wages on skill group fixed effects, years of schooling, experience, experience squared (all interacted by country of origin fixed effects) and gender identifies $\log \frac{Q_{Hc}^a}{Q_{Lc}^a}$, $\beta_H^a$, $\beta_L^a$, $\mu_1^a$, $\mu_2^a$ for all countries of origin. In what follows, I refer to (13) and close variations as the "parametric" specifications.

Under the assumption that the relative quality of skilled workers among US immigrants captures the relative quality among natives in the country origin, that is $\log \frac{Q_{H,US}^a}{Q_{L,US}^a} = \log \frac{Q_{Ha}^a}{Q_{La}^a}$, I can examine the cross-country variation in the latter. The main question of interest is the role of relative skill quality in explaining differences in relative skill efficiency. Given that workers' quality is assumed to be heterogeneous depending of the country in which they were educated, in principle one should take into account the educational composition of the population in each country when computing relative skill quantities and backing out relative efficiencies. However, if immigrants educated abroad are a sufficiently small share of the working population, the relative supply, quality and price of skills among native workers are good approximations for the corresponding population-wide quantities.[11] I rely on this approximation and compute for each country $\frac{A_{Hc}}{A_{Lc}}$ from (3), using estimates for the relative

[10] I experimented with higher degree polynomials obtaining very similar results.

[11] More precisely, the population-wide skill premium is given by

$$\frac{w_{Hc}}{w_{Lc}} = \left(\frac{A_{Hc}Q_{Hc}}{A_{Lc}Q_{Lc}}\right)^{\rho} \left(\frac{\sum_a (Q_{Hc}^a/Q_{Hc})\tilde{H}_c^a}{\sum_a (Q_{Lc}^a/Q_{Lc})\tilde{L}_c^a}\right)^{\rho-1}$$

where, for $S \in \{H, L\}$, $w_{Sc} = \sum_a w_{Sc}^a \frac{n_{Sc}^a}{n_{Sc}}$, $Q_{Sc} = \sum_a Q_{Sc}^a \frac{n_{Sc}^a}{n_{Sc}}$ and $n_{Sc}$ ($n_{Sc}^a$) is the number of workers of skill $S$ in the population (educated in country $a$). If $n_{Sc}^c \approx n_{Sc}$ for $S \in \{H, L\}$, then clearly $\frac{w_{Hc}}{w_{Lc}} \approx \frac{w_{Hc}^c}{w_{Lc}^c}$, $\frac{Q_{Hc}}{Q_{Lc}} \approx \frac{Q_{Hc}^c}{Q_{Lc}^c}$ and $\frac{\sum_a (Q_{Hc}^a/Q_{Hc})\tilde{H}_c^a}{\sum_a (Q_{Lc}^a/Q_{Lc})\tilde{L}_c^a} \approx \frac{\tilde{H}_c^c}{\tilde{L}_c^c}$.

skill quality among native workers.[12,13] From now on, I simply refer to these objects as $\frac{A_H}{A_L}$ and $\frac{Q_H}{Q_L}$.

To summarize the empirical strategy, I start from a difference-in-differences approach, where I compare, within the United States, the log wages of skilled and unskilled workers between the different countries where they were educated. I then examine whether skill premia are larger for countries of origin with a higher measured relative efficiency of skilled labor, and draw the implications for the cross-country dispersion in the latter.

## 3.3 Results

In this section I show how the relative skill bias of technology and quality of skilled labor vary across countries. I start from results relative to the broad sample and then consider the narrow one.

### 3.3.1 Broad Sample

Table IV provides summary statistics on the relative skill efficiency, skill bias of technology and skill quality in the countries with at least 100 immigrant workers per skill group in the regression sample (and the United States). The first row shows that these countries display similar patterns in terms of skill efficiency as the rest of the broad sample (compare with the third row of Table II). The second and third columns show that the skill bias of technology is substantially more dispersed across countries compared to relative skill quality. In the average country, the skill bias of technology is 31% of the US level, while relative skill quality is 88%. In Cambodia, a country with 27 times as many unskilled as skilled workers, the relative skill quality is 83% compared to the US.

Figure V illustrates how the skill bias of technology and relative skill quality vary as a function of skill supply (on a log scale). Countries with high relative supply of skilled workers have technologies more biased towards them (left panel), while the relationship with relative skill quality is almost flat (right panel).

As an alternative way to summarize these patterns, I consider the relative role of the two

[12]The Barro and Lee (2013)'s data, used to compute human capital stocks, refer to the whole population (natives and immigrants). The skill premium estimated from IPUMS data (for the countries in the narrow sample) is relative to native workers only (though including immigrants has a negligible impact on the resulting estimates).

[13]In principle, using data on the stock of migrants by country of origin, one could make some progress towards quantifying the importance of differences in the ethnic composition of the population. This approach would require, across different host countries, information on the composition by education and age of arrival of the stock of migrants from each country of origin, and assumptions on the quality of individuals whose educational career spans more than one country. Given the substantial data requirements, the additional structure that this would involve and the fact the immigrants educated abroad are a small share of the population in most countries, I chose not to follow this route.

