Skilled Labor Productivity and Cross-country Income Differences*

Lutz Hendricks† and Todd Schoellman‡

October 30, 2019

Abstract

This paper extends development accounting to an environment that features imperfectly substitutable skills and cross-country variation in skilled labor productivity. We find that human capital accounts for between one-half and three-fourths of cross-country income gaps. This finding remains robust when we consider alternative shifters of skilled labor productivity, 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.

*This paper has benefited from comments from seminar participants at the IIES, the 2018 Midwest Macroeconomics Meeting, and the 2019 Human Capital Conference at the Federal Reserve Bank of St. Louis. 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

1 Introduction

This paper studies the contribution of human capital to cross-country income differences in an environment with imperfect substitution between skill types. We allow for multiple margins of relative supply and demand to affect relative labor productivity and the skilled wage premium, including the quantity and quality of labor supply, skill-biased technology, and capital-skill complementarity.

Motivation Our work builds on a sizable development accounting literature. 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 sources of cross-country income differences.[1]

Most of the early development accounting literature focuses on the case where workers with different levels of skills are perfect substitutes (Hall and Jones, 1999; Bils and Klenow, 2000; Caselli, 2005). The typical finding is that the gaps in quality-adjusted units of the single type of labor are not large enough to account for much of cross-country income differences.

The perfect substitutes assumption is at odds with an extensive literature that documents large, systematic movements in relative wages, for example over time in the United States (Katz and Murphy, 1992), and attributes them to an underlying race between relative labor supply and relative labor demand (Goldin and Katz, 2008).[2]

This literature has motivated recent work that extends development accounting to environments where skills are imperfect substitutes. However, there is substantial disagreement about the implications. While Jones (2014), Hendricks and Schoellman (2018) and Jones (2019) find a much larger role for human capital in development accounting than earlier work, Caselli and Coleman (2006), Caselli (2016) and Caselli and Ciccone (2019) do not.

To understand the source of the disagreement it is useful to revisit the empirical argument. With imperfect substitution, the scarcity of skilled workers should drive up skill premiums in low income countries. However, empirical evidence indicates that skill premiums are roughly similar in rich and poor countries (Banerjee and Duflo, 2005), suggesting that skilled labor is relatively less productive in low income countries.

The disagreement in the literature relates to the interpretation of these skilled labor productivity differences. Jones (2014) and Hendricks and Schoellman (2018) attribute them

[1] See the surveys by Caselli (2005), Hsieh and Klenow (2010), and Caselli (2016). This work complements a parallel literature that develops quantitative theories of human capital formation, including Erosa et al. (2010), (Córdoba and Ripoll, 2013), and Cubas et al. (2016).

[2] Relative labor supply can be shifted by the number or quality of skilled versus unskilled workers. Relative labor demand can be shifted by skill-biased technical change (see Acemoglu, 1998, 2002; Gancia and Zilibotti, 2009, and Jerzmanowski and Tamura, 2019) or the accumulation of capital that complements skilled workers (see Krusell et al., 2000 and Parro, 2013).

to variation in the human capital of skilled workers. In this case, poor countries are scarce in the share and average human capital of skilled workers, magnifying cross-country differences in human capital stocks. Caselli and Coleman (2006) and Caselli (2016) attribute them to the skill bias of technology. Empirically, the two interpretations are difficult to distinguish because both have similar effects on observable wages.

The purpose of this paper is to extend development accounting to an environment that allows for both sources of cross-country variation in the relative productivity of skilled labor.

Approach We develop a model where the relative productivity of skilled labor is affected not only by human capital but also by factor-augmenting technologies, which can be exogenously given (as in Katz and Murphy, 1992), chosen endogenously from a frontier (as in Caselli and Coleman 2006), or the result of directed technical change by firms (as in Acemoglu 2007). In addition, we allow for capital-skill complementarity as in Krusell et al. (2000), so that the scarcity of equipment in low income countries affects skilled labor productivity.

We calibrate the model to standard data moments plus the new evidence on the wage gains at migration from Hendricks and Schoellman (2018) and perform development accounting. We also decompose cross-country variation in relative skilled labor productivity into the contributions of human capital and factor-augmenting technologies.

Baseline Results Our baseline model directly addresses the ongoing debate in the literature by combining elements of Jones (2014) and Caselli and Coleman (2006). The relative productivity of skilled labor varies across countries due to differences in human capital and due to the skill bias of technology, which is chosen by firms from a technology frontier. Other shifters are considered as variations or extensions of this baseline model.

We show analytically that the baseline model is equivalent to one where the skill bias of technology is fixed. Technology choice only increases the effective elasticity of substitution between skilled and unskilled labor. We call this higher elasticity the *long-run elasticity* of substitution; it mixes the traditional (short-run) elasticity of substitution with the curvature of the skill bias technology frontier. The distinction between the short-run and long-run elasticity of substitution implies that we should not restrict ourselves to elasticities in the range typically estimated in the literature, which captures changes in relative wages in response to changes in relative labor supply that are small, compared to the cross-country variation.

When calibrated, the model implies that human capital accounts for 58% to 63% of cross-country output gaps, depending on how we define unskilled and skilled labor. This result is much larger than is standard in the literature and is in line with Jones (2014) and

Hendricks and Schoellman (2018). The range of plausible results is much narrower. We are able to obtain sharp intuition for our findings by deriving a closed form solution for the share of output gaps that is due to human capital in terms of observable data moments. The solution consists of two terms that we label base and amplification.

The base term is the share that arises in a perfect substitutes model. It depends on the wage gains of immigrants, which isolate the importance of country-specific factors for workers supplying the same human capital in two countries. Since migrant wage gains are much smaller than cross-country wage gaps, the base term is large, accounting for at least 45% of output gaps. Importantly, the base term does not depend on whether the skill bias of technology is endogenous or fixed, or even whether or not skills are imperfect substitutes. Since the amplification term is nonnegative, it follows that all of the models that we study imply that human capital accounts for at least 45% of output gaps.[3]

The amplification term captures the additional contribution of human capital that arises due to imperfect substitution. Its magnitude depends critically on the elasticity of substitution between skill types. Previous work has focused on empirical estimates which imply a short-run elasticity of substitution near two, while considering a wider range as well. This produces a large central estimate but a wide range for the plausible size of the amplification term and hence the human capital share in development accounting. We focus instead the long-run elasticity of substitution. We calibrate it to fit the wage gains of skilled versus unskilled immigrants. Intuitively, unskilled immigrants gain modestly more at migration, which leads us to infer that the long-run elasticity is in the range 4-8. This leads us to find that the amplification term is smaller and more tightly pinned down.

Sources of Relative Productivity Gaps Our second objective is to decompose cross-country differences in the relative productivity of skilled labor into the contributions of human capital versus the skill bias of technology. For conventional values of the short-run elasticity of substitution between skilled and unskilled labor, our baseline model implies that human capital accounts for at most one-third of the relative productivity differences.

The intuition relies on the observation that migrant wage gains are broadly similar for skilled and unskilled workers. Our model infers that relative skill prices are similar across countries. Together with the observation that returns to schooling are broadly similar across countries, it follows that differences in relative human capital (skilled versus unskilled) are modest. This finding is consistent with Caselli and Coleman (2006) and with Rossi (2019), who also uses the economic performance of immigrants to discipline a similar decomposition.[4]

[3] While we do not have a closed form solution for the model with capital-skill complementarity, the quantitative results are similar for that case.

[4] Malmberg (2018) estimates cross-country differences in the relative productivity of skilled labor using data on the skill intensities of manufacturing imports and exports.

Robustness We show that the baseline results remain qualitatively unchanged when we treat the skill bias of technology as either exogenous, or as resulting from investment in skill-biased technical change as in Acemoglu (2007), or when we allow for capital-skill complementarity as an alternative source of variation in skilled labor productivity. In all of these cases, human capital accounts for between one-half and three-quarters of cross-country output gaps, even though it contributes only modestly to cross-country variation in relative skilled labor productivity. Our results also remain robust when we allow for larger unskilled migrant wage gains. We conclude that allowing for sources of skilled labor productivity variation other than human capital does not lead to major changes in the development accounting results.

2 Baseline Model

We perform development accounting in an environment that allows for relative wages to be affected by labor supply factors (relative employment, relative human capital) and labor demand factors (relative skill bias, relative complementarity with other inputs).

2.1 Model Specification

The baseline model combines the production function and human capital structure of Jones (2014) with the technology frontier of Caselli and Coleman (2006). There are two countries, indexed by $c \in \{p, r\}$ (poor and rich). Output per worker y_c is produced from physical capital k_c and labor L_c according to the production function

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

where the aggregate labor input is a CES aggregator of unskilled $(j = 1)$ and skilled $(j = 2)$ labor

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

with $J = 2$. The elasticity of substitution between skilled and unskilled labor is $\sigma = 1/\left(1-\rho\right) > 1$, so that $0 < \rho < 1$. Labor inputs are the product of human capital and employment: $L_{j,c} = h_{j,c} N_{j,c}$. Their supplies are taken as exogenous. The skill weights $\theta_{j,c}$ are constrained by a technology frontier, similar to Caselli and Coleman (2006) or Acemoglu (2007), given by

$$\left[\sum_{j=1}^{J} \left(\kappa_j \theta_{j,c}\right)^{\omega}\right]^{1/\omega} \leq B_c^{1/\omega} \tag{3}$$

with parameters $\omega > 0$, $B_c > 0$, and $\kappa_j > 0$. As in Caselli and Coleman (2006), we assume that

$$\omega - \rho - \omega\rho > 0 \tag{4}$$

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

Throughout, we use the following notation. $R(x_j) = x_{j,r}/x_{j,p}$ denotes the rich-to-poor country ratio of $x_{j,c}$. $S(x_c) = x_{2,c}/x_{1,c}$ denotes the skilled-to-unskilled ratio of $x_{j,c}$. Finally, $RS(x) = R(S(x)) = S(R(x))$. For example, $RS(N)$ is the relative abundance of skilled workers in the rich compared with the poor country.

Equilibrium In line with the development 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 price of capital q_c and therefore k_c/y_c. The rental prices of labor inputs, $p_{j,c}$, are determined by labor market clearing. The representative firm solves

$$\max_{k_c, L_{j,c}, \theta_{j,c}} y_c - q_c k_c - \sum_{j=1}^{J} p_{j,c} L_{j,c} \tag{5}$$

subject to (1), (2), and (3), taking factor prices as given.

Note that observable wage rates per hour are given by $w_{j,c} = h_{j,c} p_{j,c}$. Hence, the total earnings of skill j workers are given by $W_{j,c} = p_{j,c} L_{j,c} = w_{j,c} N_{j,c}$. The values of $p_{j,c}$ are not directly observable in the data.

Our setup nests the model of Caselli and Coleman (2006) as a special case when $h_{j,c} = 1$ and $B_c = 1$. It nests the model of Jones (2014) when the choice of skill bias is removed and $\theta_{j,c} = 1$.[5]

2.2 Development Accounting

We discuss how to perform development accounting when the skill bias of technology is endogenous. As is standard in the literature, we start from

$$y_c = z_c \left(k_c/y_c\right)^{\alpha/(1-\alpha)} L_c \tag{6}$$

[5] Appendix E considers an extension where firms may invest in extending the frontier (increasing B_c, as in Acemoglu 2007). We show that the development accounting results remain unchanged if the costs of investing in B_c scale appropriately with output, so that the aggregate production function features constant returns to scale.

and decompose the output gap into the contributions of TFP, physical capital, and (jointly) labor inputs and skill bias according to

$$\underbrace{\ln R\left(y\right)}_{\text{output gap}} = \underbrace{\ln R\left(z\right)}_{\text{TFP}} + \underbrace{\ln R\left(\left(k/y\right)^{\alpha/(1-\alpha)}\right)}_{\text{physical capital}} + \underbrace{\ln R\left(L\right)}_{\text{labor + skill bias}} \tag{7}$$

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

$$1 = \underbrace{\frac{\ln R\left(z\right)}{\ln R\left(y\right)}}_{\text{share}_z} + \underbrace{\frac{\ln R\left(\left(k/y\right)^{\alpha/(1-\alpha)}\right)}{\ln R\left(y\right)}}_{\text{share}_k} + \underbrace{\frac{\ln R\left(L\right)}{\ln R\left(y\right)}}_{\text{share}_L} \tag{8}$$

The literature typically defines the contribution of each input to cross-country output gaps via a counterfactual experiment. For example, the contribution of human capital is defined as the change in steady state output when human capital is increased from the poor country's to the rich country's level. To the extent that other inputs respond endogenously, their effect is counted as part of human capital's contribution. In particular, the counterfactual holds the capital-output ratio constant. This captures the induced changes in the capital stock when the saving rate is unchanged (see Hsieh and Klenow, 2010).

