A Supplemental Appendix – Cross-country income dispersion, human capital, and technology adoption

Pedro Amaral and Alberto Rivera-Padilla

A.1 Data appendix

This section provides more details on the microdata (IPUMS-International) used in the paper. We focus on countries that have individual data on education, employment status, and occupation. We use the latest available sample for each country based on these data requirements. Every sample is from a year within the last two decades (2000 or later) and the original data source is either a census or a survey. We map educational attainment to education years - between 0 and 16 - following the literature (see e.g., Lagakos, Moll, Porzio, Qian, and Schoellman 2018). Employment status in the data refers to individuals being categorized mainly as employers, own-account workers, wage/salary workers, or unpaid workers. Occupations are based on 1-digit codes following the ISCO.

As explained in the main text, one of our main goals is to combine data on occupations and employment status to create three categories of employment that are consistent with occupational choices in our model. For our cross-country analysis in Section 1, we use data from countries for which it is possible to distinguish self-employment status as own-account employment, employee, or another category.[32] Countries with a generic category for self-employment are not useful for our analysis as they do not allow us to distinguish between traditional and modern forms of entrepreneurship.[33] Moreover, we do not consider individuals who are categorized as unpaid workers, who are working for the armed forces, or who have an unspecified occupation.

Table 7 presents aggregate summary statistics of our data by region of the world. These are unweighted statistics for each region calculated by taking the average across the values obtained for each country. The table shows that we have a good number of developing and poor economies in our data.

Table 7: Aggregate Summary Statistics

Region	No. Countries	Education Years	Share of Wage/Salary Workers
Africa	16	5.1	0.41
Asia	11	7.5	0.50
Europe	8	11.2	0.77
Latin America	15	7.3	0.55
USA/Canada	2	12.9	0.85

Source: Authors' estimations using IPUMS-International data.

[32] There are a few categories for which we need to take a stand on their classification, and we use the structure of our model as a guide. Sharecroppers who are self-employed are classified as own-account workers, while sharecroppers who are employees and members of cooperatives are classified as wage and salary workers. All of these categories represent a small fraction of the population in all regions where they are observed.

[33] A particular case is the United States, which reports self-employment based on incorporated and unincorporated status. To include the United States as our key reference, we split self-employment along the occupational dimension (managers/professionals vs the rest), as explained in the main text. This yields the most comprehensive categorization of modern managers consistent with our model. For instance, defining modern managers solely as incorporated self-employed individuals would result in a smaller share of modern managers.

