Estimating Skilled Labor Efficiency Using Trade and Industry Data∗

Hannes Malmberg†

Click Here for Latest Version

Februari 12, 2018

Abstract

An important task in the growth and development literature has been to characterize the nature of cross-country labor efficiency differences. In this paper, I develop a new method to estimate the relative efficiency of skilled and unskilled labor across countries using disaggregated industry and trade data. The method assigns a country a high efficiency of skilled labor if it has low relative unit production costs in skill-intensive industries, conditional on factor prices. For estimation, I use international trade data. I first document that the share of exports in skill-intensive industries rises sharply with income levels. Interpreted through the lens of a gravity model, this pattern suggests that rich countries have low relative unit production costs in skill-intensive industries, and low effective costs of skilled labor services. I show that for standard trade elasticities, these low effective costs of skilled labor services cannot be rationalized by differences in skilled wage premia, which leads me to infer that skilled labor is relatively efficient in rich countries. Integrating these findings into a development accounting exercise in manufacturing, I find that accounting for skill-specific efficiency differences reduces the size of uniform labor efficiency differences: the difference in skill-neutral TFP differences between rich and poor countries in manufacturing falls from a factor of 4.3 to a factor of 2.6.

∗This paper was previously circulated under the name ”Human Capital and Development Accounting Revisited”. I am grateful to Timo Boppart, Axel Gottfries, John Hassler, Karl Harmenberg, Per Krusell, Erik Öberg, and Torsten Persson for extensive comments on the project. Moreover, I am grateful for helpful discussions with Ingvild Almás, Konrad Burchardi, Francesco Caselli, Saman Darougheh, Niels-Jakob Harbo Hansen, Chad Jones, Loukas Karabarbounis, David Lagakos, Bo Malmberg, Kurt Mitman, Jon de Quïdt, Jósef Sigurdsson, Lena Sommestad, David Strömberg, and David Weil.

†Postdoctoral Fellow. SIEPR, Stanford University. Email: malmberg@stanford.edu.
1 Introduction

Since the mid-1990s, an important line of research in the growth and development literature has been to analyze the variation in the efficiency of factor inputs across countries (Klenow and Rodriguez-Clare, 1997; Hall and Jones, 1999; Caselli, 2005; Jones, 2014a; Caselli, 2016; Rossi, 2017). Such efficiency differences must be present, since variations in the quantity of factor inputs are too small to explain the size of the output differences between rich and poor countries.

One important hypothesis has been that differences in the efficiency of skilled labor are particularly important to understand differences in income levels (Caselli and Coleman, 2006; Jones, 2014a; Caselli, 2016). This focus on skilled labor reflects a long tradition in development economics analyzing the various ways in which skill intensification and economic development interact, either through human capital increasing worker productivity and facilitating technology adoption (Nelson and Phelps, 1966; Lucas, 1988; Mankiw et al., 1990), or through modern economic growth being characterized by skill-complementary technologies (Galor and Weil, 1999; Goldin and Katz, 2008), potentially due to capital-skill complementarities (Fernandez-Villaverde, 2001).

To estimate efficiency differences by skill level, the standard approach in the literature has been to use aggregate data on the wages and number of workers in different skill categories.\(^1\) Such an analysis requires assumptions on the aggregate elasticity of substitution between skilled and unskilled workers, reflecting the impossibility result of Diamond et al. (1978), which states that factor efficiency differences and the elasticity of substitution cannot be separately identified using only aggregate data on prices and quantities.

Different papers have used different assumptions regarding the relevant elasticity. Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), and Caselli (2005) assume that there is perfect substitutibility between skilled and unskilled workers, while Caselli and Coleman (2006), Jones (2014a), and Caselli (2016) use estimates of the elasticity of substitution taken from US time series and panel data, e.g., from Katz and Murphy (1992) and Ciccone and Peri (2005). The lower the elasticity of substitution is, the higher is the estimated relative efficiency of skilled labor in rich countries. The reason is that a low elasticity of substitution decreases the relative price per efficiency unit of skilled labor where skilled labor is abundant. Thus, to match observed skilled wage premia, rich countries, being abundant in skilled labor, need to have a high efficiency of skilled labor.

In this paper, I propose a new method to obtain evidence on the degree of skill bias in labor efficiency differences. I use information from disaggregated trade and industry data, and the idea is that if skilled labor is efficient in rich countries, this will reduce the effective cost of skilled labor services, and thereby also the relative unit production cost in skill-intensive industries. Thus, skilled labor efficiency differences affect comparative advantage, and can be detected and quantified using data on industry level trade flows, without having to make an assumption on the aggregate

\(^1\)An exception is a recent paper by Morrow and Trefler (2017), discussed below.
elasticity of substitution.

Building on standard methods in production theory, the first step of the analysis constructs a method for measuring the factor bias of efficiency differences using data on industry productivities. Using a dual formulation, I show that, when industry productivity differences are a function of factor efficiency differences, unit production costs are functions of factor prices adjusted for factor efficiency levels. Furthermore, assuming a competitive market, factor compensation shares equal the factor price elasticities of the unit cost functions. This implies that the differences across countries in relative industry unit production costs can be approximated in terms of industry factor shares, differences in factor prices, and differences in factor efficiency levels. Given this approximation, a projection of industry unit production costs on industry factor shares identify the efficiency-adjusted factor service prices. The ratio of the observed factors service price and the efficiency-adjusted factor service price reflects the relative efficiency of a production factor.\(^2\)

The measure requires information on industry unit production costs, and, to estimate these, I use information from international trade data. A salient feature of the data is that rich countries have, on average, a much higher share of skill-intensive exports than poor countries. If we interpret this through the lens of a gravity model, this suggests that rich countries have relative low unit production costs in skill-intensive industries. For quantification, I derive a gravity regression, where country-industry fixed effects reflect relative unit production costs, which can then be used to measure skill-biased efficiency differences. This use of trade data circumvents the lack of quality-adjusted industry producer price indices that cover a large number of rich and poor countries.\(^3\)

The estimated country-industry fixed effects confirm the intuition that rich countries tend to have relatively low unit production costs in skill-intensive industries. When combined with factor share data, the unit cost estimates allow us to estimate the efficiency-adjusted skilled labor prices that are needed to rationalize these trade patterns. For the skilled labor share, I use the share of total costs paid to workers in skilled occupations: managers, professionals, and technicians. I find that, for normal ranges of trade elasticities, the efficiency-adjusted prices of skilled labor services differ more across rich and poor countries than the observed relative wages of skilled workers. Thus, I infer that labor efficiency differences across rich and poor countries are skill-biased. For the middle range of the assumed trade elasticities (\(\sigma = 7.5\)), rich countries have an estimated 7.1 times higher relative efficiency of skilled labor than poor countries. For lower assumed trade elasticities, efficiency differences become more skill biased, since larger unit cost differences are needed to rationalize the observed trade patterns.

The measure is derived assuming that relative unit production costs are only a function of factor efficiency differences. In Section ??, I show that the measure extends naturally to the case when there are also industry-country specific productivity differences. In this case, the estimated efficiency differences also reflects the component of industry-country productivity terms that is correlated to skill intensity.

The most similar approach in the literature is the one used in Levchenko and Zhang (2014) to estimate the gains of trade from comparative advantage. In Section 3.4, I redo my exercise for the subsets of countries for which direct measures of industry unit costs are available.
If we make additional assumptions, among others that aggregate production functions differ across countries only due to factor efficiency differences, the trade-based estimates can be used to perform a productivity decomposition exercise. Focusing on the manufacturing sector, I find that taking skill bias into account reduces the size of skill-neutral TFP differences between rich and poor countries from a factor of 4.3 to a factor of 2.6, with skill-specific efficiency differences instead becoming more important. The labor aggregator does not feature a constant elasticity of substitution between skilled and unskilled workers. However, having an independent measure of the skill bias, it is possible to regress relative factor shares on relative efficient supplies. If the resulting estimate is interpreted using a CES-model, this suggests an elasticity of substitution of 1.23.4

My findings of skill-biased efficiency differences are, qualitatively, similar to the results found in papers using aggregate data and an assumption of imperfect substitutability between skilled and unskilled workers (Caselli and Coleman, 2006; Jones, 2014a; Caselli, 2016; Rossi, 2017), even though the results are not fully comparable quantitatively due to my focus on manufacturing, and different skill definitions and cut-offs.5 However, when I replicate the analysis of Caselli (2016), using his published data, for the same set of countries as the ones used in my analysis, I find that the relative skilled labor efficiency is approximately 4 times higher in rich countries than in poor countries, i.e., of the same order of magnitude as in my baseline results.

The aggregate evidence, from, e.g., Caselli (2016), and the industry-based evidence in this paper, provide independent support for the existence of skill-biased efficiency differences, since they are arrived at using very different methodologies. My method rejects skill-neutral efficiency differences, since observed small variations in the skilled wage premium are inconsistent with large observed differences in the skill-intensity of trade flows, given conventional trade elasticity estimates. Caselli and Coleman (2006), Jones (2014a), Caselli (2016), and Rossi (2017), in contrast, reject skill-neutral efficiency differences, since small variations in the skilled wage premium are inconsistent with large differences in the relative number of skilled workers, for conventional estimates of the elasticity of substitution between skilled and unskilled workers.

The existence of skill bias has implications for how we interpret income differences across countries. Without skill bias, there is, in an accounting sense, no special role for skilled labor in the process of economic development, which is primarily about skill-neutral increases in productivity. With skill-biased efficiency differences, we need to understand why skilled labor efficiency in particular varies across countries. This could reflect differential barriers to the adoption of skill-

---

4This estimate is somewhat smaller than estimates in, e.g., Katz and Murphy (1992) and Ciccone and Peri (2005), but are also not directly comparable. In addition to some technical issues related to not using a CES labor aggregator, I also use an occupational rather than an educational definition of skill. Since the shares of workers in skilled occupations vary less with income than the share of highly educated workers, this assumption make the estimated elasticity of substitution smaller.

5In contrast, despite using a similar elasticity of substitution as in Caselli (2016), Morrow and Trefler (2017) find limited evidence for skill bias when using the World Input-Output Database and a college/non-college definition of skilled labor. The differences between the setups and the findings are discussed further Section ??.
complementary technologies, differences in the human capital of skilled workers (Jones, 2014a), or endogenous technology choices in response to different supplies of skilled workers (Caselli and Coleman, 2006; Caselli, 2016). These different interpretations are discussed in Section 5. Looking ahead, an important task is to further narrow down what underlies observed skilled labor efficiency differences.

The outline of the paper is as follows. Section 2 develops an industry-based notion of skill-biased efficiency differences. Section 3 develops the estimation strategy for estimating skill bias using trade data, factor share data, and factor price data. Section 4 presents the development accounting results. Section 5 discusses the interpretation of the findings and the relationship to the literature, and Section 6 concludes the paper.

**Related literature.** My paper is part of the development accounting literature, going back to Klenow and Rodriguez-Clare (1997) and Hall and C Jones (1999). This literature is surveyed in Caselli (2005), Hsieh and Klenow (2010), and C Jones (2015). There has been a number of papers revisiting the contribution of human capital in development accounting, most often in a framework featuring perfect substitutability between different types of labor services. These papers include Hendricks (2002), Erosa et al. (2010), Schoellman (2011), Manuelli and Seshadri (2014), and Hendricks and Schoellman (2017). Cubas et al. (2016) also studies quality differences in human capital between rich and poor countries.

A few papers have analyzed development accounting with imperfectly substitutable labor services. These papers include Caselli and Coleman (2006), Caselli and Ciccone (2013), B Jones (2014a), and Caselli (2016).

Beyond development accounting, my paper builds on the gravity trade literature to estimate the relative prices of skilled services (Tinbergen, 1962; Anderson et al., 1979; Eaton and Kortum, 2002; Anderson and van Wincoop, 2003; Redding and Venables, 2004; Costinot et al., 2011; Head and Mayer, 2014). A number of papers have used trade data to obtain information about productivities, including Trefler (1993) and Levchenko and Zhang (2016). Morrow and Trefler (2017) is a more recent contribution that integrates trade into development accounting. My paper also relates to the literature that uses industry data to obtain information about economic development, which includes Rajan and Zingales (1998) and Ciccone and Papaioannou (2009). In the context of trade, papers that analyze the relationship between country variables and the industrial structure of trade include Romalis (2004), Nunn (2007), Chor (2010), Cuñat and Melitz (2012), and Manova (2013). This literature is reviewed in Nunn and Trefler (2015).
2 An industry-based definition of skill biased efficiency differences

In this section, I show how disaggregated industry level data can be used to obtain information about the extent of skill-biased efficiency differences. The central idea is that a high efficiency of a factor reduces its efficiency-adjusted price, which reduces the relative unit production costs in industries that are intensive in that factor of production.

Section 2.1 gives a parametric representation of industry production functions that connects aggregate and industry level factor efficiency shifters. The key assumption is that differences in relative industry productivities across countries, i.e., comparative advantages, are only mediated by differences in relative factor efficiencies. Section 2.2 shows that, under this assumption, it is possible to derive an estimate of factor efficiencies based on projecting relative industry productivities on industry factor shares. Section 2.3 shows how this projection measure can be interpreted under a weaker set of assumptions that allows for industry-country specific productivity shifters.

2.1 Industry-level production functions and the aggregate production function

The baseline analysis makes the following assumption on how industry production functions vary across countries.

Assumption 1. For an industry $k$ and a country $i$, the industry production function $F^k_i$ can be expressed as

$$F^k_i(x^k_{i,1}, \ldots, x^k_{i,F}) = A_i F^k(Q^i_1 x^k_{i,1}, \ldots, Q^i_F x^k_{i,F}),$$

(1)

where $F^k$ is common across countries and satisfies standard conditions, $x^k_{i,f}$ is the amount of factor $f$ used in industry $k$ in country $i$, $Q^i_f$ are factor efficiency shifters, and $A_i$ is a uniform TFP shifter in country $i$.

The assumption implies that there exists a common industry production function $F^k$ across countries, and that particular production functions are obtained by adjusting this common function with TFP shifters and factor efficiency shifters, which are common across industries within a country.

Assumption 1 is similar to the one made in Trefler (1993), and it provides an industry structure that is consistent with an aggregate production function with factor-efficiency shifters, of the form commonly used in growth and development accounting. In particular, Assumption 1 implies a consistency between the industry level and the aggregate level in the following sense: if final output is produced using a constant-returns to scale aggregator $H$ of industry outputs, then there is an aggregate production function with the same factor efficiency terms as the industry production functions. This aggregation result is summarized in the following theorem.

