Skill Bias and Development: Evidence from Trade and Industry Data

Hannes Malmberg*

June 20, 2022

Abstract

Why is economic development associated with a high demand for skilled labor? This paper takes a disaggregate perspective, using revealed comparative advantage to study industry-level productivity differences. I document that beyond having a high demand for skilled labor, rich countries also have a high productivity in skill-intensive industries and skill-intensive industry segments. These findings suggest a theory where some factor in poor countries systematically impairs the productivity of skilled labor. Drawing on the literature on firms and development, I propose one such theory where skill bias arises naturally from poor countries having high barriers to the accumulation of organizational capital.

*This paper was previously circulated under the name "Human Capital and Development Accounting Revisited". I am grateful to Anmol Bhandari, Timo Boppart, Axel Gottfries, Karl Harmenberg, Per Krusell, Ellen McGrattan, Erik Öberg, Torsten Persson, and Todd Schoellman for extensive comments on the project. Moreover, I am grateful for helpful discussions with Francesco Caselli, John Hassler, Chad Jones, Loukas Karabarbounis, David Lagakos, Bo Malmberg, Kurt Mitman, Esteban Rossi-Hansberg, Lena Sommestad, David Strömberg, Daniel Trefler, and David Weil. I thank Patrick Kiernan for excellent research assistance. The input from the editor and four anonymous referees are gratefully acknowledged.

1 Introduction

Rich economies use skilled labor intensely. Their enterprises rely on managers, accountants and lawyers; their manufacturing processes rely on engineers; their farms rely on agronomists. These facts have led economists to posit that skill is intrinsically connected to economic development (Kuznets, 1966; Schultz, 1982; Lucas, 1988).

In the development accounting literature, the role of skilled labor in development has been explored through the lens of aggregate production functions (Caselli and Coleman, 2006; Hendricks and Schoellman, 2019; Rossi, 2022; Bils, Kaymak, and Wu, 2022). It has been noted that rich countries have a much higher share of skilled workers than poor countries. At the same time, they have a similar skilled wage premium. Through the lens of an aggregate production function, these facts suggest a large demand shifter for skilled labor in rich countries. Such a demand shifter is typically called a "skill-biased technology difference" or a "skill-biased efficiency difference", but these terms should be interpreted neutrally. As with their older cousin total factor productivity, they are primarily a measure of our ignorance.[1]

What economic mechanisms underpin the skill bias of development? To make progress on this question, I take a disaggregated view and analyze how skill intensity relates to productivity differences both across and within industries. To estimate relative productivity differences, I rely on industry production and trade data and the method of revealed comparative advantage. Using a gravity trade model with multiple industries and within-industry heterogeneity between traders and non-traders, I back out unit costs across the world, and I estimate productivity differences by correcting these unit costs for variations in factor prices.[2]

My main empirical findings are two disaggregate analogues to the aggregate skill bias previously identified in the literature. First, there are *skill-biased development differences between industries*: as we go from poor to rich countries, productivity rises faster in more skill-intensive industries. Second, there are skill-biased development differences *within industries*: as we go from poor to rich countries, productivity rises faster in the more skill-intensive tradable segments of industries.

The between-industry result reflects that rich countries specialize in skill-intensive exports. The degree of specialization is strong. As we go from the poorest to the richest countries, the share of merchandise

[1] Formally, the literature typically posits a constant elasticity of substitution aggregator of labor services of different types. Using elasticity estimates from, for example, Katz and Murphy (1992) and Ciccone and Peri (2005), skill-bias can be calculated as the residual demand shifter needed to rationalize price and quantity data given the assumed elasticity. In terms of interpretation, Rossi (2022) has shown that simple skill-biased quality differences struggle to explain the results, leaving us with a large residual. Another strand of the literature argues that large estimated demand shifters reflect that extant elasticity estimates understate true long-run elasticities (Bils et al., 2022; Hendricks and Schoellman, 2019).

[2] My method is most closely related to that used in Levchenko and Zhang (2016), who rely on a trade model in the style of Eaton and Kortum (2002) to back out productivities across countries.

exports that are in the most skill intensive industries goes from 5% to more than 50%. The relationship between income and skill-intensification of exports is quite tight, with an R^2-value of 0.44 when regressing the logit share of skill intensive exports on log GDP per worker. Through the lens of a gravity model, this specialization implies that rich countries have low relative unit costs in skill-intensive industries.

My within-industry results reflect two facts first observed by Waugh (2010): compared to those of rich countries, the products of poor countries are

1. ...uncompetitive abroad in the sense of making up a relatively low share of other countries' absorption,

2. ...competitive locally in the sense of making up a similar share of domestic absorption.

In a standard gravity framework, these facts are not straightforward to generate if relative unit costs between tradables and non-tradables are constant across countries, since low exports then should go together with high import penetration. This follows because low exports imply high unit costs, which make your products unattractive to foreign and domestic buyers alike. In my setup, facts 1 and 2 imply that poor countries have low relative unit cost in non-tradables, which lets them combine low export success with low import penetration.[3]

The trade analysis provides unit cost estimates at observed factor prices. To obtain my skill bias results, I adjust for factor prices – using a definition of skill bias in terms of unit costs being log submodular in skill intensity and development – and show that rich countries' low unit costs in skill-intensive production cannot be explained by low skilled wage premia. This finding reflects the well-known fact that skilled wage premia vary little with development (Caselli, Ponticelli, and Rossi, 2014).

In terms of interpretation, many existing theories of skill biased development fail to predict my findings. First, my findings of relative unit cost differences are not predicted by existing demand-based theories of skill-biased development. These theories instead focus on development increasing the expenditure share of skill-intensive products through income effects, and include models where income effects expand the non-agricultural sector (Caselli and Coleman II, 2001), expand skill-intensive sectors more generally (Comin, Danieli, and Mestieri, 2020), expand the the demand for high-quality goods (Jaimovich, Rebelo, Wong, and Zhang, 2020), or expand the production of goods where (skill-intensive) market production has a comparative advantage over home production (Buera and Kaboski, 2012). And while theories where the relative

[3]Waugh (2010) explains his findings with poor countries having a high trade costs specific to exports. In the appendix, I show that my two-technology formulation is isomorphic to such export-specific trade costs, with a comparative advantage in non-tradables being equivalent to higher export-specific costs. I also discuss whether high literal trade costs are consistent with the data.

productivity of skill-intensive sectors rises with income (Buera, Kaboski, Rogerson, and Vizcaino, 2022) successfully predict my between-industry results, they do not predict my within-industry results.[4]

Instead, my findings suggest that some factor in poor countries consistently raises the unit costs of skill-intensive production, both between and within industries. Formally, we need a mechanism that in reduced form looks like a skill-specific efficiency shifter, but not just an aggregate demand shifter as in, e.g., Caselli and Coleman (2006), but a shifter *inside* individual production functions that shifts unit costs on a disaggregate level. Essentially, we are interested in mechanisms taking the form of skilled workers producing some input "X" across various industries, with X being expensive in poor countries.

I show that a promising approach is to draw on findings from the literature on firms in development. In particular, this literature has shown that poor countries have barriers to the growth and operation of large-scale business enterprises (Hsieh and Klenow, 2014b; Hjort, Malmberg, and Schoellman, 2021; Akcigit, Alp, and Peters, 2021). Moreover, I note that this connects to skill intensification because a key task of skilled white-collar workers such as managers, engineers and business professionals is to build and maintain such business enterprises. Combining these two observations, I develop a parsimonious model of skill biased development based on the idea of organizational capital. In the model, the main activity of skilled workers is to produce and maintain organizational capital, and poor countries have high barriers to the accumulation of such capital, both in terms of a low productivity in accumulating it, and in terms of bad property rights increasing the likelihood that the capital is lost or taxed away over time.

I show that in the model, barriers to organizational capital accumulation are isomorphic to factor-augmenting productivity penalties on skilled labor. This reflects that 1) low productivity in accumulation makes skilled workers produce little organizational capital, and 2) bad property rights reduce the value of accumulation by increasing the effective depreciation rate of organizational capital. Given that accumulation barriers act like factor-augmenting shifters, they correctly predict relative cost movements: poor countries have high production costs in skill-intensive production processes, both between and within industries.

Related literature. My paper most directly contributes to the literature on skill-biased development, which studies the difference in skilled labor demand across rich and poor countries. On the cross-national level, this literature goes back to Caselli and Coleman (2006) with later contributions including Jones (2014a), Jones (2014b), Caselli (2016), Hendricks and Schoellman (2019), Rossi (2022), and Bils et al. (2022). I complement the literature's aggregate results by providing disaggregate evidence on skill-biased productivity

[4] One theory that does fit the results is that skilled labor have a higher relative quality in rich countries (Jones, 2014a). The challenge for human capital quality-based theories of skill bias, at least when quality takes a factor-augmenting form, has been to reconcile them with relatively modest differences in gains to migration of groups with different education (Rossi, 2022).

differences both across and within industries, and by showing which types of theories of skill bias that are favored by these patterrns.

In the time series dimension, skill bias has been studied under the label of skill-biased technical change. One strand of this literature focuses on inequality trends, starting with the seminal worrk of Katz and Murphy (1992), with later contributions including Berman, Bound, and Machin (1998) and Acemoglu and Autor (2011). This literature also includes non-technological explanations of skill-biased demand changes such as non-homothetic demand for skill-intensive industry outputs (Comin et al., 2020) or for higher quality goods (Jaimovich et al., 2020). The time series literature also has a strand which is closer to the structural change literature in development, and argues that skill-biased demand change is an inherent feature of the growth experience. Goldin and Katz (2008) argue that technology has increased the secular demand for skill over the last century, Buera and Kaboski (2012) develops a model where non-homothetic demand for marketable services rises skill demand, and Buera et al. (2022) show that structural change across sectors favors skilled labor.

In terms of combining skill bias with industry and trade data, the closest paper to mine is Morrow and Trefler (2022). They also focus on within-industry factor efficiency shifters, but use a different identification method, where the shifters are identified from cross-country, within-industry, variation in factor prices and quantities, assuming that all within-industry elasticities of substitution are similar to the aggregate elasticity estimates from Katz and Murphy (1992) and Acemoglu and Autor (2011).

My between-industry results builds on the quantitative Ricardian tradition, started by Eaton and Kortum (2002) and extended to multi-industry setups in Chor, 2010, Costinot, Donaldson, and Komunjer (2011), Caliendo and Parro (2015), and Levchenko and Zhang (2016). My unit cost estimation is similar to the one used in Levchenko and Zhang (2016). To define skill bias and correct for factor price differences, I use methods from the literature of complementarity (Topkis, 1998). To establish skill bias, I regress estimated relative unit costs on an interaction of log GDP per worker and skill intensity, similar to the specification used in the literature on the determinants of comparative advantage (Nunn, 2007; Cuñat and Melitz, 2012; Manova, 2013), reviewed in Nunn and Trefler (2015).[5]

My within-industry section draws on the extensive literature on firm heterogeneity and trade (Melitz,

[5] Compared to the literature on institutions and comparative advantage, one difference is that my paper interacts an industry property with log GDP per capita instead of with a specific country properties like institutional quality. My estimates thus do not capture an effect size of a country characteristic, but rather a systematic differrence between rich and poor countries. In that sense, my analysis is closer to the literature that studies differences in trade patterns between rich and poor countries. This includes papers showing that rich countries specialized in capital- and skill-intensive goods (Romalis, 2004), high-quality goods (Fajgelbaum, Grossman, and Helpman, 2011; Hallak and Schott, 2011; Hummels and Klenow, 2005), goods with a high income elasticity (Fieler, 2011), or goods with a high "complexity" (defined in different ways) (Atkin, Costinot, and Fukui, 2021; Hidalgo and Hausmann, 2009; Schetter, 2019).

2003; Melitz and Redding, 2015), in particular on the result that exporters are more skill-intensive than non-exporters (Alvarez and López, 2005; Bernard and Jensen, 2007; Verhoogen, 2008). With such heterogeneity, skill-biased productivity differences are isomorphic to export-specific trade costs, providing an explanation for the findings in Waugh (2010), which shows that price and bilateral trade data are consistent with poor countries having such export-specific costs.

The last part of the paper draws on the literature on firms in developing countries. This literature has documented that firms and establishments in poor countries are small (Bento and Restuccia, 2021; Tybout, 2000) and grow slowly (Hsieh and Klenow, 2014b). The lack of large firms has been attributed to factors such as size-dependent regulations (Guner, Ventura, and Xu, 2008), financial frictions (Buera, Kaboski, and Shin, 2011), lack of specialized human capital (Hjort et al., 2021), and restrictions on the ability to delegate (Akcigit et al., 2021).

I model these constraining factors as high barriers to the accumulation of a durable asset, which I call organizational capital. In this, I draw on the intangible and organizational capital literature which argues that high firm capabilities should be construed as assets resulting from forward-looking investments (Prescott and Visscher, 1980; Atkeson and Kehoe, 2005; Bhandari and McGrattan, 2021). In assuming that organizational capital is built and maintained by skilled labor, I connect to economic history work that links the emergence of large-scale firms to the rise of the management class (Chandler Jr, 1977), as well as later work which links skill intensity with intangible capital intensity (Bresnahan, Brynjolfsson, and Hitt, 2002). By positing (de facto) capital taxes as a barrier to organizational capital accumulation, I also connect with the literature highlighting poor property rights as a barrier to development (Acemoglu, Johnson, and Robinson, 2001; North, 1987).

2 Trade model

To guide the analysis, I use a general equilibrium model with multiple industries and international trade. The model is a multi-industry version of the stochastic productivity framework in Eaton and Kortum (2002), and is similar to other gravity trade models with industries (Chor, 2010; Levchenko and Zhang, 2016; Morrow and Trefler, 2022). To allow for heterogeneity between traders and non-traders, I assume that each industry has a non-traded and a traded segment that use different production technologies, where the traded technology has a higher skill intensity.[6]

[6] A large literature has documented differences between exporters and non-exporters, with exporters being more intensive in white-collar labor (Bernard and Jensen, 2007; Melitz, 2003; Melitz and Redding, 2015). My setup where heterogeneity captures

Below I present the model assumptions on consumer demand, production, and trade costs. These features jointly imply a gravity equation that will be used in the empirical work. The remaining model elements and the equilibrium definition are presented in the appendix.

Model. There are I countries and K industries. Each country is populated by a representative agent with preferences

$$U(Q_1^i, \ldots, Q_K^i) \tag{1}$$

over industry composites Q_k^i, where U is concave but otherwise unrestricted. Each industry composite is a CES aggregate of a unit interval of varieties $Q_k^i = \left(\int_0^1 \tilde{q}_k^i(\omega_k)^{\frac{\sigma-1}{\sigma}} d\omega_k\right)^{\frac{\sigma}{\sigma-1}}$,with a corresponding unit cost function

$$P_k^i = C_k^i(\{\tilde{p}_k^i(\omega_k)\}_{\omega^k \in [0,1]}) = \left(\int_0^1 \tilde{p}_k^i(\omega_k)^{1-\sigma} d\omega_k\right)^{\frac{1}{1-\sigma}}, \tag{2}$$

where $\tilde{p}_k^i(\omega_k)$ is the purchasing price of variety ω_k in i (the notation with $\sim$ indicates that variables vary on a variety level).

Each variety can be produced using two different production technologies, denoted by NT ("non-traded") and T ("traded"). The output of technology T in k can be shipped from country j to country i at an iceberg cost $\tau_k^{ij} \geq 1$, whereas the output of technology NT is not tradable across countries. The unit costs of the technologies have a deterministic and a stochastic component:

$$\tilde{c}_{(n,NT)}^i(\omega) = \frac{c_{(n,NT)}^i(w_u^i, w_s^i)}{\tilde{z}_{(n,NT)}^i(\omega_k)}, \tag{3}$$

$$\tilde{c}_{(n,T)}^i(\omega) = \frac{c_{(n,T)}^i(w_u^i, w_s^i)}{\tilde{z}_{(n,T)}^i(\omega_k)}, \tag{4}$$

where the cost functions have unskilled and skilled wages w_u^i and w_s^i as arguments.[7] The stochastic components $\tilde{z}_{(n,NT)}^i(\omega)$ and $\tilde{z}_{(n,T)}^i(\omega)$ have a Frechét distribution, $\mathbb{P}(\tilde{z}_{(n,NT)}^i(\omega) \leq z) = e^{-a_{NT} G_{NT}(L^i) z^{-\theta}}$ and $\mathbb{P}(\tilde{z}_{(n,T)}^i(\omega) \leq z) = e^{-a_T G_T(L^i) z^{-\theta}}$, with draws being independent across countries, varieties, and technologies. The terms $a_{NT} G_{NT}$ and $a_T G_T$ are location parameters of the Frechét distributions that depend on

industry segments with distinct technologies is in the spirit of Holmes and Stevens (2014), who show that one way to understand plant heterogeneity by exports status is in terms of there being distinct production technologies where non-trading firms produce a different, customized, good with high trading costs (Holmes and Stevens do not address white-collar labor intensity, but their theory is consistent with other trade finding in the trade literature if their large, standardized, firms have a greater need for white-collar workers in coordinating production and sales). See also Burstein and Vogel (2017) for a paper where the same industry has multiple technologies that vary in skill intensity.

[7] The restriction to unskilled and skilled labor is not needed until I correct for factor price differences in section 5. In particular, all results about unit cost variations across countries would be unaffected if I allowed for an arbitrary set of factors.

country size.[8]

Gravity equation. The purchasing price of a variety is the minimum across the local non-traded technology and the traded technology of all potential exporters

$$\tilde{p}^i_{k,t}(\omega_k) = \min\left\{\tilde{c}^i_{(k,NT)}(\omega_k), \min_j \tau^{ij}_k \tilde{c}^j_{(k,T)}(\omega_k)\right\}, \tag{5}$$

which can be substituted into (2) to obtain a cost function for the industry composite in terms of variety costs. Cost minimization implies the following gravity equation (see the appendix for derivation):

$$X^{ij}_k = \begin{cases} X^i_k \frac{a_{k,T} G_T(L^j)\left(c^j_{k,T}\tau^{ij}_k\right)^{-\theta}}{\Phi^i_k} & \text{if } i \neq j \\ X^i_k \frac{a_{k,NT} G_{NT}(L^i)\left(c^i_{k,NT}\right)^{-\theta} + a_{k,T} G_T(L^i)\left(c^i_{k,NT}\right)^{-\theta}}{\Phi^i_k} & \text{if } i = j, \end{cases} \tag{6}$$

where X^{ij}_k is the value of trade flows from j to i in industry k, X^i_k is the total absorption of industry k in country i, and $\Phi^i_k \equiv a_{k,NT} G_{NT}(L^i)\left(c^i_{k,NT}\right)^{-\theta} + \sum_j a_{k,NT} G_{NT}(L^j)\left(c^j_{k,NT}\right)^{-\theta}$ is a multilateral resistance term.

We note that for $i \neq j$, the gravity equation has a standard form. For $i = j$, there is an extra term capturing the possibility of buying goods produced by the non-traded technology. This term reduces imports and breaks the mechanic link between export success (captured by $c^j_{k,T}$) and import penetration. I return this effect in section 4.

3 Unit cost variations across industries

I use the gravity relationship (6) to estimate industry unit costs for the tradable technology. The main finding is that rich countries have low relative unit costs in skill-intensive industries. This reflects that their exports are highly concentrated in those industries, which, through the lens of the gravity equation (6), implies low unit costs.

While it takes some econometric legwork to make this reasoning formal, the basic insight is well-captured in figure 1 which relates country income to the share of exports coming from industries in the highest and

[8] To relax the independence assumption, it is possible to introduce a correlation between draws of the same technology in different countries, in line with the methodology suggested in Lind and Ramondo (2018). In this case, the reduced form aggregator becomes a nested CES, where the non-traded vs traded elasticity is lower than the elasticity between tradables from different countries. The terms G_{NT} and G_T are introduced since any stochastic productivity model needs some scaling of the number of draws with size to avoid large countries being systematically assigned low deterministic unit costs to explain their higher absolute export levels. The functional specification with G_{NT} and G_T introduces this scaling in a flexible way.

lowest quartile of skill intensity. Panel A displays the export share in the highest quartile of skill intensity, and shows that rich countries have dramatically higher export shares in skill-intensive industries, with the share going from about 5% to more than 50% of exports. Panel B shows that rich countries also export less from low-skill intensive industries. The pattern is most clear between middle income and high income countries, when the share of less skill-intensive imports falls from approximately 25% to below 10%.

3.1 Estimation equation

Let log iceberg trade costs satisfy

$$\log \tau_k^{ij} = \tau^{ij} + \tau_k^i + \sum_m \beta_{k,m} d_m^{ij} + \varepsilon_k^{ij}, \tag{7}$$

where τ^{ij} is a bilateral fixed effect, τ_k^i is a destination-industry fixed effect, and d_m^{ij} are a collection of trade cost proxies, and where $\mathbb{E}[\varepsilon_k^{ij}] = \mathbb{E}[\varepsilon_k^{ij} d_m^{ij}] = 0$ for for each combination of i, j, and k. Throughout, I also assume that the trade cost proxies are not multicollinear with the fixed effects.[9]

Proposition 1. *Suppose that trade costs satisfy (7) and consider the trade regression*

$$\log X_k^{ij} = \delta^{ij} + \alpha_k + \mu_k^i + \gamma_k^j + \sum_m \beta_{k,m} d_m^{ij} + \epsilon_k^{ij}. \tag{8}$$

Given the normalization of fixed effects $\sum_k \alpha_k = \sum_j \gamma_k^j = \sum_k \gamma_k^j = \sum_k \mu_k^i = \sum_i \mu_k^i = 0$, *the estimated origin-industry effects* $\hat{\gamma}_k^j$ *are unbiased estimates of relative unit production costs up to the trade elasticity:*

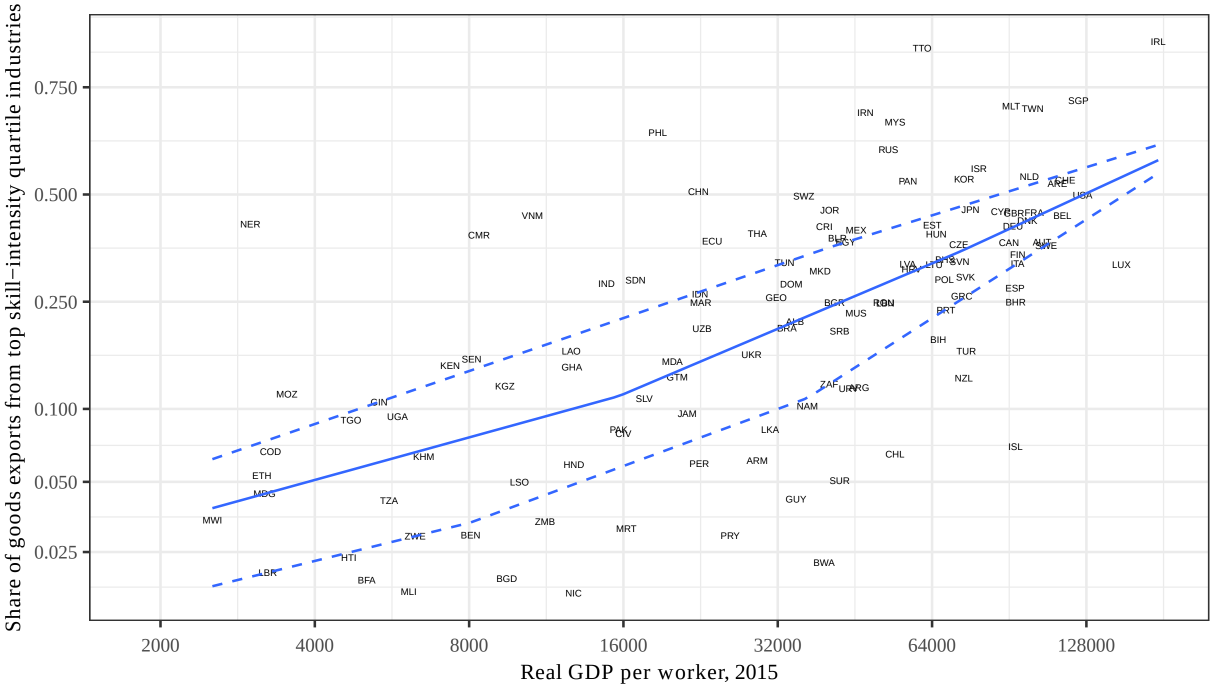
$$-\frac{\mathbb{E}[\hat{\gamma}_k^j]}{\theta} = \log\left(\frac{c_k^j/\bar{c}^j}{\bar{c}_k/\bar{c}}\right), \tag{9}$$

where $\bar{c}^j$ *is the geometric average of unit costs across* k *for a fixed* j, $\bar{c}_k$ *is the geometric average of unit costs across* j *for a fixed* k, *and* $\bar{c}$ *is the geometric average of unit costs across both* k *and* j.