Table 8: Summary Statistics by Country

Country	GDP pc	N_o	N_w	N_m	Educ. Yrs	Tech Lag
CHE	1.20	0.08	0.79	0.13	13	9.6
IRL	1.10	0.12	0.76	0.13	12	19.1
USA	1.00	0.07	0.86	0.06	13	0.0
CAN	0.87	0.05	0.84	0.11	13	4.5
ITA	0.82	0.10	0.74	0.16	11	20.5
FRA	0.81	0.06	0.85	0.10	12	15.6
ESP	0.71	0.07	0.82	0.11	11	24.0
PRT	0.59	0.06	0.80	0.14	9	27.3
GRC	0.57	0.19	0.66	0.15	11	27.7
POL	0.44	0.18	0.73	0.08	11	36.4
MYS	0.39	0.26	0.65	0.09	7	29.9
CHL	0.39	0.13	0.76	0.11	9	36.4
URY	0.39	0.24	0.60	0.15	8	44.9
PAN	0.38	0.32	0.64	0.04	9	41.6
TUR	0.34	0.33	0.60	0.07	8	43.3
MEX	0.34	0.20	0.74	0.06	10	39.7
MUS	0.31	0.14	0.79	0.07	8	35.3
CRI	0.30	0.25	0.62	0.13	8	37.2
BRA	0.28	0.23	0.68	0.08	8	40.2
THA	0.27	0.53	0.40	0.07	6	42.2
BWA	0.25	0.13	0.81	0.06	8	46.1
IRN	0.22	0.49	0.42	0.09	8	48.4
PRY	0.20	0.49	0.42	0.09	6	46.1
EGY	0.20	0.04	0.89	0.06	7	48.7
ECU	0.19	0.32	0.61	0.07	8	43.0
PER	0.19	0.42	0.51	0.06	10	43.9
MNG	0.17	0.52	0.44	0.04	10	41.1
GTM	0.16	0.39	0.46	0.15	5	39.5
JOR	0.16	0.10	0.81	0.09	11	34.0
IDN	0.15	0.53	0.40	0.07	8	49.3
SLV	0.14	0.30	0.61	0.09	6	36.8
BOL	0.13	0.48	0.43	0.09	9	33.3
MAR	0.12	0.41	0.52	0.07	4	44.9
PHL	0.11	0.30	0.59	0.11	9	43.9
NIC	0.10	0.44	0.52	0.04	5	39.0
HND	0.09	0.45	0.46	0.09	4	42.8
IND	0.08	0.37	0.53	0.10	7	49.9
GHA	0.08	0.59	0.32	0.09	7	47.7
CMR	0.07	0.66	0.32	0.03	6	46.1
SDN	0.07	0.36	0.45	0.19	4	51.8
KHM	0.06	0.71	0.27	0.02	5	38.8
SEN	0.06	0.66	0.23	0.11	3	47.3
HTI	0.05	0.81	0.16	0.04	4	53.0
NPL	0.05	0.61	0.35	0.05	5	41.9
ZMB	0.05	0.57	0.39	0.03	8	45.4
BEN	0.05	0.78	0.15	0.07	4	51.3
GIN	0.04	0.89	0.08	0.03	2	46.0
TZA	0.04	0.75	0.19	0.06	6	52.8
LSO	0.04	0.31	0.68	0.02	5	45.8
MLI	0.04	0.63	0.35	0.02	1	56.4
TGO	0.03	0.76	0.19	0.05	5	48.6
RWA	0.03	0.73	0.26	0.01	4	60.4

Notes: N_j refer to employment shares in each occupation category defined in the main text. Tech. Lag is calculated with data from Comin, Hobijn, and Rovito (2008). We only report lags for countries that have data for at least 5 of the major technologies considered. GDP per capita is expressed relative to the U.S.

Table 8 presents more detailed summary statistics by country, together with the data on technology usage lags from Comin, Hobijn, and Rovito (2008). These are the country-level estimates used in figures of Section 1.1 and in the quantitative application of the model for cross-country comparisons.

A.2 Industries technology ranking

The following table presents the ranking of industries based on usage of modern technologies.

Table 9: Technology Use Ranking by Industry in the United States

Industry	Technology use intensity	Rank
Financial services and insurance	47	1
Health and social work	40	2
Business services and real estate	40	3
Education	38	4
Manufacturing	35	5
Electricity, gas, water and waste management	34	6
Transportation, storage, and communications	31	7
Other services	26	8
Wholesale and retail trade	25	9
Construction	22	10
Mining and extraction	21	11
Hotels and restaurants	20	12
Agriculture, fishing, and forestry	18	13

Notes: Industries are classified according to IPUMS International. Technology use intensity represents the percentage of firms that use modern technologies based on ABS 2019 data. See main text for details.

A.3 Education quality adjustment

Table 10 presents the differences between raw years of schooling and quality-adjusted years of schooling for the United States and each income quartile. This is the data used in the exercise presented in Figure 14 in the main text.

We follow the same harmonization process the World Bank uses in computing its Human Capital Index in World Bank (2020b). These are harmonized scores across major international student achievement testing programs measured in Trends in International Mathematics and Science Study (TIMSS)-equivalent units from Patrinos and Angrist (2018), where 300 is minimal attainment and 625 is advanced attainment. In computing the adjustment factor, we normalize by 625. We do this for each country in our sample and then average the values in each quartile.