In line with this approach, we count the effects of induced changes in the skill bias of technology as part of the contribution of labor inputs, measured by $share_L$. Alternatively, $share_L$ could be defined as the change in steady state output, holding skill bias fixed. We consider this possibility in Section 4.

2.3 Reduced Form Labor Aggregator

One challenge for development accounting is the identification of the two elasticities that govern the substitution between skilled and unskilled labor (ρ and ω). Our first result shows that $share_L$ can be estimated without separately identifying both parameters.

Proposition 1. *Solving out the firm's optimal skill bias choices yields the reduced form labor aggregator*

$$L_c = B_c \left[\sum_{j=1}^{J} \left(\kappa_j^{-1} L_{j,c}\right)^{\Psi}\right]^{1/\Psi} \tag{9}$$

with an elasticity of substitution governed by

$$\Psi = \frac{\omega\rho}{\omega - \rho} > \rho \tag{10}$$

Proof. Section C.2 □

Proposition 1 establishes that allowing for technology choice is equivalent to increasing the elasticity of substitution while holding the skill bias of technology fixed. The reduced form skill bias parameters $\left(\kappa_j^{-1}\right)$ are common across countries and governed by the technology frontier. Variation in the level of the frontier B_c has the same effect as variation in z_c.

In other words, we are back in the world of Jones (2014) with one crucial difference: the "short-run" elasticity of substitution of the original labor aggregator $1/\left(1-\rho\right)$ no longer matters by itself. It is replaced by a higher, "long-run" elasticity $1/\left(1-\Psi\right)$ that combines the curvatures of the labor aggregator and the technology frontier. The long-run elasticity reflects two equilibrium responses to an increase in skilled labor abundance $S\left(L\right)$. The first (standard) response is that the lower skilled wage premium induces firms to substitute along the isoquant of the original CES production technology (2). The second effect is that firms choose a more skill-biased technology along the frontier (3). Not having to separately identify ρ and ω greatly simplifies the identification.

It follows directly that allowing for technology choice does *not affect* development accounting results. When calibrated to the same data moments, our baseline model implies exactly the same contribution of labor inputs to cross-country output gaps as does the model of Jones (2014), which abstracts from technology choice.

It is useful to place this result into the context of the literature. Previous work has shown that human capital can account for the majority of cross-country output gaps if skilled workers are relatively more productive in rich countries (Jones, 2014). This work assumed that all variation in relative worker productivities is due to human capital, leaving open the possibility that the role of human capital could be much smaller if other sources of relative productivity differences are considered (Caselli and Ciccone, 2019). Proposition 1 implies that this concern is unfounded when the skill bias of technology is endogenous.

2.4 Closed Form Solution

Before we proceed to calibrate the model, we derive a closed form solution for the contribution of labor inputs to cross-country output gaps, $share_L$.

Proposition 2. *The rich-to-poor country ratio of labor inputs is given by*

$$R\left(L\right) = \frac{R\left(y\right)}{wg_1} R\left(1+S(W)\right)^{1/\Psi-1} \tag{11}$$

The share of output gaps due to labor inputs is then given by

$$share_L = \underbrace{1 - \frac{\ln\left(wg_1\right)}{\ln R\left(y\right)}}_{base} + \underbrace{\left(\frac{1}{\Psi}-1\right)\frac{\ln R\left(1+S(W)\right)}{\ln R\left(y\right)}}_{amplification} \tag{12}$$

where $wg_j = R(p_j)$ denotes the wage gain due to migration and

$$\Psi = \ln(RS(W)) / \ln(RS(L)) \tag{13}$$

Proof. Section C.3 □

Note that all of the terms in (12) can be estimated from data.[6] This allows us to obtain precise intuition about how $share_L$ depends on data moments. Moreover, since the same solution applies to the model of Jones (2014) (except that the substitution elasticity is governed by ρ instead of Ψ), we gain insight into how his development accounting results differ from ours.

The solution for $share_L$ consists of two terms, which we label base and amplification. The base term is the contribution of labor inputs to output gaps with perfect substitution of skills. Intuitively, the wage gain at migration captures the importance of changing country-specific factors (capital, TFP) for a worker's wages. If wage changes are as large as GDP per worker gaps, then country-specific factors account for all of income differences. If not, the remainder of GDP per worker gaps is attributable to the gaps in average human capital between countries. For example, if workers' wages do not change at all when migrating, then we would infer that country-specific factors are irrelevant and human capital accounts for all of cross-country income differences.[7].

With imperfect skill substitution, $share_L$ is amplified when the rich country is skill abundant, so that $RS(h) > 1$. This is captured by the second term in (11) which we label amplification.[8] As in Jones (2014), we can sign this term to be positive, meaning that allowing for imperfect substitution expands the role of human capital in development accounting. Its magnitude depends on the elasticity of substitution parameter Ψ and the rich country's relative abundance of skilled labor, captured by the poor-to-rich country ratio of the unskilled labor income share, $R(1 + S(W))$.[9]

3 Quantitative Results

This section presents the development accounting implications of the baseline model, calibrated to match data moments for output gaps, labor income shares and employment by

[6] We show below that $RS(L) = RS(N)\, RS(h)$ can be estimated using data on migrant wage gains.

[7] Formally, with perfect substitution, the wage gap $R(w)$ equals the output gap $R(y)$. Since $w = ph$, we have $R(h) = R(w) / R(p) = R(y) / wg$. Hence, the term $R(y) / wg_1$ in (11) measures the cross-country human capital gap $R(h)$. Therefore, $\ln(R(y) / wg) / \ln R(y) = 1 - \ln wg / \ln R(y)$ is the contribution of labor inputs to output gaps.

[8] As $RS(h) \to 1$, the amplification term vanishes because $RS(L) \to RS(N) = RS(W)$, so that $(1/\Psi - 1) \to 0$.

[9] $1 + S(W) = \sum_j W_j / W_1$ implies $R(1 + S(W)) = 1/R\left(W_1 / \sum_j W_j\right)$.

skill, and migrant wage gains.

3.1 Data

This section summarizes key features of the data, leaving details for Appendix A. We take the rich country to be the U.S. Consistent with Hendricks and Schoellman (2018), the poor country is the median of 63 countries with $y_c/y_r < 1/4$. We consider four different lower bounds on the set of skilled workers: some secondary schooling (SHS), secondary degree (HSG), some college (SC), and college degree (CG). Table 1 shows the data moments for each skill cutoff. Data moments that do not vary across skill cutoffs are shown in Table 2. We highlight two observations that are important for the development accounting results.

1. The ratio of skilled to unskilled employment $S(N)$ varies far more across countries for the SHS skill cutoff than for the CG skill cutoff. The intuition is that rich countries have very few workers that count as unskilled under the SHS cutoff, so that $S(N)$ is large relative to the poor country. By contrast, for the CG cutoff most workers are unskilled even in rich countries, so that $S(N)$ is more similar to the poor country. This will be important for understanding how development accounting results vary across skill cutoffs.

2. Migrant wage gains are between 2 and 3.7 in all cases. They are not dramatically different for skilled versus unskilled workers. This will be important for estimating how much relative skilled human capital $S(h)$ differs across countries.

3.2 Development Accounting

We perform development accounting by applying the data moments shown in Section 3.1 to the closed form solution for $share_L$, equation (12). As shown in Table 3, $share_L$ is close to 60% for all skill cutoffs. These findings align closely with Hendricks and Schoellman (2018). Having a closed form solution for $share_L$ allows us to provide sharp intuition for our results and for why they are very different from Jones (2014).

Table 3 reveals why $share_L$ is approximately constant across skill cutoffs: variation of the base term and of the amplification term roughly balance each other. The base term ranges from 0.45 to 0.56 across skill cutoffs. Recall that the base term is equivalent to the contribution of human capital in a single skill model. It depends on the magnitude of unskilled migrant wage gains relative to the output gap. Since unskilled migrant wage gains are small (2.8 to 3.7) relative to the output gap (10.7), the contribution of human capital is large. Higher skill cutoffs are associated with smaller unskilled migrant wage gains and therefore larger base terms.

Table 1: Data Moments

	Skill Cutoff			
	SHS	HSG	SC	CG
Skilled/unskilled employment, $S(N)$				
rich	26.16	1.13	0.35	0.06
poor	0.95	0.23	0.08	0.02
rich/poor	27.45	4.86	4.45	2.72
Skilled/unskilled wage bill, $S(W)$				
rich	71.11	3.74	1.43	0.30
poor	2.59	0.77	0.32	0.11
rich/poor	27.45	4.86	4.45	2.72
Migrant wage gain, $wg = R(p)$				
unskilled	3.71	3.46	2.98	2.84
skilled	2.29	2.21	2.08	2.04
unskilled/skilled	1.62	1.57	1.43	1.39

Table 2: Data Moments Independent of Skill Cutoff

	y	k/y	Capital share
Rich	1.00	3.18	0.33
Poor	0.09	2.66	0.33
Ratio	10.70	1.19	1.00

Table 3: Closed Form Solution for share$_L$

	Skill Cutoff			
	SHS	HSG	SC	CG
$share_L$	0.60	0.54	0.57	0.54
Base term	0.15	0.34	0.48	0.51
Amplification term	0.45	0.20	0.09	0.03
$1/\Psi - 1$	0.36	0.49	0.34	0.44
$\frac{\ln R(1+S(W))}{\ln R(y)}$	1.27	0.42	0.26	0.07
$share_k$	0.04	0.04	0.04	0.04
$share_z$	0.36	0.42	0.40	0.42

Notes: The table shows the closed form solution for $share_L$ according to equation (12) and its components.

The amplification term depends on the elasticity of substitution and the skilled-to-unskilled earnings ratios. The fact that the reduced form elasticity of substitution is high (for reasons that are discussed in Section 3.5) limits the size of the amplification term. Since differences in relative employment shares $S(N)$ and therefore also in $R(1 + S(W))$ are much smaller for higher skill cutoffs, the amplification term is smaller for higher skill cutoffs. It is the offsetting variation in the base and the amplification term that generates the approximate constancy of share$_L$ across skill cutoffs.

For completeness, Table 3 also shows the fraction of output gaps due to physical capital and TFP. As commonly found in the literature, the contribution of physical capital is small (0.04), leaving more than one third of the output gap unexplained and hence attributed to TFP.

3.3 Comparison to Literature

Our results broadly agree with Jones (2014), who finds that human capital may account for a large fraction of cross-country output gaps. However, there are important differences in the calibration that affect the interpretation and robustness of the results. In particular, Jones assumes that unskilled workers are endowed with the same human capital in rich and poor countries, and he focuses on conventional values for the short-run elasticity of substitution between 1.5 and 2 (while examining a wider range as a robustness check). Table 4 shows the results if we adopt his calibration strategy using our data (which is

Table 4: Jones (2014) Calibration

Short-run Elasticity	Skill Cutoff SHS	HSG	SC	CG
1.25	5.22	1.85	1.19	0.32
1.50	2.69	1.02	0.68	0.18
2.00	1.42	0.60	0.42	0.12
3.00	0.79	0.39	0.29	0.08
4.00	0.58	0.32	0.25	0.07
5.00	0.47	0.29	0.23	0.07

Notes: The table shows $share_L$ implied by the calibration strategy of Jones (2014).

broadly consistent with his data). Even restricting attention to the elasticity of substitution between 1.5 and 2, share_L ranges from 12% to 270%.