---

6Trefler (1993) formulates this production function assumption in terms of the factor demand equation, assuming that efficiency-adjusted factor demands are equal for the same efficiency-adjusted factor prices. The two definitions are equivalent.
**Theorem 1.** Assume that there exists a set of industry strictly concave, constant returns to scale production functions $F^k$, common across countries, and a set of industry level factor augmenting efficiency differences $Q_{i,f}$ for $i = 1, \ldots, I$, $f = 1, \ldots, F$ such that

$$F^k_i(x^k_{i,f}, \ldots, x^k_{i,F}) = A_i F^k(Q_{i,f} x^k_{i,f}, \ldots, Q_{i,F} x^k_{i,F}) \quad \forall i, k,$$

and that industry outputs are aggregated using a common, strictly concave, constant returns to scale function $H$. If factor markets are competitive, there exists an aggregate production function $G_i$ for every country. Furthermore, there exists a function $G$, common across countries, such that the aggregate production function $G_i$ is given by

$$G_i(x_{i,1}, \ldots, x_{i,F}) = A_i G(Q_{i,1} x_{i,1}, \ldots, Q_{i,F} x_{i,F})$$

Proof. See Appendix A. \qed

Thus, Assumption 1 implies that industry level factor efficiency estimates can be used for aggregate productivity accounting.\(^7\)

### 2.2 Identifying factor efficiency with factor augmenting industry-level production functions

To derive an estimator of factor efficiency differences, it is helpful to restate Assumption 1 in dual form:

$$c^k_i(w_{i,1}, \ldots, w_{i,f}) = \frac{C^k_i \left( \frac{w_{i,1}}{Q_{i,1}}, \ldots, \frac{w_{i,F}}{Q_{i,F}} \right)}{A_i} \quad i = 1, \ldots, I \quad k = 1, \ldots, K \tag{2}$$

where $c^k_i$ is the unit cost of producing good $k$ in country $i$, and $w_{i,f}$ is the country $i$ cost of factor $f$. The function $C^k_i$ is the dual of the industry production function $F^k_i$. Note that factor costs enter jointly with the factor efficiency terms $Q_{i,f}$. Thus, unit cost differences are determined by efficiency-adjusted factor service prices $r_{i,f} = w_{i,f}/Q_{i,f}$.

The functional forms of the unit cost functions $c^k_i$ and $C^k_i$ are not known, so the equation system (2) cannot be estimated directly, even when having data on factor prices, factor shares, and realized unit costs.\(^8\) However, viewed as a function of factor prices, the cross-country variation in unit costs can be approximated using factor cost shares, which are observable. In particular, with competitive

---

\(^7\)The main drawback of Assumption 1 is that the productivity term $A_i$ does not depend on the industry $k$. This precludes comparative advantage effects not driven by the factor efficiency terms. This point reflects a general problem with any factor-augmenting specification of technology: with flexible comparative advantage, technology variation is more high dimensional than can be induced by low-dimensional variations due to factor efficiency variations. However, Section 2.3 shows that the factor efficiency measure derived under Assumption 1 is still interpretable when there are country-industry specific productivity terms.

\(^8\)If we are willing to make functional form assumptions in $c^k_i$, observed factor shares in conjunction with wages can be used to estimate factor-efficiency terms. This is the approach taken in Morrow and Trefler (2017), which assumes that the aggregate elasticity of substitution also holds within every industry.
factor markets, the second order approximation of the unit cost functions around the US factor prices is:

$$\log(c^k_i) = \log(c^k_{US}) - \log\left(\frac{A_i}{A_{US}}\right) + \sum_{f=1}^{F} \left(\frac{\alpha^k_{i,f} + \alpha^k_{US,f}}{2}\right) \log\left(\frac{r_{i,f}}{r_{US,f}}\right),$$

(3)

where $\alpha^k_{i,f}$ is the factor share of factor $f$ in country $i$ and industry $k$. The log deviation of unit costs of country $i$ compared to the US is the log relative TFP plus a weighted average of efficiency-adjusted relative factor service prices, using average factor shares between country $i$ and the US as weights. The approximation is exact when $\log C^k$ has a translog form (Diewert, 1976). In general, equation (3) is a second-order approximation, since the translog unit cost function can provide a second-order approximation to an arbitrary unit cost function that is homogeneous of degree 1 (Diewert, 1974). The following proposition uses (3) to derive a regression specification that identifies relative efficiency-adjusted factor prices $\frac{r_{i,f}}{r_{US,f}}$.

**Proposition 1.** Assume that there exist a set of country-specific TFP terms $A_i$, a set of factor efficiency shifters $Q_{i,f}$, and a set of country-industry TFP terms $\tilde{A}_k^k$ such that the industry unit cost functions satisfy

$$c^k_i(w_{i,1}, \ldots, w_{i,f}) = \frac{C^k_i\left(w_{i,1}/Q_{i,1}, \ldots, w_{i,f}/Q_{i,f}\right)}{A_i\tilde{A}_k^k}$$

where $\log C^k$ has a translog form for each industry, and $\log\left(\frac{\tilde{A}_k^k}{A^k_{US}}\right)$ are uncorrelated with factor shares $\frac{\alpha^k_{i,f} + \alpha^k_{US,f}}{2}$. Then, under competitive factor markets, relative efficiency-adjusted factor service prices $\frac{r_{i,f}}{r_{US,f}} = \frac{w_{i,f}/Q_{i,f}}{w_{US,f}/Q_{US,f}}$ satisfy the equation

$$\log\left(\frac{c^k_i}{c^k_{US}}\right) = -\log\left(\frac{A_i}{A_{US}}\right) + \log\left(\frac{r_{i,1}}{r_{US,1}}\right) + \sum_{f=2}^{F} \left(\frac{\alpha^k_{i,f} + \alpha^k_{US,f}}{2}\right) \log\left(\frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}}\right) - \log\left(\frac{\tilde{A}_k^k}{A^k_{US}}\right),$$

and efficiency-adjusted factor service prices in country $i$ can be identified by regressing unit costs on an industry fixed effect, a country fixed effect, and average industry factor shares of country $i$ and the US.

**Proof.** See Appendix A.

Using the estimated relative efficiency-adjusted prices $\frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}}$ in conjunction with observed factor prices, the relative factor efficiencies compared to the US can be estimated using

$$\frac{Q_{i,f}/Q_{i,1}}{Q_{US,f}/Q_{US,1}} \times \frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}} = \frac{w_{i,f}/w_{i,1}}{w_{US,f}/w_{US,1}}.$$

Thus, Proposition 1 provides a method for estimating relative factor efficiencies using industry level
2.3 Factor efficiency differences with industry-specific productivity terms

The identification in Proposition 1 assumes that the industry productivity terms $\tilde{A}^k_i$ are uncorrelated with industry factor shares. However, even when the set of $\tilde{A}^k_i$ are correlated with factor shares, the factor efficiency estimates can be interpreted. In this case, the estimates will be a combination of the factor efficiency shifters, plus the component of $\{\tilde{A}^k_i\}$ that is correlated with factor shares. This provides a natural interpretation of skill-biased efficiency differences: they reflect a combination of high human capital of skilled workers, factor-augmenting technology differences, and the component of country-industry productivities terms that is correlated with factor shares. The result is stated formally in the following proposition.9

**Proposition 2.** Suppose that the assumptions of Proposition 1 are satisfied, apart from the assumption that $\tilde{A}^k_i$ is uncorrelated with factor shares. Then, the estimated factor efficiency shifter $\hat{Q}_{i,f}$ satisfies

$$\log \hat{Q}_{i,f} = \log Q_{i,f} + \log \gamma_{i,f}^A$$

where $\log \gamma_{i,f}^A$ is the projection coefficient on factor $f$ when the vector $\{\log A^k_i\}_{k=1}^K$ is projected on the set of vectors $\{\frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2}\}_{k=1}^K$ for $f' = 1, \ldots, F$.

**Proof.** See Appendix A.

3 Estimating the relative price of skilled services

This section estimates skill-biased efficiency differences using the procedure developed in Section 2. Proposition 1 showed that this requires cross-country measures of:

- Industry factor shares
- Relative industry unit production costs
- Average relative wages of skilled and unskilled workers

The most challenging variable to measure is industry level unit production costs, since there is no dataset having industry level, quality-adjusted, producer prices from a large number of both rich and poor countries.10

---

9 Unfortunately, there is no direct analogue to the aggregation Theorem 1 when comparative advantage is flexible.  
10 The best available data set comes from the Groningen Growth and Development Center, which has done important work in constructing a data set of industry unit costs for cross-country comparisons (Inklaar and Timmer, 2008). However, their data set only covers 35 industries in 42 countries, with a limited coverage of poor countries. In Section 3.4, I show that for countries where we have both unit cost data and trade data, analyses using unit cost data and trade data yield similar results.
Table 1: Notation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td>Origin country</td>
</tr>
<tr>
<td>$j$</td>
<td>Destination country</td>
</tr>
<tr>
<td>$k$</td>
<td>Industry</td>
</tr>
<tr>
<td>$f$</td>
<td>Factor service ($f = 1$ unskilled labor services)</td>
</tr>
<tr>
<td>$x_{i,j}^k$</td>
<td>Export value of industry $k$ from country $i$ to country $j$</td>
</tr>
<tr>
<td>$r_{i,f}$</td>
<td>Factor service price of factor $f$ in country $i$</td>
</tr>
<tr>
<td>$\alpha_{i,f}^k$</td>
<td>Cost share of factor $f$ in industry $k$ in country $i$</td>
</tr>
<tr>
<td>$c_i^k$</td>
<td>Unit cost of industry $k$ in country $i$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Price elasticity of trade</td>
</tr>
</tbody>
</table>

To obtain information about relative unit production costs, I rely on indirect information from international trade data. Trade data is an attractive source of information, since it is recorded on a detailed industrial level across a wide range of countries, and, when interpreted through the lens of a gravity model, it contains implicit information about relative unit production costs. In particular, if rich countries export disproportionately in skill-intensive industries, this suggests that rich countries have low relative unit production costs in skill-intensive industries.

This section formalizes the connection between trade flows and unit costs using a gravity setup where the industry-origin fixed effect has an interpretation as a relative unit production cost.\(^{11}\) Combining the gravity-based unit cost estimates with industry factor shares data makes it possible to apply Proposition 1 to estimate efficiency-adjusted factor prices $r_{i,f} = \frac{w_{i,f}}{Q_{i,f}}$. In conjunction with observed factor prices $w_{i,f}$, these can be used to estimate factor efficiency levels.

Section 3.1 presents the environment and the regression setup. Section 3.2 presents data sources, data on factor prices and factor shares, parameter estimates, and descriptive data. Section 3.3 reports the results, and Section 3.4 discusses identification threats and conducts robustness checks.

3.1 Environment and regression setup

There are $I = 96$ countries, and each country has $K = 84$ industries, corresponding to the NAICS four-digit manufacturing industries.\(^{12}\) The value of trade flows from country $i$ to country $j$ in industry $k$ is denoted $x_{i,j}^k$. Each industry produces a good using $F = 4$ factor services: unskilled labor, skilled labor, capital, and traded intermediate inputs. The notation is summarized in Table 1.

In line with the assumptions in Proposition 1, it is assumed that log unit costs can be written

\(^{11}\)The closest analogue in the literature is the specification used in Levchenko and Zhang (2014).
\(^{12}\)The countries correspond to the 83 countries with total manufacturing exports exceeding $1bn USD, and with a GDP per worker higher than $5,000.
as
\[ \log(c^k) = \log(c^k_{US}) + a_i + \sum_{f=2}^{F} \left( \frac{\alpha_{US,f}^k + \alpha_{i,f}^k}{2} \right) \log \left( \frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}} \right), \]  
(4)
where \( a_i = \log \left( \frac{r_{i,1}}{r_{US,1}} \right) - \log \left( \frac{\tilde{A}_i}{\tilde{A}_{US}} \right) \) is the log deviation in unskilled labor service prices, adjusted for absolute productivity differences.

In addition to the unit cost equation, it is also assumed that trade flows are governed by a gravity equation:
\[ \log(x^k_{i,j}) = \delta_{i,j}^k + \mu^k_{j} + \sum_{m=1}^{M} \gamma_{m}^k d_{ij,m} - (\sigma - 1) \log(c^k_i). \]  
(5)
In this equation, the log export value from country \( i \) to country \( j \) in industry \( k \) depends on four terms. The first term is a bilateral fixed effect \( \delta_{i,j}^k \). It captures unilateral and bilateral determinants of trade flows such as the size of the two countries, their bilateral distances, common legal origins, shared language, etc. The second term is a destination-industry fixed effect \( \mu^k_{j} \), which captures the demand for good \( k \) in destination \( j \), as well as how good access country \( j \) has to industry \( k \), given its other trading partners. The third term represents industry-specific coefficients on a set of gravity terms, allowing for heterogeneous trade costs across industries. The fourth term captures that, conditional on the first three terms, exports depend negatively on origin unit production costs, with a price elasticity \( \sigma - 1 \). In Appendix B.1, I show how equation (5) can be derived from both a trade model in the style of Eaton and Kortum (2002), where trade is driven by country-variety specific productivity shocks, and from an Armington model where each country produces a unique variety of each good \( k \).

### 3.1.1 Regression specification
Combining the gravity equation (5) and the unit cost equation (4) yields:
\[ \log(x^k_{i,j}) = \tilde{\delta}_{i,j} + \tilde{\mu}_{j}^k + \sum_{m=1}^{M} \gamma_{m}^k d_{ij,m} - (\sigma - 1) \sum_{f=2}^{F} \left( \frac{\alpha_{US,f}^k + \alpha_{i,f}^k}{2} \right) \log \left( \frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}} \right). \]

Here, \( \tilde{\delta}_{i,j} = \delta_{i,j} - (\sigma - 1) \left( \log \left( \frac{r_{i,1}}{r_{US,1}} \right) - \log \left( \frac{\tilde{A}_i}{\tilde{A}_{US}} \right) \right) \) denotes a modified fixed effect that includes the trade bilateral fixed effect, the origin absolute advantage, and the origin unskilled factor service prices. The term \( \tilde{\mu}_{j}^k = \mu^k_{j} - (\sigma - 1) \log(c^k_{US}) \) denotes a modified fixed effect that includes the trade destination-industry fixed effect \( \mu^k_{j} \) and US industry unit costs.

To convert this equation into a regression specification, we note that, apart from the fixed effects, \( x^k_{i,j} \) is directly measured from international trade data, the gravity terms \( d_{ij,m} \) are obtainable from standard trade data sets, \( \alpha_{US,f}^k \) is measurable from US industry data, and \( \sigma \) is estimated in the
Direct industry factor share measures across countries are lacking, but $\alpha_{k,i,f}$ can be approximated using detailed US estimates in combination with more aggregated international measures on labor, capital, and skilled and unskilled labor shares.

The only unknown terms are $\log\left(\frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}}\right)$, which are estimated as parameters that vary on a country-factor basis. I write $\beta_{i,f}$ for this set of parameters. Thus, the equation can be turned into a regression specification where the dependent variable is $\log(x_{k,i,j}^f)$, the explanatory variables are fixed effects, gravity variables, and $\left(\sigma - 1\right)\frac{\alpha^k_{US,f} + \hat{\alpha}_{i,f}}{2}$ for $f = 2, \ldots, F$. Given the interpretation of $\beta_{i,f}$ as relative factor prices compared to the US, it is assumed that $\beta_{US,f} = 0$ for all $f$. The regression equation becomes:

$$\log(x_{k,i,j}^f) = \delta_{i,j} + \mu_j^f + \sum_{m=1}^{M} \gamma_{m}^k d_{ij,m} - \sum_{f=2}^{F} \left(\sigma - 1\right)\frac{\alpha^k_{US,f} + \hat{\alpha}_{i,f}}{2} \times \beta_{i,f} + \epsilon_{k,i,j}, \quad (6)$$

with the normalization $\beta_{US,f} = 0$ for $f = 2, \ldots, F$. There are $(4 - 1) \times 83 = 249$ parameters $\beta_{i,f}$: one for each country-factor combination, excluding unskilled labor services. With this regression specification, $\beta_{i,f}$ identifies $\log\left(\frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}}\right)$. The main regression specification is run using a log-linear OLS. In Section 3.4, I consider the effect of running a PPML specification, as suggested by Silva and Tenreyro (2006), that allows for zero trade flows.

