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

Federico Rossi*

September 2019

ONLINE APPENDIX

A Data Appendix

IPUMS Data

This section describes the construction of the key variables from IPUMS and IPUMS International. All calculations and regressions on IPUMS data make use of the provided sample weights.

Wages. Weekly and hourly wages are constructed from the variable *incwage*, which reports the respondent's wage and salary income (except for some countries where this information is not available - see the "Country-Specific Notes" below). The information on hours, when available, is from the variable *hrswork1* (*uhrswork* in the US Census), which reports the number of hours worked in all jobs in either the previous week or a typical week (again, see the "Country-Specific Notes" for some exceptions).

Educational Attainment. Educational attainment is constructed from the variable *edattaind*, which records the highest level of completed schooling. Whenever this variable does not allow me to identify all 5 levels of educational attainment, I use additional information on the number of completed years of schooling (see the "Country-Specific Notes" for the details).

Country-Specific Notes.

- **Brazil**. Separate variables on wage and self-employment incomes are not available. I use earned income as proxy for wage income for workers that declare to be wage employed, and as proxy for self-employment income for workers that declare to be self-employed.

*University of Warwick, Department of Economics, Coventry CV4 7AL, United Kingdom. Email: federico.rossi@warwick.ac.uk.

- **India**. As discussed in the paper, hours worked are not available. To identify workers attached to the labor market (which are those included in the wage regressions), I use an indicator of full-time status. Self-employment income is not available.
- **Indonesia**. Self-employment income is not available.
- **Israel**. The educational variable does not allow to identify individuals with lower secondary education (or equivalent level). I impute this level of education based on available information on years of schooling. In particular, I assign lower secondary as educational attainment to (i) individuals that report "Primary or intermediate school" as educational attainment and at least 9 years of schooling, and (ii) individuals that report secondary school as educational attainment and less than "11 or 12" years of schooling (the two are grouped together).
- **Jamaica**. The educational variable does not allow to identify individuals with lower secondary education (or equivalent level). I impute this level of education based on available information on years of schooling. As discussed in the IPUMS International documentation, the years of schooling variable appears to include pre-school, and its average is higher compared to other sources. Taking into account this, I assign lower secondary as educational attainment to individuals that report secondary school as educational attainment and 12 years of schooling or less (as opposed to 10, which would be the relevant threshold according to statutory durations). The choice of this threshold is based on the observation that the empirical distribution of reported years of schooling displays substantial peaks at 8 and 13, which correspond to the statutory durations of primary and upper secondary school plus 2. Self-employment income is not available.
- **Mexico**. Separate variables on wage and self-employment incomes are not available. I use earned income as proxy for wage income for workers that declare to be wage employed, and as proxy for self-employment income for workers that declare to be self-employed.
- **Panama**. As discussed in the paper, hours worked are not available. Wage regressions are estimated on a sample including all employed workers (with available information on wage income). Self-employment income is not available.
- **United States**. The US Census includes additional information on the number of weeks worked in the previous year, which I combine with the number of hours usually worked per week to compute annual hours (and with annual wage income to compute hourly wages). The sample of workers attached to the labor market used for the wage regressions includes employed workers with at least 30 hours worked per week and 30 weeks worked in the previous year.
- **Uruguay**. Hours worked only refer to the main occupation. Self-employment income is not available.
- **Venezuela**. Separate variables on wage and self-employment incomes are not available. I use earned income as proxy for wage income for workers that declare to be wage employed, and as proxy for self-employment income for workers that declare to be self-employed.

Additional Data for Capital-Skill Complementarity Exercise

To quantify the capital-skill complementarity term in equation (12) of the paper, I construct a country-specific estimate of the ratio between the income share of equipment and the income share of high-skill labor, for the countries in the micro-data sample. To calculate the income share of equipment, I use data from the Capital Detail File of the Penn World Table, Version 9.1 (Feenstra et al., 2015): I sum the share of capital compensation of "machinery and (non-transport) equipment" and "transport equipment", and multiply it by an assumed overall capital share of 1/3 (common across countries). I calculate the income share of high-skill labor from the IPUMS data, as $\frac{2}{3} \times \frac{w_{H,c}\tilde{H}_c}{w_{H,c}\tilde{H}_c + w_{L,c}\tilde{L}_c}$.

B Results on Average Relative Skill Efficiency

This appendix develops and computes a measure of average relative skill efficiency, which does not require the choice of particular baseline types of low- and high-skill workers. Moreover, in the spirit of the corresponding exercise in the paper, it quantifies the relative role of human capital and technology in driving the cross-country variation in this object.

Measuring Average Relative Skill Efficiency. First, I amend the notation to explicit acknowledge that factor-augmenting technologies and embodied human capital can in principle vary across educational levels within skill groups. Let $A_{H,c,n}$ and $Q_{H,c,n}$ ($A_{L,c,m}$ and $Q_{L,c,m}$) be respectively the level of factor-specific technology and the embodied human capital of high-skill (low-skill) workers belonging to educational group n (m), where $n \in \mathcal{H} = \{\text{some tertiary, tertiary}\}$ ($m \in \mathcal{L} = \{\text{primary, lower secondary, upper secondary}\}$). The human capital aggregator can be written as

$$G\left(\{A_{H,c,n}Q_{H,c,n}\tilde{H}_{c,n}\}_{n\in\mathcal{H}}, \{A_{L,c,m}Q_{L,c,m}\tilde{L}_{c,m}\}_{m\in\mathcal{L}}\right) =$$
$$= \left[\left(\sum_{n\in\mathcal{H}} A_{H,c,n}Q_{H,c,n}\tilde{H}_{c,n}\right)^{\frac{\sigma-1}{\sigma}} + \left(\sum_{m\in\mathcal{L}} A_{L,c,m}Q_{L,c,m}\tilde{L}_{c,m}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

The wage ratio between high-skill workers of type n and low-skill workers of type m is given by

$$\frac{w_{H,c,n}}{w_{L,c,m}} = \left(\frac{A_{H,c,n}Q_{H,c,n}}{A_{L,c,m}Q_{L,c,m}}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{\sum_{i\in\mathcal{H}} \frac{A_{H,c,i}Q_{H,c,i}}{A_{H,c,n}Q_{H,c,n}}\tilde{H}_{c,i}}{\sum_{j\in\mathcal{L}} \frac{A_{L,c,j}Q_{L,c,j}}{A_{L,c,m}Q_{L,c,m}}\tilde{L}_{c,j}}\right)^{-\frac{1}{\sigma}} \tag{1}$$

