Extending Development Accounting*

Lutz Hendricks[†] and Todd Schoellman[‡]

February 14, 2019

Abstract

This paper extends development accounting to an environment that features imperfectly substitutable skills and cross-country variation in the skill bias of technology. We find that human capital accounts for between 50% and 65% of cross-country income gaps. This finding remains robust when we consider alternative sources of skill-biased technology variation, alternative definitions of skilled and unskilled labor, and alternative values for the elasticity of substitution between skilled and unskilled labor. We derive closed form solutions for the contribution of human capital to output gaps in terms of observable data moments. These allow us to understand precisely which features of the data are responsible for our main results.

1 Introduction

This paper studies the contribution of human capital to cross-country income differences when countries differ in the skill bias of technology.

Motivation Our work builds on a sizable development accounting literature (see Hsieh and Klenow, 2010; Caselli, 2016). Development accounting decomposes cross-country differences in output per worker into the contributions of factor inputs and total factor productivity. Its objective is to shed light on the proximate causes of cross-country income differences.

*This paper has benefited from comments from seminar participants at the IIES and the 2018 Midwest Macroeconomics Meeting at University of Georgia. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

[†]University of North Carolina at Chapel Hill. hendricksl@protonmail.com

[‡]Federal Reserve Bank of Minneapolis. todd.schoellman@gmail.com

An early development accounting literature treated skilled and unskilled labor as perfect substitutes. This work concluded that human capital accounts for around 30% of cross-country income gaps (e.g., Hall and Jones, 1999). Recently, Jones (2014) showed that human capital can be far more important, if skills are imperfect substitutes. Hendricks and Schoellman (2018) showed that wage gains at migration provide useful information about cross-country human capital differences. They estimate that human capital accounts for around 60% of cross-country income gaps.

Virtually the entire literature on levels accounting has assumed that productivity is skill neutral. This contrasts with an extensive macro-labor literature that emphasizes the importance of skill-biased technology for the college wage premium over time (e.g., Katz and Murphy, 1992) and across countries (Caselli and Coleman, 2006). The goal of this paper is to quantify the importance of human capital for cross-country income differences in the presence of skill biased technology.

Cross-country variation in skill biased productivities creates a challenging identification problem. Skill specific human capital and skill augmenting technology have the same effect on the marginal product of labor and therefore on observable wages.

A comparison of Caselli and Coleman (2006) and Jones (2014) provides a stark illustration of this problem. Both models consist of a Cobb-Douglas aggregate production function in capital and labor inputs combined with a constant elasticity labor aggregator of the form

$$L_c = \left[\sum_{j=1}^{2} (\theta_{j,c} N_{j,c})^{\rho} \right]^{1/\rho} \tag{1}$$

where j indexes skill, $N_{j,c}$ denotes labor input of skill j in country c (measured in hours worked), $\theta_{j,c}$ denotes skill j augmenting productivity, and $\rho \in (0,1)$ governs the elasticity of substitution between skilled and unskilled labor. Caselli and Coleman (2006) and Jones (2014) calibrate this model to the same data moments (output gaps, capital income shares, and skill premiums). Both papers find large cross-country differences in the relative productivity of skilled labor $(\theta_{2,c}/\theta_{1,c})$ when ρ is set to empirically plausible values. However, the authors draw very different conclusions from their findings. While Jones attributes variation in $\theta_{j,c}$ to human capital, Caselli and Coleman attribute it to skill biased technology. Clearly, the implications for levels accounting are very different depending on which label is attached to $\theta_{j,c}$.

Approach We develop a model that allows for imperfect skill substitution and non-neutral productivity differences. Our model allows for four sources of skill bias differences across countries that are prominent in the literature:

1. exogenous skill biased technology differences as in Katz and Murphy (1992);

2. firms choosing an "appropriate technology" from a fixed technology frontier as in Caselli and Coleman (2006);

3. firms investing in skill augmenting technical change as in Acemoglu (2007);

4. capital-skill complementarity as in Krusell et al. (2000).

We calibrate the model to match data moments that are standard in the literature: cross-country output gaps, capital income shares, and skill premiums. Most of our data moments are constructed from Penn World Table 9 data (Feenstra et al., 2015). We also target migrant wage gains as in Hendricks and Schoellman (2018). The only new data moments are countries' stocks of equipment and structures and the associated income shares. These are constructed from Penn World Table and International Comparison Project data. We take the U.S. to be the stand-in rich country and compare it with the set of 63 countries with output per worker below one quarter of the U.S.

When we perform levels accounting in models with endogenous skill bias (as in Caselli and Coleman 2006 or Acemoglu 2007), we distinguish the "direct" effect of human capital (holding skill bias fixed) from its total effect, which includes induced variation in the skill bias of technology. This approach is consistent with the treatment of changes in physical capital that are induced by changes in labor inputs in the literature.

Results Our baseline model combines the labor aggregator of Jones (2014) with the technology frontier of Caselli and Coleman (2006) and the capital-skill complementarity of Krusell et al. (2000). The calibrated model implies that human capital accounts for 61% to 65% of cross-country output gaps. In contrast to Jones (2014), our results depend little on the choice of skill cutoff that divides workers into skilled and unskilled labor and not at all on the elasticity of substitution between skilled and unskilled labor.

We show analytically that our baseline model is equivalent to a model with skill-neutral productivity and an alternative, higher elasticity of substitution between skilled and unskilled labor. This "long-run" elasticity of substitution is a mixture of the conventional "short-run" elasticity of substitution and the degree of curvature between adopting or innovating technologies biased towards one type of labor. The distinction between a short-run and long-run elasticity of substitution provides a useful framework for thinking about why conventional estimates of the elasticity of substitution taken from recent U.S. experience may not be appropriate for development. It follows that the total contribution of human capital to output gaps does not depend on the (controversial) value of the "short-run" elasticity.

The model attributes at least 70% of the cross-country variation in the relative productivity of skilled labor to skill bias gaps rather than human capital gaps.[1] Nonetheless, human

[1] Our approach and findings on this point are close in spirit to Rossi (2019).

capital is far more important for cross-country output gaps than is variation in skill bias. The direct contribution of human capital, holding skill bias fixed, ranges from 49% to 61% depending on the elasticity of substitution between skilled and unskilled labor and the skill cutoff. Hence, human capital is important for output gaps, even though the relative human capital of skilled versus unskilled labor does not vary much across countries. The central reason for this result is that migrant wage gains imply that unskilled workers in rich countries are endowed with at least twice as much human capital compared with poor countries.

We consider several model variations to examine the robustness of our findings.

1. Treating cross-country variation in skill bias as exogenous reduces the total contribution of human capital to 49% to 61%. These are the same values that the baseline model with the technology frontier obtains as the direct contribution of human capital (holding skill bias fixed).

2. Abstracting from capital-skill complementarity has little effect on the quantitative results. The model again has a reduced form representation that solves out firms' skill bias choices. The reduced form model is equivalent to the model studied by Jones (2014), but with a higher elasticity of substitution between skilled and unskilled labor. It follows that the levels accounting results of Hendricks and Schoellman (2018) remain exactly valid for this case, even though the skill bias of technology is now allowed to differ across countries. We are able to derive a closed form solution for the contribution of human capital in terms of observable data moments. This provides sharper intuition for the robustness of the results across skill cutoffs.

3. Allowing firms to invest in skill biased technology, similar to Acemoglu (2007), does not change the results as long as the model has constant returns to scale.

1.1 Related Literature

To be written.

2 Model

2.1 Model Overview

We begin with a model that combines the human capital elements of Jones (2014) with the technology frontier of Caselli and Coleman (2006) and capital-skill complementarity as in Krusell et al. (2000).

The world is static. There are two countries, indexed by $c \in \{p, r\}$ (poor and rich). Output per worker is produced from capital and labor inputs, which are inelastically supplied by workers who are differentiated by skill level j. For the most part, we assume that labor comes in two skill levels (unskilled and skilled); $j \in \{1, 2\}$. But the model is written for generic $J \geq 2$. Labor input in efficiency units is the product of human capital and hours: $L_{j,c} = h_{j,c} N_{j,c}$.

2.2 Notation

It is useful to define commonly used notation at the outset.

1. $R(x) = x_r / x_p$ denotes the rich to poor country ratio of x.
2. $S(x_c) = x_{2,c} / x_{1,c}$ denotes the skilled to unskilled ratio of x.
3. The unobservable rental price per efficiency unit of labor is $p_{j,c}$. Hence, the total earnings of skill j are given by $W_{j,c} = p_{j,c} L_{j,c}$.
4. The observable wage per hour is $w_{j,c} = h_{j,c} p_{j,c}$.
5. The income ratio of two inputs is denoted by $IR_{ab} = \text{income}_a / \text{income}_b$. In particular, $IR_{L_2 L_1} = S(W)$.
6. The income share of an input is denoted by $IS_{a,c} = \text{income}_{a,c} / y_c$.

A number of useful properties of the rich-to-poor and skilled–to-unskilled ratios are worth noting. For any constant ϕ, we have

1. $R\left(x^{\phi}\right) = R(x)^{\phi}$ and $S\left(x^{\phi}\right) = S(x)^{\phi}$.
2. The order of rich-to-poor and skilled-to-unskilled ratios is interchangeable:

$$
\begin{aligned}
R\left(S\left(x^{\phi}\right)\right) &= S\left(R\left(x^{\phi}\right)\right) && (2) \\
&= \left[\frac{x_{2,r}/x_{1,r}}{x_{2,p}/x_{1,p}}\right]^{\phi} && (3)
\end{aligned}
$$

2.3 Model Details

Output per worker y_c is produced from capital and labor inputs according to the aggregate production function

$$
\begin{aligned}
y_c &= F\left(h_{j,c}, L_{j,c}, e_c, s_c, z_c\right) && (4) \\
&= s_c^{\alpha_c} \left(z_c L_c\right)^{1-\alpha_c} && (5)
\end{aligned}
$$

where

$$L_c = [(\theta_{1,c} L_{1,c})^\rho + (\theta_{2,c} Z_c)^\rho]^{1/\rho} \tag{6}$$

and

$$Z_c = \left[(\mu_e e_c)^\phi + (\mu_2 L_{2,c})^\phi\right]^{1/\phi} \tag{7}$$

s_c denotes structures per capita. The income share of structures, α_c, may vary across countries. z_c denotes neutral total factor productivity (taken as exogenous).

Aggregate labor input L_c is given by a CES aggregator of unskilled labor $L_{1,c}$ and a composite input Z_c. The elasticity of substitution is $\sigma = 1/(1-\rho) > 1$, so that $0 < \rho < 1$.

The composite input Z_c is a CES aggregator of skilled labor $L_{2,c}$ and equipment e_c. Its substitution elasticity is governed by $0 < \phi < 1$. The factor weights μ_e and μ_2 are taken as exogenous.

The factor weights in the labor aggregator $\theta_{j,c}$ are constrained by a technology frontier, similar to Caselli and Coleman (2006) or Acemoglu (2007), given by

$$CES(\theta_{j,c}, \kappa_j, \omega) = \left(\sum_j [\kappa_j \theta_{j,c}]^\omega\right)^{1/\omega} \leq B_c^{1/\omega} \tag{8}$$

with parameters $\omega > 0$, $B_c > 0$, and $\kappa_j > 0$.[2] In some versions of the model, we drop the frontier and treat the skill bias parameters as exogenous. We assume for now that B_c is fixed. When firms choose it, we assume $C(B_c) = bB_c$. This is the specification of example 1 in Acemoglu (2007).

2.4 Equilibrium

In line with the levels accounting literature, we assume that the economy is in steady state with an interest rate that is equal to the discount rate of the infinitely lived representative agent (e.g., Hsieh and Klenow 2010). This fixes the rental prices of structures and equipment, $q_{e,c}$ and $q_{s,c}$.