Table 10: Quality-adjusted years of schooling

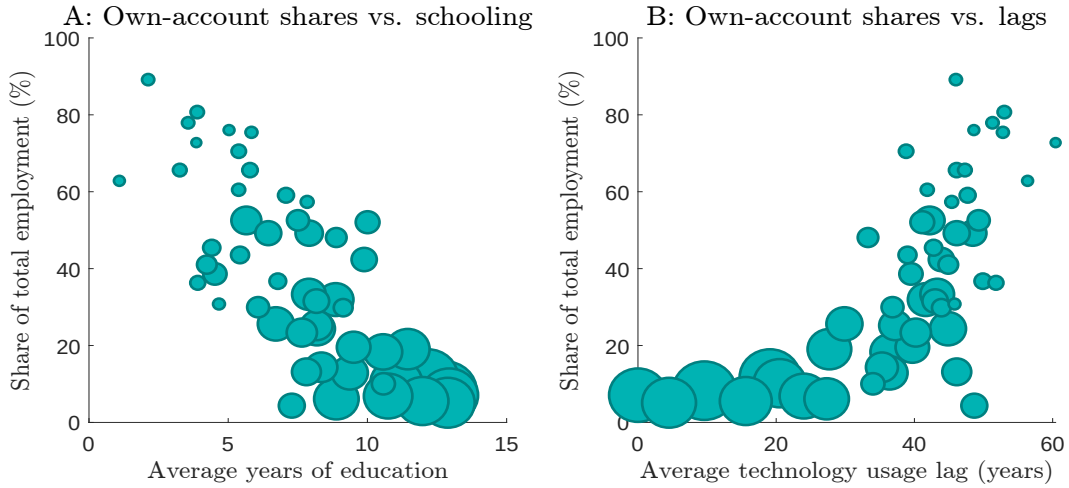
	Raw	Adjusted	Adjustment factor
USA	12.93	10.59	0.82
4th quartile	10.75	8.51	0.79
3rd quartile	7.97	5.32	0.67
2nd quartile	6.95	4.38	0.63
1st quartile	4.29	2.61	0.61

Notes: Data from Patrinos and Angrist (2018) via World Bank (2020b) and authors' calculations.

A.4 Occupational choices, education years, and technology usage lags

This section provides additional evidence on the importance of the organization of production and occupational choices in mediating between education and technology adoption. In Panel A of Figure 15 we show the fraction of own-account employment against average schooling attainment across countries. There is a strong negative correlation between these two variables. It is worth highlighting that the group of countries with large shares of own-account employment and low average education also tend to be poorer – circles are proportional to income per capita.

Figure 15: The importance of occupational choices



Panel B of Figure 15 illustrates the strong positive relationship between average technology usage lags and the fraction of own-account employment across countries. Note how countries that have around half or more of their labor force in own-account production tend to have a technology usage lag of more than 40 years.

Table 11: Own-account Employment and Technology Usage Lag

	All	Internet	PCs	Electricity	Aviation-Cargo	Tractors	Cellphones
Own-account (%)	0.363	0.150	0.198	0.676	0.462	0.386	0.159
	(0.0661)	(0.0173)	(0.0344)	(0.104)	(0.0935)	(0.0881)	(0.0157)
Observations	52	49	48	51	40	52	52

Notes: Robust standard errors are in parenthesis. Employment shares are estimated with IPUMS-International data and technology usage lags are obtained from Comin, Hobijn, and Rovito (2008).

In Table 11, we confirm the strength and statistical significance of this relationship by running simple cross-country regressions of technology usage lags on own-account employment shares. The positive correlation between both variables is weaker for digital age technologies, such as the Internet, but still meaningful (and always statistically significant). In Table 12 we do a similar estimation using educational attainment instead of own-account employment. The results confirm the negative correlation between education and technology usage lags across countries.

Table 12: Education and Technology Usage Lags

	All	Internet	PCs	Electricity	Aviation-Cargo	Tractors	Cellphones
Average Education	-3.414	-1.161	-1.744	-6.406	-3.811	-3.865	-1.270
	(0.478)	(0.111)	(0.261)	(0.725)	(0.613)	(0.778)	(0.109)
Observations	52	49	48	51	40	52	52

Notes: Robust standard errors are in parenthesis. Educational attainment is estimated with IPUMS-International data and technology usage lags are obtained from Comin, Hobijn, and Rovito (2008).

