American Economic Review 2012, 102(5): 2327–2348
http://dx.doi.org/10.1257/aer.102.5.2327

Skill Dispersion and Trade Flows†

By Matilde Bombardini, Giovanni Gallipoli, and Germán Pupato*

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 proportions theory, and numerous related empirical studies,[1] identify quantities such as the stocks of human and physical capital of countries as primary sources of comparative advantage. This paper provides evidence supporting an alternative, and empirically sizeable, source of comparative advantage: the dispersion of skills (human capital) in the working population.[2]

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 skills. Figure 1 plots the mean against the standard deviation of IALS scores for 19 countries during 1994–1998 (Figure A-1 in the online Appendix reports the distribution of IALS scores for each country vis-à-vis the United States). The coefficient of variation of the standard deviation of scores is 1.6 times larger than that of the average scores, highlighting substantial cross-country differences in the second moments.

The reasons why countries at similar stages of development differ in their skill distributions are beyond the scope of this study.[3] 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,[4] early

* Bombardini: Department of Economics, University of British Columbia and CIFAR, 997-1873 East Mall, Vancouver BC V6T 1Z1, Canada (e-mail: matildeb@mail.ubc.ca); Gallipoli: Department of Economics, University of British Columbia, 997-1873 East Mall, Vancouver BC V6T 1Z1, Canada (e-mail: gallipol@mail.ubc.ca); Pupato: Graduate School of Economics, Getulio Vargas Foundation, Praia de Botafogo 190, CEP 22250-900, Rio de Janeiro, RJ, Brazil (e-mail: pupa08@gmail.com). We would like to thank four anonymous referees, Paul Beaudry, Brian Copeland, David Green, Patrick Francois, Keith Head, Thomas Lemieux, Vadim Marmer, Francesco Trebbi, Jonathan Vogel, and seminar participants at CIFAR, EIIT, EPGE-Fundação Getulio Vargas, Queen's University, Ryerson University, University of Alberta, University of Auckland, UC Davis, UC San Diego, University of Chicago Booth, Universidad de Chile, and University of Sydney for helpful comments.

† To view additional materials, visit the article page at http://dx.doi.org/10.1257/aer.102.5.2327.

[1] Recent studies, primarily Romalis (2004), testing the predictions of the theory about commodity trade, have detected larger effects compared to tests based on factor content, namely Bowen, Leamer, and Sveikauskas (1987); Trefler (1993); Trefler (1995); and Davis and Weinstein (2001).

[2] 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.

[3] 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 IIID.

[4] The existence of peer effects, as documented for example by Hanushek et al. (2003) and Hoxby (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).

Figure 1. Mean and Dispersion in IALS Scores (1994–1998)

tracking,[5] local school financing (Benabou 1996), and the shares of private and public schools (Takii and Tanaka 2009).[6]

In the absence of previous empirical research linking skill dispersion to comparative advantage, we start by showing that relative trade flows of manufacturing goods vary with skill dispersion, i.e., countries with higher skill dispersion export relatively more in some sectors. This analysis is performed by means of a simple "atheoretical" exercise which also shows that the effect of skill dispersion is quantitatively similar to that of average skill endowments, a usual suspect in the empirical trade literature.

Although this exercise cannot explain why some industries benefit from skill dispersion, it provides a useful motivation for the next step, in which we discipline our analysis by focusing on a specific sector characteristic which interacts with skill dispersion to generate comparative advantage. In particular, we exploit cross-industry variation in the degree of complementarity of workers' skills across production tasks. In some industries, such as aerospace or engine manufacturing, production requires completing a long sequence of tasks, and poor performance at any single stage greatly reduces the value of output. These are industries with high skill complementarity (or *O-Ring*, as in Kremer 1993), where teamwork is crucial and efficiency is higher if workers of similar skills are employed at every stage of production. On the contrary, in other industries, such as apparel, teamwork is relatively less important, skills are more easily substitutable and therefore poor performance in some tasks can be mitigated by superior performance in others. The question we pose is whether countries with greater skill dispersion specialize in sectors characterized by higher substitutability of skills across tasks.

[5] Tracking refers to the practice of grouping students in different schools according to their ability. Hanushek and Wößmann (2006) show that when grouping happens before age ten, inequality in education outcomes increases at the country level.

[6] 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.

