

# Skill Dispersion and Trade Flows

Matilde Bombardini\* Giovanni Gallipoli† and Germán Pupato‡

First Draft: June 2009

This Draft: October 2009

## Abstract

Is skill dispersion a source of comparative advantage? While it is established that a country's aggregate endowment of human capital is an important determinant of comparative advantage, this paper investigates whether the *distribution* of skills in the labor force can play a role in the determination of trade flows. We develop a multi-country, multi-sector model of trade in which comparative advantage derives from (i) differences across sectors in the complementarity of workers' skills, (ii) the dispersion of skills in the working population. We first illustrate how higher dispersion in human capital can trigger specialization in sectors characterized by higher substitutability among workers' skills. Then we use industry-level bilateral trade data to show that human capital dispersion, as measured by a standard international metric, has a significant effect on trade flows. We find that the effect is of a magnitude comparable to that of aggregate endowments, as measured in the literature. The result is robust to the introduction of controls for several causes of comparative advantage, as well as to alternative measures of skill complementarity.

JEL Classification codes: F12, F16, J82.

We would like to thank Paul Beaudry, David Green, Patrick Francois, Thomas Lemieux, Vadim Marmer, Francesco Trebbi, Jonathan Vogel and seminar participants at CIFAR, University of Alberta, UC Davis, UC San Diego, University of Chicago Booth for helpful comments.

---

\*University of British Columbia, CIFAR, NBER and RCEA.

†University of British Columbia and RCEA.

‡University of British Columbia

# 1 Introduction

One of the mainstays of the theory of comparative advantage is that countries' factor endowments determine the pattern of trade. An established theoretical framework, the Heckscher-Ohlin-Samuelson factor proportion theory, and numerous related empirical studies<sup>1</sup> identify quantities such as the stocks of human and physical capital of countries as primary sources of comparative advantage. In this paper we provide evidence supporting an alternative, and empirically sizeable, source of comparative advantage: the dispersion of human capital in the working population.<sup>2</sup>

Why would the distribution of human capital matter for specialization and trade? We argue that sectors vary in the degree of complementarity among the skills of workers employed. We conjecture that for some sectors, for example engine and turbine manufacturing, it is essential to employ workers of similar skills at every stage of production,<sup>3</sup> while for other sectors, like apparel, the output of a team is sensitive to the presence of extremely skilled individuals, even if some stages of production are left to workers with lower human capital. Given that sectoral technologies may vary in this dimension, we investigate whether countries with greater skill dispersion specialize in sectors characterized by higher substitutability among workers' skills.

The hypothesis that skill dispersion may lead to specialization has been the object of work by Grossman and Maggi (2000), henceforth GM, who show that, in a two-country, two-sector model, the country with a relatively more dispersed skill distribution specialize in the sector that benefits from matching workers of different skill levels. This paper builds upon GM's insight by proposing a multi-country, multi-sector model where skill dispersion generates testable implications for the

---

<sup>1</sup>Among others, Romalis (2004), testing the predictions of the theory about commodity trade, and Bowen et al. (1987), Treffer (1993), Treffer (1995), and Davis and Weinstein (2001), testing the factor content predictions of the theory.

<sup>2</sup>Human capital is determined by many factors, among which formal education, family upbringing, underlying ability and on-the-job training. Throughout this paper we refer to human capital or skills, terms that we use interchangeably, as a set of attributes that are of productive use in the workplace.

<sup>3</sup>Using the terminology of Kremer (1993), these sectors exhibit an O-ring technology.

pattern of international trade. We present empirical evidence that diversity is in fact a strong determinant of specialization and that the dispersion of human capital, as well as its aggregate stock, matters in determining comparative advantage, a novel finding to the best of our knowledge.

A first glance at the data reveals that cross-country differences in skill dispersion are larger than differences in the average skills of workers. We employ the distribution of scores in the International Adult Literacy Survey (IALS), an internationally comparable measure of work-related skills, as a proxy for the distribution of human capital. The coefficient of variation of the standard deviation of scores is 1.64 times larger than that of the average scores. Figure 1 reports the mean and standard deviation of IALS log-scores during 1994-1998.

The reasons why countries at similar stages of development differ in their skill distribution are beyond the scope of this study;<sup>4</sup> such differences may be due to the degree of centralization in the education system and curricular control (Stevenson and Baker, 1991), the existence of elite schools, sorting and segregation,<sup>5</sup> early tracking,<sup>6</sup> local school financing (Benabou, 1996) and the share of private and public schools (Takii and Tanaka, 2009).<sup>7</sup>

Using a monopolistic competition model of international trade this paper investigates the relevance of skill dispersion for comparative advantage in a multi-sector multi-country environment. The theoretical framework is related to, but distinct from, the one presented in GM. They show that if skills are perfectly observable and the production function is symmetric, CRS and super-modular in the skills of workers,<sup>8</sup> then diversity is not a source of comparative advantage and there

---

<sup>4</sup>What is not beyond the scope of this study is a discussion of how the endogeneity of skill dispersion might affect our empirical results. See Section 3.4.

<sup>5</sup>The existence of peer effects, as documented for example by Hanushek et al. (2003) and Hoxby and Building (2000), implies that segregation and sorting might result in even higher inequality of educational outcomes. An example of this amplification mechanism is provided by Friesen and Krauth (2007).

<sup>6</sup>Tracking refers to the practice of grouping students in different schools according to their ability. Woessmann et al. (2006) show that when grouping happens before age 10, inequality in education outcomes increases at the country level.

<sup>7</sup>James (1993) argues that the mix of public and private educational services is due, for example, to the degree of religious heterogeneity within a country.

<sup>8</sup>Supermodularity implies that the marginal product of a more able worker is increasing in the ability of the

is no trade. Their main result is that comparative advantage emerges only if there are two sectors, one supermodular and one submodular.<sup>9</sup>

Unlike GM we only consider supermodular production functions which differ in the degree of complementarity of workers' skills. This choice allows us to link our analysis more easily to the existing trade literature, in which most production functions are supermodular. Our model focuses on those skills which cannot be observed ex-ante.<sup>10</sup> This focus reflects the fact (documented in section 3.2.1) that observable worker characteristics account for a minor share of total variation in IALS scores within countries, i.e. measured skill dispersion is large among workers with similar 'credentials'. Our modeling choice is also consistent with the observation that worker skills cannot be costlessly observed and that firms may take time to learn workers' skill endowments (see for example Altonji and Pierret, 2001 and Altonji, 2005). The message of the model is that the dispersion of skills among workers with otherwise identical characteristics (such as education, age or experience) affects comparative advantage. In the rest of the paper we often refer to such skills as 'residual' skills.

We consider a world where, because of frictions in the labor market, random matching in residual skills prevails between workers and firms.<sup>11</sup> The labor market frictions are similar to those assumed in Helpman and Itskhoki (2009a) and Helpman et al. (2008a; 2008b). One immediate advantage of this approach is that it can be easily applied to a world of many countries and many sectors.

---

co-worker.

<sup>9</sup>Trade emerges only conditional on the existence of a supermodular sector, where workers of identical abilities are paired together, i.e. self-matching prevails, and of a submodular sector, where the most skilled workers are paired with the least skilled co-workers, i.e. cross-matching prevails. Submodularity of the production function implies that the marginal benefit of increasing a workers' skills is decreasing in the skills of the co-worker. In this framework the country with more dispersed skill distribution specializes in the submodular sector.

<sup>10</sup>We expand on an element introduced by GM, who consider imperfect observability of skills. At the end of the paper the authors "note in passing that, with imperfect matching, trade would take place between two countries with different educational processes even if tasks were complementary in all production activities", i.e. all production functions were super-modular, which is the case we consider. We extend this model to many countries and sectors in order to derive testable implications.

<sup>11</sup>Recent microdata evidence by Brencic (2009) suggests that labor market search frictions may be considerable.

Random matching implies that the residual skills distribution prevailing in a country is inherited at the firm and industry level. This is consistent with recent international evidence (see Iranzo et al., 2008, and Lazear and Shaw, 2008) suggesting that most of wage dispersion is in fact within, rather than between, firms.

The model establishes conditions under which countries with more dispersed skill distributions specialize, and therefore export relatively more, in sectors with lower complementarity of skills in production. The empirical section of the paper examines this prediction. We adapt the empirical approach of Helpman et al. (2008c), henceforth HMR, to industry-level bilateral trade flows and augment it with our variable of interest. We show that the interaction of country skill dispersion with sectoral measures of skill substitutability is a significant and economically large determinant of exports, after controlling for a variety of trade barriers, exporting country and importing country-industry fixed effects (as dictated by the theory). We also include determinants of comparative advantage based on aggregate factor endowments as in Romalis (2004) and check that the result is not due to a correlation of country-level skill dispersion with institutional variables, like labor law rigidity and judicial quality, that other authors have found to influence trade flows.

Since the degree of substitutability across workers' skills is not directly observable, we take two distinct approaches to its measurement. First, we exploit the structure of the model, which delivers a direct link between the unobservable degree of complementarity and the observable dispersion of residual wages within industries. In our model, due to labor market frictions, workers hired by a firm are not interchangeable with workers outside such firm; therefore firms and workers engage in bargaining over the surplus. In the presence of random matching on residual skills, the resulting residual wage distribution uniquely reflects the degree of complementarity among such skills. Sectors with higher complementarity are characterized by a more compressed wage

distribution because, for example, workers with skills much higher than the average contribute to surplus relatively less, a fact reflected in their wage. In view of substantial evidence linking firm size and wages (e.g. Oi and Idson, 1999), we are careful to filter out sector-specific firm heterogeneity from our wage dispersion measures. In order to mimic random matching we also purge wages of the effect of self-selection of workers into industries. As with IALS scores, in order to bring the empirical analysis in line with the theoretical focus on unobservable skills, we purge individual workers' wages (from the US Census) of the component explained by observable characteristics, to obtain *residual* wages.

Second, we use alternative measures of the degree of skills substitutability. This set of measures is based on survey data available from the Occupational Information Network (O\*NET), which allows us to quantify the degree of teamwork, communication and interdependence between co-workers' labor inputs. These measures are not motivated by, and are independent of, our theory and provide a direct and intuitive way to proxy complementarity.

Our and GM's models are not the only ones studying the relationship between skill distribution and trade. Ohnsorge and Treffer (2007) propose a model with two-dimensional worker heterogeneity and show that, when each worker represents a bundle of two skills, the correlation of the two in the population determines comparative advantage. Grossman (2004) starts from the premise that, in some sectors, incomplete contracts make it difficult to tie remuneration to an individual worker's output. In a country with high skill dispersion highly skilled individuals prefer to sort into sectors where individual performance is easier to measure, rather than working in an industry where the common wage is dragged down by workers with relatively low skills. This type of endogenous sorting results in comparative advantage. Finally, in Bougheas and Riezman (2007) comparative advantage emerges from differential returns to skills across sectors.

Our findings relate to recent work emphasizing less traditional sources of comparative advantage. In this literature the endowment of a country, interpreted in its broadest sense, includes institutional features, such as the ability to enforce contracts (Levchenko, 2007, and Nunn, 2007), the quality of the financial system (Manova, 2008a; 2008b) and the extent of labor market frictions (Helpman and Itzhoki, 2009a, Cuñat and Melitz, 2007, Tang, 2008). We view our contribution as related to this ‘institutional endowment’ view of comparative advantage because human capital dispersion in a country is to a large extent the result of the prevailing educational system and social make-up. These, in turn, can be considered, if not immutable, a slow-moving attribute of a country.<sup>12</sup>

This paper also contributes to the large and established literature on factor endowments and comparative advantage, a topic which still receives a great deal of attention. For example, in a recent contribution to this literature, Costinot and Vogel (2009) build a model with a continuum of sectors and a continuum of skill levels and investigate the effect of trade on inequality in a rich framework.<sup>13</sup>

The paper is organized as follows. Section 2 develops a two-country multi-sector model and delivers the basic prediction about trade flows. Section A.5 in the Appendix extends the model to many countries. Section 3 presents the empirical analysis. Section 4 concludes. All proofs and a detailed data description can be found in the Appendix.

## 2 Two-Country Model

This section presents a model of trade between two countries, Home and Foreign, characterized by different skill distributions. The two countries may also vary in size, but are otherwise identical.

---

<sup>12</sup>Glaeser et al. (2004) show that education is significantly more persistent than several other institutional features, such as the form of government.

<sup>13</sup>Besides their object of interest being different from this paper’s, their assignment model yields the result that in equilibrium a sector employs only workers of a unique skill level, and as such it is not readily comparable to ours.

We denote a country by  $c$  where  $c \in \{H, F\}$ . When it does not create ambiguity we drop the country subscript. Section A.5 extends the model to a multi-country world.

## 2.1 Preferences

Each country  $c$  is populated by a measure  $L_c$  of individuals. Utility of the representative consumer depends on the consumption of a homogeneous good  $Q(0)$  and a continuum of differentiated goods  $Q(i)$  with  $i \in I$ . The utility function  $U$  is Cobb Douglas:

$$\log U = \alpha(0) \log Q(0) + \int_{i \in I} \alpha(i) \log Q(i) di$$

with  $0 < \alpha(i) < 1$  and  $\alpha(0) + \int_{i \in I} \alpha(i) di = 1$ . The aggregate  $Q(i)$  is the consumption index over the set  $\Omega(i)$  of available varieties of product  $i$  and preferences exhibit a constant elasticity of substitution  $\sigma$  across varieties of good  $i$ .<sup>14</sup> Under these preferences, demand for a given variety  $\omega$  is represented by the following equation:

$$d(\omega, i) = \frac{p(\omega, i)^{-\sigma} \alpha(i) E}{P(i)^{1-\sigma}} \quad (1)$$

where  $E$  is total expenditure,  $p(\omega, i)$  is the price of variety  $\omega$  of  $i$ , and  $P(i)$  is the ideal CES price index of aggregate  $Q(i)$ .

---

<sup>14</sup>More specifically:

$$Q(i) = \left[ \int_{\omega \in \Omega(i)} q(\omega, i)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}} \text{ with } \sigma > 1.$$

where  $q(\omega, i)$  is the quantity consumed of variety  $\omega$  of good  $i$ .



## 2.2 Production

Good  $Q(0)$  is produced under constant returns to scale by perfectly competitive firms. The technology is such that one unit of labor produces one unit of output. We choose  $Q(0)$  as our numeraire and we assume that all countries produce the numeraire good in positive quantity, which implies that the wage in sector 0,  $w(0)$ , is equal to one.

Each differentiated sector  $i$  is populated by a continuum of identical firms, each producing a different variety  $\omega$ . The market is characterized by monopolistic competition among firms, with free entry and a fixed cost of production  $f$ . The amount of output produced  $y$  depends on the skill level of each worker hired  $a$ , the measure of workers hired  $h$  and the distribution of skills across workers  $\tilde{g}(a)$ .<sup>15</sup> The distribution of skills matters for production because we assume that different levels of skills are not perfectly substitutable.<sup>16</sup> In particular, the production function of a representative firm in a sector depends on the degree of complementarity  $\lambda$  among workers' skills in that sector and takes the following form:

$$y = \left( \int a^\lambda h \tilde{g}(a) da \right)^{\frac{1}{\lambda}} \quad \text{with} \quad \frac{\sigma - 1}{\sigma} < \lambda < 1 \quad (2)$$

The parameter  $\lambda$  measures the degree of skill complementarity, since the elasticity of substitution among skills levels, for a fixed mass of workers  $h$ , is given by  $\frac{1}{1-\lambda}$ , which increases with  $\lambda$ . The larger  $\lambda$ , the more substitutable workers of different skill levels are.<sup>17</sup> The key assumption in this model is that each sector  $i$  is characterized by a different value of  $\lambda$  in production, and therefore by

---

<sup>15</sup>For simplicity we only model 'residual' skills and work under the assumption that other observable characteristics are accounted for. In this sense we do allow for selection on observables, but we do not model it explicitly. In the empirical section we carefully control for such selection.

<sup>16</sup>One possible interpretation is that the skill of each worker is a differentiated input in the production process. An alternative interpretation, along the lines of the paper by Takii and Tanaka (2009), is that each worker produces a differentiated intermediate good, in quantity proportional to her skills, and intermediate inputs are aggregated by a CES production function.

<sup>17</sup>For a fixed  $h$ , this production function is analogous to the one introduced by Grossman and Maggi (2000), p. 1261.

a different degree of complementarity among workers' skill levels. Since  $\lambda$  is the only characteristic that differentiates sectors, in the remainder of the theoretical section we drop the index  $i$  and index sectors by their parameter  $\lambda$ .

Two properties of this production function are worth discussing in detail. First, for given mass of workers  $h$ , the function is homogeneous of degree one in the skills of workers. This property stresses the relative importance of the shape, rather than location, of the distribution of skills. Second, the production function features increasing returns to the mass of workers, given the distribution of skills.<sup>18</sup> In particular,  $\lambda$  also represents the extent of increasing returns to scale (as well as the degree of complementarity), but this feature plays no substantial role in the model.<sup>19</sup> We restrict the range of  $\lambda$  to guarantee that the firm's maximization problem is concave, as described in section A.2.

### 2.3 Labor Market

We introduce labor market frictions in the spirit of Helpman and Itskhoki (2009a), although for simplicity we assume that there are no frictions in sector 0. Workers look for jobs in the homogeneous sector or in one of the differentiated good sectors. Workers are characterized by different levels of skills and skill is a continuous variable distributed in the workers' population of country  $c$  according to a density function  $g(a, c)$ . In the differentiated sectors firms pay a cost  $bh$  to randomly sample a mass  $h$  of workers from the population of workers looking for a job in that sector. The search cost  $b$  depends on labor market conditions, as described further below.