In the paper I back out $\frac{A_{H,c}Q_{H,c}}{A_{L,c}Q_{L,c}} \equiv \frac{A_{H,c,\text{tertiary}}Q_{H,c,\text{tertiary}}}{A_{L,c,\text{upper secondary}}Q_{L,c,\text{upper secondary}}}$ from (1), using data on $\{w_{H,c,i}\}_{i\in\mathcal{H}}$, $\{w_{L,c,j}\}_{j\in\mathcal{L}}$, $\{\tilde{H}_{c,i}\}_{i\in\mathcal{H}}$, $\{\tilde{L}_{c,j}\}_{j\in\mathcal{L}}$ and the fact that $\frac{w_{H,c,i}}{w_{H,c,n}} = \frac{A_{H,c,i}Q_{H,c,i}}{A_{H,c,n}Q_{H,c,n}}$ and $\frac{w_{L,c,j}}{w_{L,c,m}} = \frac{A_{L,c,j}Q_{L,c,j}}{A_{L,c,m}Q_{L,c,m}}$ for all $i \in \mathcal{H}$ and $j \in \mathcal{L}$. Here, I consider the following measure of average relative skill efficiency,

$$\overline{RSE}_c = \frac{\sum_{n\in\mathcal{H}} A_{H,c,n}Q_{H,c,n}\frac{\tilde{H}_{c,n}}{\sum_{i\in\mathcal{H}}\tilde{H}_{c,i}}}{\sum_{m\in\mathcal{L}} A_{L,c,m}Q_{L,c,m}\frac{\tilde{L}_{c,m}}{\sum_{j\in\mathcal{L}}\tilde{L}_{c,j}}}$$

that is the ratio between weighted averages of the efficiencies of the various educational groups. This measure is obviously affected by the educational composition within skill groups; to assess the importance of this, I also consider an adjusted version where educational shares within skill groups are fixed at the US levels

$$\overline{RSE}_c^{Adj} = \frac{\sum_{n\in\mathcal{H}} A_{H,c,n}Q_{H,c,n}\frac{\tilde{H}_{US,n}}{\sum_{n\in\mathcal{H}}\tilde{H}_{US,n}}}{\sum_{m\in\mathcal{L}} A_{L,c,m}Q_{L,c,m}\frac{\tilde{L}_{US,m}}{\sum_{m\in\mathcal{L}}\tilde{L}_{US,m}}}$$

Both $\overline{RSE}_c$ and $\overline{RSE}_c^{Adj}$ can be easily computed using data on wages and educational shares. In particular, notice that one can write

$$\overline{RSE}_c = \left(\frac{\bar{w}_{H,c}}{\bar{w}_{L,c}}\right)^{\frac{\sigma}{\sigma-1}} \left(\frac{\sum_{n\in\mathcal{H}} \tilde{H}_{c,n}}{\sum_{m\in\mathcal{L}} \tilde{L}_{c,m}}\right)^{\frac{1}{\sigma-1}}$$

$$\overline{RSE}_c^{Adj} = \frac{\bar{w}_{H,c}^{Adj}}{\bar{w}_{L,c}^{Adj}} \left(\frac{\bar{w}_{H,c}}{\bar{w}_{L,c}}\right)^{\frac{1}{\sigma-1}} \left(\frac{\sum_{n\in\mathcal{H}} \tilde{H}_{c,n}}{\sum_{m\in\mathcal{L}} \tilde{L}_{c,m}}\right)^{\frac{1}{\sigma-1}}$$

where $\bar{w}_{H,c}$ and $\bar{w}_{L,c}$ are weighted averages of the wages across all educational groups,

$$\bar{w}_{H,c} = \sum_{n\in\mathcal{H}} w_{H,c,n} \frac{\tilde{H}_{c,n}}{\sum_{i\in\mathcal{H}} \tilde{H}_{c,i}}$$

$$\bar{w}_{L,c} = \sum_{m\in\mathcal{L}} w_{L,c,m} \frac{\tilde{L}_{c,m}}{\sum_{j\in\mathcal{L}} \tilde{L}_{c,j}}$$

and $\bar{w}_{H,c}^{Adj}$ and $\bar{w}_{L,c}^{Adj}$ are the composition-adjusted counterparts

$$\bar{w}_{H,c}^{Adj} = \sum_{n\in\mathcal{H}} w_{H,c,n} \frac{\tilde{H}_{US,n}}{\sum_{i\in\mathcal{H}} \tilde{H}_{US,i}}$$

$$\bar{w}_{L,c}^{Adj} = \sum_{m\in\mathcal{L}} w_{L,c,m} \frac{\tilde{L}_{US,m}}{\sum_{j\in\mathcal{L}} \tilde{L}_{US,j}}$$

Table B.1 displays the resulting $\overline{RSE}_c$ and $\overline{RSE}_c^{Adj}$. For both measures, the country-specific estimates and the elasticities with respect to GDP per worker are close to the baseline relative skill efficiency shown in Table I of the paper. This reflects the fact that the baseline groups used in the paper (tertiary educated for the high-skilled, upper secondary educated for the low-skilled) are not atypical in terms of the cross-country variation in relative skill efficiency (and also relatively large across all countries, meaning that they drive a significant part of the variation in average relative skill efficiency). The composition-adjusted measures varies slightly more between rich and poor countries, as poor countries are relatively more abundant in low-educated workers within the low-skill group.

