Online Appendix to Labor Substitutability among Schooling Groups

By Mark Bils, Bariş Kaymak, and Kai-Jie Wu

This document contains two appendices. Appendix A describes the data and presents the additional empirical results mentioned in the article. Appendix B provides the mathematical derivations of the theoretical results.

* Bils: Department of Economics, University of Rochester, Rochester, NY, FRB Richmond, and NBER, mark.bils@gmail.com. Kaymak: FRB Cleveland, Cleveland, OH, and CIREQ, barkaymak@gmail.com. Wu: Department of Economics, Pennsylvania State University, State College, PA, kjwu.twn@gmail.com. Acknowledgements: We thank Hannes Malmberg for his helpful comments. Barış Kaymak acknowledges support from Chaire de la Fondation J.W. McConnell en études Américaines at Université de Montréal. The views expressed here are those of the authors and do not necessarily reflect the views of the Federal Reserve Banks of Cleveland and Richmond or the Federal Reserve System.

A. Data Appendix

1. *Educational Attainments and Mincer Returns*

Sections II through IV use cross-country panel data on schooling attainments and estimated Mincer returns. The data on educational attainment by country are from Barro and Lee (2013) posted online at `http://www.barrolee.com/`. The data include 153 countries with attainments reported at five-year intervals from 1950 to 2010. It contains population frequency distributions over 7 educational categories by broad age groups. We restrict our population sample to those ages 25 to 54. We associate each attainment category with years of schooling using UNESCO Institute for Statistics (`http://data.uis.unesco.org/`) data on the duration of educational categories (in 2010) for each country. Our benchmark case divides workers into four groups: i) completed primary or less, ii) some secondary schooling, iii) completed secondary, and iv) at least some tertiary. To measure scarcity, for each country in each year, we regress the (log) size of the population in each schooling category on the years of schooling for that category. For our 104-country sample for Section III, the average shares by group are respectively 82 percent, 9 percent, 6 percent, and 3 percent in 1965 (weighted by population), and they become 37 percent, 31 percent, 23 percent, and 8 percent in 2010.

The data on the Mincer return are obtained from Psacharopoulos and Patrinos (2018), who compile 1,120 estimates of Mincer wage equations, from micro data on workers' wages, ages and education, for 139 countries going back before 1960. We use what Psacharopoulos and Patrinos (2018) label the overall Mincerian private return (column G in Annex 2 to their paper.) In cases where multiple Mincer estimates are available for a country at the same five-year intervals, we use the average of those estimates.

2. *National Income Accounts: Penn World Table*

The growth and income accounting in II and III also requires information on real GDP per worker and capital stock per worker for each country in each year. These data are obtained from Penn World Table 9.1 (Feenstra, Inklaar and Timmer, 2015). For growth accounting, we use `rgdpna` and `rnna` for real GDP and capital stock, respectively (i.e., real variables valued at constant 2011 national prices) so that variables are comparable across time with each country. For cross-country income accounting, we use `rgdpo` and `cn` for real GDP and capital stock (i.e., real variables valued at PPPs) so that variables are comparable across countries. We then divide the levels of real income and capital stock by `emp` to obtain their per worker value.

Merging data from Barro and Lee (2013) and Psacharopoulos and Patrinos (2018) with data from the Penn World Table results in an unbalanced panel sample of 367 observations for 104 countries spanning 1960 to 2010 with data on both attainment and the Mincer return to schooling. For the growth accounting results

Table A.1—List of Countries in Sample

60 countries for growth and development accounting				
Argentina	Denmark	Iran	Panama	Sweden
Australia	Ecuador	Israel	Peru	Switzerland
Austria	Egypt	Italy	Philippines	Taiwan
Bolivia	Finland	Japan	Poland	Thailand
Brazil	France	Kenya	Portugal	Tunisia
Bulgaria	Germany	Latvia	South Korea	Turkey
Canada	Ghana	Malaysia	Romania	USA
Chile	Greece	Mexico	Slovenia	Uganda
China	Guatemala	Netherlands	South Africa	United Kingdom
Colombia	Hungary	Nicaragua	Spain	Tanzania
Costa Rica	India	Norway	Sri Lanka	Venezuela
Cyprus	Indonesia	Pakistan	Sudan	Viet Nam
44 countries for development accounting only				
Albania	Cote d'Ivoire	Ireland	Mongolia	Singapore
Algeria	Croatia	Jamaica	Morocco	Slovakia
Bangladesh	Czech Rep.	Jordan	Namibia	Tajikistan
Belgium	Dominican Rep.	Kazakhstan	Nepal	Ukraine
Belize	El Salvador	Kuwait	New Zealand	UAE
Botswana	Estonia	Kyrgyzstan	Niger	Uruguay
Cambodia	Gambia	Malawi	Paraguay	Zambia
Cameroon	Honduras	Maldives	Russia	Zimbabwe
Hong Kong	Iraq	Malta	Rwanda	