---

<sup>18</sup>This is easily seen by rewriting the production function as  $y = h^{\frac{1}{\lambda}} \left( \int a^\lambda \tilde{g}(a) da \right)^{\frac{1}{\lambda}}$ .

<sup>19</sup>We should note that it is not possible to obtain both constant returns to mass of workers and ability without confounding the quantity and quality of workers, as for example in a production function of the following type:  $y = \left( \int (ah\tilde{g}(a))^\lambda da \right)^{\frac{1}{\lambda}}$ . We give priority to maintaining constant returns to ability because we do not want to confound the degree of complementarity with differential returns to aggregate ability in different sectors. Grossman and Maggi (2000) discuss this as another case in which the distribution of ability matters and Bougheas and Riezman (2007) explicitly model this aspect in a different framework.

We make the simplifying assumption that workers ignore the full distribution of wages in all sectors (including the numeraire), except for the expected wage and the probability of sectoral unemployment.<sup>20</sup> As a result the initial distribution of residual skills in the worker population is inherited by the mass of workers looking for a job in each sector. Moreover, by definition, workers' residual skills are not observable to the firm when hiring. The combination of these assumptions yields no sorting between workers and firms. It's worth noticing that if we allowed skills to be partially observable in our model we would obtain that firms only hire workers of identical *observable* skills. Therefore we can interpret our case of unobservable skills as a residual of overall skills, once the observable component has been accounted for. For consistency with this, the empirical analysis will employ a measure of skills purged of observable characteristics.

Although the distribution of workers' skills  $\tilde{g}(a)$  could potentially be sector specific, random matching implies that every firm, in every sector  $\lambda$ , in country  $c$  inherits the residual skill distribution in the general population.<sup>21</sup>

$$\tilde{g}(a) = g(a, c)$$

## 2.4 Skill Dispersion as Comparative Advantage

Given that firms and workers match randomly with respect to unobservable skills, in this section we discuss how different skill distributions across countries generate comparative advantage. To facilitate the discussion we rewrite the production function in (2) as  $y = h^{\frac{1}{\lambda}} A(\lambda, c)$  where  $A(\lambda, c)$

---

<sup>20</sup> An alternative assumption with identical implications for the model is that workers, as in Helpman and Itskhoki (2009a) and Helpman et al. (2008a), do not know their own skills when looking for a job.

<sup>21</sup> We do not allow firms to screen workers as in Helpman et al. (2008a). We note that, contrary to the case described by Helpman et al. (2008a), with our choice of production function, firms would not want to screen workers even if the technology to screen were available, because the marginal product of an additional worker is always positive. This is the case in the static problem we are analyzing. In a dynamic framework we would expect firms to lay off unproductive workers and replace them with potentially more productive ones.

is defined as:

$$A(\lambda, c) = \left( \int a^\lambda g(a, c) da \right)^{\frac{1}{\lambda}}$$

We loosely refer to  $A(\lambda, c)$  as ‘productivity’, although clearly this is not the result of countries having access to different technologies. The magnitude of  $A(\lambda, c)$  depends on a combination of a country-specific skill distribution and a sector-specific level of complementarity across skills. We are interested in how the pattern of comparative advantage, i.e. the relative  $A$ ’s, are affected by the distribution of skills.

The general idea we explore is whether countries with lower dispersion in the distribution of skills have a comparative advantage in sectors with high degree of complementarity, i.e. where it is relatively more important to employ workers with similar skills. Since the  $A$ ’s exhibit constant returns to skills, a proportional increase in the skills of all workers increases the  $A$  by the same proportion and does not affect comparative advantage. We concentrate on comparing  $A$ ’s across countries that have the same average skills and different dispersion.<sup>22</sup> Without loss of generality countries are ordered so that, if  $c < c'$ , then country  $c'$  is characterized by a skill distribution  $g(a, c')$  that is a mean-preserving spread of the skill distribution  $g(a, c)$  in country  $c$ . We state a general condition, Property 1, for a specific pattern of comparative advantage to emerge as a result of differences in the distribution of skills.

**Property 1**  $A(\lambda, c)$  is *log-supermodular* in  $\lambda$  and  $c$ , i.e. for  $\lambda < \lambda'$  and  $c < c'$ :

$$\frac{A(\lambda, c')}{A(\lambda, c)} < \frac{A(\lambda', c')}{A(\lambda', c)} \tag{3}$$

Property 1 states that firms in countries with high skill dispersion will be relatively more pro-

---

<sup>22</sup>Note that changes in the average ability that are *not* the result of a multiplicative change in all abilities *will* affect the pattern of comparative advantage.

ductive in low complementarity sectors. As GM suggest,<sup>23</sup> a general result of this type cannot be established. Its validity is ultimately an empirical question. Therefore we verify the empirical relevance of Property 1 by employing the distributions of IALS scores observed in the data. Specifically, we construct  $A(\lambda, c)$  replacing  $g(a, c)$  with the empirical distribution of scores for 19 countries that participated in the IALS. For a grid of 100  $\lambda$ 's in the  $[0, 1]$  interval, we calculate the ratio  $\frac{A(\lambda, c')}{A(\lambda, c)}$  where country  $c'$  has higher skill dispersion (coefficient of variation of scores) than country  $c$ . We then find that, averaging across country pairs,  $\frac{A(\lambda, c')}{A(\lambda, c)}$  is increasing in  $\lambda$  for 97% of the grid points. Similar results hold if countries are ranked according to alternative measures of scores dispersion. This evidence suggests that Property 1 provides a reasonable approximation to the patterns of comparative advantage due to differences in skill dispersion. As an alternative approach, in the Appendix we also study this problem analytically and provide sufficient conditions for the theoretical validity of Property 1.<sup>24</sup>

## 2.5 Trade Flows

The previous sections established that if all sectors inherit the underlying distribution of residual skills in the country, then comparative advantage emerges from sectoral variation in the degree of skill complementarity. Under the assumption that Property 1 describes the pattern of Ricardian productivity differences, we are faced with an otherwise typical monopolistic competition model with Ricardian productivity differences across countries and sectors. Firms face an iceberg transport cost in shipping goods abroad and need to decide how much to sell in the domestic and export market. The novelty here is that, because of labor market frictions, workers hired by the firm are not interchangeable ex-post with outside workers and the firm engages in multi-lateral bargaining

---

<sup>23</sup>See p.1271.

<sup>24</sup>We show that comparative advantage can be established for any distribution if we place bounds on the degree of complementarity  $\lambda$ . Moreover, we check that Property 1 holds for any  $\lambda$  assuming specific distributions of skills (pareto, lognormal, uniform, triangular, gamma, beta and inverse gaussian).

with workers to determine wages. After computing firm-level output and revenues in all markets, free-entry allows us to calculate the equilibrium mass of firms in each sector and therefore trade flows. We provide the entire derivation in the Appendix and only report here the main predictions about trade flows.

**Proposition 1** *Under Property 1, a country with relatively higher dispersion of skills has a comparative advantage, and therefore exports relatively more to any destination, in sectors with high degree of substitutability  $\lambda$ .*

**Proof.** See Appendix. ■

As a bridge to the empirical section, section A.5 in the Appendix extends the two-country model to a multi-country world.

### 3 Empirical Analysis

The objective of this section is to assess the empirical relevance of Proposition 1 and its analogous multi-country extension presented in the Appendix. Next we discuss the estimation framework. Section 3.2 describes the data and section 3.3 reports baseline results. Finally, section 3.4 discusses identification and presents robustness checks.

#### 3.1 Estimation Framework

As a first step to design an empirical test of Proposition 1 we express the value of total exports of good  $i$  from  $H$  to  $F$  as the product of the quantity demanded of an individual variety of  $i$  from equation (1), the price and the mass of firms/varieties:

$$X_{HF}(i) = d_{HF}(i) p_{HF}(i) M_H(i) = \frac{[p_{HF}(i)]^{1-\sigma} \alpha(i) E_F}{[P_F(i)]^{1-\sigma}} M_H(i) \quad (4)$$

The price of a variety produced by a Home firm and sold in Foreign,  $p_{HF}(i)$ , depends positively on transport costs and negatively on productivity, as shown in the Appendix:

$$p_{HF}(i) = \frac{\gamma(i)\tau_{HF}}{\phi(i)A(i,H)}. \quad (5)$$

Once we substitute (5) in (4) and we take the natural logarithm, we obtain the following expression for the value of (log) exports:

$$\begin{aligned} \log X_{HF}(i) = & (\sigma - 1) \log A(i, H) + \log M_H(i) - (\sigma - 1) \log \tau_{HF} \\ & + \log \alpha(i) + \log E_F - (\sigma - 1) P_F(i) + (\sigma - 1) \log \frac{\phi(i)}{\gamma(i)} \end{aligned} \quad (6)$$

where  $A(i, H)$  captures comparative advantage of the exporting country,  $M_H(i)$  the mass of firms in the exporting country,  $\tau_{HF}$  transport costs between the two countries,  $P_F(i)$  an industry-importer specific price index,  $E_F$  the importing country total expenditure and  $\phi(i)$ ,  $\gamma(i)$  and  $\alpha(i)$  industry-specific constants. Since we consider a discrete number of industries, in the remainder of this section we use subscript  $i$  to index variables that vary across industries.

An ideal test of Proposition 1 and its multi-country analogous would require quantifying the effect of a mean-preserving spread in the distribution of residual skills in country  $H$  on its relative exports to country  $F$ , as a function of the elasticity of substitution in each sector  $i$ . These effects operate through  $A_{Hi}$  in equation 6. Although  $M_{Hi}$  is not observable, the model shows it is also a function of  $A_{Hi}$ . Therefore, in order to derive an estimation equation for  $\log X_{HF_i}$ , we assume that  $(\sigma - 1) \log A_{Hi} + \log M_{Hi}$  can be written as an additive function of industry characteristics ( $\delta_i$ ), exporter characteristics ( $\delta_H$ ), an interaction between a measure of skill substitutability in industry  $i$  ( $Substit_i$ ) and a measure of skill dispersion in country  $H$  ( $SkillDisp_H$ ), plus other unobservable

determinants of comparative advantage in country  $H$  ( $\nu_{Hi}$ ), that is,  $(\sigma - 1) \log A_{Hi} + \log M_{Hi} = \beta Substit_i \times SkillDisp_H + \delta_i + \delta_H + \nu_{Hi}$ .<sup>25</sup>

Transport costs are allowed to depend linearly on a vector of observable country-pair bilateral trade barriers ( $d_{HF}$ ) and unmeasured i.i.d. trade frictions ( $u_{HF}$ ). A set of industry-importer specific fixed effects ( $\delta_{Fi}$ ) controls non-parametrically for the price index  $P_{Fi}$ , industry constants  $\frac{\phi_i}{\gamma_i}$  and  $\delta_i$ . Finally, let  $\eta_{HF_i}$  capture measurement errors in trade flows and the effect of other unobserved determinants of  $X_{HF_i}$ .

With this specification, the estimation equation for exports takes the following form:

$$\log X_{HF_i} = \beta Substit_i \times SkillDisp_H + \gamma d_{HF} + \delta_H + \delta_{Fi} + \varepsilon_{HF_i} \quad (7)$$

where  $\varepsilon_{HF_i} = \nu_{Hi} + u_{HF} + \eta_{HF_i}$ .

The variable of interest is  $Substit_i \times SkillDisp_H$  and estimation of its coefficient  $\beta$  allows us to test Proposition 1. To see why, assume that equation (7) correctly specifies a model for the conditional expectation of  $\log X_{HF_i}$ , so that  $E[\varepsilon_{HF_i} | Substit_i \times SkillDisp_H, d_{HF}, \delta_H, \delta_{Fi}] = 0$ . Then, for any two countries  $H$  and  $G$  exporting to  $F$ , and any two industries  $i$  and  $j$ , equation (7) implies:

$$E \left[ \log \left( \frac{X_{HF_i}}{X_{GF_i}} \right) - \log \left( \frac{X_{HF_j}}{X_{GF_j}} \right) \right] = \beta \Delta_{ij} Substit \times \Delta_{HG} SkillDisp \quad (8)$$

where  $\Delta_{HG} SkillDisp \equiv SkillDisp_H - SkillDisp_G$  and  $\Delta_{ij} Substit$  is similarly defined. According to (8), Proposition 1 and its multi-country analogous imply  $\beta > 0$ .

A difficulty in implementing this test of the theory comes from the fact that the elasticity of substitution of individuals' skills at the industry level,  $Substit_i$ , is not observable in the data

---

<sup>25</sup>Note that  $A_{Hi}$  may also depend on the mean and other moments of the skill distribution of country  $H$  and these could potentially have different effect on productivity in different industries, a possibility that we explicitly consider in the empirical analysis of trade flows. These effects are summarized by  $\nu_{Hi}$ .



and we are not aware of any estimates of the elasticity of substitution for a fine disaggregation of skills. Therefore we take two different approaches to proxy for the elasticity of substitution of workers skills,  $Substit_i$ . The first is based on a theoretically-founded link between complementarity and residual wage dispersion. The second approach is to construct proxies for complementarity available from occupation-level data. Although these two approaches do not identify the elasticity of substitution,  $Substit_i$ , they allow us to rank industries in order of increasing  $Substit_i$ .

**Skill substitutability: residual wage dispersion rankings** While we provide a detailed derivation in the Appendix, here we discuss a heuristic explanation of the link between complementarity and (residual) wage dispersion. Consistent with the empirical evidence, e.g. Altonji and Pierret (2001), suggesting that firms only gradually learn about worker skills, we posit that at least part of the unobservable skills at the time of hiring are revealed to firm and worker once production begins and bargaining takes place. Hence workers of different skills receive different wages. Since our model predicts that each sector inherits the same distribution of unobservable skills, the distribution of residual wages only reflects technological differences across sectors. In particular, the degree of complementarity affects the wage of workers that are far from the average. For example, in a sector with high complementarity, a worker with high skill has a lower marginal product because her skills are very different from the average, compared with a sector with high substitutability, where high skills yield a high marginal product and high wage. Therefore sectors with low complementarity (high substitutability) have a more dispersed wage distribution. Although we do not rely on the model to structurally recover the actual value of  $Substit_i$ , we use its unambiguous prediction of a monotonic relationship between  $Substit_i$  and wage dispersion to identify a ranking of industries in terms of  $Substit_i$ .

**Skill substitutability: O\*NET rankings** In our second approach we construct proxies for complementarity using occupation-level data from O\*NET. As described in section 3.2.2, this database rates industries in three dimensions which are closely associated to skill complementarity: i) *Teamwork*: team production can naturally be thought of as a particular type of O-Ring production process (Kremer, 1993), in which the quality of final output critically depends on the successful completion of a given number of complementary tasks. (ii) *Impact on co-worker output*: a closely related way of characterizing complementarity is to quantify the extent to which a worker’s actions impact the performance of co-workers; a higher impact implies a higher degree of complementarity. (iii) *Communication/Contact*: communication and contact intensity are linked to the importance of coordinating tasks to achieve, for example, a given level of output quality; if co-workers have no need for communication or contact with each other, they are likely to have independent contributions to the final outcome. As for wage dispersion, and because we do not know the exact mapping between the O\*NET variables and  $Substit_i$ , we simply rely on O\*NET to identify a ranking of industries in terms of  $Substit_i$ .<sup>26</sup>

## 3.2 Data

Before presenting the estimation results we describe the measurement of two key explanatory variables in the empirical analysis, skill dispersion at the country level and skill substitutability at the industry level. A detailed discussion of all data can be found in the Appendix.

### 3.2.1 Skill Dispersion

We use test scores from the 1994-1998 International Adult Literacy Survey (IALS) to approximate the skill distribution within a country. Collaborators in this household survey administered a

---

<sup>26</sup>With both wage dispersion and O\*NET, regression results are qualitatively unchanged if we employ the value of the proxies instead of their ranking.

common test of work-related literacy skills to a large sample of adults between the ages of 16 and 65 in 19 countries. The IALS focuses on literacy skills that are needed for everyday tasks (e.g. working out a tip, calculating interest on a loan and extracting information), across three different dimensions of literacy: *quantitative*, *prose* and *document* literacy. We combine the results of these three tests into a single average score for each individual, measured on a scale from 0 to 500. The skill distribution is proxied by the distribution of log-scores of individuals participating in the labor market and living in the same country.

To ensure consistency with the theoretical assumption of imperfect skill observability, we construct a measure of *residual* scores dispersion within countries. For an individual  $k$  participating in the labor market of country  $H$ , we obtain the estimated residual  $\widehat{\epsilon}_{kH}$  from the following regression:

$$\log(s_{kH}) = X_{kH}\beta_H + \epsilon_{kH} \tag{9}$$

where  $s_{kH}$  is the IALS score of  $k$  and  $X_{kH}$  is a vector of individual demographic information from the IALS questionnaire. The residual  $\widehat{\epsilon}_{kH}$  is then used to compute the skill dispersion measures used for the estimation of trade flows. Analyzing the R-squared of these country-by-country regressions, we find that the variation in residual scores  $\widehat{\epsilon}_{kH}$  accounts for a minimum of 46% of the observed variation in log-scores in Canada, for a maximum of 83% in Germany and for 70% in Finland, the median country in the sample.