components of (log) skill efficiency in explaining its variance across countries and covariance with other variables of interest. In particular, I compute

$$V_A = \frac{V(\log \frac{A_H}{A_L})}{V(\log \frac{A_H Q_H}{A_L Q_L})} \quad , \quad V_Q = \frac{V(\log \frac{Q_H}{Q_L})}{V(\log \frac{A_H Q_H}{A_L Q_L})}$$

which give me the shares of the variance of log relative skill efficiency accounted for by technology skill bias and relative skill quality, and

$$Cov_A\left(\frac{\tilde{H}}{\tilde{L}}\right) = \frac{Cov(\log \frac{A_H}{A_L}, \log \frac{\tilde{H}}{\tilde{L}})}{Cov(\log \frac{A_H Q_H}{A_L Q_L}, \log \frac{\tilde{H}}{\tilde{L}})} \quad , \quad Cov_A\left(y\right) = \frac{Cov(\log \frac{A_H}{A_L}, \log y)}{Cov(\log \frac{A_H Q_H}{A_L Q_L}, \log y)}$$
$$Cov_Q\left(\frac{\tilde{H}}{\tilde{L}}\right) = \frac{Cov(\log \frac{Q_H}{Q_L}, \log \frac{\tilde{H}}{\tilde{L}})}{Cov(\log \frac{A_H Q_H}{A_L Q_L}, \log \frac{\tilde{H}}{\tilde{L}})} \quad , \quad Cov_Q\left(y\right) = \frac{Cov(\log \frac{Q_H}{Q_L}, \log y)}{Cov(\log \frac{A_H Q_H}{A_L Q_L}, \log y)}$$

which represent the shares of the covariance between log relative skill efficiency and log relative skill supply on one hand and log GDP per worker on the other driven by $\log \frac{A_H}{A_L}$ and $\log \frac{Q_H}{Q_L}$.

Table V shows the results. Starting from the baseline specification (first column), the technology term accounts for the vast majority of both the variance of log skill efficiency (92%) and of its covariance with log skill supply (95%) and log GDP (81%). The covariance between $\log \frac{A_H}{A_L}$ and $\log \frac{Q_H}{Q_L}$ is small and positive (since $V_A$ and $V_Q$ almost sum to 1). The second column considers the parametric specification of human capital stocks, with linear returns to schooling and quadratic to experience; reassuringly, this modification does not affect the results. Then, the third column uses the parametric specification with country-specific returns to education and experience, as in (13).[14] The contribution of relative skill quality increases slightly, but overall the technology term still dominates.[15]

### 3.3.2 Narrow Sample

I now turn to the corresponding results for the narrow sample. Table VI shows the patterns for skill efficiency, technology bias and skill quality. As for the broad sample, in terms of cross-country variation the technology term dwarfs the quality one. The average country displays 30% of the technology bias of the United States, and 92% of its relative skill quality. Figure VI shows that more generally the relative skill bias of technology is strongly increasing in the relative supply of skilled workers, while the relative quality term is not.

[14] Here, when computing the log skill premium I adjust the epression in (7) to be consistent with the different formulation of human capital stocks. The impact of this adjustment is negligible.

[15] While, within skill groups, returns to schooling (and, to some extent) experience are generally higher in richer countries, this raises both $\tilde{H}$ and $\tilde{L}$ having a limited impact on the relative skill supply.

An example is instructive to understand how the magnitude of the skill premium within immigrant groups is driving the result. Mexico's relative skill efficiency is 17% of the US level. If this gap was entirely reflecting higher relative skill quality among US workers, the skill premium among Mexican educated workers in the US should be 0.77 log points smaller than the one among US natives (corresponding to a wage ratio of 63%), while in reality the difference is 0.19 log points (wage ratio of 112%).

# 4 Selection

Given that my strategy consists in using immigrant workers to estimate country-specific differences in the relative quality of skilled labor, a natural concern is that emigrant workers are not randomly selected. In this section I discuss the possible consequences of selection and discuss some evidence on its importance.

It is helpful to explicitly introduce some individual-level heterogeneity in the framework of section 3.1 to illustrate the main issues. Suppose that the quality of individual $i$, of skill $S \in \{H, L\}$, having completed his education in country $a$ is $Q_S^a \varepsilon_{S,i}^a$, where $Q_S^a$ is a term common to all individuals of skill $S$ educated in $a$ and $\varepsilon_{S,i}^a$ captures the heterogeneity in unobservable skills. For analytical convenience, I assume that $\varepsilon_{S,i}^a$ follows a log-normal distribution with log-mean 0 and log-variance $(\sigma^a)^2$. Moreover, I mantain the assumption that $\varepsilon_{S,i}^a$ is uncorrelated with workers' observable characteristics (education, gender and experience).[16]

If migrants are selected on unobservable skills, $\mathbb{E}\left[\log \varepsilon_{S,i}^a | \text{migrant}\right] \neq 0$. The relative log skill quality I estimate out of US migrants using (11) would then read

$$\log Q_{H,US}^a - \log Q_{L,US}^a = \log Q_H^a - \log Q_L^a + \mathbb{E}\left[\log \varepsilon_{H,i}^a | \text{migrant}\right] - \mathbb{E}\left[\log \varepsilon_{L,i}^a | \text{migrant}\right]$$

which differs from the quantity of interest as long as $\mathbb{E}\left[\log \varepsilon_{H,i}^a | \text{migrant}\right] \neq \mathbb{E}\left[\log \varepsilon_{L,i}^a | \text{migrant}\right]$. Migrants' selection is therefore problematic to the extent that it takes place with a different degree across skill groups.

To the best of my knowledge, the migration literature offers little direct guidance on whether this form of differential selection might be relevant in practice. It has been widely established that migrants are non-randomly selected on observable and unobservable skills (Borjas, 1987), and for the vast majority of origin countries the degree of selection of emigrants to the United States appears to be positive (Feliciano, 2005). However, less is known about the relative selection by educational achievement, i.e. on how, among individuals ed-

[16]This is obviously a strong assumption, though common in the development accounting literature. See footnote 4 for a related discussion.

ucated in a given country, the degree of selection on unobservables of migrants within the low-skill group compares to the one within the high-skill group.

Since my main result is that, for most countries, the log relative quality of skilled labor inferred out of migrants is too large to account for the international gaps in skill efficiency, investigating the possibility that $\mathbb{E}\left[\log \varepsilon^a_{H,i}|\text{migrant}\right] > \mathbb{E}\left[\log \varepsilon^a_{L,i}|\text{migrant}\right]$ is of particular interest. A more positive degree of selection across skilled workers could in principle lead me to understate the importance of relative skill quality differences across countries.