Firms hire labor in competitive markets at wage rates $w_{j,c}$ per hour worked $N_{j,c}$; or $p_{j,c}$ per efficiency unit of labor $L_{j,c}$. Hence, $w_{j,c} = p_{j,c} h_{j,c}$. The representative firm solves

$$\max_{s_c, e_c, L_{j,c}, \theta_{j,c}} y_c - q_{s,c} s_c - q_{e,c} e_c - \sum_j p_{j,c} L_{j,c} \tag{9}$$

subject to the aggregate production function (4) and the frontier constraint (8).

[2] A similar technology frontier in Acemoglu (2002, section 5.2), but linear $(\omega = 1)$. This results in endogenous growth, which renders levels accounting difficult.

A competitive equilibrium is then an allocation $y_c, e_c, s_c, \theta_{j,c}$ (and B_c if chosen) and a set of prices $p_{j,c}$ that satisfy

- firm first-order conditions for e_c, s_c, $L_{j,c}$ and $\theta_{j,c}$;
- first-order condition for B_c, if chosen; otherwise $B_c = 1$;
- the aggregate production function and technology frontier.

As in Caselli and Coleman (2006), we assume that

$$\Omega = \omega - \rho - \omega\rho > 0 \tag{10}$$

This condition ensures that firms choose an interior point on the technology frontier.

2.5 Levels Accounting

Our goal is to perform levels accounting that decomposes output gaps $R(y) = y_r/y_p$ into the contributions of labor inputs $L_{j,c}$ (or $h_{j,c}$ and $N_{j,c}$ separately), physical capital (e and s), neutral productivity z_c, skill bias $\theta_{j,c}$, and induced productivity (when B_c is endogenous).

Consistent with the literature (see Hsieh and Klenow, 2010), we account for the responses of the capital stocks to changes in human capital by assuming that the model is in steady state. If the saving decision is due to an infinitely lived representative agent, then the rate of return to capital is fixed by the discount rate. This, in turn, fixes q_s and thus s/y. It also fixes q_e.

Our approach to levels accounting follows Hendricks and Schoellman 2018. Starting from

$$y_c = (s_c/y_c)^{\alpha_c/(1-\alpha_c)} z_c L_c \tag{11}$$

the output gap can be additively separated into the contributions of TFP, structures, and labor inputs jointly with equipment:

$$\underbrace{\ln R(y)}_{\text{output gap}} = \underbrace{\ln R(z)}_{\text{TFP}} + \underbrace{\ln R\left((s/y)^{\alpha/(1-\alpha)}\right)}_{\text{structures}} + \underbrace{\ln R(L)}_{\text{labor and equipment}} \tag{12}$$

For each input, the share of the output gap accounted for by each input is given by

$$1 = \underbrace{\frac{\ln R(z)}{\ln R(y)}}_{\text{share}_z} + \underbrace{\frac{\ln R\left((k/y)^{\alpha/(1-\alpha)}\right)}{\ln R(y)}}_{\text{share}_k} + \underbrace{\frac{\ln R(L)}{\ln R(y)}}_{\text{share}_L} \tag{13}$$

The contribution of L may be subdivided into the separate contributions of its components ($h_{j,c}$, $N_{j,c}$, $\theta_{j,c}$, e_c). These are defined as follows. For specificity, consider the contribution of h. Define counterfactual output in the poor country as $\hat{y}_p = F(h_{j,r}, L_{j,p}, \hat{e}_p, \hat{s}_p, z_p)$, where $\hat{e}_p$ and $\hat{s}_p$ are the capital inputs that hold the capital rental prices constant at poor country levels. $\hat{y}_p$ answers the question: How much steady state output would be produced in the poor country, if it were endowed with the rich country levels of $h_{j,c}$? Define the share of the output gap explained by human capital as

$$\text{share}_h = 1 - \ln(y_r/\hat{y}_p) / \ln(y_r/y_p) \tag{14}$$

Consistent with the literature, the contribution of labor inputs includes induced changes in structures and equipment. The contribution of s would be computed by finding e and s that yield rental prices of$q_{s,r}$ and $q_{e,p}$, holding other inputs fixed. And analogously for e.

For labor inputs, we distinguish between their "direct" effect on output, holding skill bias fixed, and their "total" effect which includes induced changes in skill bias (movements along the technology frontier).

2.6 Calibration

We investigate the empirical implications of this model when parameters are chosen to match the same data moments that were used by Hendricks and Schoellman (2018):

1. output levels y_c for $c \in \{r, p\}$;
2. labor income shares W_c/y_c;
3. skill premiums $S(w_c) = w_{2,c}/w_{1,c}$ or, equivalently, the income ratios of skilled and unskilled labor $S(W)$ where $W_{j,c} = w_{j,c}N_{j,c}$ is the income earned by labor of skill j;
4. wage gains at migration: $wg_j = R(p_j)$ for $j \in \{1, 2\}$.

In addition, we construct data for equipment/output ratios (e_c/y_c), structures/output ratios (s_c/y_c), and the income shares received by equipment and structures ($IS_{e,c}$ and $IS_{s,c}$).

This gives us 14 data moments (6 independent factor incomes shares, 2 output levels, 2 wage gains at migration, 4 capital/output ratios).

2.6.1 Data

The new data moments relate to equipment and structures inputs and their income shares. These moments are constructed as follows.

Table 1: Calibration Targets

	ln y	k/y	e/y	s/y
Rich	11.60	3.18	0.37	2.81
Poor	9.12	2.66	0.14	2.85
Rich/poor	12.00	1.19	2.62	0.98

$IS_{e,r} = 0.15$ is taken from Valentinyi and Herrendorf (2008). Together with a labor share of 1/3, this implies $IS_{s,r} = 0.18$, which is consistent with Valentinyi and Herrendorf (2008). We do not have data on equipment and structures shares for low income countries. Since we find that s/y and p_s/p_c do not vary much between rich and poor countries, we start by assuming that $\alpha_c = 0.18$ for all countries.

We construct the stocks of equipment and structures (e_c and s_c) from Penn World Table 9 (Feenstra et al., 2015) and International Comparison Project 2011 data for the year 2011. We take the rich country to be the U.S. The poor country is the median of the 63 countries with $y_c/y_r < 1/4$ (the cutoff is consistent with the other calibration targets that we take from Hendricks and Schoellman 2018).

The data moments are summarized in Table 1. We highlight two findings:

1. k/y and s/y do not increase much in y.
2. e/y is 2.6 times higher in rich versus poor countries.

2.6.2 Estimating Human Capital Gaps

Before calibrating the model, we note that some results may be obtained with far less structure than our model imposes. In particular, the ratios of rich to poor country human capital per worker, $R(h_j)$, can be estimated assuming only that workers are paid their marginal products, so that observed wages are given by $w_{j,c} = p_{j,c} h_{j,c}$, and that migrant wage gains are given by $R(p_j)$.

Claim 1. Relative human capital endowments are given by

$$R(h_j) = \frac{R(y)}{wg_j} \frac{R(IS_j)}{R(N_j)} \tag{15}$$

Proof. The derivation is immediate:

$$wg_j = R(p_j) = R(p_j L_j)\, R(L_j)^{-1} \tag{16}$$

$$= \underbrace{R\left(\frac{p_j L_j}{y}\right)}_{R(IS_j)} R(y)\, R(h_j N_j)^{-1} \tag{17}$$

Table 2: Decomposing rich-to-poor country human capital gaps

	$R(h_1)$	$R(y_1)/wg_1$	$R(IS_1)/R(N_1)$	$R(h_2)$	$R(y_2)/wg_2$	$R(IS_2)/R(N_2)$
some HS	2.1	2.9	0.71	3.7	4.7	0.8
HSG	2.7	3.1	0.88	3.6	4.8	0.75
SC	3.2	3.6	0.9	3.7	5.2	0.71
CG	3.7	3.8	0.97	3.7	5.2	0.71

Notes: The table shows the rich-to-poor country human capital ratios, $R(h_j)$, and their components according to Claim 1.

□

Claim 1 gives rich-to-poor country human capital ratios in terms of observables. Intuitively, wage gains give the ratio of unobserved skill prices $R(p_j)$. Income shares and hours imply wages per hour: $w_{j,c} = IS_{j,c}/N_{j,c}$. The ratio of these wages gives $R(w_j) = R(p_j h_j)$. Finally, relative human capital levels are given by the ratio of wages to skill prices: $R(h_j) = R(w_j)/R(p_j)$.

An alternative intuition notes that wage ratios $R(w_j)$ may be inferred from labor income share and output gap data. It remains to identify whether rich country wages are higher because of human capital gaps or because of skill price gaps: $R(w_j) = R(h_j)R(p_j)$. Wage gains at migration identify $R(p_j)$ and solve this identification problem.

Table 2 shows the values of $R(h_j)$ implied by (15). The table also shows the two ratios that determine these values: $R(y)/wg_j$ and $R(IS_j)/R(N_j)$.

A key feature of the data is that the ratio of labor incomes shares is approximately equal to the ratio of hours. Hence, the second ratio in (15) is always close to 1. Hence, the human capital ratio $R(h_j)$ is always close to the ratio of income gaps to wage gains.

From Table 2 we highlight two results:

1. The ratio of rich to poor human capital per worker is around 3 in all cases. This means that human capital must be important for output gaps – a point that we make precise below.

2. There are no dramatic differences in the relative human capital of skilled versus unskilled workers across countries. $R(S(h))$ varies by at most factor 1.8. This is important because the mechanism emphasized by Jones (2014) works off scarcity of skilled labor caused, in part, by low $S(h)$ in poor countries. If the ratio of skilled to unskilled human capital is similar across countries, the force of this mechanism is limited.

Table 3: Bounds For Human Capital Shares

	h lower bound	h upper bound	hN lower bound	hN upper bound
some HS	0.3	0.56	-0.82	0.82
HSG	0.42	0.55	0.039	0.95
SC	0.49	0.55	0.27	1.1
CG	0.55	0.56	0.49	1.1

Note: The table shows the bounds for share_h and share_L given by (18) and (19).

2.6.3 Bounding the contribution of human capital

Given only information about rich to poor human capital ratios, it is possible to bound the contribution of human capital to output gaps.

Claim 2. Assume a constant returns to scale labor aggregator of the form $L_c = F\left(h_{j,c}N_{j,c}; \theta_{j,c}\right)$. Then the contribution of h to output gaps is bounded according to

$$\frac{\ln R\left(h_1\right)}{\ln R\left(y\right)} < \text{share}_h < \frac{\ln R\left(h_2\right)}{\ln R\left(y\right)} \tag{18}$$

The joint contribution of h and N is bounded according to

$$\frac{\ln R\left(h_1 N_1\right)}{\ln R\left(y\right)} < \text{share}_h + \text{share}_N < \frac{\ln R\left(h_2 N_2\right)}{\ln R\left(y\right)} \tag{19}$$

Both bounds apply to the "direct" shares of labor inputs, holding skill bias fixed.

Proof. subsubsection B.1.3. □

The bounds implied by (18) have bite. $2 < R\left(h_j\right) < 3.75$ implies $0.31 < \text{share}_h < 0.56$; see Table 3. The lower bound is the more interesting one because the total contribution of human capital also includes that of years of schooling. As expected, the bounds on the joint contribution of h and N are too wide to be useful.

2.6.4 Identification

We normalize $h_{1,r} = 1$ so that $L_{u,r} = N_{u,r}$. Choosing units of h_2 normalizes $\mu_2 = 1$. We may normalize μ_e, κ_2 and B to 1 (they are equivalent to varying z). Choosing units of e normalizes $\kappa_1 = 1$. It also means that we need to replace the data moments e_c/y_c with $R\left(e\right)$, leaving us 13 data moments.