### 3.1.2 Intuition for the regression specification

The intuition behind the regression specification can be explained using a setup with two factors, three countries, and two industries that both have Cobb-Douglas production functions, common across countries. Thus, Vietnam consider a case where the US and Vietnam export to Japan. There are two industries: “Textile Furnishing Mills” (NAICS code 3141) and “Communications Equipment” (NAICS code 3342), both producing using a Cobb-Douglas production function with unskilled and skilled labor:

$$c_{k}^i = \left(\frac{w_{i,u}}{Q_{i,u}}\right)^{\alpha_u^k} \left(\frac{w_{i,s}}{Q_{i,s}}\right)^{\alpha_s^k} \delta = (r_{i,u})^{\alpha_u^k} (r_{i,s})^{\alpha_s^k} \quad k = Com, Text; \quad i = US, Vietnam, \quad (7)$$

where $\alpha_u^k + \alpha_s^k = 1$ for $k = Com, Text$.\(^{14}\) Table 2 shows the US factor shares in the two industries when we restrict attention to labor input. Cut and Sew Apparel has a skilled labor share of 0.45, whereas Communications Equipment has a skilled labor share of 0.83.

\(^{13}\)Some papers estimate $\sigma$ directly from trade data (Broda et al., 2006; Soderbery, 2015), exploiting short-run variations in trade prices and quantities. As I am interested in the long-run elasticity of trade, I choose a calibration approach to select $\sigma$. In Section 3.4, I explore the consequence of allowing $\sigma$ to depend on the industry.

\(^{14}\)I have not included any country-industry production terms $A_{k}^i$; in the two industry case, it is easy to show that they could be redefined as factor-efficiency terms. When there are more than two industries, but still Cobb-Douglas production functions, the factor-efficiency terms cannot be removed completely, but it is possible to adjust the factor-efficiency shifters to make the $A_{k}^i$-terms uncorrelated with the factor shares.
Starting with the derivation of relative unit costs from trade flows, the gravity equation (5), can be re-arranged to

\[
\log \left( \frac{c_{\text{Com}}/c_{\text{App}}}{c_{\text{US}}/c_{\text{US}}} \right) = \left( -\frac{1}{\sigma - 1} \right) \log \left( \frac{x_{\text{Com}}/x_{\text{App}}}{x_{\text{US,Jap}}/x_{\text{US,Jap}}} \right)
\]

The equation is obtained using a double-differencing procedure: differencing across countries removes the Japan demand effects; differencing within countries removes the bilateral fixed effect and absolute productivity advantages. The equation states that a country has low relative unit costs in an industry whenever it has high relative exports in that industry. The trade elasticity enters in the denominator: a high trade elasticity implies that smaller variations in unit costs suffice to rationalize a set of given trade patterns.

To calibrate relative unit costs, Table 3 is used, which gives the dollar export values for the relevant trade flows. The table shows that the US has large relative exports in communications equipment, and that Vietnam has large relative exports in textile furnishings. Using an assumed trade elasticity of \( \sigma = 7.5 \), the trade flows in Table 3 imply:

\[
\log \left( \frac{c_{\text{Com}}/c_{\text{App}}}{c_{\text{US}}/c_{\text{US}}} \right) = \left( -\frac{1}{7.5 - 1} \right) \times \log \left( \frac{5.5/70.5}{252/42} \right) \\
\approx 0.66
\]

Unsurprisingly, communication equipment is relatively expensive in Vietnam: given the trade flows, the cost disadvantage is estimated at 66 log points.

To obtain relative factor prices from unit cost estimates, a similar double-differencing method is applied to the (log) unit cost equation (7):

\[
\log \left( \frac{c_{\text{Ind}}/c_{\text{Ind}}}{c_{\text{US}}/c_{\text{US}}} \right) = (\alpha_{s}^{\text{Com}} - \alpha_{s}^{\text{App}}) \times \log \left( \frac{r_{\text{Ind,s}}/r_{\text{Ind,u}}}{r_{\text{US,s}}/r_{\text{US,u}}} \right).
\]

The deviations in relative unit costs across industries are driven by deviations in relative factor service prices, times the differences in factor shares across the two industries. The more different the industries are in factor shares, the less shifts in factor service prices are needed to justify the unit cost differences. Using that \( \alpha_{s}^{\text{Com}} - \alpha_{s}^{\text{App}} = 0.41 \), we obtain

\[
\log \left( \frac{r_{\text{Ind,s}}/r_{\text{Ind,u}}}{r_{\text{US,s}}/r_{\text{US,u}}} \right) = \frac{1}{0.41} \times 0.66 = 1.61.
\]

Given this set of data, Vietnam is estimated to have a 161 log points, or 5 times, higher efficiency-

\(^{15}\)Here, a gravity model without industry-dependent trade costs is analyzed, as the estimation of industry-dependent trade costs requires more than one trading destination.
adjusted relative price of skilled versus unskilled labor services. The difference in skilled wage premium is approximately 1.5, which means that the relative efficiency of skilled labor needs to be three times higher in the US. The full regression specification generalizes this logic to more industries, production factors, and countries.

Table 2: Factor shares for Textile Furnishing Mills and Communication Equipment

<table>
<thead>
<tr>
<th>Industry</th>
<th>Unskilled share</th>
<th>Skilled share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communications Equipment</td>
<td>0.27</td>
<td>0.73</td>
</tr>
<tr>
<td>Textile Furnishings Mills</td>
<td>0.68</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 3: Export values from Vietnam and USA to Japan (thousands of US dollars)

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
<th>Industry</th>
<th>Export value</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>Japan</td>
<td>Communications Equipment</td>
<td>251904.48</td>
</tr>
<tr>
<td>USA</td>
<td>Japan</td>
<td>Textile Furnishings Mills</td>
<td>42249.56</td>
</tr>
<tr>
<td>Vietnam</td>
<td>Japan</td>
<td>Communications Equipment</td>
<td>5567.40</td>
</tr>
<tr>
<td>Vietnam</td>
<td>Japan</td>
<td>Textile Furnishings Mills</td>
<td>70468.05</td>
</tr>
</tbody>
</table>

3.2 Data in trade regression

The regression equation (6) requires data on bilateral trade flows \(x_{ij}^k\), gravity terms \(d_{ij,m}\), US factor shares \(\alpha_{US,f}^k\), estimates of international factor shares \(\alpha_{i,f}^k\), and a parameter estimate for the trade elasticity \(\sigma\). To calculate relative efficiencies from \(r_s/r_u\), skilled wage premia in the manufacturing sectors are also needed.

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 at the HS 2007 six-digit industry level.

I obtain the US labor, capital, and intermediate input shares from the BEA. To separate labor into skilled and unskilled shares, I use the OES to calculate the share of payroll in each industry that goes to workers in occupations with skill levels 3 and 4 in the ISCO-08 classification. This corresponds to the major occupational groups "Managers", "Professionals", and "Technicians and Associate Professionals". I calculate the skill share as the labor share from the BEA times the share of payroll going to skilled workers, and the unskilled share as the labor share times the share of payroll going to unskilled workers. The US IO-table is used to resolve the intermediate input content into basic factor requirements and tradable intermediate inputs; this approach is described in Appendix B.2. Note that the skill definition is based on occupations rather than education; this choice is discussed in Appendix B.3.

To estimate the international factor shares \(\alpha_{i,f}^k\), I use international factor shares data on a higher level of aggregation. The World Input Output Database reports labor and intermediate input
compensation shares in manufacturing across 42 countries. I regress these compensation shares on
GDP per capita to create a predicted labor, capital, and intermediate input compensation shares
for each country (normalizing so that the US obtains the compensation shares observed in the BEA,
and that the intermediate input compensation share matches the tradable intermediate share in
the US). I similarly create a predicted split of labor compensation into skilled and unskilled labor
shares by using IPUMS data on the occupational composition and relative wages in manufacturing.
Given the lower precision of international factor share data.

The skilled wage premia in manufacturing are obtained using IPUMS data on wages for different
occupations in manufacturing. Given that IPUMS income data is only available in a limited
selection of countries, the skilled wage premia in other countries is imputed using a linear regression
of log skilled wage premia on log GDP per worker.

The regression is performed using NAICS four-digit coding, which is the coding scheme of the
OES industry data. The trade data is recorded using HS6 codes. The OES occupational data is
recorded according to SOC, and it is converted to ISCO-08 to calculate the share of payroll going
to skilled workers. The BEA data is recorded in the Input-Output coding scheme. All factor share
and trade data are converted between coding schemes using a concordance procedure described in
Appendix D.

The value of the trade elasticity $\sigma$ is taken from the literature. I look for an estimate of the
\textit{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 $\sigma$, the lower is the importance of skilled labor efficiency differences, since
it reduces the required differences in relative efficiency prices required to explain trade patterns.

I vary the trade elasticity between $\sigma = 5$ and $\sigma = 10$. This reflects a range of estimates found
in the literature. Simonovska and Waugh (2014) report $\sigma = 5$, Costinot et al. (2011) use $\sigma = 7.2$
and Eaton and Kortum find $\sigma = 9.2$ found in Eaton and Kortum (2002).\footnote{Note that the trade elasticity $\theta$ in Eaton and Kortum-style models represents the elasticities of export value with respect to price changes, whereas $\sigma$ represents the elasticity of quantity with respect to price changes. Hence, $\sigma = \theta + 1$ when we convert between the two types of parameters. $\sigma = 10$ corresponds to the higher range estimates found in Romalis (2007) when he estimates the trade effects of NAFTA. Unless stated otherwise, I report the results for $\sigma = 7.5$. The trade elasticity is a key parameter as it determines how large unit cost differences that are needed to rationalize the trade patterns.}

3.2.1 Data descriptives

Figure 1 shows the distribution across different manufacturing industries of the payroll share of
skilled labor payroll, that is, the share of payroll going to workers belonging to the occupational
groups "Managers", "Professionals", and "Technicians and Associate Professionals". This share ranges from approximately 25% to 90%. Table 4 and Table 5 show the 10 industries with the highest and lowest skilled labor payroll shares.

![Figure 1: Histogram over skilled labor share across manufacturing industries](image)

**Table 4: Top 10 Industries by skilled payroll share**

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Skill share</th>
</tr>
</thead>
<tbody>
<tr>
<td>3341</td>
<td>Computer and Peripheral Equipment Manufacturing</td>
<td>0.88</td>
</tr>
<tr>
<td>3345</td>
<td>Navigational, Measuring, Electromedical, and Control Instruments Manufacturing</td>
<td>0.76</td>
</tr>
<tr>
<td>3342</td>
<td>Communications Equipment Manufacturing</td>
<td>0.76</td>
</tr>
<tr>
<td>3254</td>
<td>Pharmaceutical and Medicine Manufacturing</td>
<td>0.75</td>
</tr>
<tr>
<td>3344</td>
<td>Semiconductor and Other Electronic Component Manufacturing</td>
<td>0.70</td>
</tr>
<tr>
<td>3343</td>
<td>Audio and Video Equipment Manufacturing</td>
<td>0.67</td>
</tr>
<tr>
<td>3364</td>
<td>Aerospace Product and Parts Manufacturing</td>
<td>0.67</td>
</tr>
<tr>
<td>3346</td>
<td>Manufacturing and Reproducing Magnetic and Optical Media</td>
<td>0.64</td>
</tr>
<tr>
<td>3391</td>
<td>Medical Equipment and Supplies Manufacturing</td>
<td>0.63</td>
</tr>
<tr>
<td>3333</td>
<td>Commercial and Service Industry Machinery Manufacturing</td>
<td>0.62</td>
</tr>
</tbody>
</table>

To gauge the relationship between income levels and export patterns in skill intensive industries, industries are divided into quintiles based on their skilled labor payroll share. Figure 2 shows the relationship between country income levels and the share of exports in the highest quintile and the two highest quintiles of skilled labor payroll share. There is a strong pattern that rich countries export more skill-intensive goods than poor countries. E.g., for the two highest skilled labor share
Figure 2: Share of export value coming from top quintile and top 2 quintiles of skill intensity
Figure 3: Share of export value coming from top quintile and top 2 quintiles of capital intensity
Table 5: Bottom 10 Industries by skilled payroll share

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Skill share</th>
</tr>
</thead>
<tbody>
<tr>
<td>3219</td>
<td>Other Wood Product Manufacturing</td>
<td>0.38</td>
</tr>
<tr>
<td>3274</td>
<td>Lime and Gypsum Product Manufacturing</td>
<td>0.37</td>
</tr>
<tr>
<td>3371</td>
<td>Household and Institutional Furniture and Kitchen Cabinet Manufacturing</td>
<td>0.37</td>
</tr>
<tr>
<td>3262</td>
<td>Rubber Product Manufacturing</td>
<td>0.37</td>
</tr>
<tr>
<td>3273</td>
<td>Cement and Concrete Product Manufacturing</td>
<td>0.36</td>
</tr>
<tr>
<td>3211</td>
<td>Sawmills and Wood Preservation</td>
<td>0.35</td>
</tr>
<tr>
<td>3141</td>
<td>Textile Furnishings Mills</td>
<td>0.33</td>
</tr>
<tr>
<td>3131</td>
<td>Fiber, Yarn, and Thread Mills</td>
<td>0.31</td>
</tr>
<tr>
<td>3151</td>
<td>Apparel Knitting Mills</td>
<td>0.29</td>
</tr>
<tr>
<td>3116</td>
<td>Animal Slaughtering and Processing</td>
<td>0.28</td>
</tr>
</tbody>
</table>

Figure 4: Skilled factor share and average fixed effect in poor and rich countries
quintiles, rich countries have approximately 75% of their exports in these quintiles, whereas poor countries have only approximately 20% of their export values in these industries.

For comparison, there is a much weaker relationship for capital intensity, which can be seen in Figure 3 that repeats the same exercise for capital intensity (defined as the capital share of value added). The export share for the highest two quintiles goes from approximately 35% to just above 50%. For the highest quintile, the relationship is even weaker.

The country-industry fixed effects (corresponding to the value $-(\sigma - 1) \log(e^K)$) in the gravity equation (5) can also be used to illustrate the relationship between income levels, skilled labor factor shares, and trade flows. These fixed effects are displayed in Figure 4. In both panels, the horizontal axis is the skilled payroll share, and the vertical axis gives the value of the fixed effect. The left panel shows the results for poor countries, and the right panel for rich countries. The graph is a bplot, and since there are approximately 7,500 fixed effects, one for every country-industry combination, and 20 points in each panel, each point represents approximately 200 fixed effects. The structure of fixed effects mirror the raw correlation in Figure 2: poor countries have a negative relationship between the skilled payroll share and the fixed effects, whereas the relationship is positive in rich countries. Thus, the fixed effects analysis shows that poor countries tend to have low relative exports, and thus high relative unit production costs, in industries with a high skilled payroll share.