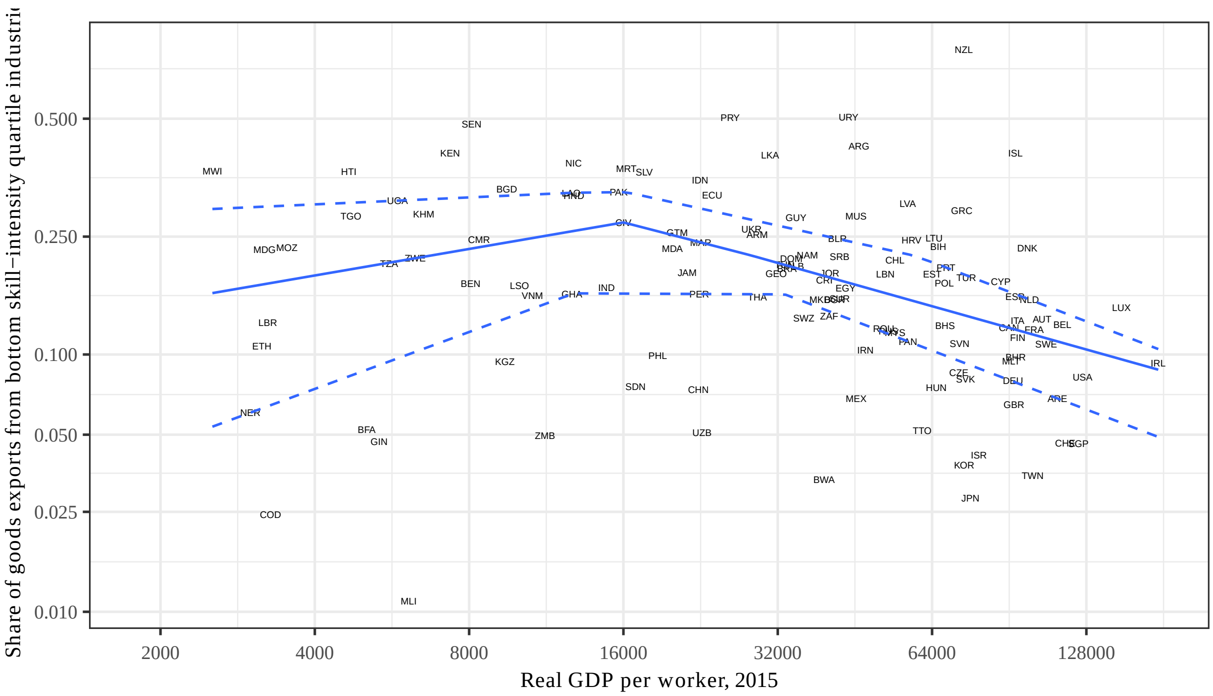
Proof. See the appendix. □

The proposition shows how to estimate comparative advantage under suitable assumptions about trade costs. The intuition is straightforward: trade flows are log-additive in destination demand, bilateral trade costs, origin population, and origin unit costs. Thus, the origin-industry fixed effect γ_k^j captures a mixture

[9]The key assumption is that within a bilateral pair, the shocks to the source country's trade costs have mean zero, so that low exports can be interpreted as reflecting high costs. In the discussion section below and the appendix, I discuss potential identification threats and conduct a number of robustness checks on my main results.



(a) Share of exports in highest quartile of skill intensity



(b) Share of exports in lowest quartile of skill intensity

Figure 1: Share of exports by skill intensity

Note: The skill quartiles are based on white-collar payroll share, calculated using the 2015 Occupational Employment Survey's information about number and average pay of workers by industry and occupation, concorded to ISIC rev 4 industry code from NAICS 2012 and to ISCO-08 from SOC-10 occupational codes. White-collar workers are defined as those with a 1-digit ISCO-08 code of 1 to 4. GDP per worker is real expenditure based GDP per worker from the Penn World Table 9.1.

of origin size L^j and unit costs provided that we control for trade costs and destination demand (via the destination fixed effect μ_k^i). The population effect is netted out by the normalization of fixed effects, leaving a term capturing the comparative advantage.

3.2 Data

Regression (8) requires data on bilateral trade flow values X_k^{ij}, bilateral measures of trade frictions d_m^{ij}, and a trade elasticity estimate θ. To relate unit costs to skill intensity, I also need to measure the skill-intensity of different industries. Unless stated otherwise, all data is from 2015.

For trade flows, I use the BACI data set which is compiled by CEPII and based on COMTRADE (Gaulier and Zignago, 2010). For each country-destination pair, it reports export values in merchandise industries at the HS 2007 six-digit industry level. The database contains 189 countries, and I include all countries with more than \$1bn in total exports and less than 50% of exports in extractive industries, leaving 123 countries in the final analysis. For the bilateral trade costs variables, I use the CEPII gravity dataset (Head and Mayer, 2014). GDP per worker data is taken from the Penn World Table 9.1 (Feenstra, Inklaar, and Timmer, 2019).

To measure skill-intensity, I use the payroll share of white-collar workers in the US. I measure this share using the Occupational Employment Survey (OES) which records average wages and number of workers by industry on a detailed occupational level. Payroll quantities are obtained by multiplying employment numbers with average wages, and the skill share is defined as the payroll share of workers with an ISCO occupation code 1 to 4. Since the OES does not cover agricultural workers, I use the ACS to obtain the skilled payroll share in agricultural industries.

For the regression, I use ISIC revision 4 coding on the four digit level. The trade data is recorded using HS 2007 six-digit codes. The Occupational Employment Survey data is reported in 2012 NAICS. All factor share and trade data are converted between coding schemes using a concordance procedure described in appendix C.1.

The value of the trade elasticity θ is taken from the literature. The estimate is chosen to reflect the long-run elasticity between different foreign varieties in the same industry. This choice reflects the nature of my regression. The regression is run between countries in different parts of the world income distribution, and aims at capturing persistent cross-country differences. Furthermore, the regression explains a source country's exports conditioned on the total industry imports of a destination country. Thus, the relevant elasticity is the long-run elasticity between different foreign varieties. The higher the θ, the smaller are the imputed relative price differences, since it reduces the unit cost differences required to explain the trade patterns. The baseline analysis uses a trade elasticity of $\theta = 6$. I also consider the effect of varying the trade elasticity between $\theta = 4$ and $\theta = 8$. This reflects a range of estimates found in the literature. Simonovska and Waugh (2014) find $\theta = 4$, Costinot et al. (2011), $\theta = 6.2$, and Eaton and Kortum (2002), $\theta = 8.2$. In the robustness checks, I also analyze the case with heterogeneous trade elasticities across industries.

Data descriptives. Figure 1, displayed earlier, shows the relationship between income levels and the skill-intensity of a country's exports. Panel A shows the share of exports that come from the top quartile of industries by skill-intensity, having a skilled payroll share above 58%, while panel B shows the share of exports that come from the bottom quartile of industries by skill-intensity, having a skilled payroll share below 36%. The plots include a fitted quantile regression.

Figure 2 shows the distribution of the skilled payroll share across the different merchandise industries. The skilled payroll share is defined as the share of total industry payroll going to skilled workers. The bulk of industries are in the range between 25% and 75%. Table 1 and 2 show the 10 industries with the highest and lowest skilled payroll shares. We see that the most skill-intensive industries are "complex" industries such as the manufacturing of computers, watches, and optical equipment. Extraction of crude petroleum and natural gas are also skill-intensive. The unskill-intensive industries are concentrated in logging, apparel, and food processing.

Table 1: Top 10 industries by skill-intensity

ISIC code	Industry	White-collar payroll share
2620	Computers and peripheral equipment	0.98
2652	Watches and clocks	0.89
2660	Irradiation, electromedical and electrotherapeutic equipment	0.89
2680	Magnetic and optical media	0.89
2630	Communication equipment	0.87
0610	Extraction of crude petroleum	0.83
0620	Extraction of natural gas	0.83
2610	Electronic components and boards	0.80
2651	Measuring, testing, navigating and control equipment	0.77
2670	Optical instruments and photographic equipment	0.76

The skill shares is the white-collar payroll share, calculated using the 2015 Occupational Employment Survey's information about number and average pay of workers by industry and occupation, concorded to ISIC rev 4 industry code from NAICS 2012 and to ISCO-08 from SOC-10 occupational codes. White-collar workers are defined as those with a 1-digit ISCO-08 code of 1 to 4.

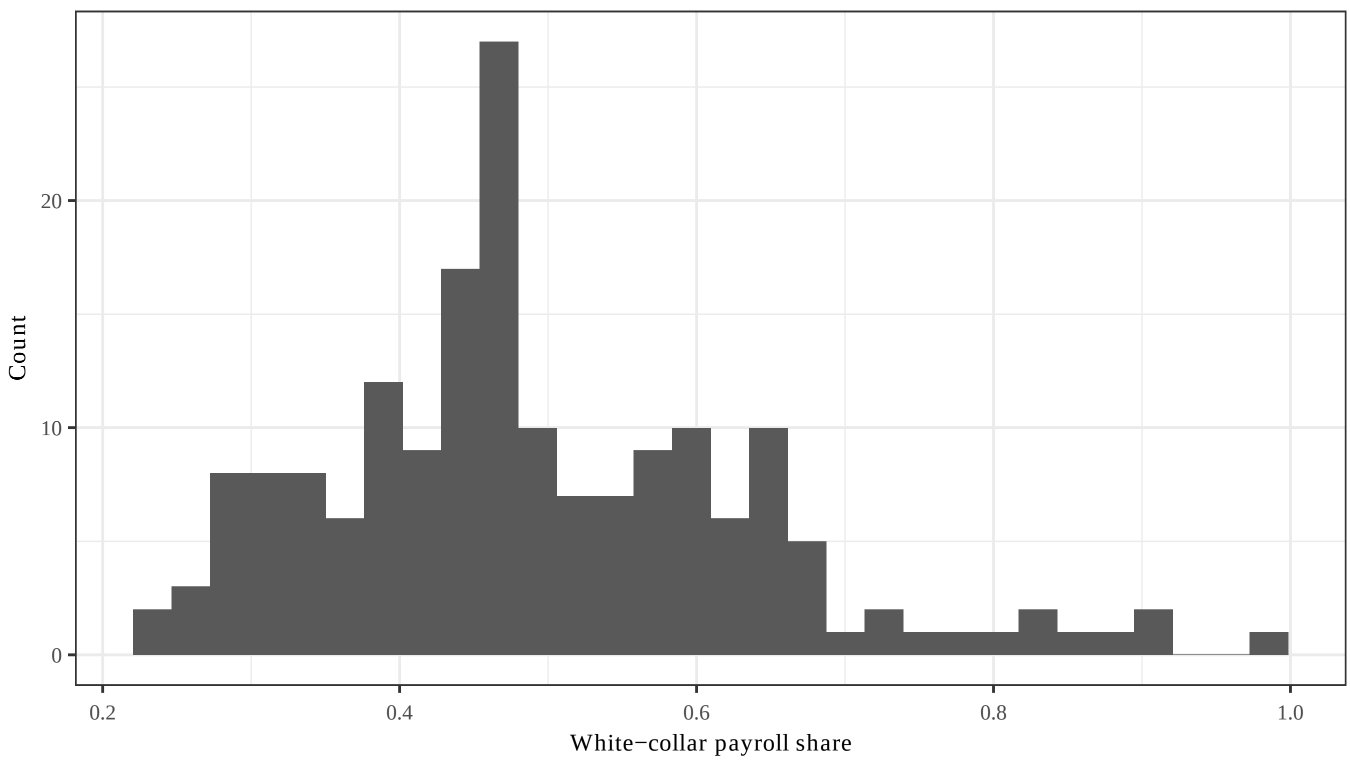
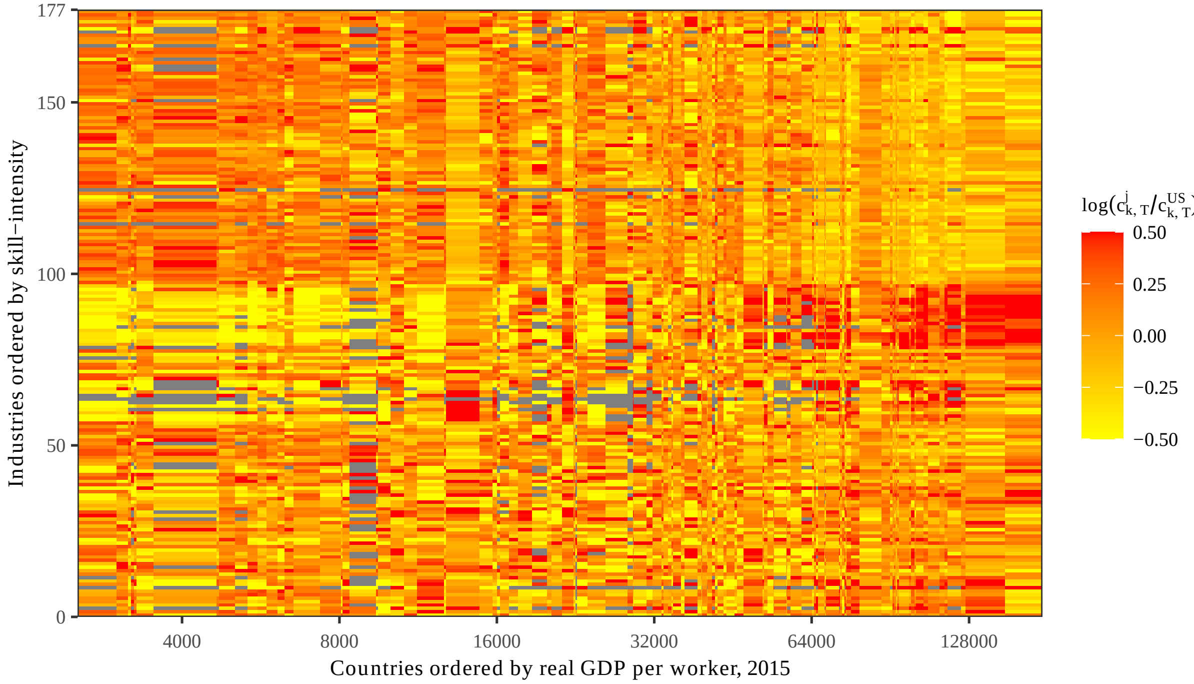
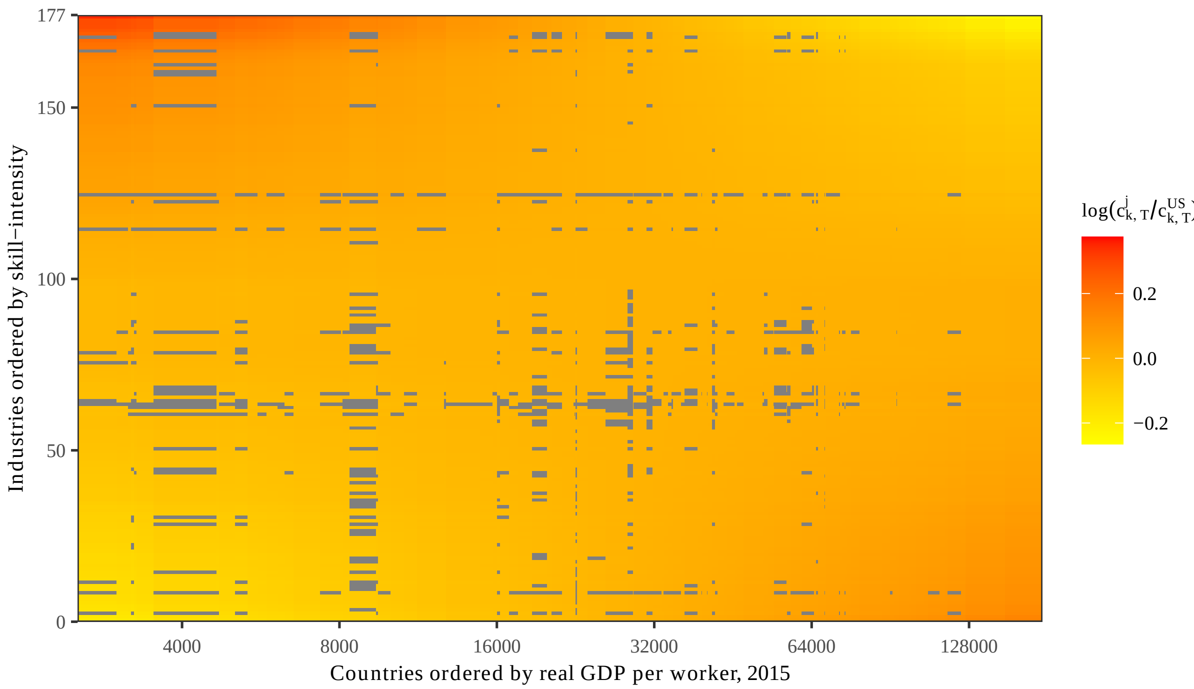


Figure 2: Histogram over skilled payroll share

Note: The white-collar payroll share is calculated using the 2015 Occupational Employment Survey's information about number and average pay of workers by industry and occupation, concorded to ISIC rev 4 industry code from NAICS 2012 and to ISCO-08 from SOC-10 occupational codes. White-collar workers are defined as those with a 1-digit ISCO-08 code of 1 to 4. GDP per worker is real expenditure based GDP per worker from the Penn World Table 9.1.



(a) Estimated relative unit costs by country and industry



(b) Smoothed relative costs by country and industry

Figure 3: Estimated relative costs

Note: Each rectangle represents a country-industry pair, with countries ordered by real GDP per worker on the horizontal axis, and industries ordered by US white-collar payroll share on the vertical axis. Panel A displays $-\frac{\hat{\gamma}^i_k}{\theta}$, where $\hat{\gamma}^i_k$ is the origin-industry fixed effect from regression (8). Panel B displays the predicted values from (10).

Table 2: Bottom 10 industries by skill-intensity

ISIC code	Industry	White-collar payroll share
1430	Knitted and crocheted apparel	0.22
1010	Processing and preserving of meat	0.23
0520	Mining of lignite	0.25
1020	Processing and preserving of fish, crustaceans and molluscs	0.26
1079	Other food products n.e.c.	0.27
1075	Prepared meals and dishes	0.28
1071	Bakery products	0.29
1074	Macaroni, noodles, couscous and similar farinaceous products	0.29
0170	Hunting, trapping and related service activities	0.29
0311	Marine fishing	0.29

The skill shares is the white-collar payroll share, calculated using the 2015 Occupational Employment Survey's information about number and average pay of workers by industry and occupation, concorded to ISIC rev 4 industry code from NAICS 2012 and to ISCO-08 from SOC-10 occupational codes. White-collar workers are defined as those with a 1-digit ISCO-08 code of 1 to 4.

3.3 Results

I estimate the regression equation (8) using OLS, clustering standard errors on an origin-industry level (see appendix for an estimate using the Pseudo-Poisson Maximum Likelihood estimator proposed by Silva and Tenreyro, 2006). The main outcomes are $-\hat{\gamma}_k^j/\theta$, the origin-industry fixed effects normalized by the trade elasticity. By (9) in proposition 1, these parameter estimates are unbiased estimates of relative unit costs by industry. There are 177 industries and 123 countries, yielding 21,771 potential combinations of industry-origin fixed effects (in the end, there are 20,492 non-missing estimates).

The results are displayed in panel A in figure 3, which shows $-\hat{\gamma}_k^j/\theta$ for every combination of countries and industries. Each rectangle represents one origin-industry fixed effect, with red representing a relatively high unit cost, and yellow representing a relatively low unit cost. To illustrate the correlation between income levels, skill intensity and unit costs, the plot has arranged countries by increasing GDP per worker on the horizontal axis, and industries by increasing skilled compensation share on the vertical axis. Doing this, we can note that unit costs are relatively high in the upper left corner and the lower right corner of the graph, illustrating that poor countries are relatively expensive in skill-intensive production while rich countries are relatively expensive in unskill-intensive production. Conversely, we see that unit costs are relatively low in the lower-left corner and upper-right corner, illustrating that poor countries have relatively low costs in unskill-intensive production and that rich countries have relatively low costs in skill-intensive production.

To formally test for such an interaction, I run the regression

$$(-\hat{\gamma}_k^j/\theta) = a^j + b_k + \eta(s_k \times \log y_j) + \epsilon_k^j, \tag{10}$$

where s_k is the skilled labor compensation share of industry k, and $\log y_j$ is the log GDP per worker of country j. Table 3 shows these regression results, which confirm that unit costs exhibit a negative interaction between skill intensity and country income levels. The regression results are also illustrated in panel B of figure 3, which plots the fitted values of (8) by country and skill intensity level.[10]

Table 3: Relative unit costs as a function of $\log(y^j) \times s_k$

	Dependent variable:
	Relative unit cost
$\log(y) \times$ Skill intensity	-0.310^{***}
	(0.015)
Observations	20,492
R^2	0.021
Adjusted R^2	0.006
Residual Std. Error	0.328 (df = 20195)
F Statistic	1.450^{***} (df = 296; 20195)
Note:	*p<0.1; **p<0.05; ***p<0.01

3.4 Discussion and robustness

The main results in table 3 relies on a gravity regression to back out the unit costs. What is required for such an approach to work? As shown in proposition 1, the main requirement is that trade cost shocks ε_k^{ij} do not vary systematically with industry for a given bilateral trading pair. This is what ensures that relative prices facing consumers in the destination country reflect relative costs in the origin country. In addition, in assuming that the price faced by consumers is unit costs plus trade costs, I abstract from potential export taxes or subsidies which will therefore end up in unit cost estimates. Last, industry-specific variations in taste in destination countries might also affect unit costs estimates.

In the appendix, I conduct a number of robustness checks. To ensure that taste variations do not affect my main findings, I show that the negative interaction term in table 3 remains even if we restrict destinations to being either only low-income countries or high-income countries. I consider an alternative econometric strategy using the Pseudo-Maximum Likelihood estimator proposed by Silva and Tenreyro (2006), and I perform the estimation allowing for different trade elasticities, using elasticity estimates from Caliendo and Parro (2015). I also go beyond using trade data and estimate the relation between unit costs, GDP per worker

[10]For standard errors, table 3 reports robust standard errors for (10). In the appendix, I show that the findings are similar if I instead estimate the interaction term directly from the trade regression, assuming that relative unit costs are a function of the interaction term, and clustering standard errors on the country-indusry level.

and skill intensity directly for the subset of countries and industries that have gross output prices in the Groningen Growth and Development (GGDC) industry-level data base (Timmer, Ypma, and van van Ark, 2007). Throughout these variations, the main pattern remains: rich countries have low relative unit costs in skill intensive industries. To ensure that this pattern is about skill and not other industry features, I also show that the estimate in table 3 remains negative once I include a number of other industry controls. Last, I use the OECD-WTO Balanced Trade in Services (BaTIS) dataset to show that rich countries' comparative advantage in skill-intensive industries extends beyond merchandise trade.

4 Unit cost variations within industries

Next, I turn to assessing unit cost variations within industries. My main result is that rich countries have relatively low costs in the tradable segments of industries, while poor countries have relatively low unit costs in the non-traded segments of industries.

In terms of identifying supply parameters, the within-industry case is more challenging than the between-industry case, where it was possible to remove dependence on demand and average bilateral costs by looking at relative exports conditional on destination demand. However, through the lens of my model, it is possible to infer a comparative disadvantage in tradables for poor countries by observing that they have a surprisingly low import penetration given how small they are in world export markets. This fact has been noted earlier in Waugh (2010), and is known to be problematic for standard gravity setups when the only heterogeneity across countries is uniform productivity differences. In my model, this fact implies a comparative advantage for poor countries in non-tradables.

Section 4.1 explains the logic of identification in a stripped-down model, section 4.2 provides a formal identification result, and section 4.3-4.4 provide data description and results. Last, section 4.5 discusses the robustness of the results and their relationship to earlier findings in the literature, in particular to Waugh (2010).