Quantifying the Role of Relative Human Capital. Average relative skill efficiency (either in the adjusted or non-adjusted form) is not multiplicative in a "relative technology" term and a "relative human capital" term (contrary to the measure considered in the paper). To have a sense of the relative importance of technology and human capital, I consider the following counterfactual values of average relative skill efficiency

$$\overline{RSE}_c^{Q} = \frac{\sum_{n\in\mathcal{H}} A_{H,US,n} Q_{H,c,n} \frac{\tilde{H}_{c,n}}{\sum_{i\in\mathcal{H}} \tilde{H}_{c,i}}}{\sum_{m\in\mathcal{L}} A_{L,US,m} Q_{L,c,m} \frac{\tilde{L}_{c,m}}{\sum_{j\in\mathcal{L}} \tilde{L}_{c,j}}}$$

$$\overline{RSE}_c^{A} = \frac{\sum_{n\in\mathcal{H}} A_{H,c,n} Q_{H,US,n} \frac{\tilde{H}_{c,n}}{\sum_{i\in\mathcal{H}} \tilde{H}_{c,i}}}{\sum_{m\in\mathcal{L}} A_{L,c,m} Q_{L,US,m} \frac{\tilde{L}_{c,m}}{\sum_{j\in\mathcal{L}} \tilde{L}_{c,j}}}$$

where $\overline{RSE}_c^{Q}$ fixes the technology terms at the US levels (keeping the human capital terms as in the data), and $\overline{RSE}_c^{A}$ fixes the human capital terms at the US levels (keeping the technology terms as in the data). The more the cross-country variation in $\overline{RSE}_c$ is driven by the relative human capital endowment of high-skill labor, the more $\overline{RSE}_c^Q$ will differ across countries (and viceversa for $\overline{RSE}_c^A$). I denote the corresponding counterfactuals for the adjusted version of average relative skill efficiency as $\overline{RSE}_c^{Adj,Q}$ and $\overline{RSE}_c^{Adj,A}$.

These counterfactuals can be computed using estimates of wages (for both migrants and natives), relative human capital endowments and educational shares. In particular, they can be written as

$$\overline{RSE}_c^Q = \left(\frac{\bar{w}_{H,US}}{\bar{w}_{L,US}}\right)^{\frac{1}{\sigma-1}} \left(\frac{\sum_{n\in\mathcal{H}} \tilde{H}_{US,n}}{\sum_{m\in\mathcal{L}} \tilde{L}_{US,m}}\right)^{\frac{1}{\sigma-1}} \frac{\sum_{n\in\mathcal{H}} w^c_{H,US,n} \frac{\tilde{H}_{c,n}}{\sum_{i\in\mathcal{H}} \tilde{H}_{c,i}}}{\sum_{m\in\mathcal{L}} w^c_{L,US,m} \frac{\tilde{L}_{c,m}}{\sum_{j\in\mathcal{L}} \tilde{L}_{c,j}}}$$

$$\overline{RSE}_c^A = \left(\frac{\bar{w}_{H,c}}{\bar{w}_{L,c}}\right)^{\frac{1}{\sigma-1}} \left(\frac{\sum_{n\in\mathcal{H}} \tilde{H}_{c,n}}{\sum_{m\in\mathcal{L}} \tilde{L}_{c,m}}\right)^{\frac{1}{\sigma-1}} \frac{\sum_{n\in\mathcal{H}} w_{H,c,n} \frac{Q_{H,US,n}}{Q_{H,c,n}} \frac{\tilde{H}_{c,n}}{\sum_{i\in\mathcal{H}} \tilde{H}_{c,i}}}{\sum_{m\in\mathcal{L}} w_{L,c,m} \frac{Q_{L,US,m}}{Q_{L,c,m}} \frac{\tilde{L}_{c,m}}{\sum_{j\in\mathcal{L}} \tilde{L}_{c,j}}}$$

where $w^c_{H,US,n}$ and $w^c_{L,US,m}$ are the wages of high- and low-skill workers belonging to educational groups n and m, working in the United States after having been educated in country c. The corresponding expressions for the composition-adjusted counterfactuals are

$$\overline{RSE}_c^{Adj,Q} = \left(\frac{\bar{w}_{H,US}}{\bar{w}_{L,US}}\right)^{\frac{1}{\sigma-1}} \left(\frac{\sum_{n\in\mathcal{H}} \tilde{H}_{US,n}}{\sum_{m\in\mathcal{L}} \tilde{L}_{US,m}}\right)^{\frac{1}{\sigma-1}} \frac{\sum_{n\in\mathcal{H}} w^c_{H,US,n} \frac{\tilde{H}_{US,n}}{\sum_{i\in\mathcal{H}} \tilde{H}_{US,i}}}{\sum_{m\in\mathcal{L}} w^c_{L,US,m} \frac{\tilde{L}_{US,m}}{\sum_{j\in\mathcal{L}} \tilde{L}_{US,j}}}$$

$$\overline{RSE}_c^{Adj,A} = \left(\frac{\bar{w}_{H,c}}{\bar{w}_{L,c}}\right)^{\frac{1}{\sigma-1}} \left(\frac{\sum_{n\in\mathcal{H}} \tilde{H}_{c,n}}{\sum_{m\in\mathcal{L}} \tilde{L}_{c,m}}\right)^{\frac{1}{\sigma-1}} \frac{\sum_{n\in\mathcal{H}} w_{H,c,n} \frac{Q_{H,US,n}}{Q_{H,c,n}} \frac{\tilde{H}_{US,n}}{\sum_{i\in\mathcal{H}} \tilde{H}_{US,i}}}{\sum_{m\in\mathcal{L}} w_{L,c,m} \frac{Q_{L,US,m}}{Q_{L,c,m}} \frac{\tilde{L}_{US,m}}{\sum_{j\in\mathcal{L}} \tilde{L}_{US,j}}}$$

I compute $\overline{RSE}_c^Q$, $\overline{RSE}_c^A$, $\overline{RSE}_c^{Adj,Q}$ and $\overline{RSE}_c^{Adj,A}$ for all countries in the broad sample. Table B.2 reports the corresponding elasticities with respect to GDP per worker.[1] For both the baseline and composition-adjusted measures, fixing the skill bias of technology to the US level (i.e., $\overline{RSE}_c^Q$ and $\overline{RSE}_c^{Adj,Q}$) almost closes the cross-country gap in average relative skill efficiency, while equalizing the relative human capital endowments ($\overline{RSE}_c^A$ and $\overline{RSE}_c^{Adj,A}$) has a limited impact. The magnitudes here should be taken with care, as for some educational groups the sample sizes at the country-of-origin-level are small (for example, there are 4 countries for which I have less than 10 primary-educated workers in the sample). However, the results in Table B.2 do suggest that the finding that technology is more important than embodied human capital in driving the dispersion in relative skill efficiency is not driven by the choice of particular baseline types of high- and low-skill labor.

