Skills, Technologies, and Development
Appendix

Juan Ignacio Vizcaino

1 ILO Data Description, Treatment, and Issues.

The ILO's database is constructed from multiple data sources, including establishment surveys, household surveys, insurance records, and administrative data sources. Data sources vary across countries, and when multiple sources are available, ILO presents all the options available. In such case, I pick data from the source I consider most reliable [1]. I discard data from sources that the ILO flags as unreliable, even if that is the only one available for a country.

The ILO data is harmonized, both at the sectoral and occupational level, allowing for comparability across countries. Occupational data is harmonized based on the International Standard Classification of Occupations (ISCO). Statistics on employment by occupation are presented in ILOSTAT according to both the categories of the latest version of the ISCO available (ISCO-08 and ISCO-88). When both versions are available, I take the latest revision (ISCO-08). I take the earlier version (ISCO-88) when it is the only one available and bridge it into the newer (ISCO-08) using the correspondence table provided by ILO [2].

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

[2] Link to ILO's occupations documentation.

B Countries Intensive in Natural Resources

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

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

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

C Hours Worked Across Development.

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

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

Average Hours Worked by Sector Across Development

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

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

Note I : B,F-S,L corresponds to the data in Bick, Fuchs-Schundeln, and Lagakos. In their paper Industry is called Manufacturing. The sectors included in it are roughly the same except for Construction, which I include in Industry and it is not clear to me if they consider it to be in Manufacturing or Services.
Note II : To compute GDP per capita and development percentiles I use as GDP measure PWT's *rgdpo*. I take the average GDP per capita for the period 2005-2014. B,F-S,L use as GDP measure PWT's *rgdpe*, and compute terciles for 2005. I follow their procedure to construct this table.
Note III : B,F-S,L focus on 49 core countries while I have data for 88 countries.

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

Average Hours Worked by Sector Across Development

-using ILO data and four broad sectors-

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

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

Average Hours Worked by Occupation Across Development

-using ILO data-

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

D Robustness Checks.

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

Different Criteria for Constructing Broad Occupational Groups.

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

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

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

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

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

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

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

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

Extreme Values.

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

The Role of the Economic Environment and Institutions.

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

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

place [33].

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

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

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

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

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

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

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

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

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

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

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

Relative Total Labor Income versus Relative Hourly Labor Income.

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

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

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

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

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

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

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

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

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

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

Workers of Both Sexes vs Males Only.

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

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

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

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

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

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

E Proofs

Proposition 1

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

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

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

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

Average labor productivity in high-skill occupations is

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

in the independent case

$$\mathbb{E}\big(Z_h|w_h z_h > w_l z_l\big) = \left(\frac{1}{\pi_h}\right) \int_0^{+\infty} \int_0^{\left(\frac{w_h}{w_l}\right) z_h} z_h g_h(z_h) g_l(z_l) dz_l dz_h.$$

Under common shape parameter

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

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

$$\mathbb{E}\big(Z_h|w_h z_h > w_l z_l\big) = \int_0^{+\infty} \left(\frac{y}{S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}}\right)^{-\left(\frac{1}{\theta}\right)} e^{-y} dy,$$

$$\mathbb{E}\big(Z_h|w_h z_h > w_l z_l\big) = \left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right)^{\left(\frac{1}{\theta}\right)} \Gamma\left(1 - \frac{1}{\theta}\right)$$

$\square$

Proposition 2 (Educational Sorting Conditional on Occupational Choice).

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

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

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

Workers choose no-schooling if:

$$\begin{aligned}
W^{ns}_h &\geqslant W^s_h - c^s \\
w_h z_h &\geqslant w_h z_h e^{\beta^s} - c^s \\
z_h &\leqslant \frac{c^s}{w_h\left(e^{\beta^s} - 1\right)} = z^\star_h,
\end{aligned}$$

and

$$\begin{aligned}
W^{ns}_h &\geqslant W^u_h - c^u \\
w_h z_h &\geqslant w_h z_h e^{\beta^u_h} - c^u \\
z_h &\leqslant \frac{c^u}{w_h\left(e^{\beta^u_h} - 1\right)} = \tilde{z}^\star_h.
\end{aligned}$$

Secondary Education is chosen if

$$\begin{aligned}
W^s_h - c^s &> W^{ns}_h \\
w_h z_h e^{\beta^s} - c^s &> w_h z_h \\
z_h &> \frac{c^s}{w_h\left(e^{\beta^s} - 1\right)} = z^\star_h,
\end{aligned}$$

and

$$\begin{aligned}
W^s_h - c^s &\geqslant W^u_h - c^u \\
w_h z_h e^{\beta^s} - c^s &\geqslant w_h z_h e^{\beta^u_h} - c^u \\
z_h &\leqslant \frac{c^u - c^s}{w_h\left(e^{\beta^u_h} - e^{\beta^s}\right)} = z^{\star\star}_h.
\end{aligned}$$

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

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

and

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

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

$$\left(\frac{c^u}{c^s}\right) > \left(\frac{\beta_h^u - 1}{\beta^s - 1}\right)$$

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

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

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

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

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

$$
\left(\frac{c^u}{c^s}\right) > \left(\frac{\beta_l^u-1}{\beta^s-1}\right),
$$

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

□

Proposition 3

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

1. If $z_l \leqslant z_l^{\star}$ workers choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right) z_l$.
2. If $z_l \in (z_l^{\star}, \tilde{z}_l^{\star\star}]$ workers choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right) z_l$.
3. If $z_l \in (\tilde{z}_l^{\star\star}, z_l^{\star\star}]$ workers choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right)\left(\frac{z_l e^{\beta^s}}{e^{\beta_h^u}}\right)$.
4. If $z_l > z_l^{\star\star}$ workers choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right)\left(\frac{z_l e^{\beta_l^u}}{e^{\beta_h^u}}\right)$.

The proof goes by contraposition.

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

Then there exists a $z_l \in [0, z_l^{\star}]$ such that

$$
w_h z_h \leqslant w_l z_l
$$

and

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

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

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

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^s = w_h z_h e^{\beta^s}$, $z_h \in (z_h^{\star}, z_h^{\star\star}]$, which together with $w_l z_l \geqslant w_h z_h \implies z_l > \left(\frac{w_h}{w_l}\right) z_h^\star = z_l^\star$, which is a contradiction.

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

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

Then there exists a $z_l \in (z_l^\star, \tilde{z}_l^{\star\star}]$ such that

$$w_h z_h \leqslant w_l z_l$$

and

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

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

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

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

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

Then there exists a $z_l \in (\tilde{z}_l^{\star\star}, z_l^{\star\star}]$ such that

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

and

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

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

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

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

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

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

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

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

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

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

$$w_h z_h e^{\beta_h^u} \leqslant w_l z_l e^{\beta_l^u}$$

and

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

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

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

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

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

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

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

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

Proposition 4

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

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

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

F Census/American Community Survey data issues and handling.

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

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

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