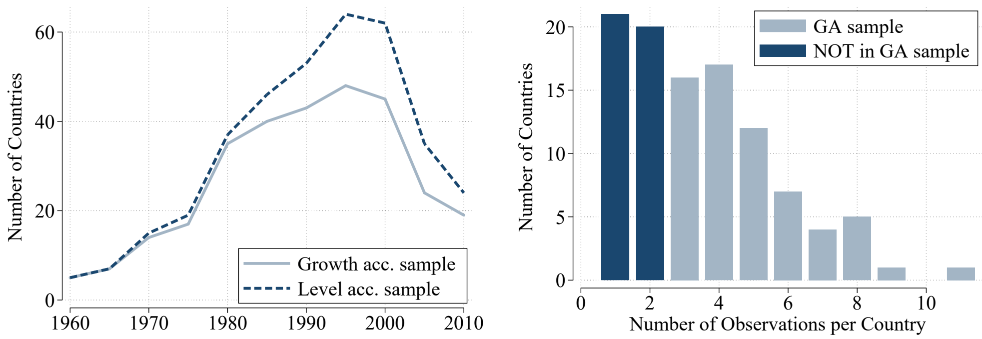
Note: Table lists the 104 countries whose Mincer return, schooling distribution, and GDP per worker are observed at least once between 1960-2010. The top panel lists the 60 countries used for growth accounting in Section II as well as development accounting in Section III. The second panel lists the other countries used only for development accounting.

in Section II, we further require a country to be observed at three or more of the intervals, yielding a smaller sample with 60 countries and 298 observations.[1] Figure A.1 describes the panel structure of our sample. Panel a depicts the number of countries for each five-year interval. It shows that our observations are mainly concentrated between 1975 and 2010. Panel b shows the frequency distributions of observations per country. Forty-five countries have less than 2 observations and hence are not used for growth accounting. The remaining countries have at least 3 observations during the sample period.

Table A.1 lists the countries represented in the empirical results, denoting each exercise for which a country could be utilized.

Figure A.2 plots the distribution of schooling scarcity (Panel a) and Mincer

[1]We drop observations through 1990 for countries that were formerly held in the Soviet Union.

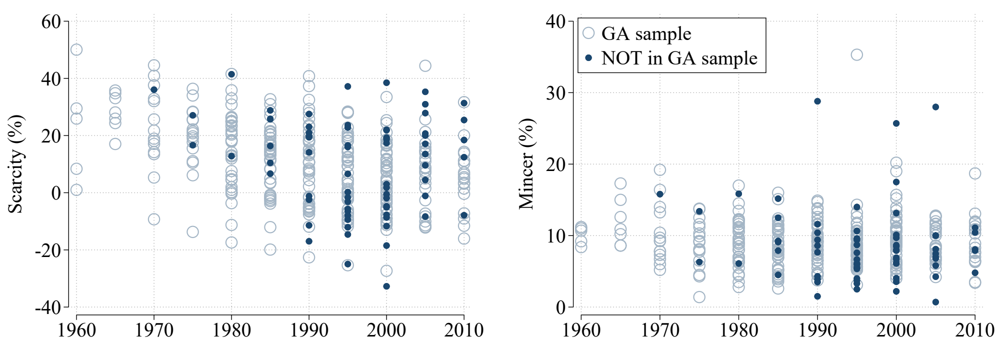


(a) Counts per Year

(b) Counts per Country

FIGURE A.1. SAMPLE COUNTS ACROSS YEARS AND COUNTRIES

Note: The light (gray) solid line in Panel a shows the total number of observations for each year in the 105-country sample used for income accounting; The dark (blue) dashed line shows the total number of observations for each year for the 60-country sample used for growth accounting (GA sample). Panel b shows the number of observations per country in the sample.