The closed form solution (12) reveals why $share_L$ declines strongly with the skill cutoff. The amplification term is determined by the difference in the unskilled wage bill share across countries $R(1 + S(W))$, which declines with the skill cutoff. Especially for low substitution elasticities, its variation dominates how $share_L$ differs across cutoffs.

Since the same closed form solution also applies to our baseline model and since we use similar data values for employment shares $N_{j,c}$ and wage bill ratios $S(W_c)$, the differences in development accounting relative to Jones (2014) must stem entirely from the estimation of $R(h_1)$ and ρ. We explore these differences in the following sections.

3.4 Estimating Human Capital Gaps

The first difference between our approach and Jones (2014) is the determination of the unskilled human capital gap $R(h_1)$. While Jones sets $R(h_1) \approx 1$, we estimate it using wage gains at migration, which yields values between 2 and 3.3.

We noted in Section 2.4 that, when skills are perfect substitutes, the human capital gap $R(h)$ can be estimated as $R(y)/wg$. A similar result holds when there are multiple skills. From $w_{j,c} = p_{j,c}h_{j,c}$, we have

$$R(h_j) = \frac{R(w_j)}{R(p_j)} \tag{14}$$

Intuitively, observed wages are higher in rich countries either due to skill price gaps or due human capital gaps: $R(w_j) = R(p_j)\,R(h_j)$. Migrant wage gains estimate skill price gaps

$(wg_j = R(p_j))$ and therefore allow us to calculate human capital gaps.[10]

Table 5 shows the human capital gaps $R(h_j)$ implied by equation (14). The table also shows the two ratios that determine these values: observable wage gaps $R(w_j)$ and migrant wage gains wg_j. We highlight a number of findings:

1. For all skill cutoffs, the fact that cross-country wage gaps exceed migrant wage gains implies that workers of all skills have more human capital in rich compared with poor countries $(R(h_j) > 1)$.

2. Higher skill cutoffs are associated with larger wage gaps $R(w_j)$, smaller migrant wage gains, and therefore larger human capital gaps $R(h_j)$.

3. In the rich country, skilled workers have relatively more human capital than in the poor country: $RS(h) = \frac{h_{2,r}/h_{1,r}}{h_{2,p}/h_{1,p}} > 1$. Since migrant wage gains are similar for skilled and unskilled workers, $RS(h)$ differs at most 1.6 fold across countries.[11] This limits the size of the amplification term in (12) (recall that amplification vanishes when $RS(h) = 1$).

The human capital gaps $R(h_j)$ directly translate into estimates of the contribution of human capital to output gaps. Any constant returns to scale labor aggregator implies that $share_h \in \left\{\frac{\ln R(h_1)}{\ln R(y)}, \frac{\ln R(h_2)}{\ln R(y)}\right\}$. For the estimated values of $R(h_j)$, the lower bounds range from 29% to 51% of output gaps. The lower bound is of particular interest because we expect that $share_L$ exceeds $share_h$ since rich countries also have higher years of schooling than poor countries.

Note that the bounds for $share_h$ do not depend on the functional form of the labor aggregator. The same bounds apply for any model where doubling all labor inputs doubles steady state output. Similarly, the estimation of $R(h_j)$ is independent of most of the model structure. It only assumes 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 $R(p_j)$.

3.5 Elasticity Implications

The second difference between our approach and Jones (2014) is the determination of the elasticity of substitution between skilled and unskilled labor. Jones abstracts from the possibility of technological adaptation in the face of large and persistent differences in

[10] The ratio of wages $R(w_j)$ can be estimated from $R(w_j) = R(W_j)/R(N_j)$. Given data for skill premiums $S(w_c)$ and employment shares by skill $N_{j,c}$, we can calculate wage bill ratios $S(W_c)$. Using data for output per worker y_c, and labor income shares $1-\alpha$ we can calculate wage bill levels as $W_{1,c} = (1-\alpha)\, y_c / (1 + S(W_c))$ and $W_{2,c} = W_{1,c}S(W_c)$.

[11] Specifically, with equal skill premiums in rich and poor countries, we have $RS(h) = 1/S(wg) \in [1.4, 1.6]$. This follows from $S(w_p) = S(p_p h_p) = S(w_r) = S(p_r h_r)$.

Table 5: Cross-country Human Capital Gaps

	Skill Cutoff			
	SHS	HSG	SC	CG
$R(h_1)$	2.00	2.00	2.45	3.35
$R(w_1)$	7.41	6.90	7.29	9.49
wg_1	3.71	3.46	2.98	2.84
$R(h_2)$	3.24	3.12	3.51	4.65
$R(w_2)$	7.41	6.90	7.29	9.49
wg_2	2.29	2.21	2.08	2.04
$RS(h)$	1.62	1.57	1.43	1.39
$share_{h_1}$	0.29	0.29	0.38	0.51

Notes: The table shows the rich-to-poor country human capital ratios, $R(h_j)$, and their components according to (14). $RS(h)$ denotes the cross-country gap in relative human capital $h_{2,c}/h_{1,c}$. $share_{h_1}$ is the lower bound for the share of cross-country output gaps due to human capital implied by a constant returns to scale labor aggregator.

the relative supply of skilled labor. This leads him to use conventional estimates of the elasticity of substitution from the literature. We allow for technological adaptation, here modeled as choosing an appropriate technology along a frontier. This leads us to calibrate an alternative, higher long-run elasticity. The higher elasticity of substitution limits the size of the amplification term in (12) and rules out very large values of $share_L$.

To understand why we find a high elasticity, consider the firm's first-order condition for labor inputs, which implies

$$RS(W) = RS(L)^{\Psi} \tag{15}$$

Since $RS(W) = RS(N) = RS(L)/RS(h)$, we have

$$RS(h) = RS(N)^{(1-\Psi)/\Psi} \tag{16}$$

so that the elasticity of substitution between skilled and unskilled labor is given by

$$\frac{1}{1-\Psi} = 1 + \frac{\ln RS(N)}{\ln RS(h)} \tag{17}$$

This may be estimated using $RS(h) = 1/S(wg)$. Table 6 shows the corresponding data

Table 6: Long-run Elasticity of Substitution

	Skill Cutoff			
	SHS	HSG	SC	CG
Elasticity	7.83	4.53	5.15	4.03
$\ln RS\,(N)$	3.31	1.58	1.49	1.00
$\ln RS\,(h)$	0.48	0.45	0.36	0.33

values for each skill cutoff. We highlight three observations:

1. $\ln RS\,(N)$ is positive for all skill cutoffs. Recall that $RS\,(N)$ is the ratio of skilled to unskilled workers in the rich country, relative to the poor country. It can be thought of as a measure of relative factor abundance. The fact that its log is greater than 0 indicates that rich countries are abundant in skilled labor.

2. Cross-country variation in relative employment $RS\,(N)$ is much larger than cross-country variation in relative human capital $RS\,(h)$. As a result, the elasticity of substitution is always high (at least 4).

3. As the skill cutoff is increased, $RS\,(N)$ declines (as previously noted) while $RS\,(h)$ is fairly stable, causing the elasticity to decline as well.

Intuitively, cross-country variation in relative skill prices is limited because $RS\,(p) = S\,(wg)$ and wage gains do not vary greatly across skill groups. Reconciling small variation in relative skill prices with large variation in relative labor inputs requires a high elasticity of substitution. Conversely, a smaller long-run elasticity of substitution would imply much larger differences in migrant wage gains between skilled and unskilled workers than we see in the data.[12]

We can now summarize why our results differ from those of Jones (2014). First, we use the wage gains of immigrants to discipline $R\,(h)$ for unskilled workers, which increases the contribution of human capital. Second, we calibrate a much smaller (long-run) elasticity of substitution. This has two effects. First, it reduces the amplification term, which rules out very large results. Second, it implies that our results are much less sensitive to skill cutoffs than are those of Jones, because a larger value of the elasticity of substitution puts less weight on the large and variable term $\ln R\,(1 + S(W))$.

[12] Specifically, for a long-run elasticity of 2, we find $RS\,(h)$ from equation (16), setting $\Psi = 0.5$. Using $S\,(wg) = 1/RS\,(h)$, we find that unskilled wage gains exceed skilled wage gains by factor 2.8 for the CG skill cutoff and by factor 27 for the SHS skill cutoff. In the data, the ratio is at most 1.6 (see Table 5).

3.6 Robustness

The previous discussion reveals that migrant wage gains play a central role for our results. The discussion in Hendricks and Schoellman (2018) addresses a number of concerns related to the interpretation of migrant wage gains as measures of cross-country skill price differences. Here, we address the possibility that unskilled wage gains may be mismeasured in their data.

This concern arises because the New Immigrant Survey and Latin Migration Project data used by Hendricks and Schoellman (2018) contain few unskilled migrants from low income countries. As a result, the migrant wage gains of this group may be understated. We explore the robustness of our findings by increasing the wage gains of the least skilled group with no secondary education to the point where their human capital no longer differs across countries, so that $R(h_1) = 1$.

The closed form solution for $share_L$ given by (12) reveals that increasing wg_1 has two opposing effects on $share_L$. First, the base term declines because higher wage gains indicate larger contributions of "country" to output gaps. Second, the amplification term increases because the elasticity of substitution is reduced according to (13). Intuitively, larger unskilled migrant wage gains imply that $R(h_1)$ declines while $R(h_2)$ is held fixed. As a result, the relative abundance of skilled labor $RS(L)$ increases. Matching the observed wage bill ratios then requires a smaller elasticity of substitution.

Quantitatively, the net result is that $share_L$ declines modestly, as shown in Table 7. Depending on the skill cutoff, it ranges from 0.54 to 0.6. The substitution elasticities decrease to values between 3.3 and 4.0. Our findings are now close to the preferred parameterizations of Jones (2014) who assumes $R(h_1) \approx 1$ and sets the elasticity of substitution based on published empirical estimates.

We have performed similar robustness checks for all of the model versions that we consider with very similar results. Details are available upon request.

3.7 Relative Skilled Labor Productivities

One contribution of our work is to allow for both relative human capital $RS(h)$ and relative skill bias $RS(\theta)$ in the same framework and to disentangle the two. Thus, we can contribute to the ongoing debate on which of these two forces explains the constancy of skill premiums across countries given the enormous differences in skilled labor supplies.

We estimate $RS(\theta)$ based on the firm's first-order condition for labor, which implies

$$RS(\theta h) = RS(N)^{(1-\rho)/\rho} \tag{18}$$

This relates the differences in the relative abundance of skilled labor $RS(N)$ to differences

Table 7: Robustness: share_L For Larger Migrant Wage Gains

	Skill Cutoff			
	SHS	HSG	SC	CG
$share_L$	0.60	0.54	0.57	0.54
Base term	0.15	0.34	0.48	0.51
Amplification term	0.45	0.20	0.09	0.03
$1/\Psi - 1$	0.36	0.49	0.34	0.44
$\frac{\ln R(1+S(W))}{\ln R(y)}$	1.27	0.42	0.26	0.07
$share_k$	0.04	0.04	0.04	0.04
$share_z$	0.36	0.42	0.40	0.42

Notes: The table shows the closed form solution for $share_L$ according to equation (12) and its components. The wage gain of the least skilled migrants is set to that $R(h_1) = 1$.

in the relative productivity of skilled labor $RS(\theta h)$. Equation (18) also applies in the case where labor-augmenting technologies are fixed exogenously.

Intuitively, skilled wage premiums appear similar across countries. Given large differences in relative labor supply $RS(N)$ and conventional estimates of the (short-run) elasticity of substitution, some combination of skill bias of technologies $(RS(\theta) > 1)$ or a relative human capital per worker advantage $(RS(h) > 1)$ is required.