4.1 Identification logic

Consider a setup with multiple countries, where each country has a single industry using labor as its only input, and where countries are identified by their sizes L as well as by their tradable and non-tradable productivities A_T and A_{NT}. In such a setting, exports and imports for a small economy can be shown to satisfy

$$\frac{Exp}{Y} = \frac{\kappa(w/A_T)^{-(\theta+1)}}{A_T} \tag{11}$$

$$\frac{Imp}{Y} = \frac{\lambda}{a_T\,(w/A_T)^{-\theta} + a_{NT}(w/A_{NT})^{-\theta} + \lambda} \tag{12}$$

where the wage level w adjusts to ensure balanced trade, and κ and λ are parameters that capture export demand and import supply, and that do not depend on w, A^T and A^{NT}.[11] Equation (11) shows exports falling with the unit cost w/A_T of tradables, while (12) shows imports rising in the unit costs of tradables w/A_T and the unit cost of non-tradables w/A_{NT}. Normalizing wages and non-tradable productivity with A_T, I obtain the following expressions

$$\frac{Exp}{Y} = \frac{\kappa}{A_T}\tilde{w}^{-(\theta+1)} \tag{13}$$

$$\frac{Imp}{Y} = \frac{\lambda}{a_T\tilde{w}^{-\theta}[1 + \frac{a_{NT}}{a_T}\tilde{A}_{NT}^{\theta}] + \lambda}, \tag{14}$$

where $\tilde{w} \equiv w/A_T$ and $\tilde{A}_{NT} \equiv A_{NT}/A_T$.

Figure 4 illustrates this setup with export/import shares of GDP on the horizontal axis, productivity adjusted wages $\tilde{w}$ on the vertical axis, with equilibrium determined by trade balance. Panel A shows the effect of reducing A_T and A_{NT} uniformly, so that A_T decreases and $\tilde{A}_{NT} = A_{NT}/A_T$ is constant. The export curve (13) shifts out, increasing both w/A_T and trade openness. Intuitively, a lower A_T increases labor requirements in tradables, which means that the economy needs to sell less tradables to respect labor market clearing. In equilibrium, this is achieved by a rising $\tilde{w}$, that is, by wages not falling enough to compensate for the lower productivity. The import share rises since higher costs make local goods less attractive.

Panel B shows the effect of increasing A_{NT} keeping A_T constant. This change shifts the import curve (14) to the left, leading to a fall in the openness of the economy and a rise in $\tilde{w}$. Intuitively, a higher A_{NT} reduces the price of local output causing consumers to switch expenditures away from imports towards locally produced goods. This reduces the openness of the economy. Moreover, in equilibrium, this switch towards locally produced goods requires a transfer of workers from the export sector, requiring a rise in $\tilde{w}$ to restrict

[11] In appendix D.1, I derive (11)-(12) and provide formulas for κ and λ. I show that κ and λ are only functions of the size of the domestic labor endowment, foreign wages, and foreign productivities, and can thus be treated as parameters under a small economy assumption where we treat foreign wages as fixed. Further, the exporting formula uses the assumption that other economies are large in the sense that exports from the economy under under study are small relative to the absorption of the destination countries.

demand for exports.

The logic in figure 4 can be used to devise an identification strategy for A_{NT}/A_T. First, we note that variations in A_{NT}/A_T can be identified from variations in the import shares provided that we know the cost of exportables $\tilde{w}$. That is, if we fix the cost of exportables, a high import share maps directly to a comparative advantage in non-tradables. Second, gravity models generally allow us to estimate the cost of exportables from export data: we simply look at how successful a country is at penetrating foreign markets. Combining these two facts points to an identification strategy: a high A_{NT}/A_T can be identified through higher-than-expected import shares conditional on export success.

4.2 Regression equation

To go from the simple to the full model, we need to take into account that there are multiple industries, trade costs, and variations in population across countries. Normalizing the gravity equation (6) by own-trade X_k^{ii}, we obtain:

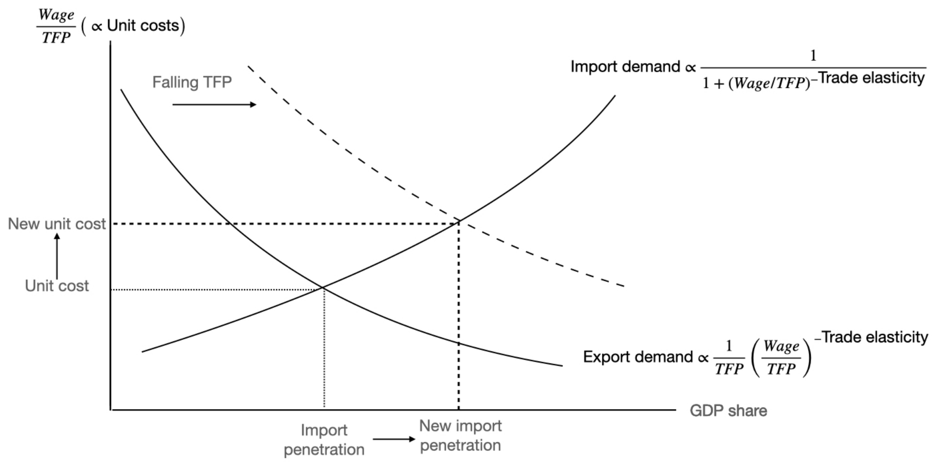
$$\log \frac{X_k^{ij}}{X_k^{ii}} = -\theta \log\left(\tau_k^{ij}\right) - \log \frac{\left(c_{k,T}^j\right)^\theta}{G_T(L^j)} + \log \frac{(c_{k,T}^i)^\theta}{G_T(L^i)} - \log\left(1 + \frac{a_{NT}}{a_T}\frac{G_{NT}(L^i)}{G_T(L^i)}\left(\frac{c_{k,T}^i}{c_{k,NT}^i}\right)^\theta\right) \qquad i \neq j. \tag{15}$$

Here, we can note here that the attractiveness of exports is high if the cost term $\log \frac{(c_{k,T})^\theta}{G_T(L)}$ is low. The same term simultaneously pushes up imports. Thus, it plays an analogous role to the productivity-adjusted wage $\tilde{w}$ in the simple model. In contrast, the last term captures that imports are depressed if countries have a comparative advantage in non-tradables, that is, if $c_{k,T}^i/c_{k,NT}^i$ is high. This term selectively depresses imports without affecting exports; thus, its role is analogous to the comparative advantage in non-tradables A_{NT}/A_T in the simple model.

Just as in the simple model, a comparative advantage in non-tradables imply low imports conditional on the attractiveness of exports. The following proposition formalizes.

Proposition 2. *Suppose trade costs satisfy*

$$\log \tau_k^{ij} = \tau_k + \sum_m \beta_{m,k} d_m^{ij} + \varepsilon_k^{ij} \tag{16}$$



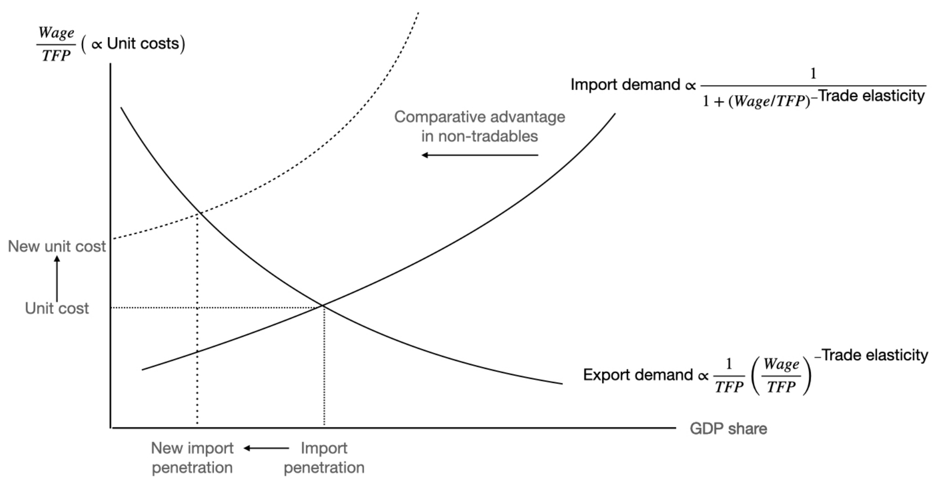


Figure 4: Comparative statics in simplified model

with $\mathbb{E}\varepsilon_k^{ij} = 0$ *for all* i, j, k*, and consider the trade regression*

$$\log \frac{X_k^{ij}}{X_k^{ii}} = \alpha_k + \textit{Export success}_k^j - \textit{Export success}_k^i - \textit{Home bias}_k^i + \sum_m \beta_{m,k} d_m^{ij} + \tilde{\varepsilon}_k^{ij}. \tag{17}$$

where d_m^{ij} *is not multicollinear with the fixed effects. Given the normalization* $\sum_j \textit{Export success}_k^j = \sum_j \textit{Home bias}_k^j = 0$ *for all* k*, we have*

$$\mathbb{E}\left[\widehat{\textit{Export success}}_k^i\right] = \log\left(G_T(L^i)\left(c_{k,T}^i\right)^\theta\right) - \overline{\log\left(G_T(L)\left(c_{k,T}\right)^\theta\right)} \tag{18}$$

$$\mathbb{E}\left[\widehat{\textit{Home bias}}_k^i\right] = \log\left(1 + \frac{a_{NT}}{a_T}\frac{G_{NT}(L^i)}{G_T(L^i)}\left(\frac{c_{k,T}^i}{c_{k,NT}^i}\right)^\theta\right) - \overline{\log\left(1 + \frac{a_{NT}}{a_T}\frac{G_{NT}(L)}{G_T(L)}\left(\frac{c_{k,T}}{c_{k,NT}}\right)^\theta\right)} \tag{19}$$

where the bar denotes averaging across countries.

Proof. See the appendix. □

The proof relies on the gravity equation (15). The term Export success$_k^j$ is an origin-fixed effect that captures the attractiveness of exports, while Home bias$_k^i$ captures the comparative advantage in non-tradables (modulo population size). To see why the identification captures the notion of "low imports conditional on exports", note that the destination fixed effect in the regression specification is $-$(Export success$_k^i$ + Home bias$_k^i$); hence, if home bias is large, this is equivalent to a surprisingly low destination fixed effect given the attractiveness of exports. Since the identification relies on the levels of exports and imports, rather than their relatives across industries as in proposition 1, the assumption on trade costs (16) is more restrictive than in section 3: there is no origin-industry fixed effect or a bilateral fixed effect (see the robustness checks for a more thorough discussion about the specification of trade costs).

Our main specification uses (17) to estimate $\widehat{\text{Home bias}}_k^i$ for all countries and industries. To test whether $c_{k,T}^i/c_{k,NT}^i$ systematically varies with development, we correct for population size by comparing $\widehat{\text{Home bias}}_k^i$ for poor and rich countries with similar populations.

4.3 Data

The data used in section 3 needs to be complemented with production data to derive own-trade X_k^{ii}. I use UNIDO INDSTAT 4 to obtain output data for manufacturing industries at a 2-digit level, and UN Detailed National Accounts to obtain production data for agriculture and forestry (ISIC code A), as well as for the mining and extraction industry (ISIC code B). To make the trade data consistent with the output data, I collapse to 2-digit industry codes, as well as ISIC codes A and B. This leaves me with 25 industries. Given gross output data, I define own-trade as $X_k^{ii} = Y_k^i - \sum_{j \neq i} X_k^{ji}$, where Y_k^i denotes gross output, and X_k^{ji} are (observable) exports. In the construction of own-trade, I exclude observations with a negative own-trade. Furthermore, I construct $X_k^i = X_k^{ii} + \sum_{j \neq i} X_k^{ij}$ for the total absorption of country i in industry k.

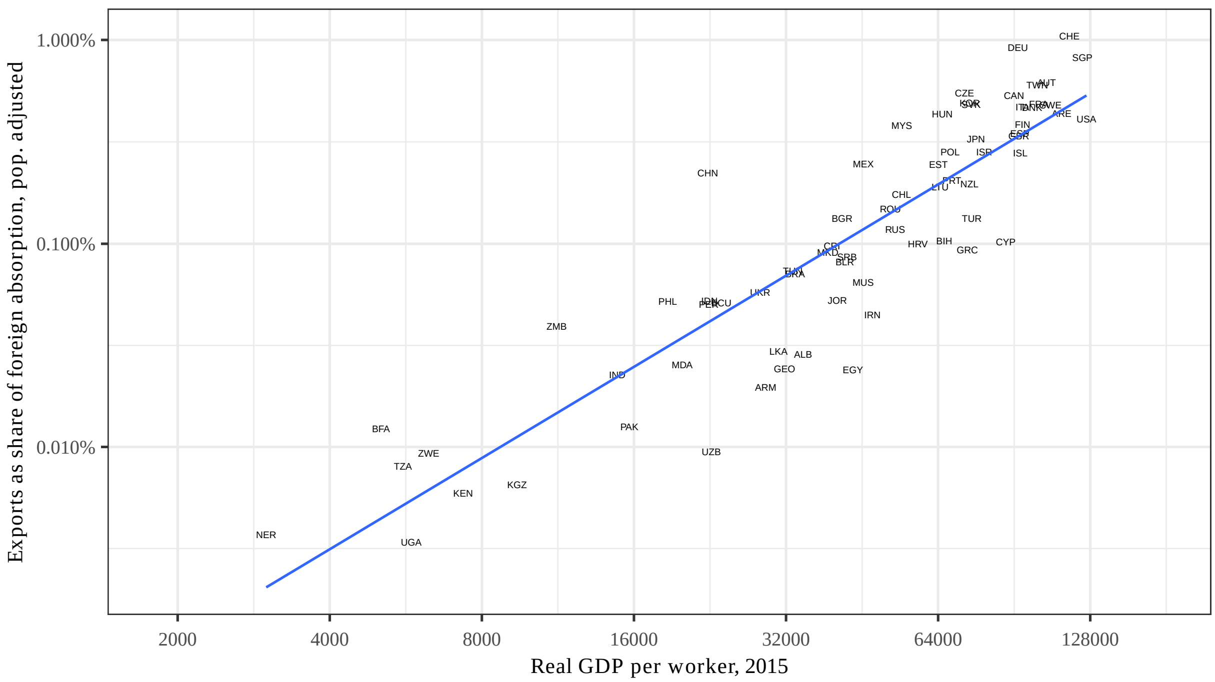
Data descriptives The discussion in section 4.1 suggests that a moment of interest is countries' relative success in penetrating foreign and domestic markets. Using the production and trade data, I can measure penetration of foreign markets as $s_{for}^i = \frac{\sum_k \sum_{j \neq i} X_k^{ji}}{\sum_k \sum_{j \neq i} X_k^k}$ and import shares as $s_{imp}^i = \frac{\sum_k \sum_{j \neq i} X_k^{ij}}{\sum_k X_k^i}$. Here, s_{for}^i has total exports of i in the numerator and total foreign absorption in the denominator, and s_{imp}^i has total imports in the numerator, and total domestic absorption in the denominator. To focus on the relationship with GDP per worker, I standardize the measures for population differences, with all values standardized to those of a country with 10 million workers.

The results are displayed in figure 5. Panel A shows that rich countries are dramatically more successful in penetrating export markets than poor countries. As we go from the poorest to the richest country, the share of foreign absorption hat is covered by a country rises more than two orders of magnitudes, from less than 0.01% to almost 1%.. The relationship between income and export success is also very tight: it has an R^2-value of 0.95. This fact is what we should expect from poor countries having a low productivity: in the language of figure 4, a lower productivity shifts out the export demand curve, which increases the productivity-adjusted wage level and pushes down a country's export success.

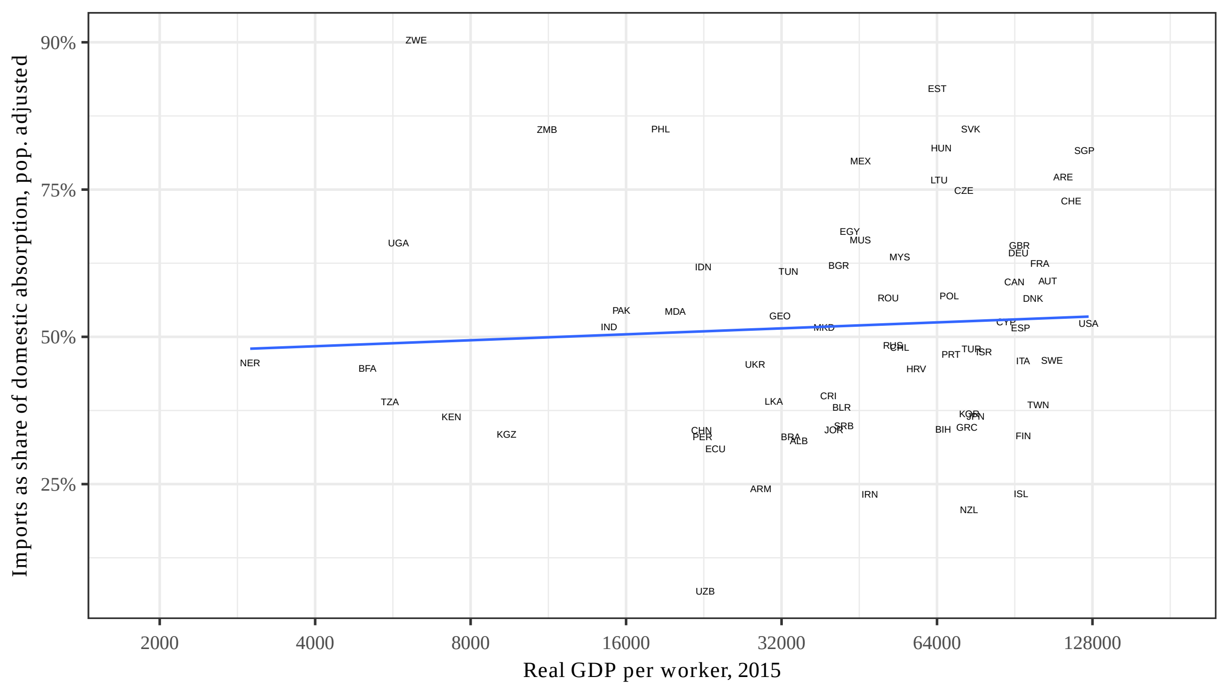
In Panel B, we see that despite large variations in export success, there is no systematic trend in the share of domestic output in domestic absorption. Even if there is considerable variation, the average import share is stable at around 50% of domestic absorption. This is contrary to the effect of uniform productivity differences in figure 4, which predicts that countries with low productivity should have a high import penetration. Instead, the data is prima facie more consistent with a shift in A_T that is larger than the shift in A_{NT}, so that poor countries have a comparative advantage in non-tradables, representing a joint shift out of the export demand curve and a shift in of the import demand curve. The next section shows that the regression results confirm this conjecture.

4.4 Results

Regression (17) provides one estimate for each country-industry combination. With 25 industries and 89 countries, this implies up to 2225 values for the home bias term $\widehat{\text{Home bias}}_k^i$. To illustrate the structure of the



(a) Export penetration as share of foreign absorption



(b) Import penetration as a share of domestic absorption

Figure 5: Export and import penetration

Note: Real GDP per worker is from Penn World Table 9.1. The export share is total exports divided by the sum of all foreign absorption, and the import share is total imports divided by domestic absorption, with absorption defined in (**??**). The export and import shares are adjusted for population by regressing logit export/import shares on log population, with the plots showing the counterfactual value if a country would have had 10 million in employment.

estimates, figure 6 displays the estimated home bias for one industry: food manufacturing. Panel A reports individual estimates, and panel B is a smoothed version. The values of home bias are encoded in color, with GDP per worker on the horizontal axis and total employment on the vertical axis. The employment dimension is included since proposition 2 shows that home bias is a function of size L^i in addition to the comparative advantage of non-tradables $c^i_{k,T}/c^i_{k,NT}$, our quantity of interest.

Panel A shows that home bias is low in large countries or countries with high incomes, like India, China, and the Netherlands. Panel B displays the same pattern by plotting the contour of a linear regression of home bias on (log) GDP per worker and employment: the diagonal contours capture that home bias is falling both with size and income level. The negative relationship between home bias and income suggests that rich countries have a comparative advantage in tradable production. Indeed, by (19). the home bias term is an unbiased estimate of $\log\left(1+\frac{G_{NT}(L^i)}{G_T(L^i)}\frac{c^i_{k,T}}{c^i_{k,NT}}\right)$, so a negative relationship between income and home bias for a fixed L^i implies that the relative price of tradables, $c^i_{k,T}/c^i_{k,NT}$, falls with income.

To test whether this pattern is general across industries, I run the regression

$$\widehat{\text{Home bias}}^i_k = F_k(L^i) + \xi_k \times \log \text{GDP per worker}^i \tag{20}$$

for every industry, where $F_k(L^i)$ is a flexible control for the size of the country, and ξ_k captures whether home bias is lower in rich countries conditional on L^i. Figure 7 reports the result of this regression, and we see that $\hat{\xi}_k$ is consistently negative and significant for every industry but one, suggesting that $c^i_{k,T}/c^i_{k,NT}$ is consistently lower for rich countries.

4.5 Interpretation and robustness

My findings are closely related to those of Waugh (2010). He also observes that standard gravity logic struggles to explain how poor countries can have similar import penetration to rich countries despite being so unsuccessful in export markets. He then shows that to rationalize the data, poor countries need to have something looking like high trade costs that are specific to exports.[12] Anything acting as such a trade cost breaks the link between export success and import penetration: if exports fail due to high trade costs, import penetration can still be low as long as a country's goods are competitive in their local market.

My findings provide an interpretation for the content of such export-specific trade costs. First, the non-traded technology in my model acts like an effective export cost. This is because, in Eaton and Kortum

[12] Waugh also discusses the possibility of import-specific costs, but concludes that they are inconsistent with the similar prices of tradable good across different countries.

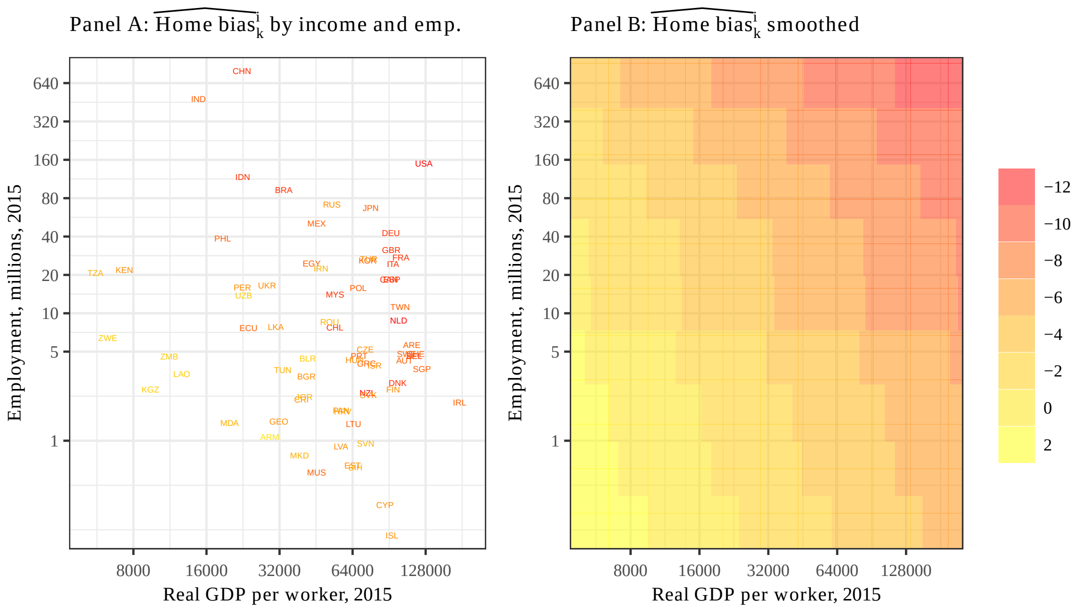


Figure 6: Home bias for food manufacturing

Notes: Real GDP per worker and employment are from Penn World Table 9.1. The colors in Panel A code for $\widehat{\text{Home bias}}_k^i$ for country i with k =Food manufacturing. Panel B plots the predicted values from a regression that predicts home bias with log GDP per worker and discretized employment levels.