The hypothesis that skill dispersion may lead to specialization has been the object of theoretical work by Grossman and Maggi (2000), henceforth GM. They show that in a two-country, two-sector model with perfectly observable talent and competitive labor markets, the country with a relatively more dispersed skill distribution specializes in the sector that benefits from matching workers of different skill levels. In Bombardini, Gallipoli, and Pupato (2009) we build on this insight and propose a multi-country, multi-sector model where skill dispersion generates testable implications for the pattern of international trade. Section II shows that the key difference between the two approaches resides in the role of observable (the focus in GM) versus unobservable skills (our focus), that is, the portion of skills that is not ex-ante observable during hiring. While in GM comparative advantage emerges as the result of perfect assortative (or cross) matching, we explore the alternative case of imperfect matching due to unobservability of certain skill dimensions. In the absence of sorting in unobservable skills between firms and workers, firms in every sector inherit the country's unobservable skill distribution.[7] Then, comparative advantage emerges from the combination of a sector's degree of skill complementarity and a country's skill dispersion.

A stylized example with two countries and two sectors, depicted in Figure 2, clarifies the intuition for our mechanism. Each sector employs only two workers, who perform symmetric tasks in the production process, and whose skills (a_1 and a_2) are measured on the axes. Technologies in the two sectors are represented by isoquants. For simplicity assume one of the sectors to be the limit case of perfect skill substitutability, corresponding to a linear isoquant Q_{PS}. Isoquant Q_{IS} represents a sector with imperfect substitutability of skills. Each country corresponds to one point: country C has two workers with the same average skills as country C' (they both lie on a line with constant mean skills). Skills in country C, however, are more dispersed relative to country C'. One can immediately verify that output in sector PS is the same for both countries because only aggregate skills matter in the presence of perfect skill substitutability. However, in the sector with skill complementarity (IS) output is higher in the country with lower skill dispersion, C'. The less dispersed country has a comparative advantage in the sector with higher skill complementarity.

The empirical counterpart of unobservable skills can be residually approximated by purging IALS scores of the effect of a variety of observable individual characteristics, such as education, age, and gender, to create what we refer to as "residual" skill dispersion. We investigate empirically the prediction that countries with more dispersed residual skill distributions specialize in sectors with lower skill complementarity in production. We adapt the empirical approach of Helpman, Melitz, and Rubinstein (2008) to industry-level bilateral trade flows and augment it with our variable of interest. The analysis shows that the interaction of exporter skill dispersion with sectoral measures of skill substitutability is a significant and economically large determinant of exports, while controlling for bilateral trade barriers, exporter and importer-industry fixed effects. We also include determinants of comparative advantage based on aggregate factor endowments, in the spirit of Romalis (2004),

[7] This assumption is consistent with evidence, in Altonji and Pierret (2001), that firms take time to learn about many dimensions of workers' skills and that sorting, both across industries and occupations, does not seem to depend, for the most part, on unobservable worker characteristics, as documented by Blackburn and Neumark (1992).

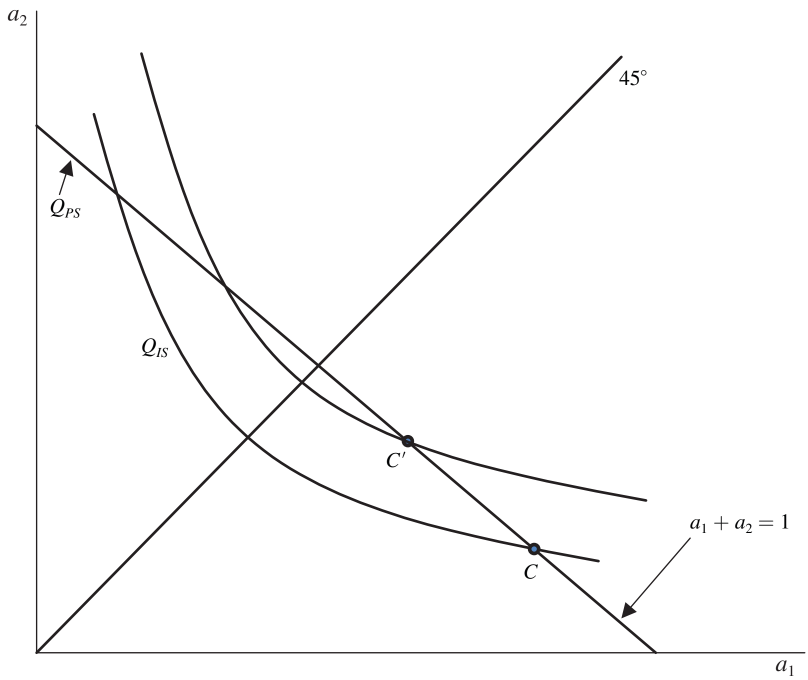

Figure 2. Comparative Advantage: Two Countries and Two Sectors

and institutional quality as in Nunn (2007) and show that their effects on trade flows are of the same statistical magnitude as that of skill dispersion.