I discuss one piece of evidence suggesting that this is unlikely to be a major concern: the propensity to migrate conditional on skill group is much higher for the high-skilled than for the low-skilled. Figure VII plots the share of skilled workers among US emigrants for each country of origin against the share of skilled workers in the country of origin population. Almost all countries are above the 45 degree line, showing that emigrants are substantially more likely to be high-skilled.[17] If, as suggested by the literature, migrants' selection on unobservables (conditional on observables) is positive, such a pattern would imply that skilled migrants are relatively more negatively selected. To see why, suppose there exists a skill-specific threshold $t^a_S$ such that workers migrate if $Q^a_S \varepsilon^a_{S,i} \geqslant t^a_S$. The within-skill group share of emigrants is then $1 - \Phi\left(\frac{\log t^a_S - \log Q^a_S}{\sigma^a}\right)$, where $\Phi(.)$ is the standard Normal's cumulative distribution function. The fact that $1 - \Phi\left(\frac{\log t^a_H - \log Q^a_H}{\sigma^a}\right) > 1 - \Phi\left(\frac{\log t^a_L - \log Q^a_L}{\sigma^a}\right)$, implied by Figure VII, means that $\log t^a_H - \log Q^a_H < \log t^a_L - \log Q^a_L$. It follows that

$$\mathbb{E}\left[\log \varepsilon^a_{H,i}|\text{migrant}\right] = \frac{\sigma^a \phi\left(\frac{\log t^a_H - \log Q^a_H}{\sigma^a}\right)}{1 - \Phi\left(\frac{\log t^a_H - \log Q^a_H}{\sigma^a}\right)} < \mathbb{E}\left[\log \varepsilon^a_{L,i}|\text{migrant}\right] = \frac{\sigma^a \phi\left(\frac{\log t^a_L - \log Q^a_L}{\sigma^a}\right)}{1 - \Phi\left(\frac{\log t^a_L - \log Q^a_L}{\sigma^a}\right)}$$

Intuitively, if migrating is relatively easier for high-skilled individuals, the low-skilled ones who do migrate should be comparatively better selected on unobservable skills.

I conclude this section by noting that the threat of selection is less severe when comparing countries other than the United States. This is because, while US estimates are based on native workers, for all other countries the relative skill quality is inferred from immigrants only. Cross-country gaps in these estimates capture actual gaps in relative skill quality as long as there is not a pattern of *differential relative selection* of skilled and unskilled migrants across countries, while a common degree of relative selection in unconsequential. As it is evident from Figures V and Figure VI, the result that technology skill bias (as opposed to relative skill quality) is the key factor varying between skill-poor and skill-abundant countries is not driven by the United States only.

[17]Similar patterns hold when expressing units in terms of baseline equivalent workers as opposed to counting persons.

# 5 Conclusions

In this paper I re-visit the question of how the relative productivity of skilled and unskilled labor differs across countries. I show that, according to various sources, the skill premium varies little across countries, implying large gaps in relative skill efficiency. In the second part of the paper, I show that skill premia within immigrant groups are not consistent with the view that differences in relative skill quality play a quantitatively important role. The variation in relative skill efficiency is instead more likely to be related to technological factors.

These results have important implications when considering the relative role of human capital and technology in accounting for cross-country differences in output per worker. Malmberg (2017) suggests that gaps in skill efficiency are an important component of differences in economic performance. My findings imply that one should be careful in attributing these gains to human capital. Moreover, if we accept the view that the factor-bias of adopted technologies is very different between rich and poor country, this gives credit to the possibility that rich countries' technologies might not be appropriate for firms in poor countries, as argued by Acemoglu and Zilibotti (2001).

Indeed, my results emphasize the importance of understanding the determinants of technological skill bias. A common view is that differences in the technology mix reflect the optimal responses of firms to the abundance or scarcity of skilled labor (Caselli, 2016). It would be useful to have a sense of the quantitative importance of this mechanism, and of whether other institutional, cultural or geographical factors might contribute to explain why poorer countries adopt less skill-biased technologies.

The approach of this paper can be extended in various directions. For example, it would be interesting to explore the relative role of technology and human capital in explaining the differential evolution of the skill premium over time in the United States and Europe. Moreover, a similar exercize could be performed within countries, in order to explore how relative skill quality vary across regions with different characteristics. I hope to address some of these open issues in future work.

# References


**Acemoglu, Daron**, "Why Do New Technologies Complement Skills? Directed Technical Change And Wage Inequality," *The Quarterly Journal of Economics*, November 1998, *113* (4), 1055–1089.

_ , "Directed Technical Change," *Review of Economic Studies*, October 2002, *69* (4), 781–809.

_ **and Fabrizio Zilibotti**, "Productivity Differences," *The Quarterly Journal of Economics*, May 2001, *116* (2), 563–606.

**Barro, Robert J. and Jong Wha Lee**, "A New Data Set of Educational Attainment in the World, 1950-2010," *Journal of Development Economics*, 2013, *104* (C), 184–198.

**Borjas, George J.**, "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 1987, *77* (4), 531–553.

**Bowlus, Audra J. and Chris Robinson**, "Human Capital Prices, Productivity, and Growth," *American Economic Review*, December 2012, *102* (7), 3483–3515.

**Bratsberg, Bernt**, "School Quality and Returns to Education of U.S. Immigrants," *Economic Inquiry*, April 2002, *40* (2), 177–198.

**Card, David**, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, September 2001, *69* (5), 1127–1160.

**Caselli, Francesco**, *Technology Differences over Space and Time*, Princeton University Press, 2016.

_ **and Wilbur John Coleman**, "The World Technology Frontier," *American Economic Review*, June 2006, *96* (3), 499–522.