Table 1 ranks 19 countries according to the coefficient of variation (CV) of IALS scores, and also reports their rank by mean, standard deviation (St Dev) and standard deviation of residual IALS (St Dev Res). The figures show different dispersion in countries at similar stages of development: for example, we observe a more spread distribution of skills in the US, UK and Canada, than in

Sweden, the Netherlands and Germany.<sup>27</sup>

### 3.2.2 Substitutability

In this section we describe the construction of the two rankings of skill substitutability at the industry level, based on residual wage dispersion and O\*NET indices.

**Residual Wage Dispersion** We use the 5% Public Use Microdata Sample (PUMS) files of the 2000 Census of Population in the United States to construct industry-specific measures of wage dispersion to identify a ranking of industries in terms of the unobserved elasticity of substitution. An advantage of our approach is that we can match individual wage observations to a detailed industry classification, accounting for the entire manufacturing sector<sup>28</sup>. This procedure results in 63 industries for which both wage dispersion and international trade flows can be computed, at a level of aggregation between the 3 and 4 digit levels of the 1997 North American Industry Classification System (NAICS).

As with IALS scores, we focus on residual wage dispersion. We start by removing variation in wages driven by individual characteristics on which firms can typically condition employment decisions. We also adapt the correction method proposed in Dahl (2002) to address the possibly non-random selection of workers into multiple industries. In essence, this procedure controls for selection effects using differences in the probabilities of being observed in a given industry due to exogenous variation, such as the state of birth of two people who are otherwise similar in terms of education, experience, household structure, race and gender. Details are provided in the Appendix.

---

<sup>27</sup>Brown et al. (2007) report similar variation in skill distributions in a comprehensive study using IALS, the 1995, 1999 and 2003 Trends in International Maths and Science Study (TIMSS), the 2000 and 2003 Programme for International Student Assessment (PISA) and the 2001 Progress in International Reading Literacy Study (PIRLS).

<sup>28</sup>This is not feasible for IALS data, since individual observations are assigned a broad sectoral classification (e.g. agriculture, mining, manufacturing, construction, etc), while international trade data is available only for manufacturing industries.

For an individual  $k$  employed in industry  $i$ , we obtain the estimated residual  $\widehat{\xi}_{ki}$  from the following regression:

$$\log(w_{ki}) = Z_{ki}\beta_i + \xi_{ki} \quad (10)$$

where  $w_{ki}$  is the weekly wage of  $k$  and  $Z_{ki}$  is a vector of observable characteristics (age, gender, etc.). Note that we run these regressions separately for each industry to allow for differences in the return to observable characteristics across industries.<sup>29</sup>

Several studies have shown that firm size affects workers' wages.<sup>30</sup> This implies that wage dispersion might also reflect variation in the distribution of firm size across different industries. Although the model does not incorporate firm heterogeneity, we purge residual wage dispersion of the effect of firm heterogeneity in order to isolate the degree of complementarity. Since the Census does not provide the size of the establishment at which individual workers are employed, we regress measures of dispersion of  $\widehat{\xi}_{ki}$  on the coefficient of variation of firm size within industry  $i$ ,  $FirmDisp_i$ . The residuals from this regression are employed to construct  $WageDisp_i$ , a ranking of industries in Table 2a, where we report the top and bottom 5. For example, in terms of the standard deviation of residual wages, the three lowest ranked sectors are railroad, ship building and aerospace. The three highest ranked are apparel accessories, bakeries and cut and sew apparel.

The use of U.S. estimates as proxies for within-industry wage dispersion (and skill substitutability) in other countries is warranted if they have access to similar production technologies.<sup>31</sup> Equal access to technology implies that the elasticity of substitution in any given industry will be constant across countries. As a result, the ranking of industries according to wage dispersion will be the same

<sup>29</sup>Regression results are available upon request.

<sup>30</sup>See Oi and Idson (1999).

<sup>31</sup>The assumption that industry-specific characteristics computed for the United States also apply to industries in other countries is not an unusual one in the recent empirical trade literature on comparative advantage. Examples include the measurement of financial vulnerability (Manova, 2008b), the importance of relationship-specific investment (Nunn, 2007), firm-specific skill intensity (Tang, 2008) and the variance of firm-specific shocks (Cuñat and Melitz, 2007).

within each country, a hypothesis that is not easy to verify due to the scarcity of publicly available microdata with similar sector classification. However, we do perform this exercise for the U.S. and Canada. We compute the sectoral dispersion of wage residuals in Canada to verify whether the ranking is similar to the one prevailing in the US.<sup>32</sup> To maximize comparability, we are careful to control for the same set of observable characteristics of workers in both countries when computing the residuals, use similar sampling criteria and the same industry classification. Figure 2 shows industry rankings in terms of the standard deviation of the wage residuals in the two countries. The positive slope of the fitted line is significant at the 1% level. Clearly, the sectoral ranking of residual dispersion in the US is strongly correlated to the one observed in Canada. Sectors like computers and clothing exhibit higher dispersion in both countries, compared to sectors like machinery and paper manufacturing.

**O\*NET survey-based measures of complementarity** Sponsored by the Employment and Training Administration of the United States Department of Labor, O\*NET provides detailed information on job requirements and worker attributes for 965 occupations in the U.S. Information on 277 descriptors including abilities, work styles, work context, interests, experience and training, is annually updated by ongoing surveys of each occupation’s worker population and occupational experts.

As anticipated in section 3.1, our complementarity rankings are based on four selected O\*NET (Version 12.0) questions capturing different aspects of skill complementarity: (1) *Teamwork*: How important are interactions that require you to work with or contribute to a work group or team to perform your current job?<sup>33</sup> (2) *Impact*: How do the decisions an employee makes impact the results

---

<sup>32</sup>We use the Canadian Labor Force Survey data for May 2000. Details of this exercise are available upon request.

<sup>33</sup>An alternative measure of teamwork can be obtained from the Detailed Work Activities (a supplemental file to O\*NET). Reported results are qualitatively unchanged when this measure is used.

of co-workers, clients or the company? (3) *Communication*: How important is communicating with supervisors, peers or subordinates to the performance of your current job? (4) *Contact*: How much contact with others (by telephone, face-to-face, or otherwise) is required to perform your current job? Respondents were asked to rate these questions on a scale from 1 to 5. The O\*NET database provides average scores for each occupation.

In constructing industry-level proxies of complementarity, O\*NET scores were matched to the 2000 Census microdata.<sup>34</sup> In this way, because occupational structures vary across industries, we obtain a different distribution of scores for each industry. Using the median score for each industry we generate  $O^*NET_i$ , a ranking of sectors in terms of substitutability.<sup>35</sup> Industries with higher  $O^*NET_i$  exhibit lower skill substitutability. Table 2a reports the ranking in terms of  $Contact_i$  for the top and bottom 5 industries as ranked according to residual wage dispersion (other O\*NET variables produce similar rankings). The table shows that among the lowest ranked sectors in terms of wage dispersion appear the top ranked sectors in terms of O\*NET measures. These are the low substitutability sectors. Similarly, among the highest ranked sectors in terms of  $WageDisp_i$  we find the bottom  $O^*NET_i$  sectors (those sectors with high substitutability). This reflects the fact that, as shown in table 2b,  $O^*NET_i$  and  $WageDisp_i$  are inversely correlated. Although weakly significant, correlation signs among substitutability rankings are consistent with the expected pattern.

### 3.3 Baseline Results

This section discusses results of the empirical analysis of trade flows using specification (7). The dependent variable in tables 3 to 5 is the log of exports from country  $H$  to country  $F$  in industry  $i$ . Our data set contains the value of exports in year 2000 from 19 exporters to 145 importers in 63

---

<sup>34</sup>This is possible since the occupational classifications in both O\*NET and the Census are based on the Standard Occupational Classification.

<sup>35</sup>The results reported in the empirical section are robust to reweighting by hours worked and to using mean scores instead of medians as complementarity proxies.

industries. We first report results when  $Substit_i$  is proxied by a wage dispersion ranking  $WageDisp_i$  and later show similar quantitative findings when we utilize survey-based complementarity rankings from O\*NET.

### 3.3.1 Results with Substitutability proxied by Wage Dispersion Rankings

Table 3 reports estimates of the impact of skill dispersion as proxied by the dispersion of *residual* IALS test scores (defined in section 3.2.1): we identify this effect through an interaction with *residual* wage dispersion rankings (defined in section 3.2.2). For comparability, all tables report standardized coefficients of the explanatory variables. The measures of dispersion employed in table 3 are: the standard deviation in columns (1) and (4), the 95-5 interpercentile range in columns (2) and (5), and the Gini mean difference in columns (3) and (6). Columns (1)-(3) add exporter, importer and industry dummies to our variables of interest; columns (4)-(6) include theoretically consistent exporter and importer-industry dummies, along with a vector of bilateral trade barriers described in the Appendix. We find that  $WageDisp_i \times SkillDisp_H$  has a positive and significant effect on exports. We note that the magnitudes of the coefficient are stable across specifications and measures of dispersion. The standardized coefficient of  $WageDisp_i \times SkillDisp_H$  varies between 7.4% and 8.2% in the six specifications.

We employ the estimated coefficients to gauge the economic magnitude of this source of comparative advantage. The standardized coefficient of  $WageDisp_i \times SkillDisp_H$  is similar across specifications. Our baseline estimate is 0.079 (column 4, table 3). Consider two countries, the US and Canada, and two sectors, ‘computers’ and ‘plastics’. These countries and sectors are chosen because, going from plastics’  $WageDisp$  interacted with Canada’s  $SkillDisp$  to computers’s  $WageDisp$  interacted with the US  $SkillDisp$ , the interaction  $WageDisp_i \times SkillDisp_H$  increases by approximately one standard deviation. Since the standard deviation of log exports is 2.204, the



relative ratios of US and Canada’s exports (to an average importer  $c$ ) in the two sectors are given by  $e^{0.079 \times 2.204}$ , that is:

$$\frac{\frac{X_{USc}(computers)}{X_{CANADAc}(computers)}}{\frac{X_{USc}(plastics)}{X_{CANADAc}(plastics)}} = 1.19$$

This implies that, averaging across destination markets, the US exports of computers relative to Canada are 19% higher than the US exports of plastics relative to Canada (this number varies between 17.7% and 19.8% depending on the measure of dispersion and specification).

Section D in the Appendix shows that similar results are obtained if raw wages and raw scores are employed in building measures of dispersion, that is wages and scores before we filter out the effect of observables.

### 3.3.2 Results with Substitutability proxied by O\*NET rankings

Next, we report estimates of the effect of skill dispersion on trade flows using four alternatives measures of skill complementarity constructed from the O\*NET database. Table 4 replicates the structure of columns (4)-(6) of Table 3, in terms of the set of fixed effects included and trade barriers used as controls. The variable of interest is the interaction of  $SkillDisp_H$  (measured by the standard deviation of residual scores) and the corresponding O\*NET ranking:  $Teamwork_i$ ,  $Impact_i$ ,  $Communic_i$  and  $Contact_i$ . Note that since O\*NET rankings are proxying for complementarity, the expected sign of the interaction is negative (i.e. countries with a higher skill dispersion export relatively less in industries with high skill complementarity). This is confirmed in every specification of table 4 at the 1% significance level. The estimates of the effect of skill dispersion are quantitatively very similar to the ones generated using  $WageDisp_i$ . In unreported regressions we check that these results are qualitatively unchanged if (i)  $SkillDisp_H$  is measured as either the 95-5 interpercentile range or the Gini mean difference of residual scores; (ii) importer-industry fixed effects are replaced

by importer and industry fixed effects; (iii) trade barriers are not included in the estimation and (iv) O\*NET rankings are computed using the mean score of occupations in the industry rather than the median.

### 3.4 Identification and Robustness

In this section we discuss some potential issues related to the identification of the effects quantified in tables 3 and 4. For parsimony we present our robustness analysis only for the residual wage dispersion as a proxy for substitutability. All the results presented below also hold when using the O\*NET rankings of substitutability.<sup>36</sup>

#### 3.4.1 The Extensive Margin of Trade: Selection

Tables 3 and 4 report estimation results which do not take into account the fact that a substantial fraction of bilateral trade flows are zero and that trade flows reflect both an intensive margin (the amount exported by each firm) and an extensive margin (the number of firms exporting, possibly zero). The estimation of (7) requires excluding observations for countries which do not trade in specific industries. These amount to 66.5% of the sample. As discussed in HMR, selection of trading partners induces a negative correlation between observed and unobserved trade barriers ( $d_{HF}$  and  $u_{HF}$ ) that might bias OLS estimates in (7), including  $\beta$ .

In order to correct for selection bias, we implement a two-step estimation procedure: in the first step we account for the discrete export decision using a linear probability model and obtain the predicted probabilities of observing positive exports,  $\widehat{\varphi}_{HF_i}$ ; in the second stage, equation (7) is estimated including a flexible polynomial of degree four in  $\widehat{\varphi}_{HF_i}$  to control for selection bias.<sup>37</sup> For

---

<sup>36</sup> Estimation results are available from the authors.

<sup>37</sup> We favor using a linear probability model in the first stage since its two most common alternatives, probit and logit models, suffer different problems in the current application. The probit model with fixed effects yields inconsistent estimates. In turn, estimating a fixed effects logit becomes computationally very costly due to the large number of

identification not to rely on the non-linearity of  $\widehat{\varphi_{HF_i}}$  one needs to identify a source of variation which affects the discrete choice of engaging in exports without changing the intensity of trade flows. HMR argue that cross-country variation in start-up regulation costs likely relates to the decision to export, and it has no bearing on the intensive margin. The economic rationale lies in the fact that start-up costs in the exporting country, as well as in the importing one, affect fixed rather than variable costs of trade. Different forces can be at work and the nature and strength of this effect may depend on characteristics of both exporting and importing countries. For example, HMR find that start-up regulation costs are an effective predictor of the extensive export decision and that the interaction between home and foreign regulation costs has a negative gradient on the likelihood to export. On the other hand, De Groot et al. (2004) show that *differences* in institutional factors, including differences in regulation and red tape, have large effects on trade flows; their work unveils an alternative channel through which regulation can affect trade, and stresses the importance of ‘similarity’ in institutional frameworks.

An analysis of the first-stage bilateral export decisions (see table 6) uncovers strong effects of regulation costs. We use exporter-importer interactions of three proxies of regulation costs: the number of days ( $RegDays_H \times RegDays_F$ ), number of legal procedures ( $RegProc_H \times RegProc_F$ ) and relative cost, as a percentage of GDP per capita ( $RegProc_H \times RegProc_F$ ), for an entrepreneur to start operating a business.<sup>38</sup> We find that these proxies are significant predictors of selection into exporting.<sup>39</sup>

In table 5 we report the second stage obtained using the selection correction. To facilitate

---

fixed effects required by our specification of equation (7).

<sup>38</sup>To test the overidentifying restrictions we performed a Hausman test comparing second stage estimates using all three instruments to the corresponding estimates using only a subset of them. We tested all possible combinations of exclusion restrictions and in no case could we reject the null hypothesis that they are valid and, therefore, estimates with different restrictions only differ as a result of sampling error.

<sup>39</sup>In fact, as might be expected, we find that regulatory costs tend to have a direct negative effect on export choices, but also that relative differences across countries do matter, and can lead to positive interaction effects. Additional details available from the authors.

comparison, column (1) of table 5 is identical to column (4) of table 3, which is the baseline result using  $WageDisp_i$ , i.e. we employ only one of the three measures of dispersion, the standard deviation.<sup>40</sup> Columns (2)-(6) report the second stage of the selection-corrected estimation. Column (2) documents the robustness of the effect associated to the interaction  $WageDisp_i \times SkillDisp_H$ : the standardized coefficient is largely unchanged at 0.071.

### 3.4.2 Omitted Determinants of Comparative Advantage

A second potential source of bias is due to the omission of other determinants of comparative advantage, possibly correlated to our variable of interest. Suppose that the true model includes an additional term  $n_i Z_H$ . If  $WageDisp_i$  were correlated with  $n_i$  and  $SkillDisp_H$  were correlated with  $Z_H$ , the OLS estimate of  $\beta$  in equation (7) would be inconsistent. As an example, industries with lower dispersion of wages tend to be capital intensive. Similarly, exporters with low skill dispersion tend to be relatively abundant in aggregate physical capital.<sup>41</sup> In this case, comparative advantage driven by skill dispersion is correlated with comparative advantage deriving from standard factor proportions theory.

Columns (3) to (5) of table 5 show that the estimated effect of the interaction  $WageDisp_i \times SkillDisp_H$  is robust to a number of controls for other potential determinants of comparative advantage. Column (3) introduces controls for standard Heckscher-Ohlin sources of comparative advantage: the interaction of factor endowment of a country (in particular human capital,  $SkillEndow_H$  and physical capital,  $KEndow_H$ ) and factor intensity of the sector (human capital  $SkillIntens_i$  and physical capital,  $KIntens_i$ ), as in Romalis (2004). Looking at 95% confidence intervals, the impact on trade flows of our variable of interest  $WageDisp_i \times SkillDisp_H$  is quanti-

<sup>40</sup>The same qualitative results emerge if we employ the other two measures of dispersion.