A.5 Validation of model predictions

In this section, we provide evidence that validates key predictions of the model, particularly those in Proposition 1. We use microdata to show that, at the establishment level, there is a positive and statistically significant correlation between manager education and: (i) the average education of non-management workers; (ii) the percentage of the workforce that uses computer regularly; and (iii) the establishment size. To do this, we use data from the 2005 Enterprise Surveys from India. We use India to remain consistent with the evidence presented in Section 1.2. Moreover, the 2005 surveys are useful because they contain information on the variables needed to assess the model's predictions.

Table 13: Evidence on Model Predictions from Establishment Data in India

	Non-managers education	Workforce computer use	Establishment size
	(1)	(2)	(3)
top manager with college education	0.643***	7.528***	0.435***
	(0.143)	(1.009)	(0.066)
Observations	1,966	2,204	2,080
avg. education management workers	0.394***	1.101***	0.067***
	(0.027)	(0.165)	(0.011)
Observations	1,788	1,786	1,780

Notes: the table presents results from multiple regressions at the establishment level. Column (1) refers to the average years of education for non-management workers. Column (2) refers to the percentage of the workforce that uses computer regularly. Column (3) refers to the average number of permanent workers in a year (in logs). All cases control for establishment age, location, and industry. Robust standard errors are reported. *$p < 0.1$, ** $p < 0.05$, ***$p < 0.01$. Source: Authors' estimations using WBES data.

Table 13 presents results from regressions of the three outcomes mentioned above on two alternative measures of managers' human capital. First, using a categorical variable, we consider a dummy equal to one if the top manager of the entire firm has college education. Second, we use the average years of education reported for all management workers in the establishment. In every regression we control for establishment age, industry, and location. The results indicate that the coefficient of interest is positive and statistically significant in all specifications.

A.6 Sensitivity analysis

In this section we conduct various exercises related to sensitivity analysis.

A.6.1 Elasticity of targets with respect to parameters

We start by asking how sensitive the moments used to calibrate the model are to each of the model's parameters. Table 14 reports the elasticities of the various targets (in the columns) with respect to the parameters (in the rows) at our benchmark calibration, where all targets are exactly matched.

Table 14: Sensitivity Matrix

	% workers	% managers	E to GDP	Mincer (all)	Mincer (trad.)	SD(log. empl.)	VAR(earnings)
η	-0.260	**4.287**	-1.340	-1.158	-0.002	-0.409	-2.964
σ	0.296	-1.650	-0.416	-0.082	-0.002	0.212	**2.048**
ψ	**-0.139**	-0.073	-0.027	-0.085	-0.002	-0.024	-0.004
θ	0.024	-0.301	**0.978**	0.001	0.022	-0.020	0.129
ϕ	-0.030	-0.029	0.048	**1.014**	1.014	-0.010	0.002
λ	-0.205	0.489	-0.022	-0.045	-0.002	**0.508**	0.061
ρ	0.059	0.004	0.012	0.054	**1.038**	0.003	0.045

Notes: Values are elasticities (in percentage) from target (column) with respect to parameter (row). The bold numbers represent the largest change in absolute value for each parameter.

The targets we selected are all sensitive to the model's parameters, and the elasticities support the identification strategy discussed in the main text. That is, although the parameters are jointly calibrated with a simulated method of moments, each parameter is particularly informative about one targeted moment. For example, the share of managers has a relatively large response to the value of the parameter η, which governs the returns to scale in modern production. Similarly, the variance of earnings is particularly sensitive to the value of σ, which itself governs the variance of individual ability.

We can further assess identification using the elasticity matrix in Table 14. The matrix has full column rank, implying that the parameters are locally identified. Moreover, its singular values are well separated from zero, suggesting that different parameters affect the targeted moments in sufficiently distinct ways. In particular, there is no direction in which changes in parameters leave the moments almost unchanged, indicating that the parameters are reasonably well disciplined by the chosen set of moments.