### 3.3 Results

The main results are displayed in abridged form in Table 6. The table displays the estimate of log efficiency-adjusted factor service prices $\beta_{i,f} = \log \left( \frac{r_{i,f}}{r_{US,f}} \right)$ for the different production factors, and for four randomly selected countries in every income group. The prices are relative to unskilled labor, with US prices normalized to 1. The standard errors are clustered on a country-industry level.

The table shows that poor countries tend to have high relative prices of factors other than unskilled labor. The tendency is strongest for the relative price of skilled labor. The systematic patterns across the income distribution for skilled labor is summarized in Table 9, which shows the relative price of skilled labor services across income groups for different trade elasticities. The richest group of countries is normalized to 1. The relative price of skilled labor in the poorest countries range from 5 to 62 times higher than in the richest countries, depending on the assumed trade elasticity; a higher trade elasticity means that less relative price differences are needed to justify the trade patterns.

Figure 5 graphically illustrates the regression results. The vertical axis is a log scale of relative prices of skilled and unskilled labor services, where the relative price of skilled labor services in rich countries is again normalized to 1. There is a negative, approximately linear, relationship between income and the relative price of skilled labor services. There are a few exceptions of poor countries
with low estimated relative prices of skilled services (e.g., the Philippines and Costa Rica), and
some exceptions with rich countries that have relatively high estimated skill prices (e.g., the oil
countries and Macau).

Table 6: Trade regression results in an abridged form

<table>
<thead>
<tr>
<th>Output per worker: $5000-15,000</th>
<th>Skill</th>
<th>Capital</th>
<th>Tradable intermediates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zambia</td>
<td>3.65 (0.00)</td>
<td>0.73 (0.59)</td>
<td>0.37 (0.97)</td>
</tr>
<tr>
<td>Ghana</td>
<td>5.55 (1.45)</td>
<td>1.61 (0.46)</td>
<td>1.55 (0.68)</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>4.16 (1.23)</td>
<td>0.72 (0.37)</td>
<td>1.20 (0.67)</td>
</tr>
<tr>
<td>Peru</td>
<td>4.12 (0.96)</td>
<td>1.57 (0.38)</td>
<td>1.58 (0.66)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output per worker: $15,000-$30,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ukraine</td>
</tr>
<tr>
<td>Brazil</td>
</tr>
<tr>
<td>Syrian Arab Republic</td>
</tr>
<tr>
<td>Uruguay</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output per worker: $30,000-$50,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latvia</td>
</tr>
<tr>
<td>Mexico</td>
</tr>
<tr>
<td>Lithuania</td>
</tr>
<tr>
<td>Algeria</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output per worker: $50,000-$75,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malta</td>
</tr>
<tr>
<td>Greece</td>
</tr>
<tr>
<td>Cyprus</td>
</tr>
<tr>
<td>Japan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output per worker: &gt;$75,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
</tr>
<tr>
<td>France</td>
</tr>
<tr>
<td>Bahrain</td>
</tr>
<tr>
<td>Australia</td>
</tr>
<tr>
<td>USA</td>
</tr>
</tbody>
</table>

| Observations                      | 396645 |
| R2                                | 0.68  |

To obtain the skill bias of technology, the regression estimates are complemented with data on
skilled wage premia. Table 8 shows the relative efficiency adjusted prices of skilled labor services,
the observed relative factor prices, and the relative efficiency of skilled and unskilled workers.
Here, we see that the relative skilled labor efficiency is approximately 8 times higher in the poorest
countries, compared to in the richest countries, when σ = 7.5.
<table>
<thead>
<tr>
<th>GDP per worker range</th>
<th>$\sigma = 5$</th>
<th>$\sigma = 7.5$</th>
<th>$\sigma = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$15,000</td>
<td>49.17</td>
<td>10.99</td>
<td>5.65</td>
</tr>
<tr>
<td>$15,000-$30,000</td>
<td>27.70</td>
<td>7.72</td>
<td>4.38</td>
</tr>
<tr>
<td>$30,000-$50,000</td>
<td>11.73</td>
<td>4.55</td>
<td>2.99</td>
</tr>
<tr>
<td>$50,000-$75,000</td>
<td>1.64</td>
<td>1.36</td>
<td>1.25</td>
</tr>
<tr>
<td>&gt;$75,000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 7: Relative price of skilled and unskilled labor services for different trade elasticities

Figure 5: GDP per worker and efficiency-adjusted relative price of skilled labor services

3.4 Robustness checks

This section performs robustness checks of the estimation of efficiency shifters. The robustness checks are informed by Propositions 1 and 2, on which the estimation exercise was based. Proposition 1 showed that, if relative industrial productivities were driven by relative factor efficiencies, competitive factor markets implied that the projection of relative unit costs on factor shares correctly identified relative efficiency-adjusted factor service prices. Proposition 2 further showed that, with additional country-industry specific productivity shifters, the projection estimator of a factor
#### Table 8: Skilled labor efficiencies ($\sigma = 7.5$)

<table>
<thead>
<tr>
<th>GDP per worker range</th>
<th>$\frac{w_s}{w_u}$</th>
<th>$\frac{r_s}{r_u}$</th>
<th>$Q_s/Q_u = \frac{w_s}{r_s}/\frac{w_u}{r_u}$</th>
<th>$Q_s/Q_u$ (normalized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$15,000</td>
<td>3.51</td>
<td>10.99</td>
<td>0.32</td>
<td>0.14</td>
</tr>
<tr>
<td>$15,000-$30,000</td>
<td>2.97</td>
<td>7.72</td>
<td>0.39</td>
<td>0.17</td>
</tr>
<tr>
<td>$30,000-$50,000</td>
<td>2.63</td>
<td>4.55</td>
<td>0.58</td>
<td>0.26</td>
</tr>
<tr>
<td>$50,000-$75,000</td>
<td>2.39</td>
<td>1.36</td>
<td>1.75</td>
<td>0.80</td>
</tr>
<tr>
<td>&gt;$75,000</td>
<td>2.20</td>
<td>1.00</td>
<td>2.20</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The shifter was interpretable as the sum of the factor-efficiency shifter and the component of productivity shifters correlated with that particular factor. Hence, the robustness discussion will focus on whether unit costs and factor shares are correctly measured, and whether the factor shares correspond to the elasticity of output with respect to their corresponding factor (which is the implication of the competitive factor market assumption that is used in the propositions).  

### 3.4.1 Unit cost estimation

The relative unit costs are estimated using the gravity specification:

$$\log(x_{i,j}^k) = \delta_{i,j} + \mu_j^k + \sum_{m=1}^{M} \gamma_m^k d_{ij,m} - (\sigma - 1) \log(c_i^k) + \varepsilon_{i,j}^k.$$  

I first test for the effect of having heterogeneous trade elasticities. To do this, I modify the regression specification (6) by changing the term $\sigma$ to an industry-dependent trade elasticity $\sigma_k$, and by adding an extra control variable $\gamma_i(\sigma_k - 1)$. The second term is an adjustment for the fact that heterogeneous trade elasticities imply that absolute advantages lead to changes in relative trade flows. This effect means that, in general, an adjustment is needed for the fact that if a country is cheap in all industries, it will export disproportionately in industries with a high trade elasticity. Appendix C.1 shows that, under some structural assumptions, this problem can be addressed by including a term $\gamma_i(\sigma_k - 1)$. I use industry-specific trade elasticities $\sigma_k$ from Caliendo and Parro (2015).  

---

17 It was also noted in Section 2 that the aggregation Theorem 1 does not apply directly when there were industry-specific productivity shifters. Essentially, with flexible comparative advantages, the true aggregate production function becomes $A_i G(Q_{i,1},x_{i,1},\ldots,Q_{i,F},x_{i,F}; \tilde{A}_i^1,\ldots,\tilde{A}_i^K)$, and generally depends on the whole set of country-industry productivity terms. Thus, even though the estimated $Q_{i,f}$ are interpretable when there are industry-specific productivity shifters, there are multiple ways to use them in the aggregate exercise. For the development accounting exercise in this paper, the average estimate of $Q_{i,f}$ in an income group is used as the efficiency shifter in this income group, meaning that the systematic component is attributed to the factor-efficiency shifters. A more careful study of this problem would involve characterizing aggregation with flexible comparative advantages. Since country-industry productivity differences are likely reasons for the estimated dispersions in factor-efficiencies for given income level, this is an interesting area of future study.

18 For computational reasons, I also replace the bilateral fixed effect $\delta_{i,j}$ with a source country fixed effect $\delta_i$. In
and homogeneous trade elasticities. For comparability, I normalize $\sigma_k$ so that the trade-weighted median elasticity is equal to the homogeneous trade elasticities $\sigma = 5, 7.5, 10$. Heterogeneous trade elasticities make the pattern across rich and poor countries somewhat weaker.

Table 9: Relative price of skilled and unskilled labor services with heterogeneous trade elasticities

<table>
<thead>
<tr>
<th>GDP per worker range</th>
<th>$\sigma = 5$</th>
<th>$\sigma = 7.5$</th>
<th>$\sigma = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,000-15,000</td>
<td>21.02</td>
<td>6.51</td>
<td>3.87</td>
</tr>
<tr>
<td>15,000-30,000</td>
<td>8.47</td>
<td>3.72</td>
<td>2.58</td>
</tr>
<tr>
<td>30,000-50,000</td>
<td>7.88</td>
<td>3.56</td>
<td>2.50</td>
</tr>
<tr>
<td>50,000-75,000</td>
<td>1.21</td>
<td>1.13</td>
<td>1.09</td>
</tr>
<tr>
<td>&gt;75,000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Furthermore, I test the robustness with respect to zero trade flows. In the baseline regression, I run a log-linear specification and drop observations with zero trade flows. For robustness, I rerun the regression using a Poisson Pseudo-Maximum Likelihood (PPML) estimator (Silva and Tenreyro, 2006) of the gravity model, which allows for zeros. It is unfortunately not computationally feasible to perform a PPML estimation for the whole sample and the complete set of fixed effects. Thus, the sample is restricted to countries with more than $5bn in manufacturing exports, and to industries with more than 0.5% of world trade. This leaves 64 countries and 49 industries. Furthermore, the number of fixed effects and parameters are reduced by running the regression:

$$\log(x_{i,j}^k) = \tilde{\delta}_{i,j} + \mu^k + \gamma_j - \sum_{f=2}^{F} \left[ (\sigma - 1) \frac{\alpha_{US,f}^k + \hat{\alpha}_{i,f}^k}{2} \right] \times \beta_{i,f} + \varepsilon_{i,j}^k.$$  

Compared to the main regression specification (6), the destination-industry fixed effects are replaced by industry fixed effects and destination fixed effects. To distinguish the effects of the changed estimation method from the changed regression specification, I compare the PPML estimates with a log-linear specification run on the same countries and using both the baseline specification, and a specification with a more limited set of fixed effects. Table 10 shows the result, where the first column gives the result from the baseline regression, the second and third column give the results from the restricted specification using log-linear and PPML respectively. Note that the results become larger when zero trade flows are allowed for and a PPML approach is used. Since the results in the first two columns are quite similar, with the baseline regression even giving somewhat stronger results, this is unlikely to be driven by the regression specification.

Figure 6 shows the PPML estimates in a graphical form. The figure shows that with a PPML estimation, the estimates align closely to a log linear relationship between relative prices and GDP per worker. Given the PPML estimates, the regression of log relative prices on log GDP per worker has an $R^2$-value of 0.34. For the log-linear specification, the corresponding figure is only 0.12.

Appendix C.1, I show that this does not lead to large differences in the baseline estimation, which suggest that it is not central to the estimation.
Figure 6: Estimates of relative efficiency-adjusted skill prices using PPML estimator
Table 10: Relative price estimates with PPML gravity estimation

<table>
<thead>
<tr>
<th>GDP/worker range</th>
<th>Log-linear</th>
<th>Log-linear (restricted)</th>
<th>PPML (restricted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$30,000</td>
<td>4.92</td>
<td>4.38</td>
<td>127.38</td>
</tr>
<tr>
<td>$30,000-$50,000</td>
<td>3.26</td>
<td>3.09</td>
<td>11.37</td>
</tr>
<tr>
<td>$50,000-$75,000</td>
<td>1.64</td>
<td>1.50</td>
<td>3.67</td>
</tr>
<tr>
<td>&gt;$75,000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

A likely reason for the larger size of the PPML estimates is that fitting the extensive margin of skill-intensive trade requires larger imputed skill-price differences, to capture that poor countries export few skill-intensive goods.\(^{19}\)

Lastly, I also perform my exercise using unit cost data for the set of countries where these are available. For this, I use data from the Groningen Growth and Development Center (GGDC), which has constructed a unit cost measure for 34 industries across 42 countries (Inklaar and Timmer, 2008). A natural consistency check is whether my trade data method yields similar conclusions as a unit cost based method on this set of countries.

Following correspondence with one of the creators of the data set, I exclude financial services, business services, real estate, government, health services and education. For these industries, it is difficult to obtain data on output quantities which makes it difficult to make cross-country comparisons in unit costs. I also exclude "private households with employed persons" as this variable is missing for a large number of countries. After my exclusions, I am left with a total of 27 industries and 34 countries with a complete set of observations.\(^{20}\)

The GGDC index covers both tradable and non-tradable industries, and manufacturing as well as services. Using the GGDC data set, I can run a unit cost regression to estimate relative factor service prices.

\[
\log(c^k_i) = \delta_i + \mu_k + \sum_{f=2}^{F} \alpha_{US,f}^k + \alpha_{i,f}^k \beta_{i,f}. 
\]

Here, \(\delta_i\) is a country-fixed effect, \(\mu_k\) is an industry-fixed effect, and \(\beta_{i,f}\) identifies the country-factor relative factor service price differences.

Figures 7 shows the relationship between estimated log relative skilled service prices and log

\(^{19}\)In this specification, the intensive margin trade elasticity \((\sigma = 7.5)\) was used both for the log-linear specification and for the PPML. The use of a higher trade elasticity for the PPML estimation would lower the estimates. However, a full exploration of the degree to which the trade elasticity should be adjusted when zeros are included would require a model with structural zeros, as in, e.g., ?.

\(^{20}\)To obtain factor shares, I use BEA data for the US. I define the labor share as the labor compensation over gross output, and the intermediate share as the intermediate good compensation over gross output. I calculate the skill share by multiplying the labor share with the share of payroll going to skilled workers with an occupational skill level of 3 or 4. I define the capital share as one minus the other factor shares. In contrast to the baseline specification, I do not resolve the non-traded intermediate inputs using an input output table given the lower granularity of the data. The foreign input shares are constructed as in the baseline specification.
GDP per worker. Like the trade based estimates, they exhibit a strong negative correlation with log GDP per worker. The slope parameter of log relative skilled service prices is similar to the one found for the trade data when $\sigma = 5$. Thus, when both types of data exist, the trade data method and the unit cost method paint a similar picture of the relationship between relative skilled service prices and income per worker.

### 3.5 Factor shares measurement

Proposition 1 showed that unit costs should be projected on factor shares to recover efficiency-adjusted factor prices. In addition to assuming that factor shares are correctly measured, the proposition was based on the assumptions that factor markets were competitive, i.e., that factor shares equal unit cost factor elasticities. In this section, I analyze the consequence of relaxing the assumption of competitive factor markets.