The model then contains 14 calibrated parameters:

1. 2 productivity parameters z_c;

2. 2 structures shares α_c;
3. 3 human capital levels $h_{j,c}$, with $h_{1,r} = 1$;
4. 4 capital stocks e_c and s_c;
5. 1 frontier curvature ω;
6. 2 substitution elasticities, governed by ρ, ϕ.

Hence, we lack one data moment to identify all model parameters.

2.7 Reduced form labor aggregator

Fortunately, we can identify enough model parameters to perform levels accounting. Central to this result is that the model admits a reduced form production function of the form

$$y_c = k_c^{\alpha} \left(z_c \tilde{L}_c \right)^{1-\alpha} \tag{20}$$

where the labor aggregator is given by

$$\tilde{L} = B \left([L_1/\kappa_1]^{\Psi} + [Z/\kappa_2]^{\Psi} \right)^{1/\Psi} \tag{21}$$

with an elasticity of substitution governed by $\Psi = \omega\rho / (\omega - \rho) > \rho$. The inequality stems from the symmetry condition (10) which implies that $\omega > \rho$ and therefore $\Psi > \rho$.

Proof. subsubsection A.4.3 □

The basic intuition is that firms have two ways of substituting between skilled and unskilled labor. In addition to the standard movement along an isoquant, firms can also move along the technology frontier. When skilled labor is abundant, firms invest in skilled labor augmenting technology and $S(\theta)$ increases. This has a similar effect as raising the elasticity of substitution between skills. As a result, the substitution elasticity of the reduced form aggregator (40) is greater than that of the original aggregator (33) (since $\Psi > \rho$).

The result that the reduced form production function features a constant elasticity of substitution depends on the functional form of the technology frontier. It implies implies that the optimal skill bias ratio is a constant elasticity function of relative inputs:

$$S(\theta)^{\omega-\rho} = S(\kappa)^{-\omega} (Z/L_1)^{\rho} \tag{22}$$

This follows directly from the firm's first-order conditions for $\theta_{j,c}$ (see subsection A.2).

Table 4: Levels accounting summary

	1.50	1.75	2.00	2.50	Total
some HS	0.49	0.53	0.56	0.59	0.65
HSG	0.52	0.55	0.58	0.61	0.61
SC	0.56	0.59	0.61	NaN	0.62
CG	NaN	NaN	NaN	NaN	NaN

Note: The table shows the total share of human capital in cross-country output gaps. Columns show substitution elasticities $\sigma = 1/(1-\rho)$.

The reduced form model contains 13 calibrated parameters; one fewer than the baseline model. The reason is that the two separate elasticities governed by ρ and ω have been replaced by the reduced form elasticity Ψ. The model is therefore identified using the data moments listed in subsection 3.4 (a calibration algorithm is given in subsubsection A.4.4).

An important implication is that the total effect of labor inputs no longer depends on the elasticity of substitution between skilled and unskilled labor governed by ρ. However, the direct effect of labor inputs still depends on ρ.

2.8 Quantitative Results

Table 4 summarizes levels accounting results across skill cutoffs (rows) and substitution elasticities (columns). The last column shows the total share of human capital (including the effect via endogenous skill bias) which does not depend on the elasticity of substitution. The model cannot be calibrated for the college graduate skill cutoff and for high substitution elasticities. These cases would violate the symmetry condition (10).

We find a narrow range for the total effect of human capital, ranging from 61% to 65% of output gaps. These figures are close to the findings of Hendricks and Schoellman (2018) who study a model without skill bias differences across countries and without capital-skill complementarity.

The direct contribution of human capital to output gaps ranges from 49% to 61%. It increases with the elasticity of substitution and with the skill cutoff.

Table 5 show the relative skill bias gaps $R(S(\theta))$ for each case. Cases with direct effects below 50% imply very large cross-country variation in the relative productivity of skilled labor. By comparison, differences in the relative human capital of skilled labor tend to be modest (at most a factor of 1.8; see Table 2). It follows that at least 70% of the cross-country variation in the relative productivity of skilled labor is due to skill bias rather than relative human capital differences.[3]

[3] A natural decomposition of relative skilled labor productivity gaps is given by $\ln R(S(\theta h)) = \ln R(S(\theta)) + \ln R(S(h))$.

Table 5: Relative skill bias

	1.50	1.75	2.00	2.50
some HS	316.90	34.72	11.49	3.80
HSG	8.14	2.94	1.77	1.07
SC	4.26	1.88	1.24	NaN
CG	NaN	NaN	NaN	NaN

Note: The table shows $R(S(\theta))$.

2.8.1 Intuition

To gain intuition for the results shown in Table 5, we derive a solution for the share of human capital in cross-country income gaps in terms of data moments and e/L_s.

Claim 3. The "total" share of human capital in cross-country output gaps is given by

$$\text{share}_L = \frac{\ln R(L_1)}{\ln R(y)} + \frac{1}{\Psi} \frac{\ln R\left(1 + S(W)\left(1 + [e/L_2]^{\phi}\right)\right)}{\ln R(y)} \tag{23}$$

where

$$\Psi = \frac{\ln(R(S(W))) - \ln\left(R\left(\frac{1}{1+1/IR_{se}}\right)\right)}{\ln(R(Z/L_1))} \tag{24}$$

and

$$R(Z) = R(e)\, R\left((1 + IR_{se})^{1/\phi}\right) \tag{25}$$

with ϕ given by (31) (see Claim 4). Here, we abuse notation slightly. In (23), $R(e/L_2)$ compares the counterfactual value, which holds q_e constant, with the poor country value.

Proof. The labor aggregator may be written as

$$L_c = L_{1,c}\left[1 + (Z/L_1)^{\Psi}\right]^{1/\Psi} \tag{26}$$

Using (86) and (71) we have

$$L_c = L_{1,c}\left[1 + S(W)\left(1 + [e/L_2]^{\phi}\right)\right]^{1/\Psi} \tag{27}$$

Hence we have (23). The solution for Ψ follows from (86) and (71). The solution for $R(Z)$ follows from (71). □

The human capital share is determined by three factors.

1. $R(L_1)$: The rich-to-poor ratio of unskilled labor inputs increases sharply with the skill cutoff. The intuition is that rich countries have very few workers with less than some

high school education. Hence, $R(L_1)$ is small for the "some high school" skill cutoff. But even rich countries have many workers with less than a college degree. Since these workers have more human capital than poor country workers, $R(L_1)$ is large for the CG skill cutoff.

2. The reduced form elasticity of substitution governed by Ψ is always large because it has to reconcile large cross-country variation in Z/L_1 with small differences in observed skill premiums. Hence, $1/\Psi$ is always close to 1 and does not vary much with the skill cutoff.

3. The final term in (23) decreases with the skill cutoff. The intuition is again that rich countries have very few workers with less than some high school education. Hence, $S(W)$ is very large relative to poor countries for the "some HS" skill cutoff. For the CG skill cutoff, rich countries still have far more workers than poor countries, and therefore $S(W)$ is much higher. But the level of $S(W)$ is small in both countries because college graduates earn only a small fraction of total labor income. Hence, $1 + S(W)$ is close to 1.

It is the offsetting variation in $R(L_1)$ and the final term that generates the rough constancy of share$_L$. This intuition can be made more precise for a model without capital-skill complementarity for which a closed for solution for share$_L$ can be derived (see subsection 3.7).

2.9 Elasticity Implications

A challenge for models of human capital is to reconcile two empirical observations that appear at odds with each other.

1. Relative labor inputs $S(N_c)$ vary greatly between rich and poor countries. Yet skill premiums $S(w_c)$ are similar, suggesting that skilled and unskilled labor are highly substitutable.

2. However, empirical estimates of the elasticity of substitution between skilled and unskilled labor are not high. Most estimates cluster around 2 (e.g., Ciccone and Peri 2005).

In a model with fixed skill bias, these observations are difficult to reconcile. To see this problem, note that the observable skill premium is given by

$$S(w_c) = S(p_c h_c) = S(\theta_c h_c)^{\rho} S(N_c)^{\rho-1} \tag{28}$$

Data suggest that observable skill premiums (the LHS of (28)) are similar across countries. Therefore, if $\theta_{j,c} = 1$, we have $h_s/h_u \propto (N_s/N_u)^{(1-\rho)/\rho}$ and $p_s/p_u \propto (N_s/N_u)^{(\rho-1)/\rho}$. With $\sigma = 1/(1-\rho) \approx 2$, we have $(1-\rho)/\rho \approx 1$ and $p_s/p_u \propto N_s/N_u$.

It follows that skill price premiums $S(p_c)$ vary greatly across countries. But this is inconsistent with the observations that wage gains at migration do not vary dramatically with skill (Hendricks and Schoellman, 2018). To see this, note that the ratio of wage gains at migration for skilled versus unskilled workers is given by $S(p_r)/S(p_p)$.

Our model offers a solution to the apparent contradiction. Empirical estimates reveal the "short-run" elasticity of substitution, governed by ρ, holding skill bias fixed. But what governs cross-country variation in skill prices is the "long-run" elasticity, governed by Ψ. The long-run elasticity is larger than the short-run elasticity because it accounts for movements along the technology frontier.

2.10 Robustness: Exogenous Skill Bias

As a first robustness check, we consider the model without the technology frontier. We are agnostic as to how the values of $\theta_{j,c}$ is determined. We make three main points:

1. Given the data moments described in subsubsection A.4.4, the model is only identified for given ρ.
2. For given ρ the model implies the same calibrated parameters as the model with the technology frontier (up to normalizations that do not affect levels accounting).
3. Therefore, the direct effects of the model with the technology frontier shown in Table 4 are now the total effects of labor inputs. Labor inputs therefore account for 49% to 61% of cross-country output gaps.

2.10.1 Identification

We show that the model is only identified for given ρ. We may normalize several model parameters. To see this clearly, it is useful to rewrite the labor aggregator as

$$L_c = \theta_{1,c}\left[(L_{1,c})^{\rho} + \left(\frac{\theta_{2,c}}{\theta_{1,c}}Z_c\right)^{\rho}\right]^{1/\rho} \tag{29}$$

and the Z aggregator as

$$Z_c = \mu_e\left[(e_c)^{\phi} + \left(\frac{\mu_2}{\mu_e}L_{2,c}\right)^{\phi}\right]^{1/\phi} \tag{30}$$

It is now apparent that we may normalize the following parameters:

1. $\mu_e = 1$ and $\theta_{1,r} = 1$ (they are equivalent to varying z).

2. Choosing units of h_2 normalizes $\mu_2 = 1$.
3. Choosing units of e normalizes $\theta_{2,r} = 1$. This normalization means that we no longer observe e_c but only $R(e)$.
4. We normalize $h_{1,r} = 1$ so that $L_{u,r} = N_{u,r}$.
5. Finally, some normalization of $\theta_{1,p}$ is needed to define TFP (i.e., to distinguish varying z from varying both $\theta_{j,c}$ in proportion). One such normalization is the technology frontier.

This leaves us with 14 calibrated parameters (2 z_c, 2 α_c, 3 $h_{j,c}$, 2 equipment levels e_c, 2 structures levels, and ρ, ϕ, $\theta_{2,p}/\theta_{1,p}$). Since we only have 13 data moments, one model parameter is not identified.

However, a subset of the parameters is identified. We find that these parameters are the same as in the baseline model, up to a normalization that defines TFP and does not affect the levels accounting results.

Claim 4. The curvature of the Z aggregator ϕ is given by

$$R(IR_{se}) = R(L_2/e)^{\phi} \tag{31}$$

It is independent of ρ and of how the $\theta_{j,c}$ are determined.