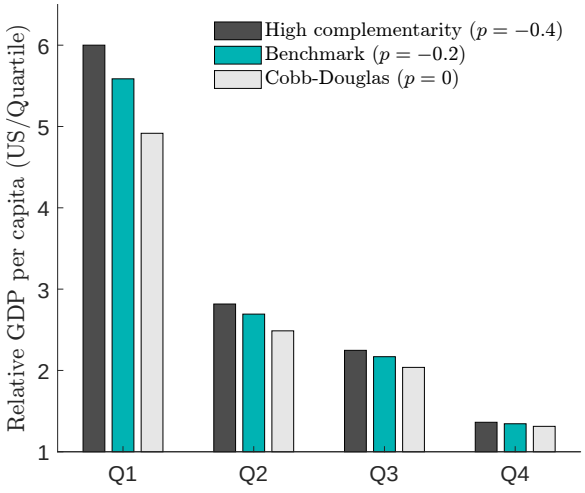
A.6.2 Complementarity between worker and manager productivity

Recall that our estimate for p is not calibrated to match a target, and was instead obtained from Abowd, Kramarz, Pérez-Duarte, and Schmutte (2018), who estimate it directly from U.S. data described in Section 3. It is therefore important to understand how sensitive the cross-country income dispersion we obtain is to this parameter. To do this we consider two alternative values for p, one above and one below what we use in our benchmark calibration ($p = -0.208$). For a production function with more substitutability, we consider the standard Cobb-Douglas ($p = 0$), and for one with stronger complementarity we use $p = -0.4$. In both cases we recalibrate all the other parameters to continue to hit the same targets. Figure 16 compares the amplification we get under the three

different calibrations when we only change education.[34]

While it is not the case that skill complementarity in production is necessary to obtain a substantial amount of income dispersion – as the Cobb-Douglas case shows – it does generate much more amplification (for a given level of human capital). Models of cross-country income dispersion that fail to feature this channel therefore risk underestimating income differences. Moreover, complementarity in skills gets amplified by the fact that more productive units, in terms of human capital, also operate with better technologies, and the extent to which they do depends crucially on the parameter η, which we now turn to.

Figure 16: Sensitivity to p



A.6.3 Convexity of adoption costs

In Section 2, we used Equation (5) to argue that η, the parameter governing the convexity of adoption costs, is key for amplification: given human capital complementarity in production $f(x, y)$, the smaller η, the more productive the technology a manager adopts. Here, we confirm that this channel is quantitatively important for cross-country income differences.

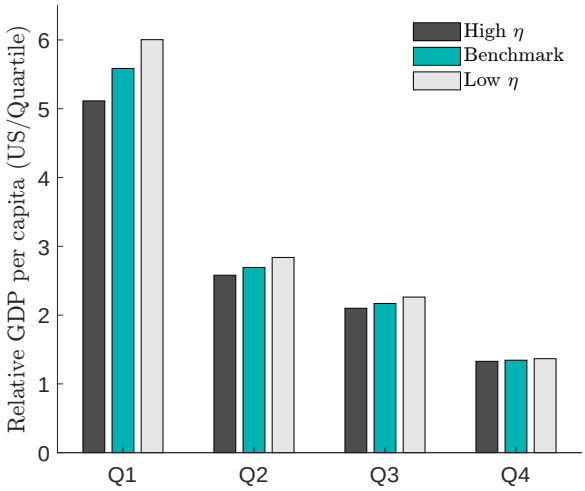
Unlike what is the case for p, the value of η is chosen to match (joint) targets. As we experiment with values 10 percent above and below our benchmark value of $\eta = 1.5$, we opted for keeping all the same targets, resulting in an over-identified system.[35] Note, from Table 14 that all but one of the targets is highly sensitive to η, especially the share of modern managers. As discussed above, we think the value of η is pinned down credibly by our calibration strategy. Our goal in this exercise is simply to ascertain how important this channel is in driving cross-country income differences.

As Figure 17 shows, this parameter is rather important, and quantitatively similar to skill complementarity.

[34] Had we changed barriers at the same time, the calibrated values for τ that would match the share of traditional workers in each specification would be different, which would confound the pure effect of changing p.

[35] In practice, all targets are exactly matched, except for the standard deviation of log employment and the variance of log earnings, which are within 10 percent of their target values.