The main focus of the paper is on residual skill dispersion. One reason for this choice is that, in the median country in our sample, residual dispersion accounts for 70 percent of overall dispersion. The second reason is that data constraints do not allow us to implement a theory-based test of GM (see Section IIID). However, we expand the analysis by also assessing the effect of predicted skill dispersion, a proxy for variation in observable skills, on trade flows. Although not formally grounded in GM's theory, this exercise confirms the significance and robustness of the effect of skill dispersion on comparative advantage.

As the degree of substitutability of skills is not directly observable, we take two distinct approaches to its measurement. First, we exploit a theoretical result—established in Bombardini, Gallipoli, and Pupato (2009)—linking the unobservable degree of complementarity to the observed dispersion of residual wages within industries. In a setting with labor market frictions and random matching on residual skills, residual wage dispersion within industries increases in the degree of skill substitutability. Sectors with higher complementarity are characterized by a more compressed wage distribution because, for example, workers with higher-than-average skills contribute relatively less to surplus, a fact reflected in their wage. As with IALS scores, in order to bring the empirical analysis in line with this theory, we use US census data to construct a measure of residual wages by purging the effect of observable characteristics from individual wages. In order to mimic random matching, we spend considerable effort addressing the possible non-random selection in unobservable characteristics across industries using a method proposed in Dahl (2002). Furthermore, in view of substantial evidence linking firm size and wages (e.g., Oi and Idson 1999), we filter out sector-specific firm heterogeneity from our residual wage dispersion measures.

Second, we use an alternative set of proxies for skill substitutability based on data from the Occupational Information Network (O*NET), which allow us to quantify the degree of teamwork, communication, and interdependence between coworkers' labor inputs. These measures do not rely on the theoretical structure of Bombardini, Gallipoli, and Pupato (2009) and provide a direct and intuitive way to proxy for complementarity.

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, b) and the extent of labor market frictions (Helpman and Itskhoki 2010; Cuñat and Melitz 2010; Tang forthcoming). 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.[8]

The paper is organized as follows. Section I provides preliminary evidence that skill dispersion matters as much as average skills in determining trade flows. Section II describes the theoretical background. Sections III and IV inspect the mechanism put forward in Section II. Section V concludes. A detailed data description can be found in the online Appendix.

I. Preliminary Evidence: The Importance of Second Moments

This section provides preliminary evidence that skill dispersion within a country shapes its pattern of international trade. We present an atheoretical exercise that aims at quantifying the overall effect of IALS dispersion on comparative advantage without the need to specify any particular mechanism driving specialization, a task that will be the concern of following sections in the paper. Importantly, the impact of skill dispersion is assessed against that of skill abundance, the first moment of the skill distribution, which is a natural benchmark in the trade literature.

More specifically, the question addressed in this section is: what is the effect of marginal changes in skill dispersion (as well as skill mean) on the relative exports of any two manufacturing industries? The exercise is implemented through an OLS regression of export volumes on interactions of the first and second moments of an exporter's skill distribution with a full set of industry dummies

$$\log X_{HFi} = \sum_{i\in S} \alpha_i^{mean} I_i \times SkillMean_H + \sum_{i\in S} \alpha_i^{disp} I_i \times SkillDisp_H + \mathbf{d}'_{HF}\gamma + \delta_H + \delta_{Fi} + \varepsilon_{HFi}, \tag{1}$$

[8] Glaeser et al. (2004) show that education is significantly more persistent than several other institutional features, such as the form of government.

where $\log X_{HFi}$ is the logarithm of the value of exports from country H to country F in industry i; $SkillMean_H$ and $SkillDisp_H$ are the mean and standard deviation of the distribution of log IALS scores in exporter H, and the I's are dummy variables for each of the S sectors (except an excluded baseline industry). Although not explicitly derived, the fixed effects included in this specification can be rationalized in a model of monopolistic competition with trade frictions, where bilateral trade flows depend on: the industry's price index and total expenditure level in the importing country (captured by importer-industry fixed effects, δ_{Fi}), the exporter's size (accounted for by exporter fixed effects, δ_H) and bilateral trade barriers (represented by $\mathbf{d}_{HF}$, a vector of observable bilateral trade frictions).[9] We estimate (1) employing the value of bilateral trade flows from 19 exporters to 145 importers in 63 industries in the year 2000. A detailed data description is provided in Section IIIB and the online Appendix. For comparability, average skill and skill dispersion are standardized across exporters. Estimation of (1) allows us to gauge the impact of mean skill and skill dispersion on the relative exports of any two exporting countries, say H and G, to an average third country F in industries i and j. For example, focusing on skill dispersion, we can write