Previous work explored versions of equation (18) with one of these two possibilities ruled out, eliminating the identification challenge (Caselli and Coleman, 2006; Jones, 2014). Substantial disagreement has ensued (Caselli and Ciccone, 2019; Jones, 2019).

Our approach is to use the new evidence from the wage gains of migrants to discipline $RS(h)$. Implicitly, the remainder is attributed to $RS(\theta)$. One way to think about the model with endogenous technology choice is that we calibrate the value of Ψ (and implicitly the curvature of the technology frontier ω) to induce firms to choose $RS(\theta)$ as an optimal response to $RS(h)$ and $RS(N)$. However, we can also measure $RS(\theta)$ directly without this structure.

Table 8 shows our results. For typical estimates of the (short-run) elasticity of substitution, the relative skill bias $RS(\theta)$ exceeds two. It is larger for low skill cutoffs (because they imply larger $RS(N)$, which increases the right-hand side of equation (18)) and smaller values of ρ (which also increases the right-hand side of equation (18)).

Table 9 shows the fraction of cross-country variation in the relative productivity of skilled

Table 8: Relative Skill Bias Rich vs. Poor

Short-run Elasticity	Skill Cutoff			
	SHS	HSG	SC	CG
1.25	3.50×10^5	355.98	274.06	39.14
1.50	463.99	15.08	13.83	5.30
2.00	16.91	3.10	3.11	1.95
3.00	3.23	1.41	1.47	1.18
4.00	1.86	1.08	1.15	1.00
5.00	1.41	0.95	1.01	0.92

Notes: The table shows $RS(\theta)$.

labor $RS(\theta h)$ that is due to human capital, defined as $\ln RS(h) / \ln RS(\theta h)$. Since relative skilled labor endowments $RS(h)$ do not vary with σ, this fraction varies inversely with the relative skill bias ratios shown in Table 8. For conventional values of the elasticity of substitution (between 1.5 and 2), at most one-third of the cross-country variation in relative skilled labor productivity is due to human capital. However, the fraction rises rapidly as the elasticity increases.

These findings agree with the previous work of Rossi (2019), who also decomposes cross-country variation in the relative productivity of skilled labor into the contributions of relative human capital and technological skill bias. He uses the returns to schooling of foreign-educated immigrants as the extra moment to provide identification (rather than wage gains of immigrants) and concludes that 90% of the variation can be attributed to technology. If we focus on the same definition of skill (some college or more) and elasticity of substitution (1.5), we find a very similar share of 88%.

4 Exogenous Skill Bias

We study a version of the model presented in Section 2 with one departure: instead of allowing firms to choose skill bias parameters from the technology frontier, we take the values of $\theta_{j,c}$ as exogenous. Our goal is to see how far we can go with a more minimal structure that does not impose optimal endogenous technology choice.[13]

[13] We continue to refer to $\theta_{j,c}$ as technological skill bias, recognizing that its variation may be due to factors other than technology, as argued by Caselli and Ciccone (2019).

Table 9: Fraction of Relative Skilled Labor Productivity Differences Due to Human Capital

Short-run Elasticity	Skill Cutoff			
	SHS	HSG	SC	CG
1.25	3.7	7.1	6.0	8.3
1.50	7.3	14.2	12.0	16.5
2.00	14.6	28.3	24.1	33.0
3.00	29.3	56.7	48.2	66.0
4.00	43.9	85.0	72.3	99.1
5.00	58.5	113.4	96.4	132.1

Notes: The table shows $100 \times \ln RS\,(h)\,/\ln RS\,(\theta h)$.

4.1 Development Accounting Approach

Development accounting assesses how each factor input affects steady state output. As pointed out by Jones (2019), the effect of changing labor inputs is not uniquely determined when skilled labor and technological skill bias are complements; it depends on the reference country's technology. Our baseline model sidesteps this issue because the skill bias of technology is endogenous and depends on labor endowments. Now that the skill bias of technology is taken as fixed, we confront this issue by considering two definitions of $share_L$:

1. $share_L^{poor}$ fixes the skill bias of technology at the poor country level. Intuitively, this corresponds to the effect of increasing poor country labor inputs to rich country levels.
2. $share_L^{rich}$ fixes the skill bias of technology at the rich country level. Intuitively, this corresponds to the effect of reducing rich country labor inputs to poor country levels.

4.2 Closed Form Solution

We derive a closed form solution for $share_L^{poor}$ and $share_L^{rich}$ in terms of observable data moments.

Proposition 3. *The share of labor inputs evaluated at poor country skill bias is given by*

$$share_L^{poor} = \underbrace{1 - \frac{\ln\left(wg_1\right)}{\ln R\left(y\right)}}_{base} + \underbrace{\frac{\frac{1}{\rho}\ln\frac{1+S(W_p)RS(L)^{\rho}}{1+S(W_p)} - \ln R\left(1+S\left(W\right)\right)}{\ln R\left(y\right)}}_{amplification} \tag{19}$$

When evaluated at rich country skill bias, the share of labor inputs is given by

$$share_L^{rich} = \underbrace{1 - \frac{\ln(wg_1)}{\ln R(y)}}_{base} + \underbrace{\frac{\frac{1}{\rho}\ln\frac{1+S(W_r)}{1+S(W_r)RS(L)^{-\rho}} - \ln R(1+S(W))}{\ln R(y)}}_{amplification} \tag{20}$$

Proof. Section D.1. □

Both expressions resemble the closed form solution for $share_L$ obtained from the model with endogenous skill bias (12).[14] $share_L$ is comprised of two parts. The base term is the same as in the model with endogenous skill bias. It represents the contribution of labor inputs with a single skill. The second term represents the amplification due to imperfect skill substitution.

As in the baseline model, the amplification term depends on the elasticity of substitution between skilled and unskilled labor (now governed by ρ) and on the labor income ratios $S(W_c)$. The amplification term is small when skilled labor is "unimportant" in the sense of earning little income, so that $S(W_c)$ is small. It tends to be large when countries differ greatly in their relative labor endowments, so that $RS(L)$ is large.

4.3 Quantitative Results

We calibrate the model to match the same data moments that were used in the calibration of the baseline model. However, $share_L$ now depends on the short-run elasticity of substitution. The data moments are therefore not sufficient to perform development accounting. Following Jones (2014), we explore a range of values for ρ.

Table 10 shows the share of output gaps accounted for by labor inputs, evaluated at poor country skill bias values ($share_L^{poor}$). For conventional values of the elasticity of substitution between 1.5 and 2 (Ciccone and Peri, 2005), $share_L^{poor}$ ranges from 50% to 57%. For lower elasticities, especially for the SHS skill cutoff, $share_L^{poor}$ can drop below 50%. However, it is worth keeping in mind that these cases imply extremely large cross-country differences in skill bias (see Table 8).[15] As the elasticity approaches the value implied by the model with the technology frontier (Ψ), $share_L^{poor} \to share_L$.

Table 11 shows the corresponding results when the contribution of labor inputs is evaluated using rich country skill bias parameters. For substitution elasticities in the conventional range between 1.5 and 2, $share_L^{rich}$ ranges from 59% to 74%. Across all cells, the range is only modestly wider.

[14] In fact, when $\rho = \Psi$, given by (13), $share_L^{poor} = share_L = share_L^{rich}$ because $S(W_p)RS(L)^{\Psi} = S(W_r)$.

[15] For any given value of ρ, the model implies the same values for $h_{j,c}$ and $RS(\theta)$ as the model with the technology frontier. Since changing all $\theta_{j,c}$ by a common factor is equivalent to varying z_c, the skill bias parameters are only identified up to country specific constants.

Table 10: Development Accounting with Poor Country Skill Bias

Short-run Elasticity	Skill Cutoff			
	SHS	HSG	SC	CG
1.25	0.44	0.48	0.50	0.56
1.50	0.50	0.51	0.52	0.56
2.00	0.56	0.54	0.55	0.57
3.00	0.60	0.57	0.58	0.58
4.00	0.61	0.59	0.59	0.58
5.00	0.62	0.60	0.60	0.58
Endog. θ	0.63	0.59	0.60	0.58

Notes: The table shows $share_L^{poor}$ for selected values of the elasticity of substitution between skilled and unskilled labor (rows) and for selected skill cutoffs (columns). Each column represents a skill cutoff. The last row shows the contribution of labor inputs when skill bias is endogenous, $share_L$, taken from Table 3.

To understand the diverging patterns between $share_L^{rich}$ and $share_L^{poor}$, it is useful to remember from Table 1 that the relative abundance of skilled labor $RS\left(N\right)$ is much larger with lower skill cutoffs such as SHS. In order to fit the targets, our calibration infers much larger gaps between rich and poor countries in labor augmenting technologies $RS\left(\theta\right)$ in this case. Thus, development accounting results become much more sensitive to whether we use poor or rich country technologies as the benchmark. Equations (19) and (20) show that this effect interacts with ρ, so that the divergence is larger as the elasticity of substitution moves away from the value we calibrated in the endogenous skill bias case.

5 Capital-skill Complementarity

In this section, we consider capital-skill complementarity as an alternative source of cross-country variation in skilled labor productivity. The model specification is based on Krusell et al. (2000).

Table 11: Development Accounting with Rich Country Skill Bias

Short-run Elasticity	Skill Cutoff SHS	HSG	SC	CG
1.25	0.75	0.71	0.71	0.61
1.50	0.74	0.68	0.68	0.60
2.00	0.72	0.65	0.65	0.59
3.00	0.69	0.62	0.62	0.59
4.00	0.67	0.60	0.61	0.58
5.00	0.65	0.59	0.60	0.58
Endog. θ	0.63	0.59	0.60	0.58

Notes: The table shows $share_L^{rich}$ for selected values of the elasticity of substitution between skilled and unskilled labor (columns) and for selected skill cutoffs (rows).

5.1 Model Specification

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

$$y_c = s_c^\alpha \left(z_c L_c\right)^{1-\alpha} \tag{21}$$

where

$$L_c = \left[(\theta_{1,c} L_{1,c})^\rho + (\theta_{2,c} Z_c)^\rho\right]^{1/\rho} \tag{22}$$

and

$$Z_c = \left[(\mu_e e_c)^\phi + (\mu_2 L_{2,c})^\phi\right]^{1/\phi} \tag{23}$$

with parameters $\alpha, \rho \in (0,1)$, $\phi < 1$, and $\mu_e, \mu_2 > 0$.

s_c denotes structures per capita. L_c is given by a CES aggregator of unskilled labor $L_{1,c}$ and a composite input Z_c, which is in turn a CES aggregator of skilled labor $L_{2,c}$ and equipment e_c. The skill bias parameters $\theta_{j,c}$ are constrained by the technology frontier (3) with $B_c = 1$ taken as fixed. The baseline model emerges as a special case when $\mu_e = 0$ so that $Z_c = L_{2,c}$.

As before, 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. This fixes the rental prices of equipment $q_{e,c}$ and structures $q_{s,c}$ and therefore also s_c/y_c. The representative firm solves

$$\max_{s_c, e_c, L_{j,c}, \theta_{j,c}} y_c - q_{s,c} s_c - q_{e,c} e_c - \sum_{j=1}^{J} p_{j,c} L_{j,c} \tag{24}$$

subject to (21), (22), (23), and the frontier constraint (3).

5.2 Reduced Form Labor Aggregator

Similar to the baseline model, we are able to derive a reduced form labor aggregator that substitutes out the firm's optimal skill bias choices.