_ **, Jacopo Ponticelli, and Federico Rossi**, *A New Data Set on Mincerian Returns* Technology Differences across Space and Time, Princeton University Press, 2016.

**Ciccone, Antonio and Giovanni Peri**, "Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950-1990," *The Review of Economics and Statistics*, November 2005, *87* (4), 652–663.

**Feliciano, Cynthia**, "Educational selectivity in U.S. Immigration: How do immigrants compare to those left behind?," *Demography*, 2005, *42* (1), 131–152.

**Hall, Robert E. and Charles I. Jones**, "Why Do Some Countries Produce So Much More Output per Worker than Others?," *The Quarterly Journal of Economics*, February 1999, *114* (1), 83–116.

**Hendricks, Lutz and Todd Schoellman**, "Student abilities during the expansion of US education," *Journal of Monetary Economics*, 2014, *63* (C), 19–36.

**Jones, Benjamin F.**, "The Human Capital Stock: A Generalized Approach," *American Economic Review*, November 2014, *104* (11), 3752–3777.

**Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman**, "Life-Cycle Human Capital Accumulation Across Countries: Lessons From U.S. Immigrants," NBER Working Papers 21914, National Bureau of Economic Research January 2016.

**Malmberg, Hannes**, "Human Capital and Development Accounting Revisited," Working Paper 2017.

**Schoellman, Todd**, "Education Quality and Development Accounting," *Review of Economic Studies*, 2012, *79* (1), 388–417.

_ , "Early Childhood Human Capital and Development," *American Economic Journal: Macroeconomics*, July 2016, *8* (3), 145–174.

# Tables

Table I: Returns to Skill, Education, Gender and Experience - US Census

| Skill | | Edu | | Gender | | Experience | |
|---|---|---|---|---|---|---|---|
| Group | Coeff | Group | Coeff | Group | Coeff | Group | Coeff |
| Low Skill | 0 | Primary or less | 0 | Male | 0 | 0 to 4 | 0 |
| High Skill | 0.305<br>(0.005) | Some Secondary | 0.111<br>(0.005) | Female | -0.251<br>(0.001) | 5 to 9 | 0.289<br>(0.001) |
| | | Secondary | 0 | | | 10 to 14 | 0.451<br>(0.001) |
| | | Some Tertiary | 0.178<br>(0.001) | | | 15 to 19 | 0.540<br>(0.001) |
| | | Tertiary | 0.615<br>(0.001) | | | 20 to24 | 0.586<br>(0.001) |
| | | | | | | 25 to 29 | 0.610<br>(0.001) |
| | | | | | | 30 to 34 | 0.639<br>(0.001) |
| | | | | | | 35 to 39 | 0.648<br>(0.001) |
| | | | | | | 40+ | 0.636<br>(0.002) |

*Notes*: The Table shows the estimated returns to skill, education, gender and experience in the 2000 US Census. The dependent variable is log wage per hour. The sample includes 4282320 observations. The coefficients on low skill, primary education or less, secondary education, male and 0 to 4 experience are normalised to 0. Observations are weighted according to the provided sample weights. Robust standard errors are shown in parentheses.

Table II: Skill Premium, Supply and Efficiency across Countries - Broad Sample

| | N | Mean | Min | Max | Std. Dev. | Corr w/ $\tilde{H}/\tilde{L}$ | Corr w/ $y$ |
|---|---|---|---|---|---|---|---|
| $w_H/w_L$ | 80 | 2.30 | 0.81 | 11.55 | 1.43 | -0.27 | -0.35 |
| $\tilde{H}/\tilde{L}$ | 80 | 1.36 | 0.02 | 12.71 | 1.80 | 1 | 0.40 |
| $(A_H Q_H)/(A_L Q_L)$ | 80 | 0.26 | 0.02 | 1.25 | 0.20 | 0.63 | 0.26 |

*Notes*: The Table shows summary statistics for the skill premium, relative skill supply and efficiency across the countries in the broad sample. Relative skill efficiency is normalised such that it takes value 1 for the United States.

Table III: Skill Premium, Supply and Efficiency across Countries - Narrow Sample

| Country | $w_H/w_L$ | $\tilde{H}/\tilde{L}$ | $(A_H Q_H)/$ $(A_L Q_L)$ |
|---|---|---|---|
| Indonesia | 2.12 | 0.15 | 0.08 |
| Brazil | 2.22 | 0.33 | 0.14 |
| Venezuela | 1.24 | 0.38 | 0.09 |
| Uruguay | 1.81 | 0.41 | 0.14 |
| India | 2.57 | 0.42 | 0.19 |
| Mexico | 2.09 | 0.45 | 0.17 |
| Trinidad and Tobago | 1.68 | 0.48 | 0.14 |
| Jamaica | 1.46 | 0.82 | 0.17 |
| Panama | 2.06 | 0.94 | 0.27 |
| Canada | 1.57 | 1.73 | 0.30 |
| Israel | 1.62 | 2.32 | 0.38 |
| United States | 1.36 | 12.71 | 1 |
| Average | 1.82 | 1.76 | 0.26 |

*Notes*: The Table shows the the skill premium, relative skill supply and efficiency across the countries in the narrow sample. Relative skill efficiency is normalised such that it takes value 1 for the United States.

Table IV: Relative Technology and Skill Quality across Countries - Broad Sample

| | N | Mean | Min | Max | Std. Dev. | Corr w/ $\tilde{H}/\tilde{L}$ | Corr w/ $y$ |
|---|---|---|---|---|---|---|---|
| $(A_H Q_H)/(A_L Q_L)$ | 42 | 0.28 | 0.04 | 1.25 | 0.22 | 0.58 | 0.17 |
| $A_H/A_L$ | 42 | 0.31 | 0.05 | 1.30 | 0.22 | 0.57 | 0.10 |
| $Q_H/Q_L$ | 42 | 0.88 | 0.72 | 1.27 | 0.11 | 0.15 | 0.43 |

*Notes*: The Table shows summary statistics for relative skill efficiency, the relative skill bias of technology and relative skill supply across the countries in the broad sample. Only countries with at least 100 unskilled and 100 skilled workers in the regression sample are included. All variables are normalised such that they take value 1 for the United States.