$$E\left[\log\left(\frac{X_{HFi}}{X_{GFi}}\right) - \log\left(\frac{X_{HFj}}{X_{GFj}}\right)\right] = (\alpha_i^{disp} - \alpha_j^{disp})\Delta_{HG} SkillDisp, \tag{2}$$

where $\Delta_{HG} SkillDisp \equiv SkillDisp_H - SkillDisp_G$. Regardless of its sign, the larger the difference $\alpha_i^{disp} - \alpha_j^{disp}$ (in absolute value), the stronger the impact of skill dispersion on relative exports of i and j.[10] Reporting $|\alpha_i^{disp} - \alpha_j^{disp}|$ for each possible industry pair and for both moments is cumbersome (there are two sets of 62 α coefficients), therefore we summarize the estimation results by providing an average of those differences across *all* possible industry pairs. In this sense, the mean difference $MD(\alpha^{disp}) \equiv \frac{1}{S(S-1)} \sum_{i \in S} \sum_{j \in S} |\alpha_i^{disp} - \alpha_j^{disp}|$ captures the average effect of skill dispersion.

Within this framework we perform three different exercises. The first evaluates the importance of within-country skill dispersion vis-à-vis skill mean. Table 1 reports the bounds of 95 percent confidence intervals for $MD(\alpha^{mean})$ and $MD(\alpha^{disp})$, the estimated mean difference of the effects of mean and standard deviation of log IALS scores.[11] Column 1 indicates that both moments contribute to shaping the pattern of specialization across industries, with quantitatively similar effects.

The second exercise extends the specification in equation (1) by including standardized measures of the thickness of the left and right tails of the skill distribution in country H. Each of these measures is interacted with a full set of industry dummies, just like we do for the mean and standard deviation of skills. These "thickness-of-tails" measures correspond to the shares of the country's population that belong,

[9] The estimation framework is analogous to Manova (2008b) and, with the exception of our focus on a breakdown of trade flows by sectors, to Helpman, Melitz, and Rubinstein (2008).

[10] Notice that, while the choice of baseline industry clearly affects the individual estimates of the α's, it is inconsequential in terms of the object of interest, $\alpha_i^{disp} - \alpha_j^{disp}$, which is the pairwise difference in those estimated coefficients.

[11] Confidence intervals are computed using the Delta method.

TABLE 1—UNRESTRICTED EFFECTS ON RELATIVE TRADE FLOWS: 95 PERCENT CONFIDENCE INTERVALS

	(1)	(2)	(3)	(4)
Average log IALS	0.24–0.30	0.25–0.33	0.22–0.27	0.20–0.25
Standard deviation log IALS	0.15–0.20	0.24–0.30		
Population share in IALS 1st quintile		0.20–0.25		
Population share in IALS 5th quintile		0.42–0.54		
Standard deviation predicted log IALS			0.11–0.15	
Standard deviation residual log IALS			0.11–0.17	0.12–0.16

Notes: This table reports 95 percent confidence intervals for the mean difference of the 62 coefficients associated with interactions of standardized features of an exporter's log IALS score distribution (first column) and a full set of industry dummies from an OLS regression of equation (1). Standard errors are calculated using the Delta method.

respectively, to the top and bottom quintiles of the *world* IALS distribution.[12] The goal is to verify that the estimated effect of skill dispersion in column 1 is not solely driven by the tails of the distribution. In addition, we can assess whether cross-country differences in the sets of very (un)skilled individuals have an independent effect on trade (beyond their contribution to skill dispersion and mean).[13] The results in column 2 of Table 1 confirm this and show that confidence intervals of $MD(\alpha^{disp})$ and $MD(\alpha^{mean})$ overlap, pointing to statistically equivalent impacts of the first two moments of the skill distribution.

The third exercise employs the same approach to quantify and compare the impact on trade of two sources of skill dispersion, namely dispersion in observable skills and dispersion in unobservable skills. This exercise attempts to capture the role of variation due to easily observable "credentials," like education, as opposed to those residual skills that employers find harder to identify before a worker has been hired. This atheoretical framework allows us to assess the importance of both sources of skill dispersion (the decomposition of predicted and residual skills is discussed in Section IIIB). In this exercise, we interact each of three moments, average skills, the standard deviation of predicted skills and that of residual skills, with a full set of industry dummies and again report the MD's associated with each set of estimated coefficients. Column 3 of Table 1 shows that both types of skill dispersion matter for specialization and their effects have similar magnitude. Column 4 also shows that, whether or not predicted skill dispersion is included, the coefficients on residual dispersion are unaffected.