<sup>41</sup>In our dataset, the correlation between the coefficient of dispersion of residual wages and physical capital intensity across industries is -0.511. In turn, the correlation between the standard deviation of residual IALS scores and physical capital abundance across exporters is -0.524.

tatively similar to the interaction  $SkillIntens_i \times SkillEndow_H$  and in the same order of magnitude as  $KIntens_i \times KEndow_H$ . In column (4) we control for the interaction between  $WageDisp_i$  and institutional features of countries that might be correlated with  $SkillDisp_H$ . Our concern is that, to the extent that  $WageDisp_i$  displays a similar pattern to other characteristics of sectors that make them benefit from those institutional features, our interaction of interest could be capturing alternative channels that have been found empirically relevant in the literature. In particular, we interact  $WageDisp_i$  with  $LaborRigid_H$  (a measure of labor law rigidity in country  $H$ ) and with  $JudicQual_H$  (a measure of judicial quality). These alternative controls do not substantially affect the magnitude of our variable of interest. In column (5) we introduce the share of individual wages that are top-coded within an industry,  $TopCode_i$ , interacted with  $SkillDisp_H$ , to show that our result is not driven by the fact that some sectors rely on ‘superstars’ (those sectors that have a high share of top-coded wages). This suggests that more than one aspect of the dispersion of the distribution of wages is driving the result, and that the overall shape of the distribution seems to be better captured by broader measures of dispersion.

### 3.4.3 Reverse Causality

Finally,  $WageDisp_i$  and  $SkillDisp_H$  might be partly influenced by the pattern of international trade, potentially resulting in reverse causality. We explore this possibility by examining the relationship between each of these two variables and the error term  $\varepsilon_{HF_i}$ . The orthogonality condition needed for consistent estimation of  $\beta$  in equation (7) is:

$$E(WageDisp_s \times SkillDisp_c \times \varepsilon_{HF_i}) = 0 \quad \forall s, c \quad (11)$$

By the Law of Iterated Expectations, a sufficient condition to obtain identification is:

$$E(WageDisp_s \times \varepsilon_{HF_i} | SkillDisp_c) = 0 \quad \forall s, c \quad (12)$$

which requires that, for every exporter in our sample, within-industry wage dispersion be uncorrelated with unobserved determinants of trade. For example, a violation of (12) would arise if  $\varepsilon_{HF_i}$  contained the unobserved share of exporting firms in a given sector in  $H$  and the proportion of exporters varied across industries and importers. In a model with heterogeneous firms, Helpman et al. (2008a) show that within-industry wage dispersion is a function of the proportion of firms exporting in the industry since, on average, exporters pay higher wages than non-exporters.<sup>42</sup> However, as shown in HMR, the correction for self-selection into the export market discussed in section 3.4.1 effectively removes this potential bias.

Furthermore, since we measure wage dispersion at the industry level using U.S. data, we can check the robustness of our estimates by removing the U.S. from our set of exporters. To the extent that the U.S. wage structure is not significantly affected by bilateral trade flows between other countries, this procedure substantially decreases the likelihood of feedback effects running from trade flows to  $WageDisp_s$ . Column (6) in table 5 shows that, also in this case, the coefficient of our interaction of interest maintains the same magnitude and significance.

An alternative sufficient condition that guarantees (11), and therefore identification of  $\beta$ , is

$$E(SkillDisp_c \times \varepsilon_{HF_i} | WageDisp_s) = 0 \quad \forall s, c$$

which means that, for every sector, skill dispersion in every exporting country is uncorrelated with the error term  $\varepsilon_{HF_i}$ . This condition is satisfied if unobserved exporting opportunities captured in

---

<sup>42</sup>Exporters do pay higher wages. See, for example, Bernard et al. (1995) and Bernard and Jensen (1997).

$\varepsilon_{HF_i}$  are not significantly related to the dispersion, and overall distribution, of residual skills in a country. There are several reasons to believe that this is plausible. First, the unobserved exporting opportunities  $\varepsilon_{HF_i}$  must occur at levels other than exporter or importer-industry, which are already captured by our set of dummies. Moreover, since our skill dispersion measures pre-date trade flows by several years, the link between  $\varepsilon_{HF_i}$  and  $SkillDisp_c$  introduces bias only if: (i)  $\varepsilon_{HF_i}$  is a highly persistent shock to exporting opportunities which is not captured by our dummies and also affects the long-term, ‘residual’ skill distribution, and (ii) the skill distribution reacts very quickly in response to export shocks. In this respect Glaeser et al. (2004) show that the education system is a slow-changing characteristic of a country. However, skill dispersion is not only the product of the formal education system, but may change after school through on-the-job training. A number of papers have established the relatively limited impact of on-the-job training on the overall level of human capital.<sup>43</sup> Nevertheless, we explicitly account for the possibility that re-training is triggered by exporting opportunities through the inclusion, in the derivation of residual skills, of a control for whether a worker was re-trained in the previous year.

## 4 Conclusions

Relative differences in the distribution of production factors are central to the classical theory of international trade. The Heckscher-Ohlin-Samuelson factor proportion model stresses the idea that differences in factor endowments play a major role in predicting trade flows. Comparative advantage is associated with relatively abundant factors of production: the aggregate endowment of some important factor can be a driving force in determining international specialization. In this paper we push this idea further and argue that the whole distribution of factors, and not

---

<sup>43</sup>See discussion in Carneiro and Heckman (2003) and Adda et al. (2006).

just their aggregate endowment, can help rationalize observed trade flows. We focus on skills, and use industry-level trade data to show that factors' dispersion accounts for as much as factors' endowment in the determination of trade flows.

We develop a theoretical framework where, because of frictions in the labor market and ex-ante unobservable skills, workers and firms are randomly matched. The skill distribution matters for different sectors because some industries are more capable to substitute workers of different skills than others. All sectors inherit the distribution of (unobserved) skills in the country's population and, as a result, firms in sectors with higher complementarity are relatively more productive in countries with lower skill dispersion. Our model provides an observable proxy for the otherwise unobservable degree of complementarity among workers' skills, that is the dispersion of residual wages at the industry level. We also use other proxies for skill substitutability which do not rely on our theoretical framework and are based on the O\*NET occupational survey. Detailed data on industry-level bilateral trade flows reveal that countries with higher residual skill dispersion specialize in low complementarity sectors. This empirical finding is robust to a variety of controls and indicates that the dispersion of human capital is not only statistically significant, but also quantitatively large: in fact we find that the magnitude of its effect on trade flows is comparable to that of the aggregate endowment of human capital. The two alternative measures of substitutability employed produce results that are qualitatively and quantitatively very similar.

Our analysis focuses on the impact of *residual* skill dispersion: in the model this means analyzing *unobservable* skills; in the empirical analysis it translates into purging skills and wages of all characteristics that are observable to the econometrician. To the extent that a substantial component of residual wages and skills is observed by firms and workers, but unobservable to the econometrician, two remarks about the interpretation of our evidence are in order. First, our empir-



ical results are not inconsistent with a model of observable skills like GM's: we find that countries with high skill dispersion specialize in sectors with high wage dispersion. In our model wage dispersion only reflects the degree of complementarity, and not compositional effects. Conversely, in GM, any differences in the sectoral wage distribution is due exclusively to industries employing workers of different skills: the supermodular sector employs similar workers, the submodular sector employs workers at the tails of the skill distribution. We expect that a multi-country, multi-sector extension of GM could be consistent with the empirical evidence that this paper presents. We are not aware of such an extension and we believe it would be non-trivial.<sup>44</sup> Second, we hypothesize that a Heckscher-Ohlin-Samuelson model of factor proportions with a large number of factors and different factor intensities across sectors would potentially yield testable implications similar to our model. Our results indicate that such a model should encompass a much finer level of disaggregation of factors than Heckscher-Ohlin-Samuelson-type models and their empirical tests have employed so far.<sup>45</sup>

Finally, the theoretical framework developed in this paper has implications for the impact of trade on overall wage inequality, which are beyond the scope of this study. Our model, taken at face value, implies that a more disperse skill distribution does not just have a direct effect on the income distribution, but also an indirect effect, as countries with higher skill dispersion specialize in sectors with high wage dispersion. Although we consider this effect intriguing, we are aware that our static, stylized description of the labor market is not sophisticated enough to account for alternative determinants of overall inequality.

---

<sup>44</sup>As previously noted, trade does not emerge in GM with supermodular sectors and observable skills. Therefore such an extension with  $n$  sectors would have to feature  $n - 1$  submodular industries, that exhibit different degrees of submodularity.

<sup>45</sup>Tests of the factor proportions theory typically involve a dichotomous classification of workers into production and non-production, or college and non-college educated.

## References

- Acemoglu, D., Antràs, P., Helpman, E., 2007. Contracts and technology adoption. *American Economic Review* 97 (3), 916–943.
- Adda, J., Dustmann, C., Meghir, C., Robin, J., Street, G., 2006. Career progression and formal versus on-the-job training. IZA Discussion Paper 2260.
- Altonji, J., 2005. Employer learning, statistical discrimination and occupational attainment. *American Economic Review*, 112–117.
- Altonji, J., Pierret, C., 2001. Employer learning and statistical discrimination. *Technology* 313.
- Antweiler, W., Trefler, D., 2002. Increasing returns and all that: a view from trade. *American Economic Review*, 93–119.
- Aumann, R. J., Shapley, L. S., 1974. Values of non-atomic games. Princeton University Press Princeton.
- Beaudry, P., Green, D., Sand, B., 2007. Spill-Overs from Good Jobs. NBER working paper 13006.
- Behrens, K., Lamorgese, A., Ottaviano, G., Tabuchi, T., 2009. Beyond the Home Market Effect: Market Size and Specialization in a Multi-Country World. Mimeo, University of Tokyo.
- Benabou, R., 1996. Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance. *The American Economic Review*, 584–609.
- Bernard, A., Jensen, J., 1997. Exporters, skill upgrading, and the wage gap. *Journal of International Economics* 42 (1-2), 3–31.
- Bernard, A., Jensen, J., Lawrence, R., 1995. Exporters, jobs, and wages in US manufacturing: 1976-1987. *Brookings Papers on Economic Activity. Microeconomics*, 67–119.
- Botero, J., Djankov, S., La Porta, R., Lopez-de Silanes, F., Shleifer, A., 2004. The Regulation of Labor. *Quarterly Journal of Economics* 119 (4), 1339–1382.
- Bougheas, S., Riezman, R., 2007. Trade and the distribution of human capital. *Journal of International Economics* 73 (2), 421–433.
- Bowen, H. P., Leamer, E. E., Sveikauskas, L., 1987. Multicountry, multifactor tests of the factor abundance theory. *The American Economic Review*, 791–809.
- Brencic, V., 2009. Do employers respond to the costs of continued search? *Oxford Bulletin of Economics and Statistics*.
- Brown, G., Micklewright, J., Schnepf, S., Waldmann, R., 2007. International surveys of educational achievement: how robust are the findings? *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 170 (3), 623–646.
- Carneiro, P., Heckman, J., 2003. Human Capital Policy. NBER Working Paper 9495.
- Costinot, A., Vogel, J., 2009. Matching and Inequality in the World Economy. Mimeo, MIT.

- Cuñat, A., Melitz, M. J., 2007. Volatility, Labor Market Flexibility, and the Pattern of Comparative Advantage. NBER Working Paper 13062.
- Dahl, G., 2002. Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica*, 2367–2420.
- Davis, D. R., Weinstein, D. E., 2001. An account of global factor trade. *American Economic Review*, 1423–1453.
- De Groot, H., Linders, G., Rietveld, P., Subramanian, U., 2004. The institutional determinants of bilateral trade patterns. *Kyklos* 57 (1), 103–123.
- Feenstra, R., Lipsey, R., Deng, H., Ma, A., Mo, H., Drive, O., 2005. World trade flows: 1962-2000. NBER working paper 11040.
- Friesen, J., Krauth, B., 2007. Sorting and inequality in Canadian schools. *Journal of Public Economics* 91 (11-12), 2185–2212.
- Glaeser, E., La Porta, R., Lopez-de Silanes, F., Shleifer, A., 2004. Do institutions cause growth? *Journal of Economic Growth* 9 (3), 271–303.
- Grossman, G. M., 2004. The Distribution of Talent and the Pattern and Consequences of International Trade. *Journal of Political Economy* 112 (1), 209–239.
- Grossman, G. M., Maggi, G., 2000. Diversity and trade. *American Economic Review*, 1255–1275.
- Hanushek, E., Kain, J., Markman, J., Rivkin, S., 2003. Does peer ability affect student achievement? *Journal of Applied Econometrics*, 527–544.
- Heckman, J., Scheinkman, J., 1987. The importance of bundling in a Gorman-Lancaster model of earnings. *The Review of Economic Studies*, 243–255.
- Helpman, E., Itskhoki, O., 2009a. Labor Market Rigidities, Trade and Unemployment. Harvard manuscript.
- Helpman, E., Itskhoki, O., 2009b. Labor Market Rigidities, Trade and Unemployment: A Dynamic Model. Mimeo, Harvard University.
- Helpman, E., Itskhoki, O., Redding, S. J., 2008a. Inequality and unemployment in a global economy. Mimeo, Harvard University.
- Helpman, E., Itskhoki, O., Redding, S. J., 2008b. Wages, Unemployment and Inequality with Heterogeneous Firms and Workers. Mimeo, Harvard University.
- Helpman, E., Krugman, P. R., 1985. Market structure and foreign trade. MIT press Cambridge, Mass.
- Helpman, E., Melitz, M., Rubinstein, Y., 2008c. Estimating Trade Flows: Trading Partners and Trading Volumes\*. *Quarterly Journal of Economics* 123 (2), 441–487.
- Hoxby, C., Building, L., 2000. Peer effects in the classroom: Learning from gender and race variation. NBER working paper 7867.

- Iranzo, S., Schivardi, F., Tosetti, E., 2008. Skill dispersion and firm productivity: an analysis with employer-employee matched data. *Journal of Labor Economics* 26 (2), 247–285.
- James, E., 1993. Why do different countries choose a different public-private mix of educational services? *Journal of Human Resources*, 571–592.
- Kaufmann, D., Kraay, A., Mastruzzi, M., 2003. Governance matters III: Governance indicators for 1996-2002. World Bank Working Paper No.3106.
- Kremer, M., 1993. The O-ring theory of economic development. *The Quarterly Journal of Economics*, 551–575.
- Lazear, E., Shaw, K., 2008. *The Structure of Wages: An International Comparison*. University of Chicago Press.
- Levchenko, A. A., 2007. Institutional Quality and Trade. *Review of Economic Studies* 74 (3), 791–819.
- Manova, K., 2008a. Credit constraints, equity market liberalizations and international trade. *Journal of International Economics* 76 (1), 33–47.
- Manova, K., 2008b. Credit Constraints, Heterogeneous Firms, and International Trade. Stanford University mimeo.
- Nunn, N., 2007. Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade\*. *The Quarterly Journal of Economics* 122 (2), 569–600.
- Ohnsorge, F., Treffer, D., 2007. Sorting It Out: International Trade with Heterogeneous Workers. *Journal of Political Economy* 115 (5), 868–892.
- Oi, W., Idson, T., 1999. Firm size and wages. *Handbook of labor economics* 3 (2), 2165–2214.
- Romalis, J., 2004. Factor proportions and the structure of commodity trade. *American Economic Review*, 67–97.
- Ross, S. A., 1981. Some stronger measures of risk aversion in the small and the large with applications. *Econometrica: Journal of the Econometric Society*, 621–638.
- Stevenson, D., Baker, D., 1991. State control of the curriculum and classroom instruction. *Sociology of Education*, 1–10.
- Stole, L. A., Zwiebel, J., 1996. Intra-firm bargaining under non-binding contracts. *The Review of Economic Studies*, 375–410.
- Takii, K., Tanaka, R., 2009. Does the diversity of human capital increase gdp? a comparison of education systems. Forthcoming *Journal of Public Economics*.
- Tang, H., 2008. Labor Market Institutions, Firm-specific Skills, and Trade Patterns. Mimeo, Hong Kong University of Science and Technology.
- Trefler, D., 1993. International factor price differences: Leontief was right! *Journal of Political Economy*, 961–987.

- Trefler, D., 1995. The case of the missing trade and other mysteries. *The American Economic Review*, 1029–1046.
- Tuijnman, A. (Ed.), 2000. *Literacy in the Information Age: Final Report of the International Adult Literacy Survey*. OECD Paris.
- Woessmann, L., Hanushek, A., Str, P., 2006. Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Economic Journal* 116 (510), C63–C76.