Proof. Take the rich/poor ratio of the income ratio of skilled labor to equipment from (62) to obtain (31). $R(e)$ is observable. $R(L_2)$ is pinned down by wage gains. Hence, this equation identifies ϕ. This only depends on the equipment/skilled labor aggregator. □

Claim 5. The Z aggregator and its inputs (e and $L_{2,c}$) are identified up to a common scale factor, regardless of ρ and of how the $\theta_{j,c}$ are chosen.

Proof. subsubsection A.3.2 □

The scale factor stems from a choice of units (of equipment and skilled labor and thus of Z). Hence, we can identify the relative scarcity of skilled labor versus equipment, $S(L_2/e)$, but not the relative scarcity of their composite Z versus unskilled labor. This depends on ρ, as the next Claim shows.

2.10.2 Levels accounting

We show that, for any given ρ, the model implies the same parameters as the model with the technology frontier up to a normalization that does not affect levels accounting.

Claim 6. For a given ρ and normalizing $\theta_{1,c}$ to define what neutral TFP z is, and choosing units of skilled labor to normalize $\theta_{2,r} = 1$, the model implies the same calibrated parameters and thus labor aggregator, regardless of how $\theta_{j,c}$ are chosen (with or without technology frontier).

Proof. subsubsection A.3.3 □

It follows that the contribution of human capital to output gaps, $R(L)$, holding $\theta_{j,c}$ fixed, is independent of how $\theta_{j,c}$ are chosen. It means that the "direct" share$_L$ is the same for the cases of fixed $\theta_{j,c}$ and endogenous $\theta_{j,c}$.

3 Robustness: No Capital-Skill Complementarity

As a further robustness check, we now consider a model with a single capital stock k_c and no capital-skill complementarity. This is a special case of the baseline model where $\mu_e = 0$ so that $Z_c = L_{2,c}$.

Qualitatively and quantitatively, the results of the baseline model remain valid. A summary is as follows:

1. The model is only identified for given ρ.
2. However, a reduced form model may be derived that is identified. The reduced form labor aggregator has the same elasticity of substitution as in the baseline model (governed by Ψ).
3. Therefore, the total share of labor inputs in output gaps is independent of ρ, but the direct share does depend on ρ.
4. Quantitatively, we find total shares of labor inputs near 60% for all skill cutoffs.
5. We find direct shares of labor inputs between 45% and 60%.

3.1 Model

Output per worker y_c is produced according to the production function

$$y_c = k_c^{\alpha} \left(z_c L_c\right)^{1-\alpha} \tag{32}$$

where

$$L_c = CES\left(L_{j,c}, \theta_{j,c}, \rho\right) = \left[\sum_{j=1}^{J} \left(\theta_{j,c} L_{j,c}\right)^{\rho}\right]^{1/\rho} \tag{33}$$

The skill bias parameters $\theta_{j,c}$ are constrained by the technology frontier 8. This nests Caselli and Coleman (2006) when $h_j = 1$ and $B_c = 1$. It nests Jones (2014) when the choice of skill bias is removed and $\theta_j = 1$.

3.2 Equilibrium

Firms hire capital at rental price q_c and labor at rental prices $p_{j,c}$. The representative firm solves

$$\max_{k_c, L_{j,c}, \theta_{j,c}, B_c} y_c - q_c k_c - \sum_j p_{j,c} L_{j,c} - C(B_c) \tag{34}$$

subject to (32), (33), and (8).

A competitive equilibrium is an allocation $y_c, k_c, \theta_{j,c}, B_c$ and a set of prices $p_{j,c}$ that satisfy

- firm first-order conditions for k_c, $L_{j,c}$ and $\theta_{j,c}$;
- first-order condition for B_c, if chosen; otherwise $B_c = 1$;
- the aggregate production function.

3.3 Levels accounting

We only consider the case where the model has constant returns to scale (except in subsection 3.9 where we allow firms to invest in B_c). This is the case when B_c is fixed or when b scales appropriately with output. We then use the usual expression for levels accounting (see Hendricks and Schoellman 2018):

$$y_c = z_c \left(k_c / y_c\right)^{\alpha/(1-\alpha)} L_c \tag{35}$$

The contributions of TFP, physical capital, and human capital to output gaps are defined analogously to the baseline model:

$$\underbrace{\ln R\left(y\right)}_{\text{output gap}} = \underbrace{\ln R\left(z\right)}_{\text{TFP}} + \underbrace{\ln R\left(\left(k/y\right)^{\alpha/(1-\alpha)}\right)}_{\text{physical capital}} + \underbrace{\ln R\left(L\right)}_{\text{labor inputs}} \tag{36}$$

The share of the output gap accounted for by each input is given by

$$1 = \underbrace{\frac{\ln R\left(z\right)}{\ln R\left(y\right)}}_{\text{share}_z} + \underbrace{\frac{\ln R\left(\left(k/y\right)^{\alpha/(1-\alpha)}\right)}{\ln R\left(y\right)}}_{\text{share}_k} + \underbrace{\frac{\ln R\left(L\right)}{\ln R\left(y\right)}}_{\text{share}_L} \tag{37}$$

Labor inputs vary for three reasons: h, N, θ. For notational convenience, define

$$G(a,b,c) = CES(h_{j,a}N_{j,b},\theta_{j,c};\rho) \tag{38}$$

as the labor input with human capital per worker from country a, hours worked from country b, and skill bias from country c, so that $L_c = G(c,c,c)$. Then we may decompose cross-country differences in labor inputs using

$$R(L) = \underbrace{\frac{G(r,r,r)}{G(r,r,p)}}_{\text{contribution } \theta}\underbrace{\frac{G(r,r,p)}{G(r,p,p)}}_{N}\underbrace{\frac{G(r,p,p)}{G(p,p,p)}}_{h} \tag{39}$$

Then the share of each labor input is defined in the obvious way. For example, share$_\theta$ = $\ln(G(r,r,r)/G(r,r,p))/\ln R(y)$. Hence, share$_L$ = share$_h$ + share$_N$ + share$_\theta$.

As in the baseline case, there is a "direct" share of labor inputs that holds skill bias constant, and a "total" share that includes induced changes in skill bias.

3.4 Calibration

We calibrate the model using the same data moments that we used for the baseline model. However, we replace the separate moments for equipment and structures with moments for the total capital stock (taken from Hendricks and Schoellman 2018). This leaves us with 8 data moments: 2 outputs, 4 factor incomes shares, 2 wage gains at migration.

Since the model has 9 calibrated parameters ($h_{j,c}$, α_c, z_c, ρ, ω where $h_{1,r}$ is normalized to 1), the model is only identified for given ρ.

3.5 Reduced form production function

Fortunately, similar to the baseline model, we can derive a reduced form model where the labor aggregator is given by

$$\tilde{L}_c = CES\left(L_{j,c},\kappa_j^{-1},\Psi\right) = \left[\sum_j \left(\kappa_j^{-1}L_{j,c}\right)^{\Psi}\right]^{1/\Psi} \tag{40}$$

Proof. subsubsection B.3.4 □

The aggregate production function looks exactly like that of Jones (2014). The skill bias parameters are common across countries and governed by the technology frontier $\left(\kappa_j^{-1}\right)$. The substitution elasticity now depends on the curvatures of the original labor aggregator

(33) and of the technology frontier. Variation in the level of the frontier B has the same effect as variation in z.

The intuition why a CES reduced form labor aggregator exists parallels the baseline model. It is again the case the the optimal skill bias ratio is a constant elasticity function of relative labor inputs:

$$S(\theta)^{\omega-\rho} = S\left(\kappa^{-\omega} L^{\rho}\right) \tag{41}$$

(see subsubsection B.3.1). Allowing firms to invest in skill biased technology when skilled labor is scarce has the same effect as increasing the elasticity of substitution between skilled and unskilled labor.

It follows that the model has the same implications for levels accounting as the model of Jones (2014) when calibrated to the data moments listed in subsection 3.4. Therefore, the levels accounting results of Hendricks and Schoellman (2018) remain valid.

3.6 Levels Accounting Implications

We gain intuition about the levels accounting results by solving in closed form for the direct and the total effect of labor inputs as functions of observable data moments.

Claim 7. The direct effect of human capital is given by

$$DE = R(L_1) \frac{\left[1 + S(W)_p \left(S\left(R\left(L\right)\right)\right)^{\rho}\right]^{1/\rho}}{\left[1 + S(W)_p\right]^{1/\rho}} \tag{42}$$

where $R(L_1)$ is given by (15).

Proof. Immediate from (109). □

If the rich country is abundant in skilled labor in the sense that $S(R(L)) \geq 1$ (or $R(L_2) > R(L_1)$), the direct effect of human capital is bounded below by $R(L_1)$ and increasing in ρ. This is intuitive and, in fact, implied by the "non-parametric" bounds of Table 3. The fact that DE is increasing in ρ is implied by $S(R(L)) \geq 1$.

Claim 8. The "total" share of human capital is given by

$$\text{share}_L = \frac{\ln(R(L_1))}{\ln(R(y))} + \frac{1}{\Psi} \frac{\ln(R(1+S(W)))}{\ln(R(y))} \tag{43}$$

where $R(L_1)$ is given by (15) and $1/\Psi = \ln(S(R(L))) / \ln(R(S(W)))$.

Table 6: Human Capital Shares

	1.50	1.75	2.00	2.50	Total
some HS	0.46	0.50	0.53	0.56	0.63
HSG	0.48	0.51	0.52	0.55	0.59
SC	0.51	0.53	0.54	0.56	0.60
CG	0.56	0.56	0.57	0.57	0.58

Proof. Using (137), the reduced form labor aggregator (40) may be written as

$$L_c = L_{1,c}\left[1 + S\left(L\right)^{\Psi}\right]^{1/\Psi} \tag{44}$$

$$= L_{1,c}\left[1 + S(W)\right]^{1/\Psi} \tag{45}$$

Hence,

$$\ln\left(R\left(L\right)\right) = \ln\left(R\left(L_1\right)\right) + \frac{1}{\Psi}\ln\left(R\left(1 + S(W)\right)\right) \tag{46}$$

Also using (137), we have the solution for Ψ. □

Intuitively, the reduced form elasticity must reconcile the observed variation in relative labor inputs, $S\left(R\left(L\right)\right) = R\left(S\left(L\right)\right)$, with the observed variation in skill premiums or labor income shares, $R\left(S(W)\right)$. This result is useful for gaining intuition about how the human capital share varies across skill cutoffs.

3.7 Quantitative Results

Table 6 shows the share of output gaps accounted for by human capital, defined in (39). Each row represents a skill cutoff. Each column shows the direct contribution of labor inputs for a fixed substitution elasticity $1/\left(1-\rho\right)$. The last column shows the "total" contribution of labor inputs, which does not depend on ρ.

Consistent with 7, a larger elasticity, governed by ρ, implies a larger share due to human capital. As in the baseline model, the human capital shares below one-half imply very large differences in skill bias across countries. For cutoff HSD and elasticity 1.5, $R\left(\theta_s/\theta_u\right) = 550$. For cutoff HSG, the ratio is still 20. See Table 7.

The last column of Table 6 shows the total effect of human capital: the sum of the direct effect and of the effect on skill bias. It does not vary with ρ. It is always close to 60% and coincides with the findings of Hendricks and Schoellman (2018) obtained from a model without skill bias variation.

To gain intuition why the skill cutoff does not matter much (in contrast to Jones 2014), Table 8 shows the components of share_L from (43).

Table 7: Relative Skill Bias Rich vs. Poor

	1.50	1.75	2.00	2.50
some HS	557.7	57.6	18.5	6.0
HSG	20.0	6.3	3.6	2.0
SC	17.2	5.9	3.5	2.0
CG	6.4	3.1	2.1	1.5

Note: The table shows $R(S(\theta))$.