II. Theoretical Background: Why Dispersion Matters

The previous section shows that skill dispersion has an impact on trade flows, but does not explain why. This section highlights a mechanism through which skill dispersion matters for specialization, hinging on the degree of complementarity of skills across tasks in the production process. In Bombardini, Gallipoli, and Pupato

[12] The top and bottom quintiles of the world IALS distribution define two thresholds. For each country we compute the share of individuals above the top and below the bottom threshold. Notice that, in any country, these shares can be higher or lower than 20 percent.

[13] This exercise is particularly important in light of our analysis in Section G of the online Appendix, where we decompose the cross-country variation in skill dispersion and assess the importance of various parts of the skill distribution. Because we find that differences in the left tail of the distribution are the largest driver of the variation in skill dispersion, it is particularly important to verify, as we do in column 2 of Table 1, that, holding the thickness of the left tail constant, skill dispersion still has the same effect on trade flows.

(2009) we develop a monopolistic competition model with variable transport costs in which countries are characterized by different skill distributions.[14] All sectors feature symmetric supermodular production functions, but vary in the degree of complementarity of skills across tasks. More specifically, output y depends on the skill a of employed workers, the mass $h(a)$ of workers with given skill a, and a parameter λ measuring skill complementarity, so that $y = \left(\int a^{\lambda} h(a)\, da\right)^{1/\lambda}$ with $\lambda < 1$. Sectors with low λ benefit relatively more from a less dispersed skill distribution. The model in Bombardini, Gallipoli, and Pupato (2009) features labor market frictions in the spirit of Helpman and Itskhoki (2010). Workers decide to look for a job in an industry only knowing the average industry wage and its unemployment rate. By definition, any residual skill is not ex-ante observable to hiring firms. As a result, the distribution of residual skills of the set of workers looking for jobs in each industry will resemble the country's distribution, leading to no sorting along this dimension between workers and firms. Extending the model to account for the observable component of individual skills would result in firms only hiring workers of identical *observable* skills, but there would still be no sorting on residual skills. The model is static and, given labor market frictions, once workers are hired, bargaining between firm and workers determines wages, as described in detail in Bombardini, Gallipoli, and Pupato (2009) and discussed in Section IIIA.

Random matching on unobservable skills implies that, in equilibrium, the residual skill distribution prevailing in a country is passed on to every industry and firm.[15] Therefore output can be rewritten as a function of the mass of workers employed and a "productivity" factor $A(\lambda, c)$ defined as $A(\lambda, c) = \left(\int a^{\lambda} g(a, c)\, da\right)^{1/\lambda}$, where $g(a, c)$ is the distribution of skills in country c. The variation of $A(\lambda, c)$ across countries and industries is the unique determinant of comparative advantage and relative trade flows in the model. Of particular interest for the purpose of the empirical analysis in Section III is the case in which a country c', with identical mean but higher dispersion of skills than country c, has a comparative advantage in sectors with lower complementarity (high λ). This requires that, for any $\lambda' > \lambda$,

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

Inequality (3) simply states that countries with high skill dispersion are relatively more productive in low-complementarity sectors. Bombardini, Gallipoli, and Pupato (2009) examine (3) analytically and provide sufficient conditions on skill distributions and complementarity that ensure its validity. Here we present a simple numerical exercise to verify the empirical relevance of (3) using score distributions from IALS. $A(\lambda, c)$ is computed by replacing $g(a, c)$ with the empirical IALS distribution for each of the 19 participant countries. Given a grid of 100 λ's in the $[0, 1]$ interval, we calculate the ratio $\frac{A(\lambda, c')}{A(\lambda, c)}$ for every pair of countries (c, c') where c' has higher skill dispersion than c, according to the coefficient of variation of scores.

[14] Similar results can be obtained in a competitive environment. See Bombardini, Gallipoli, and Pupato (2012).

[15] This is consistent with recent international evidence (see Iranzo, Schivardi, and Tosetti 2008 and Lazear and Shaw 2008) suggesting that most of wage dispersion is in fact within, rather than between, firms.

We find that, averaging across pairs, $\frac{A(\lambda, c')}{A(\lambda, c)}$ is increasing in λ for 97 percent of the grid points. This result implies that if the empirical IALS distributions were used to simulate our model, they would generate a pattern of comparative advantage in which countries with higher skill dispersion export relatively more in industries with low complementarity.