[1] Given that average relative skill efficiency is not multiplicative in a "relative technology" term and a "relative human capital" term, the elasticities of $\overline{RSE}_c^Q$ and $\overline{RSE}_c^A$ with respect to GDP per worker do not sum to the corresponding elasticity of $\overline{RSE}_c$ (the same being true for the composition-adjusted quantities).

Table B.1: Average Relative Skill Efficiency

Country	$\overline{RSE}$	$\overline{RSE}^{Adj}$
India	0.083	0.042
Indonesia	0.005	0.004
Jamaica	0.012	0.010
Brazil	0.132	0.082
Venezuela	0.108	0.099
Uruguay	0.128	0.115
Panama	0.131	0.096
Mexico	0.075	0.049
Trinidad and Tobago	0.019	0.018
Israel	0.145	0.138
Canada	0.757	0.733
United States	1	1
Elasticity wrt GDP p.w.	1.199**	1.401***
	(0.413)	(0.387)

Notes: The Table shows the two measures of average relative skill efficiency defined in the text across the countries in the micro-data sample. Both measures are normalised such that they take value 1 for the United States. The last row shows the coefficient of a regression of the log of each variable on log GDP per worker (standard errors in parentheses).

Table B.2: Average Relative Skill Efficiency - Counterfactuals

	$\overline{RSE}^{Q}$	$\overline{RSE}^{A}$	$\overline{RSE}$	$\overline{RSE}^{Adj,Q}$	$\overline{RSE}^{Adj,A}$	$\overline{RSE}^{Adj}$
Elasticity wrt GDP p.w.	0.021	0.836***	0.889***	0.041***	0.975***	1.017***
	(0.017)	(0.191)	(0.188)	(0.013)	(0.196)	(0.193)

Notes: The Table shows the elasticity (standard errors in parentheses) with respect to GDP per worker of several counterfactual measures of average relative skill efficiency, as defined in the text. The elasticities are estimated in the broad sample (65 countries).

C Results with Mincerian-Based Skill Premia

As shown in Table I of the paper, constructing skill premia and relative supplies using a Mincerian return of 10% common across countries leads to an overstatement of the cross-country variation in relative skill efficiency. This is because the skill premium is in fact somewhat higher in poor countries, as opposed to constant across countries as implied by a common Mincerian return. This section documents that this finding also applies to the collections of Mincerian returns used in the literature, and proposes a possible explanation based on the non-linearity of returns with respect to years of schooling.

Table C.1 shows skill premia and relative supplies based on alternative country-specific estimates of the Mincerian return (relative supplies are constructed here using educational attainment data for the working age population). The first two columns use the estimates collected in Caselli et al. (2016) (I use the values for the 2000s). Columns 3-4 use the Mincerian returns estimated in Montenegro and Patrinos (2014), a collection that, while not yet widely used in the development accounting literature, offers more comparable estimates from harmonized household surveys. Columns 5-6 use Mincerian returns estimated from the IPUMS and IPUMS International data used for the core analysis of the paper.

For all sources, the resulting skill premium is flatter across countries compared to the baseline in Table I of the paper. Even when using the same (IPUMS) data, the elasticities of the Mincerian-based skill premium with respect to GDP per worker and the relative supply are between 1/3 and 1/2 of the corresponding baseline elasticities; the estimates of the relative supply are instead similar. As a result of a less variable skill premium and an equally variable relative supply, Mincerian-based estimates of relative skill efficiency display more variability across countries compared to the baseline (see Table C.2).

Table C.1: Skill Premium and Relative Supply from Mincerian Coefficients

	Caselli, Ponticelli and Rossi (2016)		Montenegro and Patrinos (2014)		IPUMS	
Country	w_H/w_L	$\tilde{H}/\tilde{L}$	w_H/w_L	$\tilde{H}/\tilde{L}$	w_H/w_L	$\tilde{H}/\tilde{L}$
India	1.405	0.191	1.323	0.176	1.519	0.211
Indonesia	1.578	0.084	1.498	0.080	1.481	0.079
Jamaica	3.165	0.104	1.559	0.088	1.432	0.085
Brazil	1.874	0.259	1.772	0.246	1.690	0.235
Venezuela	1.550	0.316	1.350	0.298	1.526	0.314
Uruguay	1.610	0.670	1.534	0.644	1.213	0.515
Panama	1.730	0.460	1.623	0.440	1.542	0.423
Mexico	1.571	0.279	1.649	0.292	1.495	0.266
Trinidad and Tobago	-	-	-	-	1.343	0.129
Israel	1.623	0.696	-	-	1.259	0.663
Canada	1.428	1.793	1.623	1.819	1.323	1.774
United States	1.636	1.596	1.642	1.596	1.452	1.610
Elasticity wrt GDP p.w.	-0.018	0.950***	0.079**	1.033***	-0.034	0.902***
	(0.082)	(0.200)	(0.032)	(0.226)	(0.035)	(0.244)
Elasticity wrt $\tilde{H}/\tilde{L}$	-0.095	1	0.028	1	-0.030	1
	(0.067)	-	(0.029)	-	(0.028)	-

Notes: The Table shows the skill premium and relative skill supply based on country-specific estimates of the Mincerian coefficient, across the countries in the micro-data sample. The column headings indicate the source of the Mincerian coefficients, with *IPUMS* referring to estimates from the same IPUMS and IPUMS International samples used for the main analysis of the paper. The last two rows show the coefficient of a regression of the log of each variable on log GDP per worker and log relative skill supply (standard errors in parentheses).