Table 8: Closed form solution for share$_L$

	shareL	L1 term	$1/\Psi$	S(W) term	S(W) rich	S(W) poor
some HS	0.78	-0.82	1.3	1.3	58	1.9
HSG	0.75	0.039	1.6	0.44	3.7	0.66
SC	0.76	0.27	1.8	0.27	1.4	0.29
CG	0.67	0.49	2.5	0.07	0.3	0.1

- The term that involves $R(L_1)$ increases strongly with the skill cutoff. Intuition: Rich countries have very few workers with less than some high school education relative to poor countries. But all countries have a substantial fraction of workers with less than a college degree.

- The reduced form curvature does not vary much with the cutoff. Intuition: Regardless of cutoff, the lack of correlation between skill premiums and relative labor inputs requires a high substitution elasticity.

- The final term that involves $R(1+S(W))$ decreases with the cutoff. Intuition: similar to why the first term decreases, but in reverse. Rich countries have very few workers with less than some HS. Hence, $S(W)$ is very large relative to poor countries. For the CG cutoff, rich countries still have far more workers than poor countries, and therefore $S(W)$ is much higher. But the level of $S(W)$ is small in both countries. Hence, $1+S(W)$ is close to 1.

- It is the offsetting variation in $R(L_1)$ and the final term that generates the rough constancy of share$_L$.

The model of Jones (2014) implies an analogous expression for the share of human capital. Since

$$L_c = L_{1,c}\left[1+S(W)_c\right]^{1/\rho} \tag{47}$$

we have

$$\text{share}_L = \frac{\ln(R(L_1))}{\ln(R(y))} + \frac{1}{\rho}\frac{\ln(R(1+S(W)))}{\ln(R(y))} \tag{48}$$

He obtains very different results from ours because his calibration approach differs. First, Jones assumes that $R(L_1)$ is smaller by fixing $R(h_1)$ at 1 or close to 1. This reduces share$_L$. Second, instead of the reduced form curvature parameter Ψ, his model of fixed skill bias implies that the second term in (48) is multiplied by the short term curvature $1/\rho$. This is why the value of ρ is important for his conclusions. It is also apparent why the cutoff is important, especially when ρ is small.

3.8 Exogenous Skill Bias

As further robustness check, we consider a model without capital-skill complementarity and with exogenous skill bias. Similar to the baseline case, we find that the share due to labor inputs equals the "total" share of labor inputs from the model with the technology frontier.

Claim 9. For given ρ, data on labor income shares and wage gains identify $h_{j,c}$ and $\theta_{j,c}$, regardless of how the $\theta_{j,c}$ are determined. $\theta_{j,c}$ are determined up to a normalization for $\theta_{1,c}$ which distinguishes neutral TFP from skill bias changes.

Proof. We need to determine 4 $h_{j,c}$ and 4 $\theta_{j,c}$.

1. We may normalize $h_{j,r} = 1$ by choosing units of h. We know $R(h_j)$ from (15). Hence we know all $h_{j,c}$.
2. We may normalize $\theta_{1,c} = 1$, essentially defining TFP z_c.
3. The ratio of labor income shares from (102) gives $S(\theta_c L_c)$. Since $L_{j,c}$ are known, this can be solved for $S(\theta_c)$ and thus $\theta_{2,c}$.

□

The intuition is as follows. Wage gains give rich/poor L_j ratios. Skill premiums (or wage bill ratios) give skilled/unskilled θL ratios (the relative scarcity of skilled labor in efficiency units). Whether we attribute that scarcity to skill bias or to labor inputs is a harmless normalization. Only $S(\theta L)$ matters, not $S(\theta)$ and $S(L)$ separately. That point is perhaps not obvious, but shown next.

The reason why Claim 9 is important, in spite of all the normalizations that it imposes: Identifying $h_{j,c}$ and $\theta_{j,c}$ up to these normalizations is sufficient to perform levels accounting. This is shown next.

3.8.1 Levels Accounting Implications

Claim 10. For given ρ, the contribution of human capital to output gaps is identified using data on output gaps, wage gains, and labor income shares. The contribution is independent of how the $\theta_{j,c}$ are determined.

Proof: subsubsection B.1.4

This claim is really implied by the previous claim that all models imply the same $h_{j,c}$ and $\theta_{j,c}$, up to a harmless normalization. But it is useful to spell this out directly. Estimating the contribution of θ requires a normalization that distinguishes TFP from skill bias variation. Without more structure on the model, there is no meaningful way of doing this. We can say, however:

Claim 11. For a given normalization that fixes the scale of the $\theta_{j,c}$, such as $\theta_{1,c} = 1, \forall c$, the contribution of θ is identified, regardless of how $\theta_{j,c}$ are determined.

Proof: subsubsection B.1.5

Without more structure on the model, the normalization is arbitrary. Only when we put more structure on the model, via a technology frontier, can we compare outcomes across values of ρ. It follows that the effect of human capital is given by the direct effects in Table 6.

3.9 Endogenous Technology Frontier

Finally, we consider a model where firm can expend resources to shift the technology frontier outwards, as in Acemoglu 2007.

Firms again solve (34), now choosing B_c as well. We now think of B_c as expenditure for changing skill weights. κ_j is the cost of changing the skill weight $\theta_{j,c}$. We can again normalize all $\kappa_j = 1$. We need $\omega > 1$ to ensure that optimal skill weights are finite.

The only change compared with the fixed frontier case is the level of B_c and thus in the level of skill bias. Effectively, we make B_c endogenous but change nothing conditional on B_c. All the relative quantities and prices are the same as in the baseline model. What changes is that the level of productivity responds to labor supplies.

We show the following main points:

1. Whether the model implies increasing or constant returns to scale depends on the functional form of the technology that allows firms to increase B_c.

2. If the model has constant returns to scale (the cost of changing B scales appropriately with output), the results of the fixed frontier model remain valid.

3. If the model has increasing returns to scale, share_h is magnified.

3.9.1 Reduced form production function

The reduced form human capital aggregator is given by

$$L_c = \left(\left(1-\alpha\right) z_c^{1-\alpha} k_c^{\alpha} \omega^{-1} b^{-1}\right)^{\frac{1}{\omega+\alpha-1}} \tilde{L}_c^{\frac{\omega}{\omega+\alpha-1}} \tag{49}$$

where $\tilde{L}_c$ is again given by (40). The reduced form production function is then given by

$$y_c = \left(k_c^{\alpha}\left(Az_c\tilde{L}_c\right)^{1-\alpha}\right)^{\frac{\omega}{\omega+\alpha-1}} \tag{50}$$

where $A = \left(\frac{1-\alpha}{\omega b}\right)^{1/\omega}$ is a constant.

Proof. subsubsection B.6.3 and subsubsection B.6.4 □

The key differences relative to the baseline model's (20):

1. We have increasing returns to scale due to a scale effect. Larger population (or higher productivity; or more capital) increases the benefits from productivity, but not the cost. Optimal B_c increases. This amplifies the effects of anything that increases output equally (given by the exponent in (50)).
2. The scale effect is eliminated if the cost of investing in B_c scales appropriately with output. To see this, note that $A = \left(\frac{1-\alpha}{\omega b}\right)^{1/\omega}$ in (50). If $b = y^{\varphi}$ (but taken as given by firms), then the production function reverts to the one for the fixed frontier (20) for the right value of φ. In that case, investment in the frontier has no impact on levels accounting.

3.9.2 Levels accounting

For levels accounting, rewrite (50) as

$$y_c = \left[\left(k_c/y_c\right)^{\alpha/(1-\alpha)} Az_c\tilde{L}_c\right]^{\frac{\omega}{\omega-1}} \tag{51}$$

Proof. subsubsection B.6.5 □

Now the only difference relative to the case where B_c is fixed is the exponent $\omega/\left(\omega-1\right)$. To perform levels accounting, it is necessary to know the values of ω and ρ, not just the reduced form elasticity governed by Ψ. Identifying both values requires an additional data moment.

Relative to the contribution of labor inputs to output gaps is amplified by a constant factor, $\omega/\left(1-\omega\right)$.

Claim 12. The share of output gaps accounted for by labor inputs is given by

$$\text{share}_L = \frac{\omega}{\omega - 1} \frac{\ln R\left(\tilde{L}\right)}{\ln R\left(y\right)} \tag{52}$$

where $\tilde{L}_c$ takes on the same value as in the model without investment in the frontier.

Proof. Let $A_c = \left(k_c/y_c\right)^{\alpha/(1-\alpha)} Az_c$ collect all country specific terms other than labor inputs. Then $y_c = \left[A_c \tilde{L}_c\right]^{\frac{\omega}{\omega-1}}$ and

$$\frac{\omega - 1}{\omega} \ln R\left(y\right) = \ln R\left(A\right) + \ln R\left(\tilde{L}\right) \tag{53}$$

This implies (52). Since the calibrated values of $h_{j,c}$ and Ψ do not depend on whether or not B_c is endogenous (see subsection B.4), the labor aggregator is the same as in the model with fixed B_c. □

The empirical literature suggests that departures from constant returns to scale are not large. If this is the case, then allowing firms to invest in shifting the technology frontier does not imply substantial changes to our levels accounting results.

4 Conclusion

To be written.

References

ACEMOGLU, D. (2002): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 1–67.

——— (2007): "Equilibrium bias of technology," *Econometrica*, 75, 1371–1409.

CASELLI, F. (2016): *Technology Differences over Space and Time*, Princeton University Press.

CASELLI, F. AND W. J. COLEMAN (2006): "The World Technology Frontier," *American Economic Review*, 96, 499–522.

CICCONE, A. AND G. PERI (2005): "Long-run substitutability between more and less educated workers: Evidence from US states, 1950-1990," *Review of Economics and Statistics*, 87, 652–663.

FEENSTRA, R. C., R. INKLAAR, AND M. P. TIMMER (2015): "The next generation of the Penn World Table," *American economic review*, 105, 3150–82.

HALL, R. E. AND C. I. JONES (1999): "Why do some countries produce so much more output per worker than others?" *Quarterly Journal of Economics*, 114, 83–116.

HENDRICKS, L. AND T. SCHOELLMAN (2018): "Human Capital and Development Accounting: New Evidence From Immigrant Earnings," *Quarterly Journal of Economics*, 133, 665–700.

HSIEH, C.-T. AND P. J. KLENOW (2010): "Development Accounting," *American Economic Journal: Macroeconomics*, 2, 207–223.

JONES, B. F. (2014): "The Human Capital Stock: A Generalized Approach," *American Economic Review*, 104, 3752–77.

KATZ, L. F. AND K. M. MURPHY (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107, 35–78.

KRUSELL, P., L. E. OHANIAN, J.-V. RIOS-RULL, AND G. L. VIOLANTE (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68, 1029–1053.

ROSSI, F. (2019): "The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation," .

VALENTINYI, A. AND B. HERRENDORF (2008): "Measuring factor income shares at the sectoral level," *Review of Economic Dynamics*, 11, 820–835.

A Appendix: Capital-skill Complementarity

A.1 Equipment and Structures Data

We construct equipment and structures stocks by combining data from the Penn World Table 9 (Feenstra et al., 2015) and the International Comparison Project 2011. All data are constructed for year 2011, which is the latest and most comprehensive benchmark year for the ICP.

From the PWT, we obtain:

1. output per worker y as `cgdpo/emp`.
2. capital per worker k as `ck/emp`.
3. the price levels of capital pl_k `and consumption` pl_c.
4. the value of the equipment stock at local prices as `Kc_Mach + Kc_TraEq`.
5. the value of the structures stock at local prices as `Kc_Struc + Kc_Other` (from the capital detail file).

From ICP we obtain the PPP prices (series S03) of equipment (classification C20 Machinery and equipment) and structures (classification C21 Construction).