Our theoretical analysis differs from GM's in three dimensions. First, we focus on the set of skills which are not easily observable ex-ante, so that random matching prevails along this dimension. This focus reflects the fact (documented in Section IIIB) that observable worker characteristics account for a smaller share of total variation in IALS scores within countries, i.e., measured skill dispersion is large among workers with similar "credentials." Second, we do not assume the existence of submodular sectors, i.e., sectors which benefit from cross-matching of skills. We posit instead that all sectors benefit from assortative matching, albeit to different degrees, which makes it easier to link our analysis to the existing trade literature, in which most production functions are supermodular.[16] The role of unobservable skills in the presence of supermodular technologies is only briefly discussed in GM.[17] Third, we provide a framework that is suitable for empirical testing as we model multiple countries, multiple sectors and transport costs, smoothing out the otherwise knife-edge predictions of Ricardian-type models.[18]

III. Inspecting the Mechanism: Residual Skills and Substitutability

This section presents evidence in support of the specific mechanism discussed above, linking residual skill dispersion to trade flows. First, we discuss the estimation framework. Section IIIB describes the data and Section IIIC reports baseline results. Section IIID discusses identification and presents robustness checks.

A. *Estimation Framework*

To test whether skill dispersion matters for trade flows through the specific channel of skill substitutability, we build on specification (1) and interact $SkillDisp_H$, a measure of skill dispersion in country H, with $Substit_i$, a measure of skill substitutability in industry i:

$$\log X_{HFi} = \beta Substit_i \times SkillDisp_H + \gamma d_{HF} + \delta_H + \delta_{Fi} + \varepsilon_{HFi}. \quad (4)$$

[16] It is worth stressing that in the presence of observable skills and symmetric supermodular production functions there is no basis for comparative advantage even if countries vary in the degree of skill dispersion. Each sector only employs workers of similar ability. Comparative advantage emerges only in the presence of a submodular sector where firms actively seek to match workers of different skill levels.

[17] In fact, we expand on an element introduced by GM: at the end of the paper they "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.

[18] Our and GM's models are not the only ones studying theoretical links between skill distributions and trade, although comparative advantage emerges as a result of substantially different mechanisms. Ohnsorge and Trefler (2007); Grossman (2004); Bougheas and Riezman (2007); and Costinot and Vogel (2010) are prominent examples of this literature.

The variable of interest is $Substit_i \times SkillDisp_H$ and estimation of its coefficient β allows us to test the prediction that, everything else equal, a country with a more dispersed skill distribution, exports relatively more in sectors with high substitutability of workers' skills. To see why, assume that equation (4) correctly specifies a model for the conditional expectation of $\log X_{HFi}$, so that $E[\varepsilon_{HFi} | Substit_i \times SkillDisp_H, d_{HF}, \delta_H, \delta_{Fi}] = 0$. Then, for any two countries H and G exporting to F, and any two industries i and j, equation (4) implies

$$(5) \qquad E\left[\log\left(\frac{X_{HFi}}{X_{GFi}}\right) - \log\left(\frac{X_{HFj}}{X_{GFj}}\right)\right] = \beta \Delta_{ij} Substit \times \Delta_{HG} SkillDisp,$$

where $\Delta_{HG} SkillDisp \equiv SkillDisp_H - SkillDisp_G$ and $\Delta_{ij} Substit$ is similarly defined. Our theoretical framework implies that $\beta > 0$. As in other studies of comparative advantage, our approach does not aim at explaining the overall volume, but rather the pattern of trade, i.e., differences in the composition of trade flows across countries. This initial specification is extended in Section IIID to account for alternative sources of comparative advantage that may be correlated with skill dispersion.

A difficulty in implementing a test of our hypothesis comes from the fact that the elasticity of substitution of individuals' skills at the industry level is not directly observable and we are not aware of any estimates of the elasticity of substitution across finely disaggregated skills. We take two different approaches to proxying for the elasticity of substitution of workers skills. The first is based on a theoretically-founded link between complementarity and residual wage dispersion. In the second approach we use proxies for complementarity available from occupation-level microdata.

Skill Substitutability: Residual Wage Dispersion Rankings.—What follows is a heuristic explanation of the link between complementarity and (residual) wage dispersion.[19] Consistent with empirical evidence, e.g., Altonji and Pierret (2001), suggesting that firms learn only gradually about worker skills, we posit that part of unobservable skills are revealed after hiring. Due to frictions, we assume wages are determined by multilateral bargaining within the firm. At the bargaining stage each worker receives a wage that corresponds to her average marginal product (the Shapley value), therefore workers of higher skills receive higher wages. To the extent that each sector inherits the country-specific distribution of residual skills, the variation in the distribution of residual wages only reflects technological differences across sectors. Therefore wage dispersion is driven by the degree of skill complementarity. For example, in a sector with high complementarity and a stronger need for a homogeneous labor force, high skill workers have lower marginal product, relative to high substitutability sectors, because their skills are far from the average. In general, sectors with low complementarity (high substitutability) will exhibit more dispersed wage distributions. We cannot rely on our theory to structurally recover actual values of skill substitutability, but we can use its

[19] A complete derivation is available in Bombardini, Gallipoli, and Pupato (2009).

unambiguous prediction of a monotonic relationship between skill substitutability and residual wage dispersion to identify a ranking.