Figure 17: Sensitivity to η



A.6.4 Occupation-specific taste shocks

Recall that λ represents the dispersion (scale) of the occupation-specific preference shock. Again, unlike what is the case for p, and just like what happens with η, the value of λ is chosen to match (joint) targets. As we experiment with values 10 percent above and below our benchmark value of $\lambda = 0.497$, we keep all the same calibration targets, resulting in an over-identified system. Nonetheless, all parameters except the standard deviation of log employment are matched exactly.[36]

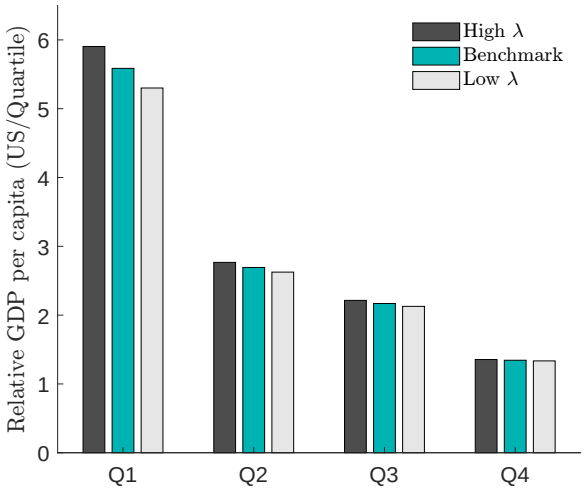
As λ increases and the dispersion of the occupation-specific preference shocks rises, the odds that an individual draws a shock that is large enough to bump them into an occupation that is not the most pecuniarily attractive one increase. Raising λ blunts pecuniary motives and reduces the importance of human capital in occupational choice, bringing occupational shares, for each human capital type, closer together. As Table 14 shows, λ is "well-identified" by the moments we choose to target. Not only the two occupational shares, but also the standard deviation of log-employment are all sensitive to it.

As Figure 18 shows, income dispersion is not very sensitive to the choice of λ.[37] To the extent that it is, a higher λ increases income dispersion. Recall from our discussion in Section 3, that our calibration does a fairly reasonable job at capturing untargeted moments of the establishment distribution. It can deliver a large number of small establishments, but it falls short in generating a long enough right tail of the establishment size distribution. It fails to capture very large and very productive establishments. Everything else being the same, the more right-skewed the establishment distribution is in the U.S. benchmark, the larger the dispersion will be, since these sort of establishments will be much smaller in less-educated, barrier-ridden, economies. Increasing λ, it turns out, makes for a more right-skewed distribution. Not because the larger establishments become larger or more productive, but because they constitute a larger share, one that is closer to the data. Therefore, we argue that our choice of targets and the resulting calibration strike a fair balance, were we to directly target the share of large establishments, the resulting amplification would be even larger.

[36] The target for the standard deviation of log employment is 1.3. We get 1.22 for the low λ and 1.39 for the high λ.

[37] As in the previous section, we only change education, and not barriers, because the calibrated values for τ that would match the share of traditional workers for each λ would be different, which would confound the pure effect of changing λ.

Figure 18: Sensitivity to λ



A.6.5 Sensitivity to U.S. managerial share

As Table 8 shows, the U.S. has a relatively low managerial share compared to other advanced economies. Since we calibrate our model economy to this value, it is natural to wonder how this choice affects results, compared to targeting the average occupational shares of the richest countries in our sample. We picked the 4 richest countries (Switzerland, Ireland, USA and Canada) to average over, as opposed to a larger sample like the top quartile since this would include countries that are a lot poorer than the U.S., like Malaysia, for example, that has a share of own-account workers of 26 percent. These new calibration targets are a managerial share of 11 percent and a salaried worker share of 81 percent. All the other targets remain the same. As a result of this calibration, the parameter η, which we have highlighted as the most important for amplification (see Equation 5), increases from 1.5 in the benchmark to 2.5, implying less amplification.