We define the stock of equipment as $e_c = (\text{Kc\_Mach+KcTraEq})\,/\text{emp}/PPP_{C20}$ and the stock of structures as $s_c = (\text{Kc\_Struc+Kc\_Other})\,/\text{emp}/PPP_{C21}$.

Before computing the calibration targets, we drop countries with missing output or employment data or with population (`pop`) < 1m. We also drop 6 countries with capital or consumption prices above 10 times the sample median. Finally, we drop 7 countries for which the discrepancy between k and $e + s$ is above 20%.

A.2 Firm first-order conditions

$s:$

$$\alpha y_c / s_c = q_{s,c} \tag{54}$$

$e:$

$$\frac{\partial y}{\partial L}\frac{\partial L}{\partial Z} Z^{1-\phi}\mu_e^{\phi} e^{\phi-1} = q_e \tag{55}$$

$L_2:$

$$\frac{\partial y}{\partial L}\frac{\partial L}{\partial Z} Z^{1-\phi}\mu_s^{\phi} L_s^{\phi-1} = p_{2,c} \tag{56}$$

L_1 :

$$\frac{\partial y}{\partial L} L^{1-\rho} \theta_{1,c}^{\rho} L_{1,c}^{\rho-1} = p_{1,c} \tag{57}$$

If there is a technology frontier, we also have

$\theta_{1,c}$:

$$\frac{\partial y}{\partial L} L^{1-\rho} \theta_{1,c}^{\rho-1} L_1^{\rho} = \lambda \omega \kappa_1^{\omega} \theta_{1,c}^{\omega-1} \tag{58}$$

$\theta_{2,c}$:

$$\frac{\partial y}{\partial L} L^{1-\rho} \theta_{2,c}^{\rho-1} Z^{\rho} = \lambda \omega \kappa_2^{\omega} \theta_{2,c}^{\omega-1} \tag{59}$$

where

$$\frac{\partial y}{\partial L} = (1-\alpha)\, y/L \tag{60}$$

$$\frac{\partial L}{\partial Z} = L^{1-\rho} \theta_2^{\rho} Z^{\rho-1} \tag{61}$$

A.3 Derivations for Arbitrary $\theta_{j,c}$

The following derivations hold regardless of how $\theta_{j,c}$ are chosen, including the cases where the $\theta_{j,c}$ are exogenous or chosen from a technology frontier. These expressions are used in subsequent derivations.

A.3.1 Income ratios and shares

Directly from the first order conditions for e and L_2:

$$IR_{se} = \left(\frac{\mu_2 L_2}{\mu_e e}\right)^{\phi} \tag{62}$$

Income of skilled versus unskilled labor: Start from

$$S(p) = \frac{\partial L}{\partial Z} \frac{Z^{1-\phi} \mu_2^{\phi} L_2^{\phi-1}}{L^{1-\rho} \theta_1^{\rho} L_u^{\rho-1}} \tag{63}$$

Therefore

$$S(W) = S(pL) = L^{1-\rho}\theta_{2,c}^{\rho} Z^{\rho-1} \frac{Z^{1-\phi}(\mu_2 L_2)^{\phi}}{L^{1-\rho}(\theta_{1,c}L_1)^{\rho}} \tag{64}$$

$$= \frac{Z^{1-\phi}\theta_{2,c}^{\rho} Z^{\rho-1}(\mu_2 L_2)^{\phi}}{(\theta_{1,c}L_1)^{\rho}} \tag{65}$$

$$= \left(\frac{\mu_2 L_2}{Z}\right)^{\phi}\left(\frac{\theta_2 Z}{\theta_1 L_1}\right)^{\rho} \tag{66}$$

Equipment income share:

$$IS_e = \frac{q_e e}{y} \tag{67}$$

$$= [(1-\alpha)\, y/L]\left[L^{1-\rho}\theta_2^{\rho} Z^{\rho-1}\right]\left[Z^{1-\phi}(\mu_e e)^{\phi}\right]/y \tag{68}$$

$$= (1-\alpha)\left[\frac{\theta_2 Z}{L}\right]^{\rho}\left[\frac{\mu_e e}{Z}\right]^{\phi} \tag{69}$$

A.3.2 Identification of Z Aggregator

Proof. [Claim 5]

We have $R(L_j)$ from wage gains (see (17)) and ϕ from (31).

Use (62) to solve for $\left[\frac{\mu_2 L_2}{\mu_e e}\right]^{\phi} = IR_{se}$. From the definition of Z we have

$$\left[\frac{Z}{\mu_2 L_2}\right]^{\phi} = 1 + \left[\frac{\mu_2 L_2}{\mu_e e}\right]^{-\phi} = 1 + 1/IR_{se} \tag{70}$$

Hence

$$Z/(\mu_2 L_2) = [1 + 1/IR_{se}]^{1/\phi} \tag{71}$$

and $Z/(\mu_e e) = [1 + IR_{se}]^{1/\phi}$ are identified given the normalization $\mu_2 = \mu_e = 1$. The ratio of rich versus poor Z is given by

$$R(Z) = R(Z/(\mu_e e))\, R(e) \tag{72}$$

$$= R\left([1 + IR_{se}]^{1/\phi}\right) R(e) \tag{73}$$

□

A.3.3 Identification, given ρ

Proof. [Claim 6]

With all $\theta_{j,r} = 1$, (66) gives $(Z_r/L_{1,r})^\rho$ and thus Z_r. Using $R(Z)$, we have Z_p. Using the results from Claim 5, all of the parameters of the Z aggregator are now identified. This includes e_c and $h_{2,c}$. Knowing $L_{1,p}$ and Z_p, we can solve (66) for $S(\theta_p)$. Given the normalization $\theta_{1,p}$, we have $\theta_{j,p}$. This fixes all the parameters required to calculate the labor aggregator. The structures shares then imply α_c and outputs imply z_c. □

A.4 Model With Technology Frontier

A.4.1 Implications of optimal skill bias choice

Claim 13. Optimal skill bias choice implies

$$S(\theta)^\omega = S(\kappa)^{-\omega}\left(\frac{Z}{L_1}\frac{\kappa_1}{\kappa_2}\right)^\Psi \tag{74}$$

Proof. From (22) we have

$$S(\theta)^\omega = S(\kappa)^{-\frac{\omega^2}{\omega-\rho}}\left(\frac{Z}{L_1}\right)^\Psi \tag{75}$$

$$= S(\kappa)^{\Psi-\frac{\omega^2}{\omega-\rho}}\left(\frac{Z}{L_1}\frac{\kappa_1}{\kappa_2}\right)^\Psi \tag{76}$$

Finally, note that

$$\Psi - \frac{\omega^2}{\omega-\rho} = \frac{\omega\rho-\omega^2}{\omega-\rho} = -\omega \tag{77}$$

□

Claim 14. Optimal skill bias choice (22) implies

$$\left(\frac{\theta_2 Z}{\theta_1 L_1}\right)^\rho = \left(\frac{Z}{L_1}\frac{\kappa_1}{\kappa_2}\right)^\Psi \tag{78}$$

Proof. From (22) we have

$$S(\theta)^\rho = S(\kappa)^{-\frac{\omega\rho}{\omega-\rho}}(Z/L_1)^{\frac{\rho^2}{\omega-\rho}} \tag{79}$$

Then

$$\left(\frac{\theta_2 Z}{\theta_1 L_1}\right)^\rho = \left(\frac{Z}{L_1}\right)^\rho S(\kappa)^{-\frac{\omega\rho}{\omega-\rho}}(Z/L_1)^{\frac{\rho^2}{\omega-\rho}} \tag{80}$$

and note that $\rho + \rho^2/(\omega-\rho) = \Psi$. □

From the definition of the labor aggregator:

$$\left(\frac{L}{\theta_2 Z}\right)^{\rho} = \left(\frac{\theta_1 L_1}{\theta_2 Z}\right)^{\rho} + 1 \tag{81}$$

$$= 1 + \left(\frac{L_1}{Z}\frac{\kappa_2}{\kappa_1}\right)^{\Psi} \tag{82}$$

$$= L_c/Z_c \tag{83}$$

Also

$$\left[\frac{L}{\theta_1 L_1}\right]^{\rho} = 1 + \left[\frac{\theta_2 Z}{\theta_1 L_1}\right]^{\rho} \tag{84}$$

$$= 1 + \left[\frac{\kappa_1}{\kappa_2}\frac{Z}{L_1}\right]^{\Psi} \tag{85}$$

A.4.2 Income ratios and shares: optimal skill bias

Ratio of skilled to unskilled labor income: Apply (78) to (66):

$$S(W) = \left(\frac{\mu_2 L_2}{Z}\right)^{\phi}\left(\frac{Z}{L_1}\frac{\kappa_1}{\kappa_2}\right)^{\Psi} \tag{86}$$

Equipment income share: Applying (82) to (69) yields

$$IS_e = \frac{q_e e}{y} = (1-\alpha)\left[1 + \left(\frac{L_1}{Z}\frac{\kappa_2}{\kappa_1}\right)^{\Psi}\right]^{-1}\left[\frac{\mu_e e}{Z}\right]^{\phi} \tag{87}$$

A.4.3 Reduced form labor aggregator

The technology frontier together with (74) imply

$$B^{\omega} = (\kappa_1\theta_1)^{\omega}\left[1 + \left(\frac{\kappa_2\theta_2}{\kappa_1\theta_1}\right)^{\omega}\right] \tag{88}$$

$$= [\kappa_1\theta_1]^{\omega}\left[1 + \left[\frac{\kappa_1}{\kappa_2}\frac{Z}{L_1}\right]^{\Psi}\right] \tag{89}$$

Labor aggregator with (78) gives

$$L = \theta_1 L_1 \left[1 + \left[\frac{\kappa_1}{\kappa_2} \frac{Z}{L_1} \right]^{\Psi} \right]^{1/\rho} \tag{90}$$

$$= L_1 \left[1 + \left[\frac{\kappa_1}{\kappa_2} \frac{Z}{L_1} \right]^{\Psi} \right]^{1/\rho} \times \frac{B}{\kappa_1} \left[1 + \left[\frac{\kappa_1}{\kappa_2} \frac{Z}{L_1} \right]^{\Psi} \right]^{-1/\omega} \tag{91}$$

$$= L_1 \left[1 + \left[\frac{\kappa_1}{\kappa_2} \frac{Z}{L_1} \right]^{\Psi} \right]^{1/\Psi} \frac{B}{\kappa_1} \tag{92}$$

which implies the labor aggregator (21). That last equality uses $1/\rho - 1/\omega = 1/\Psi$.

A.4.4 Calibration Algorithm

The model now consists of the following 10 equations:

1. 2 per capita output levels given by the aggregate production function (20) with labor aggregator (21);
2. 2 skill premiums or, equivalently, labor income ratios given by (86)
3. 2 structures shares given by (54);
4. 2 equipment shares given by which is given by (87);
5. 2 wage gains at migration given by (17) imply $R(h_j)$.

Algorithm:

1. Set α_c to match structures income shares.
2. Calibrate $R(h_j)$ to match wage gains.
3. Calibrate parameters of Z aggregator up to a scale factor (see Claim 4 and 5).
4. Now we can solve for Ψ using the rich/poor ratio of (86) and (71):

$$S(W) = \left[\frac{Z}{L_1} \right]^{\Psi} \left[\frac{\mu_2 L_2}{Z} \right]^{\phi} \tag{93}$$

$$= \left[\frac{Z}{L_1} \right]^{\Psi} \left[1 + 1/IR_{se} \right]^{-1} \tag{94}$$

5. Also from (86) calculate (Z/L_1) and Z (given that L_1 is known by normalization).

6. Recover levels of e_c and $L_{2,c}$ from the Z parameters up to scale factor and the actual Z.
7. Compute $\tilde{L}$ from (**??**).
8. Find z_c to match y_c.