Proposition 4. *Substituting out the firm's optimal skill bias choices yields the reduced form labor aggregator*

$$L_c = B_c \left(\left[L_{1,c} / \kappa_{1,c} \right]^{\Psi} + \left[Z_c / \kappa_{2,c} \right]^{\Psi} \right)^{1/\Psi} \tag{25}$$

with $\Psi = \frac{\omega \rho}{\omega - \rho}$ *as in the baseline model.*

Proof. Section F.2.1 □

5.3 Development Accounting

Development accounting proceeds analogously to the baseline model. Starting from

$$y_c = \left(s_c / y_c \right)^{\alpha/(1-\alpha)} z_c L_c \tag{26}$$

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

$$\underbrace{\ln R\left(y\right)}_{\text{output gap}} = \underbrace{\ln R\left(z\right)}_{\text{TFP}} + \underbrace{\ln R\left(\left(s/y\right)^{\alpha/(1-\alpha)}\right)}_{\text{structures}} + \underbrace{\ln R\left(L\right)}_{\text{labor and equipment}} \tag{27}$$

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

$$1 = \underbrace{\frac{\ln R\left(z\right)}{\ln R\left(y\right)}}_{\text{share}_z} + \underbrace{\frac{\ln R\left(\left(s/y\right)^{\alpha/(1-\alpha)}\right)}{\ln R\left(y\right)}}_{\text{share}_s} + \underbrace{\frac{\ln R\left(L\right)}{\ln R\left(y\right)}}_{share_{L+e}} \tag{28}$$

The joint contribution of labor inputs and equipment has a closed form solution in terms of data moments (see Section F.2.2). It may be subdivided into the separate contributions of its components ($h_{j,c}$, $N_{j,c}$, e_c). These are defined as the changes in steady state output

Table 12: Additional Calibration Targets

	s/y	e/y
Rich	2.81	0.37
Poor	2.85	0.14
Ratio	0.98	2.62

that result from changing each input from its poor country value to its rich country value, holding the rental prices of equipment and structures fixed. The counterfactual output changes depend on the fixed equipment rental prices. We therefore define two versions of each input's share. Superscript "poor" fixes q_e at the poor country's level. Superscript "rich" fixes them at the rich country's level.

As in the baseline model, the development accounting implications depend on the reduced form curvature parameter Ψ, but not on the separate values of ρ and ω.

5.4 Calibration

We calibrate the model using the same data moments that were used for the baseline model. However, we replace the moments related to capital inputs with separate moments for equipment and structures. Specifically, 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 $IS_{s,c}$; see Section A.1).[16] These data moments are summarized in Table 12. We find that s/y is similar for rich and poor countries, while e/y is 2.6 times higher in rich versus poor countries.

In total, we have 14 data moments (6 independent factor incomes shares, 2 output levels, 2 wage gains at migration, 4 capital/output ratios). However, choosing units of e to normalize $\kappa_1 = 1$ means that we need to replace the data moments e_c/y_c with $R(e)$.[17] This leaves us with 13 data moments that can be used to calibrate the model's 13 parameters (z_c; α; $h_{j,c}$, where $h_{1,r} = 1$; e_c and s_c; Ψ, ϕ).

5.5 Development Accounting Results

Table 13 summarizes the development accounting implications. Across skill cutoffs, labor inputs and equipment jointly account for around three-quarters of cross-country output

[16] Due to data limitations, we assume that the income shares of structures and equipment do not differ across countries, but explore the robustness of our findings when this assumption is relaxed.

[17] We also normalize $h_{1,r} = 1$ so that $L_{u,r} = N_{u,r}$. We set $\mu_2 = 1$ by choosing units of h_2. We may normalize μ_e, κ_2 and B_c to 1 as varying them has the same effect as varying z_c.

Table 13: Development Accounting with Capital-skill Complementarity

	Skill Cutoff			
	SHS	HSG	SC	CG
$share_L^{poor}$	0.65	0.61	0.62	0.58
$share_L^{rich}$	0.68	0.67	0.70	0.65
$share_{L+e}$	0.78	0.75	0.76	0.74
Elasticity	4.77	2.51	2.17	1.37

gaps. Using poor country equipment prices, human capital accounts for around 60% of output gaps. Using the lower rich country equipment prices, $share_L^{rich}$ is moderately higher, ranging from 65% to 70%. Since skilled labor and equipment are complements, increasing labor inputs has larger effects on output when equipment is abundant.

The reduced form elasticities of substitution $1/(1-\Psi)$ are much smaller than in the model without capital-skill complementarity. The intuition is based on the observation that Z/L_1 varies more across countries than L_2/L_1. At the same time, the relative income share of Z versus L_1 varies less than $S(W)$. Hence, a smaller elasticity reconciles cross-country variation in factor incomes and factor income shares. For the higher skill cutoffs, the elasticities of substitution are in line with conventional estimates for $1/(1-\rho)$.

For completeness, Table 14 summarizes the shares of output gaps accounted for by other inputs evaluated at poor country equipment prices. Structures make essentially no contribution. Equipment contributes about 8%. The contribution of TFP is given by $1 - share_s - share_{L+e}$ and therefore amounts to about 22%.[18] The complementarity of skilled labor and equipment implies that jointly increasing both inputs has a larger effect on output than increasing each input separately. This explains why $share_{L+e}$ is almost ten percentage points larger than $share_L + share_e$.

We also explore a version of the model where the skill bias of technology is taken as exogenous. For conventional values of the elasticity of substitution between skilled and unskilled labor, we find that $share_L$ ranges from 52% to 74%. Details are relegated to Section F.3.[19]

Finally, Table 15 shows the fraction of the cross-country variation in relative skilled labor productivities that is due to human capital, defined as $\ln RS(h) / \ln RS(\theta h)$. Even if atten-

[18] The results are very similar when we use rich country equipment prices instead. The contribution of equipment is defined as the steady state output change induced by changing q_e from $q_{e,p}$ to $q_{e,r}$ holding q_s fixed (its level does not matter).

[19] These findings are robust to reasonable variations in the equipment stocks or equipment income shares that we use as calibration targets. For example, $share_L$ remains above one-half even if we reduce the poor country's equipment stock to one quarter of its estimated value. $share_L$ remains above 0.46 when the the poor country's equipment income share is reduced by half.

Table 14: Development Accounting with Poor Country Equipment Prices

	Skill Cutoff			
	SHS	HSG	SC	CG
$share_L$	0.65	0.61	0.62	0.58
$share_e$	0.07	0.06	0.05	0.06
$share_s$	0.00	0.00	0.00	0.00
$share_z$	0.22	0.25	0.24	0.26

Table 15: Fraction of Relative Skilled Labor Productivity Variation Due to Human Capital

Short-run	Skill Cutoff			
Elasticity	SHS	HSG	SC	CG
1.25	3.8	8.7	8.9	33.3
1.50	7.9	19.1	21.2	n/a
2.00	16.9	48.3	67.8	n/a
3.00	38.9	203.7	n/a	n/a
4.00	68.7	n/a	n/a	n/a
5.00	111.3	n/a	n/a	n/a

Notes: The table shows $100 \times \ln RS(h) / \ln RS(\theta h)$. This is not defined for cases where the rich country's technology is less skill biased than the poor country's technology.

tion is restricted to conventional values of the elasticity, the fraction due to human capital ranges from 8% to 68%. The reason is that cross-country variation in the abundance of equipment reduces the variation in relative skill bias needed to account for the constancy of skill premiums across countries.

6 Conclusion

We evaluate the contribution of human capital in accounting for cross-country income differences in an environment with imperfect substitution between skill types and a variety of factors that shift relative labor supply or demand. The literature has found it challenging to identify the answer to this question. We provide additional discipline by incorporating information on the wage gains of migrants and show that this reduces the range of plausible variation substantially, to between one-half to three-quarters of cross-country income differences. 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.

References

ACEMOGLU, D. (1998): "Why do new technologies complement skills? Directed technical change and wage inequality," *The Quarterly Journal of Economics*, 113, 1055–1089.

——— (2002): "Directed technical change," *Review of Economic Studies*, 69, 781–810.

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

BANERJEE, A. AND E. DUFLO (2005): "Growth theory through the lens of development," in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, Elsevier, vol. 1A, 473–554.

BARRO, R. J. AND J. W. LEE (2013): "A new data set of educational attainment in the world, 1950–2010," *Journal of Development Economics*, 104, 184 – 198.

BILS, M. AND P. J. KLENOW (2000): "Does Schooling Cause Growth?" *The American Economic Review*, 90, 1160–1183.

CASELLI, F. (2005): "Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, Elsevier, vol. 1B, chap. 9.

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

CASELLI, F. AND A. CICCONE (2019): "The Human Capital Stock: A Generalized Approach Comment," *American Economic Review*, 109, 1155–74.

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.

CÓRDOBA, J. C. AND M. RIPOLL (2013): "What explains schooling differences across countries?" *Journal of Monetary Economics*, 60, 184–202.

CUBAS, G., B. RAVIKUMAR, AND G. VENTURA (2016): "Talent, labor quality, and economic development," *Review of Economic Dynamics*, 21, 160–81.

EROSA, A., T. KORESHKOVA, AND D. RESTUCCIA (2010): "How important is human capital? a quantitative theory assessment of world income inequality," *The Review of Economic Studies*, 77, 1421–1449.

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

GANCIA, G. AND F. ZILIBOTTI (2009): “Technological change and the wealth of nations,” *Annu. Rev. Econ.*, 1, 93–120.

GOLDIN, C. AND L. F. KATZ (2008): *The Race between Education and Technology*, Harvard University Press.

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.

JERZMANOWSKI, M. AND R. TAMURA (2019): “Directed Technological Change & Cross Country Income Differences: A Quantitative Analysis,” *Journal of Development Economics*, 141.

JONES, B. (2019): “The Human Capital Stock: A Generalized Approach: Reply,” *American Economic Review*, 109, 1175–95.

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.

MALMBERG, H. (2018): “How does the efficiency of skilled labor vary across rich and poor countries? An analysis using trade and industry data,” Manuscript. Institute for International Economic Studies.

PARRO, F. (2013): “Capital-Skill Complementarity and the Skill Premium in a Quantitative Model of Trade,” *American Economic Journal: Macroeconomics*, 5, 72–117.

ROSSI, F. (2019): “The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation,” Manuscript. University of Warwick.

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

Online Appendix

A Data

The raw data on migrant wage gains and output gaps are taken from Hendricks and Schoellman (2018). Labor inputs in efficiency units are given by $N_{j,c} = \sum_{s \in j} e^{\phi t_s} N_{s,c}^{BL}$ where $N_{s,c}^{BL}$ denotes the employment share of school group s taken from Barro and Lee (2013). s takes on seven values (no school, some primary, primary completed, some secondary, secondary completed, some tertiary, tertiary completed). We set the school durations to $t_s = [0, 3, 6, 9, 12, 14, 16]$.

Based on the evidence collected by Banerjee and Duflo (2005), we assume that skill premiums are the same in rich and poor countries. Specifically, we assume a Mincer return of $\phi = 0.1$. For each country, we normalize $N_c = \sum_{j=1}^{J} N_{j,c} = 1$.

Labor income shares $IS_{j,c}$ are constructed as follows. The labor income of school group s (up to an arbitrary, country specific scale factor) is defined as $W_{s,c} = e^{\phi t_s} N_{s,c}^{BL}$. The share of skill group j is then given by: $IS_{j,c} = \frac{\sum_{s \in j} W_{s,c}}{\sum_s W_{s,c}} (1 - \alpha)$. With equal Mincer returns in rich countries, relative earnings and relative labor inputs vary across countries by the same amount: $RS(W) = RS(N)$.

A.1 Equipment and Structures Data

Calibrating the model with capital-skill complementarity requires additional data moments related to equipment and structures that are constructed as follows. $IS_{e,r} = 0.15$ is taken from Valentinyi and Herrendorf (2008). Together with a labor share of 0.33, 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 relative price of structures versus consumption are similar for rich and poor countries, we set $IS_{s,c} = 0.18$ 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 data. 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%.

B Baseline Model

The following derivations apply to all models without capital-skill complementarity. Skill bias can be exogenous or chosen from a technology frontier.

B.1 Notation

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

1. The income share of an input is denoted by $IS_{a,c} = \text{income}_{a,c}/y_c$.

2. The income ratio of two inputs is denoted by $IR_{a/b,c} = \text{income}_{a,c}/\text{income}_{b,c}$. In particular, $IR_{L_2L_1} = S\left(W\right)$.