To explore the consequence of imperfect competition, I assume that there is a simple constant markup across all goods. If there is a markup, the estimated labor and materials shares become larger as total costs are lower than the value of gross output. The capital share becomes correspondingly smaller, as this is defined as the residual given the other factor shares. A labor share
reduces the differences in \( \log \left( \frac{r_{i,\text{skill}}/r_{i,\text{unskill}}}{r_{US,\text{skill}}/r_{US,\text{unskill}}} \right) \), since smaller skilled labor price variations are needed to rationalize the trade data. The results are summarized in Table 11 for \( \sigma = 7.5 \) and different assumptions on the markup.

<table>
<thead>
<tr>
<th>GDP ranges</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$15,000</td>
<td>8.36</td>
<td>7.10</td>
<td>6.17</td>
</tr>
<tr>
<td>$15,000-$30,000</td>
<td>6.40</td>
<td>5.55</td>
<td>4.91</td>
</tr>
<tr>
<td>$30,000-$50,000</td>
<td>3.44</td>
<td>3.12</td>
<td>2.88</td>
</tr>
<tr>
<td>$50,000-$75,000</td>
<td>1.35</td>
<td>1.32</td>
<td>1.29</td>
</tr>
<tr>
<td>&gt;$75,000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

4 Development accounting

In this section, I use the estimates from Section 3 to decompose productivity differences in manufacturing. The aim is to find the size of uniform TFP differences in manufacturing between rich and poor countries. I focus on manufacturing as the trade estimates are based on manufacturing trade data. Thus, the only required extrapolation is from the traded sector to the overall manufacturing sector, and not from the manufacturing sector to the rest of the economy.

4.1 Chained TFP calculations

Proposition 1 provided conditions for the existence of an aggregate production with factor-augmenting shifters:

\[
Y_i = A_i G(Q_{aU}, Q_{aS}, Q_k K),
\]

where the aggregate factor shifters agree with the industry-level factor shifters.\(^{21}\) Given that the trade analysis identifies the efficiency-adjusted relative prices \( r_{f}/r_1 = \frac{w_f/Q_f}{w_1/Q_1} \), the trade-based estimates of \( r_f/r_1 \) can be used to back out factor-efficiency adjusters via the equation

\[
\frac{w_f}{w_1} = \frac{r_f Q_f}{r_1 Q_1},
\]

where \( w_f \) are the observed factor prices. By using an independent method to estimate a baseline factor efficiency \( Q_1 \) (in our case, this will be the quality of unskilled labor), it is possible to obtain estimates of \( Q_{i,1}, \ldots, Q_{i,F} \) for all countries. This gives us the efficiency-adjusted factor inputs across all countries.

\(^{21}\)Keeping with the standard for development accounting exercises, I use a value added production function and not a gross output production function.
In the development accounting literature, the standard approach to decompose productivity differences has been to make a functional form assumption on $G$. Standard assumptions have been that $G$ is a Cobb-Douglas aggregator of labor and capital with linear aggregation of labor (Hall and Jones, 1999; Caselli, 2005; Jones, 2015), or a Cobb-Douglas aggregator of capital and labor with a CES aggregator of labor (Caselli and Coleman, 2006; Jones, 2014a; Caselli, 2016).

In this current case, the problem with applying these two approaches is that the labor share is not constant in manufacturing. To address this problem, I move beyond assuming a functional form for the aggregate production function, and instead, I employ a non-parametric approach. In particular, when aggregate output per worker satisfies

$$\frac{Y_i}{L_i} = A_i G \left( \frac{Q_{i,1} x_{i,1}}{L_i}, \ldots, \frac{Q_{i,F} x_{i,F}}{L_i} \right),$$

then the second order approximation to output deviations under competitive markets is

$$\log \left( \frac{Y_i}{Y_i'} \right) = \log \left( \frac{A_i}{A_i'} \right) + \sum_{f=1}^{F} \left( \frac{\alpha_{i,f} + \alpha_{i',f}}{2} \right) \log \left( \frac{Q_{i,f} x_{i,f}}{Q_{i',f} x_{i',f}} \right),$$

where $\frac{\alpha_{i,f} + \alpha_{i',f}}{2}$ denotes the average factor share of factor $f$ in countries $i$ and $i'$ (Diewert, 1974). By re-arranging the equation, the TFP-difference between the two countries can be estimated using data on relative output per workers, relative efficiency-adjusted factor supplies, and factor shares.

I use this result to create a chained TFP calculation exercise between different income groups $g = 1, \ldots, G$. For each income group $g$, the difference in TFP compared to the next income group $g - 1$ is defined as

$$\log \left( \frac{\bar{A}_g}{\bar{A}_{g-1}} \right) = \log \left( \frac{\bar{Y}_g}{\bar{Y}_{g-1}} \right) - \sum_{f=1}^{F} \left( \frac{\bar{\alpha}_{f,g} + \bar{\alpha}_{f,g-1}}{2} \right) \log \left( \frac{\bar{Q}_{f,g} x_{f,g}}{\bar{Q}_{f,g} x_{f,g-1}} \right),$$

where all factor shares with $g$ subscripts are simple averages, and all other variables with $g$ subscripts are geometric averages.

By using this expression, it is possible to decompose the differences in TFP across different income groups without placing strong functional form assumptions on the aggregate manufacturing sector.

4.2 Data and estimates

In my accounting exercise, I use three factors of production: unskilled labor, skilled labor, and capital services. Thus, I need information about manufacturing output per worker $y_{man}$, capital per worker $k_i$, share of skilled and unskilled workers $s_i$ and $s_i'$, efficiency adjusters $Q_{u,i}$, $Q_{s,i}$, and $Q_{k,i}$, factor shares $\alpha_{u,g}$, $\alpha_{s,g}$, and $\alpha_{k,g}$. I define the income groups by GDP per worker in 2005,
using an output-based PPP from PWT 9.0.

<table>
<thead>
<tr>
<th>GDP per output ranges</th>
<th>Number of countries in the WIOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,000-15,000</td>
<td>3</td>
</tr>
<tr>
<td>15,000-30,000</td>
<td>4</td>
</tr>
<tr>
<td>30,000-50,000</td>
<td>10</td>
</tr>
<tr>
<td>50,000-75,000</td>
<td>8</td>
</tr>
<tr>
<td>$&gt;$75,000</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 12: Output per worker groups and number of countries

The definition of the income groups is displayed in Table 12. The subdivisions are chosen with reference to data availability in the WIOD, which is the source of the data on factor shares, capital intensity, and manufacturing output per worker. The subdivisions reflect a trade-off between having as large a number of countries as possible in each bin, while keeping the bins as small as possible. The small set of poor countries in the WIOD makes it potentially problematic to use the trade estimates for this set of countries, as the trade estimates are estimated with error. Thus, for the trade estimates, I use the average of all observations in each respective income group.

To measure manufacturing output per worker, I use the WIOD data to obtain total manufacturing employment and value added in local prices for 2005. I deflate the local currency manufacturing output using a manufacturing producer price PPP exchange rate. I create it by combining the PWT 9.0 2005 market exchange rate, and the production-side manufacturing price level constructed in Inklaar and Timmer (2013).

To construct the employment share of unskilled and skilled labor, I use census data and labor force survey data from IPUMS for the set of countries with 2005 GDP per worker greater than $5,000, and any available dataset after 2000 which records industry and occupation of workers. I calculate the share of manufacturing workers that have an occupation in any of the ILO categories “Professionals”, “Technicians and associate professionals”, or “Legislators, senior officials and managers”. These are the occupations which have skill level 3 or 4 according to ILO’s definition, and this definition is consistent with the skill definition in the trade data. Note that I use an occupational definition of skill. This choice is discussed in Appendix B.3.

For each country, I use the WIOD to measure the labor share in the manufacturing sector as total labor compensation divided by value added, and then I define the labor income share $\alpha_{l,g}$ for the income group as the average of country labor shares within that income group. The capital share $\alpha_{k,g}$ is defined as 1 minus the average labor share in group $g$.

To split the labor share into compensation of skilled and unskilled workers, I combine information on the relative supply of skilled and unskilled workers with data on the skilled wage premium. The method for deriving the relative supply of skilled workers was described before. The skilled premium is also measured using IPUMS data. Unfortunately, less countries record income data in the IPUMS. Thus, my sample is restricted to Brazil, Canada, Dominican Republic, India, Indonesia-
sia, Jamaica, Mexico, Panama, South Africa, and the USA. Thus, I calculate the average wage premia in each income group by regressing the log wage premia on log output per worker and using the average predicted value for each range of GDP values. Using the relative supply of skilled and unskilled workers $s_g/u_g$ and the skilled wage premia $w_{s,g}/w_{u,g}$, the relative factor share of skilled and unskilled workers labor can be defined by

$$\frac{\alpha_{s,g}}{\alpha_{u,g}} = \frac{s_g}{u_g} \frac{w_{s,g}}{w_{u,g}}.$$ 

I measure the relative efficiency of skilled versus unskilled labor $Q_s/Q_u$ by

$$\frac{w_s}{w_u} = \frac{Q_s}{Q_u} \frac{r_s}{r_u},$$

where I obtain $r_s/r_u$ from the trade data estimates. To calculate the group averages, I log both sides of the expression and take the averages over countries in each income group.

It is theoretically possible to use the trade data estimates to estimate the relative efficiency of capital and unskilled labor

$$\frac{w_k}{w_u} = \frac{Q_k}{Q_u} \frac{r_k}{r_u},$$

However, since the user cost of capital $w_k$ is not directly observed, I set $Q_k$ to be constant across countries.\textsuperscript{22}

Lastly, I define the quality of unskilled labor $Q_u$ through a standard Mincerian method where I use the average education length among unskilled workers in manufacturing, and then measure $Q_u$ by converting workers to unskilled equivalent units assuming a Mincerian return of 10%. I use the IPUMS data to obtain data on the educational level among manufacturing workers, and define the length of primary school as 6 years, the length of secondary school as 6 years, and the length of college as 3 years.

\textbf{4.3 Results}

In Table 13, I report the key summary statistics of the development accounting exercise (for $\sigma = 7.5$). The first column shows the average overall output per worker for the countries in the WIOD in each income range. The second column shows manufacturing output per worker. This shows that

\textsuperscript{22}It is possible to back out $w_k$ indirectly from the factor shares by setting $w_k$ to a level that is needed to rationalize the observed capital share. I have decided to not use this approach as it leads to a compounding of an existing measurement problem, namely that I define the capital elasticity of output as 1 minus the labor share. Even though poor countries have somewhat higher capital-output ratios than rich countries in manufacturing, their high capital share could be due to a high profit share. If profits are driving the capital share, this will, in general, lead to an overestimation of the importance of capital, which, given the high capital-output ratio in poor countries, increase estimated TFP differences. However, this potential error is compounded if we use an indirect method to impute $w_k$, as a high $\alpha_k$ leads us to estimate a high $w_k$, and consequently a high efficiency of capital. For this reason, I use a constant $Q_k$ in my analysis.
differences in manufacturing output are of approximately the same size as the differences in overall output. The third column shows the capital output ratio in manufacturing, which is somewhat higher in lower income groups. The share of skilled workers in manufacturing is increasing with income, and goes from approximately 10% of the number of workers to 30% of workers. The wage premium is higher in poor countries, but not dramatically higher – only a factor of 1.5. In contrast, the relative efficiency-adjusted price of skilled labor services is about 12 times higher. Combining these two facts imply that the relative quality of skilled and unskilled labor is about 7 times lower in poor countries compared to in rich countries. The last column shows human capital calculated by linear aggregation, i.e. by taking converting skilled workers to unskilled workers using the skilled wage premium, and then multiplying by the quality of unskilled workers.

Table 13: Summary statistics for development accounting exercise

<table>
<thead>
<tr>
<th>GDP/worker</th>
<th>(y)</th>
<th>(y_{man,f})</th>
<th>(K_m/Y_m)</th>
<th>(s_m)</th>
<th>(w_s^m/w_u^m)</th>
<th>(x_s^m/x_u^m)</th>
<th>(Q_s^m/Q_u^m)</th>
<th>(h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,000-15,000</td>
<td>8856</td>
<td>7380</td>
<td>3.04</td>
<td>0.11</td>
<td>3.54</td>
<td>0.97</td>
<td>3.65</td>
<td>2.76</td>
</tr>
<tr>
<td>15,000-30,000</td>
<td>23686</td>
<td>9403</td>
<td>3.50</td>
<td>0.15</td>
<td>2.96</td>
<td>0.68</td>
<td>4.35</td>
<td>2.93</td>
</tr>
<tr>
<td>30,000-50,000</td>
<td>41244</td>
<td>20783</td>
<td>2.46</td>
<td>0.19</td>
<td>2.62</td>
<td>0.40</td>
<td>6.55</td>
<td>3.06</td>
</tr>
<tr>
<td>50,000-75,000</td>
<td>62280</td>
<td>41263</td>
<td>2.55</td>
<td>0.25</td>
<td>2.39</td>
<td>0.12</td>
<td>19.92</td>
<td>3.71</td>
</tr>
<tr>
<td>&gt;75,000</td>
<td>82997</td>
<td>81199</td>
<td>1.39</td>
<td>0.29</td>
<td>2.08</td>
<td>0.08</td>
<td>26.00</td>
<td>3.65</td>
</tr>
</tbody>
</table>

In Table 14, I report the variables that are directly used in the accounting exercise. Note that the capital share \(\alpha_k\) is considerably higher for poor countries than for rich countries, making the assumption of a Cobb-Douglas aggregator between capital and labor inappropriate for this situation.

Table 15 displays the result of the development accounting exercise. I write \(\Delta_m = \log(y_{man,g}) - \log(y_{man,g-1})\) for the difference in manufacturing output between two income groups. I write \(\gamma_x = \frac{\alpha_x \alpha_{x-1} \log(xQ_x)}{2}\) for the Divisia index of the efficiency adjusted difference in factor input between two groups. For the linear human capital aggregator \(h\), I define \(\alpha_h = 1 - \alpha_k\) as the total labor share. The last three columns represent three different ways of measuring TFP. The first TFP measure does not adjust for labor composition at all, and only subtracts the capital contribution. The second TFP measure is the standard TFP measure from the development accounting literature, which uses a linear aggregator of labor input. The third TFP measure with imperfect substitutability between different labor services.

The second to last line sums each column, and shows the relative contribution of different factor inputs, as well as total log TFP differences. The last line is the exponent of the second to last line. Output differences are a factor of 11. When there is no correction for labor quality, a factor of 5 depends on uniform TFP differences. This is reduced to 4.4 using the traditional correction, but only to approximately 2.6 when imperfect substitutability is taken into account. Thus, the introduction of imperfect substitutability decreases the importance of uniform TFP differences in explaining manufacturing productivity differences.

32
Lastly, Table 16 shows how the results depend on the trade elasticity $\sigma$. The first column shows the calculated TFP ratios between rich and poor countries. We see that the higher $\sigma$ is, the larger are the required TFP differences. When $\sigma = 5$, there is only a factor 1.7 in TFP differences. When $\sigma = 10$, the TFP ratio is instead 2.9. The mechanism underlying this is that a higher $\sigma$ reduces the estimated difference in efficiency of skilled labor across countries. This is shown in the second column. For $\sigma = 5$, there is a 37 times difference in the efficiency of skilled workers, but for $\sigma = 10$, this is reduced to only approximately 4.