B Appendix: No Capital-skill Complementarity

B.1 Derivations: All models

B.1.1 Wages

Wage per efficiency unit of labor from the firm's FOC:

$$p_{j,c} = \frac{\partial y_c}{\partial L_{j,c}} = (1-\alpha)\, z_c^{1-\alpha} k_c^{\alpha} L_c^{-\alpha} \frac{\partial L_c}{\partial L_{j,c}} \tag{95}$$

$$\frac{\partial L_c}{\partial L_{j,c}} = \theta_{j,c}^{\rho} L_c^{1-\rho} L_{j,c}^{\rho-1} \tag{96}$$

This holds because $Q^{1/\rho-1} = L^{1-\rho}$. This implies

$$p_{j,c} = (1-\alpha)\, z_c^{1-\alpha} k_c^{\alpha} L_c^{1-\rho-\alpha} \theta_{j,c}^{\rho} L_{j,c}^{\rho-1} \tag{97}$$

This implies relative skill prices per unit of human capital of

$$S(p_c) = S(\theta_c)^{\rho}\, S(L_c)^{\rho-1} \tag{98}$$

The ratio of observed wages per hour is given by

$$S(w_c) = S(p_c h_c) = S(\theta L)^{\rho}\, S(N)^{-1} \tag{99}$$

Wage gains at migration are given by the ratios of rich to poor country skill prices. From (101), this may be written as

$$wg_j = R(y)\, R(\theta_j L_j / L)^{\rho}\, R(L_j)^{-1} \tag{100}$$

B.1.2 Income Shares

A useful implication for wage bills:

$$W_{j,c} = p_{j,c} L_{j,c} = (1-\alpha)\, y_c \left(\frac{\theta_{j,c} L_{j,c}}{L_c}\right)^{\rho} \tag{101}$$

Since $\rho > 0$, the share of labor type j in labor income increases in the ratio of type j effective labor to total effective labor. Wage bill ratio:

$$S(W) = S\,(pL) = S\,(\theta L)^{\rho} \tag{102}$$

B.1.3 Bounds on contributions of human capital

Contribution of h is defined as

$$\frac{G\,(r,p,p)}{G\,(p,p,p)} = R\,(h_1)\, \frac{\left[(\theta_{1,p} N_{1,p})^{\rho} + \left(\theta_{2,p} N_{2,p} \frac{h_{2,r}}{h_{1,r}}\right)^{\rho}\right]^{1/\rho}}{\left[(\theta_{1,p} N_{1,p})^{\rho} + \left(\theta_{2,p} N_{2,p} \frac{h_{2,p}}{h_{1,p}}\right)^{\rho}\right]^{1/\rho}} > R\,(h_1) \tag{103}$$

where the inequality follows from the assumption that skilled labor has relative more human capital in the rich country: $\frac{h_{2,r}}{h_{1,r}} > \frac{h_{2,p}}{h_{1,p}}$. This yields the lower bound.

The upper bound is derived analogously:

$$\frac{G\,(r,p,p)}{G\,(p,p,p)} = R\,(h_2)\, \frac{\left[\left(\theta_{1,p} N_{1,p} \frac{h_{1,r}}{h_{2,r}}\right)^{\rho} + (\theta_{2,p} N_{2,p})^{\rho}\right]^{1/\rho}}{\left[\left(\theta_{1,p} N_{1,p} \frac{h_{1,p}}{h_{2,p}}\right)^{\rho} + (\theta_{2,p} N_{2,p})^{\rho}\right]^{1/\rho}} < R\,(h_2) \tag{104}$$

The joint contribution of h and N is defined as $G\,(r,r,p)\,/G\,(p,p,p)$. The lower bound is given by

$$\frac{G\,(r,r,p)}{G\,(p,p,p)} = R\,(h_1 N_1)\, \frac{\left[(\theta_{1,p})^{\rho} + \left(\theta_{2,p} \frac{h_{2,r} N_{2,r}}{h_{1,r} N_{1,r}}\right)^{\rho}\right]^{1/\rho}}{\left[(\theta_{1,p})^{\rho} + \left(\theta_{2,p} \frac{h_{2,p} N_{2,p}}{h_{1,p} N_{1,p}}\right)^{\rho}\right]^{1/\rho}} > R\,(h_1 N_1) \tag{105}$$

where the inequality follows from the assumption that skilled labor is relatively abundant in the rich country: $S\,(L_r) > S\,(L_p)$. The upper bound is derived analogously.

Note that these bounds only depend on the labor aggregator having constant returns to scale, not on the CES functional form.

B.1.4 Level Accounting Implications

Proof. [Claim 10]

Define the direct effect of human capital in accordance with (39) as $DE = G(r, r, p) / G(p, p, p)$ or

$$DE = \frac{\left[\sum_j (\theta_{j,p} L_{j,r})^\rho\right]^{1/\rho}}{\left[\sum_j (\theta_{j,p} L_{j,p})^\rho\right]^{1/\rho}} \tag{106}$$

$$= \frac{\theta_{1,p} L_{1,p} \left[R(L_1)^\rho + \left(\frac{\theta_{2,p}}{\theta_{1,p}} \frac{L_{2,r}}{L_{1,p}}\right)^\rho\right]^{1/\rho}}{\theta_{1,p} L_{1,p} \left[1 + S(W)_p\right]^{1/\rho}} \tag{107}$$

$$= \frac{\left[R(L_1)^\rho + \left(\frac{\theta_{2,p}}{\theta_{1,p}} \frac{L_{2,p}}{L_{1,p}} R(L_2)\right)^\rho\right]^{1/\rho}}{\theta_{1,p} L_{1,p} \left[1 + S(W)_p\right]^{1/\rho}} \tag{108}$$

$$= \frac{\left[R(L_1)^\rho + S(W)_p (R(L_2))^\rho\right]^{1/\rho}}{\left[1 + S(W)_p\right]^{1/\rho}} \tag{109}$$

This uses (102) to replace $S(\theta L)^\rho$ with $S(W)$. For fixed ρ, wage gains determine $R(L_j)$ and therefore DE is entirely pinned down by data on wage gains and income shares. □

B.1.5 All models imply the same indirect effect

Proof. [Claim 11]

Again, following (39), write the indirect effect as $IE = G(r, r, r) / G(r, r, p)$ or

$$IE = \frac{\left[\sum (\theta_{j,r} L_{j,r})^\rho\right]^{1/\rho}}{\left[\sum (\theta_{j,p} L_{j,r})^\rho\right]^{1/\rho}} \tag{110}$$

$$= \frac{\theta_{1,r} L_{1,r} \left[1 + S(W)_r\right]^{1/\rho}}{\theta_{1,r} L_{1,r} \left[\left(\frac{\theta_{1,p}}{\theta_{1,r}}\right)^\rho + \left(\frac{\theta_{2,p}}{\theta_{1,r}} \frac{L_{2,r}}{L_{1,r}}\right)^\rho\right]^{1/\rho}} \tag{111}$$

Write the second term in the denominator as

$$\frac{\theta_{2,p} L_{2,r}}{\theta_{1,p} L_{1,r}} = \frac{\theta_{2,p} L_{2,p}}{\theta_{1,p} L_{1,p}} \frac{L_{2,r}}{L_{2,p}} \frac{L_{1,p}}{L_{1,r}} \frac{\theta_{1,p}}{\theta_{1,r}} \tag{112}$$

Therefore

$$IE = \frac{\left[1 + S(W)_r\right]^{1/\rho}}{\frac{\theta_{1,p}}{\theta_{1,r}} \left[1 + S(W)_p \left(\frac{R(L_2)}{R(L_1)}\right)^\rho\right]^{1/\rho}} \tag{113}$$

Normalize $R(\theta_1)$ somehow to distinguish between neutral productivity z_c and skill bias. Finally, observed that wage gains fix $R(L_j)$. □

B.2 Calibration: Exogenous skill bias

This normalizes $\theta_{j,r} = 1$.

1. Set α to capital share.
2. Obtain $R(L)$ from wage gains using (17).
3. For the rich country, $S(\theta) = 1$, so we have $L_{2,r}$ from (102) and therefore all $L_{j,c}$.
4. From the same equation, recover poor's $S(\theta) = IR_{su}^{1/\rho}/S(L)$. This gives all $\theta_{j,c}$ up to a normalization for $\theta_{1,p}$.
5. Scale $\theta_{1,p}$ to match a technology frontier (see below).
6. Find z_c to match y_c.

B.2.1 Scale $\theta_{j,c}$ to match a frontier.

Purpose: Give scale to those parameters. Such that levels accounting implications are the same as for model with frontier.

1. $S(\kappa^\omega) = S(L^\rho)/S(\theta^{\omega-\rho})$. We normalize $S(\theta) = 1$ for the rich country.
2. Solve
$$(\omega-\rho)\ln(S(\theta)) = \ln\left(S\left(\kappa^{-\omega}\right)\right) + \rho\ln(S(L)) \tag{114}$$
for the poor country for ω.
3. Solve for $S(\kappa)$.
4. Find κ_j to match the level of the technology frontier where $\theta_{j,r} = 1$ by normalization.

B.3 Derivations: Fixed or endogenous frontier

B.3.1 Skill bias ratio

Derivation of (41):

$$\frac{\partial y_c}{\partial \theta_{j,c}} = \lambda\omega\kappa_j^\omega\theta_{j,c}^{\omega-1} \tag{115}$$

where λ is the Lagrange multiplier on the technology frontier constraint and

$$\frac{\partial y_c}{\partial \theta_{j,c}} = k_c^{\alpha} z_c^{1-\alpha} L_c^{1-\rho-\alpha} L_{j,c}^{\rho} \theta_{j,c}^{\rho-1} \tag{116}$$

This implies (41).

B.3.2 Skill bias levels

Claim 15. Optimal skill bias levels are given by

$$\theta_{1,c}^{\omega} = \frac{B_c}{\kappa_1^{\omega} \Lambda_c} \tag{117}$$

with

$$\Lambda_c = \sum_j \left(\frac{\kappa_1}{\kappa_j} \frac{L_j}{L_1} \right)^{\Psi} \tag{118}$$

This treats B_c as given, but it also holds in models where B_c is chosen.