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. $R\left(x^{\phi}\right) = R\left(x\right)^{\phi}$ and $S\left(x^{\phi}\right) = S\left(x\right)^{\phi}$.

2. The order of rich-to-poor and skilled-to-unskilled ratios is interchangeable:

$$R\left(S\left(x^{\phi}\right)\right) = S\left(R\left(x^{\phi}\right)\right) \tag{29}$$

$$= \left[\frac{x_{2,r}/x_{1,r}}{x_{2,p}/x_{1,p}}\right]^{\phi} \tag{30}$$

B.2 CES Results

It is useful to state a number of known properties of cost minimization with CES production. These results will be used repeatedly in the derivations below.

Consider the generic cost minimization problem

$$\min_{x_j} \sum_{j=1}^{J} p_j x_j + \lambda \left[\bar{y} - \left[\sum_{j=1}^{J} (\gamma_j x_j)^\rho \right]^{1/\rho} \right] \tag{31}$$

The cost-minimizing input ratios are given by

$$\left[\frac{x_i}{x_j}\right]^{1-\rho} = \left[\frac{\gamma_i}{\gamma_j}\right]^{\rho} \frac{p_j}{p_i} \tag{32}$$

The ratio of factor incomes is then given by

$$\frac{p_i x_i}{p_j x_j} = \left[\frac{\gamma_i x_i}{\gamma_j x_j}\right]^{\rho} \tag{33}$$

$$= \left[\frac{\gamma_i}{\gamma_j}\right]^{\frac{1}{1-\rho}} \left[\frac{p_j}{p_i}\right]^{\frac{\rho}{1-\rho}} \tag{34}$$

The income share of each input is given by

$$\frac{p_j x_j}{\bar{y}} = \left[\frac{\gamma_j x_j}{\bar{y}}\right]^{\rho} \tag{35}$$

The minimized cost per unit of output is given by

$$p_y = \left[\sum_{j=1}^{J} (\gamma_j p_j)^{\frac{\rho}{1-\rho}} \right]^{\frac{1-\rho}{\rho}} \tag{36}$$

B.3 Firm First-order Conditions

The firm's first-order conditions for labor inputs are given by

$$p_{j,c} = (1-\alpha)\, z_c^{1-\alpha} k_c^{\alpha} L_c^{1-\rho-\alpha} \theta_{j,c}^{\rho} L_{j,c}^{\rho-1} \tag{37}$$

If skill bias is endogenous, the first-order condition for $\theta_{j,c}$ is given by

$$\frac{\partial y_c}{\partial \theta_{j,c}} = \lambda_c \omega \kappa_j^{\omega} \theta_{j,c}^{\omega-1} \tag{38}$$

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

$$\frac{\partial y_c}{\partial \theta_{j,c}} = (1-\alpha)\, k_c^{\alpha} z_c^{1-\alpha} L_c^{1-\rho-\alpha} L_{j,c}^{\rho} \theta_{j,c}^{\rho-1} \tag{39}$$

From (33), the wage bill ratio is given by

$$S(W_c) = S\left(p_c L_c\right) = S\left(\theta_c L_c\right)^{\rho} \tag{40}$$

Since $\rho > 0$, an increase in the relative supply of type j labor increases its income share.

C Endogenous Skill Bias

The derivations in this section apply for the model with endogenous skill bias.

C.1 Optimal Skill Bias Choice

The first-order conditions (38) imply the optimal skill bias ratio

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

Proposition 5. *Optimal skill bias levels are given by*

$$\theta_{1,c}^{\omega} = \frac{B_c}{\kappa_1^{\omega} \Lambda_c} \tag{42}$$

with

$$\Lambda_c = \sum_{j=1}^{J} \left(\frac{\kappa_1}{\kappa_j} \frac{L_{j,c}}{L_{1,c}}\right)^{\Psi} \tag{43}$$

This holds whether or not B_c is chosen by firms.

Proof. Starting from the technology frontier, we have

$$B_c^{\omega} = \sum_{j=1}^{J} \left(\kappa_j \theta_{j,c}\right)^{\omega} \tag{44}$$

$$= \left(\kappa_1 \theta_{1,c}\right)^{\omega} \sum_{j=1}^{J} \left(\frac{\kappa_j}{\kappa_1} \frac{\theta_{j,c}}{\theta_{1,c}}\right)^{\omega} \tag{45}$$

Substituting in the condition for optimal relative skill bias (41) yields

$$B_c^{\omega} = (\kappa_1\theta_{1,c})^{\omega} \sum_j \left(\frac{\kappa_j}{\kappa_1}\right)^{\omega} \left[\left(\frac{L_{j,c}}{L_{1,c}}\right)^{\rho} \left(\frac{\kappa_1}{\kappa_j}\right)^{\omega}\right]^{\frac{\omega}{\omega-\rho}} \tag{46}$$

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

$$B_c^{\omega} = (\kappa_1\theta_{1,c})^{\omega} \sum_{j=1}^{J} \left(\frac{\kappa_1}{\kappa_j}\frac{L_{j,c}}{L_{1,c}}\right)^{\Psi} = (\kappa_1\theta_{1,c})^{\omega} \Lambda_c \tag{47}$$

□

Proposition 6. *When skill bias is endogenous, the skill premium is given by*

$$S(p_c) = (S(L_c))^{\Psi-1} S(\kappa)^{-\Psi} \tag{48}$$

Proof. From (32) we have

$$S(p_c) = S(L_c)^{\rho-1} S(\theta)^{\rho} \tag{49}$$

Applying the optimal skill bias ratio (41)

$$S(\theta_c^{\rho}) = S\left(\kappa^{-\Psi} L_c^{\frac{\rho^2}{\omega-\rho}}\right) \tag{50}$$

yields

$$S(p_c) = S\left(\kappa^{-\Psi}\right) S(L_c)^{\rho+\frac{\rho^2}{\rho-\omega}-1} \tag{51}$$

Using $\rho + \frac{\rho^2}{\omega-\rho} = \frac{\rho\omega}{\omega-\rho} = \Psi$ gives (48). □

C.2 Reduced Form Labor Aggregator

Proof. (Proposition 1)

The following hold regardless of how B_c is determined (endogenous or fixed). The definition of the labor aggregator (22) implies

$$L_c = \theta_{1,c}L_{1,c}\left(\sum_{j=1}^{J}\left[\frac{\theta_{j,c}}{\theta_{1,c}}\frac{L_{j,c}}{L_{1,c}}\right]^{\rho}\right)^{1/\rho} \tag{52}$$

Substituting in the condition for the optimal choice of relative skill bias (50) yields

$$L_c = \theta_{1,c}L_{1,c}\left(\sum_{j=1}^{J}\left[\frac{L_{j,c}}{L_{1,c}}\right]^{\frac{\rho^2}{\omega-\rho}}\left[\frac{\kappa_j}{\kappa_1}\right]^{-\Psi}\left[\frac{L_{j,c}}{L_{1,c}}\right]^{\rho}\right)^{1/\rho} \tag{53}$$

The exponent on labor inputs is given by

$$\frac{\rho^2}{\omega-\rho}+\rho=\frac{\omega\rho}{\omega-\rho}=\Psi \tag{54}$$

Then the summation term becomes Λ_c, defined in (43), and we have

$$L_c=\theta_{1,c}L_{1,c}\Lambda_c^{1/\rho} \tag{55}$$

Then using (42), we have

$$L_c=B_c\kappa_1^{-1}\Lambda_c^{-1/\omega}L_{1,c}\Lambda_c^{1/\rho} \tag{56}$$

Note that

$$1/\rho-1/\omega=\frac{\omega-\rho}{\omega\rho}=1/\Psi \tag{57}$$

so that

$$\begin{aligned} L_c &= B_c\left(1/\kappa_1\right)L_{1,c}\Lambda_c^{1/\Psi} && (58)\\ &= B_c\left(1/\kappa_1\right)L_{1,c}\left[\sum_{j=1}^{J}\left(\kappa_j^{-1}L_{j,c}\right)^{\Psi}\right]^{1/\Psi}\kappa_1/L_{1,c} && (59)\end{aligned}$$

□

C.3 Closed Form Solution

Proof. (Proposition 2)

Using the labor aggregator (9) with $\kappa_j=1$, we have

$$\begin{aligned} R\left(L\right) &= \frac{L_{1,p}\left[R\left(L_1\right)^{\Psi}+S\left(L_r\right)^{\Psi}R\left(L_1\right)^{\Psi}\right]^{1/\Psi}}{L_{1,p}\left[1+S\left(L_p\right)^{\Psi}\right]^{1/\Psi}} && (60)\\ &= R\left(L_1\right)R\left(1+S\left(L\right)^{\Psi}\right)^{1/\Psi} && (61)\end{aligned}$$

Since (40) also applies to the reduced form labor aggregator, we have

$$S(W_c)=S\left(L_c\right)^{\Psi} \tag{62}$$

Using this to replace $S\left(L\right)^{\Psi}$ in (61) with $S(W)$ yields

$$R\left(L\right)=R\left(L_1\right)R\left(1+S(W)\right)^{1/\Psi} \tag{63}$$

If $S(L_r) > S(L_p)$, then $R(1+S(W)) > 1$ and $R(L) > R(L_1)$. The ratio of unskilled labor inputs is given by

$$R(L_1) = \frac{R(W_1)}{wg_1} \tag{64}$$

$$= R(y)\frac{R(W_1)}{wg_1 R((1-\alpha)y)} \tag{65}$$

$$= \frac{R(y)}{wg_1}\frac{1}{R(1+S(W))} \tag{66}$$

Substituting this into (63) and rearranging yields (11).

The solution for Ψ follows from (62) which implies $RS(W) = RS(L)^{\Psi}$. □

D Exogenous Skill Bias

D.1 Closed Form Solution

Proof. (Proposition 3)

Define $contrib_L^{poor}$ as the increase in L due to replacing $L_{j,p}$ with $L_{j,r}$, holding $\theta_{j,p}$ fixed:

$$contrib_L^{poor} = \frac{\left[\sum_j (\theta_{j,p}L_{j,r})^{\rho}\right]^{1/\rho}}{\left[\sum_j (\theta_{j,p}L_{j,p})^{\rho}\right]^{1/\rho}} \tag{67}$$

$$= \frac{\theta_{1,p}L_{1,p}\left[R(L_1)^{\rho} + \left(\frac{\theta_{2,p}}{\theta_{1,p}}\frac{L_{2,r}}{L_{1,p}}\right)^{\rho}\right]^{1/\rho}}{\theta_{1,p}L_{1,p}\left[1+S(W)_p\right]^{1/\rho}} \tag{68}$$

$$= \frac{\left[R(L_1)^{\rho} + \left(\frac{\theta_{2,p}}{\theta_{1,p}}\frac{L_{2,p}}{L_{1,p}}R(L_2)\right)^{\rho}\right]^{1/\rho}}{\left[1+S(W)_p\right]^{1/\rho}} \tag{69}$$

$$= \frac{\left[R(L_1)^{\rho} + S(W)_p\left(R(L_2)\right)^{\rho}\right]^{1/\rho}}{\left[1+S(W)_p\right]^{1/\rho}} \tag{70}$$

This uses (40) to replace $S(\theta L)^{\rho}$ with $S(W)$. Pulling out $R(L_1)$ yields

$$contrib_L^{poor} = R(L_1)\left[\frac{1+S(W_p)RS(L)^{\rho}}{1+S(W_p)}\right]^{1/\rho} \tag{71}$$

Replacing $R(L_1)$ using (66) gives

$$contrib_L^{poor} = \frac{R(y)}{wg_1} \frac{1}{R(1+S(W))} \left[\frac{1+S(W_p)RS(L)^\rho}{1+S(W_p)} \right]^{1/\rho} \tag{72}$$

To see that $contrib_L^{poor} \in (R(L_1), R(L_2))$, note that

$$contrib_L^{poor} = R(L_2) \frac{\left[RS(L)^{-\rho} + S(W_p)\right]^{1/\rho}}{\left[1+S(W)_p\right]^{1/\rho}} \tag{73}$$

If $R(L_2) > R(L_1)$, then $R(L_1) < contrib_L^{poor} < R(L_2)$; otherwise $R(L_2) < contrib_L^{poor} < R(L_1)$.