In a world where technology adoption is barrier-free ($\tau = 0$), lowering schooling years from U.S. levels to the average in the lowest quartile, implies an income factor difference of 4.1, as opposed to 5.6 in the benchmark. While lower, this is still a very high value in the context of the literature considering we are just varying education and not changing the cost of adoption (τ).[38]

If, in addition to changing schooling, we also calibrate τ to match the share of own-account workers in the poorest quartile, the income dispersion increases from 9 to 10.5. This happens because, despite the higher η, τ needs to increase from 0.4 in the benchmark to 4.8 under this new calibration to match the large own-account share. The model thus requires adoption costs that are 5.8 times higher in poor countries, compared to the U.S. While this might seem like a large increase, it is reasonable in light of the literature.[39] Notably, a substantially larger value of τ is still required to match the

[38] This compares well, for example, with the magnitudes obtained by Buera, Hopenhayn, Shin, and Trachter (2026). When varying distortions only ($\xi = 0.2$) in their calibration for India, they obtain an income factor difference of $2.5 = 1/0.4$ (see their Figure 3 and the discussion around it).

[39] Consider, again, Buera, Hopenhayn, Shin, and Trachter (2026). They attempt to capture the income gap between the U.S. and India by changing not only the degree of distortions, but also the technology adoption cost (among other parameters), which they calibrate to be 17 times higher in India (see their Table III) to get an income ratio of 6.7. While the models are not directly comparable, it does put our factor of 5.8 in a more charitable light, especially given the larger income gap we obtain.

share of own-account workers in the poorest quartile in the absence of schooling gaps.

A.7 Profit maximization in modern production

Recall from equation (3) that the problem of a modern sector manager is:

$$\Pi_m(y;\boldsymbol{\omega},\tau) = \max_{v(x),l(x)} \int_x \left(v(x) \left[\alpha_m y^p + (1-\alpha_m)x^p\right]^{1/p} l(x)^\alpha - l(x)\omega(x) - (1+\tau)\frac{v(x)^\eta}{\eta} \right) dx,$$

Letting $f(x,y) = \left[\alpha_m y^p + (1-\alpha_m)x^p\right]^{1/p}$, the F.O.C. w.r.t. $v(x)$ is:

$$f(x,y)l(x)^\alpha = (1+\tau)v(x)^{\eta-1},$$

and therefore, the optimal technology level is:

$$v(x,y) = \left(\frac{f(x,y)l(x)^\alpha}{1+\tau} \right)^{\frac{1}{\eta-1}}. \tag{9}$$

Plugging this back into the objective:

$$\Pi_m(y;\boldsymbol{\omega},\tau) = \max_{l(x)} \int_x \left(\left[\frac{(f(x,y)l(x)^\alpha)^\eta}{1+\tau} \right]^{\frac{1}{\eta-1}} - l(x)\omega(x) - \frac{1}{\eta}\left[\frac{(f(x,y)l(x)^\alpha)^\eta}{1+\tau} \right]^{\frac{1}{\eta-1}} \right) dx,$$

$$\Pi_m(y;\boldsymbol{\omega},\tau) = \max_{l(x)} \int_x \left(\frac{\eta-1}{\eta}\left[\frac{(f(x,y)l(x)^\alpha)^\eta}{1+\tau} \right]^{\frac{1}{\eta-1}} - l(x)\omega(x) \right) dx,$$

The F.O.C. w.r.t. $l(x)$ is:

$$\omega(x) = \alpha \left(\frac{[f(x,y)]^\eta}{1+\tau} \right)^{\frac{1}{\eta-1}} l(x)^{\frac{1+\eta(\alpha-1)}{\eta-1}}$$

Solving for $l(x)$, we get the optimal demand for workers of type x:

$$l^d(x,y) = \left(\frac{\alpha}{\omega(x)} \right)^{\frac{\eta-1}{\eta(1-\alpha)-1}} \left(\frac{f(x,y)^\eta}{1+\tau} \right)^{\frac{1}{\eta(1-\alpha)-1}} \tag{10}$$