*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 IIIB, 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 coworker output:* A closely related way of characterizing complementarity is to quantify the extent to which a worker's actions impact the performance of coworkers; 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 coworkers 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 skill substitutability, we simply rely on O*NET to identify a ranking of industries in terms of skill substitutability.[20]

B. *Data*

A detailed data description can be found in the online Appendix. Here we discuss the measurement of two key explanatory variables in the empirical analysis, skill dispersion at the country level and skill substitutability at the industry level.

Residual Skill Dispersion.—We use test scores from the 1994–1998 International Adult Literacy Survey (IALS)[21] to approximate the skill distribution within a country. Collaborators in this household survey administered a 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 score 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:

$$\log(s_{kH}) = \mathbf{X}'_{kH}\boldsymbol{\beta}_H + \epsilon_{kH}, \tag{6}$$

[20] For both wage dispersion and O*NET, regression results are qualitatively unchanged if we employ the value of the proxies instead of their ranking.

[21] International Adult Literacy Survey (IALS), http://www.statcan.gc.ca/pub/89-588-x/4152887-eng.htm.

Table 2—IALS Log-Scores

Coefficient of variation rank	Exporter	Mean		Standard deviation		Residual standard deviation	
		Rank	Value	Rank	Value	Rank	Value
1	Denmark	3	5.671	1	0.150	1	0.128
2	Germany	6	5.654	2	0.162	4	0.147
3	Netherlands	4	5.666	3	0.167	2	0.136
4	Norway	2	5.684	4	0.171	3	0.145
5	Finland	5	5.666	5	0.181	5	0.151
6	Sweden	1	5.717	6	0.184	6	0.153
7	Czech Republic	7	5.636	7	0.190	7	0.168
8	Hungary	15	5.546	8	0.204	8	0.176
9	Belgium	8	5.632	9	0.221	10	0.187
10	New Zealand	10	5.597	10	0.240	13	0.211
11	United Kingdom	11	5.595	11	0.262	17	0.234
12	Ireland	14	5.569	12	0.266	12	0.209
13	Switzerland	13	5.573	13	0.269	9	0.186
14	Canada	9	5.628	14	0.274	11	0.187
15	Italy	16	5.499	15	0.285	15	0.224
16	United States	12	5.587	16	0.289	14	0.215
17	Chile	19	5.355	17	0.302	16	0.224
18	Slovenia	17	5.446	18	0.314	18	0.246
19	Poland	18	5.415	19	0.333	19	0.284

where s_{kH} is the IALS score of k and $\mathbf{X}_{kH}$ is a vector of individual demographic information from the IALS questionnaire: education, age, gender, immigrant status, and on-the-job training (details in online Appendix A). The residual $\widehat{\epsilon_{kH}}$ is then used to compute the skill dispersion measures used for the estimation of trade flows. Analyzing the R^2 of these country-by-country regressions, we find that the variation in residual scores $\widehat{\varepsilon_{kH}}$ accounts for a minimum of 46 percent of the observed variation in log-scores in Canada, for a maximum of 83 percent in Germany, and for 70 percent in Finland, the median country in the sample.

Table 2 ranks 19 countries according to the coefficient of variation of IALS scores, and also reports their rank by mean, standard deviation, and standard deviation of residual IALS. The table shows how countries at similar stages of development differ substantially in the degree of skill dispersion: the United States and the United Kingdom display a more dispersed skill distribution than Sweden and Germany.[22]

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 percent Public Use Microdata Sample[23] (PUMS) files of the 2000 Census of Population in the United States to construct industry-specific measures of wage dispersion and identify a ranking of industries with respect to the unobserved elasticity of substitution. An advantage of our

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

[23] See US Census Bureau (2003).

approach is that we can match individual wage observations to a detailed industry classification, accounting for the entire manufacturing sector. 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.[24]

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

$$\log(w_{ki}) = \mathbf{Z}'_{ki}\boldsymbol{\beta}_i + \xi_{ki}, \tag{7}$$

where w_{ki} is the weekly wage of k and $\mathbf{Z}_{ki}$ is a vector of observable characteristics (education, age, gender, and race, see online Appendix A for details). Note that we run these regressions separately for each industry to allow for differences in the return to observable characteristics across industries.[25]

Several studies have shown that firm size affects wages (Oi and Idson 1999). This implies that wage dispersion might also reflect variation in the distribution of firm size across different industries. Therefore 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. The residuals from this regression are employed to construct a ranking of industries in terms of wage dispersion (in Table 3 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. Although these rankings are constructed using US data, in online Appendix C we show that rankings based on Canadian data are highly correlated.