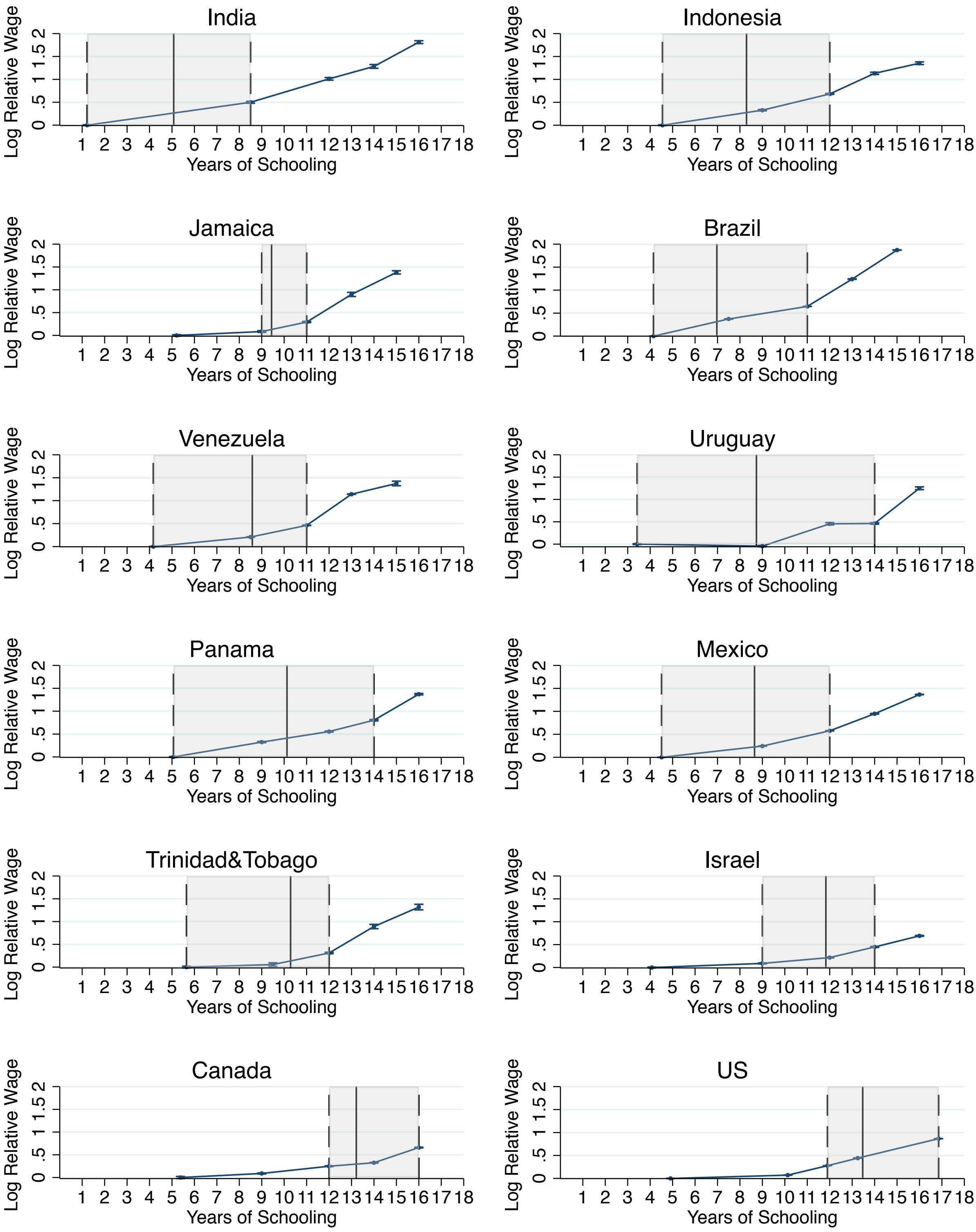
Table C.2: Relative Skill Efficiency from Mincerian Coefficients

	$(A_H Q_H)\,/\,(A_L Q_L)$		
Country	Caselli, Ponticelli and Rossi (2016)	Montenegro and Patrinos (2014)	IPUMS
India	0.009	0.006	0.020
Indonesia	0.002	0.002	0.003
Jamaica	0.031	0.003	0.003
Brazil	0.039	0.030	0.033
Venezuela	0.033	0.019	0.044
Uruguay	0.168	0.133	0.060
Panama	0.098	0.073	0.083
Mexico	0.027	0.034	0.030
Trinidad and Tobago	-	-	0.005
Israel	0.186	-	0.111
Canada	0.839	1.254	0.918
United States	1	1	1
Elasticity wrt GDP p.w.	1.845***	2.303***	1.700***
	(0.318)	(0.426)	(0.503)

Notes: The Table shows estimates of relative skill efficiency based on country-specific estimates of the Mincerian coefficient, across the countries in the micro-data sample. The column headings indicate the source of the Mincerian coefficients, with *IPUMS* referring to estimates from the same IPUMS and IPUMS International samples used for the main analysis of the paper. The last row shows the coefficient of a regression of the log of each variable on log GDP per worker (standard errors in parentheses).

Why do Mincerian returns understate the variation in the skill premium? Figure C.1 illustrates a possible reason. The 12 panels plot, for all countries in the micro-data sample, the estimated log wage gap with respect to primary educated workers against the average years of schooling corresponding to each level of educational attainment. The Mincerian return is the slope of the best linear fit for this relationship. However, for most countries returns to years of education appear to be convex, and particularly low at low levels of educational attainment. As illustrated by the vertical lines in each graph (identifying the 25th percentile, average and 75th percentile of years of schooling in each sample), the sample used for the estimation of Mincerian returns is biased towards low-education workers in poor countries, and towards high-education workers in rich countries. Since low-education workers have relatively low returns, this difference in sample composition tends to reduce the poorer countries' Mincerian returns, therefore understating the cross-country variation in the Mincerian-based skill premium.

Figure C.1: Returns to Education and Years of Schooling



Notes: The figure plots for all countries in the micro-data sample the average log wage (normalised by the average log wage of primary educated workers) by education groups agains the average years of schooling in each group. The vertical lines identify the average (solid line), 25th percentile (left dashed line) and 75th percentile (right dashed line) of years of schooling for the workers in the wage regression sample. For this calculation, workers are assigned the average years of schooling of their educational group. The graphs are ordered (left to right, up to bottom) by the country's GDP per worker.

D Further Material on Relative Skill Efficiency across Sectors

D.1 Sector-Level Elasticity of Substitution

The derivation of the mapping between aggregate and sector-level elasticity of substitution can be found in Oberfield and Raval (2014). In this section I provide the exact definition of the terms entering in equation (11) of the paper. The calibration of the sector-level elasticity is entirely based on US data, since the available estimates of the aggregate elasticity are from the US. In light of this, in what follows I omit the country subscript.