Table 14: Productivity accounting variables

<table>
<thead>
<tr>
<th>log($y_{manuf}$)</th>
<th>log($uQ_u$)</th>
<th>log($sQ_s$)</th>
<th>log($K/L$)</th>
<th>log($h$)</th>
<th>$\alpha_u$</th>
<th>$\alpha_s$</th>
<th>$\alpha_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.91</td>
<td>0.65</td>
<td>-0.12</td>
<td>10.02</td>
<td>1.02</td>
<td>0.24</td>
<td>0.11</td>
<td>0.65</td>
</tr>
<tr>
<td>9.15</td>
<td>0.64</td>
<td>0.44</td>
<td>10.40</td>
<td>1.09</td>
<td>0.33</td>
<td>0.18</td>
<td>0.49</td>
</tr>
<tr>
<td>9.94</td>
<td>0.63</td>
<td>1.08</td>
<td>10.84</td>
<td>1.12</td>
<td>0.32</td>
<td>0.20</td>
<td>0.48</td>
</tr>
<tr>
<td>10.63</td>
<td>0.72</td>
<td>2.64</td>
<td>11.56</td>
<td>1.32</td>
<td>0.35</td>
<td>0.29</td>
<td>0.37</td>
</tr>
<tr>
<td>11.30</td>
<td>0.67</td>
<td>2.97</td>
<td>11.64</td>
<td>1.31</td>
<td>0.33</td>
<td>0.30</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Table 15: Decomposition of productivity differences

<table>
<thead>
<tr>
<th>log($y_m$)</th>
<th>$\Delta y_m$</th>
<th>$\gamma_u$</th>
<th>$\gamma_s$</th>
<th>$\gamma_k$</th>
<th>$\gamma_h$</th>
<th>$(\Delta y_m - \gamma_k)$</th>
<th>$(\Delta y_m - \gamma_k - \gamma_h)$</th>
<th>$(\Delta y_m - \gamma_k - \gamma_u - \gamma_s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.91</td>
<td></td>
<td>0.24</td>
<td>-0.00</td>
<td>0.08</td>
<td>0.22</td>
<td>0.03</td>
<td>-0.00</td>
<td>-0.05</td>
</tr>
<tr>
<td>9.15</td>
<td></td>
<td>0.79</td>
<td>-0.00</td>
<td>0.12</td>
<td>0.22</td>
<td>0.02</td>
<td>0.58</td>
<td>0.56</td>
</tr>
<tr>
<td>9.94</td>
<td></td>
<td>0.69</td>
<td>0.03</td>
<td>0.38</td>
<td>0.31</td>
<td>0.11</td>
<td>0.38</td>
<td>0.27</td>
</tr>
<tr>
<td>10.63</td>
<td></td>
<td>0.68</td>
<td>-0.02</td>
<td>0.10</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.65</td>
<td>0.66</td>
</tr>
<tr>
<td>11.30</td>
<td></td>
<td>2.40</td>
<td>0.01</td>
<td>0.68</td>
<td>0.76</td>
<td>0.15</td>
<td>1.63</td>
<td>1.48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>log($y_m$)</th>
<th>$\Delta y_m$</th>
<th>$\gamma_u$</th>
<th>$\gamma_s$</th>
<th>$\gamma_k$</th>
<th>$\gamma_h$</th>
<th>$(\Delta y_m - \gamma_k)$</th>
<th>$(\Delta y_m - \gamma_k - \gamma_h)$</th>
<th>$(\Delta y_m - \gamma_k - \gamma_u - \gamma_s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.00</td>
<td></td>
<td>1.01</td>
<td>1.98</td>
<td>2.15</td>
<td>1.16</td>
<td>5.12</td>
<td>4.40</td>
<td>2.57</td>
</tr>
</tbody>
</table>

Table 16: Development accounting results for different $\sigma$

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>$Q_s^{rich}/Q_s^{poor}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1.84</td>
</tr>
<tr>
<td>7.5</td>
<td>2.57</td>
</tr>
<tr>
<td>10</td>
<td>2.99</td>
</tr>
</tbody>
</table>

4.3.1 Elasticity of substitution

The paper is motivated by the challenge of separately identifying the elasticity of substitution and skill biased efficiency differences. The exercise above has identified the skill bias. What is the elasticity of substitution between skilled and unskilled workers?
Unfortunately, it is not possible to define a single number, given the flexible functional form assumptions on the aggregate production function (a single elasticity of substitution requires separability between labor and capital and that labor is aggregated using a CES specification). However, it is still possible to define a notion of elasticity of substitution motivated by a CES specification. In particular, under a CES aggregator, relative supplies and prices are related by

\[
\frac{w_s}{w_u} = a_s \frac{Q_s}{Q_u} \left( \frac{Q_s}{Q_u} \right)^{-1/\eta} \iff \log \left( \frac{w_s}{w_u} \right) = \log(a_s) - \left( 1 - \frac{1}{\eta} \right) \log \left( \frac{Q_s}{Q_u} \right). \tag{8}
\]

which means that if \( 1 - \frac{1}{\eta} \) is defined as the regression estimate when regressing the log relative factor shares of skilled and unskilled workers on their relative effective supplies, the elasticity of substitution \( \eta \) can be estimated. The results are displayed in Table 17.

The middle estimate of 1.23 estimate is somewhat smaller than estimates in, e.g., Katz and Murphy (1992) and Ciccone and Peri (2005), but are not directly comparable. In addition to some technical issues related to not using a CES labor aggregator, I also use an occupational rather than an educational definition of skill. Since the shares of workers in skilled occupations vary less with income than the share of highly educated workers, this assumption make the estimated elasticity of substitution smaller.

Table 17: Estimated elasticities of substitution for different trade elasticities

<table>
<thead>
<tr>
<th>( \sigma = 5 )</th>
<th>( \sigma = 7.5 )</th>
<th>( \sigma = 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity of substitution</td>
<td>1.13</td>
<td>1.23</td>
</tr>
</tbody>
</table>

5 Comparison and interpretation

In this section, I compare my findings to existing results in the literature, and discuss how the existence of skill-biased labor efficiency differences should be interpreted.

5.1 Comparison to literature

Most earlier studies have used aggregate data on the relative wages and quantities of skilled workers to estimate labor efficiency differences. The logic of this estimation procedure can be explained by noting that, with a constant elasticity of substitution labor aggregator, equation (8) can be manipulated to obtain:

\[
\frac{Q_s}{Q_u} = \left( \frac{s}{u} \right)^{\frac{1}{\eta-1}} \left( \frac{w_s}{w_u} \right)^{\frac{\eta}{\eta-1}}, \tag{9}
\]

where \( \frac{Q_s}{Q_u} \) is the relative efficiency of skilled and unskilled workers, \( \frac{w_s}{w_u} \) is the skilled wage premium, and \( \frac{s}{u} \) is the relative shares of skilled and unskilled workers. A high \( \frac{s}{u} \) in rich countries implies that
is high, whereas a low $\frac{w_s}{w_u}$ in rich countries implies that $\frac{Q_s}{Q_u}$ is low. The relative importance of the two effects is regulated by the elasticity of substitution $\eta$. The dependence of the elasticity of substitution illustrates the impossibility result of Diamond et al. (1978): it is not possible to identify skill-biased efficiency differences without making assumptions on the elasticity of substitution between skilled and unskilled workers.

Equation (9) can be used to summarize previous results in the literature. If there is perfect substitutability, $\eta = \infty$, then the estimated skill bias coincides with the observed skilled wage premium. Since the wage premium only differs modestly across rich and poor countries, the same is true for $\frac{Q_s}{Q_u}$. This is the finding in Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), and Caselli (2005). In contrast, if we assume that the cross-country relevant elasticity is in line with findings from US time series and panel data, e.g., from Katz and Murphy (1992) and Ciccone and Peri (2005), then there is a much smaller elasticity of substitution $\eta$, approximately between 1 and 2. In this case, the large differences in the relative numbers of skilled workers, $\frac{s_u}{u}$, implies that efficiency differences are skill-biased. For example, if $\eta = 1.5$ is used, as in Caselli (2016), we obtain a term $(\frac{s_u}{u})^2$ on the right-hand side of equation (9), and estimated skill-bias differences grow as the square of relative supply differences. Thus, exercises using a low $\eta$ tend to find skill-biased efficiency differences (Caselli and Coleman, 2006; Jones, 2014a; Caselli, 2016; Rossi, 2017).\(^{23}\)

In finding skill-biased efficiency differences, my results, qualitatively, agrees with findings using aggregate data and a relatively low elasticity of substitution. Since my paper uses an occupational definition of skill, it is not possible to directly compare the results quantitatively. However, it is possible to gauge orders of magnitudes by redoing the analysis in Caselli (2016) for the same set of countries as that used in my exercise. I replicate the findings using the method description from the book and his online-published raw data. Table 18 summarizes the results by income group, with the richest group normalized to 1, as in Table 8. The ratio between the richest and poorest group is approximately 4, which is of the same order of magnitude as the skill bias found in my baseline specification.

Thus, the industry based and aggregate methods agree on the existence of skill-biased efficiency differences. Furthermore, they provide independent support for this existence. The industry method rejects skill-neutral efficiency differences since small variations in the skilled wage premium

\(^{23}\)An exception is Morrow and Trefler (2017), which finds limited evidence for skill-biased efficiency differences despite using a relatively low elasticity of substitution: $\eta = 1.67$. Their estimate is part of a larger trade and general equilibrium model. To identify identify the skill-bias of efficiency differences, they assume that the aggregate elasticity of substitution holds within every industry, and they use factor input and price data from the World Input Output Database. Their estimate coincides with the one in Caselli (2016) when there is one industry; in general, it also contains a term correcting for industry composition. Their finding of low skill bias not seem to be due to the multi-industry setup, since the estimates based on aggregate supply and wage data yield very similar results. Thus, it seems to be due to using information from the WIOD, implying a smaller set of countries, a different skill cut-off (college versus non-college), and a different estimated skilled wage premium. This suggests a value in performing the aggregate analysis with even more attention to data choices and definition of skill groups. Rossi (2017) represents work in this direction by using IPUMS census data on skilled labor supply and wages; he finds that efficiency differences are skill-biased.
Table 18: Skill bias estimates in Caselli (2016)

<table>
<thead>
<tr>
<th>Income per worker</th>
<th>Skill bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$15,000</td>
<td>0.28</td>
</tr>
<tr>
<td>$30,000-$30,000</td>
<td>0.39</td>
</tr>
<tr>
<td>$50,000-$50,000</td>
<td>1.51</td>
</tr>
<tr>
<td>$75,000-$75,000</td>
<td>0.90</td>
</tr>
<tr>
<td>&gt;$75,000</td>
<td>1.00</td>
</tr>
</tbody>
</table>

are inconsistent with large differences in the trade composition given conventional trade elasticity estimates. The aggregate method rejects skill-neutral efficiency differences since small variations in the skilled wage premium are inconsistent with large differences in the relative number of skilled workers, for conventional estimates of the elasticity of substitution between skilled and unskilled workers.

5.2 Interpretation of skill-biased efficiency differences

The presence of skill biased efficiency differences has implications for how we interpret cross-country income differences. With skill-neutral efficiency differences, the skilled-unskilled dimension is, in an accounting sense, orthogonal to the process of economic development, which is about uniform shifts in labor productivity. In particular, the extremely tight relationship between income levels and the number of skilled workers – Figure 8 shows the relative number of workers in skilled occupations by income level – is not strictly a necessary feature of economic development. With perfect substitutability and skill-neutral efficiency differences, there is nothing on the production side of the economy preventing countries from being skill-intensive, yet poor, in the upper left corner of the figure, or unskill-intensive, yet rich, in the lower right corner of the figure. In practice, we do not observe such cases.

In contrast, with imperfect substitutability and skill-biased efficiency differences, the process of skill intensification is more tightly connected to the process of economic development. In particular, with imperfect substitutability, the strong relationship between income levels and the number of skilled workers can be interpreted as a supply response to high skilled labor efficiency. With imperfect substitutability, a hypothetical country in the upper left corner would have very low, if not negative, skilled wage premium, whereas a hypothetical country in the lower right corner would have very high skilled wage premium. Thus, in any model with endogenous choice of occupation, the number of skilled workers would tend to increase when there are skill-biased efficiency differences, and Figure 8 captures this process.

However, the exact interpretation of skill-biased efficiency differences depends on their source. There have been multiple interpretations proposed in the literature. Caselli and Coleman (2006) and Caselli (2016) assume that skilled labor efficiency differences reflect differences in skill-augmenting
technologies across rich and poor countries. They also propose a model where the degree of skill bias endogenously could reflect the high relative supply of skilled workers in rich countries.

Jones (2014a), in contrast, interpret skill-specific efficiency differences as reflecting differences in the human capital of skilled workers across countries. Under this interpretation, human capital can play a much larger role in explaining income differences than in development accounting exercises which aggregate labor input using relative wages, since the high human capital of skilled workers in rich countries is not fully reflected in their relative wages, due to imperfect substitutability. In a companion paper, he also discusses how human capital differences can arise as a function of varying degrees of specialization of skilled workers across countries (Jones, 2014b), which is, arguably a hybrid between a technology and a human capital and a technology model: human capital differences are the proximate explanation of income differences, but they are, partly, driven by differences in the technologies and institutions for allowing specialization.

Different interpretations are difficult to disentangle, since they are isomorphic in price and quantity data. Findings in Caselli and Ciccone (2017) and Rossi (2017) suggest that large uniform differences in the human capital of skilled workers are hard to reconcile with evidence for migration data. It is less easy to use migration data to distinguish between more direct technological theories, and the view of Jones (2014b) that there are differences in specialized, complementary skills, across countries, since the set of workers to match with change upon migration. An important task for future research is to clarify the source of skilled labor efficiency differences.

6 Concluding remarks

This paper has proposed a new method to estimate the skill bias of labor efficiency differences using disaggregated industry and trade data. The basic idea of the paper is that skill-specific labor efficiency differences will lower the effective price of skilled labor services, and, consequently, the relative unit production cost in skill-intensive industries. This implies that skill bias has implications for comparative advantage, which can be detected and quantified using industry level trade data.

The analysis shows that the share of skill-intensive exports rises rapidly with income levels. Seen through the lens of a gravity model, this suggests that rich countries have low relative unit production costs in skill-intensive industries, and that the effective price of skilled labor services is low in rich countries. I show that, for normal ranges of trade elasticities, these low effective prices of skilled labor services cannot be explained by low skilled wage premia, and I infer that labor efficiency differences are skill-biased. Making an additional set of assumptions, it is possible to integrate the trade findings into a development accounting exercise, and then I find that the existence of skill-biased efficiency differences reduces the estimated size of uniform TFP differences in manufacturing from a factor of 4.3 to a factor of 2.6.

The previous standard approach to estimate skilled labor efficiency differences in the literature has relied on using aggregated data on the wages and quantity of labor in different skill categories.
Figure 8: Share of workers in skilled occupations
Identification has relied on making an assumption about the aggregate substitution elasticity between skilled and unskilled workers, and the resulting skill bias has reflected the assumptions on the elasticity of substitution. The higher the assumed elasticity of substitution, the lower the estimated degree of skill bias. The results in my paper are most similar to those in papers that assume a relatively low elasticity of substitution, which have also found skill bias in labor efficiency differences. My results complement these results by providing new evidence for skill bias using an independent strategy of identification.