Figure 1: Mean and Dispersion in IALS log-scores (1994-1998)

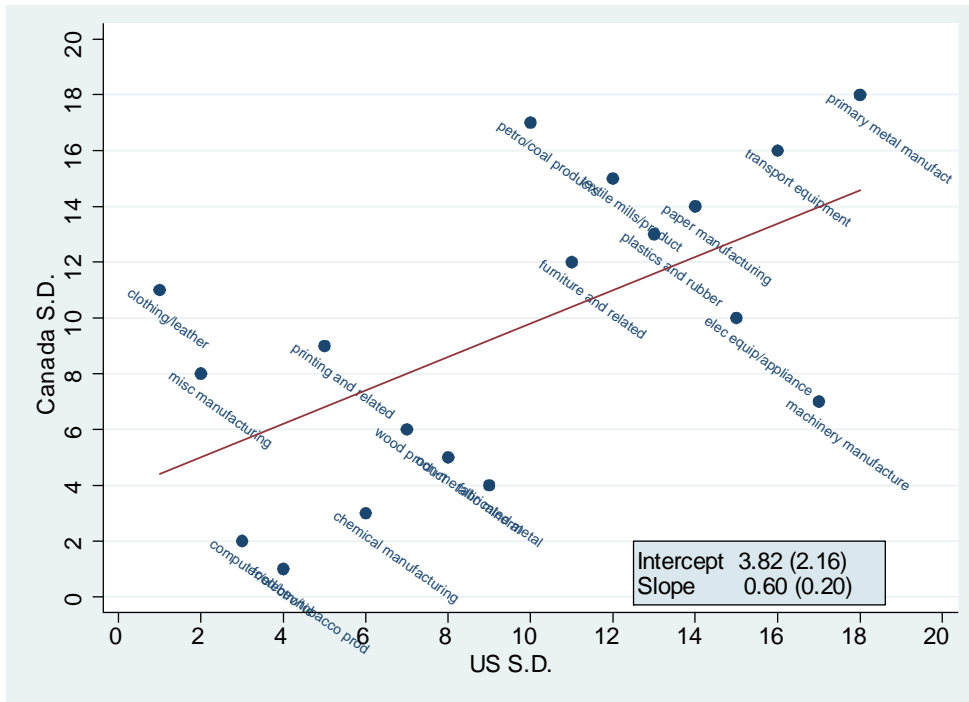


Figure 2: Industry Rankings in terms of Standard Deviation of Residual Wages

Table 1 - IALS log-scores

CV Rank	Exporter	<i>Mean</i>		<i>St Dev</i>		<i>St Dev Res</i>	
		Rank		Rank		Rank	
1	Denmark	3	5.671	1	0.150	1	0.134
2	Germany	6	5.654	2	0.162	3	0.151
3	Netherlands	4	5.666	3	0.167	2	0.149
4	Norway	2	5.684	4	0.171	4	0.156
5	Finland	5	5.666	5	0.181	5	0.162
6	Sweden	1	5.717	6	0.184	6	0.168
7	Czech Republic	7	5.636	7	0.190	7	0.171
8	Hungary	15	5.546	8	0.204	9	0.192
9	Belgium	8	5.632	9	0.221	10	0.199
10	New Zealand	10	5.597	10	0.240	13	0.222
11	United Kingdom	11	5.595	11	0.262	19	0.318
12	Ireland	14	5.569	12	0.266	12	0.222
13	Switzerland	13	5.573	13	0.269	8	0.189
14	Canada	9	5.628	14	0.274	11	0.216
15	Italy	16	5.499	15	0.285	16	0.256
16	United States	12	5.587	16	0.289	14	0.227
17	Chile	19	5.355	17	0.302	15	0.230
18	Slovenia	17	5.446	18	0.314	17	0.260
19	Poland	18	5.415	19	0.333	18	0.312

Table 2a - Substitutability Rankings

	<i>WageDisp<sub>i</sub></i> <i>St Dev Res</i> Rank	<i>O*NET<sub>i</sub></i> <i>Contact<sub>i</sub></i> Rank
Lowest <i>Substit<sub>i</sub></i>		
Railroad rolling stock	1	60
Ship and boat building	2	40
Aircraft, aerospace products and parts	3	28
Engines, turbines, and power trans. equipment	4	42
Nonferrous metals, exc. aluminum	5	59
Highest <i>Substit<sub>i</sub></i>		
Leather tanning and products, except footwear	59	21
Seafood and other miscellaneous foods, n.e.c.	60	31
Apparel accessories and other apparel	61	2
Bakeries	62	32
Cut and sew apparel	63	1

Table 2b - Correlations of *WageDisp<sub>i</sub>* and *O\*NET<sub>i</sub>*

	<i>WageDisp<sub>i</sub></i>		<i>O*NET<sub>i</sub></i>			
	<i>StDev</i> <i>Mean</i>	<i>St DevRes</i>	<i>Contact<sub>i</sub></i>	<i>Communic<sub>i</sub></i>	<i>Impact<sub>i</sub></i>	<i>Teamwork<sub>i</sub></i>
<i>StDev</i> <i>Mean</i>	1					
<i>St Dev Res</i>	0.8497 <i>0.000</i>	1				
<i>Contact<sub>i</sub></i>	-0.2061 <i>0.1052</i>	-0.1756 <i>0.1688</i>	1			
<i>Communic<sub>i</sub></i>	-0.1414 <i>0.2689</i>	-0.0755 <i>0.5565</i>	0.5818 <i>0.000</i>	1		
<i>Impact<sub>i</sub></i>	-0.2414 <i>0.0567</i>	-0.097 <i>0.4496</i>	0.668 <i>0.000</i>	0.7467 <i>0.000</i>	1	
<i>Teamwork<sub>i</sub></i>	-0.1606 <i>0.2087</i>	-0.1666 <i>0.1919</i>	0.7943 <i>0.000</i>	0.614 <i>0.000</i>	0.7254 <i>0.000</i>	1

p-values in italics



Table 3 - Residual Wage Dispersion Rankings and Residual Score Dispersion

	(1)	(2)	(3)	(4)	(5)	(6)
Measure of Dispersion	St Dev	95-5 IPR	Gini MD	St Dev	95-5 IPR	Gini MD
$WageDisp_i \times SkillDisp_H$	0.079** (0.015)	0.082** (0.018)	0.081** (0.021)	0.079** (0.013)	0.078** (0.016)	0.074** (0.019)
Trade Barriers	No	No	No	Yes	Yes	Yes
Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer FE	Yes	Yes	Yes	No	No	No
Industry FE	Yes	Yes	Yes	No	No	No
Importer-Industry FE	No	No	No	Yes	Yes	Yes
Observations	58124	58124	58124	58124	58124	58124
R-squared	0.54	0.54	0.54	0.7	0.7	0.7

The dependent variable is the natural logarithm of exports from country  $H$  to country  $F$  in industry  $i$ . Standardized beta coefficients are reported. †, \* and \*\* indicate the coefficient is significant at the 10%, 5% and 1% levels. Bootstrap standard errors clustered by importer-exporter pair in parenthesis (50 replications).

Table 4 - O\*NET Rankings and Residual Score Dispersion (St Dev)

Measure of Complementarity	(1) $O^*NET_i =$ $Teamwork_i$	(2) $O^*NET_i =$ $Impact_i$	(3) $O^*NET_i =$ $Communic_i$	(4) $O^*NET_i =$ $Contact_i$
$O^*NET_i \times$ $SkillDisp_H$	-0.071** (0.017)	-0.071** (0.018)	-0.073** (0.02)	-0.063** (0.015)
Trade Barriers	Yes	Yes	Yes	Yes
Exporter FE	Yes	Yes	Yes	Yes
Imp-Ind FE	Yes	Yes	Yes	Yes
Observations	58124	58124	58124	58124
R-squared	0.70	0.70	0.70	0.70

Bootstrap standard errors clustered by importer-exporter pair in parenthesis (50 replications).

Table 5 - Selection and Other Controls

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline (St Dev)	HMR Selection	Heckscher Ohlin	Institution Controls	Top Coding	Without US
$WageDisp_i \times SkillDisp_H$	0.079** (0.013)	0.071** (0.016)	0.096** (0.024)	0.049* (0.02)	0.104** (0.025)	0.077** (0.014)
$KIntens_i \times KEndow_H$			0.266** (0.073)			
$SkillIntens_i \times SkillEndow_H$			0.122** (0.017)			
$WageDisp_i \times JudicQual_H$				-0.075** (0.027)		
$WageDisp_i \times LaborRigid_H$				-0.035* (0.014)		
$TopCode_i \times SkillDisp_H$					-0.069* (0.033)	
Trade Barriers	Yes	Yes	Yes	Yes	Yes	Yes
Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer-Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	58124	52455	41301	51166	52455	48129
R-squared	0.69	0.69	0.73	0.7	0.69	0.68

The dependent variable is the natural logarithm of exports from country  $H$  to country  $F$  in industry  $i$ . Standardized beta coefficients are reported. †, \* and \*\* indicate the coefficient is significant at the 10%, 5% and 1% levels. Bootstrap standard errors clustered by importer-exporter pair in parenthesis (50 replications). Column (6) is the same specification of column (2) excluding the observations involving US as exporter. The regression includes a polynomial in the probability to export, obtained from the first stage, which is significant and we do not report.

Table 6 - First Stages of Table 5

	(1)	(2)	(3)	(4)	(5)
	HMR Selection	Heckscher Ohlin	Institution Controls	Top Coding	Without US
$WageDisp_i \times SkillDisp_H$	0.02** (0.004)	0.043** (0.005)	0.022** (0.006)	0.042** (0.005)	0.019** (0.004)
$RegCosts_H \times RegCosts_F$	0.013** (0.004)	0.001 (0.006)	0.01* (0.005)	0.013** (0.004)	0.006 (0.003)
$RegDays_H \times RegDays_F$	0.014* (0.006)	0.02 (0.012)	0.013 (0.008)	0.014* (0.006)	0.012* (0.005)
$RegProc_H \times RegProc_F$	0.029** (0.009)	0.077** (0.018)	0.032** (0.011)	0.029** (0.009)	0.031** (0.011)
$KIntens_i \times KEndow_H$		0.023 <sup>†</sup> (0.012)			
$SkillIntens_i \times SkillEndow_H$		-0.013* (0.005)			
$WageDisp_i \times JudicQual_H$			0.018 <sup>†</sup> (0.01)		
$WageDisp_i \times LaborRigid_H$			-0.01* (0.004)		
$TopCode_i \times SkillDisp_H$				-0.054** (0.008)	
Trade Barriers	Yes	Yes	Yes	Yes	Yes
Exporter FE	Yes	Yes	Yes	Yes	Yes
Importer-Industry FE	Yes	Yes	Yes	Yes	Yes
Observations	132867	94794	125874	132867	124740
R-squared	0.58	0.59	0.58	0.58	0.58

Columns (1)-(5) report the first stage estimation results corresponding to Columns (2)-(6) of Table D. The dependent variable is a dummy that is one if exports from country  $H$  to country  $F$  in industry  $i$  are positive and zero otherwise. Standardized beta coefficients are reported.

<sup>†</sup>, \* and \*\* indicate the coefficient is significant at the 10%, 5% and 1% levels. Bootstrap standard errors clustered by importer-exporter pair in parenthesis (50 replications). The measure of dispersion employed is the standard deviation of residual wages and residual scores.

Column (6) is the same specification of column (2) excluding the observations involving US as exporter. All estimations were performed with a linear probability model.

Table 7 - Additional Variables

Variable	Obs	Mean	Std. Dev	Min	Max
Exports dummy	173565	0.335	0.472	0	1
Exports volume ( $X_{HFi}$ )	58124	7.866	2.204	0	17.906
Language	2755	0.193	0.395	0	1
Legal	2755	0.217	0.412	0	1
Religion	2755	0.196	0.257	0	0.973
Land Border	2755	0.019	0.135	0	1
Currency Union	2755	0.002	0.047	0	1
Distance	2755	4.136	0.806	0.882	5.661
FTA	2755	0.017	0.131	0	1
Colonial Ties	2755	0.022	0.146	0	1
Gatt / WTO	2755	1.489	0.578	0	2
Island	2755	0.291	0.494	0	2
Landlock	2755	0.309	0.509	0	2
$RegProc_F$	112	9.679	3.491	2	19
$RegDays_F$	112	49.402	38.593	2	203
$RegCosts_F$	112	90.065	165.785	0	1268.4
$RegProc_H$	19	5.947	2.818	2	10
$RegDays_H$	19	23.842	16.433	3	61
$RegCosts_H$	19	7.874	7.190	0	22.9
$SkillEndow_H$	14	-3.435	0.402	-4.522	-2.957
$JudicQual_H$	18	0.832	0.115	0.615	0.972
$LaborRigid_H$	19	0.473	0.155	0.205	0.667
$KEndow_H$	14	-0.530	0.662	-1.377	0.925
$SkillIntens_i$	61	0.381	0.116	0.166	0.757
$KIntens_i$	61	0.859	0.464	0.235	2.535
$TopCode_i$	63	0.009	0.005	0.004	0.030

# A Appendix - Proofs, derivations and extensions

## A.1 Conditions under which Property 1 holds

In this section we report analytical conditions which guarantee that Property 1 holds. We show that comparative advantage can be established for any distribution if we place bounds on the degree of complementarity  $\lambda$ . Moreover, we perform comparative statics assuming specific distributions of skills.

Our first approach yields a general result based on restrictions on the degree of complementarity and on the upper bound of the support of the skill distribution.<sup>46</sup>

**Proposition A-1** *Property 1 holds, i.e. a country  $c'$  with a more dispersed skill distribution than country  $c$  has a comparative advantage in sectors with lower complementarity (higher  $\lambda$ ) under the following sufficient conditions:*

(i) *Skill is bounded from above by  $a_{\max}$*

(ii) *The degree of complementarity is low enough:  $\lambda > \bar{\lambda}$  where  $\bar{\lambda}$  is defined by the following condition*

$$\log a_{\max} = \frac{2\bar{\lambda} - 1}{(1 - \bar{\lambda})\bar{\lambda}}$$

**Proof.** By definition of log-supermodularity we need to prove that, if  $g(a, c')$  is a mean-preserving spread of  $g(a, c)$  then:

$$\frac{\partial \log A(\lambda, c)}{\partial \lambda} \leq \frac{\partial \log A(\lambda, c')}{\partial \lambda}.$$

The partial derivative has the following expression:

$$\frac{\partial \log A(\lambda, c)}{\partial \lambda} = \frac{1}{\lambda} \frac{\int a^\lambda \log a g(a, c) da}{\int a^\lambda g(a, c) da} - \frac{1}{\lambda^2} \log \left( \int a^\lambda g(a, c) da \right) \quad (\text{A-1})$$

A mean-preserving spread of  $g(a, c)$  increases the second term of the right-hand side of (A-1) by definition, since  $a^\lambda$  is a concave function. A sufficient condition for the first term of (A-1) to increase with a mean-preserving spread in  $g(a, c)$  is that  $k(a) = a^\lambda \log a$  is a convex function which is verified if its second derivative with respect to  $a$  is positive for every value of  $a$ . i.e.  $\log a < \frac{2\lambda-1}{(1-\lambda)\lambda}$ . Since the right-hand side of this inequality is continuous and increasing in  $\lambda$ , it is equal to zero for  $\lambda = \frac{1}{2}$  and  $\lim_{\lambda \rightarrow 1} \frac{2\lambda-1}{(1-\lambda)\lambda} = \infty$  then, if  $a$  is bounded above by  $a_{\max}$ , then there exists a value  $\bar{\lambda} < 1$  such that  $\log a_{\max} = \frac{2\bar{\lambda}-1}{(1-\bar{\lambda})\bar{\lambda}}$ . If  $\lambda > \bar{\lambda}$  then  $\frac{\partial \log A(\lambda, c)}{\partial \lambda}$  increases with a mean preserving spread of  $g(a, c)$ . ■

In our second approach to studying Property 1 we relax the conditions on complementarity at the cost of concentrating on specific distributions. We can only consider continuous distributions that are characterized by at least two parameters (in order to be able to consider mean-preserving increases in dispersion) and are defined on a positive support.

---

<sup>46</sup>Imposing an upper bound on  $a$  is realistic because it means we do not admit the existence of infinitely productive workers.

**Proposition A-2** *If skills are distributed according to a Pareto or Log-normal distribution then, if country  $c$  and  $c'$  are characterized by skill distributions  $g(a, c)$  and  $g(a, c')$  such that  $g(a, c')$  has equal mean and higher variance than  $g(a, c)$  and if  $\lambda < \lambda'$  then Property 1 holds, i.e. country  $c'$  has a comparative advantage in  $\lambda'$ .*

- (i) *Pareto Distribution* - Under the assumption that skills follow a Pareto distribution with mean  $\mu$  and standard deviation  $\sigma$ ,  $A$  takes the following expression:<sup>47</sup>

$$A = \frac{\mu^2 + \sigma^2 - \sigma\sqrt{\mu^2 + \sigma^2}}{\mu} \left( \frac{\sigma + \sqrt{\mu^2 + \sigma^2}}{\sigma + \sqrt{\mu^2 + \sigma^2} - \lambda\sigma} \right)^{\frac{1}{\lambda}}.$$

Since  $A$  is twice differentiable in  $\sigma$  and  $\lambda$ , the result in Proposition 3 is equivalent to  $A$  being log-supermodular in  $\lambda$  and  $\sigma$ , that is  $\frac{\partial^2 \log A}{\partial \sigma \partial \lambda} > 0$ . The expression for the cross partial derivative is the following:

$$\frac{\partial^2 \log A}{\partial \sigma \partial \lambda} = \frac{\sigma \left( \sqrt{\mu^2 + \sigma^2} - \sigma \right)}{\sqrt{\mu^2 + \sigma^2} \left[ \sigma(1 - \lambda) + \sqrt{\mu^2 + \sigma^2} \right]} \quad (\text{A-2})$$

and  $\lambda < 1$  so  $A$  is log-supermodular in  $\lambda$  and  $\sigma$ .