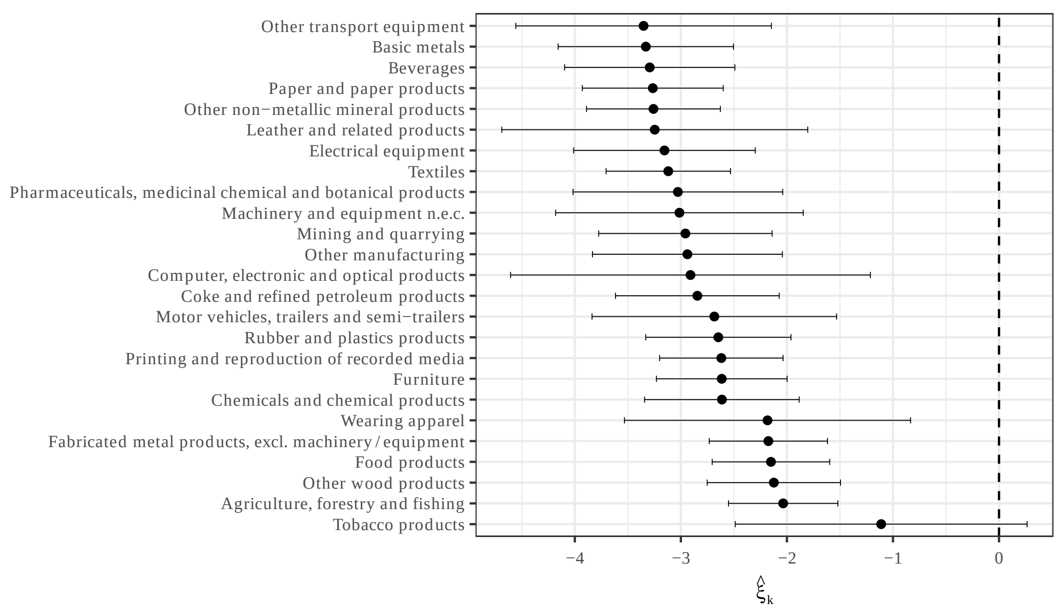


Figure 7: Slope $\hat{\xi}_k$ of home bias relative to GDP per worker

The graph displays $\hat{\xi}_k$ from the regression (20) run separately for each industry, with log employment discretized into 11 categories. The error bars represent 1.96 times the standard error.

language, the non-tradable technology only provides "draws" to local consumers. Thus, similar to with an export-specific trading costs, the existence of the non-traded technology increases the relative cost faced by global versus local consumers. Moreover, these effective "export costs" are higher in poor countries, since a comparative advantage in non-tradables expands the cost advantage of local producers for local consumers.[13]

The appendix formalizes the mathematical equivalence between my model and a model with export-specific trade costs. I also examine if literal trade costs provide a good explanation of my data. I find that such trade costs need to be larger than what seems warranted by unit cost patterns for industries where unit costs are available. In particular, for agriculture, implied unit costs in poor countries from trade data are only moderately higher than observed unit costs from the FAO. While this implies somewhat higher export costs, these are not large enough to explain my findings. Conversely, the similarity of import price indices across countries suggest that poor countries do not have dramatically higher import costs.

5 Skill-biased development differences

The results in section 3 and 4 show that when we compare poor and rich countries, unit cost differences are skill-biased between and within industries in the following sense: between industries, rich countries have relatively low unit costs in skill-intensive industries compared to in non-skill intensive industries, and within industries, rich countries have relatively low unit costs in the more skill-intensive tradable segment compared to in the less skill-intensive non-tradable segment.

In this section, I go one step further to establish that industry-level *productivity* differences across countries are also skill-biased between and within industries. The key extra step is to distinguish variations in unit cost *functions,* which reflect productivity, from variations in *realized* unit costs, which are the measures provided in section 3 and 4. To perform this extra step, I need a strategy to correct for variations in factor prices across countries.

My argument has three steps. First, I use the concept of log-submodularity to provide a general definition of skill-biased development differences between and within industries in terms of arbitrary unit cost functions. Second, I use this definition to derive a test for skill bias that can be performed using only the realized unit costs from section 3 and 4 together with data on skilled wage premia from countries at different income levels. Third, I apply this test and show that development differences indeed are skill-biased between and

[13]An alternative perspective that similarly generates effective export costs is if high-quality goods have lower trade costs and poor countries have a comparative disadvantage in high-quality goods as in Fajgelbaum et al. (2011). Even though my model does not explicitly incorporate quality, this would similarly imply that poor countries have a comparative disadvantage in skill-intensive production, since high-quality goods are typically more skill-intensive (Jaimovich et al., 2020).

within industries.

5.1 Definition

Consider a setup where a collection of countries $\mathcal{Y}$ are ordered by development level and a collection of industries $\mathcal{K}$ are ordered by skill-intensity, and where each industry consists of a traded and a non-traded segment, T and NT. In each segment, production is done using unskilled and skilled labor and a constant returns to scale production function. Production functions vary across countries and are characterized by the set of unit cost functions

$$c_t(w_u, w_s; k, y) \quad t \in \{NT, T\}, k \in \mathcal{K}, y \in \mathcal{Y}, \tag{21}$$

where the elasticities of the unit cost functions with respect to skilled wages satisfy

$$\frac{\log c_t(w_u, w_s; k', y)}{\partial \log w_s} \geq \frac{\log c_t(w_u, w_s; k, y)}{\partial \log w_s} \qquad \forall w_u, w_s, y, k' > k \tag{22}$$

$$\frac{\log c_T(w_u, w_s; k', y)}{\partial \log w_s} \geq \frac{\log c_{NT}(w_u, w_s; k, y)}{\partial \log w_s}, \quad \forall w_u, w_s, y, k \tag{23}$$

which captures that skill-intensity is higher in industries with a high k, as well as in the tradable segment within industries. Given this setup of unit cost functions, I make the following definition of skill-bias between and within industries.

Definition 3. Given cost functions (21) satisfying (22) and (23), development is said to be...

1. ...skill-biased between industries if

$$\log \frac{c_t(w_u, w_s; k', y')}{c_t(w_u, w_s; k, y')} \leq \log \frac{c_t(w_u, w_s; k', y)}{c_t(w_u, w_s; k', y)} \qquad \forall k' > k, y' > y$$

2. ...skill-biased within industries if

$$\log \frac{c_T(w_u, w_s; k, y')}{c_{NT}(w_u, w_s; k, y')} \leq \log \frac{c_T(w_u, w_s; k, y)}{c_{NT}(w_u, w_s; k, y)}. \quad \forall k, y' > y$$

Definition 3 states that skill-biased development is equivalent to unit cost functions being log-submodular in development and skill intensity. The concept of log-submodularity has emerged as a general way to define complementarity in a number of areas of economics.[14] Here, the concept similarly offers a natural way of characterizing skill bias as a complementarity between development and skill-intensive production processes.

5.2 Testing for skill bias in practice

To apply definition 3, I use the trade-based unit cost estimates from 3 and 4. I combine these estimates with ILO data from across the world on skilled wage premia w_s^i/w_u^i, which allows me to adjust unit cost estimates for variation in factor prices.

Skill bias between industries. First, I test whether the estimated relative unit costs $\log \widehat{\frac{c_{k,T}^i/\bar{c}_T^i}{\bar{c}_{k,T}/\bar{c}_T}}$, estimated in section 3, are consistent with skill-biased development between industries. Since skill bias is defined for fixed factor prices, a key challenge is how to adjust for differences in skilled wage premia across countries. The proposition below shows how this can be done given smoothed versions of relative unit costs and skilled wage premia.

Proposition 4. *Suppose that there is a continuum of countries $\mathcal{Y} = [\underline{y}, \bar{y}]$ ranked by income, that skilled wage premia across countries are given by some differentiable function $\omega(y)$. Define doubly normalized unit costs as $\log \tilde{c}(y,k) \equiv \log \frac{c(1,\omega(y);k,y)/\bar{c}(1,\omega(y);\bullet,y)}{\bar{c}(1,\overline{\omega};k,\bullet)/\bar{c}(1,\overline{\omega})}$ where $\bar{c}(1,\omega(y);\bullet,y)$, $\bar{c}(1,\overline{\omega(y)};k,\bullet)$, and $\bar{c}(1,\overline{\omega(y)})$ are geometric averages of unit costs across industries, countries, or both (with averaging across countries taking into account varying skilled wage premia). Then c exhibits skill-bias between industries only if*

$$\frac{\partial \log \tilde{c}(y,k')/\tilde{c}(y,k)}{\partial y} \leq \frac{d \log \omega(y)}{dy}(s_{k'}(y) - s_k(y)), \tag{24}$$

where $s_k(y) \equiv \left.\frac{\partial \log c[1,w_s/w_u;k,y]}{\partial \log w_s}\right|_{\frac{w_s}{w_u}=\omega(y)}$ is the elasticity of costs with respect to skilled labor. If there is a continuum of industries and s_k^y, c, and $\tilde{c}$ are twice differentiable, the condition is:

$$\frac{\partial \log \tilde{c}(y,k)}{\partial y \partial k} \leq \frac{d \log \omega(y)}{dy} \frac{\partial s_k^y}{\partial k}. \tag{25}$$

The condition is invariant under any monotone, differentiable, transformation of the index variable k.

Proof. See the appendix. □

The key to proposition 4 is that skill bias requires relative unit costs in skill-intensive industries to fall *faster* with development than what is predicted by changes in skill premia alone. The pure effect of skill

[14] See, for example, Shimer and Smith (2000) for an application to matching, Costinot (2009) for an application to international trade, and Davis and Dingel (2020) for an application to urban economics.

premia on relative prices is captured by the right-hand side of (24): if the skill premium rises, the relative unit cost of two industries changes in line with the difference in their skill shares. Since skill bias requires that relative unit costs $\tilde{c}(y,k')/\tilde{c}(y,k)$ fall faster with development than this quantity, I obtain (24). To obtain the second expression (25), I simply let $k'-k \to 0$.

To make proposition 4 actionable, I use that condition (25) is invariant to any change in industry labeling. In particular, it can be applied if k is indexed by US skill intensity, in which case we obtain

$$\frac{\partial^2 \log \tilde{c}(y, s^{US})}{\partial \log y \partial s^{US}} \leq \frac{d \log \omega(y)}{d \log y} \frac{\partial s^y}{\partial s^{US}}. \tag{26}$$

which shows that skill bias can be assessed from a smooth function that expresses relative unit costs as a function of income levels and US skill shares. The term $\partial s^y/\partial s^{US}$ captures how much the skill share at an income level y changes when we traverse a certain distance in skill share space using a US base. The size of $\partial s^y/\partial s^{US}$ then captures whether the skill share distribution is more dispersed in poor countries than in rich countries, with a larger dispersion of skill intensities ($\partial s^y/\partial s^{US} > 1$) meaning that changes in factor prices have a larger effect on relative unit costs. In the special case when industries have the same skill intensity in all countries, we have $\partial s^y/\partial s^{US} = 1$ and (26) does not require information about factor shares outside of the US. If skill intensities are different across the world, we have $\partial s^y/\partial s^{US} \neq 1$ and varying across countries and industries. To adjust for this, I consider robustness checks and test whether (26) hold for a wide range of $\partial s^y/\partial s^{US}$.[15]

Condition (26) requires smoothed versions of $\log \widehat{\frac{c^i_{k,T}/\bar{c}^i_T}{\bar{c}_{k,T}/\bar{c}_T}}$ and $\log \frac{w^i_s}{w^i_u}$. To estimate these, I use the functional form assumption $\log \tilde{c}(y_i, s_k^{US}) = \beta^{y,s}(\log y_i \times s_k^{US})$ and $\log \tilde{\omega}(y_i) = \alpha + \beta^{y,w_s/w_u} \times \log y_i$, which imply that $\frac{\partial^2 \log \tilde{c}}{\partial \log y \partial s^{US}} = \beta^{y,s}$ and $\frac{d \log \tilde{\omega}}{d \log y} = \beta^{y,w_s/w_u}$ are constants. I identify these constants using the regressions

$$\log \widehat{\frac{c^i_{k,T}/\bar{c}^i_T}{\bar{c}_{k,T}/\bar{c}_T}} = \beta^{y,s} \log y_i \times s_k^{US} + \varepsilon_k^{i,c} \tag{27}$$

$$\log \frac{w^i_s}{w^i_u} = \alpha + \beta^{y,w_s/w_u} \log y_i + \varepsilon_k^{i,c}. \tag{28}$$

The wage premia are plotted in figure 8 and the results are reported in table 4. The first line replicates the result in table 3 and shows that the cross-derivative of unit costs between income level and skill-intensity is -0.31. The second column shows that the skilled wage premium vary little with income levels. The

[15] A second extension is if industries use more inputs than only skilled and unskilled labor. In general, this weakens the ability of skilled wage premia variations to affect the relative unit costs of industries with different payroll shares of skilled labor, because the effect of skilled wage premium variations on relative unit costs is given by the cost share of skilled labor, not by the payroll share of skilled labor.

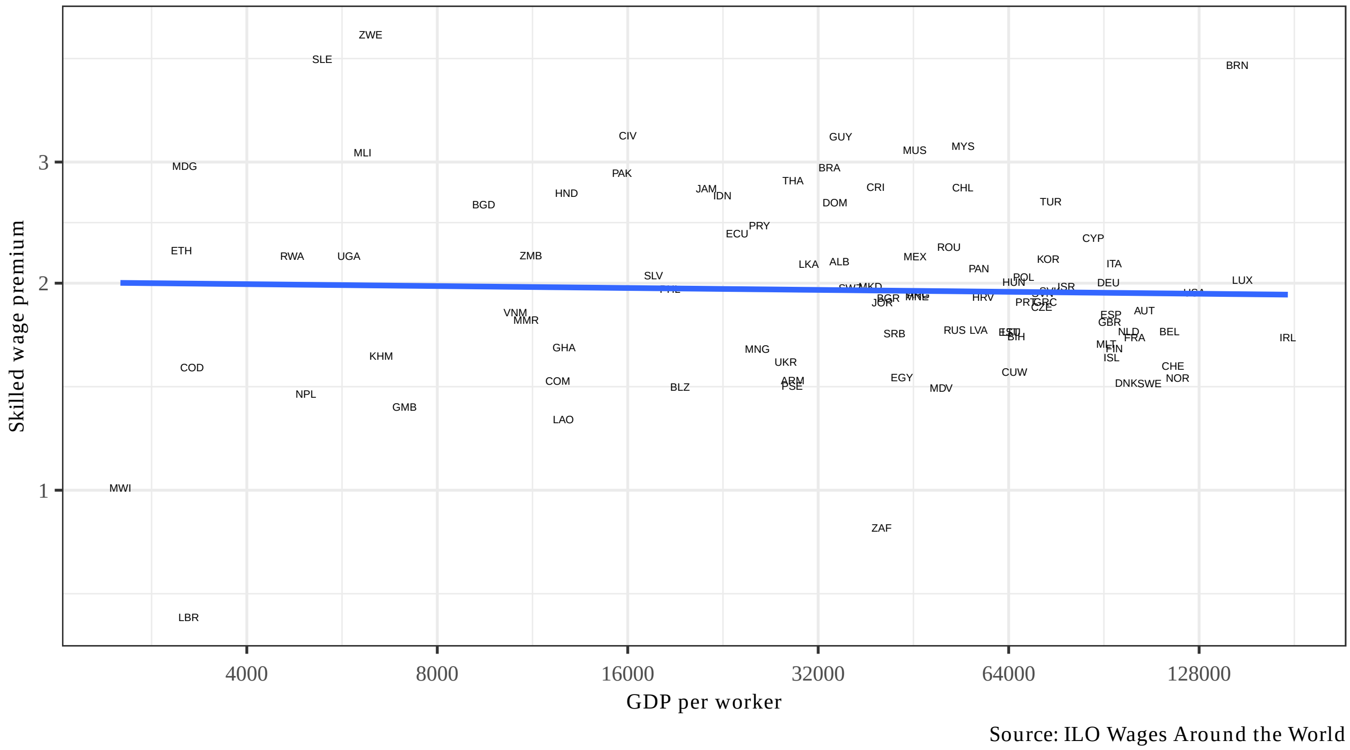


Figure 8: Skilled wage premium and GDP per worker

Wage data is from ILO Wages Around the World, which reports local wages by one-digit occupational group for 2010 to 2021. Skilled workers are defined as those with a one-digit ISCO code of 1 to 4. The wage premium is defined as the average wage of these occupations divided by the average wage of other occupations, averaged across the years 2012 to 2018.

last three columns show $\frac{\partial^2 \log \tilde{c}(y, s^{US})}{\partial \log y \partial s^{US}} - \frac{d \log \omega(y)}{d \log y} \frac{\partial s^y}{\partial s^{US}}$ for different values of $\frac{\partial s^y}{\partial s}$. We note that the expression is consistently negative and vary little with $\partial s^y / \partial s^{US}$, reflecting that skilled wage premia vary little with development. I conclude that the estimated unit cost functions are consistent with skill bias between industries.

Table 4: Testing for skill bias between industries

Parameters	values
$\beta^{y,s}$	-0.31
$\beta^{y,w_s/w_u}$	-0.01
$\beta^{y,s} - 0.5 \times \beta^{y,w_s/w_u}$	-0.31
$\beta^{y,s} - 1.0 \times \beta^{y,w_s/w_u}$	-0.30
$\beta^{y,s} - 2.0 \times \beta^{y,w_s/w_u}$	-0.29

Skill bias within industries. To evaluate skill bias within industries, I rely on skilled wage premium data and the estimates $\widehat{\text{Home bias}}^i_k$ from section 4. Assuming that $c_{k,T}/c_{k,NT}$ is a smooth function of $\log y$, proposition 4 implies that the relative unit costs are consistent with skill bias if

$$\frac{\partial \log \tilde{c}_{k,T}(y)/\tilde{c}_{k,NT}(y)}{\partial \log y} \le \frac{\partial \log \tilde{\omega}(y)}{\partial \log y}(s_{k,T} - s_{k,NT}), \tag{29}$$

which captures that skill bias requires observed relative unit costs $\tilde{c}_{k,T}/\tilde{c}_{k,NT}$ to fall faster with development than the change in the skilled wage premium, times the difference in skill shares between tradable and non-tradable industries. As in the case with between industry skill bias, the right-hand side captures the change in relative unit cost that comes from changes in the skilled wage premium alone. If we write $HB_k(y, L)$ for a smoothed version of the estimates $\widehat{\text{Home bias}}_k^i$ (as a function of income and employment), a sufficient condition for (29) is

$$\frac{1}{\theta}\frac{HB_k(y, L)}{\partial \log y} \le \frac{\partial \log \tilde{\omega}(y)}{\partial \log y}. \tag{30}$$

That is, the estimates $HB_k(y, L)$ has to fall faster with income than the skilled wage premium does.[16] From the section 4 results in figure 7, we have that $\frac{\partial HB_k(y,L)}{\partial \log y}$ ranges from -3.5 to -1 across different industries, implying that $\frac{1}{\theta}\frac{\partial HB_k(y,L)}{\partial \log y}$ ranges from approximately -0.6 to -0.15. From table 4, we see that $\frac{\partial \log \tilde{\omega}(y)}{\partial \log y} \approx -0.01$, so (30) is satisfied for all industries.

6 Interpretation

My results are related to the literature on skill-biased development. This literature studies why rich countries have a higher demand for skilled labor, and has proposed a number of mechanisms for this, including the structure of consumer demand, the structure of industry productivities, the quality of human capital, and endogenous technology differences (Buera and Kaboski, 2012; Caselli and Coleman II, 2001; Comin et al., 2020; Jaimovich et al., 2020).

My findings show that not only is aggregate demand for skilled labor high in rich countries, but rich countries also have low production costs in skill-intensive industries, as well as in the more skill-intensive tradable segment within industries. This suggests that there are cost-shifters on skilled labor that systematically lower the unit costs in skill-intensive production processes. Is there a theory of skill bias and development that implies these results? I show that one promising theory can be obtained by drawing on

[16] Condition (30) first uses that $\widehat{Home\ bias}_k^i$ identifies $\log\left(1 + \frac{G_{k,NT}(L^i)}{G_{k,T}(L^i)}\left[\frac{\tilde{c}_{k,T}^i}{\tilde{c}_{k,NT}^i}\right]^\theta\right)$ and that $\partial \log\left(1 + \frac{G_{k,NT}(L)}{G_{k,T}(L)}\left[\frac{\tilde{c}_{k,T}}{\tilde{c}_{k,NT}}\right]^\theta\right)/\partial \log y \le \theta \frac{\partial \log \tilde{c}_{k,T}/\tilde{c}_{k,NT}}{\partial \log ty}$. Second, it uses that $s_{k,T} - s_{k,NT} \le 1$ because both values are shares.

the literature that documents barriers to firm growth in developing countries (Akcigit et al., 2021; Hsieh and Klenow, 2014a). In particular, if we interpret the findings of this literature as reflecting barriers to the accumulation of organizational capital in poor countries, my results follow naturally if the accumulation and maintenance of such capital is the primary activity of skilled workers. Mathematically, connecting skill with organizational capital means that accumulation barriers to organizational capital systematically push up unit costs in skill-intensive production processes, both within and across industries.

Section 6.1 outlines a set of formal properties that lets a theory fit my findings; section 6.2 discusses existing theories in light of these requirements; section 6.3 outlines a theory based on barriers to the accumulation of organizational capital, and section 6.4 discusses the key assumptions of the model, the robustness to relaxing auxiliary assumptions, and a stylized quantitative exercise.

6.1 Explaining skill bias between and across industries

The trade findings show that unit costs in poor countries are systematically higher in skill-intensive production processes both within and across industries. Mathematically, a parsimonious way to rationalize this is to posit that unit costs take the form

$$c^i_{t,k}(w_u, w_s) = \frac{1}{A_i} C_{t,k}\left(w_u, \frac{e^{\tau^i}}{A^i_s} w_s\right), \tag{31}$$

with e^{τ^i}/A^i_s being high in poor countries. The specification assumes that unit costs in a segment (t,k) in i reflects a productivity term that is common across production units within a country A_i, a cost function that is common across countries $C_{t,k}$, and a term $\frac{1+\tau^i_s}{A^i_s}$ that shifts up the effective cost of skilled labor across all production units in a country. To a first order

$$d\log c^i_{t,k} = -d\log A_i + d\log w_u + \frac{\partial \log C_{t,k}}{\partial \log w_s}\left[\left(d\tau^i - d\log A^i_s\right) + d\log\left(w_s/w_u\right)\right],$$

which shows that the term $e^{\tau^i_s}/A^i_s$ increases unit costs in line with the skill intensity $\frac{\partial \log C_{t,k}}{\partial \log w_s}$.

On the surface, $e^{\tau^i_s}/A^i_s$ is similar to a skill-biased efficiency term in an aggregate production function, but in terms of interpretation, it is quite different. An aggregate skill-shifter can reflect any factor that shifts demand for skilled labor. In contrast, $e^{\tau^i_s}/A^i_s$ needs to be a factor that systematically increases unit production costs in skill intensive production processes both across and within industries. The canonical case for such a setup would be that skilled labor consistently produces some input X and that $e^{\tau^i_s}/A^i_s$ captures a country either having low productivity in making this input, or some wedge on its use. From this perspective, an interpretation of my findings would map model objects to real objects so that skilled labor consistently produces X, the term e^{τ_s}/A_s is low in poor countries, and e^{τ_s}/A_s is a factor that increases the cost of using/producing X.

6.2 Theories of skill-bias in the literature

Below, I discuss existing theories of skill bias and how they relate to the structure in (31).[17]

Demand-based theories Demand-based theories of skill biased development posit that economic growth increases the demand for skilled-labor by shifting expenditure towards output with more skill-intensive production. They include settings where income effects expand the non-agricultural sector (Caselli and Coleman II, 2001), expand skill-intensive sectors more generally (Comin et al., 2020), expand the the demand for high-quality goods (Jaimovich et al., 2020), or expand the production of goods where (skill-intensive) market production has a comparative advantage over home production (Buera and Kaboski, 2012). Since demand-based theories of skill bias concern consumer expenditures rather than production costs, they do not directly predict the unit cost patterns found in the data.[18]

Correlation between industry productivity and skill intensity Buera et al. (2022) propose a theory where skill-bias reflects a rapid productivity growth in the relatively skill-intensive industries. If the elasticity of substitution between high-skilled and low-skilled industries is greater than unity, this bias in Hicks-neutral productivity growth increases the relative demand for skilled labor. This theory correctly predicts the between-industry pattern found in section 3, but does not speak to the within-industry patterns found in section 4.

Skill-biased human capital quality Jones (2014a) and Bils et al. (2022) propose that skill-bias could reflect skill-biased differences in human capital quality. The idea is that if skilled workers provide more efficiency units of skilled labor in rich countries, this increases the demand of skilled workers provided that the elasticity of substitution between skilled and unskilled workers is more than unity. Quality differences in human capital *do* predict the patterns found in my trade data. This can be seen from the fact that they

[17] I cover the main theories of skill bias apart from endogenous technological differences (Caselli and Coleman, 2006), which does not have as clear an interaction with industry cost patterns.

[18] One demand-based theory that could potentially speak to my result but that has not been explored in the literature on aggregate skill bias is that scale economies and non-homotheticicities generate productivity advantage in skill intensive industries in rich countries, in line with the literature on home-market effects such as in Linder (1961) and Krugman (1980), later explored in Hanson and Xiang (2004), Fajgelbaum et al. (2011), and Costinot, Donaldson, Kyle, and Williams (2019).

satisfy the formal conditions in (31): if we interpret the generic input X to be skilled labor services and A_s to be the quality of skilled workers – so that one skilled worker produces A_s units of skilled services – then we have a theory where skilled workers produce X within every industry, and where poor countries have a lower productivity in converting skilled workers into X. The main issue with quality differences as an explanation has been that findings from migration data limits the potential size of such skill-biased quality differences, at least if they take a simple skill-augmenting form (Rossi, 2022).