Using rich country skill bias, we have

$$\begin{aligned} contrib_L^{rich} &= \frac{\left[\sum_j (\theta_{j,r} L_{j,r})^\rho\right]^{1/\rho}}{\left[\sum_j (\theta_{j,r} L_{j,p})^\rho\right]^{1/\rho}} && (74) \\ &= \frac{\theta_{1,r} L_{1,r} \left[1 + \left(\frac{\theta_{2,r}}{\theta_{1,r}} \frac{L_{2,r}}{L_{1,r}}\right)^\rho\right]^{1/\rho}}{\theta_{1,r} L_{1,r} \left[R(L_1)^{-\rho} + \left(\frac{\theta_{2,r}}{\theta_{1,r}} \frac{L_{2,r}}{L_{1,r}} \frac{L_{2,p}}{L_{2,r}}\right)^\rho\right]^{1/\rho}} && (75) \\ &= \frac{\left[1+S(W)_r\right]^{1/\rho}}{\left[R(L_1)^{-\rho} + S(W)_r R(L_2)^{-\rho}\right]^{1/\rho}} \end{aligned}$$

Pulling out $R(L_1)$ and replacing it using (66) yields (20). □

E Investment in the Frontier

We consider a model where firms can expend resources to shift the technology frontier outwards, as in Acemoglu (2007). The representative firm solves

$$\max_{k_c, L_{j,c}, \theta_{j,c}, B_c} y_c - q_c k_c - \sum_j p_{j,c} L_{j,c} - C(B_c) \tag{76}$$

subject to (1), (2), and (3), taking factor prices as given. We assume the linear cost function $C(B_c) = b_c B_c$ as in Acemoglu (2007)'s example 1. The firm takes $b_c > 0$ as given. We assume $\omega > 1$ to ensure that optimal skill weights are finite. We normalize all $\kappa_j = 1$.

Compared with the fixed frontier case studied in Section 2, the only change is the endogeneity of B_c. Conditional on its value all quantities and prices are the same as in the baseline model.

If we treat b_c as a parameter, the model has increasing returns to scale. We show that, in this case, $share_L$ is magnified by the factor $\frac{\omega}{\omega-1}$ compared with the baseline model. If b_c scales appropriately with y_c so that the model has constant returns to scale, we show that the development accounting results of the baseline model remain unchanged.

E.1 Reduced Form Labor Aggregator

Proposition 7. *The labor aggregator is given by*

$$\hat{L}_c = \left((1-\alpha) z_c^{1-\alpha} k_c^{\alpha} \omega^{-1} b_c^{-1}\right)^{\frac{1}{\omega+\alpha-1}} \tilde{L}_c^{\frac{\omega}{\omega+\alpha-1}} \tag{77}$$

where $\tilde{L}_c$ is the reduced form labor aggregator with a fixed frontier, given by (9).

Proof. The firm's problem may be written as

$$\max_{k_c, L_{j,c}, \theta_{j,c}} y_c - q_c k_c - \sum_j p_{j,c} L_{j,c} - b_c \sum_j (\kappa_j \theta_{j,c})^{\omega} \tag{78}$$

The firm's first-order condition for $\theta_{j,c}$ is again given by (38), except that now $\lambda_c = b_c$ so that

$$\theta_{j,c}^{\omega-\rho} = X_{j,c} L_{j,c}^{\rho} L_c^{1-\alpha-\rho} \tag{79}$$

where

$$X_{j,c} = \frac{(1-\alpha) z_c^{1-\alpha} k_c^{\alpha}}{b_c \omega \kappa_j^{\omega}} \tag{80}$$

Together with $1 + \rho/(\omega-\rho) = \omega/(\omega-\rho)$ this implies

$$\theta_{1,c} L_{1,c} = X_{1,c}^{\frac{1}{\omega-\rho}} L_{1,c}^{\frac{\omega}{\omega-\rho}} L_c^{\frac{1-\alpha-\rho}{\omega-\rho}} \tag{81}$$

From (55), we have

$$\theta_{1,c} L_{1,c} = L_c \Lambda_c^{-1/\rho} \tag{82}$$

$$= L_c \left(L_1 \kappa_1^{-1} / \tilde{L}_c\right)^{\Psi/\rho} \tag{83}$$

Setting both expressions for $\theta_{1,c} L_{1,c}$ equal and noting that $\Psi/\rho = \omega/(\omega-\rho)$, we have

$$L_c^{1-\frac{1-\alpha-\rho}{\omega-\rho}} = \left(\kappa_1 \tilde{L}_c\right)^{\frac{\omega}{\omega-\rho}} X_{1,c}^{\frac{1}{\omega-\rho}} \tag{84}$$

Since

$$1 - \frac{1-\alpha-\rho}{\omega-\rho} = \frac{\omega+\alpha-1}{\omega-\rho} \tag{85}$$

we have

$$L_c = \left(\kappa_1^{\omega} X_{1,c}\right)^{\frac{1}{\omega+\alpha-1}} \tilde{L}_c^{\frac{\omega}{\omega+\alpha-1}} \tag{86}$$

□

E.2 Reduced Form Production Function

Proposition 8. *The reduced form production function is given by*

$$y_c = \left(k_c^{\alpha}\left(\hat{A}_c z_c \tilde{L}_c\right)^{1-\alpha}\right)^{\frac{\omega}{\omega+\alpha-1}} \tag{87}$$

where $\tilde{L}_c$ is given by (9) and $\hat{A}_c = \left(\frac{1-\alpha}{\omega b_c}\right)^{1/\omega}$ is a constant.

Proof. Substituting the reduced form labor aggregator (86) into the production function, we have

$$y_c = k_c^{\alpha}\left(z_c L_c\right)^{1-\alpha} \tag{88}$$

$$= k_c^{\alpha} z_c^{1-\alpha}\left(\kappa_1^{\omega} X_{1,c}\right)^{\frac{1-\alpha}{\omega+\alpha-1}}\left(\tilde{L}_c\right)^{\frac{\omega(1-\alpha)}{\omega+\alpha-1}} \tag{89}$$

$$= \tilde{A}_c\left(z_c^{1-\alpha} k_c^{\alpha}\right)^{1+\frac{1-\alpha}{\omega+\alpha-1}}\left(\tilde{L}_c\right)^{\frac{\omega(1-\alpha)}{\omega+\alpha-1}} \tag{90}$$

where

$$\tilde{A}_c = \left(\frac{1-\alpha}{\omega b_c}\right)^{\frac{1-\alpha}{\omega+\alpha-1}} \tag{91}$$

collects all constant terms. Then

$$y_c = \left(k_c^{\alpha}\left(\hat{A}_c z_c \tilde{L}_c\right)^{1-\alpha}\right)^{\frac{\omega}{\omega+\alpha-1}} \tag{92}$$

This is true because the exponent on $z_c^{1-\alpha} k_c^{\alpha}$ is

$$1 + \frac{1-\alpha}{\omega+\alpha-1} = \frac{\omega}{\omega+\alpha-1} \tag{93}$$

□

If b_c is fixed, the model has increasing returns to scale due to scale effects. Increasing any factor input or increasing TFP raises the benefits from investing in B_c, but not the cost. The optimal level of B_c increases, amplifying the effect on output. The amplification is governed by the exponent $\frac{\omega}{\omega+\alpha-1}$.

The scale effect is eliminated if the cost of investing in B_c scales appropriately with output. Specifically, if $b_c = y_c^{1-\alpha}$, the production function reverts to the one for the fixed frontier, except that the TFP level z_c is multiplied by a constant. In that case, investment in the frontier has no impact on development accounting.

E.3 Development Accounting

Proposition 9. *The reduced form production function (87) satisfies*

$$y_c = \left[(k_c/y_c)^{\alpha/(1-\alpha)} \hat{A}_c z_c \tilde{L}_c \right]^{\frac{\omega}{\omega-1}} \tag{94}$$

Proof. Write (87) as

$$y_c = \left[(k_c/y_c)^{\alpha} \left(\hat{A}_c z_c \tilde{L}_c \right)^{1-\alpha} \right]^{\frac{\omega}{\omega+\alpha-1}} y_c^{\frac{\alpha\omega}{\omega+\alpha-1}} \tag{95}$$

and note that the exponent on y_c becomes

$$1 - \frac{\alpha\omega}{\omega+\alpha-1} = \frac{\omega+\alpha-1-\alpha\omega}{\omega+\alpha-1} \tag{96}$$

$$= (\alpha-1)\frac{1-\omega}{\omega+\alpha-1} \tag{97}$$

Then

$$y_c = \left[(k_c/y_c)^{\alpha} \left(\hat{A}_c z_c \tilde{L}_c \right)^{1-\alpha} \right]^{\phi} \tag{98}$$

with $\phi = \frac{\omega}{\omega+\alpha-1} \times \frac{\omega+\alpha-1}{(1-\alpha)(\omega-1)}$. Simplify exponents to arrive at equation (94). □

Now the only difference relative to the case where B_c is fixed is the exponent $\omega/(\omega-1)$. To perform development 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 model with a fixed frontier, the contribution of labor inputs to output gaps is amplified by a constant factor, $\omega/(1-\omega)$.

Proposition 10. *The share of cross-country output gaps accounted for by labor inputs is given by*

$$share_L = \frac{\omega}{\omega-1}\frac{\ln R\left(\tilde{L}\right)}{\ln R(y)} \tag{99}$$

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

Proof. Let $A_c = (k_c/y_c)^{\alpha/(1-\alpha)} \hat{A}_c z_c$ collect all country specific terms other than labor inputs.

Then $y_c = \left[A_c \tilde{L}_c\right]^{\frac{\omega}{\omega-1}}$ and

$$\frac{\omega-1}{\omega} \ln R(y) = \ln R(A) + \ln R\left(\tilde{L}\right) \tag{100}$$

This implies (99). Since the calibrated values of $h_{j,c}$ and Ψ do not depend on whether or not B_c is endogenous, the labor aggregator is the same as in the model with fixed B_c. $\square$

F Capital-skill Complementarity

F.1 Preliminaries

This section contains results that are used in subsequent derivations. They hold for endogenous and exogenous skill bias.