- (ii) *Log-Normal Distribution* - If the distribution of skills  $a$  is lognormal on the support  $[0, \infty]$  with mean  $\mu$  and standard deviation  $\sigma$  then  $A$  takes the following form:

$$A = e^{\log \mu - \frac{1-\lambda}{2} \log \left( \frac{\sigma^2}{\mu^2} + 1 \right)}$$

It is easy to show that under this distribution,  $A$  is log-supermodular since the following expression is always positive:

$$\frac{\partial^2 \log A}{\partial \sigma \partial \lambda} = \frac{\sigma}{\mu^2 + \sigma^2}$$

While Proposition A-2 establishes an analytical result, we have also numerically computed the  $A$ 's for the following distributions: uniform, triangular, gamma, beta and inverse gaussian. For all

---

<sup>47</sup>The Pareto distribution is characterized by a shape parameter  $k$  and location parameter  $a_{\min}$ , i.e. the cumulative distribution of ability is given by  $G(a) = 1 - \left(\frac{a_{\min}}{a}\right)^k$  with  $a_{\min} > 0$  and  $k > 2$ . We could have written  $A$  as a function of those parameters:

$$A = a_{\min} \left( \frac{k}{k - \lambda} \right)^{\frac{1}{\lambda}}$$

Since we are interested in a mean-preserving increase in variance, we express the  $A$  as a function of  $\mu$  and  $\sigma$ , which are related to shape and location parameters according to the following equations:

$$\begin{aligned} a_{\min} &= \frac{\mu^2 + \sigma^2 - \sigma\sqrt{\mu^2 + \sigma^2}}{\mu} \\ k &= \frac{\sigma + \sqrt{\mu^2 + \sigma^2}}{\sigma} \end{aligned}$$

these distributions, and for a wide range of parameters, we cannot find a violation of the ranking in (3).<sup>48</sup>

## A.2 The Firm Problem

This section analyzes the problem of a representative Home firm in a given sector. Analogous expressions can be derived for a Foreign firm. Firms can sell in the domestic market or export, facing a transport cost. The transport cost  $\tau$  is of the iceberg type, so that firms have to ship  $\tau > 1$  units of good in order for one unit to arrive. We denote by  $x_{cc'}$  a variable  $x$  originating in market  $c$  and destined for market  $c'$ . We drop the sector index to simplify notation.

Total revenues of a firm in Home are given by:

$$r_H = B_H y_{HH}^{\frac{\sigma-1}{\sigma}} + B_F y_{HF}^{\frac{\sigma-1}{\sigma}} \tau^{\frac{1-\sigma}{\sigma}} \quad (\text{A-3})$$

where  $B_c = P_c Q_c^{\frac{1}{\sigma}}$  for  $c = H, F$ . For a profit-maximizing firm marginal revenues have to be equal across markets. Rearranging the equality of marginal revenue condition leads to the following:

$$\frac{y_{HH}}{y_{HF}} = \left( \frac{B_H}{B_F} \right)^{\sigma} \tau^{\sigma-1} \quad (\text{A-4})$$

From (A-4)  $y_{HH}$  can be expressed as a function of  $y_{HF}$  and replaced in (A-3) to find:

$$r_H = B_F y_{HF}^{-\frac{1}{\sigma}} \tau^{\frac{1-\sigma}{\sigma}} (y_{HH} + y_{HF}) \quad (\text{A-5})$$

From (A-5) and its analogous for  $y_{HH}$  we can find the two following equations:

$$y_{HH} = r_H^{-\sigma} B_H^{\sigma} (y_{HH} + y_{HF})^{\sigma} \quad (\text{A-6})$$

$$y_{HF} = r_H^{-\sigma} B_F^{\sigma} (y_{HH} + y_{HF})^{\sigma} \tau^{1-\sigma} \quad (\text{A-7})$$

Adding up (A-6) and (A-7) and rearranging them leads to the following expression for total revenues:

$$r_H = y_H^{\frac{\sigma-1}{\sigma}} \Gamma_H \quad (\text{A-8})$$

where  $y_H = y_{HH} + y_{HF}$ ,  $\Gamma_H = (B_H^{\sigma} + B_F^{\sigma} \tau^{1-\sigma})^{\frac{1}{\sigma}}$ . The firm must then simply choose the total amount of output to produce and therefore how many workers to employ. In this decision it takes into account how much workers are paid.

Because of the presence of search frictions, once workers are hired they are not interchangeable with outside workers and we assume that the firm and all workers employed engage in bargaining

---

<sup>48</sup> A violation of the ranking can be engineered using a result by Ross (1981). The intuition is the following. Ross (1981) shows that, if we adopt the Arrow-Debreu definition of risk aversion, then, starting from a given lottery, we might find the counterintuitive result that a more risk-averse individual is willing to pay less than a less risk-averse individual to avoid an increase in risk in the sense of a mean-preserving spread. We can view our  $A$  as the certainty equivalent of lottery  $g$  for an individual with Bernoulli utility  $u(a) = a^{\lambda}$ ,  $0 < \lambda < 1$ . Individuals with lower  $\lambda$  are more risk averse in the Arrow-Pratt sense. In our case we can show, using the example proposed by Ross (1981) that, with a mean-preserving spread, the certainty equivalent of a more risk averse individual drops proportionately by less than for a less risk averse individual. Details are available from the authors.



to share the surplus created. We assume that the intra-firm bargaining is of the type described by Stole and Zwiebel (1996), with the workers having unemployment as outside option, which we assume yields a payoff of zero. Stole and Zwiebel show that the bargaining solution yields payoffs that correspond to the Shapley value. See section A.6 for a detailed derivation. The bargaining outcome for a firm with revenues  $r$  is given by  $sr$ , where:

$$s = \frac{\sigma\lambda}{\sigma(1+\lambda) - 1}. \quad (\text{A-9})$$

Given the expression for total revenues in (A-8), the firm static problem<sup>49</sup> reduces to choosing how many workers to hire ( $h$ ) to maximize profits  $\pi$ :

$$\max_h \pi = s \left[ A(\lambda, H) h^{\frac{1}{\lambda}} \right]^{\frac{\sigma-1}{\sigma}} \Gamma_H - bh - f. \quad (\text{A-10})$$

This is a concave problem because of the restriction placed on  $\lambda$  in (2). The first order condition of problem (A-10) can be written as a function of revenues  $r_H$  as follows:

$$s \frac{\sigma-1}{\sigma\lambda b} r_H = h_H$$

This first order condition, together with the zero profit condition deriving from free entry:

$$sr_H - bh_H - f = 0$$

delivers total revenues and employment:

$$\begin{aligned} r_H &= \frac{f\sigma\lambda}{s(\sigma\lambda - \sigma + 1)} \\ h_H &= \frac{f(\sigma-1)}{b(\sigma\lambda - \sigma + 1)} \end{aligned}$$

Given the production function, the expression for total output produced by a Home firm follows:

$$y_H = A(\lambda, H) \phi \quad (\text{A-11})$$

where  $\phi(\lambda) = \left[ \frac{f(\sigma-1)}{b(1+(\lambda-1)\sigma)} \right]^{\frac{1}{\lambda}}$ . Intuitively, output is increasing in productivity  $A$ , the size of the fixed cost  $f$ , and the elasticity of demand  $\sigma$ , while it decreases with the hiring cost  $b$ .<sup>50</sup> We assume that differences in productivity between Home and Foreign firms in a given sector are not too large, that is:

$$\frac{1}{\tau} \leq \frac{A(\lambda, H)}{A(\lambda, F)} \leq \tau \quad \forall \lambda \quad (\text{A-12})$$

<sup>49</sup>For a dynamic extension of this type of framework see Helpman and Itskhoki (2009b).

<sup>50</sup>The hiring cost depends on tightness of the labor market  $x$ , and is assumed to take the same form as in Helpman and Itskhoki (2009a) and Helpman et al. (2008a):  $b = \delta_0 x^{\delta_1}$ . We refer to these papers for a discussion. We similarly obtain that in equilibrium the hiring cost is constant across sectors, i.e.  $b = \delta_0^{\frac{1}{1+\delta_1}}$ .

otherwise the amount produced is zero. Under condition (A-12) we can derive how much output is produced for the domestic and export market. We employ (A-6) and (A-7) and their analogues for the Foreign firm to find the relative output of firms *selling* in the same market:

$$\frac{y_{HH}}{y_{FH}} = \frac{r_H^{-\sigma} (y_{HH} + y_{HF})^\sigma}{r_F^{-\sigma} (y_{FF} + y_{FH})^\sigma \tau^{1-\sigma}} \quad (\text{A-13})$$

$$\frac{y_{FF}}{y_{HF}} = \frac{r_F^{-\sigma} (y_{FF} + y_{FH})^\sigma}{r_H^{-\sigma} (y_{HH} + y_{HF})^\sigma \tau^{1-\sigma}} \quad (\text{A-14})$$

The expressions above can be simplified using the fact that total revenues are constant in a given sector:  $r_H = r_F = r$ . Together with (A-11) and its foreign equivalent, (A-13) and (A-14) deliver the amount of output sold by a Foreign and a Home firm in every market. The amounts of output sold in the two markets by a Home firm are given by:

$$y_{HH} = \frac{\phi A(\lambda, H)}{1 - \rho^2} \left[ 1 - \rho \left( \frac{A(\lambda, H)}{A(\lambda, F)} \right)^{\sigma-1} \right] \quad (\text{A-15})$$

$$y_{HF} = \frac{\phi \rho A(\lambda, H)}{1 - \rho^2} \left[ \left( \frac{A(\lambda, H)}{A(\lambda, F)} \right)^{\sigma-1} - \rho \right] \quad (\text{A-16})$$

where  $\rho = \tau^{1-\sigma}$ , while the corresponding Foreign firm expressions are:

$$y_{FF} = \frac{\phi A(\lambda, F)}{1 - \rho^2} \left[ 1 - \rho \left( \frac{A(\lambda, H)}{A(\lambda, F)} \right)^{1-\sigma} \right] \quad (\text{A-17})$$

$$y_{FH} = \frac{\phi \rho A(\lambda, F)}{1 - \rho^2} \left[ \left( \frac{A(\lambda, F)}{A(\lambda, H)} \right)^{\sigma-1} - \rho \right] \quad (\text{A-18})$$

We focus attention here on the relative revenues (i.e. value of output sold) of a Home and Foreign firm in a *given* market, for example Foreign. We derive relative revenues by expressing it first as a function of relative output  $\frac{r_{HF}}{r_{FF}} = \left( \frac{y_{HF}}{y_{FF}} \right)^{\frac{\sigma-1}{\sigma}} \tau^{\frac{1-\sigma}{\sigma}}$ , and then replacing the expressions for  $y_{HF}$  and  $y_{FF}$ :

$$\frac{r_{HF}}{r_{FF}} = \left( \frac{A(\lambda, H)}{\tau A(\lambda, F)} \right)^{\sigma-1} \quad (\text{A-19})$$

Intuitively, relative revenues increase in relative productivity, as predicted by comparative advantage. As standard with iso-elastic demand, the producer price is constant across markets and for a Home firm is equal to  $p_H = \frac{\gamma}{\phi A_H}$  where  $\gamma(\lambda) = f^{\frac{\sigma\lambda + \sigma - 1}{\sigma\lambda - \sigma + 1}}$ . The consumer price in the export market is the producer price multiplied by  $\tau$ :

$$p_{HF} = \frac{\gamma\tau}{\phi A_H} \quad (\text{A-20})$$

### A.3 Derivation of the Mass of Firms

In the previous section we derived the amount of output sold by each firm in the domestic and export market. In order to determine trade flows we need to calculate the equilibrium mass of firms

for country  $c$  and sector  $\lambda$ ,  $M_c(\lambda)$ . Having determined the revenues of a firm in each market, the mass of firms in each country has to be such that, total expenditure on good  $\lambda$  in a given country is equal to total revenues accruing to all firms operating in that market. The two equations below express these equilibrium conditions for sector  $\lambda$ :

$$\alpha(\lambda) L_H = M_H(\lambda) r_{HH}(\lambda) + M_F(\lambda) r_{FH}(\lambda) \quad (\text{A-21})$$

$$\alpha(\lambda) L_F = M_F(\lambda) r_{FF}(\lambda) + M_H(\lambda) r_{HF}(\lambda) \quad (\text{A-22})$$

It is convenient to rewrite conditions (A-21) and (A-22) as a function of output, rather than of revenues:

$$\alpha L^H = M_H \frac{\gamma}{\phi A(\lambda, H)} y_{HH} + M_F \frac{\gamma}{\phi A(\lambda, F)} y_{FH},$$

$$\alpha L^F = M_F \frac{\gamma}{\phi A(\lambda, F)} y_{FF} + M_H \frac{\gamma}{\phi A(\lambda, H)} y_{HF}.$$

The solution to this linear system is given by the following expressions for  $M_H$  and  $M_F$ :

$$M_H = A(\lambda, H) \frac{\alpha \phi (L_H y_{FF} - L_F y_{FH})}{\gamma (y_{FF} y_{HH} - y_{HF} y_{FH})}, \quad (\text{A-23})$$

$$M_F = A(\lambda, F) \frac{\alpha \phi (L_F y_{HH} - L_H y_{HF})}{\gamma (y_{FF} y_{HH} - y_{HF} y_{FH})}. \quad (\text{A-24})$$

First, we show that the denominator of  $M_H$  and  $M_F$  is always positive. Define Home productivity advantage  $z(\lambda) = \frac{A(\lambda, H)}{A(\lambda, F)}$ . The denominator is positive if and only if  $\frac{y_{HH}}{y_{FH}} > \frac{y_{HF}}{y_{FF}}$ , a condition we can rewrite as  $z^\sigma \frac{1}{\rho} > z^\sigma \rho$  and that is always satisfied since  $\rho < 1$ .

We remark that, similarly to other models of monopolistic competition with trade costs (Helpman and Krugman, 1985), the presence of a home-market effect requires that we restrict the degree of asymmetry in country sizes to prevent all firms from locating in one country. Define relative population in Home as  $\eta \equiv \frac{L_H}{L_F}$ . The mass of Home firms  $M_H$  is positive if and only if  $L_H y_{FF} - L_F y_{FH} > 0$ . This condition places a lower bound on the relative population, since  $M_H > 0$  if and only if:

$$\eta > \frac{\rho \left( \frac{1}{z^{\sigma-1}} - \rho \right)}{1 - \frac{\rho}{z^{\sigma-1}}} = \eta_{low} \quad (\text{A-25})$$

Equivalently,  $M_F$  is positive if and only if  $L_F y_{HH} - L_H y_{HF} > 0$ , a condition that places an upper bound on the relative population  $\eta$ , i.e.  $M_F > 0$  if and only if:

$$\eta < \frac{1 - \rho z^{\sigma-1}}{\rho (z^{\sigma-1} - \rho)} = \eta_{up} \quad (\text{A-26})$$

Both  $\eta_{low}$  and  $\eta_{up}$  are positive under the condition that we imposed in order to guarantee that a positive amount of output is produced for every market:  $\rho < z^{\sigma-1} < \frac{1}{\rho}$ . We impose throughout the restrictions that  $\eta_{low} < \eta < \eta_{up}$ . If the condition is violated for some industries, we expect to observe no production and no exports.<sup>51</sup> If the condition is satisfied, then the following proposition

<sup>51</sup> As equations (A-25) and (A-26) establish, the conditions for a positive mass of firms depend on size, but also on

establishes a link between comparative advantage and equilibrium entry.

**Proposition A-3** *Under the condition that country sizes are sufficiently similar, i.e.  $\eta_{low} < \eta < \eta_{up}$ , the equilibrium mass of firms in country  $H$  relative to country  $F$  in sector  $\lambda'$  is higher than in sector  $\lambda$  if and only if country  $H$  has a comparative advantage in sector  $\lambda'$ , i.e.*

$$\frac{A(\lambda, H)}{A(\lambda, F)} < \frac{A(\lambda', H)}{A(\lambda', F)} \iff \frac{M_H(\lambda)}{M_F(\lambda)} < \frac{M_H(\lambda')}{M_F(\lambda')}$$

**Proof.** We define the mass of Home relative to Foreign firms in sector  $\lambda$  as  $m(\lambda) \equiv \frac{M_H}{M_F}$ . We investigate how  $m$  changes with  $z$ , assuming that we are operating in the parameter space where  $\eta_{low} < \eta < \eta_{up}$ . We rewrite the relative mass of firms, using (A-23), (A-24), the expressions for Home firm outputs, (A-15) and (A-16), and the corresponding expressions for the Foreign firm:

$$m = \frac{z^{1-\sigma} (1 + \eta) \rho - (\eta + \rho^2)}{z^{\sigma-1} (1 + \eta) \rho - (1 + \eta \rho^2)}$$

The first derivative of  $m$  with respect to  $z$  takes the following form:

$$\frac{\partial m}{\partial z} = \frac{1}{z^{\sigma+2}} (\sigma - 1) (1 + \eta) \frac{-2z^{1+\sigma} (1 + \eta) \rho + z^{2\sigma} (\eta + \rho^2) + z^2 (1 + \eta \rho^2)}{(\rho z^{\sigma-1} - \eta \rho^2 + \eta \rho z^{\sigma-1} - 1)^2}$$

This derivative is positive if the numerator is positive and the denominator can be divided in two parts, which we show are both positive. The first part, denoted by  $\psi_1$  is:

$$\psi_1 = -z^{1+\sigma} (1 + \eta) \rho + z^{2\sigma} (\eta + \rho^2),$$

while the second part denoted by  $\psi_2$  is:

$$\psi_2 = -z^{1+\sigma} (1 + \eta) \rho + z^2 (1 + \eta \rho^2).$$

It is straightforward to show that  $\psi_1 > 0$  if and only if  $\eta > \eta_{low}$  and that  $\psi_2 > 0$  if and only if  $\eta < \eta_{up}$ , conditions we have imposed throughout. ■

#### A.4 Proof of Proposition 1

Since trade flows are completely determined by the amount sold in the export market by each firm and by the number of firms. We denote the value of total sales by firms from country  $c$  in market  $c'$ , as  $X_{cc'}$ . Relative total sales of good  $\lambda$  by Home and Foreign firms in a given market, for example Foreign, are then equal to:

$$\frac{X_{HF}(\lambda)}{X_{FF}(\lambda)} = \frac{r_{HF}(\lambda) M_H(\lambda)}{r_{FF}(\lambda) M_F(\lambda)} \tag{A-27}$$

---

comparative advantage. If a country is relatively more productive it can afford to be smaller in size and still have a positive mass of firms. In this sense our model also predicts an extensive margin of trade (whether we observe or not trade between two countries) based on comparative advantage, albeit a very stark one. Differently from models with heterogeneous firms, e.g. Helpman et al. (2008c), in this setup the assumption of identical firms implies that either firms exist and export or they do neither.