Define the high-skill share of wage payments in sector s as

$$a_s = \frac{w_{H,s}\tilde{H}_s}{w_{H,s}\tilde{H}_s + w_{L,s}\tilde{L}_s}$$

the sectoral share of wage payments as

$$\theta_s = \frac{w_{H,s}\tilde{H}_s + w_{L,s}\tilde{L}_s}{\sum_s w_{H,s}\tilde{H}_s + w_{L,s}\tilde{L}_s}$$

and the average high-skill share of wage payments as

$$a = \sum_s \theta_s a_s$$

The heterogeneity index χ is defined as

$$\chi = \sum_s \frac{(a_s - a)^2}{a(1-a)}\theta_s$$

Let the capital share of and the elasticity of substitution between capital and the labor bundle in sector s be, respectively, α_s and ζ_s. The weighted averages $\bar{\alpha}$ and $\bar{\zeta}$ are defined as

$$\bar{\alpha} = \sum_s \frac{(a_s - a)^2\theta_s}{\sum_j (a_j - a)^2\theta_j}\alpha_s$$

$$\bar{\zeta} = \sum_s \frac{(a_s - a)^2\theta_s\alpha_s}{\sum_j (a_j - a)^2\theta_j\alpha_j}\zeta_s$$

D.2 Sector-Level Relative Skill Efficiency: Additional Results

Table D.1: Skill Premium and Relative Skill Supply - Sector-Level

Panel A	w_H/w_L				
	Aggregate	Agriculture	Manufacturing	Low-Skill Services	High-Skill Services
India	2.230	2.280	2.114	1.869	1.531
Indonesia	1.957	2.984	3.015	2.356	1.511
Jamaica	2.969	4.098	3.710	3.580	2.425
Brazil	3.412	3.895	3.806	3.384	3.010
Venezuela	2.490	2.331	2.692	2.383	2.393
Uruguay	2.218	3.382	2.751	2.304	1.783
Panama	2.262	3.068	2.293	2.518	1.956
Mexico	2.202	2.504	2.434	2.311	1.938
Trinidad and Tobago	2.741	2.518	3.615	3.443	2.088
Israel	1.606	1.292	1.532	1.376	1.633
Canada	1.508	1.417	1.415	1.453	1.460
United States	1.802	1.690	1.842	1.771	1.868
Elasticity wrt GDP p.w.	-0.138	-0.274**	-0.209*	-0.159	-0.016
	(0.077)	(0.106)	(0.102)	(0.105)	(0.080)

Panel B	$\tilde{H}/\tilde{L}$				
	Aggregate	Agriculture	Manufacturing	Low-Skill Services	High-Skill Services
India	0.205	0.035	0.156	0.213	1.145
Indonesia	0.070	0.006	0.040	0.036	0.395
Jamaica	0.067	0.005	0.038	0.027	0.223
Brazil	0.158	0.014	0.083	0.075	0.419
Venezuela	0.257	0.038	0.195	0.131	0.633
Uruguay	0.363	0.089	0.214	0.194	0.924
Panama	0.313	0.024	0.152	0.221	0.843
Mexico	0.226	0.025	0.128	0.131	0.645
Trinidad and Tobago	0.101	0.027	0.056	0.050	0.228
Israel	0.596	0.234	0.461	0.292	1.104
Canada	1.539	0.625	0.961	0.965	3.447
United States	1.397	0.554	0.786	0.864	2.807
Elasticity wrt GDP p.w.	0.911***	1.459***	0.900***	0.843**	0.530*
	(0.244)	(0.366)	(0.272)	(0.316)	(0.268)

Notes: The Table shows the aggregate and sector-specific skill premium (Panel A) and relative supply of skilled labor (Panel B) across the countries in the micro-data sample. The last row of each panel show the coefficient of a regression of the log of each variable and log GDP per worker (standard errors in parentheses).

Table D.2: Sectoral Shares of Employment by Education Level

Panel A: Primary	Agriculture	Manufacturing	Low-Skill Services	High-Skill Services
India	0.707	0.148	0.105	0.039
Indonesia	0.555	0.196	0.230	0.019
Jamaica	0.286	0.199	0.409	0.105
Brazil	0.251	0.243	0.333	0.173
Venezuela	0.229	0.216	0.422	0.133
Uruguay	0.166	0.235	0.399	0.200
Panama	0.380	0.187	0.316	0.117
Mexico	0.272	0.312	0.314	0.102
Trinidad and Tobago	0.123	0.345	0.385	0.147
Israel	0.054	0.406	0.283	0.257
Canada	0.083	0.401	0.312	0.205
United States	0.103	0.379	0.322	0.196

Panel B: Lower Secondary	Agriculture	Manufacturing	Low-Skill Services	High-Skill Services
India	0.494	0.200	0.228	0.077
Indonesia	0.298	0.284	0.326	0.091
Jamaica	0.150	0.217	0.478	0.155
Brazil	0.051	0.238	0.400	0.311
Venezuela	0.045	0.235	0.492	0.228
Uruguay	0.094	0.247	0.447	0.213
Panama	0.076	0.237	0.458	0.228
Mexico	0.078	0.337	0.369	0.216
Trinidad and Tobago	0.082	0.303	0.414	0.201
Israel	0.033	0.388	0.314	0.264
Canada	0.046	0.280	0.441	0.233
United States	0.039	0.270	0.460	0.231

Panel C: Upper Secondary	Agriculture	Manufacturing	Low-Skill Services	High-Skill Services
India	0.378	0.195	0.263	0.165
Indonesia	0.115	0.268	0.284	0.333
Jamaica	0.055	0.189	0.461	0.294
Brazil	0.029	0.188	0.353	0.430
Venezuela	0.022	0.188	0.432	0.359
Uruguay	0.051	0.167	0.383	0.398
Panama	0.039	0.187	0.451	0.322
Mexico	0.031	0.235	0.353	0.381
Trinidad and Tobago	0.031	0.258	0.360	0.350
Israel	0.030	0.268	0.314	0.389
Canada	0.029	0.266	0.387	0.318
United States	0.025	0.301	0.339	0.334

Panel D: Some Tertiary	Agriculture	Manufacturing	Low-Skill Services	High-Skill Services
India	0.316	0.182	0.248	0.254
Indonesia	0.037	0.156	0.187	0.620
Jamaica	0.007	0.124	0.266	0.603
Brazil	0.014	0.137	0.233	0.616
Venezuela	0.012	0.158	0.244	0.586
Uruguay	0.048	0.211	0.295	0.446
Panama	0.018	0.133	0.386	0.462

Mexico	0.020	0.182	0.301	0.497
Trinidad and Tobago	0.015	0.205	0.224	0.556
Israel	0.015	0.257	0.176	0.551
Canada	0.021	0.220	0.318	0.442
United States	0.015	0.206	0.289	0.490

Panel E: Tertiary	Agriculture	Manufacturing	Low-Skill Services	High-Skill Services
India	0.140	0.155	0.207	0.497
Indonesia	0.035	0.164	0.145	0.657
Jamaica	0.009	0.107	0.159	0.725
Brazil	0.014	0.117	0.148	0.720
Venezuela	0.007	0.136	0.192	0.665
Uruguay	0.037	0.094	0.108	0.760
Panama	0.015	0.105	0.236	0.644
Mexico	0.016	0.176	0.179	0.629
Trinidad and Tobago	0.013	0.158	0.172	0.657
Israel	0.009	0.203	0.108	0.680
Canada	0.012	0.121	0.175	0.692
United States	0.010	0.146	0.148	0.695

Notes: The Table shows the sectoral employment shares by education level. Each panel refers to a different level of educational attainment.