F.1.1 Firm first-order conditions

The firm's first-order conditions are:

$$s : \alpha y_c / s_c = q_{s,c} \tag{101}$$

$$e : \frac{\partial y}{\partial L} \frac{\partial L}{\partial Z} Z^{1-\phi} \mu_e^{\phi} e^{\phi-1} = q_e \tag{102}$$

$$L_2 : \frac{\partial y}{\partial L} \frac{\partial L}{\partial Z} Z^{1-\phi} \mu_s^{\phi} L_s^{\phi-1} = p_{2,c} \tag{103}$$

$$L_1 : \frac{\partial y}{\partial L} L^{1-\rho} \theta_{1,c}^{\rho} L_{1,c}^{\rho-1} = p_{1,c} \tag{104}$$

where

$$\frac{\partial y}{\partial L} = (1-\alpha)\, y/L \tag{105}$$

$$\frac{\partial L}{\partial Z} = L^{1-\rho} \theta_2^{\rho} Z^{\rho-1} \tag{106}$$

If there is a technology frontier, we also have

$$\theta_{1,c} : \frac{\partial y}{\partial L} L^{1-\rho} \theta_{1,c}^{\rho-1} L_1^{\rho} = \lambda \omega \kappa_1^{\omega} \theta_{1,c}^{\omega-1} \tag{107}$$

$$\theta_{2,c}: \frac{\partial y}{\partial L} L^{1-\rho} \theta_{2,c}^{\rho-1} Z^{\rho} = \lambda \omega \kappa_2^{\omega} \theta_{2,c}^{\omega-1} \tag{108}$$

which implies that the optimal skill bias ratio is a constant elasticity function of relative inputs:

$$S(\theta)^{\omega-\rho} = S(\kappa)^{-\omega} (Z/L_1)^{\rho} \tag{109}$$

F.1.2 Income ratios and shares

Applying the generic CES expression (33) yields the income ratios of skilled labor to equipment

$$IR_{L_2/e} = \left(\frac{\mu_2 L_2}{\mu_e e}\right)^{\phi} \tag{110}$$

and of Z versus L_1

$$IR_{Z/L_1} = \left(\frac{\theta_2 Z}{\theta_1 L_1}\right)^{\rho} \tag{111}$$

The income ratio of skilled versus unskilled labor is then given by

$$S(W) = IR_{L_2/e} IR_{Z/L_1} = \left(\frac{\mu_2 L_2}{Z}\right)^{\phi} \left(\frac{\theta_2 Z}{\theta_1 L_1}\right)^{\rho} \tag{112}$$

The income share of equipment is given by $IS_e = IS_L IR_{Z/L} IR_{e/Z}$. Again applying the generic CES expressions yields

$$IS_e = (1-\alpha) \left[\frac{\theta_2 Z}{L}\right]^{\rho} \left[\frac{\mu_e e}{Z}\right]^{\phi} \tag{113}$$

F.2 Endogenous Skill Bias

F.2.1 Reduced form labor aggregator

Proof. (Proposition 4)

We may think of the firm as solving its problem in two steps. First, the firm chooses $L_{2,c}/e_c$ to minimize the cost of Z. This is a standard CES cost minimization problem with the solution

$$\left[\frac{L_2}{e}\right]^{1-\phi} = \frac{q_e}{p_2} \left[\frac{\mu_2}{\mu_e}\right]^{\phi} \tag{114}$$

and the unit cost

$$p_Z = \left[(\mu_e q_e)^{\frac{\phi}{1-\phi}} + (\mu_2 p_2)^{\frac{\phi}{1-\phi}}\right]^{\frac{1-\phi}{\phi}} \tag{115}$$

In the second step, the firm solves

$$\max_{L_{1,c},Z_c,\theta_{j,c},s_c} s^{\alpha}\left[z_c L_c\right]^{1-\alpha} - q_s s - p_1 L_1 - p_Z Z \tag{116}$$

subject to the labor aggregator (22) and the frontier constraint (3). This problem has the same structure as the one solved by the firm in the baseline model, except that the firm chooses structures instead of capital and Z instead of L_2. It follows directly that the labor aggregator takes on the same form as in the baseline model. □

F.2.2 Joint Contribution of Labor Inputs and Equipment

We derive a closed form solution for the joint contribution of labor inputs and equipment to cross-country output gaps, $share_{L+e}$.

Proposition 11. *The joint contribution of labor inputs and equipment to cross-country output gaps is given by*

$$share_{L+e} = \underbrace{1 - \frac{\ln\left(wg_1\right)}{\ln R\left(y\right)}}_{base} + \underbrace{\frac{\frac{1}{\Psi}\ln R\left(1 + IR_{Z/L_1}\right) - \ln R\left(1 + S(W)\right)}{\ln R\left(y\right)}}_{amplification} \tag{117}$$

where the reduced form curvature is given by

$$\Psi = \frac{\ln R\left(IR_{Z/L_1}\right)}{\ln R\left(Z/L_1\right)} \tag{118}$$

and the curvature of the Z aggregator is given by

$$\phi = \frac{\ln R\left(IR_{L_2/e}\right)}{\ln R\left(L_2/e\right)} \tag{119}$$

In terms of observable data moments, the reduced form curvature may be written as

$$\Psi = \frac{\ln RS\left(W\right) + \ln R\left(1 + IR_{e/L_2}\right)}{\ln RS\left(L\right) + \frac{1}{\phi}\ln R\left(1 + IR_{e/L_2}\right)} \tag{120}$$

Throughout, $IR_{a/b,}$ denotes the ratio of incomes received by inputs a and b.

Proof. The labor aggregator may be written as

$$L_c = L_{1,c}\left[1 + \left(Z_c/L_{1,c}\right)^{\Psi}\right]^{1/\Psi} \tag{121}$$

Applying the generic CES expressions for income shares and income ratios to the reduced form labor aggregator yields

$$\left(\frac{Z}{L_1}\frac{\kappa_1}{\kappa_2}\right)^{\Psi} = S(W)\left(1+IR_{e/L_2}\right) \tag{122}$$

$$= IR_{Z/L_1} \tag{123}$$

where κ_j may be normalized to one. Using (123) we have

$$L_c = L_{1,c}\left[1+S(W_c)\left(1+IR_{e/L_2,c}\right)\right]^{1/\Psi} \tag{124}$$

Taking logarithms and replacing $R(L_1)$ using (11) yields (117).

The solution for Ψ is obtained by taking the rich-to-poor country ratio of (123) in logarithms which yields

$$\Psi = \frac{\ln R\left(S(W)\left(1+IR_{e/L_2}\right)\right)}{\ln R\left(Z/L_1\right)} \tag{125}$$

$$= \frac{\ln R\left(IR_{Z/L_1}\right)}{\ln R\left(Z/L_1\right)} \tag{126}$$

where $R(Z)$ follows from

$$R(Z) = R\left(Z/\left(\mu_e e\right)\right)R(e) \tag{127}$$

$$= R\left(\left[1+IR_{s/e}\right]^{1/\phi}\right)R(e) \tag{128}$$

$$= R\left(\left[1+IR_{e/L_2}\right]^{1/\phi}\right)R(L_2) \tag{129}$$

The solution for Ψ can be expressed in a form that is closer to the baseline model. From (123), we have

$$\frac{Z}{L_1} = \frac{Z}{L_2}S(L) = S(L)\left(1+IR_{e/L_2}\right)^{1/\phi} \tag{130}$$

Therefore

$$\Psi = \frac{\ln R\left(S(W)\right) + \ln R\left(1+IR_{e/L_2}\right)}{\ln RS(L) + \frac{1}{\phi}\ln R\left(1+IR_{e/L_2}\right)} \tag{131}$$

□

The data moments used in the calibration imply that skilled labor and equipment are complements ($\phi < 0$).[20] This is consistent with U.S. time series evidence (see Krusell et al. 2000).

Since we assume that the income share of equipment is the same in rich and poor countries,

[20]The numerator in (119) is positive because $R(IS_e) = 1$ and $R(IS_{L_2}) > 1$. The denominator is negative because equipment stocks vary across countries more than labor inputs. Hence $\phi < 0$.

IR_{e/L_2} is lower in rich compared with poor countries. Together with $\phi < 0$, it follows that the long-run elasticity of substitution between skilled and unskilled labor is lower than in the baseline model. This increases the amplification term.

The expression for $share_{L+e}$ is similar in structure to the baseline model's (12). The base term is the same, again reflecting the contribution of human capital in a single skill model. The amplification term now depends on the ratio of incomes received by Z (by skilled labor and equipment jointly) to unskilled labor. When equipment is "unimportant," so that $IR_{e/L_2} \approx 0$, the values of Ψ and $R(L)$ approach those of the baseline model.

F.3 Exogenous Skill Bias

Our final model treats variation in skill bias $\theta_{j,c}$ across countries as exogenous. Except for dropping the technology frontier, the model is identical to the one described in Section 5.1.

F.3.1 Development Accounting

We define the contribution of labor inputs to cross-country output variation as the change in steady state output that results from increasing $L_{j,p}$ to $L_{j,r}$, holding capital rental prices and skill bias $\theta_{j,c}$ constant. It follows that $share_L$ depends on the fixed levels of q_e (but not on q_s) and now also on those of $\theta_{j,c}$. We consider two cases:

1. $share_L^{poor}$ fixes skill bias and q_e at poor country levels. This corresponds to increasing labor inputs in the poor country.
2. $share_L^{rich}$ fixes skill bias and q_e at rich country levels. This corresponds to reducing labor inputs in the rich country.

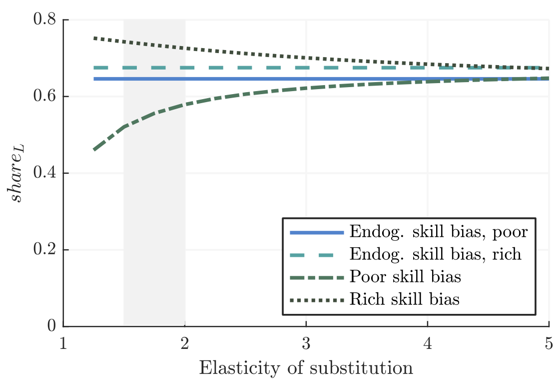
Relative to the model with the technology frontier, one additional parameter needs to be calibrated because counterfactual output depends on the values of ρ and ω, not only on the reduced form curvature Ψ. The development accounting results therefore require fixed values of ρ.

F.3.2 Quantitative Results

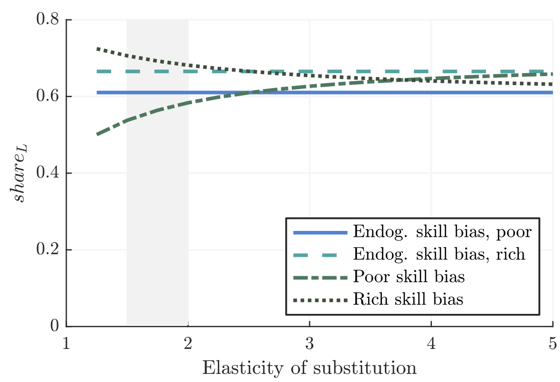
Figure 1 provides a compact visual summary of the results. When poor country $\theta_{j,c}$ and q_e are used, the results are very similar to the baseline model. $share_L^{poor}$ is smaller than $share_L$ when $\rho < \Psi$. It increases with the elasticity of substitution and the skill cutoff. Values below 0.5 are associated with very large cross-country differences in relative skill bias (at least factor 10^5).

Figure 1: $share_L$: Capital-skill Complementarity

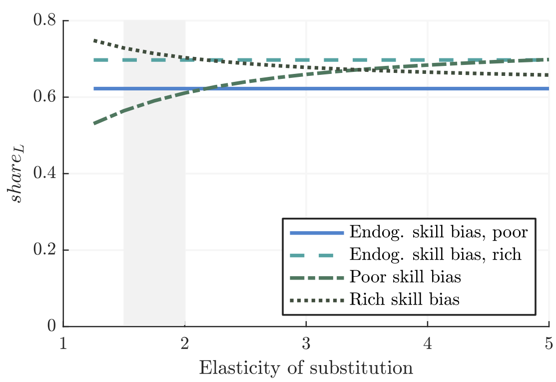
(a) SHS Skill Cutoff



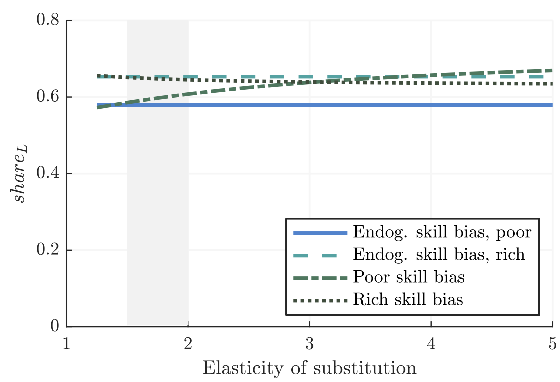
(b) HSG Skill Cutoff



(c) SC Skill Cutoff



(d) CG Skill Cutoff



With rich country $\theta_{j,c}$ and q_e, $share_L^{rich}$ is higher than $share_L$ when $\rho < \Psi$. Its value decreases with the elasticity of substitution and the skill cutoff. For conventional values of the elasticity, we find $share_L^{rich}$ between 0.64 and 0.74 (compared with 0.59 to 0.74 in the baseline model).