The existence of skill-biased efficiency differences suggest that the role of skilled labor is important for the process of economic development. The exact interpretation of the results depends on the source of skilled labor efficiency differences. One view is that they are due to technological differences, or due to endogenous technological responses to different supplies of skilled workers (Caselli and Coleman, 2006; Caselli, 2016; Rossi, 2017). Another hypothesis is that they reflect differences in human capital among skilled workers (Jones, 2014a), potentially arising from differences in specialization among skilled workers across different countries (Jones, 2014b). Migration data suggest that uniform efficiency differences among skilled workers are unlikely to fully explain the differences in skilled labor efficiencies (Caselli and Ciccone, 2017; Rossi, 2017), whereas the importance of skill differences in the form of specialization into complementary specialities is an open question. Thus, an important task for future research is to clarify the exact source of skilled labor efficiency differences.

Acknowledgments

IPUMS International was used for this research project. The author wishes to acknowledge the statistical offices that provided the underlying data making this research possible: National Institute of Statistics and Censuses, Argentina; National Institute of Statistics, Bolivia; Institute of Geography and Statistics, Brazil; Statistics Canada, Canada; National Institute of Statistics, Chile; National Administrative Department of Statistics, Colombia; National Institute of Statistics and Censuses, Costa Rica; Office of National Statistics, Cuba; National Statistics Office, Dominican Republic; National Institute of Statistics and Censuses, Ecuador; Department of Statistics and Censuses, El Salvador; Institute of Statistics and Informatics, Haiti; Statistical Institute, Jamaica; National Institute of Statistics, Geography, and Informatics, Mexico; National Institute of Information Development, Nicaragua; Census and Statistics Directorate, Panama; General Directorate of Statistics, Surveys, and Censuses, Paraguay; National Institute of Statistics and Informatics, Peru; U.S. Bureau of the Census, Puerto Rico; Central Statistical Office, Trinidad and Tobago; Bureau of the Census, United States; National Institute of Statistics, Uruguay; and National Institute of Statistics, Venezuela.
References


Romalis, J. (2007). NAFTA’s and CUSFTA’s Impact on International Trade. [http://dx.doi.org/10.1162/rest.89.3.416](http://dx.doi.org/10.1162/rest.89.3.416).


A Appendix to Section 2

A.1 Proof of Theorem 1

I prove the theorem by construction. Define a function

\[ G(\tilde{x}_1, \ldots, \tilde{x}_f) = \max_{\{\tilde{x}_f\}} H(F^1(\tilde{x}_1, \ldots, \tilde{x}_f), \ldots, F^K(\tilde{x}_1, \ldots, \tilde{x}_f)) \]

s.t. \[ \sum_{k=1}^{K} \tilde{x}_k^f \leq \tilde{x}_f \forall f. \]

Such a function exists given the assumptions made on the production function.

Furthermore, define

\[ r_{i,f} = \frac{w_{i,f}}{Q_{i,f}} \]

where \( w_{i,f} \) is the price of factor \( f \) in country \( i \).

If we write \( \chi_i(\{x_{i,f}\}) = \{x_{i,f}^k : \sum_{k=1}^{K} x_{i,f}^k \leq x_{i,f}\} \) for the set of feasible factor allocation for economy \( i \), we can see that the aggregate production function in country \( i \) is given by

\[ G_i(x_{i,1}, \ldots, x_{i,f}) = \max_{\{x_{i,f}^k \in \chi_i(\{x_{i,f}\})\}} H(F^1_i(x_{i,1}, \ldots, x_{i,f}, \ldots, F^K_i(x_{i,1}, \ldots, x_{i,f})) \]

\[ = \max_{\{x_{i,f}^k \in \chi_i(\{x_{i,f}\})\}} H(\tilde{A}_i F^1_i(Q_{i,1} x_{i,1}, \ldots, Q_{i,f} x_{i,f}), \ldots, \tilde{A}_i F_i(Q_{i,1} x_{i,1}, \ldots, Q_{i,f} x_{i,f})) \]

\[ = \tilde{A}_i G(Q_{i,1} x_{i,1}, \ldots, Q_{i,f} x_{i,f}). \]

Thus, the aggregate production function \( G_i \) can be written on a factor augmenting form (the assumptions on \( H \) ensures that the optimization problems have unique solutions).

A.2 Proof of Proposition 1

Under the assumptions on the unit cost function and competitive factor markets, the log relative unit cost between country \( i \) and the US can be written as

\[ \log \left( \frac{c_i^k}{c_{US}^k} \right) = -\log \left( \frac{A_i \tilde{A}_i^k}{A_{US} \tilde{A}_{US}^k} \right) + \sum_{f=1}^{F} \left( \frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2} \right) \times \log \left( \frac{r_{i,f}}{r_{US,f}} \right) \]

\[ = -\log \left( \frac{A_i}{A_{US}} \right) + \log \left( \frac{r_{i,1}}{r_{US,1}} \right) + \sum_{f=2}^{F} \left( \frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2} \right) \times \log \left( \frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}} \right) - \log \left( \frac{\tilde{A}_i}{\tilde{A}_{US}} \right), \]

where the second line uses that factor shares sum to 1. Since \( \log \left( \frac{\tilde{A}_i}{\tilde{A}_{US}} \right) \) is independent of \( \frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2} \), then \( \log \left( \frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}} \right) \) can be identified by regressing \( \log (c_i^k) \) on a country fixed effect, an industry
fixed effect, and \( \frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2} \) (if the parameter on US factor shares \( \alpha_{US,f}^k \) is constrained to be zero).

A.3 Proof of Proposition 2

Express the linear decomposition of \( \log(\tilde{A}_i^k / A_{US}^k) \) as

\[
\log(\frac{\tilde{A}_i^k}{A_{US}^k}) = \sum_{f=2}^{F} \gamma_{i,f} \left( \frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2} \right) + \varepsilon_{i,f}^k,
\]

where \( \varepsilon_{i,f}^k \) is uncorrelated to \( \frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2} \) for every country \( i \). By substituting in this expression into the regression specification, the estimated efficiency-adjusted relative factor price becomes

\[
\log(\frac{r_{i,f}}{r_{i,1}} / \frac{r_{US,f}}{r_{US,1}}) + \gamma_{i,f}.
\]

B Estimating the relative price of skilled services

B.1 Theoretical derivation of gravity equation

In this section, I show how my gravity specification can be derived from theoretical trade models. I first derive the specification from an Armington style trade model, and then from an Eaton and Kortum style trade model.

B.1.1 Armington model

There are \( K \) industries and \( I \) countries, indexed \( i \) for source countries and \( j \) for destination countries. Each country admits a representative household with preferences

\[
U_j = \left( \sum_{i=1}^{I} \sum_{k=1}^{K} (a_j^k)^{1/\sigma} (q_{ij}^k)^{\sigma-1} \sigma \right)^{-1/\sigma} \quad j = 1, \ldots, I; \ \sigma > 1
\]

where \( q_{ij}^k \) are goods from industry \( k \) produced in country \( i \) and consumed in country \( j \), \( \sigma \) captures the elasticity of substitution between different varieties, and \( a_j^k \) is a country-specific taste term. The taste term is a reduced form way of capturing differences in tastes across countries, including potential non-homotheticities in preferences. The representative consumer maximizes (10) subject to a constraint

\[
\sum_{i=1}^{I} \sum_{k=1}^{K} P_{i,j}^k q_{i,j}^k \leq Y_j
\]

where \( P_{i,j}^k \) is the price of good \( k \) produced in country \( i \) and bought in country \( j \). \( Y_j \) is income in country \( j \).
Each variety is produced using a constant returns to scale production function with the unit cost function
\[ c^k_i = C^k(r_{i,1}, \ldots, r_{i,F}) \] (11)
where \( r_{i,f} \) is the price of factor service \( f \) in country \( i \).

Trade costs take an iceberg form and to consume one unit of a good from country \( i \), a country \( j \) consumer has to buy \( d_{i,j} \geq 1 \) goods from country \( i \). The cost term \( d_{i,j} \) satisfies
\[
\begin{align*}
d_{i,j} & \geq 1 \\
d_{i,i} & = 1 \quad \forall i = 1, \ldots, I \\
d_{i,j}d_{j,l} & \geq d_{i,l}.
\end{align*}
\]

Output markets are competitive, which implies that prices satisfy
\[ P^k_{i,j} = c^k_i d_{i,j}. \] (12)

Each country has a supply of factor service flows
\[ e_{j,f} \geq 0 \quad i = 1, \ldots, I; \ f = 1, \ldots, F, \]
and country income is given by
\[ Y_j = \sum_{f=1}^{F} r_{j,f} e_{j,f} \] (13)

An equilibrium is a set of consumption quantities \( q^k_{i,j} \), production quantities \( Q^k_i \), factor service prices \( r_{i,f} \), unit costs \( c^k_i \), output prices \( P^k_{i,j} \), and incomes \( Y_j \) such that:

1. \( \{q^k_{i,j}\} \) solves the consumer problem given output prices and incomes.
2. Output market clears
\[ Q^k_i = \sum_{j=1}^{I} q^k_{i,j}d_{i,j} \forall i, k \]
3. \( c^k_i \) and \( P^k_{i,j} \) satisfy (11) and (12) respectively
4. Income is given by (13)
5. Factor markets clear
\[ e_{i,f} = \sum_k Q^k_i \frac{\partial c^k_i}{\partial r_{i,f}} \]
I will not solve the complete equilibrium, but will only solve for the regression specification relating industry export values to unit costs. In the data, export values between $i$ and $j$ in industry $k$ are presented excluding trade costs (FOB). This corresponds to $P_{i,i}^k q_{i,j}^k$, i.e. the domestic price in $i$ of good $k$ produced in $i$. Using the competitive output market assumption, this quantity is $c_i^k d_{i,j}^k$.

Consumer optimization implies that for any country-industry pairs $(i, k)$, $(i', k')$

$$
\left( a_j^k \right)^{1/\sigma} \left( d_{i,j}^k \right)^{-1/\sigma} = \frac{P_{i,j}^k}{P_{i',j}^{k'}} \\
\sum_{k=1}^{K} \sum_{i=1}^{I} q_{i,j}^k P_{i,j}^k = Y_j
$$

Re-arranging the terms gives us

$$
P_{i,i}^k q_{i,j}^k = Y_j \frac{a_j^k (P_{i,j}^k)^{1-\sigma}}{\sum_{j',k'} a_{j'}^{k'} (P_{i,j'}^{k'})^{1-\sigma} P_{i,j'}^{k'}}
$$

Taking logarithms, writing total exports $x_{i,j}^k = P_{i,i}^k q_{i,j}^k$, and substituting in (11) for prices gives me

$$
\log(x_{i,j}^k) = \delta_{i,j} + \mu_{j}^k - \left( \sigma - 1 \right) \log(c_i^k)
$$

where

$$
\delta_{i,j} = \log(Y_j) - \log \left( \sum_{j',k'} a_{j'}^{k'} (c_{i,j'}^{k'} d_{i,j'}^k)^{1-\sigma} \right) - \log(d_{i,j})
$$

$$
\mu_{j}^k = \log(a_{j}^k).
$$

Here, $\delta_{i,j}$ captures all terms that only depend on the bilateral relationship: the income of the buying country, the market access term of the buying country, and all bilateral trading costs between the two countries. $\mu_{j}^k$ captures industry-specific demand effects in the buying country.

### B.1.2 Eaton and Kortum model

To derive an industry based gravity equation using an Eaton and Kortum framework, I construct a model close to Chor (2010), who analyzed industry-level trade in an Eaton and Kortum setup. There are $I$ countries where $i$ is an index for a source country and $j$ is an index for a destination country. The model has $K$ goods which are produced domestically, and the production of each good $k$ uses a range of internationally traded intermediate good varieties.
Each country has a representative consumer with preferences

\[ U_j = \left( \sum_{k=1}^{K} a_j^k (Q_j^k)^{\xi - 1} \right)^{\frac{\xi}{\xi - 1}} \quad \xi > 1. \]

Each final good \( k \) is a composite of internationally traded varieties \( q_i^k(z) \) with \( m \in [0,1] \). The price of final good \( k \) in country \( i \) is

\[ P_j^k = \left( \int_0^1 p_j^k(m)^{1 - \eta} dm \right)^{\frac{1}{1 - \eta}}, \quad \eta > \xi > 1, \]

where \( p_j^k(m) \) is the country \( j \) price of variety \( m \) in industry \( k \). The assumption on the elasticity of substitution means that different varieties are more substitutable than goods from different industries.

As varieties are internationally traded, the price \( p_j^k(m) \) paid for a variety will reflect the cheapest available variety for country \( j \). When I specify the cost function for varieties, I am therefore interested in the unit cost of offered varieties from country \( i \) to country \( j \), which I write \( p_{i,j}^k(m) \). The price \( p_j^k(m) \) is obtained by minimizing over potential source countries \( i \).

The offered price \( p_{i,j}^k(m) \) will depend on a deterministic component of costs in country \( i \) and industry \( k \), on trade costs between country \( i \) and \( j \), and on a stochastic productivity shock to this particular variety. The deterministic component of costs is

\[ c_i^k = C^k(r_{i,1}, \ldots, r_{i,F}) \]  

where \( r_{i,f} \) denotes the factor service price of factor \( f \) in country \( i \). Trade costs take an iceberg form and to obtain one unit of an intermediate good from country \( i \), a country \( j \) producer has to buy \( d_{i,j} \geq 1 \) intermediate goods from country \( i \). The cost term \( d_{i,j} \) satisfies

\[
\begin{align*}
    d_{i,j} &\geq 1 \\
    d_{i,i} &= 1 \quad \forall i = 1, \ldots, I \\
    d_{i,j}d_{j,l} &\geq d_{i,l}.
\end{align*}
\]

The offered price is

\[ p_{i,j}^k(m) = \frac{c_i^k d_{i,j}}{z_i^k(m)} \]  

where \( z_i^k(m) \sim Frechét(\theta) \) is a country-industry-variety specific productivity shock which is Frechét distributed with a parameter \( \theta \). A random variable \( Z \) is Frechét-distributed with parameter \( \theta \) if

\[ P(Z \leq z) = e^{-z^{-\theta}}. \]
I will not solve a full equilibrium for this model, but only derive the gravity trade equation that results from the model. For each variety $m$ in industry $k$, country $j$ obtains an offer $p_{i,j}^k(m)$ from each country $i$ given by equation (16). The probability distribution of this offer is

$$P(p_{i,j}^k(m) \leq p) = P\left(\frac{c_i^kd_{i,j}}{p} \leq z^k_i(m)\right) = 1 - e^{-\left(\frac{c_i^kd_{i,j}}{p}\right)^{-\theta}} = 1 - e^{-\left(c_i^kd_{i,j}\right)^{-\theta}p^\theta}$$

The best price $p_i^k(m)$ for country $i$ is the minimum of all offers $\min_i p_{i,j}^k(m)$ and has distribution

$$G(p) = P\left(\min_i p_{i,j}^k(m) \leq p\right) = 1 - P(\max_i p_{i,j}^k(m) > p) = 1 - \prod_i P(p_{i,j}^k(m) > p) = 1 - \prod_i (1 - P(p_{i,j}^k(m) \leq p)) = 1 - e^{-\sum_i (c_i^kd_{i,j})^{-\theta}p^\theta}$$

I write

$$\Phi_j^k = \sum_i \left(c_i^kd_{i,j}\right)^{-\theta}. \tag{17}$$

This expression summarizes country $j$’s access to industry $k$. It is decreasing in production costs in industry $k$ and in the bilateral trading costs $d_{i,j}$.