The result follows directly since we have proven that both components of relative sales (A-27), relative revenues per firm  $\frac{r_{FH}(\lambda)}{r_{HH}(\lambda)}$  and relative mass of firms  $\frac{M_F(\lambda)}{M_H(\lambda)}$  are increasing in relative productivity  $\frac{A(\lambda, F)}{A(\lambda, H)}$  (see (A-19) and Proposition A-3) and relative productivity depends the degree of complementarity  $\lambda$  (proxied by the dispersion of wages according to Proposition A-5) and the dispersion of skills according to the discussion in section 2.4.

## A.5 Multi-Country Model

The goal of this section is to generalize the model to many countries and provide the conditions under which the main result of the two-country model holds, i.e. countries with relatively higher dispersion of skills have a comparative advantage, and therefore export relatively more, in sectors where the dispersion of wages is higher.

Without loss of generality we consider three countries, so that  $c \in \{H, F, G\}$ . Following HMR, we allow transport costs to be country-pair specific and asymmetric, i.e.  $\tau_{HF} \neq \tau_{FH}$ . We fix as destination market country  $F$  and express the value of exports of good  $\lambda$  by country  $H$  relative to country  $G$  as follows:

$$\frac{X_{HF}(\lambda)}{X_{GF}(\lambda)} = \frac{r_{HF}(\lambda) M_H(\lambda)}{r_{GF}(\lambda) M_G(\lambda)}$$

While the determination of relative revenues of individual firms  $r_{HF}/r_{GF}$  is straightforward, the equilibrium mass of firms can be computed, but not easily characterized, with more than two asymmetric countries. This is a known problem in the home-market effect literature.<sup>52</sup> Therefore in the following proposition we limit ourselves to imposing that the relative mass of firms be non-decreasing in relative productivity. This is reasonable if we believe that, in equilibrium, entry is relatively higher in sectors where a country has a comparative advantage.

**Proposition A-4** *Under Property 1, if the relative mass of firms  $\frac{M_H(\lambda)}{M_G(\lambda)}$  is non-decreasing in relative productivity  $\frac{A(\lambda, H)}{A(\lambda, G)}$  then a country with relatively higher dispersion of skills has a comparative advantage, and therefore exports relatively more to any destination, in sectors with higher degree of substitutability  $\lambda$ .*

**Proof.** Since the derivation is analogous to the two-country case we simply report the expression relative revenues:<sup>53</sup>

$$\frac{r_{HF}}{r_{GF}} = \left( \frac{A(\lambda, H)}{A(\lambda, G)} \right)^{\sigma-1} \left( \frac{\tau_{HF}}{\tau_{GF}} \right)^{1-\sigma}.$$

It follows that, if the relative mass of firms is non-decreasing in relative productivity, relative exports are higher in comparative advantage sectors, similarly to the two-country case in Proposition 1. ■

## A.6 Derivation of the Shapley Value

In this section we provide details on how to derive the share of revenues accruing to the firm and the wages paid to workers. Stole and Zwiebel (1996) have proved the equivalence of their bargaining

<sup>52</sup>Behrens et al. (2009) show that the home-market effect intuition does not easily generalize to the case of more than two countries. Our case of multiple countries with productivity differences further complicates the problem and is beyond the scope of this paper.

<sup>53</sup>Details are available from the authors upon request.

solution to the Shapley value of the corresponding cooperative game not only for the case of identical workers, but also for the case of heterogeneous workers,<sup>54</sup> therefore we calculate the Shapley value directly.<sup>55</sup> The Shapley value of the firm is heuristically derived as its marginal contribution averaged over all possible orderings of employees and the firm itself. The case of heterogeneous employees is easy to handle under our assumption of a continuum of workers because no matter how the firm is ordered, it is preceded by a mass of workers whose skill distribution mirrors the overall skill distribution in the workers population, so the only variable we have to keep track of is the mass of workers preceding the firm, define it  $n$ , which varies from zero to  $h$ . As discussed in Acemoglu et al. (2007), since the firm is an essential input its marginal contribution is equal to revenues when  $n$  workers are employed in production  $r_H(n) = \left(n^{\frac{1}{\lambda}} A\right)^{\frac{\sigma-1}{\sigma}} \Gamma_H$ . The Shapley value of the firm  $S_{firm}$  is therefore:

$$S_{firm} = \int_0^h \frac{1}{h} \left(n^{\frac{1}{\lambda}} A\right)^{\frac{\sigma-1}{\sigma}} \Gamma_H dn = sr_H$$

where  $s$  is defined by (A-9). As discussed in Acemoglu et al. (2007) the share of revenues accruing to the firm depends on the curvature of the revenue function, due to characteristics of the demand function ( $\sigma$ ) and the production function ( $\lambda$ ).

In a similar fashion we calculate the Shapley value of a worker of skill  $a$ , by averaging its marginal contribution across all possible orderings. When a mass  $n$  of workers is employed, revenues of the firm are:

$$r(n) = \Gamma_H \left[ \int_a a^\lambda n(a, c) da \right]^{\frac{\sigma-1}{\sigma\lambda}}$$

where  $n(a, c) = ng(a, c)$ . The marginal contribution of a worker of skills  $a$  is given by the marginal revenue from an increase in the mass of workers of skill  $a$ ,  $n(a, c)$ , conditional on the firm being ordered before the worker (otherwise the marginal contribution is null):

$$\frac{\partial r(n)}{\partial n(a)} = \Gamma_H \frac{\sigma-1}{\sigma\lambda} \left[ \int_a n\tilde{g}(a) a^\lambda da \right]^{\frac{\sigma-1}{\sigma\lambda}-1} a^\lambda$$

The Shapley value and wage of worker of skill  $a$  in industry  $\lambda$  is:

$$w(a, \lambda) = \frac{1}{h} \int_0^h \frac{n}{h} \frac{\partial r(n)}{\partial n(a)} dn = \Gamma_H A(\lambda, c)^{\frac{\sigma-1-\sigma\lambda}{\sigma}} \frac{\sigma-1}{\sigma-1+\sigma\lambda} h^{\frac{\sigma-1}{\sigma\lambda}-1} a^\lambda$$

Since the average wage also differs across sectors, we normalize wages by the average wage in the sector  $E[w(a, \lambda)]$ . The normalized wage is denoted by  $\tilde{w}(a, \lambda) = \frac{w(a, \lambda)}{E[w(a, \lambda)]}$  and takes the following form:

$$\tilde{w}(a, \lambda) = \frac{a^\lambda}{E(a^\lambda)}$$

---

<sup>54</sup>See their Theorems 8 and 9, p. 393.

<sup>55</sup>The analogous of the Shapley value for a continuum of players is derived in Aumann and Shapley (1974).

## A.7 Wage Distribution and Complementarity

In this section we take the theoretical model as a guide to finding a proxy for the degree of complementarity, which is not directly observable and for which we have no available estimates. This section establishes a one-to-one link between the degree of complementarity and the dispersion of wages in sector  $\lambda$ , which can be measured in the data.

As discussed above, we assume that at the bargaining and production stage workers' skills are revealed, so that workers of different skills receive different wages as a result of intra-firm bargaining. Although the assumption that skill is perfectly revealed only at the production and bargaining stage is stark, we believe it captures some realistic features of the hiring process, where workers' skills in particular tasks are difficult to assess until they start working. Moreover, even if skills were partially revealed at the production stage, as long as the portion revealed were constant across sectors, this would not substantially change the implications we are about to discuss.

The previous section shows the calculation of the Shapley value for a worker of skill  $a$ . Since the average wage also differs across sectors, we normalize the wage of a worker of skill  $a$  in sector  $\lambda$  by the average wage in the sector. The normalized wage is  $\tilde{w}(a, \lambda) = \frac{a^\lambda}{E(a^\lambda)}$ , which reflects the marginal product of a worker of skill  $a$  when added to the production team and depends on  $\lambda$ . The higher the substitutability across workers the larger the marginal product of a worker with high skills. In contrast, if  $\lambda$  is low, i.e. complementarity is high, a worker of high skills has a relatively lower marginal product because her skills are very different from the average skills of her team-mates. An implication of this wage structure is that workers with identical skills, but employed in different sectors, generally receive different wages, as returns to skills vary across industries.<sup>56</sup>

Keeping in mind that the distribution of skills is the same in every industry, the distribution of wages within a sector depends, in our framework, exclusively on technological factors that determine the marginal product of workers with different skills. It therefore does not reflect compositional differences across sectors. The following proposition establishes that there is a one-to-one correspondence between the dispersion of wages and the degree of complementarity.

**Proposition A-5** *For any non-degenerate distribution of skills  $g(a, c)$ , the following three measures of dispersion of sectoral wages are strictly increasing in the degree of substitutability of workers' skills,  $\lambda$ : (i) the Coefficient of Variation; (ii) the Gini Coefficient and (iii) the Inter-Percentile Ratio<sup>57</sup>*

**Proof.** We consider three measures of wage dispersion: ■

- (i) the Coefficient of Variation of wages  $w(a, \lambda)$ , directly related to the variance of the normalized wage  $\tilde{w}(a, \lambda)$ ,  $Var(\tilde{w}(a, \lambda))$ , which is given by:

$$Var(\tilde{w}(a, \lambda)) = \frac{E(a^{2\lambda})}{E(a^\lambda)^2} - 1, \quad (\text{A-28})$$

<sup>56</sup>The point is made by Heckman and Scheinkman (1987), who show that returns to unobservable characteristics are different across sectors.

<sup>57</sup>The Interpercentile-Ratio,  $IPR_{kj}$ , is defined as  $IPR_{kj} = \frac{w_k}{w_j}$ , where  $w_k$  ( $w_j$ ) is the wage of the worker at the  $k^{th}$  ( $j^{th}$ ) percentile of the sectoral wage distribution and  $j < k$ .

- (ii) the Gini Coefficient, defined with respect to the Lorenz Curve for normalized wages at the sector level  $\Lambda(w, \lambda)$ ,
- (iii) the Inter-Percentile Ratio  $IPR_{kj}$  defined as:

$$IPR_{kj} = \frac{w_k}{w_j},$$

where  $w_k$  ( $w_j$ ) is the wage of the worker at the  $k^{th}$  ( $j^{th}$ ) percentile of the sectoral wage distribution and  $j < k$ .

(i) *Coefficient of Variation*

Since the variance of normalized wages is equal to the square of the coefficient of variation we prove the result for the former. We start by rewriting (A-28) in an explicit form, dropping the country index  $c$  to simplify notation:

$$Var(\tilde{w}(a, \lambda)) = \frac{\int a^{2\lambda} \tilde{g}(a) da}{\left(\int a^\lambda \tilde{g}(a) da\right)^2} - 1 \quad (\text{A-29})$$

The derivative of (A-29) with respect to  $\lambda$  is non-negative if and only if the following inequality is satisfied:

$$\left(\int_0^\infty a^{2\lambda} \log a \tilde{g}(a) da\right) \left(\int_0^\infty a^\lambda \tilde{g}(a) da\right) \geq \left(\int_0^\infty a^\lambda \log a \tilde{g}(a) da\right) \left(\int_0^\infty a^{2\lambda} \tilde{g}(a) da\right) \quad (\text{A-30})$$

The left-hand side of (A-30), which we denote by  $\Phi_L$  can be rewritten as:

$$\Phi_L = \int_0^\infty \int_0^\infty a^{2\lambda} \log a \tilde{g}(a) b^\lambda \tilde{g}(b) da db$$

We can divide the region of integration in two parts, delimited by the 45 degree line in the plane  $[0, \infty] \times [0, \infty]$ . It follows that  $\Phi_L$  can be rewritten as:

$$\Phi_L = \int_0^\infty \left(\int_0^a b^\lambda \tilde{g}(b) db\right) a^{2\lambda} \log a \tilde{g}(a) da + \int_0^\infty \left(\int_a^\infty b^\lambda \tilde{g}(b) db\right) a^{2\lambda} \log a \tilde{g}(a) da \quad (\text{A-31})$$

We change the order of integration in the second component of  $\Phi_L$  so that we can rewrite (A-31) it as:

$$\Phi_L = \int_0^\infty \left(\int_0^a b^\lambda \tilde{g}(b) db\right) a^{2\lambda} \log a \tilde{g}(a) da + \int_0^\infty \left(\int_0^b a^{2\lambda} \log a \tilde{g}(a) da\right) b^\lambda \tilde{g}(b) db \quad (\text{A-32})$$

Finally, a change of variable in the second component of (A-32) allows us to express  $\Phi_L$  as:

$$\Phi_L = \int_0^\infty \left(\int_0^a b^\lambda \tilde{g}(b) db\right) a^{2\lambda} \log a \tilde{g}(a) da + \int_0^\infty \left(\int_0^a b^{2\lambda} \log b \tilde{g}(b) db\right) a^\lambda \tilde{g}(a) da$$

If the same decomposition is performed on the right-hand side of (A-30) we can rewrite the



inequality as follows:

$$\int_0^\infty \left( \int_0^a a^\lambda b^\lambda \left[ (a^\lambda - b^\lambda) (\log a - \log b) \right] \tilde{g}(b) \tilde{g}(a) db \right) da \geq 0$$

which is always satisfied since  $(a^\lambda - b^\lambda) (\log a - \log b) \geq 0$ .

(ii) *Gini Coefficient*

We proceed by deriving the Lorenz Curve for sectoral normalized wages and showing that increasing  $\lambda$  produces a downward shift in the curve at all points. This is a sufficient condition for the Gini coefficient to increase with an increase in  $\lambda$ . The Lorenz Curve  $\Lambda(w, \lambda)$  of normalized wages in sector  $\lambda$  is given by the following expression:

$$\Lambda(w, \lambda) = \frac{\int_0^w a^\lambda \tilde{g}(a) da}{\int_0^\infty a^\lambda \tilde{g}(a) da}$$

The first derivative with respect to  $\lambda$ ,  $\frac{\partial \Lambda(w, \lambda)}{\partial \lambda}$  is non-positive if and only if the following condition is satisfied  $\forall w$ :

$$\left( \int_0^w a^\lambda \log a \tilde{g}(a) da \right) \left( \int_0^\infty b^\lambda \tilde{g}(b) db \right) \leq \left( \int_0^w a^\lambda \tilde{g}(a) da \right) \left( \int_0^\infty b^\lambda \log b \tilde{g}(b) da \right)$$

The region of integration can be divided into two part on both sides of the inequality, so that the inequality can be rewritten as follows:

$$\begin{aligned} & \left( \int_0^w \left( \int_0^w b^\lambda \tilde{g}(b) db \right) a^\lambda \log a \tilde{g}(a) da \right) + \int_0^w \left( \int_w^\infty b^\lambda \tilde{g}(b) db \right) a^\lambda \log a \tilde{g}(a) da \leq \\ & \left( \int_0^w \left( \int_0^w b^\lambda \log b \tilde{g}(b) db \right) a^\lambda \tilde{g}(a) da \right) + \int_0^w \left( \int_w^\infty b^\lambda \log b \tilde{g}(b) db \right) a^\lambda \tilde{g}(a) da \end{aligned}$$

Simplifying and factorizing leads to the following inequality:

$$\int_0^w \int_w^\infty b^\lambda a^\lambda (\log a - \log b) \tilde{g}(b) \tilde{g}(a) db da \leq 0$$

which is always satisfied since the range of integration of  $a$  is  $[0, w]$  while the range of integration of  $b$  is  $[w, \infty]$ .

(iii) *Inter-Percentile Ratio*

It is straightforward to show that  $IPR_{kj}$  increases with  $\lambda$  since for any percentile the ratio of wages is given by:

$$IPR_{kj} = \left( \frac{a_k}{a_j} \right)^\lambda$$

where  $a_k(a_j)$  is the skill of the worker at the  $k^{th}(j^{th})$  percentile.

## B Appendix - Main variables

### B.1 Measuring Skill Dispersion

The IALS microdata used for this paper was compiled by Statistics Canada using the original data sets collected between 1994 and 1998 in each of the participating countries. Tuijnman (2000) describes the three dimensions of literacy used to approximate skills. *Prose literacy* represents the knowledge and skills needed to understand and use information from texts including editorials, news stories, brochures and instruction manuals. *Document literacy* represents the knowledge and skills required to locate and use information contained in various formats, including job applications, payroll forms, transportation schedules, maps, tables and charts. *Quantitative literacy* represents the knowledge and skills required to apply arithmetic operations, either alone or sequentially, to numbers embedded in printed materials, such as balancing a chequebook, figuring out a tip, completing an order form or determining the amount of interest on a loan from an advertisement.