D.3 Details on Counterfactual Exercise

As stated in the text of the paper, I impose the following additional assumptions for the counterfactual exercise:

(A1) The sectoral production function is

$$Y_{c,s} = Z_{c,s} K_{c,s}^{\alpha} \left[(A_{H,c,s} H_{c,s})^{\frac{\tilde{\sigma}-1}{\tilde{\sigma}}} + (A_{L,c,s} L_{c,s})^{\frac{\tilde{\sigma}-1}{\tilde{\sigma}}} \right]^{\frac{(1-\alpha)\tilde{\sigma}}{\tilde{\sigma}-1}}$$

where $K_{c,s}$ is physical capital in sector s and $Z_{c,s}$ is a technological term capturing total factor productivity.

(A2) The prices of the sectoral goods are exogenous to the allocation of labor (as for small open economies).

(A3) Labor and physical capital are not mobile across sectors.

(A4) Conditional on educational attainment, human capital does not vary across sectors.

Let $w_{H,c,s,n}$ and $w_{L,c,s,m}$ be the wages paid to skilled and unskilled workers employed in sector s and belonging to educational groups n and m, where $n \in \{\text{some tertiary, tertiary}\}$ and $m \in \{\text{primary, lower secondary, upper secondary}\}$. Moreover, let $\widetilde{H}^{\text{bodies}}_{c,s,n}$ and $\widetilde{H}^{\text{hours}}_{c,s,n}$ ($\widetilde{L}^{\text{bodies}}_{c,s,m}$ and $\widetilde{L}^{\text{hours}}_{c,s,M}$) be the numbers of high-skill (low-skill) workers and total hours worked in sector s for educational group n (m), with $\widetilde{H}^{\text{bodies}}_{c,n}$ and $\widetilde{H}^{\text{hours}}_{c,n}$ ($\widetilde{L}^{\text{bodies}}_{c,m}$ and $\widetilde{L}^{\text{hours}}_{c,M}$) being the corresponding totals across sectors. One can then write $w_{H,c,n}$ and $w_{L,c,m}$ as the geometric averages of these wages across sectors

$$\begin{aligned} w_{H,c,n} &= \prod_s \left(w_{H,c,s,n} \right)^{\frac{\widetilde{H}^{\text{bodies}}_{c,s,n}}{\widetilde{H}^{\text{bodies}}_{c,n}}} \\ w_{L,c,m} &= \prod_s \left(w_{L,c,s,m} \right)^{\frac{\widetilde{L}^{\text{bodies}}_{c,s,m}}{\widetilde{L}^{\text{bodies}}_{c,m}}} \end{aligned} \tag{2}$$

where the weights $\frac{\widetilde{H}^{\text{bodies}}_{c,s,n}}{\widetilde{H}^{\text{bodies}}_{c,n}}$ and $\frac{\widetilde{L}^{\text{bodies}}_{c,s,m}}{\widetilde{L}^{\text{bodies}}_{c,m}}$ are the education-specific shares of workers employed in sector s (since the unit of observation for the computation of $w_{H,c,n}$ and $w_{L,c,m}$ is the worker, not weighted by hours worked).[2]

For every educational group n and m, I construct counterfactual allocations of employment and hours across sectors based on the (education-specific) sectoral employment and hours shares in the US,

$$\widetilde{H}^{C,\text{bodies}}_{c,s,n} = \left(\frac{\widetilde{H}^{\text{bodies}}_{n,US,s}}{\widetilde{H}^{\text{bodies}}_{n,US}}\right)\widetilde{H}^{\text{bodies}}_{n,c} \qquad \widetilde{L}^{C,\text{bodies}}_{c,s,m} = \left(\frac{\widetilde{L}^{\text{bodies}}_{m,US,s}}{\widetilde{L}^{\text{bodies}}_{m,US}}\right)\widetilde{L}^{\text{bodies}}_{m,c}$$

$$\widetilde{H}^{C,\text{hours}}_{c,s,n} = \left(\frac{\widetilde{H}^{\text{hours}}_{n,US,s}}{\widetilde{H}^{\text{hours}}_{n,US}}\right)\widetilde{H}^{\text{hours}}_{n,c} \qquad \widetilde{L}^{C,\text{hours}}_{c,s,m} = \left(\frac{\widetilde{L}^{\text{hours}}_{m,US,s}}{\widetilde{L}^{\text{hours}}_{m,US}}\right)\widetilde{L}^{\text{hours}}_{m,c}$$

Denote the equilibrium wages prevailing at this counterfactual allocation as $w^C_{H,c,s,n}$ and $w^C_{L,c,s,m}$. Under assumptions (A1)-(A4), these counterfactual wages can be written as[3]