(a) Scarcity

(b) Mincer Return

FIGURE A.2. SCARCITY AND MINCER RETURN IN MAIN SAMPLES

Note: Figures depict the sample distribution of scarcity and Mincer return across time in our sample. Each light (gray) circle indicates an observation in the 60-country sample used for growth accounting; each dark (blue) dot indicates an observation in the 105-country sample but not used for growth accounting. The plot thickens with time as data become available on more countries.

return (Panel b) in our merged sample. Each light (gray) circle indicates an observation in the 60-country growth accounting sample, and each dark (blue) dot indicates an additional observation of the 105-country income accounting sample. As described in Section II.A, scarcity trends downward due to increased educational attainment worldwide, while Mincer return stays stable over time. These patterns are consistent in both samples.

3. Employment-Based Measure of School Scarcity

In our calculations of scarcity, we rely on the population shares of schooling from Barro and Lee (2013), whereas the wage equation in (7) stipulates the relative shares of schooling in the work force. The International Labor Organization (ILO, `https://ilostat.ilo.org/`) provides employment share by schooling for 121 countries and for the years 1990 to 2018. For most countries, however, data are only available after 2002, which is too recent to line up with the estimated Mincer returns from Psacharopoulos and Patrinos (2018) and hence is not applicable to our main analysis. We therefore use schooling population shares from Barro and Lee (2013) in our analysis.

To gauge the difference between the two measures, we compare employment-based scarcity calculated from the ILO to the population-based scarcity calculated from Barro and Lee (2013) for recent years. Figure A.3 shows that the two measures closely align, being concentrated along the dashed 45-degree line with an almost perfect correlation of 0.98, implying that substituting population shares of schooling for employment shares imparts no significant bias.

4. Employment by sector and education

The ILO provides employment by area (rural and urban) and education (less than basic, basic, intermediate, and advanced) for 135 countries and for years 2010 to 2021. In Section II.C, we use rural employment as a proxy for agricultural employment, defining the bottom group as basic education or below. Basic education in the ILO includes workers with primary and lower-secondary education. For more information, see the Education and Mismatch Indicators (EMI): `https://ilostat.ilo.org/resources/concepts-and-definitions/description-education-and-mismatch-indicators/`.

5. Measures of School Quality

In Section III.C, we consider two measures of schooling quality across countries. First, we employ Schoellman's (2012) estimates of a country's schooling quality in 2000 based on US earnings of immigrants who received all or most of their schooling in their country of birth. Schoellman's (2012) supplementary data (reported in his Table A1) include school quality estimates for 131 countries in 2000. We merge this sample with the Penn World Table 9.1 to obtain GDP per

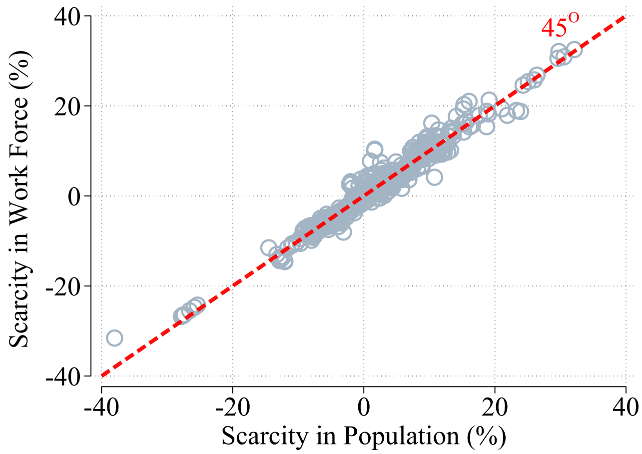


Figure A.3. Employment and Population Measures of Scarcity

Note: Schooling scarcity in population (x) reflects authors' calculations based on Barro and Lee (2013). Scarcity in work force is calculated based on data from the International Labor Organization (ILO). Dashed red line depicts the 45° line.

worker. The merged sample has 116 countries and is used in Section III.C. There are 51 countries for which we also have estimates of scarcity and Mincer returns in the home countries in 2000.

Our second measure is based on standardized test scores across countries, more precisely on the gradient of the test score with respect to years of schooling by country. The testing is overseen by the Programme for International Student Assessment (PISA). These tests are given to students age 15 in three areas: mathematics, science, and reading. We construct two school quality measures based on the micro-level data from the 2015 wave of the test, as discussed in Section III.C.[2] (We discuss our preferred measure in the text, the alternate only in Section A.5 below.) These data are available from the OECD (`https://www.oecd.org/pisa/data/`). We map these test scores to their implications for wages based on the relationship between wage rates and a standardized test score in the US as estimated by Lange (2007) for the 1979 cohort of the National Longitudinal Survey of Youth (NLSY, see `https://www.bls.gov/nls/nlsy79.htm`) using Armed Forces Qualification Tests. There are 51 countries from the PISA data for which we can obtain schooling scarcity and Mincer returns from our main sample.