Country $j$ chooses to buy a variety from the country with the lowest price. The probability that country $i$ offers the lowest price is

$$\pi_{i,j}^k = P(p_{i,j}^k(z) \leq \min_i p_{i,j}^k(z)) = \frac{(c_i^kd_{i,j})^{-\theta}}{\Phi_j^k}.$$

If $x_j^k$ is the total amount of intermediate inputs bought by country $j$ in industry $k$, the trade flow matrix is

$$x_{i,j}^k = \pi_{i,j}^k x_j^k = \frac{(c_i^kd_{i,j})^{-\theta}}{\Phi_j^k} x_j^k. \tag{18}$$

Equation (18) requires that the share of import value coming from country $i$ only depends on the share of inputs for which $i$ is the supplier. This property holds as the Frechet distribution has a
desirable property called max-stability, which ensures that the best offered price \( p_{i,k}(z) \) to country \( i \) is independent of the source of the best offer (see Eaton and Kortum (2002) for a derivation in this particular case, and Mattsson et al. (2014) for a more general discussion of this property of random variables). This means that the total expenditure on imports from one country will be fully determined by the share of varieties \( \pi_{k}^{n,i} \) bought from that country. The reason is that all countries offer identical distributions of variety prices conditioned on them offering the best prices.

Taking the logarithm of both sides of equation (18) gives me

\[
\log(x_{i,j}^{k}) = \delta_{i,j} + \mu_{j}^{k} - \theta \log(c_{i}^{k})
\]

where \( \delta_{i,j} = -\theta \log(d_{i,j}) \) and \( \mu_{j}^{k} = \log(X_{j}^{k}) - \log(\Phi_{j}^{k}) \). Thus, the model implies a gravity equation of the right form. Note that when using Eaton and Kortum elasticity estimates \( \theta \), there needs to be added a 1 to convert them to the corresponding Armington elasticity estimates \( \sigma \).

### B.2 Treatment of intermediate inputs

In my main specification, I include the cost share of traded intermediate inputs \( \alpha_{US,\text{int}}^{k} \). The corresponding estimate \( \beta_{i,\text{int}} \) identifies \( \log \left( \frac{r_{i,\text{int}}}{r_{US,\text{int}}} \right) \). This estimate gives the difference between the US and country \( i \) in the relative cost of intermediate input and unskilled labor services.

In my interpretation of this parameter, I assume that intermediate inputs are traded. I interpret \( r_{i,\text{int}} \) as a product of an international price of intermediate inputs \( r_{\text{int}} \), which is constant across countries, and a country-specific barrier to international intermediate input markets \( \tau_{i} \), which varies across countries.

With this interpretation,

\[
\beta_{i,\text{int}} = \log(\tau_{i}/\tau_{US}) - \log \left( \frac{r_{i,1}}{r_{US,1}} \right).
\]

\( \beta_{i,\text{int}} \) varies across countries for two reasons. First, countries differ in their access to international intermediate goods markets \( \tau_{i} \). Bad access to international markets (high \( \tau_{i} \)) gives a high revealed price of intermediate input services (high \( \beta_{i,\text{int}} \)). Second, countries differ in their prices of unskilled labor services \( \log \left( \frac{r_{i,1}}{r_{US,1}} \right) \). Countries with a low price of unskilled services have a high revealed price of intermediate input services. This has an intuitive interpretation: relatively inexpensive unskilled labor services make internationally traded intermediate inputs relatively expensive.

To implement this approach, we need to separate the traded from the non-traded component of intermediate inputs. The intuition is that the intermediate input share in an industry \( k \) should be resolved into contributions from different factor services, using the input-output structure to determine the factor shares of industry \( k \)’s intermediate inputs.
To calculate the share of traded intermediate inputs, I assume that non-services are traded.\textsuperscript{24} I use the BEA information in the IO table to obtain information about capital, labor, and intermediate input shares in different industries, and the OES survey data to decompose the labor share into payments to skilled and unskilled workers.

I write $N_T$ for the number of traded goods and $N_{NT}$ for the number of non-traded goods. The input-output table $L$ is an $(N_T + N_{NT}) \times (N_T + N_{NT})$ matrix. For each good $k = 1, \ldots, N_T + N_{NT}$, I measure its factor shares including its intermediate input share, and I use these measured factor shares to define the first-stage factor shares $\tilde{\alpha}_f^k$. This is the same as normal factor shares with one difference. For intermediate inputs, we define $\tilde{\alpha}_f^k$ as the share of inputs that come from non-tradeable intermediates. In the first stage, I am interested in the cost shares of different factors and of tradable inputs. For each industry, $1 - \sum_{f=1}^{F} \tilde{\alpha}_f^k$ gives the share of costs in industry $k$ going to nontraded factor inputs. These first-stage factor shares are the building blocks of the factor shares $\alpha_f^k$ that will be obtained by resolving the cost share of nontraded intermediate inputs into conventional factors and tradable inputs.

I find the factor shares $\alpha_f^k$ of tradable goods recursively by first finding the factor shares of nontradable goods. I define two matrices $L_T$ and $L_{NT}$ where $L_T$ is an $N_T \times N_{NT}$ matrix giving the input uses of nontraded intermediate inputs in the traded sector, and $L_{NT}$ is an $N_{NT} \times N_{NT}$ matrix giving the cost shares from nontraded inputs in the nontraded sector.

I solve the system recursively. The factor shares of nontraded goods are

$$\alpha_{NT} = \tilde{\alpha}_{NT} + (L_{NT}) \alpha_{NT} \iff \alpha_{NT} = (I - L_{NT})^{-1} \tilde{\alpha}_{NT}$$

where $\alpha_{NT}$ is an $N_{NT} \times F$ matrix, $\tilde{\alpha}_{NT}$ is an $N_{NT} \times F$ matrix, and $L_{NT}$ is an $N_{NT} \times N_{NT}$ matrix. The final matrix $\alpha_{NT}$ gives the factor shares of nontraded services in terms of standard factor shares and traded input shares. All nontraded input shares have been resolved into these constituent parts. Having solved for the factor shares of nontraded goods, the factor shares of traded goods are

$$\alpha_T = \tilde{\alpha}_T + (L_T) \alpha_{NT}.$$  

Using this modified definition of factor shares, I can re-estimate my baseline specification. In Figure 9, I compare the estimates for the estimated skilled service coefficient to my baseline estimation. The new results are very similar to my baseline estimates. The reason is that even though resolving the nontraded factors increases the skilled share in all industries (as I move the skilled component of inputs from the intermediate input share to the skill share), the resolving of nontraded factors does little to alter the relative skill shares across industries, which are the bases of my estimation.

\textsuperscript{24}There is moderate trade in some services such as entertainment, financial services, and transportation, but the distinction captures the large differences in traded shares between services and goods in the US input-output table.
B.3 Occupational vs schooling based skill cutoff

I define the share of unskilled and skilled workers $u$ and $s$ as the shares of people working in an unskilled and skilled occupation, respectively. This contrasts to the approach taken in Caselli and Coleman (2006), B Jones (2014a), and Caselli (2016) who define the share of skilled workers as the share of individuals having an educational attainment above a pre-specified threshold (for example, primary education and above, high school and above, or college and above).

I choose an occupational rather than an educationally based definition since my main analysis concerns industry production functions rather than an aggregate production function. When modeling aggregate production opportunities, it is appropriate to define the function over variables connected to the properties of the workers – e.g., their education level – and assume that the occupational distribution is an equilibrium object determined under the hood of the aggregate production function. In contrast, when modeling an industry production function, it is, often, more appropriately defined in terms of needing a particular set of tasks, which means that an occupational definition is more appropriate.

Thus, I measure the share of skilled workers in line with the ILO’s ISCO-08 definitions of skill requirements and major occupational groups. The ILO defines 10 major occupational groups and four skill levels. The occupational groups and their respective skill levels are presented in Figure 10. I use the ILOSTAT database to obtain $s$ as the share of the labor force working as managers, professionals, or technicians and associated technicians, i.e. skill categories 3 and 4 (I define the armed forces as primarily unskilled). I define the unskilled share as $u \equiv 1 - s$. 

Figure 9: Comparison of estimated relative skilled service prices with different input measurements
Figure 10: Mapping of ISCO-08 major groups to skill levels

<table>
<thead>
<tr>
<th>ISCO-08 major groups</th>
<th>Skill level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Managers</td>
<td>3 + 4</td>
</tr>
<tr>
<td>2 Professionals</td>
<td>4</td>
</tr>
<tr>
<td>3 Technicians and Associate Professionals</td>
<td>3</td>
</tr>
<tr>
<td>4 Clerical Support Workers</td>
<td>2</td>
</tr>
<tr>
<td>5 Services and Sales Workers</td>
<td></td>
</tr>
<tr>
<td>6 Skilled Agricultural, Forestry and Fishery Workers</td>
<td></td>
</tr>
<tr>
<td>7 Craft and Related Trades Workers</td>
<td></td>
</tr>
<tr>
<td>8 Plant and Machine Operators, and Assemblers</td>
<td></td>
</tr>
<tr>
<td>9 Elementary Occupations</td>
<td>1</td>
</tr>
<tr>
<td>0 Armed Forces Occupations</td>
<td>1 + 2 + 4</td>
</tr>
</tbody>
</table>

Figure 11 compares the results from an education based and occupation based definition of the skill share. Figure 11 shows that for poor countries, the share of high school educated workers and the share of skilled workers approximately coincide. For rich countries, there are much more high school educated workers than skilled workers. This is evidence that the mapping between educational attainment and skill level is different in rich and poor countries, and that the educational cutoff for being in a skilled occupation is lower in poor countries.
Figure 11: High school and above and share of skilled occupations

C Robustness

C.1 Heterogeneous trade elasticities

This section explains how the specification with heterogeneous trade elasticities in Section 3.4 is derived.

The starting point is to modify the Eaton and Kortum-based model in Appendix B.1. I posit that the unit production cost for a variety in industry \( k \) in country \( i \) is given by

\[
\tilde{c}^k_i(\omega) = c^k_i \tilde{A}_{i}^{\theta_k} A_i z^k_i(\omega),
\]

where \( A_i \) is a TFP term, and \( z^k_i(\omega) \sim \text{Frechet}(\theta^k) \). The first difference from the main model from Appendix B.1 is that the dispersion parameter \( \theta^k \) is allowed to vary across industries. Since the dispersion parameter regulates the trade elasticity of an industry in an Eaton and Kortum framework, this implies that trade elasticities become heterogeneous across industries. The second difference is the term \( \tilde{A}^{\theta_k}_i \). This term is added to allow for an absolute advantage term that does not change relative exports across industries.\(^{25}\)

\(^{25}\)This term needs to be added since the model otherwise gets problematic implications. For example, if Catalunya and Spain split, an estimated Eaton and Kortum model would predict that their TFP terms would have fallen, since they now both have smaller share in all other countries' export bundles. This can be rationalized by thinking of the
The derivation is very similar to the model in Appendix B.1, and the key equation is
\[ X_{i,j}^k = X_j^k \pi_{i,j}^k, \]
where \( X_j^k \) is total imports of country \( j \) in industry \( k \), and
\[ \pi_{i,j}^k = \mathbb{P}(\tau_{i,j}^k \bar{c}_i^k(\omega) \leq \min_{i'} \tau_{i',j}^k \bar{c}_{i'}^k(\omega)) \]
is the probability that a variety produced in country \( i \) is the cheapest available for country \( j \), once trade costs are taken into account.

Resolving \( \pi_{i,j}^k \) using our distributional assumptions yields
\[ X_{i,j}^k = X_j^k \bar{A}_i \sum_{i'} \bar{A}_{i'} A_{i'}^{\theta_k} (c_{i',j}^k)^{-\theta_k}. \]

We then use that
\[ \log(c_i^k) = \log(c_{US}^k) + \log \left( \frac{r_{i,1}}{r_{US,1}} \right) + \sum_{f=2}^F \left( \frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2} \right) \log \left( \frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}} \right), \]
and that
\[ \log(\tau_{i,j}^k) = \sum_{m=1}^M \gamma^k d_{i,j}^m, \]
where \( d_{i,j}^m \) is a gravity variable of type \( m \), e.g. distance or a dummy for a regional trade agreement between country \( i \) and \( j \).

Then, it is possible to write
\[ \log(X_{i,j}^k) = \log(\bar{A}_i) + \theta_k \delta_i + \mu_j^k + \sum_{m=1}^M \gamma^k d_{i,j}^m - \sum_{f=2}^F \theta_k \left( \frac{\alpha_{i,f}^k + \alpha_{US,f}^k}{2} \right) \beta_{i,f} \]
technology term as depends on country size since size determines the number of draws. However, if this effect is uniform across industries, the model would then predict that trade should decrease in all industries with high trade elasticities. It is problematic that the model’s comparative advantage predictions are not invariant under country subdivision, and this is an artifact of all trade effects in Eaton and Kortum style models going through unit costs. Including a term \( \bar{A}_i^{\theta_k} \) is a simple way to address this problem, since it allows for a size effect in trade that is not heterogeneous across sectors.

For computational reasons, I do not include a bilateral fixed effect in the trade cost specification, as this would necessitate estimating an interaction between \( \delta_{i,j} \) and the trade elasticity \( \theta_k \).
where

\[ \tilde{\delta}_i = \log(A_i) - \log \left( \frac{r_{i,1}}{r_{US,1}} \right) \]

\[ \mu^k_j = \log(X^k_j) - \log \left( \sum_{i'} A^{\theta_k}_{i'} \left( c^{k}_{i'j} \right)^{-\theta_k} \right) - \theta_k \log(c^k_{US}) \]

\[ \tilde{\gamma}^k = \theta_k \gamma^k \]

\[ \beta_{i,f} = \log \left( \frac{r_{i,f}/r_{i,1}}{r_{US,f}/r_{US,1}} \right) \].

Thus, apart from removing the bilateral fixed effect, we should modify the main specification in equation 6 by adding a source country fixed effect, a source country fixed effect interacted with industry trade elasticities, and by premultiplying the factor shares with the industry trade elasticities.

To check that removing the bilateral fixed effect does

D Appendix: Concordance construction

To generate concordances and map data across coding systems, I create a general mathematical framework to treat the problem. Here, describe how the general system works, and then I show how I use it to convert our particular data.

The basic building block of our 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 19.

In Table 19, 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 code 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 being two sets A,B with measures \( w_A, w_B \) giving the mass on each code, and a concordance being a correspondence

\[ \phi: A \Rightarrow B. \]

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).

D.1 Transform quantity variables using two-weighted concordances

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 $\Phi$ as well, together with the weights $\mu_A$ and $\mu_B$. If

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

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

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

E Transform property variables using two-weighted concordances

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 $\phi$ by the equation

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

### E.1 Construct a two-side weighted concordance from a one-sided weighted concordance

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 $\phi$ and a measure $\mu_A$ on $A$ and want to create a measure $\mu_B$ on $B$. Then I define the measure on $B$ as.

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

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