6.3 Organizational capital and skill bias

Next, I consider a model where skill bias arises from barriers to the accumulation of organizational capital. To accommodate organizational capital accumulation, I use an infinite horizon setup, and to study the relationship between unit cost and skill intensity, I solve for the steady-state industry output price and skilled labor compensation share. Below, I outline the setup and show that the model implies reduced form costs in line with (31). In the appendix, I provide a more complete treatment, and show how to relate the dynamic setup to a static setup of the type used in section 2-4.

I consider an industry where firms produce a homogeneous output according to the production function

$$y_a = \ell_a^{1-\alpha} k_a^{\alpha}, \tag{32}$$

where ℓ_a is the number of production workers, k_a is the amount of organizational capital, and α is the output share of organizational capital, where the blue color indicates that this parameter varies across industries, and a indexes the age of a firm.[19] Furthermore, I assume that organizational capital is accumulated and maintained by skilled workers:

$$k_{a+1} = (1-\delta)k_a + s_a^I A_K \tag{33}$$

$$\pi k_a \leq s_a^M A_K, \tag{34}$$

where s_a^I and s_a^M is the number of skilled workers that the firm uses for investment and maintenance of organizational capital at age a. The term A_K captures the efficiency of using skilled labor for these purposes, with the red color capturing that this parameter varies across countries.

Firms are born without organizational capital, sell their output at a fixed price P and maximize discounted profits given an exogenous interest rate r and death risk τ. The parameter τ also varies across countries and

[19] In the appendix, I consider the case when final production also uses skilled labor.

is assumed to depend on a country's property rights. These assumptions imply that the firm solves

$$\max_{\{\ell_a, k_{a+1}, s_a^I, s_a^M\}} \sum_{a \geq 0} \left(\frac{1-\tau}{1+r}\right)^a \left[P y_a - w_u \ell_a - w_s (s_a^I + s_a^M)\right] \tag{35}$$

subject to $x_0 = 0$ and (32)-(34).

With free entry and zero profits, steady-state P satisfies

$$P = \left[\frac{(\pi + \delta + [r+\tau]/[1+r])}{A_K} \frac{w_s}{\alpha}\right]^{\alpha} \left[\frac{w_u}{1-\alpha}\right]^{1-\alpha} \tag{36}$$

Thus, in reduced form, the industry price level is isomorphic to that found in a static setup where industries have a Cobb-Douglas production function over skilled and unskilled labor, with a productivity shifter $\frac{A_K}{\pi+\delta+(r+\tau)/(1+r)}$ on skilled labor. The denominator reflects different costs associated with operating a unit of organizational capital: the maintenance cost πw_s, the replacement cost $\delta w_s x$, and the capital cost $\frac{(r+\tau)}{1+r} w_s$, where the capital cost is incurred on the initial outlay w_s, with τ capturing the additional capital cost coming from the risk of firm death.

Furthermore, if we write S and L for the aggregate number of skilled and production workers in the industry, it can be shown that their relative compensation satisfies

$$\frac{w_s}{w_u} \frac{S^I + S^M}{L} = \frac{\alpha}{1-\alpha} [1-\rho] \quad \rho \equiv \frac{r \times (1-\tau)/(1+r)}{\pi + \delta + \tau + r \times (1-\tau)/(1+r)}. \tag{37}$$

Thus, the relative compensation of skilled versus unskilled labor is proportional to $\alpha/(1-\alpha)$, precisely as with a static Cobb-Douglas production in which skilled labor enters directly into production with a share α. The term ρ captures the share of organizational capital expenditure which reflects holding costs of capital rather than direct payments to skilled labor, with a high ρ reducing the compensation to skilled labor.

Jointly, (36) and (37) imply that a low A_X and a high τ disproportionately raise prices in industries that have a high skilled labor compensation share. Thus, we can obtain a theory for between and within industry skill bias if we posit that A_X is low and τ is high in poor countries.

6.4 Interpretation and robustness

Interpretation of organizational capital and the role of skilled workers The first key assumption in (33)-(35) is that skilled workers primarily engage in accumulating and maintaining a set of durable assets, which I call "organizational capital".

Table 5: Top 10 skilled occupations in merchandise sector

ISCO-08 code	Occupation	Share
31	Science and engineering associate professionals	0.26
33	Business and administration associate professionals	0.16
21	Science and engineering professionals	0.14
24	Business and administration professionals	0.06
42	Customer services clerks	0.06
13	Production and specialised services managers	0.05
25	Information and communications technology professionals	0.05
41	General and keyboard clerks	0.05
12	Administrative and commercial managers	0.04
43	Numerical and material recording clerks	0.04

Note: The white-collar payroll share is calculated using the 2015 Occupational Employment Survey's information about number and average pay of workers by industry and occupation, concorded to ISIC rev 4 industry code from NAICS 2012 and to ISCO-08 from SOC-10 occupational codes. White-collar workers are defined as those with a 1-digit ISCO-08 code of 1 to 4. Merchandise industries are define as those with a two-digit ISIC revision 4 code 1 through 33.

In making this assumption, I conceive of skilled workers to be those with a one-digit occupational code between 1 and 4 in the ISCO-08 system. From table 5, which lists the ten largest skilled occupations in the merchandise producing sector, we see that skilled workers are dominated by engineers (40%), business and administrative professionals (22%), and different forms of clerks (15%) and managers (9%). For organizational capital, I conceive it to be what management science calls "structural and relational capital" (Edvinsson and Malone, 1997; Maddocks and Beaney, 2002), consisting in:

1. Capabilities, routines, methods, procedures and methodologies embedded in an organization,
2. Company's relationships with its customers, vendors, and other important constituencies,
3. Protected commercial rights such as patents, copyrights and trademarks.

Given these interpretations, my assumption is that the occupations listed in table 5 are mostly concerned with creating and keeping track of the durable assets described in bullet points 1-3. While a formal task analysis lies beyond the scope of the paper, we can informally evaluate this assumption by looking at what the workers in the different occupations do. Engineers and business professionals design and oversee production processes, work with accounting and sales, and develop and maintain company relationships. The clerks ensure that these activities are coordinated and documented. All of these activities arguably map naturally to the accumulation and maintenance of organizational capital.[20]

[20]Beyond the merchandise sector, the overall correspondence remains good, with the major exception being health care workers and teachers. These occupations constitute a large share of skilled workers by the ISCO definition, but their tasks arguably lie closer to production than to the accumulation of organizational capital. The treatment of this class of workers do

Interpretation of the parameters A_K and τ The second key assumption is that poor countries have a combination of low A_K and a high τ. To motivate these assumptions, I draw on the literature on firms and development. Starting with A_K, I conceive of it as capturing various factors that reduce the amount of organizational capital that can be accumulated and maintained by a given amount of skilled labor input. The literature has multiple explanations for why poor countries might have a low A_K. The most direct measure is the Cost of Doing Business Survey, which documents that poor countries have much longer lead times in regular business activities such as obtaining licenses (World Bank, 2008). This corresponds closely to a low A_K since skilled labor needs more time to achieve the same results. Other paperr have highlighted that poor countries suffer from bad contracting possibilities and limits on delegation, both of which reduce the productivity of complex organizations (Acemoglu, Antràs, and Helpman, 2007; Akcigit et al., 2021; Nunn, 2007). Finally, there may be a shortage of specialized human capital needed to run complex organizations (Hjort et al., 2021).

The term τ captures the risk that a firm goes out of business in any given period, and I conceive of it as capturing poor property rights. These bad property rights can either mean the risk of outright expropriation, but might also capture situations where ex post changes in rules make running a business more difficult, or where bribes and other payments are needed to keep the organization running. The key thing is that τ reduces the firm's ability to appropriate the returns to organizational capital. A large literature documents limits to property rights in poor countries (Acemoglu et al., 2001; Djankov, Glaeser, Perotti, and Shleifer, 2020).

Quantification. While it is beyond the scope of this paper to fully calibrate an organizational capital model to fit the trade data, below is a brief quantitative discussion which sketches how the model can be taken to data, which parameters that determine the outcome, and the magnitude of the quantities involved.

To do this, I assume that countries only differ in property rights τ which is some function of log GDP per worker $\log y$, income, in which case we obtain

$$\frac{\partial \log P}{\partial s \partial \log y} = \frac{1}{\pi + \delta + \frac{r+\tau}{1+r}} \frac{\partial \alpha}{\partial s} \frac{\partial \tau}{\partial \log y} \tag{38}$$

where s is the skilled labor payroll share.[21] Hence, the responsiveness to log unit costs from distortions in

not affect my empirical analysis, since they constitute a negligible share of workers in merchandise industries. However, their different task content does suggests a conceptual value in treating them as a separate category from other white-collar workers. Interestingly, the share of teachers and health care workers vary much less across countries than the share of other white-collar workers.

[21] I use that $\frac{w_s}{w_u} \frac{S^I + S^M}{L} = \frac{\alpha}{1-\alpha} [1 - \rho]$ with $\rho = \frac{r \times (1-\tau)/(1+r)}{\pi+\delta+\tau+r\times(1-\tau)/(1+r)}$. to derive that the skilled labor compensation share

the capital accumulation decision is the inverse of the user cost of capital. The intuition for this is that a low user cost of capital implies a high effective duration of the assets, which makes distortions to the carrying cost of capital more consequential.

To gauge the size of differences in τ needed to explain the trade findings given a set of parameter values, I use (38) together with $\delta = 0.06$, $\pi = 0.06$, $r = 0.05$, use $\frac{\partial \log P}{\partial s \partial \log y} = -0.38$ from table 3, and consider an approximation around $\tau = 0$ and $\alpha = 0.5$. The depreciation rate is close to the 5.8% used for organizational capital depreciation in Bhandari and McGrattan (2021), and r is close to the after-tax return of 5.15% on business capital calculated in Gomme, Ravikumar, and Rupert (2011). The term π has no equally clear data counterpart, and my calibration assumes that skilled labor splits equally between building and maintaining organizational capital. With these numbers, I obtain $\partial\alpha/\partial s \approx 1.03$ and

$$\frac{\partial \tau}{\partial \log y} = \frac{\frac{\partial \log P}{\partial s \partial \tau}}{\frac{1}{\pi+\delta+\frac{r}{1+r}}\frac{\partial \alpha}{\partial s}} = \frac{-0.31}{\frac{1}{0.06+0.06+0.05/(1+0.05)}\frac{\partial \alpha}{\partial s}} = \frac{-0.31}{6.16} \approx -0.050.$$

Hence, in this calibration, the expropriation risk needs to fall approximately 5 percentage points per log point of GDP to explain the data by τ alone, with approximately 20 percentage points lower expropriation risks in rich versus poor countries. While high values of π, δ and r increases the required differences in τ, the required differences are smaller if poor countries also have a lower A_K.[22]

Looking ahead, it is an interesting question how far barriers to organizational capital accumulation could go quantitatively in explaining my trade findings. Such an exercise would need a less stylized model as well as a more careful mapping of parameters to the data and considerations of other inputs that only unskilled and skilled labor. In particular, it would be important to go from an abstract τ to a realistic assessment of effective capital taxation levels in poor countries.

7 Conclusion

In this paper, I have used information from international trade data to explore the relationship between development and skill intensification. My main finding is that the aggregate skill-biased development doc-

satisfies $s = \alpha\frac{1-\rho}{1-\rho\alpha}$. This can be used to derive $\frac{\partial \alpha}{\partial s} = \left(\frac{\partial s}{\partial \alpha}\right)^{-1}$.

[22]A lower A_K could for example reflect skill-biased quality differences in human capital. To illustrate how this affects the calculation, we can note that Bils et al. (2022) estimate that one log point increase in y increases the human capital return of one year of schooling by 1.5 percentage points. On average, white collar workers in the world have approximately 5 more years of education than blue-collar workers, so this quality effect could explain 7.5 percentage points in difference per unit of GDP, reducing the required expropriation risk to $\frac{-0.31+0.075}{\frac{1}{0.06+0.06+0.05/(1+0.05)}\frac{\partial \alpha}{\partial s}} \approx 0.038$, that is a 3.8 percentage points decrease in expropriation risk per unit of log GDP per worker.

umented in the literature is accompanied by skill-biased development between and within industries: with development, unit costs fall faster in skill-intensive industries and in the skill-intensive tradable segment of industries. These patterns suggest a theory where some factor in poor countries shifts unit costs in a way that interacts with skill intensity. One such theory is that poor countries have bad property rights and high barriers to the accumulation of organizational capital, which can explain the unit cost patterns if the role of skilled workers is to accumulate and maintain such capital.

My findings show that a theory focusing on barriers to organizational capital accumulation can simultaneously address issues in the literatures on skill bias, firms and development, and international trade. Looking ahead, an interesting task is whether such a theory can help us understand other facets of development. There are two promising applications. First, it would be interesting to use an organizational capital perspective to study the consequences of bad property rights. While bad property rights are often cited as a culprit behind low development (Acemoglu et al., 2001), the quantitative importance of bad property rights is limited if they only work through physical capital accumulation, since differences in physical capital intensities account for little of income differences (Hsieh and Klenow, 2010). With an organizational capital perspective, bad property rights could be more important quantitatively, while at the same time help explain why poor countries have low skilled labor demand, small firms, and a comparative advantage in low-skill industries. A second interesting application would be to use an organizational capital perspective to study different aspects of structural change. For example, it is well-known that structural change into services is dominated by the growth of high-skill services such as finance and other business services (Buera and Kaboski, 2012). This pattern might be connected to barriers to organizational capital accumulation, since these industries work with the production and maintenance of organizational capital, and will be less important in economies with a lower organizational capital intensity.

References

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4, pages 1043–1171. Elsevier, 2011.

Daron Acemoglu, Simon Johnson, and James A Robinson. The colonial origins of comparative development: An empirical investigation. *American economic review*, 91(5):1369–1401, 2001.

Daron Acemoglu, Pol Antràs, and Elhanan Helpman. Contracts and Technology Adoption. *American Economic Review*, 97(3):916–943, jun 2007. ISSN 0002-8282. 10.1257/aer.97.3.916. URL https://www.aeaweb.org/articles.php?f=s&doi=10.1257/aer.97.3.916.

Ufuk Akcigit, Harun Alp, and Michael Peters. Lack of selection and limits to delegation: firm dynamics in developing countries. *American Economic Review*, 111(1):231–75, 2021.

Roberto Alvarez and Ricardo A López. Exporting and performance: evidence from chilean plants. *Canadian Journal of Economics/Revue canadienne d'économique*, 38(4):1384–1400, 2005.

Andrew Atkeson and Patrick J Kehoe. Modeling and measuring organization capital. *Journal of political Economy*, 113(5):1026–1053, 2005.

David Atkin, Arnaud Costinot, and Masao Fukui. Globalization and the ladder of development: Pushed to the top or held at the bottom? Technical report, National Bureau of Economic Research, 2021.

World Bank. *Doing business 2009*. The World Bank, 2008.

Pedro Bento and Diego Restuccia. On average establishment size across sectors and countries. *Journal of Monetary Economics*, 117:220–242, 2021.

Eli Berman, John Bound, and Stephen Machin. Implications of skill-biased technological change: international evidence. *The quarterly journal of economics*, 113(4):1245–1279, 1998.

Andrew B Bernard and J Bradford Jensen. Firm structure, multinationals, and manufacturing plant deaths. *The Review of Economics and Statistics*, 89(2):193–204, 2007.

Anmol Bhandari and Ellen R McGrattan. Sweat equity in us private business. *The Quarterly Journal of Economics*, 136(2):727–781, 2021.

Mark Bils, Bariş Kaymak, and Kai-Jie Wu. Labor substitutability among schooling groups. Technical report, National Bureau of Economic Research, 2022.

Timothy F Bresnahan, Erik Brynjolfsson, and Lorin M Hitt. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The quarterly journal of economics*, 117(1): 339–376, 2002.

Francisco J Buera and Joseph P Kaboski. The Rise of the Service Economy. *American Economic Review*, 102(6):2540–2569, oct 2012. ISSN 0002-8282. 10.1257/aer.102.6.2540. URL https://www.aeaweb.org/articles?id=10.1257/aer.102.6.2540.

Francisco J Buera, Joseph P Kaboski, and Yongseok Shin. Finance and Development: A Tale of Two Sectors. *American Economic Review*, 101(5):1964–2002, aug 2011. ISSN 0002-8282. 10.1257/aer.101.5.1964. URL https://www.aeaweb.org/articles.php?doi=10.1257/aer.101.5.1964.

Francisco J Buera, Joseph P Kaboski, Richard Rogerson, and Juan I Vizcaino. Skill-biased structural change. *The Review of Economic Studies*, 89(2):592–625, 2022.

Ariel Burstein and Jonathan Vogel. International trade, technology, and the skill premium. *Journal of Political Economy*, 125(5):1356–1412, 2017.

L. Caliendo and F. Parro. Estimates of the Trade and Welfare Effects of NAFTA. *The Review of Economic Studies*, 82(1):1–44, jan 2015. ISSN 0034-6527. 10.1093/restud/rdu035. URL `https://academic.oup.com/restud/article-lookup/doi/10.1093/restud/rdu035`.

Francesco Caselli. *Technology Differences Over Space and Time*. Princeton University Press, 2016.

Francesco Caselli and Wilbur John Coleman. The World Technology Frontier. *American Economic Review*, 96(3):499–522, jun 2006. ISSN 0002-8282. 10.1257/aer.96.3.499. URL `https://www.aeaweb.org/articles?id=10.1257/aer.96.3.499`.

Francesco Caselli and Wilbur John Coleman II. The us structural transformation and regional convergence: A reinterpretation. *Journal of political Economy*, 109(3):584–616, 2001.

Francesco Caselli, Jacopo Ponticelli, and Federico Rossi. A new dataset on mincerian returns. *Technology differences over space and time*, pages 108–118, 2014.

Alfred D Chandler Jr. *The Visible Hand: The Managerial Revolution in American Business*. Harvard University Press, 1977.

Davin Chor. Unpacking sources of comparative advantage: A quantitative approach. *Journal of International Economics*, 82(2):152–167, nov 2010. ISSN 00221996. 10.1016/j.jinteco.2010.07.004. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022199610000668`.

Antonio Ciccone and Giovanni Peri. Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950-1990. *Review of Economics and Statistics*, 87(4):652–663, nov 2005. ISSN 0034-6535. 10.1162/003465305775098233. URL `http://www.mitpressjournals.org/doi/abs/10.1162/003465305775098233`.

Diego A Comin, Ana Danieli, and Martí Mestieri. Income-driven labor market polarization. Technical report, National Bureau of Economic Research, 2020.

A. Costinot, D. Donaldson, and I. Komunjer. What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas. *The Review of Economic Studies*, 79(2):581–608, sep 2011. ISSN 0034-6527. 10.1093/restud/rdr033. URL `http://restud.oxfordjournals.org/content/early/2011/09/28/restud.rdr033.short`.

Arnaud Costinot. An elementary theory of comparative advantage. *Econometrica*, 77(4):1165–1192, 2009.

Arnaud Costinot, Dave Donaldson, Margaret Kyle, and Heidi Williams. The more we die, the more we sell? a simple test of the home-market effect. *The quarterly journal of economics*, 134(2):843–894, 2019.

Alejandro Cuñat and Marc J. Melitz. Volatility, Labor Market Flexibility, and the Pattern of Comparative Advantage. *Journal of the European Economic Association*, 10 (2):225–254, apr 2012. ISSN 15424766. 10.1111/j.1542-4774.2011.01038.x. URL http://doi.wiley.com/10.1111/j.1542-4774.2011.01038.x.

Donald R Davis and Jonathan I Dingel. The comparative advantage of cities. *Journal of International Economics*, 123:103291, 2020.

Simeon Djankov, Edward L Glaeser, Valeria Perotti, and Andrei Shleifer. Measuring property rights institutions. Technical report, National Bureau of Economic Research, 2020.

Jonathan Eaton and Samuel Kortum. Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779, sep 2002. ISSN 0012-9682. 10.1111/1468-0262.00352. URL http://doi.wiley.com/10.1111/1468-0262.00352.

Leif Edvinsson and Michael S Malone. *Intellectual capital*. HarperBusiness,, 1997.

Pablo Fajgelbaum, Gene M Grossman, and Elhanan Helpman. Income distribution, product quality, and international trade. *Journal of political Economy*, 119(4):721–765, 2011.

Robert C Feenstra, Robert Inklaar, and Marcel P Timmer. Penn world table version 9.1 (april 13, 2015), 2019.

Ana Cecilia Fieler. Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation. *Econometrica*, 79(4):1069–1101, 2011.

Guillaume Gaulier and Soledad Zignago. BACI: International Trade Database at the Product-Level (the 1994-2007 Version) (accessed January 20, 2017). *SSRN Electronic Journal*, oct 2010. ISSN 1556-5068. 10.2139/ssrn.1994500. URL http://www.ssrn.com/abstract=1994500.

Claudia Dale. Goldin and Lawrence F. Katz. *The race between education and technology*. Belknap Press of Harvard University Press, 2008. ISBN 9780674035300.

Paul Gomme, B Ravikumar, and Peter Rupert. The return to capital and the business cycle. *Review of Economic Dynamics*, 14(2):262–278, 2011.

Nezih Guner, Gustavo Ventura, and Yi Xu. Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, 11(4):721–744, oct 2008. ISSN 10942025. 10.1016/j.red.2008.01.005. URL http://www.sciencedirect.com/science/article/pii/S1094202508000070.

J. C. Hallak and P. K. Schott. Estimating Cross-Country Differences in Product Quality. *The Quarterly Journal of Economics*, 126(1):417–474, feb 2011. ISSN 0033-5533. 10.1093/qje/qjq003. URL http://qje.oxfordjournals.org/cgi/doi/10.1093/qje/qjq003.

Gordon H Hanson and Chong Xiang. The home-market effect and bilateral trade patterns. *American Economic Review*, 94(4):1108–1129, 2004.

Keith Head and Thierry Mayer. Gravity Equations: Workhorse,Toolkit, and Cookbook. In *Handbook of International Economics*, volume 4, pages 131–195. Elsevier B.V., 2014. ISBN 9780444543141. 10.1016/B978-0-444-54314-1.00003-3. URL http://dx.doi.org/10.1016/B978-0-444-54314-1.00003-3.

Lutz Hendricks and Todd Schoellman. Skilled labor productivity and cross-country income differences. Technical report, 2019.

César A Hidalgo and Ricardo Hausmann. The building blocks of economic complexity. *Proceedings of the national academy of sciences*, 106(26):10570–10575, 2009.

Jonas Hjort, Hannes Malmberg, and Todd Schoellman. The missing middle managers: Labor costs, firm structure, and development. 2021.

Thomas J Holmes and John J Stevens. An alternative theory of the plant size distribution, with geography and intra-and international trade. *Journal of Political Economy*, 122(2):369–421, 2014.

C.-T. Hsieh and P. J. Klenow. The Life Cycle of Plants in India and Mexico. *The Quarterly Journal of Economics*, 129(3):1035–1084, may 2014a. ISSN 0033-5533. 10.1093/qje/qju014. URL http://qje.oxfordjournals.org/content/129/3/1035.full.

Chang-Tai Hsieh and Peter J. Klenow. Development Accounting. *American Economic Journal: Macroeconomics*, 2(1):207–23, 2010. URL http://ideas.repec.org/a/aea/aejmac/v2y2010i1p207-23.html.

Chang-Tai Hsieh and Peter J Klenow. The life cycle of plants in india and mexico. *The Quarterly Journal of Economics*, 129(3):1035–1084, 2014b.

David Hummels and Peter J Klenow. The variety and quality of a nation's exports. *American economic review*, 95(3):704–723, 2005.

Nir Jaimovich, Sergio Rebelo, Arlene Wong, and Miao Ben Zhang. Trading up and the skill premium. *NBER Macroeconomics Annual*, 34(1):285–316, 2020.

Benjamin F. Jones. The Human Capital Stock: A Generalized Approach. *American Economic Review*, 104(11):3752–3777, 2014a. ISSN 0002-8282. 10.1257/aer.104.11.3752. URL http://pubs.aeaweb.org/doi/abs/10.1257/aer.104.11.3752.

Benjamin F Jones. The Knowledge Trap: Human Capital and Development Reconsidered. jun 2014b. URL http://www.nber.org/papers/w14138.

L. F. Katz and K. M. Murphy. Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The*

Quarterly Journal of Economics, 107(1):35–78, feb 1992. ISSN 0033-5533. 10.2307/2118323. URL `http://qje.oxfordjournals.org/cgi/doi/10.2307/2118323`.