[2]See the OECD "PISA 2015 Results in Focus" (`https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf`).

An Alternative PISA-Based School Quality Measure . — Our benchmark quality measure implicitly assumes that the test score prior to schooling is zero. If richer countries have better pre-school training, then the measure is biased up for these countries. To relax this assumption, we construct a second measure of quality.

In the PISA data, students from the same country will be in different grades when taking the test if the age at which they begin school depends on the month of birth or differs across schools, for instance, across regions. We use this variation in schooling to construct an alternative measure. In each country, we regress the test score on the grade year in which the test was taken, controlling for gender. We restrict the sample to native-born students who never repeated a grade. The coefficient on the grade year gives the test-score return to a year of schooling and forms the basis of our second measure. As with our first measure, we divide the resulting per-school-year score by the standard deviation of US test scores and valorize it at 15 percent.

Figure A.4 contrasts the alternative PISA-based measure against log GDP per worker. On average, a one log point higher income is associated with a 1.4 percentage-point increase in ϕ_q (with a much larger standard error of 0.7). The gradient is similar to that of Schoellman's immigrant-based measure, and hence their implied ϕ_b variation and human capital are alike (see Panel a of Figure 9). We prefer the benchmark measure because of its lower standard error.

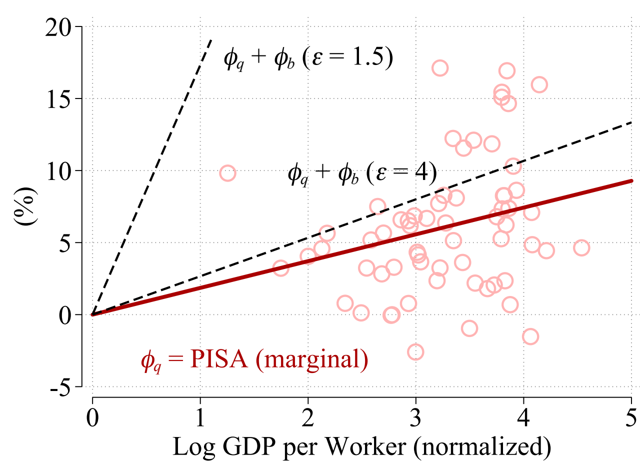


Figure A.4. PISA-Based School Quality Measure vs. Implied $\tilde{\phi}_z(\varepsilon)$

Note: Figure plots the alternative PISA-based school quality measure against log GDP per worker. Variables are normalized to set the predicted value for the poorest country to zero. The solid red line depicts the OLS fitted values for school quality. The dashed black lines depict the projections of $\tilde{\phi}_z(\varepsilon)$ on income for $\varepsilon = 1.5$ and 4. Data on income per worker are from PWT 9.1. $\tilde{\phi}_z(\varepsilon)$ reflects authors' calculation based on data from Barro and Lee (2013) and Psacharopoulos and Patrinos (2018).

6. *Scarcity under Different Grouping Rules*

Our benchmark reflects the four schooling groups as listed above. But we consider the sensitivity of measured scarcity to the following alternative groupings:

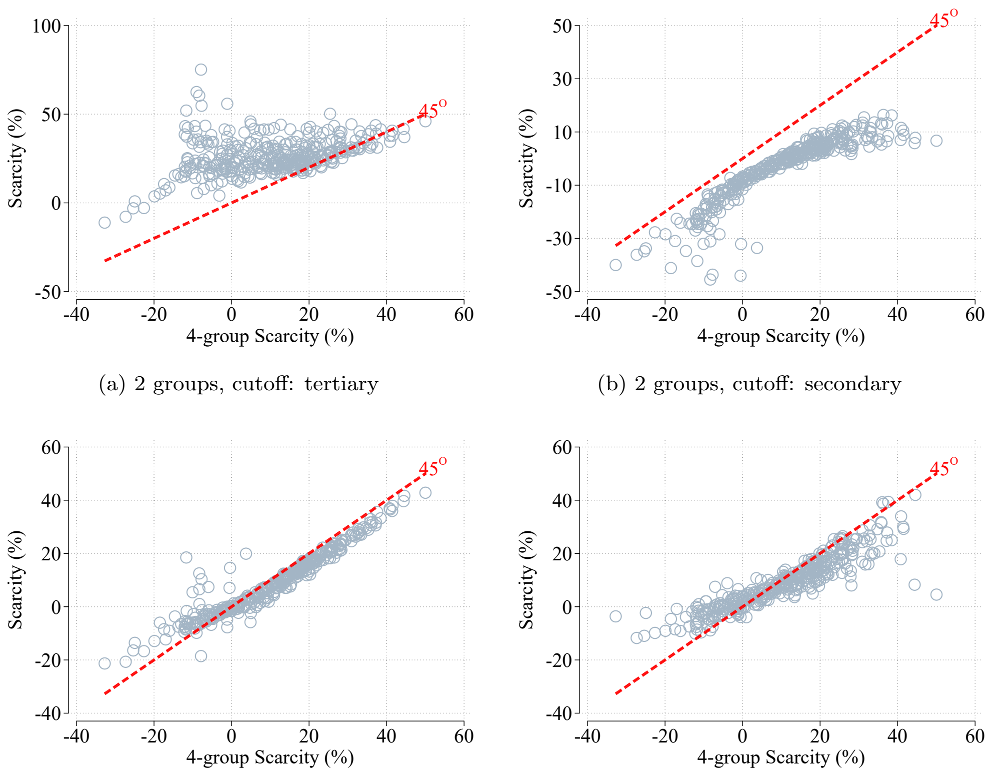
(a) 2 Groups : less than tertiary; some tertiary or more.

(b) 2 Groups : less than secondary; some secondary or more.

(c) 3 Groups: less than secondary; some or complete secondary; some tertiary and above.

(d) 6 Groups: less than complete primary; complete primary; some secondary; complete secondary; some tertiary; complete tertiary.

Figure A.5 compares schooling scarcity calculated under each alternative (y-axis) to our benchmark with 4 groups (x-axis). The dashed red lines depict the 45° lines. Scarcity measures based on three or six groupings, Panels c and d, are each similar to that from our four-group case. The groupings into two levels, Panels a and b, both diverge from our benchmark, but in different directions reflecting the choice of cutoff. With fewer categories, a larger share of schooling variations is manifested within-group, muting variations in measured scarcity. This is especially true for only two groups, most notably when the cutoff is further from median schooling as in Panel a. The takeaway from Figure A.5 is that distinctions between primary and secondary and between secondary and tertiary schooling can both be important components of scarcity. Note that grouping multiple schooling categories into one group implicitly assumes perfect substitution among them. We analyze the implications of grouping choices for substitutability and development accounting more generally in Section IV.A.



(a) 2 groups, cutoff: tertiary

(b) 2 groups, cutoff: secondary

(c) 3 groups

(d) 6 groups

Figure A.5. Scarcity with Differing Number of Groups versus Benchmark 4 Group Scarcity.

Note: Schooling scarcity in population (x) reflects authors' calculations based on Barro and Lee (2013). Red dash lines depict the 45° lines.

B. Theoretical Appendix

1. Derivation of the Long-Run Elasticity of Substitution

As in Caselli and Coleman (2006) we consider an economy with a large number of competitive firms, with labor and capital supplied elastically. The representative firm solves the optimization problem:

$$\max_{\{L_i,A_i\},K} K^{\alpha}H^{1-\alpha} - \sum_{i\in S} w_i L_i - RK,$$

subject to the technological frontier

$$\sum_{i\in S} \left(\gamma_i\, A_i\right)^{\omega} \leq B,$$

where effective labor input, H, aggregates labor over skill groups

$$H = \Big[\sum_{i\in S}\Big(A_i q_i L_i\Big)^{\frac{\varepsilon_{SR}-1}{\varepsilon_{SR}}}\Big]^{\frac{\varepsilon_{SR}}{\varepsilon_{SR}-1}}. \tag{22}$$

An equilibrium consists of factor prices $\{w_i\}_{i\in S}$ and R and allocations $\{L_i, A_i\}_{i\in S}$ and K such that input markets clear subject to firms' having optimized at those prices.

We next show the condition for a symmetric equilibrium with an interior solution. It enables us to characterize the equilibrium with the first-order conditions of a representative firm. Then we derive the long-run elasticity of substitution, which parallels Hendricks and Schoellman's (2023) treatment.

Symmetric Equilibrium with an Interior Solution. — We want to show that $\omega-\varepsilon_{SR}+1 > 0$ is a sufficient condition for a symmetric equilibrium with an interior solution. A symmetric equilibrium means all firms choose the same technology bundles, and an interior solution means $A_i > 0$ for all $i \in S$.