Table V: Variance and Covariance Decomposition

| | Baseline | Parametric | Parametric + Country-Specific Returns |
|---|---|---|---|
| $V_A$ | 0.92 | 0.91 | 0.87 |
| $V_Q$ | 0.04 | 0.04 | 0.06 |
| $Cov_A\left(\frac{\tilde{H}}{\tilde{L}}\right)$ | 0.95 | 0.94 | 0.89 |
| $Cov_Q\left(\frac{\tilde{H}}{\tilde{L}}\right)$ | 0.05 | 0.06 | 0.11 |
| $Cov_A\,(y)$ | 0.81 | 0.79 | 0.74 |
| $Cov_Q\,(y)$ | 0.19 | 0.21 | 0.26 |
| N | 42 | 42 | 42 |

*Notes*: The Table shows results for the decompositions of the skill efficiency variance (first panel) and its covariance with log relative skill supply (second panel) and with log GDP (second panel). Only countries with at least 100 unskilled and 100 skilled workers in the regression sample are included. The second column refers to the specification with linear return to years of schooling and quadratic to experience, while the third refers to the specification where returns are country-specific.

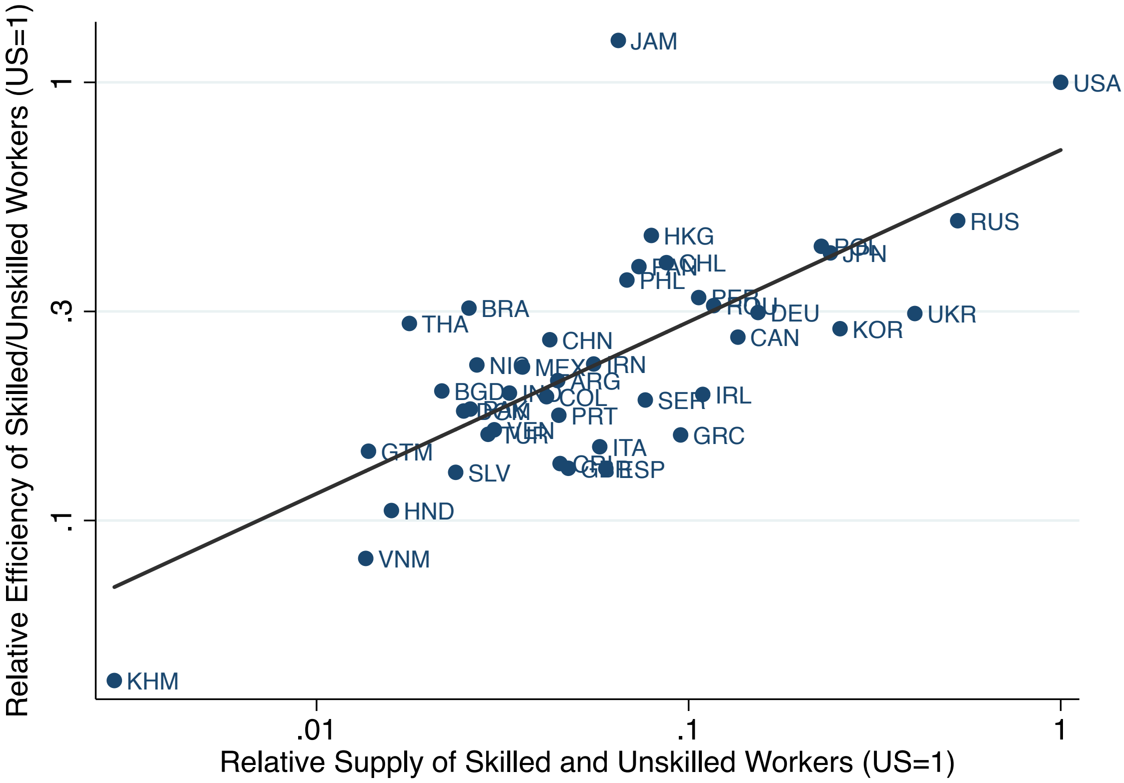
Table VI: Relative Technology and Skill Quality across Countries - Narrow Sample

| Country | $\frac{A_H Q_H}{A_L Q_L}$ | $\frac{A_H}{A_L}$ | $\frac{Q_H}{Q_L}$ |
|---|---|---|---|
| Indonesia | 0.08 | 0.10 | 0.85 |
| Brazil | 0.14 | 0.16 | 0.88 |
| Venezuela | 0.09 | 0.11 | 0.79 |
| Uruguay | 0.14 | 0.15 | 0.89 |
| India | 0.19 | 0.21 | 0.95 |
| Mexico | 0.17 | 0.20 | 0.83 |
| Trinidad and Tobago | 0.14 | 0.16 | 0.88 |
| Jamaica | 0.17 | 0.18 | 0.96 |
| Panama | 0.27 | 0.30 | 0.88 |
| Canada | 0.30 | 0.31 | 0.97 |
| Israel | 0.38 | 0.34 | 1.13 |
| United States | 1 | 1 | 1 |
| Average | 0.26 | 0.27 | 0.92 |

*Notes*: The Table shows summary statistics for relative skill efficiency, the relative skill bias of technology and relative skill supply across the countries in the narrow sample. All variables are normalised such that they take value 1 for the United States.