$$\begin{aligned}
w^C_{H,c,s,n} &= w_{H,c,s,n}\left(\frac{\widetilde{H}^{C,\text{hours}}_{c,s,n}}{\widetilde{H}^{\text{hours}}_{c,s,n}}\right)^{-\alpha}\left[\frac{1+\left(\frac{A_{L,c,s,m}Q_{L,c,m}\widetilde{L}^{C,\text{hours}}_{c,s,m}}{A_{H,c,s,n}Q_{H,c,n}\widetilde{H}^{C,\text{hours}}_{c,s,n}}\right)^{\frac{\tilde{\sigma}-1}{\tilde{\sigma}}}}{1+\left(\frac{A_{L,c,s,m}Q_{L,c,m}\widetilde{L}^{\text{hours}}_{c,s,m}}{A_{H,c,s,n}Q_{H,c,n}\widetilde{H}^{\text{hours}}_{c,s,n}}\right)^{\frac{\tilde{\sigma}-1}{\tilde{\sigma}}}}\right]^{\frac{1-\alpha\tilde{\sigma}}{\tilde{\sigma}-1}} \\
w^C_{L,c,s,m} &= w_{L,c,s,m}\left(\frac{\widetilde{L}^{C,\text{hours}}_{c,s,m}}{\widetilde{L}^{\text{hours}}_{c,s,m}}\right)^{-\alpha}\left[\frac{1+\left(\frac{A_{H,c,s,n}Q_{H,c,n}\widetilde{H}^{C,\text{hours}}_{c,s,n}}{A_{L,c,s,m}Q_{L,c,m}\widetilde{L}^{C,\text{hours}}_{c,s,m}}\right)^{\frac{\tilde{\sigma}-1}{\tilde{\sigma}}}}{1+\left(\frac{A_{H,c,s,n}Q_{H,c,n}\widetilde{H}^{\text{hours}}_{c,s,n}}{A_{L,c,s,m}Q_{L,c,m}\widetilde{L}^{\text{hours}}_{c,s,m}}\right)^{\frac{\tilde{\sigma}-1}{\tilde{\sigma}}}}\right]^{\frac{1-\alpha\tilde{\sigma}}{\tilde{\sigma}-1}}
\end{aligned} \tag{3}$$

I set $\alpha = 1/3$, and compute counterfactual wages for all educational groups using (3). With these at hand, I compute the counterfactual skill premium and relative skill supply as

$$\frac{w^C_{H,c}}{w^C_{L,c}} = \frac{\prod_s \left(w^C_{H,c,s,\text{tertiary}}\right)^{\frac{\widetilde{H}^{C,\text{bodies}}_{c,s,\text{tertiary}}}{\widetilde{H}^{\text{bodies}}_{c,\text{tertiary}}}}}{\prod_s \left(w^C_{L,c,s,\text{upper sec}}\right)^{\frac{\widetilde{L}^{C,\text{bodies}}_{c,s,\text{upper sec}}}{\widetilde{L}^{\text{bodies}}_{c,\text{upper sec}}}}}$$

$$\frac{\widetilde{H}^C_c}{\widetilde{L}^C_c} = \frac{\frac{w^C_{H,1,c}}{w^C_{H,c,\text{tertiary}}}\widetilde{H}^{\text{hours}}_{1,c} + \ldots + \frac{w^C_{H,N,c}}{w^C_{H,c,\text{tertiary}}}\widetilde{H}^{\text{hours}}_{N,c}}{\frac{w^C_{L,1,c}}{w^C_{L,c,\text{upper sec}}}\widetilde{L}^{\text{hours}}_{1,c} + \ldots + \frac{w^C_{L,M,c}}{w^C_{L,c,\text{upper sec}}}\widetilde{L}^{\text{hours}}_{M,c}}$$

that can be used to calculate RSE^C_c, as discussed in the paper. Table D.3 displays the baseline and counterfactual skill premium and relative supply. Compared to the corresponding baseline quantities, the counterfactual skill premium is more variable across countries (its elasticity with respect to GDP per worker approximately doubles). Intuitively, the counterfactual reallocation results, for poorer countries, in higher employment shares in sectors with high relative skill efficiency, which contribute to a higher overall skill premium. The counterfactual relative supply of skilled labor is instead similar

[2]As discussed in the paper, $w_{H,c,n}$ and $w_{L,c,m}$ are computed on a restricted sample of workers (wage workers with relatively high labor market attachment). As the sectoral shares of employment in this subsample are not exactly equal to $\frac{\widetilde{H}^{\text{bodies}}_{c,s,n}}{\widetilde{H}^{\text{bodies}}_{c,n}}$ and $\frac{\widetilde{L}^{\text{bodies}}_{c,s,m}}{\widetilde{L}^{\text{bodies}}_{c,m}}$, which are calculated over the whole population of employed workers, the expressions in (2) hold up to a small approximation error.

[3]Assumption (A4) allows me to abstract from any change in the sector-level average human capital following the reallocation of labor across sectors.

to the baseline one, since relative wages within skill groups are less impacted by the reallocation across sectors. Overall, this results in a marginally lower cross-country variation in the counterfactual aggregate relative skill efficiency compared to the baseline case, as shown in Table IV in the paper.

Table D.3: Counterfactual Skill Premium, Supply and Efficiency across Countries

	Baseline		Counterfactual	
Country	w_H/w_L	$\tilde{H}/\tilde{L}$	w_H^C/w_L^C	$\tilde{H}^C/\tilde{L}^C$
India	2.230	0.205	2.927	0.165
Indonesia	1.957	0.070	2.664	0.060
Jamaica	2.969	0.067	3.271	0.065
Brazil	3.412	0.158	3.508	0.152
Venezuela	2.490	0.257	2.655	0.234
Uruguay	2.218	0.363	2.277	0.357
Panama	2.262	0.313	2.400	0.301
Mexico	2.202	0.226	2.302	0.220
Trinidad and Tobago	2.741	0.101	2.803	0.099
Israel	1.606	0.596	1.575	0.598
Canada	1.508	1.539	1.518	1.558
United States	1.802	1.397	1.809	1.400
Elasticity wrt GDP p.w.	-0.138	0.911***	-0.243***	0.982***
	(0.077)	(0.244)	(0.062)	(0.237)

Notes: The Table shows the skill premium and relative skill supply across the countries in the micro-data sample. *Baseline* refers to the baseline specification, while *Counterfactual* refers to the counterfactual exercise where sectoral shares of employment are equalised across countries. The last row shows the coefficient of a regression of the log of each variable on log GDP per worker (standard errors in parentheses).

References

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

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer, "The Next Generation of the Penn World Table," *American Economic Review*, 2015, *105* (10), 3150–3182.

Montenegro, Claudio E. and Harry Anthony Patrinos, "Comparable estimates of returns to schooling around the world," Policy Research Working Paper Series 7020, The World Bank September 2014.

Oberfield, Ezra and Devesh Raval, "Micro Data and Macro Technology," NBER Working Papers 20452, National Bureau of Economic Research, Inc September 2014.