First we denote $D_i = {A_i}^{\omega}$ and rewrite the firm's optimization problem over technologies, for given values of $K > 0$ and $L_i > 0$, for all $i \in S$, as:

$$\max_{\{D_i\}_S} K^{\alpha}\Big[\sum_{i\in S} D_i^{\frac{\varepsilon_{SR}-1}{\omega\varepsilon_{SR}}}\Big(q_i L_i\Big)^{\frac{\varepsilon_{SR}-1}{\varepsilon_{SR}}}\Big]^{\frac{(1-\alpha)\varepsilon_{SR}}{\varepsilon_{SR}-1}} - \sum_{i\in S} w_i L_i - RK,$$

subject to

$$\sum_{i\in S} {\gamma_i}^{\omega} D_i \leq B.$$

The constraint set is convex without additional restrictions on parameters. Now suppose $\omega - \varepsilon_{SR} + 1 > 0$. Then $(\varepsilon_{SR} - 1)/\omega\varepsilon_{SR} < 1$ because $\varepsilon_{SR} > 1$. Under this condition, the objective function is strictly quasi-concave, so the existence and uniqueness of a global maximizer is guaranteed. Additionally, because the marginal profit of investing in D_i goes to infinity when D_i goes to zero, the solution must have $A_i > 0$ for all $i \in S$. The symmetry of equilibrium is directly implied because all firms face the same optimization problem with unique solutions.

LONG-RUN ELASTICITY OF SUBSTITUTION . — Rearranging the first-order condition with respect to A_i for each $i \in S$ gives:

$$A_i = \gamma_i^{\frac{-\omega\varepsilon_{SR}}{\omega\varepsilon_{SR}-\varepsilon_{SR}+1}} \Big(q_i\, L_i\Big)^{\frac{\varepsilon_{SR}-1}{\omega\varepsilon_{SR}-\varepsilon_{SR}+1}} Q^{\frac{\varepsilon_{SR}}{\omega\varepsilon_{SR}-\varepsilon_{SR}+1}}, \tag{23}$$

where $Q = (1-\alpha)K^{\alpha}H^{1/\varepsilon_{SR}-\alpha}/(\lambda\omega)$ and λ is the Lagrangian multiplier. Note that (23) can also be written as:

$$\Big(\gamma_i\, A_i\Big)^{\omega} = \Big(A_i\, q_i\, L_i\Big)^{\frac{\varepsilon_{SR}-1}{\varepsilon_{SR}}} Q,$$

for each $i \in S$. Summing up both sides of the equation across skill groups, we have

$$Q = BH^{\frac{1-\varepsilon_{SR}}{\varepsilon_{SR}}}.$$

Substituting for Q in (23) and letting $b_i = B^{\frac{1}{\omega}}/\gamma_i$, gives the optimal choice of technology:

$$A_i = \left(\frac{q_i\, L_i}{H}\right)^{\frac{\varepsilon_{SR}-1}{\omega\varepsilon_{SR}-\varepsilon_{SR}+1}} b_i^{\frac{\omega\varepsilon_{SR}}{\omega\varepsilon_{SR}-\varepsilon_{SR}+1}}. \tag{24}$$

Plugging the optimal technology choice into the labor input aggregator (22), we get:

$$H = \left[\sum_{i\in S} \Big(q_i\, b_i\, L_i\Big)^{\frac{\omega\varepsilon_{SR}-\omega}{\omega\varepsilon_{SR}-\varepsilon_{SR}+1}}\right]^{\frac{\varepsilon_{SR}-1}{\varepsilon_{SR}}} H^{\frac{1-\varepsilon_{SR}}{\omega\varepsilon_{SR}-\varepsilon_{SR}+1}}.$$

Rearranging the equation to solve for H, we can rewrite the aggregator as:

$$H = \left[\sum_{i \in S} \Big(q_i \, b_i \, L_i \Big)^{\frac{\omega\varepsilon_{SR} - \omega}{\omega\varepsilon_{SR} - \varepsilon_{SR} + 1}} \right]^{\frac{\omega\varepsilon_{SR} - \varepsilon_{SR} + 1}{\omega\varepsilon_{SR} - \omega}}$$

This gives the long-run elasticity of substitution:

$$\varepsilon = \frac{\omega\varepsilon_{SR} - \varepsilon_{SR} + 1}{\omega - \varepsilon_{SR} + 1}.$$

Under the assumption $\omega - \varepsilon_{SR} + 1 > 0$, this long-run elasticity is finite and positive.