We employ the logarithm of scores (in conjunction with the log of wages) because the standard deviation of the logarithm of a random variable is scale invariant. When extracting residual scores in equation (9), using log-scores on the left-hand side is consistent with the common practice of obtaining residual wages from a regression of log-wages, as in equation (10). The results of the empirical analysis are qualitatively similar if we use levels instead of logs.

Only individuals participating in the labor market are included in the estimation of equation (10). These individuals were either: (i) employed or unemployed at some time in the 12 months previous to the survey or (ii) not searching for a job due to skill upgrading (school or work programs) or a temporary disability.

The right-hand side vector  $X_{kH}$  in equation (9) includes a number of observable individual characteristics. Education is among them: we include indicators for 7 levels of educational attainment as defined by the International Standard Classification of Education (ISCED). The levels considered in IALS are: ISCED 0 Education preceding the first level; ISCED 1 Education at the first level; ISCED 2 Education at the second level, first stage; ISCED 3 Education at the second level, second stage; ISCED 5 Education at the third level, first stage (leads to an award not equivalent to a first university degree); ISCED 6 Education at the third level, first stage (leads to a first university degree or equivalent); ISCED 7 Education at the third level, second stage (leads to a postgraduate university degree or equivalent); ISCED 9 Education not definable by level. The vector  $X_{kH}$  also includes 5 age intervals 16-25, 26-35, 36-45, 46-55 and 56-65, gender, immigrant status and participation in adult education or training programs 12 months prior to the survey date. The latter filters out the effect of skill upgrading on individual residual scores. As explained in section 3.4, this is an important issue for the identification of the effect of skill dispersion on trade flows as (unobserved) trade shocks might have an impact on aggregate skill dispersion by changing incentives for skill upgrading at the individual level. Residual scores  $\widehat{\epsilon}_{kH}$  are constructed as  $\widehat{\epsilon}_{kH} = \log(s_{kH}) - X_{kH}\widehat{\beta}_H$ , where  $\widehat{\beta}_H$  is estimated by OLS.

As a result of focusing on log-scores, the scale of measurement of IALS scores does not affect the standard deviation of  $\widehat{\epsilon}_{kH}$  or  $\log(s_{kH})$ . Also note that, since  $X_{kH}$  in (9) contains a constant, the distribution of  $\widehat{\epsilon}_{kH}$  has the same (zero) mean in each country. For this reason, we do not normalize the standard deviation (or any inter-percentile range) by the mean in order to make cross-country

comparisons of residual scores dispersion.

## B.2 Measuring Wage Dispersion

Wage inequality measures are computed from a sample of full-time manufacturing workers, 16-65 years old, not living in group quarters, reporting positive wages and industry affiliation.<sup>58</sup> Following Dahl (2002), individuals were considered as ‘full-time employed’ if in 1999 they: (i) were not enrolled full time in school, (ii) worked for pay for at least ten weeks, and (iii) earned an annual salary of at least 2,000 dollars. We focus on the log of weekly wages, calculated by dividing wage and salary income by annual weeks worked. We use weekly wages as opposed to hourly wages, because it requires fewer assumptions to calculate it. In the 2000 Census, hours worked are reported as ‘usual hours’. Using this variable to convert weekly wages into hourly wages would almost certainly result in the introduction of a source of measurement error. Incomes for top-coded values are imputed by multiplying the top code value (\$175,000) by 1.5.<sup>59</sup>

In equation (10), vector  $Z_{ki}$  includes indicators for 4 categories of educational attainment,<sup>60</sup> a quartic polynomial in age, race and gender dummies (plus their interaction), Hispanic and immigrant dummies (plus their interaction) and state of residence dummies. Residual wages are constructed as  $\widehat{\xi}_{ki} = \log(w_{ki}) - Z_{ki}\widehat{\beta}_i$ , where  $\widehat{\beta}_i$  is estimated by OLS.

Correcting for self-selection into industries is important in estimating equation (10), as the assumption that workers do not selectively search for jobs according to comparative advantage or unobservable tastes is relevant for Proposition A-5. In the presence of self-selection the distribution of residual wages in any given industry would reflect not only the degree of skill substitutability in production but also workers’ skill composition. For this reason, we use a selection estimator proposed by Dahl (2002). In equation (10), correcting for self-selection is complicated by the fact that individuals could choose to search for a job in any of the 63 industries of the manufacturing sector, potentially making the error mean, i.e.  $E(\xi_{ki} | k \text{ is observed in } i)$ , a function of the characteristics of all the alternatives. In this case, Dahl (2002) argues that under a specific sufficiency assumption,<sup>61</sup> the error mean is only a function of the probability that a person born in the same state as  $k$  would make the choice that  $k$  actually made (i.e. selecting into industry  $i$ ), which can be estimated. The sufficiency assumption can be relaxed by including functions of additional selection probabilities; for this reason,  $Z_{ki}$  includes a cubic polynomial in the estimated first-best selection probability and in the highest predicted probability for  $k$ . Identification in this approach is based on the exclusion of state of birth by industry of employment interactions from equation (10).

To estimate selection probabilities, we group individuals into cells defined by state of birth<sup>62</sup> and a vector of discrete characteristics: 4 categories of education attainment, 4 age intervals (16-30, 31-40, 41-50, 51-65), race, gender and 2 binary indicators of family status (family/non-family household and presence of own child 18 or younger in the household). As in Dahl (2002), for every

<sup>58</sup>Manufacturing employment excludes workers in private non-profit and government organizations.

<sup>59</sup>Since top codes vary by state, we follow Beaudry et al. (2007) and impose a common top-code value of \$175,000.

<sup>60</sup>These are: (i) High school dropout, (ii) high school graduate, (iii) some college but no degree, (iv) college degree or higher.

<sup>61</sup>See Dahl (2002), page 2378.

<sup>62</sup>As in Beaudry et al. (2007), we keep immigrants in the analysis by dividing the rest of the world into 14 regions (or ‘states’ of birth).

individual  $k$ , we estimate his selection probability into each industry  $j$  using the proportion of individuals within  $k$ 's cell that are observed working in  $j$ , denoted by  $\widehat{p}_{kj}$ . Individual  $k$ 's estimated first-best selection probability is  $\widehat{p}_{ki}$  and  $k$ 's highest predicted probability is given by  $\widehat{p}_{kj^*}$ , where  $j^*$  is such that  $\widehat{p}_{kj^*} = \max\{\widehat{p}_{kj}\} \forall j$ .

For the empirical analysis, the Census industry classification was matched to NAICS. It was not possible to match the trade data to Census codes directly, since the former is originally coded according to the Standard International Trade Classification (SITC rev.2). However, it is possible to use NAICS as a bridge between the two classifications. We construct a one-to-one mapping between the Census classification and NAICS by re-coding two or more 4 digit NAICS codes into a single industry (which does not necessarily match a 3 digit level). This re-coding also involves cases where two Census codes map perfectly into two NAICS codes -although originally there was no one-to-one matching between them. Importantly, the resulting mapping (available upon request) exhausts all manufacturing sectors in NAICS. Finally, the trade data was matched to wage inequality data using a concordance between SITC rev. 2 and NAICS, available through the NBER online database.

## C Appendix - Additional Data

In this Appendix we provide a description of additional data sources used in the empirical analysis. Descriptive statistics for each variable can be found in table 8.

*Bilateral export volumes at the industry level:* From Feenstra et al. (2005), for the year 2000. Sector-level bilateral exports data are categorized at the 4-digit SITC (4-digit rev. 2) level. The mapping from SITC to NAICS required the concordance available at the NBER website.<sup>63</sup>

*Bilateral trade barriers:* From HMR. This is a set of exporter-importer specific geographical, cultural and institutional variables. 1) *Distance*, the distance (in km.) between importer's and exporter's capitals (in logs). 2) *Land border*, a binary variable that equals one if and only if importer and exporter are neighbors that meet a common physical boundary. 3) *Island*, the number of countries in the pair that are islands. 4) *Landlocked*, the number of countries in the pair that have no coastline or direct access to sea. 5) *Colonial ties*, a binary variable that equals one if and only if the importing country ever colonized the exporting country or vice versa. 6) *Legal system*, a binary variable that equals one if and only if the importing and exporting countries share the same legal origin. 7) *Common Language*, a binary variable that equals one if and only if the exporting and importing countries share a common language. 8) *Religion*, computed as  $(\% \text{ Protestants in exporter} \times \% \text{ Protestants in importer}) + (\% \text{ Catholics in exporter} \times \% \text{ Catholics in importer}) + (\% \text{ Muslims in exporter} \times \% \text{ Muslims in importer})$ . 9) *FTA*, a binary variable that equals one if exporting and importing countries belong to a common regional trade agreement, and zero otherwise. 10) *GATT/WTO*, the number of countries in the pair that belong to the GATT/WTO.

*Start-up regulation costs:* From HMR. We use exporter-importer interactions of three proxies of regulation costs: the number of days ( $RegDays_H \times RegDays_F$ ), number of legal procedures ( $RegProc_H \times RegProc_F$ ) and relative cost as a percentage of GDP per capita ( $RegProc_H \times RegProc_F$ ), for an entrepreneur to start operating a business.

<sup>63</sup><http://www.nber.org/lipsey/sitc22naics97/>

*Factor endowments:* Physical capital endowment,  $KEndow$ , and human capital endowment,  $SkillEndow$ , are taken from Antweiler and Treffer (2002). A country’s stock of physical capital is the log of the average capital stock per worker. The stock of human capital is the natural log of the ratio of workers that completed high school to those that did not. The measures used are from 1992, the closest year of which data are available. There’s no data on factor endowments for four countries in our sample: Switzerland, Czech Republic, Hungary and Poland.

*Factor intensities:* From Nunn (2007). Originally coded as 1997 I-O industries, the mapping to NAICS required a concordance available from the Bureau of Economic Analysis.<sup>64</sup> Physical capital intensity,  $KIntens$ , is the total real capital stock divided by value added of the industry in the United States in 1996. Skill intensity,  $SkillIntens$ , is the ratio of non-production worker wages to total wages at the industry level in the United States in 1996. There’s no data on factor intensities for two industries: ‘Furniture and related products manufacturing’ and ‘Sawmills and wood preservation’.

*Proportion of top-coded wages:* From the 2000 Census of Population in the U.S. For each industry,  $TopCode$  is calculated as the proportion of workers earning a wage exceeding the top code value of \$175,000.

*Firm size dispersion:* From the 1997 Census of manufacturing in the U.S. For each industry, we calculate  $FirmDisp$ , the coefficient of variation in the average shipments per establishment across bins defined by employment size. The employment bins defined in the Census are: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000-2499 and 2500 employees or more.

*Quality of the judicial system:* From Nunn (2007)  $JudicQual$  is based on the “rule of law” measures originally from Kaufmann et al. (2003).

*Labor law rigidity:* From Tang (2008)  $LaborRigid$  is an index that summarizes firing and employment contract adjustment costs combined with measures of the power of labor unions. These measures are originally from Botero et al. (2004).

## D Appendix - Additional results with raw wage rankings and raw scores

Table A-1 reports estimates of the impact of skill dispersion as proxied by the dispersion of (raw) test scores: we identify this effect through an interaction with a (raw) wage dispersion ranking.<sup>65</sup> We show results based on three alternative measures of dispersion: the 95-5 interpercentile range divided by the average in column (1), the Gini relative mean difference (i.e. twice the Gini coefficient) in column (2) and the coefficient of variation in column (3).<sup>66</sup> Columns (1)-(3) add exporter, importer and industry dummies to our variables of interest; columns (4)-(6) include theoretically consistent exporter and importer-industry dummies, along with a vector of bilateral trade barriers

<sup>64</sup>[http://www.bea.gov/industry/xls/1997import\\_matrix.xls](http://www.bea.gov/industry/xls/1997import_matrix.xls)

<sup>65</sup>Raw measures are not purged of the effect of observable characteristics.

<sup>66</sup>We note that all three measures have a common structure in that the numerator is a measure of dispersion (the 95-5 interpercentile range, the standard deviation and the Gini mean difference) while the denominator is the average of the variable. Since we are using the logarithm of variables, the reason why we employ measures of dispersion divided by the average is not for rescaling, but rather to parsimoniously control for the effect that the interaction of the averages might have on trade flows.

described above.

In all specifications the estimated interaction  $WageDisp_i \times SkillDisp_H$  shows a positive effect on exports, significant throughout at the 5% level. The reported coefficients imply that a one standard deviation increase in the value of the interaction raises log exports by anywhere between 3.5% and 6.5% standard deviations.<sup>67</sup>

Table A-2 reproduces the structure of table A-1 in terms of controls, but it separately reports the effect of the interaction  $WageDisp_i \times SkillDisp_H$  (where the measure of dispersion is not divided by the average), as well as those of the interaction of average scores and average wages,  $WageMean_i \times SkillMean_H$ , and of the other two interactions,  $WageDisp_i \times SkillMean_H$  and  $WageMean_i \times SkillDisp_H$ . The interaction of the averages is expected to capture standard factor proportions effects: on average, countries with more skilled workers specialize in sectors that employ skilled workers and have higher average wages. The interaction  $WageMean_i \times SkillDisp_H$  is a flexible way to control for possible bias, due to differences in sectoral average wages, in the estimated effect of our interaction of interest. The interaction  $WageDisp_i \times SkillMean_H$  plays a similar role.<sup>68</sup> In general, columns (1)-(6) suggest that the coefficient of  $WageDisp_i \times SkillDisp_H$  is robust to the inclusion of all interactions: all estimates are similar to the ones in table A-1 and, when trade barriers and importer industry dummies are included, significant at the 5% level. As for the other interactions, as expected  $WageMean_i \times SkillMean_H$  has a strong and positive impact on trade flows. Moreover  $WageMean_i \times SkillDisp_H$  is consistently positive, significant and large, while  $WageDisp_i \times SkillMean_H$  is positive, but not always significant, particularly in columns (1)-(3). We note that the magnitudes of the impact of our variable of interest are similar in tables A-1 and A-2 to the ones in table 3 and 4, indicating a substantial degree of robustness in our results.

---

<sup>67</sup>In regressions we do not report, we interacted all three measures of dispersion for wages and scores with one another obtaining results qualitatively and quantitatively similar to columns (1)-(6).

<sup>68</sup>This interaction relates to the theoretical prediction that increases in average skills not resulting from proportional changes also have an effect on comparative advantage. This effect depends on the degree of complementarity, approximated by  $WageDisp_i$ .

Table A-1 - Normalized Raw Scores and Wage Rankings

Measure of Dispersion	(1)	(2)	(3)	(4)	(5)	(6)
	$\frac{\text{St Dev}}{\text{Mean}}$	$\frac{95-5 \text{ IPR}}{\text{Mean}}$	Gini RMD	$\frac{\text{St Dev}}{\text{Mean}}$	$\frac{95-5 \text{ IPR}}{\text{Mean}}$	Gini RMD
$WageDisp_i \times SkillDisp_H$	0.059** (0.017)	0.035* (0.016)	0.046* (0.018)	0.065** (0.017)	0.039* (0.016)	0.043* (0.018)
Trade Barriers	No	No	No	Yes	Yes	Yes
Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer FE	Yes	Yes	Yes	No	No	No
Industry FE	Yes	Yes	Yes	No	No	No
Importer-Industry FE	No	No	No	Yes	Yes	Yes
Observations	58124	58124	58124	58124	58124	58124
R-squared	0.54	0.54	0.54	0.7	0.69	0.69

The dependent variable is the natural logarithm of exports from country  $H$  to country  $F$  in industry  $i$ . Standardized beta coefficients are reported. †, \* and \*\* indicate the coefficient is significant at the 10%, 5% and 1% levels. Standard errors clustered by importer-exporter pair in parenthesis.

Table A-2 - Non-Normalized Interactions

	(1)	(2)	(3)	(4)	(5)	(6)
Measure of Dispersion	St Dev	95-5 IPR	Gini MD	St Dev	95-5 IPR	Gini MD
$WageDisp_i \times SkillDisp_H$	0.068* (0.03)	0.018 (0.027)	0.059 (0.038)	0.11** (0.029)	0.055* (0.025)	0.106** (0.036)
$WageMean_i \times SkillMean_H$	8.109** (0.525)	8.986** (0.513)	9.513** (0.619)	8.548** (0.433)	9.242** (0.428)	9.682** (0.512)
$WageMean_i \times SkillDisp_H$	0.37** (0.037)	0.41** (0.033)	0.479** (0.045)	0.349** (0.030)	0.377** (0.027)	0.438** (0.037)
$WageDisp_i \times SkillMean_H$	0.547 (0.493)	0.123 (0.485)	0.564 (0.57)	1.376** (0.493)	0.654 <sup>†</sup> (0.469)	1.468** (0.555)
Trade Barriers	No	No	No	Yes	Yes	Yes
Exporter FE	Yes	Yes	Yes	Yes	Yes	Yes
Importer FE	Yes	Yes	Yes	No	No	No
Industry FE	Yes	Yes	Yes	No	No	No
Importer-Industry FE	No	No	No	Yes	Yes	Yes
Observations	58124	58124	58124	58124	58124	58124
R-squared	0.54	0.54	0.54	0.70	0.70	0.70

The dependent variable is the natural logarithm of exports from country  $H$  to country  $F$  in industry  $i$ . Standardized beta coefficients are reported. <sup>†</sup>, \* and \*\* indicate the coefficient is significant at the 10%, 5% and 1% levels. Standard errors clustered by importer-exporter pair in parenthesis.