Proof. Start from the technology frontier

$$B_c^{\omega} = \sum_j (\kappa_j \theta_j)^{\omega} \tag{119}$$

$$= (\kappa_1 \theta_1)^{\omega} \sum_j \left(\frac{\kappa_j}{\kappa_1} \frac{\theta_j}{\theta_1} \right)^{\omega} \tag{120}$$

Sub in the condition for optimal relative skill bias (41):

$$B_c^{\omega} = (\kappa_1 \theta_1)^{\omega} \sum_j \left(\frac{\kappa_j}{\kappa_1} \right)^{\omega} \left[\left(\frac{L_j}{L_1} \right)^{\rho} \left(\frac{\kappa_1}{\kappa_j} \right)^{\omega} \right]^{\frac{\omega}{\omega-\rho}} \tag{121}$$

and note that $\omega - \frac{\omega^2}{\omega-\rho} = \frac{-\rho\omega}{\omega-\rho} = -\Psi$. This implies

$$B_c^{\omega} = (\theta_1 \kappa_1)^{\omega} \sum_j \left(\frac{\kappa_1}{\kappa_j} \frac{L_j}{L_1} \right)^{\Psi} = (\theta_1 \kappa_1)^{\omega} \Lambda_c \tag{122}$$

□

B.3.3 Skill premium with optimal relative skill bias

$$S(w) = S(ph) = S(Nh)^{\rho-1} S(h) S(\theta)^{\rho} \tag{123}$$

Apply optimal skill bias ratio (41):

$$S\left(\theta^{\rho}\right)=S\left(\kappa^{-\Psi}L^{\frac{\rho^{2}}{\omega-\rho}}\right) \tag{124}$$

to obtain

$$S\left(w\right)=S\left(N\right)^{-1}S\left(Nh\right)^{\rho}S\left(L\right)^{\frac{\rho^{2}}{\omega-\rho}}S\left(\kappa\right)^{-\Psi} \tag{125}$$

Then simplify exponents using $\rho+\frac{\rho^{2}}{\omega-\rho}=\frac{\rho\omega}{\omega-\rho}=\Psi$ and $L=Nh$. This implies

$$S\left(w\right)=S\left(N^{-1}\kappa^{-\Psi}L^{\Psi}\right) \tag{126}$$

or

$$S\left(p\right)=\left(S\left(L\right)\right)^{\Psi-1}S\left(\kappa\right)^{-\Psi} \tag{127}$$

B.3.4 Reduced form production function

Here, it does not matter where B_c comes from (endogenous or not). Start from the definition of the labor aggregator (6)

$$L_{c}=\theta_{1,c}L_{1,c}\left(\sum_{j}\left[\frac{\theta_{j,c}}{\theta_{1,c}}\frac{L_{j,c}}{L_{1,c}}\right]^{\rho}\right)^{1/\rho} \tag{128}$$

Sub in the condition for the optimal choice of relative skill bias (124) to obtain

$$L_{c}=\theta_{1,c}L_{1,c}\left(\sum_{j}\left[\frac{L_{j,c}}{L_{1,c}}\right]^{\frac{\rho^{2}}{\omega-\rho}}\left[\frac{\kappa_{j}}{\kappa_{1}}\right]^{-\Psi}\left[\frac{L_{j,c}}{L_{1,c}}\right]^{\rho}\right)^{1/\rho} \tag{129}$$

and simplify exponents. Notably, the exponent on labor inputs is given by

$$\frac{\rho^{2}}{\omega-\rho}+\rho=\frac{\omega\rho}{\omega-\rho}=\Psi \tag{130}$$

Then the summation term becomes Λ_c and we have

$$L_{c}=\theta_{1,c}L_{1,c}\Lambda_{c}^{1/\rho} \tag{131}$$

Then using (117), we have

$$L_{c}=B_{c}\kappa_{1}^{-1}\Lambda_{c}^{-1/\omega}L_{1,c}\Lambda_{c}^{1/\rho} \tag{132}$$

Note that

$$1/\rho - 1/\omega = \frac{\omega - \rho}{\omega\rho} = 1/\Psi \tag{133}$$

so that

$$L_c = B_c (1/\kappa_1) L_{1,c} \Lambda_c^{1/\Psi} \tag{134}$$

$$= B_c (1/\kappa_1) L_{1,c} \left[\sum_j \left(\kappa_j^{-1} L_j \right)^{\Psi} \right]^{1/\Psi} \kappa_1 / L_1 \tag{135}$$

So we have (20)

B.4 Calibration: Fixed or endogenous frontier

Claim 16. We can solve for $h_{j,c}$ and Ψ using only $R(h_j)$ from wage gains and data on labor income ratios $S(W)$. This is true regardless how B_c is determined.

Proof. Write (126) as $S(\theta) = S(W)^{1/\rho} / S(L)$. From (41), we have $S(\theta) = S(L)^{\frac{\rho}{\omega - \rho}}$. Combine the two expressions for $S(\theta)$ to obtain

$$S(L)^{\frac{\rho}{\omega-\rho}+1} = S(W)^{1/\rho} \tag{136}$$

Note that the exponent on $S(L)$ equals $\omega / (\omega - \rho)$ to find

$$S(L) = S(W)^{1/\Psi} \tag{137}$$

Given $L_{j,c} = h_{j,c} N_{j,c}$, we can solve (137) together with $R(h_j)$ and $h_{1,r} = 1$ for the five unknowns $h_{j,c}$ and Ψ. □

B.5 Calibration Algorithm: Fixed Frontier

1. Use Claim 16 to solve for Ψ and $h_{j,c}$.
   (a) Based on (137), solve

   $$R(S(L)) = S(R(L)) \tag{138}$$

   $$= R(S(W))^{1/\Psi} \tag{139}$$

   for Ψ.
   (b) Solve (137) for $S(L)$ and then for $L_{j,c}$ and $h_{j,c}$.
2. With κ_j normalized to 1, solve for $\tilde{L}_c$.

3. Set α_c to match capital shares.

4. Find z_c to match y_c.

5. Set frontier curvature $\omega = \Psi\rho / (\Psi - \rho)$.

6. For fixed ρ, find $\theta_{j,c}$ from $L_{j,c}$ using (117).

Perform levels accounting in the standard way.

B.6 Derivations: Investment in the Frontier

B.6.1 Skill bias choice

FOC:

$$z_c^{1-\alpha} k_c^{\alpha} (1-\alpha) L_c^{-\alpha} Q^{1/\rho - 1} L_{j,c}^{\rho} \theta_{j,c}^{\rho-1} = b \kappa_j^{\omega} \omega \theta_{j,c}^{\omega - 1} \tag{140}$$

where $L_c = Q^{1/\rho}$. Therefore,

$$L_c^{-\alpha} Q^{1/\rho - 1} = L_c^{-\alpha+1-\rho} \tag{141}$$

This yields

$$\theta_j^{\omega-\rho} = (1-\alpha) z_c^{1-\alpha} k_c^{\alpha} \omega^{-1} \kappa_j^{-\omega} L_{j,c}^{\rho} L_c^{1-\alpha-\rho} \tag{142}$$

Alternative derivation, using what we have for the general case:

FOC for B_c: $\partial y / \partial B = 0 \implies \lambda = b$. Combine with first order conditions for $\theta_{j,c}$ to obtain (142) (when the marginal cost of B_c is b).

B.6.2 Optimal level of B_c:

From (117):

$$B_c = \kappa_1^{\omega} \Lambda_c \theta_{1,c}^{\omega} \tag{143}$$

It should be possible to derive something reasonable for B_c, but (142) implies

$$\theta_{j,c}^{\omega} = \left[\frac{(1-\alpha) k^{\alpha} z^{1-\alpha}}{\omega \kappa_j^{\omega} b} \right]^{\frac{\omega}{\omega-\rho}} L_j^{\Psi} L_c^{(1-\alpha-\rho)\frac{\omega}{\omega-\rho}} \tag{144}$$

Substituting that into (143) is too complicated to make progress. Better to derived reduced form production function without solving for B_c.

B.6.3 Reduced form labor aggregator

Start from (131): $L_c = \theta_{1,c} L_{1,c} \Lambda_c^{1/\rho}$

Use the FOC for $\theta_{1,c}$ (142), written as

$$\theta_{1,c}^{\omega-\rho} = X L_{1,c}^{\rho} L_c^{1-\alpha-\rho} \tag{145}$$

where

$$X = \frac{(1-\alpha)\, z^{1-\alpha} k^{\alpha}}{b\omega\kappa_j^{\omega}} \tag{146}$$

Or

$$\theta_{1,c} L_{1,c} = X^{\frac{1}{\omega-\rho}} L_{1,c}^{\frac{\omega}{\omega-\rho}} L_c^{\frac{1-\alpha-\rho}{\omega-\rho}} \tag{147}$$

The exponent on L_1 is $1 + \frac{\rho}{\omega-\rho} = \frac{\omega}{\omega-\rho}$.

Substituting back into (131) we have

$$L_c = \Lambda_c^{1/\rho} X^{1/(\omega-\rho)} L_1^{\frac{\omega}{\omega-\rho}} L_c^{\frac{1-\alpha-\rho}{\omega-\rho}} \tag{148}$$

The exponent on labor input becomes

$$1 - \frac{1-\alpha-\rho}{\omega-\rho} = \frac{\omega+\alpha-1}{\omega-\rho} \tag{149}$$

Hence

$$L_c^{\frac{\omega+\alpha-1}{\omega-\rho}} = L_{1,c}^{\frac{\omega}{\omega-\rho}} X^{1/(\omega-\rho)} \Lambda_c^{1/\rho} \tag{150}$$

Next,

$$\Lambda_c = \sum_j (L_j/\kappa_j)^{\Psi} \times (\kappa_1/L_1)^{\Psi} \tag{151}$$

$$= (\kappa_1/L_1)^{\Psi}\, \tilde{L}_c^{\Psi} \tag{152}$$

Pull κ_1 and $L_{1,c}$ out of the $\Lambda_c^{1/\rho}$ term:

$$\Lambda_c^{1/\rho} = (\kappa_1/L_{1,c})^{\Psi/\rho}\, \tilde{L}_c^{\Psi/\rho}$$

That gives an exponent on $L_{1,c}$ of $\frac{\omega}{\omega-\rho} - \frac{\Psi}{\rho} = 0$. So we have

$$L_c^{\frac{\omega+\alpha-1}{\omega-\rho}} = \kappa_1^{\frac{\Psi}{\rho}} X^{1/(\omega-\rho)} \tilde{L}_c^{\Psi/\rho} \tag{153}$$

$$L_c = (\kappa_1^{\omega} X)^{\frac{1}{\omega+\alpha-1}}\, \tilde{L}_c^{\frac{\omega}{\omega+\alpha-1}} \tag{154}$$

The exponent on κ_1 is $\frac{\omega}{\omega-\rho}\frac{\omega-\rho}{\omega+\alpha-1}$. The exponent on $\tilde{L}_c$ is the same.

B.6.4 Reduced form production function

$$y_c = k_c^\alpha \left(z_c L_c\right)^{1-\alpha} \tag{155}$$

$$= k_c^\alpha z_c^{1-\alpha} \left(\kappa_1^\omega X\right)^{\frac{1-\alpha}{\omega+\alpha-1}} \left(\tilde{L}_c\right)^{\frac{\omega(1-\alpha)}{\omega+\alpha-1}} \tag{156}$$

$$= \tilde{A} \left(z_c^{1-\alpha} k_c^\alpha\right)^{1+\frac{1-\alpha}{\omega+\alpha-1}} \left(\tilde{L}_c\right)^{\frac{\omega(1-\alpha)}{\omega+\alpha-1}} \tag{157}$$

where

$$\tilde{A} = \left(\frac{1-\alpha}{\omega b}\right)^{\frac{1-\alpha}{\omega+\alpha-1}} \tag{158}$$

collects all constant terms. Then

$$y_c = \left(k_c^\alpha \left(A z_c \tilde{L}_c\right)^{1-\alpha}\right)^{\frac{\omega}{\omega+\alpha-1}} \tag{159}$$

where $A = \left(\frac{1-\alpha}{\omega b}\right)^{1/\omega}$. This is true because the exponent on $z_c^{1-\alpha} k_c^\alpha$ is

$$1 + \frac{1-\alpha}{\omega+\alpha-1} = \frac{\omega}{\omega+\alpha-1} \tag{160}$$

B.6.5 Derivation of levels accounting

Start from 50 and divide by k:

$$y_c^\eta = \left[\left(k_c/y_c\right)^\alpha \left(A z_c \tilde{L}_c\right)^{1-\alpha}\right]^{\frac{\omega}{\omega+\alpha-1}} \tag{161}$$

where

$$\eta = 1 - \frac{\alpha\omega}{\omega+\alpha-1} = \frac{\omega+\alpha-1-\alpha\omega}{\omega+\alpha-1} \tag{162}$$

$$= (\alpha-1)\frac{1-\omega}{\omega+\alpha-1} \tag{163}$$

Then

$$y_c = \left[\left(k_c/y_c\right)^\alpha \left(A z_c \tilde{L}_c\right)^{1-\alpha}\right]^\phi \tag{164}$$

with $\phi = \frac{\omega}{\omega+\alpha-1} \times \frac{\omega+\alpha-1}{(1-\alpha)(\omega-1)}$. Simplify exponents to arrive at (51).