Now we can derive the wage-schooling relationship when technology choices are endogenized. Equating group s_i's wage to its marginal product gives:

$$w_i = \frac{\partial Y}{\partial H} H^{\frac{1}{\varepsilon}} \Big(A_i \, q_i \Big)^{\frac{\varepsilon_{SR} - 1}{\varepsilon_{SR}}} L_i^{\frac{-1}{\varepsilon_{SR}}}.$$

Substituting for the optimal technology choice, equation (24), yields:

$$w_i = \frac{\partial Y}{\partial H} H^{\frac{1}{\varepsilon}} \Big(q_i \, b_i \Big)^{\frac{\varepsilon - 1}{\varepsilon}} L_i^{\frac{-1}{\varepsilon}}, \tag{25}$$

which is equivalent to the first-order condition derived from the long-run aggregator.

2. *Technology Choice and Wage Shares*

In equilibrium, the efficiency of workers in schooling group i can be written as a function of the group's quality q_i, its technology frontier b_i, and its earnings as a share of the total labor income of the economy. Substituting the optimal choice of technology (24) into the definition of e_i gives

$$e_i = A_i \, q_i = q_i \, b_i \left[\frac{q_i \, b_i \, L_i}{H} \right]^{\frac{\varepsilon_{SR} - 1}{\omega\varepsilon_{SR} - \varepsilon_{SR} + 1}}.$$

Then we substitute in the long-run labor aggregator H and relabel the parameters.

$$e_i = q_i\, b_i \left[\frac{q_i\, b_i\, L_i}{\left[\sum_{j \in S} \left(q_j\, b_j\, L_j \right)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}} \right]^{\frac{1}{\omega}\left(\frac{\varepsilon-1}{\varepsilon}\right)} = q_i\, b_i \left[\frac{\left(q_i\, b_i\, L_i \right)^{\frac{\varepsilon-1}{\varepsilon}}}{\sum_{j \in S} \left(q_j\, b_j\, L_j \right)^{\frac{\varepsilon-1}{\varepsilon}}} \right]^{\frac{1}{\omega}}$$

$$= q_i\, b_i \left(\frac{w_i\, L_i}{\sum_{j \in S} w_j\, L_j} \right)^{\frac{1}{\omega}} = q_i\, b_i \left(\frac{w_i\, L_i}{\overline{w} L} \right)^{\frac{1}{\omega}}$$

The last row is implied by the long-run wage equation (25).

3. *Immigrant Mincer Return and Cross-Country Human Capital*

There are two sources of efficiency associated with a schooling level: human capital accumulated from the schooling (ϕ_q) and the level of technology accessible with that schooling (ϕ_b). In this paper, we follow Schoellman (2012) by using the Mincer returns that he estimates for immigrants in the United States as a measure of ϕ_q for the immigrants' country of origin. The intuition is that technology reflects a worker's current location, while human capital from schooling was determined by the efficiency of schooling in the country where that investment took place, that being the worker's home country.

To see this, consider the following aggregator extended from (6), where workers in the United States from different home countries $c \in C$ are perfect substitutes provided they have the same educational attainment.

$$H_{US} = \left[\sum_{i \in S} \left(\sum_{c \in C} b_{i,US}\, q_{i,c}\, L_{i,c} \right)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}.$$

Note that immigrant workers work with US technology, so they share a common technology frontier, $b_{i,US}$. On the other hand, the human capital gain from schooling, $q_{i,c}$, depends on the workers' country of origin because immigrant workers accumulated human capital in their home country. The wage for such an immigrant worker is:

$$w_{i,c} = \left(\frac{\partial Y}{\partial H_{US}} H_{US}{}^{\frac{1}{\varepsilon}} \right) \left(\sum_{c} b_{i,US}\, q_{i,c}\, L_{i,c} \right)^{\frac{-1}{\varepsilon}} b_{i,US}\, q_{i,c}. \tag{26}$$

Let m_c^{US} be the Mincer return estimated in the US labor market *across* immigrants from country c. Taking natural logs and projecting on s_i for both sides of equation (26) gives

$$m_c^{US} = \underbrace{\left[\phi_{b,US} + \frac{1}{\varepsilon}\tilde{x}_{US}\right]}_{\zeta} + \phi_{q,c}, \tag{27}$$

where $\tilde{x}_{US}$ is the scarcity of more-educated workers in the US in terms of efficiency units, obtained by projecting $-\ln\left(\sum_c b_{i,US}\, q_{i,c}\, L_{i,c}\right)$ on s_i. Equation (27) shows that the cross-country variation in $\phi_{q,c}$ can be captured by the cross-(home)-country variation of immigrants' Mincer return m_c^{US}. In the main text we refer to m_c^{US} as m^{US}, keeping the country subscript implicit.