Paul Krugman. Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5):950–959, 1980.

Simon Kuznets. *Modern Economic Growth: Rate, Structure and Spread*. Yale University Press, 1966.

Andrei A. Levchenko and Jing Zhang. The evolution of comparative advantage: Measurement and welfare implications. *Journal of Monetary Economics*, 78: 96–111, apr 2016. ISSN 03043932. 10.1016/j.jmoneco.2016.01.005. URL `http://www.sciencedirect.com/science/article/pii/S0304393216000088`.

Nelson Lind and Natalia Ramondo. Trade with correlation. Technical report, National Bureau of Economic Research, 2018.

Staffan Burenstam Linder. *An essay on trade and transformation*. Almqvist & Wiksell Stockholm, 1961.

Robert E. Lucas. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42, jul 1988. ISSN 03043932. 10.1016/0304-3932(88)90168-7. URL `http://linkinghub.elsevier.com/retrieve/pii/0304393288901687`.

Jon Maddocks and Mark Beaney. See the invisible and intangible. *Knowledge Management*, 16(1):16–17, 2002.

K. Manova. Credit Constraints, Heterogeneous Firms, and International Trade. *The Review of Economic Studies*, 80(2):711–744, apr 2013. ISSN 0034-6527. 10.1093/restud/rds036. URL `http://restud.oxfordjournals.org/lookup/doi/10.1093/restud/rds036`.

Marc J Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica*, 71(6):1695–1725, 2003.

Marc J Melitz and Stephen J Redding. Heterogeneous Firms and Trade *. *Handbook of International Economics*, 4:1–54, 2015. 10.1016/B978-0-444-54314-1.00001-X.

Peter M Morrow and Daniel Trefler. How do endowments determine trade? quantifying the output mix, factor price and skill-biased technology channels (forthcoming). *Journal of International Economics*, 2022.

Douglass C North. Institutions, transaction costs and economic growth. *Economic inquiry*, 25(3):419–428, 1987.

Nathan Nunn. Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade. *The Quarterly Journal of Economics*, 122(2):569–600, 2007. URL http://ideas.repec.org/a/tpr/qjecon/v122y2007i2p569-600.html.

Nathan Nunn and Daniel Trefler. Chapter 5: Domestic Institutions as a Source of Comparative Advantage. In *Handbook of International Economics*, volume 4, pages 263–315. 2015. URL http://www.sciencedirect.com/science/article/pii/B9780444543141000057.

Edward C Prescott and Michael Visscher. Organization capital. *Journal of political Economy*, 88(3):446–461, 1980.

John Romalis. Factor Proportions and the Structure of Commodity Trade. *American Economic Review*, 94 (1):67–97, mar 2004.

Federico Rossi. The relative efficiency of skilled labor across countries: Measurement and interpretation. *American Economic Review*, 112(1):235–66, 2022.

Ulrich Schetter. A structural ranking of economic complexity. *CID Research Fellow & Graduate Student Working Paper*, (119), 2019.

Theodore William Schultz. *Investing in people: The economics of population quality.* Univ of California Press, 1982.

Robert Shimer and Lones Smith. Assortative matching and search. *Econometrica*, 68(2):343–369, 2000.

J. M. C. Santos Silva and Silvana Tenreyro. The Log of Gravity. *Review of Economics and Statistics*, 88(4):641–658, nov 2006. ISSN 0034-6535. 10.1162/rest.88.4.641. URL http://www.mitpressjournals.org/doi/10.1162/rest.88.4.641.

Ina Simonovska and Michael E. Waugh. The elasticity of trade: Estimates and evidence. *Journal of International Economics*, 92(1):34–50, jan 2014. ISSN 00221996. 10.1016/j.jinteco.2013.10.001. URL http://www.sciencedirect.com/science/article/pii/S0022199613000986.

Marcel Timmer, Gerard Ypma, and Bart van van Ark. PPPs for Industry Output: A New Dataset for International Comparisons. *GGDC Research Memorandum*, 2007. URL https://ideas.repec.org/p/gro/rugggd/gd-82.html.

Donald M Topkis. *Supermodularity and complementarity.* Princeton university press, 1998.

James R Tybout. Manufacturing firms in developing countries: How well do they do, and why? *Journal of Economic literature*, 38(1):11–44, 2000.

Eric A. Verhoogen. Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector [*]. *Quarterly Journal of Economics*, 123 (2):489–530, may 2008. ISSN 0033-5533. 10.1162/qjec.2008.123.2.489. URL http://qje.oxfordjournals.org/lookup/doi/10.1162/qjec.2008.123.2.489.

Michael E Waugh. International Trade and Income Differences. *American Economic Review*, 100(5):2093–2124, dec 2010. ISSN 0002-8282. 10.1257/aer.100.5.2093. URL `https://www-aeaweb-org.ezp.sub.su.se/articles.php?doi=10.1257/aer.100.5.2093`.

Appendix

A Appendix to section 2

A.1 Full model specification

There are I countries, K industries, two segments NT and T within each industry and 2 production factors. The different countries are characterized by a set of trade trade costs $\{\tau_k^{ij}\}_{k\in\{1,\ldots,K\}}^{i,j\in\{1,\ldots,I\}}$, factor supplies $\{Z_f^i\}_{f\in\{1,2\}}^{i\in\{1,\ldots,I\}}$, and deterministic unit cost functions $\{c_{k,NT}^i(\cdot), c_{k,T}^i(\cdot)\}_{k\in\{1,\ldots,K\}}^{i\in\{1,\ldots,I\}}$ which are functions of $\mathbf{w}^i = (w_1^i, w_2^i)$. I write $L^i = Z_1^i + Z_2^i$ for the aggregate labor supply of a country, with Z_1^i being unskilled labor and Z_2^i being skilled labor. Preferences are given by (1), and I write $E(P_1^i, \ldots, P_K^i; U)$ for the corresponding expenditure function. Technology is defined by the cost functions (2), (3), (4), and (5), which together express the unit cost of an industry aggregator in terms of the deterministic unit cost functions alone:

$$P_k^i(c_{k,NT}^i(\mathbf{w}^i), \{c_{k,T}^j(\mathbf{w}^j)\}) = \gamma \left(\int_0^1 \min\left\{ \frac{c_{k,NT}^i(\mathbf{w}^i)}{z_{k,NT}^i(\omega^k)}, \min_j \frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k)} \right\}^{1-\sigma} d\omega^k \right)^{\frac{1}{1-\sigma}}, \tag{39}$$

By writing the unit cost function with these arguments, we show that the model can be expressed as countries producing two intermediate inputs in every industry with unit costs $c_{k,NT}^j(\mathbf{w}^i)$ and $c_{k,T}^j(\mathbf{w}^i)$. For each variety, the non-traded intermediate inputs are transformed into local non-traded varieties with productivity $z_{k,NT}(\omega^k)$ and into traded varieties in country i with productivities $z_{k,T}(\omega^k)/\tau_k^{ij}$. By integrating out the varieties, we obtain that the price index in each country is a function of the deterministic unit costs functions alone.

An equilibrium consists of a set of utility levels U^i, a set of factor prices $\{w_1^i, w_2^i\}$ and a set of industry quantities and prices $\{Q_1^i, \ldots, Q_K^i\}$ and $\{P_1^i, \ldots, P_K^i\}$ such that the P_k^i's satisfy (39) and

$$Q_k^i = \frac{\partial E(P_1^i, \ldots, P_K^i; U^i)}{\partial P_k^i} \quad i \in \{1, \ldots, I\}, k \in \{1, \ldots, K\}$$
$$Z_f^j = \sum_k \left\{ Q_k^j \frac{\partial P_k^j}{\partial c_{k,NT}^j(\mathbf{w}^j)} \frac{\partial c_{k,NT}^j}{\partial w_f^j} + \sum_i Q_k^i \frac{\partial P_k^i}{\partial c_{k,T}^j(\mathbf{w}^i)} \frac{\partial c_{k,T}^j}{\partial w_f^j} \right\}, \quad i \in \{1, \ldots, I\}, f \in \{1, 2\}.$$

The first expression uses that quantity demanded is the gradient of the expenditure function. The second expression captures the total factor use of f in j across all industries k and destinations i. The expression relies on Shepherd's lemma to express factor use per unit output by the derivative of P_k^j with respect to factor prices. In doing so, I use that the cost function is almost everywhere differentiable as a function of deterministic unit costs, with all sourcing decisions implicitly defined in this derivative (because the derivative is zero almost surely for all varieties where consumption is zero). The first term captures the use of factor f in non-tradables, and is given by local industry consumption Q_k^j, times the use of the non-traded intermediate per unit of (j, k) output, times the use of factor f per unit of non-tradables. The second term captures the use of factor f in tradables, and is given by consumption in the destination country i, times the use of the country j traded intermediate in country i, times the use of factor f in producing a unit of traded intermediates.

A.2 Gravity equation

Given an equilibrium, the value of bilateral trade flows satisfy

$$X_k^{ij} = X_k^i \left(\frac{\partial \log P_k^i}{\partial \log c_{k,NT}^j} + \frac{\partial \log P_k^i}{\partial \log c_{k,T}^j} \right) \tag{40}$$

where $X_k^i \equiv Q_k^i P_k^i$ is the value of industry k consumption in country i. To obtain this expression, I treat expenditure in industry k as an output of an industry with a unit cost function P_k^i, and use that use that for constant returns to scale production functions, expenditure shares are given by the elasticities of the unit cost function with respect to costs. Hence, the expression says that the value of total shipments from j to i in k is the value of total consumption in i of k, times the share spent on $j's$ non-tradable and tradable output respectively.

Writing $I_{k,T}^{ij}(\omega^k) = \mathbb{I}\left[\frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k)} < \min \left\{ \frac{c_{k,NT}^i(\mathbf{w}^i)}{z_{k,NT}^i(\omega^k)}, \min_j \frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k)} \right\} \right]$ for the indicator function that the tradable output of j is the cheapest supplier of i in industry k, we obtain

$$
\begin{aligned}
\frac{\partial \log P_k^i}{\partial \log c_{k,T}^j} &= \frac{1}{1-\sigma} \frac{\partial}{\partial \log c_{k,t}^j} \log \left(\int_0^1 \left(\min \left\{ \frac{c_{k,NT}^i(\mathbf{w}^i)}{z_{k,NT}^i(\omega^k)}, \min_j \frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k)} \right\}^{1-\sigma} \right) d\omega^k \right) \\
&= \frac{\int_0^1 \left(\frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k)} \right)^{1-\sigma} I_{k,T}^{ij}(\omega^k) d\omega^k}{\left(P_k^i\right)^{1-\sigma}}
\end{aligned}
$$

where we use that the partial derivative with respect to $c_{k,T}^j$ is only non-zero if country j's tradables are the cheapest option. Using that

$$
\tilde{z}_{k,NT}^j \equiv z_{k,NT}^j / [a_{NT} G_{NT}(L^j)]^{1/\theta}, \quad \tilde{z}_{k,T}^j = z_{k,T}^j / [a_T G_T(L^j)]^{1/\theta}
$$

have Frechét distributions with scale parameter 1. In this case, the indicator function can be expressed as $I_{k,T}^{ij}(\omega^k) = \mathbb{I}\left[\frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{\tilde{z}_{k,T}^j(\omega^k)[a_T G_T(L^j)]^{1/\theta}} < \min \left\{ \frac{c_{k,NT}^i(\mathbf{w}^i)}{\tilde{z}_{k,NT}^i(\omega^k)[a_{NT} G_{NT}(L^j)]^{1/\theta}}, \min_j \frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{\tilde{z}_{k,T}^j(\omega^k)[a_T G_T(L^j)]^{1/\theta}} \right\} \right]$, and standard extreme value results imply that the distribution of $\min \left\{ \frac{c_{k,NT}^i(\mathbf{w}^i)}{z_{k,NT}^i(\omega^k) G_{NT}(L^j)^{1/\theta}}, \min_j \frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k) G_T(L^j)^{1/\theta}} \right\}$ is the same as the distribution of $\frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k)}$ conditional on this being the minimum. Moreover, since $\left(P_i^k\right)^{1-\sigma} = \mathbb{E} \min \left\{ \frac{c_{k,NT}^i(\mathbf{w}^i)}{\tilde{z}_{k,NT}^i(\omega^k)[a_{NT} G_{NT}(L^j)]^{1/\theta}}, \min_j \frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{\tilde{z}_{k,T}^j(\omega^k)[a_T G_T(L^j)]^{1/\theta}} \right\}^{1-\sigma}$, we obtain that

$$
\begin{aligned}
\frac{\partial \log P_k^i}{\partial \log c_{k,T}^j} &= \mathbb{P}\left(\frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k) G_T(L^j)^{1/\theta}} < \min \left\{ \frac{c_{k,NT}^i(\mathbf{w}^i)}{z_{k,NT}^i(\omega^k) G_{NT}(L^j)^{1/\theta}}, \min_j \frac{\tau_k^{ij} c_{k,T}^j(\mathbf{w}^j)}{z_{k,T}^j(\omega^k) G_T(L^j)^{1/\theta}} \right\} \right) \\
&= \frac{a_T G_T(L^j)(\tau_k^{ij} c_{k,t}^j(\mathbf{w}^j))^{-\theta}}{a_T G_{NT}(L^i)(c_{k,t}^i(\mathbf{w}^i))^{-\theta} + \sum_j a_T G_T(L^j)(c_{k,t}^j(\mathbf{w}^j))^{-\theta}}. \qquad (41)
\end{aligned}
$$

By a similar logic, we obtain that the expenditure share on non-tradables is

$$
\frac{\partial \log P_k^i}{\partial \log c_{k,NT}^j} = \begin{cases} 0 & \text{if } i \neq j \\ \frac{a_{NT} G_{NT}(L^i)(c_{k,NT}^i(\mathbf{w}^i))^{-\theta}}{a_T G_{NT}(L^i)(c_{k,t}^i(\mathbf{w}^i))^{-\theta} + \sum_j a_T G_T(L^j)(c_{k,T}^j(\mathbf{w}^j))^{-\theta}} & \text{if } i = j. \end{cases} \qquad (42)
$$

Inserting (41) and (42) into (40) yields the gravity equation (6) in the main text.

B Appendix to section 3

B.1 Proof of proposition 1

To derive the proposition, we start from the logarithm of the gravity equation (6) when $i \neq j$, and note that we can express trade flows as placed by the $\bullet$-sign. The expression implies that we can write

$$
\begin{aligned}
\log X_k^{ij} =& \log a_K + \log G_T(L^j) - \theta \log \tau_k^{ij} - \theta \log c_{k,T}^j - \log \Phi_k^i(\{\varepsilon_k^{ij}\}) \\
=& \underbrace{-\theta\tau^{ij} + \log G_T(L^j) - \theta \log c_{\bullet,T}^j - \theta \log \tau_\bullet^i - \log \Phi_\bullet^i(\{\varepsilon_k^{ij}\})}_{\delta^{ij}} \\
& \underbrace{\log a_K - \theta \log \frac{c_{k,T}^\bullet}{c_{\bullet,T}^\bullet} - \theta \log \tau_k^\bullet / \tau_\bullet^\bullet - \log \Phi_k^\bullet(\{\varepsilon_k^{ij}\}) / \Phi_\bullet^\bullet(\{\varepsilon_k^{ij}\})}_{\alpha_k} \\
& \underbrace{\left(- \log \frac{\Phi_k^i(\{\varepsilon_k^{ij}\})/\Phi_\bullet^i(\{\varepsilon_k^{ij}\})}{\Phi_k^\bullet(\{\varepsilon_k^{ij}\})/\Phi_\bullet^\bullet(\{\varepsilon_k^{ij}\})} - \theta \log \frac{\tau_k^i/\tau_\bullet^i}{\tau_k^\bullet/\tau_\bullet^\bullet} + \varepsilon_k^{ij} \frac{\operatorname{Cov}\left(\log \frac{\Phi_k^i(\{\varepsilon_k^{ij}\})/\Phi_\bullet^i(\{\varepsilon_k^{ij}\})}{\Phi_k^\bullet(\{\varepsilon_k^{ij}\})/\Phi_\bullet^\bullet(\{\varepsilon_k^{ij}\})}, \varepsilon_k^{ij}\right)}{\operatorname{Var} \varepsilon_k^{ij}} \right)}_{\mu_k^i} + \\
& \underbrace{\left(-\theta \log \frac{c_{k,T}^j / c_{\cdot,T}^j}{c_{k,T}^\cdot / c_{\cdot,T}^\cdot} \right)}_{\gamma_k^j} + \\
& \sum_m (-\theta) \beta_{m,k} d_m^{ij} + \\
& \left(-\theta \varepsilon_k^{ij} - \varepsilon_k^{ij} \frac{\operatorname{Cov}\left(\log \frac{\Phi_k^i(\{\varepsilon_k^{ij}\})/\Phi_\bullet^i(\{\varepsilon_k^{ij}\})}{\Phi_k^\bullet(\{\varepsilon_k^{ij}\})/\Phi_\bullet^\bullet(\{\varepsilon_k^{ij}\})}, \varepsilon_k^{ij}\right)}{\operatorname{Var} \varepsilon_k^{ij}} \right),
\end{aligned}
\tag{43}
$$

Suppose that where $\bullet$ denotes the geometric mean of a variable taken over the set of indices that are referred to by $\bullet$. This formulation implies that we can express trade flows as

$$
\log X_k^{ij} = \delta^{ij} + \alpha_k + \mu_k^i + \gamma_k^j + \sum_m \tilde{\beta}_{k,m} d_m^{ij} - \theta \varepsilon_k^{ij}, \tag{44}
$$

where the expressions for the parameters are given in (43).

$$
\sum_k a_k = \sum_j \mu_k^j = \sum_k \mu_k^j = \sum_j \gamma_k^j = \sum_k \gamma_k^j = 0, \tag{45}
$$

with ε_k^{ij} being uncorrelated with the regressors (it is uncorrelated with the fixed effect given the assumption of zero mean for each combination of i, j, k, and uncorrelated with the gravity proxies by assumption). To see that (45) hold, I note that

$$\sum_i \log c^{\bullet}_{k,T} = I \log \exp\left(\frac{1}{I}\sum_i \log c^i_{k,T}\right) = \sum_i \log c^i_{k,T}$$

$$\sum_k \log c^i_{\bullet,T} = K \exp\left(\frac{1}{K}\sum_k c^i_{k,T}\right) = \sum_k \log c^i_{k,T}$$

$$\sum_k \log c^{\bullet}_{\bullet} = K \exp\left(\frac{1}{KI}\sum_{i,k} c^i_{k,T}\right) = \frac{1}{I}\sum_{i,k} \log c^i_{k,T} = K \log c^{\bullet}_{\bullet,T},$$

$$\sum_i \log c^{\bullet}_{\bullet} = I \exp\left(\frac{1}{KI}\sum_{i,k} c^i_{k,T}\right) = \frac{1}{K}\sum_{i,k} \log c^i_{k,T} = I \log c^{\bullet}_{\bullet,T}$$

Using these equalities, (45) follows from (43). Given these restrictions, $\mathbb{E}\hat{\gamma}_k^j = -\theta \log \frac{c^j_{k,T}/c^j_{\cdot,T}}{c^{\cdot}_{k,T}/c^{\cdot}_{\cdot,T}}$ follows from the unbiasedness of OLS.[23]

C Appendix to section 3

C.1 Concordance description

To generate concordances and map data across coding systems, I create a general mathematical framework to treat the problem, described below. The basic building block of the concordance system is a many-to-many concordance between coding systems A and B where I have weights on both A and B. I call such concordances two-weighted concordances. An example of such a concordance is provided in Table 6. In Table 6, note that each code in system A can be converted to multiple B codes (in this example, code 2 in System A maps to both code "b" and "c" in System B). The converse is also true: both codes 4 and 5 map to code "e". The weights code how important the respective industries are. This could, for example, be the total value of shipments, total trade value, etc. Notice that the weights are both on A and B, and that they are constant whenever they stand for the same industry. I can define this mathematically as there

[23]The one worry is that the regression matrix does not have full rank, so that the estimator $\hat{\gamma}_k^j$ is not unique. However, we can show that the restrictions (45)are sufficient for a unique estimator. Indeed, posit an arbitrary null element (using that the distance proxies are not multi-collinear with the fixed effects):

$$0 = \tilde{\delta}^{ij} + a_k + \tilde{\mu}_k^i + \tilde{\gamma}_k^j$$

for all i, j, k. Summing over k, we obtain $\tilde{\delta}^{ij} = 0$ for all i, j, and summing over (i, j) implies that $\tilde{a}_k = 0$ for all k. Given $\tilde{\delta}^{ij} = a_k = 0$, we can sum over i and j respectively to obtain $\tilde{\mu}_k^i = \tilde{\gamma}_k^j = 0$. Hence, there are no elements in the null space of the regression matrix that satisfy (45).

Table 6: Example concordance table

A	B	A_w	B_w
1	a	10	70
2	b	20	50
2	c	20	100
3	c	15	40
4	d	5	70
5	d	25	70
6	e	30	90

being two sets A and B, with measures w_A and w_B giving the mass on each code, and a concordance being a correspondence

$$\phi : A \rightrightarrows B.$$

I will write results in terms of this mathematical definition, but also in terms of examples to show the working of the system. I will go through three operations relating to two-weighted concordances:

1. How to transform quantity variables such as total industry sales using a two-weighted concordance

2. How to transform property variables such as capital share using a two-weighted concordance

3. How to create a two-weighted concordance using an unweighted concordance and a weighting scheme for one of the variables (e.g. when I want to create a two-weighted concordance between HS and SITC and only have total trade in HS codes).

Starting with quantity variables, suppose that I have export values denoted in industry code A. I then want to allocate it across different codes in industry code B given a weighting scheme on B. In this case, for each element $a \in A$, I allocate the export values in industry a across industries $b \in B$ in proportion to their weights w_b. The quantity attributed to element $b \in B$ is then the sum of the contributions from all elements in A to b. I can write this in terms of the mathematical representation Φ as well, together with the weights μ_A and μ_B. If

$$f_A : A \to \mathbb{R}$$

is an arbitrary quantity measure on A I convert it to B by

$$f_B(b) = \sum_{a \in \Phi^{-1}(b)} f_A(a) \times \frac{\mu_B(b)}{\sum_{b' \in \Phi(a)} \mu_B(b')}.$$

The situation is different when I have so-called property variables, for example capital share, skill share or other industry-level properties. The difference can be illustrated with an example. In the previous part, I considered the problem of mapping trade data from A to B. Then, the reasonable thing is to split it up the value a across $b \in \Phi(a)$ according to the weights w_b. However, suppose that I want to map the capital share from a to b. Then, we should not split up the capital share across $b \in \Phi(a)$. If b and b' have the same pre-image a, they should have the same capital share as a. Thus, property variables translate across coding systems in a fundamentally different way from quantity variables. I define the transformation scheme for property variables by saying that for each code $b \in B$ in the target system, I define its property as a weighted average of the properties that its pre-images $a \in A$, where I use the weights on A as a weighting scheme. For example, in our example concordance, I would attribute c a property which is the weighted average of 2,3 in System A, using the measures $\mu_A(\{2\}) = 20$ and $\mu_A(\{3\}) = 15$ as weights. More formally, if I have a property measure

$$g_A : A \to \mathbb{R}$$

defined on A, then I translate it to B using ϕ by the equation

$$g_B(b) = \frac{\sum_{a \in \phi^{-1}(y)} g_A(a)\mu_A(a)}{\sum_{a \in \phi^{-1}(b)} \mu_A(a)}.$$

Above I defined how you translate between different coordinate systems if you have a two-sided weighted concordance. However, sometimes I only have a one-sided concordance. For example, if I have total trade data in HS 2007 six-digit and want to create a concordance between HS 2007 6-digit and NAICS 2007 it might be that I do not have data to create a natural weighting scheme for the NAICS 2007 coding scheme. For this case, I have a procedure to create a two-sided weighted concordance from a one-sided weighted concordance. It is quite similar to the quantity transformation above. Suppose that I have a concordance ϕ and a measure μ_A on A and want to create a measure μ_B on B. Then I define the measure on B as.

$$\mu_B(b) = \sum_{a \in \phi^{-1}(b)} \frac{\mu_A(a)}{|\phi^{-1}(\phi(a))|}.$$

That is, I split the weights on $a \in A$ equally on all $b \in B$ to which a maps.

C.2 Robustness checks

This section presents the different robustness checks to the results in section 3. The main results are presented in table 7.