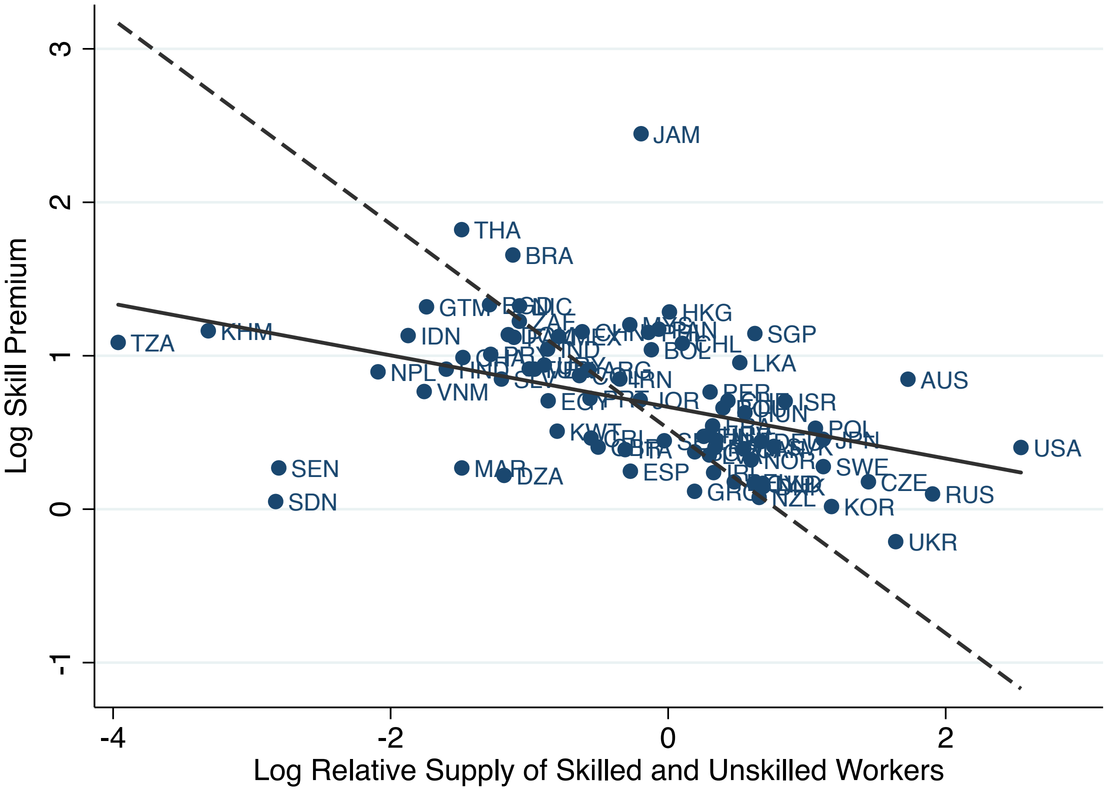
# Figures

Figure I: Relative Efficiency and Relative Supply of Skilled Labor - Broad Sample



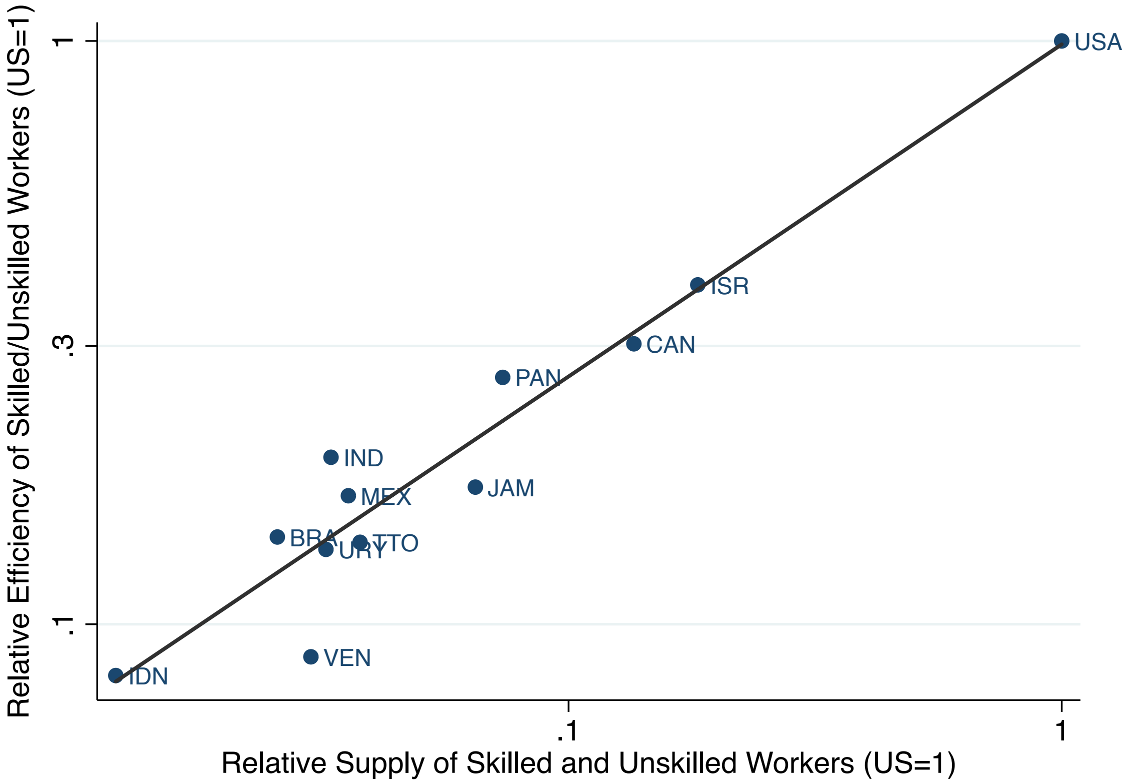

*Notes*: The figure plots on a log scale the relative efficiency and relative supply of skilled workers for countries in the broad sample. Both variables are normalized so that they take value 1 for the United States. The solid line represents the best exponential fit.