4. *Growth Accounting Equations with Nested-CES Production*

Let G be the compound bottom group that consists of the first N schooling groups. Consider the nested-CES labor aggregator:

$$H = \left[\left(z_G L_G\right)^{\frac{\varepsilon-1}{\varepsilon}} + \tilde{H}\left(z_{N+1}L_{N+1}, ..., z_S L_S\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{28}$$

where $\tilde{H}(\cdot)$ is constant return to scale, $L_G = \sum_{j\leq N} L_j$, $z_G \equiv Z_G/L_G$ and

$$Z_G \equiv \left[\sum_{j\leq N}\left(z_j L_j\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}. \tag{29}$$

The wage for each schooling group $i > N$ is

$$w_i = \left(\frac{\partial Y}{\partial H}\right) H^{\frac{1}{\varepsilon}} \tilde{H}^{\frac{-1}{\varepsilon}} \tilde{H}_i' z_i. \tag{30}$$

On the other hand, the wage for each group $j \leq N$ is:

$$w_j = \left(\frac{\partial Y}{\partial H}\right) H^{\frac{1}{\varepsilon}} Z_G^{\frac{\varepsilon-\sigma}{\varepsilon\sigma}} z_j^{\frac{\sigma-1}{\sigma}} L_j^{\frac{-1}{\sigma}}. \tag{31}$$

Taking the average among the first N groups, we get:

$$\overline{w}_G \equiv \left(\frac{1}{L_G}\right)\sum_{j\leq N} w_j L_j = \left(\frac{\partial Y}{\partial H}\right) H^{\frac{1}{\varepsilon}} z_G^{\frac{\varepsilon-1}{\varepsilon}} L_G^{\frac{-1}{\varepsilon}}. \tag{32}$$

Now apply (30) and (32). We can write the overall average wage as:
(33)

$$\overline{w} = \frac{1}{L}\left(\overline{w}_G L_G + \sum_{i>N} w_i L_i\right) = \frac{1}{L}\left(\frac{\partial Y}{\partial H}\right) H^{\frac{1}{\varepsilon}} \left[(z_G L_G)^{\frac{\varepsilon-1}{\varepsilon}} + \tilde{H}^{\frac{-1}{\varepsilon}} \sum_{i>N} \tilde{H}'_i z_i L_i\right]$$
$$= \frac{1}{L}\left(\frac{\partial Y}{\partial H}\right) H^{\frac{1}{\varepsilon}} \left[(z_G L_G)^{\frac{\varepsilon-1}{\varepsilon}} + \tilde{H}^{\frac{\varepsilon-1}{\varepsilon}}\right],$$

where $\sum_{i>N} \tilde{H}'_i z_i L_i = \tilde{H}$ because $\tilde{H}(\cdot)$ is constant return to scale (Euler's homogeneous function theorem). Combining the previous two equations, we have:

$$\frac{\overline{w}L}{\overline{w}_G L_G} = \left[1 + \left(\frac{\tilde{H}}{z_G L_G}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right].$$

Now turning back to the aggregator (28), we can get an equation parallel to (10)

$$(34) \quad h \equiv \frac{H}{L} = z_G\left(\frac{L_G}{L}\right)\left[1 + \left(\frac{\tilde{H}}{z_G L_G}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}} = z_G\left(\frac{\overline{w}}{\overline{w}_G}\right)^{\frac{\varepsilon}{\varepsilon-1}}\left(\frac{L}{L_G}\right)^{\frac{1}{\varepsilon-1}}.$$

Likewise, combining equations (31) and (32), we have:

$$\frac{\overline{w}_G L_G}{w_1 L_1} = \left[1 + \sum_{2\le j\le N}\left(\frac{z_j L_j}{z_1 L_1}\right)^{\frac{\sigma-1}{\sigma}}\right],$$

which, in turn, gives:

(35)

$$z_G = \frac{Z_G}{L_G} = z_1\left(\frac{L_1}{L_G}\right)\left[1 + \sum_{2\le j\le N}\left(\frac{z_j L_j}{z_1 L_1}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} = z_1\left(\frac{\overline{w}_G}{w_1}\right)^{\frac{\sigma}{\sigma-1}}\left(\frac{L_G}{L_1}\right)^{\frac{1}{\sigma-1}}.$$

Combining (34) and (35) gives equation (18).