Plugging this back into the objective once more:

$$\Pi_m(y;\boldsymbol{\omega},\tau) = \int_x \left(\frac{\eta-1}{\eta} \left(\frac{f(x,y)^\eta}{1+\tau} \left(\frac{\alpha}{\omega(x)} \right)^{\frac{\alpha\eta(\eta-1)}{\eta(1-\alpha)-1}} \left(\frac{f(x,y)^\eta}{1+\tau} \right)^{\frac{\alpha\eta}{\eta(1-\alpha)-1}} \right)^{\frac{1}{\eta-1}} - \omega(x) \left(\frac{\alpha}{\omega(x)} \right)^{\frac{\eta-1}{\eta(1-\alpha)-1}} \left(\frac{f(x,y)^\eta}{1+\tau} \right)^{\frac{1}{\eta(1-\alpha)-1}} \right) dx.$$

$$\Pi_m(y;\boldsymbol{\omega},\tau) = \int_x \left(\frac{\eta-1}{\eta} \left(\frac{f(x,y)^\eta}{1+\tau} \right)^{\frac{1}{\eta(1-\alpha)-1}} \left(\frac{\alpha}{\omega(x)} \right)^{\frac{\alpha\eta}{\eta(1-\alpha)-1}} - \omega(x) \left(\frac{\alpha}{\omega(x)} \right)^{\frac{\eta-1}{\eta(1-\alpha)-1}} \left(\frac{f(x,y)^\eta}{1+\tau} \right)^{\frac{1}{\eta(1-\alpha)-1}} \right) dx.$$

$$\Pi_m(y;\boldsymbol{\omega},\tau) = \int_x f(x,y)^{\frac{\eta}{\eta(1-\alpha)-1}} \left[\frac{1}{(1+\tau)\omega(x)^{\eta\alpha}} \right]^{\frac{1}{\eta(1-\alpha)-1}} \left[\left(\frac{\eta-1}{\eta} \right) \alpha^{\frac{\alpha\eta}{\eta(1-\alpha)-1}} - \alpha^{\frac{\eta-1}{\eta(1-\alpha)-1}} \right] dx.$$

Letting $C(\omega(x),\tau) = \left[\frac{1}{(1+\tau)\omega(x)^{\eta\alpha}} \right]^{\frac{1}{\eta(1-\alpha)-1}} \left[\left(\frac{\eta-1}{\eta} \right) \alpha^{\frac{\alpha\eta}{\eta(1-\alpha)-1}} - \alpha^{\frac{\eta-1}{\eta(1-\alpha)-1}} \right]$, this is

$$\Pi_m(y;\boldsymbol{\omega},\tau) = \int_x f(x,y)^{\frac{\eta}{\eta(1-\alpha)-1}} C(\omega(x),\tau) dx.$$

A.8 Proof of Proposition 1 part (ii)

Proof. Recall that in the main text we defined the ratio of demand for workers of type x_1 to that for workers of type x_0 by any manager of type y, $R(y, x_1/x_0)$, as:

$$\frac{l^d(x_1,y)}{l^d(x_0,y)} = \left(\frac{w(x_0)}{w(x_1)}\right)^{\frac{\eta-1}{\eta(1-\alpha)-1}} \left(\frac{f(x_1,y)}{f(x_0,y)}\right)^{\frac{\eta}{\eta(1-\alpha)-1}}.$$

To show that $R(y, x_1/x_0)$ is increasing in type y whenever $x_1 > x_0$ and $\eta(1-\alpha) > 1$, it is enough to show that the sectoral human capital ratio $f(x_1, y)/f(x_0, y)$ is increasing in y if and only if $\rho < 0$. Differentiating the ratio with respect to y, the condition for the derivative to be positive—when $x_1 > x_0$—is:

$$\begin{aligned} & f_y(x_1,y)f(x_0,y) > f_y(x_0,y)f(x_1,y) \\ & \iff \frac{f_y(x_1,y)}{f_y(x_0,y)} > \frac{f(x_1,y)}{f(x_0,y)} \\ & \iff \left(\frac{f(x_1,y)}{f(x_0,y)}\right)^{1-\rho} > \frac{f(x_1,y)}{f(x_0,y)} > 1 \\ & \iff 1-\rho > 1 \quad \Rightarrow \quad \rho < 0. \end{aligned}$$

□