Figure II: Skill Premium and Skill Supply - Broad Sample



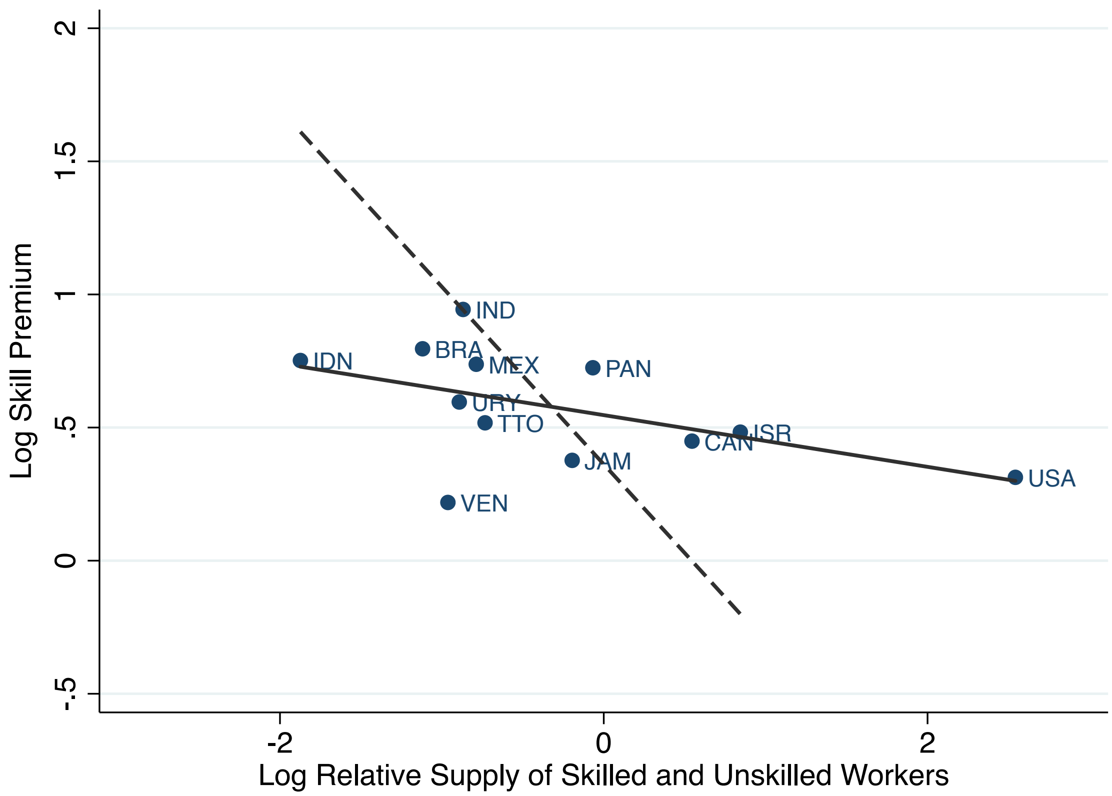

*Notes*: The figure plots the log skill premium and the log and relative supply of skilled workers for countries in the broad sample. The solid line represents the best linear fit. The dashed line has the slope of the predicted relationship (-0.67) in a counterfactual where skill efficiency and supply are uncorrelated.

Figure III: Relative Efficiency and Relative Supply of Skilled Labor - Narrow Sample



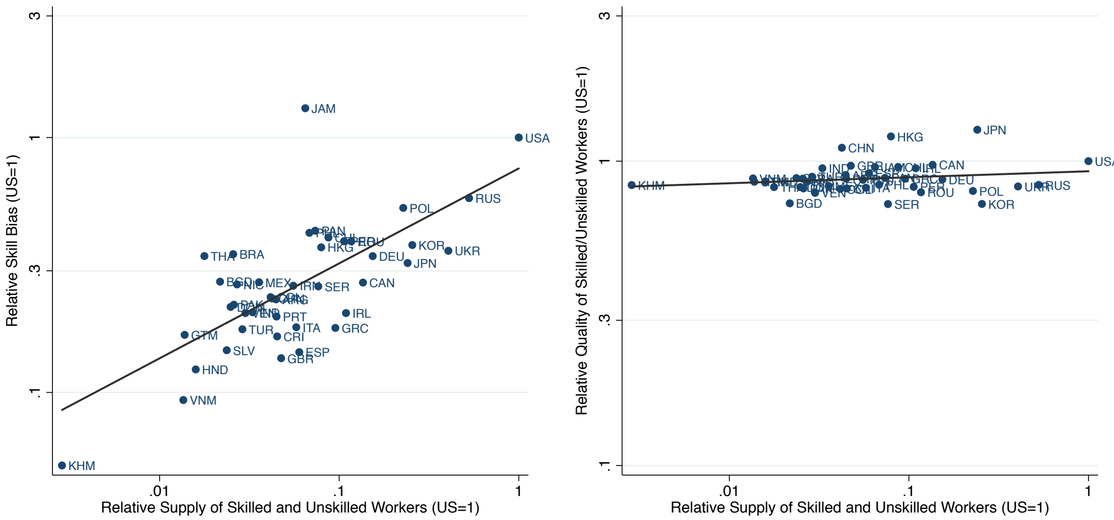

*Notes*: The figure plots on a log scale the relative efficiency and relative supply of skilled workers for countries in the narrow sample. Both variables are normalized so that they take value 1 for the United States. The solid line represents the best exponential fit.