Alternative unit cost estimates In this section, I consider how robust the findings in table 3 are to variations in how unit costs are backed out. In all cases, I run a regression

$$\log\left(\frac{c_k^i/\bar{c}^i}{\bar{c}_k/\bar{c}}\right) = \eta \times \left(\log y_i \times s^k\right),$$

with different assumptions on how the left-hand side is estimated. I consider four variations: using a Poisson Pseudo-Maximum Likelihood (PPML) estimator instead of OLS, allowing for different trade elasticities θ across industries, considering different subsets of destinations in the regression, and using unit cost data directly instead of unit costs backed out from gravity regressions.

To explore the effects of using different subsamples, I split the sample into low and high income countries, and run the regression using only low or high income destinations. The results are very similar to the -0.31 found for the full set of destinations. To obtain the PPML estimator for the origin-industry fixed effects, I use the Poisson regression from the fixglm-function in the fixest package from R; the results are mildly weaker than for the OLS specification, but similar in magnitude.

To have different trade elasticities, I need to make stronger assumptions about the structure of trade costs. In particular, if the gravity equation satisfies

$$\log X_k^{ij} = \log X_k^i a_{k,T} - \theta_k \log[G_T^k(L^j)c_{k,T}^j] - \log \Phi_k^i - \theta_k \log \tau_k^{ij},$$

the variation across industries in θ_k means that one cannot remove a common component of bilateral trade costs with a bilateral fixed effect. Hence, I assume that $\log \tau_k^{ij} = \tau_k^i + \sum_m \beta_{m,k} d_m^{ij} + \epsilon_k^{ij}$, in which case the origin-fixed effect estimates $\theta_k \log G_T^k(L^j)c_{k,T}^j$ if we have origin and destination fixed effects together with a battery of trade cost parameters with industry specific effects. Writing $\hat{\gamma}_k^j$ for the estimated origin fixed effects, I run the regression

$$-\frac{\hat{\gamma}_k^j}{\theta_k} = \alpha^i + \xi_k + \xi_k \log L^i + \eta(\log y^i \times s_k^i),$$

where I include an industry-specific slope on country size to capture that the scale effect G_T^k might differ across industries. For trade elasticities, I use the estimates from Caliendo and Parro (2015) which estimates elasticities on a two-digit level ISIC revision 3 level. I exclude estimates that are smaller than 0.5, and adjust so that the median elasticity is $\bar{\theta}_k = 6$ to be comparable with my uniform elasticity assumptions. We see that the estimated interaction term is similar to that in the main model. [24]

For unit cost comparisons, I use the GGDC 35-industry database which provides estimates of gross output prices in 42 major economies. I concord their industrial classification to ISIC rev 4 2-digit, and calculate the US payroll share of skilled workers by industry, and run the regression

$$\log c_k^i = a^i + b_k + \eta \log y_i \times \log s^k + \epsilon_k^i.$$

Restricting my attention to merchandise industries for comparability with the trade estimates, I find a similar effect in the GGDC data as in the trade data – -0.43 vs -0.31 in the trade model. The fit is surprisingly good in that unit costs exploiting a completely different source of variation generates such a similar result.

Alternative standard errors The fixed effects estimator used for (8) does not supply standard errors for the estimated fixed effects, which means that the standard errors in the interaction regression (10) does not take into account the uncertainty in the underlying estimates. Due to this, I also consider an alternative estimation strategy where I substitute in (10) into (8) and estimate the interaction term directly from the trade regression, clustering standard errors on the origin-industry level to take into that the source of variation occurs on this level. We see that this approach yields a very similar result to my main approach.

Using aggregate exports As a complement to the fixed effect analysis, the last row of table 7 reports the results of the following regression

$$\log X_k^j = a^j + b_k + \eta \times [(-\theta) \times \log y^j \times s_k] + \epsilon_k^j,$$

where X_k^j is aggregate export values. Using $\theta = 6$, we find that there is a negative and highly significant interaction also in this specification, which is of the same approximate magnitude as the main estimation. This specification illustrates that similar magnitudes can be obtained from a relatively reduced form analysis of the trade data as that obtained via a more theory-consistent, fixed-effect based, gravity estimation.

[24] Only code 28 – machinery – has a positive elasticity smaller than 0.5 (it has $\theta_k = 0.22$). It is a skill-intensive industry and including it makes my interaction estimates much larger, because such a low elasticity means that implied unit costs differences from trade flows become very large, imply a very negative interaction between GDP and skill intensity to fit unit cost variations.

Table 7: Robustness results

	Poor dest.	Rich dest.	PPML	Different θ	Unit costs	Clustered	Total exports
Skill int. x log(y)	-0.315	-0.318	-0.257	-0.205	-0.425	-0.377	-0.238
	(0.017)	(0.016)	(0.015)	(0.057)	(0.118)	(0.015)	(0.015)
Num.Obs.	18 516	20 292	20 492	4236	666	893 003	21 017
R2	0.087	0.055	0.245	0.618	0.640	0.696	0.748

Other industry features Table 8 includes other industry features in the regression (10) to test whether the interaction between skill and income, running the regression

$$-\frac{\hat{\gamma}_k^i}{\theta} = a^i + b_k + \eta \times (\log y^i \times s_k) + \sum_m \xi_m \log y^i \times f_{m,k} + \epsilon_k^i$$

for other industry features $f_{m,k}$. I perform the regression using the the compensation share of capital, the compensation share of intermediate inputs, and the share of inputs that are not traded on an exchange or reference-priced in trade journals (this measure was introduced by Nunn, 2007 as a proxy for the dependence of an industry on contracting institutions). We note that including these other industry features has a negligible impact on the interaction term of interest.

Table 8: Interaction regression with industry controls

	Dependent variable:
	(- Origin-industry f.e.) / Trade elasticity
$\log(y)$ × Skill intensity	-0.356^{***}
	(0.015)
$\log(y)$ × Capital intensity	0.360^{***}
	(0.026)
$\log(y)$ × Intermediate intensity	-0.208^{***}
	(0.028)
$\log(y)$ × Contracting intensity	-0.070^{*}
	(0.037)
Observations	20,492
R^2	0.279
Adjusted R^2	0.268
Residual Std. Error	0.316 (df = 20192)
F Statistic	26.138^{***} (df = 299; 20192)
Note:	*p<0.1; **p<0.05; ***p<0.01

Service trade To analyze trade in services, I draw on the OECD-WTO Balanced Trade in Services (BaTIS) dataset. This dataset provides harmonized estimates of trade flows in services for 206 reporters for the 12 main EBOPS 2010 service categories for the period 2005-2019. The twelve main categories are listed in table 9. I measure comparative advantage for a country in terms of its skill-intensive services relative total exports, where I define the skill-intensive categories as "Insurance and pension services", "Financial services", and Charges for the use of intellectual property n.i.e.". The results are displayed in figure 9, where we see that the share of service exports in skill-intensive industries rises dramatically with aggregate income: the share of skill-intensive services increases almost two orders of magnitude as we go from the poorest to the richest countries.[25]

Table 9: EBOPS 2010: 12 main categories

Description
Manufacturing services on physical inputs owned by others
Maintenance and repair services n.i.e.
Transport
Travel
Construction
Insurance and pension services
Financial services
Charges for the use of intellectual property n.i.e.
Telecommunications, computer, and information services
Other business services
Personal, cultural, and recreational services
Government goods and services n.i.e.

D Appendix to section 4

D.1 Small open economy equilibrium

Starting from the baseline model with only one type of labor and one industry, a small open economy equilibrium for a country i is defined relative to an exogenous wage vector $\{w_{-i}\}$ for the other economies in the world. To capture that the economy is small, we assume that $L^j \gg L^i$ for $j \neq i$, and we treat $\{w^{-i}\}$ as fixed; that is, we do not demand that $\{w_{-i}\}$ clears the world market. We also assume that τ^{ij} is large in the sense that domestic attractiveness of goods $a_T G_T(L^i)(w^i/A^T)^{-\theta}$ and $G_L(L^i)a_{NT}(w^i/A^i_{NT})^{-\theta}$ are not trivial in size relative to $G_T(L^j)\left(\tau^{ij}\right)^{-\theta}$.

[25] The EBOPS categories are final use categories that do not map neatly into industry categories. I selected the financial services and intellectual properties as unambiguously skill-intensive industries. The results are similar if I also include "Telecommunications, computer, and information services" as a skill-intensive service industry.

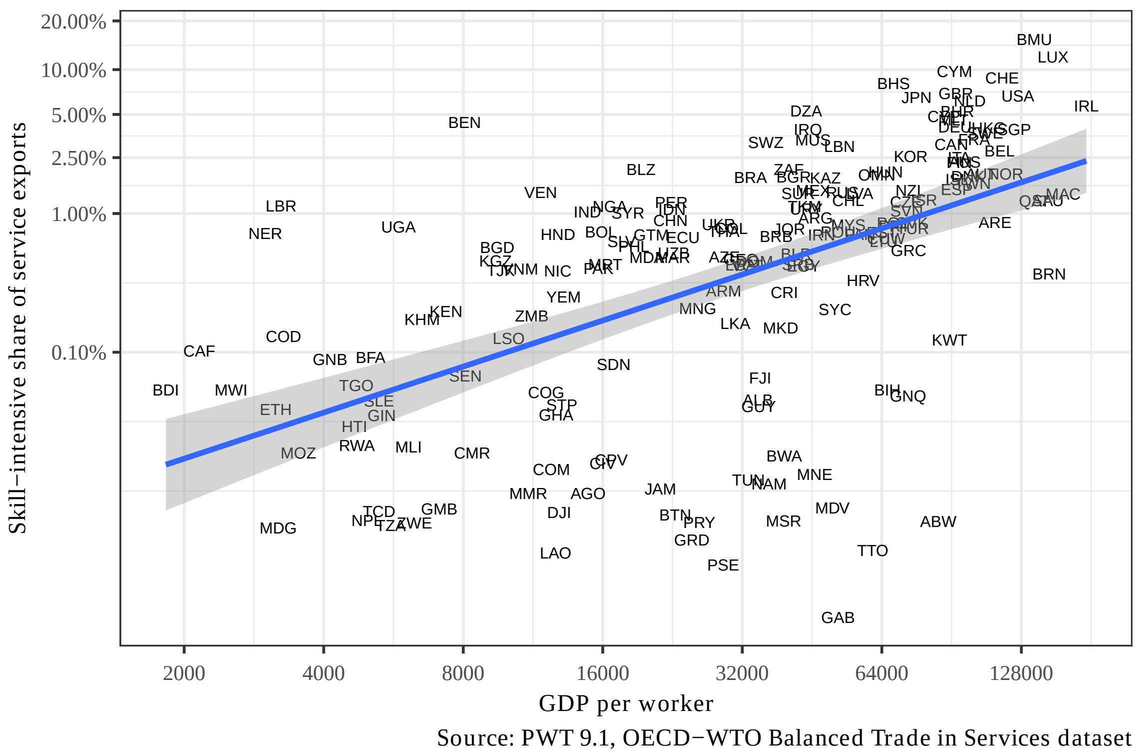


Figure 9: Share of service exports in skill-intensive categories

Given these assumptions, the gravity equation (6) implies that total exports and imports satisfy

$$Exp^i = \sum_{j\neq i} w^j L^j a_T \frac{G_T(L^i)(w^i \tau^{ji}/A_T^i)^{-\theta}}{a_T G_T(L^i)(w^i \tau^{ji}/A_T^i)^{-\theta} + G_L(L^j) a_{NT}(w^j/A_{NT}^j)^{-\theta} + \sum_{j'\neq i} a_T G_T(L^{j'})(w^{j'}\tau^{j'i}/A_T^j)^{-\theta}}$$

$$Imp^i = w^i L^i \frac{\sum_{j\neq i} G_T(L^j)(w^j \tau^{ji})^{-\theta}}{a_T G_T(L^i)(w^i/A^T)^{-\theta} + G_L(L^i) a_{NT}(w^i/A_{NT}^i)^{-\theta} + \sum_{j'\neq i} a_T G_T(L^{j'})(w^{j'}\tau^{j'i}/A_T^j)^{-\theta}}$$

Using that $G_T(L^i)/L^j \approx 0$ for all $j \neq i$, we obtain that

$$\frac{Exp^i}{Y^i} = \frac{Exp^i}{w^i L^i} \approx \kappa \frac{1}{A_T^i} \left(\frac{w^i}{A_T^i} \right)^{-\theta-1}$$

where we use that aggregate income Y^i is total wage payments $w^i L^i$ when there is only one input, and where we note that $\kappa = \frac{G_T(L^i)}{L^i} \sum_{j\neq i} \frac{w^j (\tau^{ji})^{-\theta}}{G_L(L^j)/L^j a_{NT}(w^j/A_{NT})^{-\theta} + \sum_{j'\neq i} a_T G_T(L^{j'})(w^{j'}\tau^{j'i})^{-\theta}/L^j}$ does not depend on the wage or the technology parameters of i. Similarly, we obtain that

$$\frac{Imp^i}{Y^i} = \frac{Imp^i}{w^i L^i} = \frac{\lambda}{a_T G_T(L^i)(w^i \tau^{ji})^{-\theta} + G_L(L^i) a_{NT}(w^i/A_{NT})^{-\theta} + \lambda},$$

where $\lambda = \sum_{j' \neq i} a_T G_T(L^{j'})(w^{j'}\tau^{j'i}/A_T^j)^{-\theta} \sum_{j' \neq i} a_T G_T(L^{j'})(w^{j'}\tau^{j'i}/A_T^j)^{-\theta}$ again does not depend on the wage or technology parameters in i (here, the assumption of a relatively large set of trade costs τ^{ij} ensures that the import share is not 1, which would happen if λ were much larger than $a_T G_T(L^i)(w^i\tau^{ji})^{-\theta} + G_L(L^i)a_{NT}(w^i/A_{NT})^{-\theta}$).

Since κ and λ do not depend on A_T^i, A_{NT}^i, and w^i, the equations can be used to solve for w^i and for conducting comparative statics with respect to A_T^i and A_{NT}^i.

D.2 Proof of proposition 2

The gravity specification 6 implies

$$\begin{aligned}
\log\frac{X_k^{ij}}{X_k^{ii}} &= \log\left(\frac{a_{k,T}G_T(L^j)\left(c_{k,T}^j\tau_k^{ij}\right)^{-\theta}}{a_{k,NT}G_{NT}(L^i)\left(c_{k,NT}^i\right)^{-\theta} + a_{k,T}G_T(L^i)\left(c_{k,NT}^i\right)^{-\theta}}\right)\\
&= \log\left(\frac{G_T(L^j)\left(c_{k,T}^j\tau_k^{ij}\right)^{-\theta}}{G_T(L^i)\left(c_{k,T}^i\right)^{-\theta}}\right) - \log\left(1 + \frac{a_{k,NT}}{a_{k,T}}\frac{G_{NT}(L^i)}{G_T(L^i)}\left(\frac{c_{k,NT}^i}{c_{k,T}^i}\right)^{-\theta}\right)\\
&= \underbrace{-\theta\log\tau^k - \log\left(1 + \frac{a_{k,NT}}{a_{k,T}}\frac{G_{NT}(L^{US})}{G_T(L^{US})}\left(\frac{c_{k,NT}^{US}}{c_{k,T}^{US}}\right)^{-\theta}\right)}_{\alpha_k} +\\
&\quad \underbrace{\log\frac{G_T(L^j)\left(c_{k,T}^j\right)^{-\theta}}{G_T(L^{US})\left(c_{k,T}^{US}\right)^{-\theta}}}_{\gamma_k^j} - \underbrace{\log\frac{G_T(L^i)\left(c_{k,T}^i\right)^{-\theta}}{G_T(L^{US})\left(c_{k,T}^{US}\right)^{-\theta}}}_{\gamma_k^i}\\
&\quad (-1)\underbrace{\left[\log\left(1 + \frac{a_{k,NT}}{a_{k,T}}\frac{G_{NT}(L^i)}{G_T(L^i)}\left(\frac{c_{k,NT}^i}{c_{k,T}^i}\right)^{-\theta}\right) - \log\left(1 + \frac{a_{k,NT}}{a_{k,T}}\frac{G_{NT}(L^{US})}{G_T(L^{US})}\left(\frac{c_{k,NT}^{US}}{c_{k,T}^{US}}\right)^{-\theta}\right)\right]}_{\mu_k^i} +\\
&\quad \sum_m \beta_{m,k}d_m^{ij} - \theta\varepsilon_k^{ij}.
\end{aligned}$$

This shows that trade flows normalized by own trade can be written as

$$\log\frac{X_k^{ij}}{X_k^{ii}} = \alpha_k + \gamma_k^j - \gamma_k^i - \mu_k^i + \sum_m \beta_{m,k}d_m^{ij} - \theta\varepsilon_k^{ij},$$

where $\gamma_k^{US} = \mu_k^{US} = 0$. If we For identification, we note that ε_k^{ij} is uncorrelated to the explanatory variables, since it is uncorrelated to d_m^{ij} and $\mathbb{E}\varepsilon_k^{ij} = 0$ for every i, j, k implies that it is uncorrelated to the variables associated with the fixed effects. For identification, it remains to show that the null space does not contain variations in γ_k^i and μ_k^i conditional on being in the subspace with $\gamma_k^{US} = \mu_k^{US} = 0$.

We consider a candidate element in the null-space, not considering the parameters $\beta_{m,k}$ given the assumptions that the distance variables are linearly independent from the fixed effects.

$$0 = \tilde{\alpha}_k + \tilde{\gamma}_k^j - \tilde{\gamma}_k^i - \tilde{\mu}_k^i \quad \forall k, i \neq j$$

Considering for $i = US$ and $j = US$, we obtain

$$\tilde{\alpha}_k = -\tilde{\gamma}_k^j \quad \forall k, j \neq US$$
$$\tilde{\alpha}_k = \tilde{\gamma}_k^i + \tilde{\mu}_k^i \quad \forall k, i \neq US.$$

This implies that for $i \neq US, j \neq US$, we have $0 = \tilde{\alpha}_k - \tilde{\alpha}_k - \tilde{\alpha}_k$, so $\tilde{\alpha}_k = \tilde{\gamma}_k^j = \tilde{\mu}_k^i = 0$ as desired.

D.3 Equivalence of non-tradable technology and export-specific trade costs

Suppose that bilateral trade costs satisfy

$$\tau_k^{ij} = f_k(\{d^{ij}\})\tau_{k,exp}^j,$$

where $\{d_k^{ij}\}$ is a measurable set of trading costs, and $\tau_{k,exp}^j$ is a cost associated with shipping goods from j to any other country. In that case, (6) implies that bilateral trade flows satisfy

$$X_k^{ij} = \begin{cases} X_k^i \frac{a_{k,T} G_T(L^j)\left(c_{k,T}^j f_k(\{d^{ij}\})\tau_{k,exp}^j\right)^{-\theta}}{\Phi_k^i} & \text{if } i \neq j \\ X_k^i \frac{a_{k,NT} G_{NT}(L^i)\left(c_{k,NT}^i\right)^{-\theta} + a_{k,T} G_T(L^i)\left(c_{k,NT}^i\right)^{-\theta}}{\Phi_k^i} & \text{if } i = j, \end{cases}$$

From this expression, we get that trade flows from two separate origins j and l to the same destination satisfies

$$\frac{X_k^{ij}}{X_k^{il}} = \left(\frac{f_k(\{d^{ij}\})}{f_k(\{d^{il}\})}\right)^{-\theta} \frac{G_T(L^j)}{G_T(L^k)} \left(\frac{c_{k,T}^j \tau_{k,exp}^j}{c_{k,T}^i \tau_{k,exp}^i}\right)^{-\theta}.$$

Thus, we note that export success conditional on observable trade costs satisfies

$$\text{Export success}_k^i = G_T(L^i)[c_{k,T}^i \tau_{k,exp}^i]^{-\theta}.$$

Furthermore, we note that trade flows relative to own-trade satisfies

$$\frac{X_k^{ij}}{X_k^{ii}} = f_k(\{d^{ij}\})^{-\theta} \frac{G_T(L^j)}{G_T(L^i)} \left(\frac{c_{k,T}^j \tau_{k,exp}^j}{c_{k,T}^i \tau_{k,exp}^i} \right)^{-\theta} \left(\tau_{k,exp}^i\right)^{-\theta} \left(1 + \frac{a_{k,NT} G_{MT}(L^i)}{a_{k,T} G_T(L^j)} \left(c_{k,T}^i / c_{k,NT}^i\right)^{\theta}\right)^{-1}$$

$$= f_k(\{d^{ij}\})^{-\theta} \frac{\text{Export success}_k^j}{\text{Export success}_k^i} \left(\tau_{k,exp}^i\right)^{-\theta} \left(1 + \frac{a_{k,NT} G_{MT}(L^i)}{a_{k,T} G_T(L^j)} \left(c_{k,T}^i / c_{k,NT}^i\right)^{\theta}\right)^{-1}.$$

This expression shows that conditional on observable trade costs and relative export success, export-specific trade costs are isomorphic to the comparative advantage term $1 + \frac{a_{k,NT} G_{MT}(L^i)}{a_{k,T} G_T(L^j)} \left(c_{k,T}^i / c_{k,NT}^i\right)^{\theta}$ in terms of their effect on trade flows.

D.4 Adjusting for differences in trade costs

Different export costs. One way to rationalize the data is that poor countries have systematically higher export costs as in Waugh (2010). To assess this hypothesis, I compare trade-implied unit costs to observed unit costs when both are available. If poor countries have systematically higher export costs and we back out unit costs assuming that they do not, their trade-implied unit costs should be high relative to their actual unit costs.

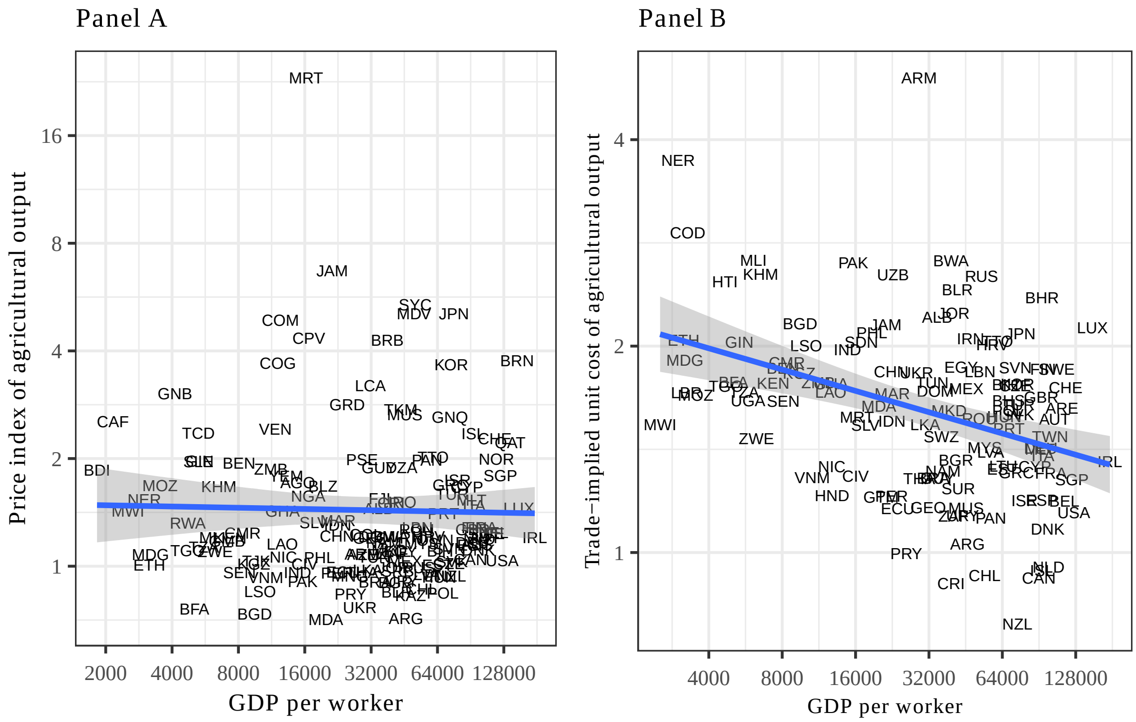
To test this, I consider export data from the agricultural sector. First, I calculate implied unit costs by using the estimates of $\widehat{\text{Export success}}_k^i$ from regression (17) and adjust for country size. I then compare how much these measures vary compared to a producer price index constructed by Kiernan (2022), who uses FAO-provided farm-gate prices across many crops in both poor and rich countries. The results are displayed in figure 10. We see that there is some evidence that export costs are higher for poor countries than for rich countries: observed unit costs are similar across countries (panel A), but the estimated unit costs from trade data are approximately 0.5 log points higher in poor countries than in rich countries.

However, quantitatively, these estimated trade costs differences are much smaller than those needed to rationalize my findings. In particular, we know that if the home bias term should not vary across countries, we need that

$$\Delta\text{Destination fixed effect} = \theta\Delta\text{Unit costs}.$$

In 4, I found that this was violated because unit costs were much higher in poor countries, while estimated destination fixed effects were not. To see how adjusting for trade costs changes this picture, we note that

$$\theta(\Delta\text{Unit cost} + \Delta\log\text{export costs}) = -\Delta\text{Origin fixed effects}.$$



Source: Penn World Table 9.1, FAOSTAT

Source: Penn World Table 9.1, Authors' estimate

Figure 10: Trade-implied unit costs and observed unit costs in agriculture

The main estimation set $\theta\Delta \log \text{export costs} = 0$ (in the sense of only adjusting for observable trade barriers and not allowing for a separate effect of income on trade costs). From the agricultural data, I can derive that $\theta \times$ (log export costs) have a semi-elasticity of approximately 0.6 with GDP per worker. Using this estimate, we can adjusts unit cost estimates for trade costs, and compare them to estimated destination fixed effect. This is done in figure 11, where I display destination fixed effects, unadjusted differences in unit costs (that is, simply the negative origin fixed effects), and adjusted differences in unit costs (where I add in the income-dependent export costs from the agricultural data). For each country, the fixed effect is averaged across all industries. We see the familiar patterns that estimated unit costs are higher in poor countries, while destination fixed effects are similar. Adjusting for different export costs makes the picture somewhat weaker, but can only explain a minority of the gap.

Import costs. A second hypothesis is that import costs are higher in poor countries. This would help explain why their imports are "too low" relative to their limited import success. To explore the plausibility of this hypothesis, I do not have as direct a test as using implied unit costs from the agricultural data. However, it is possible to gauge how the price index of imports from the Penn World Table varies as a

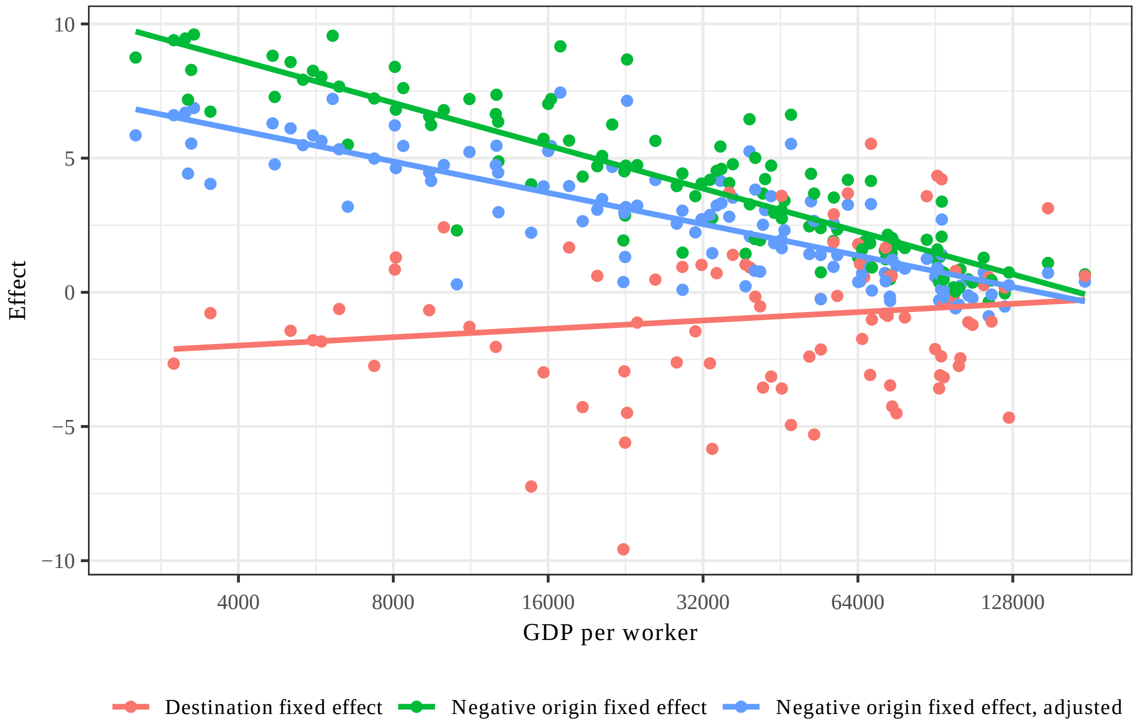


Figure 11: Destination and origin fixed effects, adjusted for trade costs

function of country income. These results are presented in panel A of figure 12. A more narrow but more interpretable, data pattern is to look at the price index of equipment – which are intensely traded – in the International Comparison Project. This is displayed in panel B. In both cases, there is no large variation in price indices across countries: poor countries have approximately a 10% lower price index of exports, and 10% higher price index in equipment. Neither of these facts suggest that dramatically higher import costs in poor countries can explain my findings.

E Appendix to section 5

Proof of proposition 4. Suppose that the set of unit costs c exhibits skill bias. Then, we have that

$$\log \frac{c[1,\omega;k',y']}{c[1,\omega;k,y']} \leq \log \frac{c[1,\omega;k',y]}{c[1,\omega;k,y]} \quad \forall \omega, k' > k, y' > y \tag{46}$$

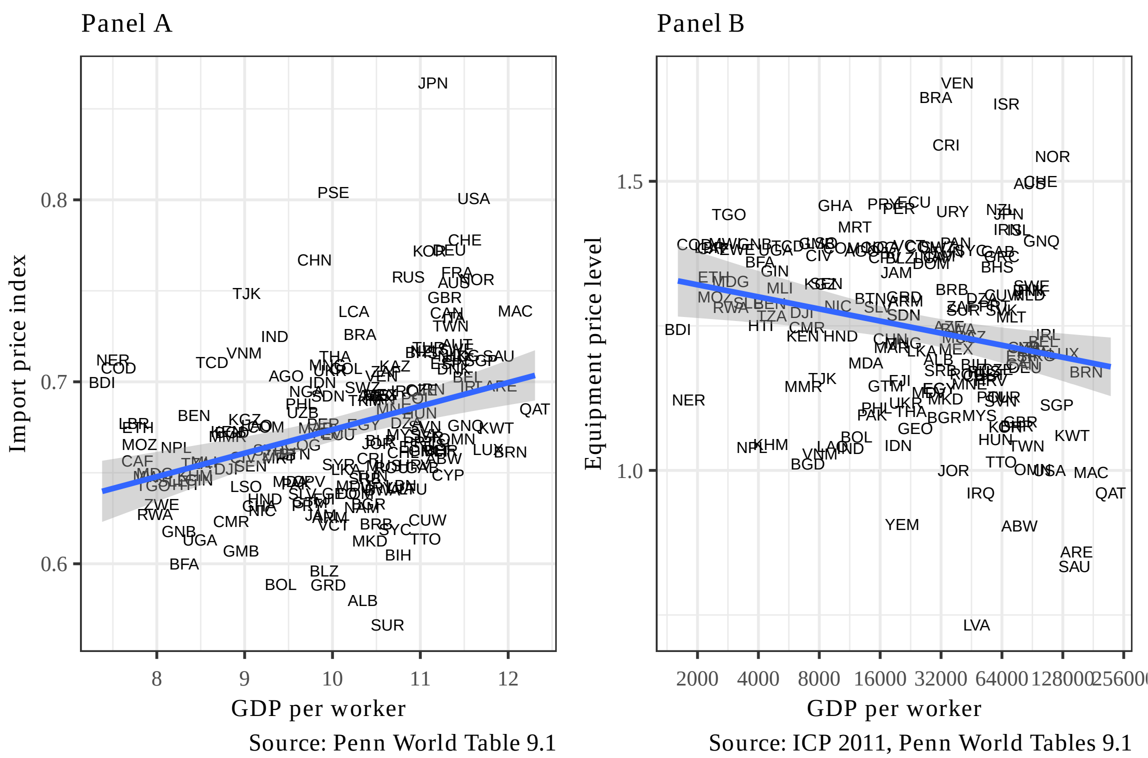


Figure 12: Price indices for imports and equipment

where ω is some arbitrary skilled wage premium. Furthermore, we note that this expression implies

$$\log \frac{c[1,\omega;k',y']/\bar{c}[1,\omega;\bullet,y']}{c[1,\overline{\omega(y)};k',\bullet]/\bar{c}[1,\overline{\omega(y)};\bullet,\bullet']} - \log \frac{c[1,\omega;k,y']/\bar{c}[1,\omega;\bullet,y']}{\bar{c}[1,\overline{\omega};k,\bullet]/\bar{c}[1,\overline{\omega};\bullet,\bullet]} \leq$$
$$\log \frac{c[1,\omega;k',y]/\bar{c}[1,\omega;\bullet,y]}{c[1,\overline{\omega};k',\bullet]/\bar{c}[1,\overline{\omega};\bullet,\bullet]} - \log \frac{c[1,\omega;k,y]/\bar{c}[1,\omega;\bullet,y]}{c[1,\omega;k,y]/\bar{c}[1,\omega;\bullet,\bullet]}$$

which is equivalent to (46), since $\bar{c}[1,\omega;\bullet,y']$ cancel out on the left-hand side, $\bar{c}[1,\bar{\omega};\bullet,y]$ cancel on the right-hand side, $\bar{c}[1,\bar{\omega};\bullet,\bullet]$ cancel on both sides, and $c[1,\bar{\omega};k',\bullet]$ and $c[1,\bar{\omega};k,\bullet]$ occur symmetrically on both sides of the expression. I write this as

$$\log \frac{\hat{c}(\omega;k',y')}{\hat{c}(\omega;k,y')} \leq \log \frac{\hat{c}(\omega;k',y)}{\hat{c}(w;k,y)}, \tag{47}$$

where $\log \hat{c}$ are doubly normalized unit costs. In this notation, the proposition makes a claims about the derivative of $\log \tilde{c}(y,k) \equiv \log \hat{c}[\omega(y);k,y]$ given that (47) holds. Now, (47) implies that $\frac{\partial}{\partial y} \log \frac{\hat{c}[\omega;k',y]}{\hat{c}[\omega;k,y]} \leq 0$. This means that

$$
\begin{aligned}
\frac{\partial \log \tilde{c}(y,k')/\tilde{c}(y,k)}{\partial y} &= \frac{d}{dy}\log\frac{\hat{c}(\omega(y);y,k')}{\hat{c}(\omega(y);y,k)} \\
&= \left(\frac{\partial \log \hat{c}[\omega(y);y,k']}{\partial \log \omega} - \frac{\partial \log \hat{c}[\omega(y);y,k]}{\partial \log \omega}\right)\frac{d\log\omega}{dy} + \frac{\partial}{\partial y}\log\frac{\hat{c}(\omega(y);y,k')}{\hat{c}(\omega(y);y,k)} \\
&\leq \left(s_{k'}(y) - s_k(y)\right)\frac{d\log\omega}{dy}
\end{aligned}
$$

where the last step uses that $\frac{\partial}{\partial y}\log\frac{\hat{c}[\omega;k',y]}{\hat{c}[\omega;k,y]} \leq 0$ and that the skill share is the elasticity of the unit cost function with respect to the skilled wage premium.Assuming that k is a continuum and that $\tilde{c}$ is twice differentiable and $s_k(y)$ is differentiable in k for every y, we can divide both sides by $k' - k$ and let $k' \to k$ to obtain

$$
\frac{\partial \log \tilde{c}(y,k)}{\partial y \partial k} \leq \frac{\partial s}{\partial k}\frac{d\log\omega}{dy}.
$$

Last, note that if we have any parameter change $k = g(\tilde{k})$ with g differentiable and strictly monotonic, writing $\tilde{c}^*$ and $s^*_{\tilde{k}}$ for the functions in the new coordinates, we have:

$$
\begin{aligned}
\frac{\partial \log \tilde{c}^*(y,\tilde{k})}{\partial y \partial \tilde{k}} &= \left.\frac{\partial s_k(y)}{\partial k}\right|_{k=g(\tilde{k})} g'(\tilde{k}) \\
\frac{\partial s^*_{\tilde{k}}(y)}{\partial \tilde{k}} &= \left.\frac{\partial s_k(y)}{\partial k}\right|_{k=g(\tilde{k})} g'(\tilde{k}).
\end{aligned}
$$

Since $g'(\tilde{k})$ cancel on both sides, we obtain the same inequality, which concludes the proof.

F Appendix to section 6

F.1 Formal industry model setup and derivation

To define an industry equilibrium, we assume that in every period t, there is a measure of $\mu_t \geq 0$ firms being founded (for readability, we drop the industry index k in the derivation). At the date of founding, the value of the firm V_t is defined by the discounted value of future profits, and satisfies $V_t \leq 0$ with equality if $\mu_t > 0$. The measure of firms by age a at time t is $\mu_{t-a}(1-\tau)^a$. A steady state equilibrium is defined by a price P, a birth rate of firms μ, age-specific decision functions of firms $\{\ell_a, s^I_a, s^M_a, k_a\}$ and aggregate quantities L,

S^I, S^M, S, K such that $\{\ell_a, s_a^I, s_a^M, x_a\}$ solves the firm's problem, firm profits are zero

$$\sum_{a\geq 0} \frac{(1-\tau)^a}{(1+r)^a} \left(P\ell_a^{1-\alpha}k_a^\alpha - w_u\ell_a - w_s[s_a^I + s_a^M]\right) = 0,$$

and aggregate quantities satisfy:

$$\begin{aligned} L &= \mu \sum_{a\geq 0} (1-\tau)^a \ell_a \\ S^I &= \mu \sum_{a\geq 0} (1-\tau)^a s_a^I \\ S^M &= \mu \sum_{a\geq 0} (1-\tau)^a s_a^M \\ K &= \mu \sum_{a\geq 0} (1-\tau)^a k_a \\ S &= S^I + S^M. \end{aligned}$$

To solve for the steady-state, we note that the first order condition of the firm's problem implies

$$(1-\alpha_a)P\left(\frac{k_a}{\ell_a}\right)^{\alpha_k} \leq w_u \qquad \forall a \geq 0$$

with equality if $\ell_a > 0$. Hence, the amount of production workers satisfy

$$\ell_a = \begin{cases} 0 & \text{if } k_a = 0 \\ k_a \left[\frac{(1-\alpha)P}{w_u}\right]^{1/\alpha} & \text{if } k_a > 0. \end{cases}$$

If we define the constant $\chi \equiv \left[\frac{(1-\alpha)P}{w_u}\right]^{-1/\alpha_k}$, we have that $k = \ell_a\chi$ for both $x_a = 0$ and $x_a > 0$, since $\ell_a = 0$ when $k_a = 0$. Combining $k = \ell_a\chi$ with $s_a^m = \pi k_a/A_K$ and $s_a^I = \frac{ka+1-(1-\delta)k_a}{A_K}$, we can rewrite the firm's objective as

$$\max_{\{\ell_a\}} \sum_{a\geq 0} \left(\frac{1-\tau}{1+r}\right)^a \left[P\ell_a\chi^\alpha - w_u\ell_a - \frac{w_s\chi}{A_K}\left(\ell_a(\pi+\delta) + \ell_{a+1} - \ell_a\right)\right].$$

Furthermore, we note that

$$\sum_{a\geq 0}\left(\frac{1-\tau}{1+r}\right)^a(\ell_{a+1}-\ell_a) = \sum_{a\geq 1}\left(\frac{1-\tau}{1+r}\right)^a \ell_a\left(1-\frac{1-\tau}{1+r}\right)-\ell_0$$
$$= \frac{r+\tau}{1+r}\sum_{t\geq 1}\left(\frac{1-\tau}{1+r}\right)^a \ell_a$$

where the last line uses $\ell_0 = 0$. Putting everything together, we obtain

$$\max_{\ell_a}\sum_{a\geq 1}\left(\frac{1-\tau}{1+r}\right)^a \ell_a\left[P\chi^\alpha - w_u - \frac{w_s\chi}{A_K}\left(\pi+\delta+\frac{r+\tau}{1+r}\right)\right].$$

For an equilibrium to exist, the expression inside the square brackets need to be zero. Writing $w_u = \chi^\alpha(1-\alpha)P$, we obtain

$$\begin{aligned}
P\chi^\alpha\alpha = \frac{w_s x}{A_X} &\implies P = \frac{\left(\pi+\delta+\frac{r+\tau}{1+r}\right)}{A_K}\frac{w_s}{\alpha}\chi^{1-\alpha}\\
&\implies P^{\frac{1}{\alpha}} = \frac{\left(\pi+\delta+\frac{r+\tau}{1+r}\right)}{A_K}\frac{w_s}{\alpha}\left[\frac{(1-\alpha_k)}{w_u}\right]^{\frac{-(1-\alpha)}{\alpha}}\\
&\implies P = \left[\frac{\left(\pi+\delta+\frac{r+\tau}{1+r}\right)}{A_K}\frac{w_s}{\alpha}\right]^\alpha\left[\frac{w_u}{1-\alpha}\right]^{1-\alpha},
\end{aligned} \tag{48}$$

as required. This also implies that

$$\chi = \left[\frac{(1-\alpha)P}{w_u}\right]^{-1/\alpha_k} = \frac{\alpha}{1-\alpha}\left(\frac{w_s}{w_u}\right)^{-1}\frac{A_X}{\pi+\delta+\frac{r+\tau}{1+r}}$$

To derive skill intensity, we can substitute $k_a = \ell_a\chi$ to obtain

$$\begin{aligned}
S^I &= \chi\mu\sum_{a\geq 0}(1-\tau)^a\frac{\ell_{a+1}-(1-\delta)\ell_a}{A_X} &&= \frac{\chi(\tau+\delta)}{A_X}\times L\\
S^M &= \sum_{a\geq 0}(1-\tau)^a\frac{\pi\ell_a}{A_X} &&= \frac{\pi\chi}{A_X}\times L
\end{aligned}$$

where the first equation uses that $\sum_{a\geq 0}(1-\tau)^a[\ell_{a+1}-\ell_a] = \sum_{a\geq 0}\tau\ell_a - \ell_0$ together with $\ell_0 = 0$. Hence, the skill intensity satisfies

$$
\begin{aligned}
\frac{w_s}{w_u}\frac{S^I+S^M}{L} &= \frac{w_s}{w_u}\frac{\chi}{A_X}(\tau+\delta+\pi) \\
&= \frac{\alpha}{1-\alpha}\left[1 - \frac{r\times(1-\tau)/(1+r)}{\pi+\delta+\tau+r\times(1-\tau)/(1+r)}\right].
\end{aligned}
$$

To incorporate this model into a static trade model used in section 2, we assume that each industry k in every country has a continuum of varieties $\omega_k \in [0,1]$, where each variety has two potential technologies with different shares of organizational capital and productivity shocks $z^i_{k,T}(\omega_k), z^i_{k,NT}(\omega_k)$ for the traded and non-traded technology respectively. Furthermore, we assume that there is no growth and that each country has a representative household with preferences $\sum_t \beta^t u(c_t)$. There is no trade in capital. In steady-state, each country will then have a common interest rate r. In steady-state, the minimum price at which a variety will be active for production in i is if it can fetch a price $\tilde{p}^i_{k,t}(\omega_k) = \frac{1}{z_{k,t}(\omega_k)}\left[\frac{\left(\pi+\delta+\frac{r+\tau^i}{1+r}\right)}{A_K}\frac{w^i_s}{\alpha}\right]^{\alpha_{k,t}}\left[\frac{w^i_u}{1-\alpha}\right]^{1-\alpha_{k,t}}$. Hence, the model will behave like an Eaton and Kortum style model where deterministic unit costs are given by (48).

F.2 Robustness to auxiliary assumptions

To show robustness of my conclusions, suppose an alternative setup where firms have a production technology

$$
y_a = [\ell_a^\xi s_a^{1-\xi}]^{1-\alpha} k_a^\alpha,
$$

where ℓ is the number of unskilled production workers, s is the number of skilled production workers, k is the amount of organizational capital, and α is the output share of organizational capital. As before, I assume that organizational capital is accumulated and maintained by skilled workers:

$$
k_{a+1} \leq (1-\delta)k_a + s^k_I A_K \tag{49}
$$

$$
\pi k_a \leq s^M_a A_K, \tag{50}
$$

For any k_a, the firm statically solves

$$
\max_{\ell_a, s_a} P[\ell_a^\xi s_a^{1-\xi}]^{1-\alpha} k_a^\alpha - w_u\ell_a - w_s\ell_s,
$$

which implies

$$\frac{P\xi(1-\alpha)\left[l_a^{\xi}s_a^{1-\xi}\right]^{1-\alpha}k_a^{\alpha}}{l_a} = w_u$$
$$\frac{P(1-\xi)(1-\alpha)\left[l_a^{\xi}s_a^{1-\xi}\right]^{1-\alpha}k_a^{\alpha}}{l_s} = w_s$$

which, if $l_a > 0$, implies

$$s_a = \frac{1-\xi}{\xi}\frac{w_u}{w_s}l_a$$

$$P\xi(1-\alpha)\left(\frac{1-\xi}{\xi}\frac{w_u}{w_s}\right)^{(1-\xi)(1-\alpha)}(k_a/l_a)^{\alpha} = w_u \Longrightarrow k_a^{\alpha} = \ell_a^{\alpha}\psi^{\alpha} \quad \psi = \left[\frac{w_u}{P\xi(1-\alpha)}\left(\frac{1-\xi}{\xi}\frac{w_u}{w_s}\right)^{-(1-\xi)(1-\alpha)}\right]^{1/\alpha}$$

Hence, the solution is given by

$$\max_{\{l_a\}_{a\geq 0}}\sum_{a\geq 0}\left(\frac{1-\tau}{1+r}\right)^{a}\left\{\frac{\ell_a w_u}{\xi(1-\alpha)} - \frac{l_a}{\xi}w_u - \frac{\psi w_s}{A_K}\left[\pi l_a + (\ell_{a+1} - \ell_a + \delta\ell_a)\right]\right\}$$

subject to the constraint $l_0 = 0$. Noting that

$$\sum_{a\geq 0}\left(\frac{1-\tau}{1+r}\right)^{a}\frac{\psi w_s}{A_K}(\ell_{a+1} - \ell_a) = \sum_{a\geq 1}\left(\frac{1-\tau}{1+r}\right)^{a}\ell_a\left(1 - \frac{1-\tau}{1+r}\right) - \ell_0 = \frac{r+\tau}{1+r}\sum_{t\geq 1}\left(\frac{1-\tau}{1+r}\right)^{a}\ell_a\frac{\psi w_s}{A_K},$$

we obtain

$$\max_{\{l_a\}_{a\geq 0}}\sum_{a\geq 0}\left(\frac{1-\tau}{1+r}\right)^{a}\left\{\frac{\ell_a w_u}{\xi(1-\alpha)} - \frac{l_a w_u}{\xi} - \frac{\psi w_s}{A_K}\left[\pi + \frac{r+\tau}{1+r} + \delta\right]l_a\right\}$$

Hence, the zero profit condition is

$$\frac{\alpha}{\xi(1-\alpha)}w_u = \psi\frac{w_s}{A_K}\left[\pi + \frac{r+\tau}{1+r} + \delta\right] \Longleftrightarrow \psi = \left(\frac{w_s}{w_u}\right)^{-1}\frac{\alpha}{\xi(1-\alpha)}\frac{A_K}{\left[\pi + \frac{r+\tau}{1+r} + \delta\right]}$$

which implies

$$P = \left(\frac{\alpha}{\xi}\right)^{-\alpha}\frac{1}{\xi(1-\alpha)}\left(\frac{1-\xi}{\xi}\frac{}{w_s}\right)^{-(1-\xi)(1-\alpha)}w_u^{\xi(1-\alpha)}w_s^{\alpha+(1-\xi)(1-\alpha)}\left(\frac{1}{A_K}\left[\pi + \frac{r+\tau}{1+r} + \delta\right]\right)^{\alpha}. \qquad (51)$$

The relative payroll share of skilled and unskilled labor is

$$\frac{w_s}{w_u}\frac{(\tau+\delta+\pi)k_a/A_K+s_a}{l_a} = \frac{w_s}{w_u}\left[(\tau+\delta+\pi)\psi/A_K+\frac{w_u}{w_s}\frac{1-\xi}{\xi}\right]$$

$$= \frac{1}{(1-\alpha)\xi}\left[\alpha\frac{\tau+\delta+\pi}{\pi+\frac{r+\tau}{1+r}+\delta}+(1-\xi)(1-\alpha)\right]. \quad (52)$$

$$= \frac{\alpha}{(1-\alpha)\xi}\frac{\tau+\delta+\pi}{\pi+\frac{r+\tau}{1+r}+\delta}+\frac{(1-\xi)}{\xi} \quad (53)$$

The expression (51) shows that the price index increases with α, and the expression shows that (52) shows that so does the ratio of skilled and unskilled labor compensation. Thus, the prediction is robust even if skilled labor is used in final production.