Figure IV: Skill Premium and Skill Supply - Narrow Sample



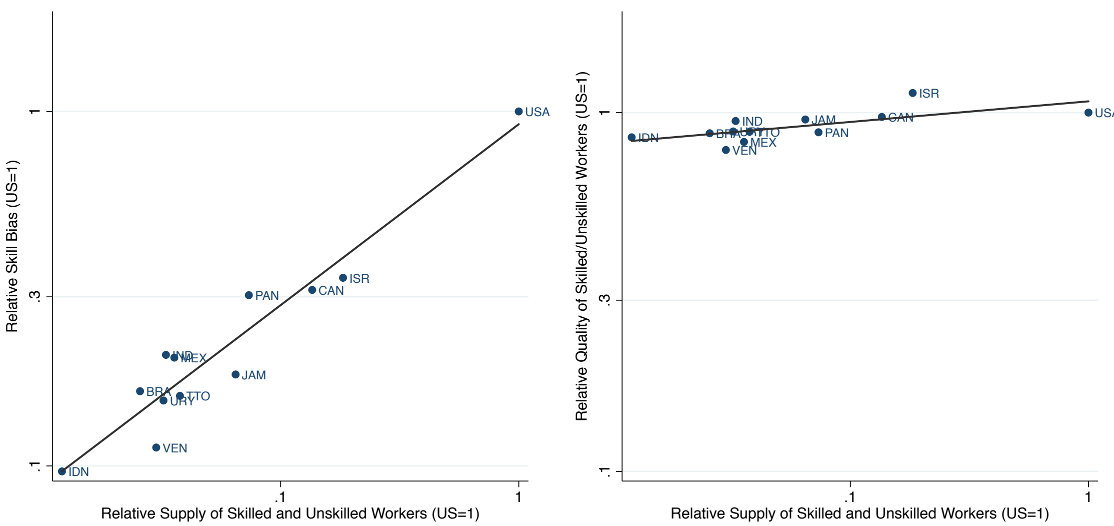

*Notes*: The figure plots the log skill premium and the log and relative supply of skilled workers for countries in the narrow sample. The solid line represents the best linear fit. The dashed line has the slope of the predicted relationship (-0.67) in a counterfactual where skill efficiency and supply are uncorrelated..

Figure V: Technology Skill Bias, Skill Quality and Skill Supply - Broad Sample



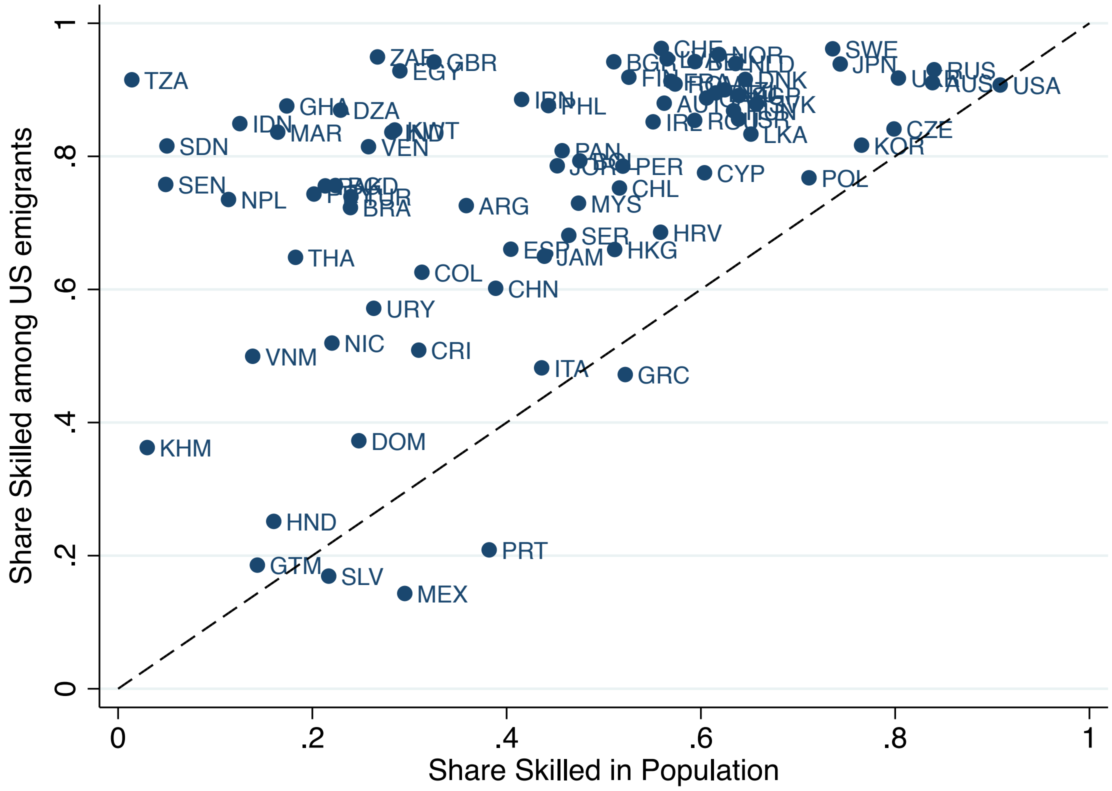

*Notes*: The left graph plots (on a log scale) the relative skill bias of technology again the relative supply of skilled labor. The right graph plots (on a log scale) the relative quality of skill labor again its relative supply. Only countries in the broad sample, with at least 100 unskilled and 100 skilled workers in the regression sample are included. All variables are normalised such that they take value 1 for the United States. The lines show the best exponential fits.

Figure VI: Technology Skill Bias, Skill Quality and Skill Supply - Narrow Sample




*Notes*: The left graph plots (on a log scale) the relative skill bias of technology again the relative supply of skilled labor. The right graph plots (on a log scale) the relative quality of skill labor again its relative supply. Only countries in the narrow sample are included. All variables are normalised such that they take value 1 for the United States. The lines show the best exponential fits.

Figure VII: Share of Skilled Workers among Emigrants by Country of Origin




*Notes*: The figure plots the share of skilled workers among emigrants to the US against the one in the country of origin. Only emigrants entirely educated in their country of origin are included. Skilled workers are defined as having completed secondary education or more.