*O*NET Survey-Based Measures of Complementarity.*—Sponsored by the Employment and Training Administration of the United States Department of Labor, O*NET[26] provides detailed information on job requirements and worker attributes for 965 occupations in the United States. 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.

Our complementarity rankings are based on four selected O*NET (Version 12.0) questions capturing different aspects of skill complementarity: (i) *Teamwork*: How

[24] In essence, this procedure controls for selection effects using differences in the probability 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 online Appendix.

[25] Regression results are available upon request.

[26] O*NET, http://www.doleta.gov/programs/onet/.

Table 3—Substitutability Rankings

	Wage dispersion$_i$ residual standard deviation rank	O^*NET_i $Contact_i$ rank
Lowest substitutability ($Substit_i$)		
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 substitutability ($Substit_i$)		
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

important are interactions that require you to work with or contribute to a work group or team to perform your current job?[27] (ii) *Impact*: How do the decisions an employee makes impact the results of coworkers, clients or the company? (iii) *Communication*: How important is communicating with supervisors, peers, or subordinates to the performance of your current job? (iv) *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 one to five. 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 through a common occupational classification (the Standard Occupational Classification). In this way, as occupational structures vary across industries, we obtain a different distribution of scores for each industry. Using the median score[28] for each industry we generate O^*NET_i, a ranking of sectors in terms of substitutability.[29] Industries with higher O^*NET_i exhibit lower skill substitutability. Table 3 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 wage dispersion we find the bottom O^*NET_i sectors (those sectors with high substitutability). This reflects the fact that, as shown in online Appendix Table A-1, the rankings based on occupational surveys, O^*NET_i, and the rankings based on residual wage dispersion are inversely correlated.

[27] 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.

[28] We employ average scores to break ties based on the medians.

[29] The results are robust to reweighting by hours worked and to using mean scores instead of medians as complementarity proxies.

Table 4—Residual Wage Dispersion Rankings and Residual Score Dispersion

Measure of dispersion	Standard deviation (1)	95-5 inter-percentile ratio (2)	Gini mean difference (3)	Standard deviation (4)	95-5 inter-percentile ratio (5)	Gini mean difference (6)
Wage dispersion$_i$ $\times$ *Skill dispersion*$_H$	0.017** (0.004)	0.015** (0.004)	0.016** (0.004)	0.016** (0.004)	0.013** (0.004)	0.014** (0.004)
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	58,124	58,124	58,124	58,124	58,124	58,124
R^2	0.54	0.54	0.54	0.70	0.70	0.70

Notes: The dependent variable is the natural logarithm of exports from country H to country F in industry i. Standardized beta coefficients are reported. Bootstrap standard errors clustered by importer-exporter pair in parenthesis (50 replications).

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

C. *Baseline Results*

This section discusses results of the empirical analysis of trade flows using specification (4). We report results employing first wage dispersion rankings and then O*NET rankings. Unless otherwise noted, the method of estimation is OLS. For comparability, all tables report standardized coefficients of the explanatory variables.

Results with Substitutability Proxied by Wage Dispersion Rankings.—Table 4 reports estimates of the impact of skill dispersion as proxied by the dispersion of residual IALS test scores (defined in Section IIIB): we identify this effect through an interaction with residual wage dispersion rankings (defined in Section IIIB). The measures of dispersion employed in Table 4 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 online Appendix. We find that the interaction of skill substitutability and skill dispersion has a positive and significant effect on exports. Note that the magnitudes of the coefficient are stable across specifications and measures of dispersion. The standardized coefficient of the interaction varies between 1.3 percent and 1.7 percent in the 6 specifications. The quantitative relevance of this channel is discussed in Section IV alongside that of other sources of comparative advantage.

*Results with Substitutability Proxied by O*NET Rankings.*—Next, we report estimates of the effect of skill dispersion on trade flows using four alternative measures of skill complementarity constructed from the O*NET database. Table 5 replicates the structure of columns 4–6 of Table 4, in terms of the set of fixed effects included and trade barriers used as controls. The variable of interest is the interaction of skill

Table 5—O*NET Rankings and Residual Score Dispersion (*standard deviation*)

Measure of complementarity	$O{*}NET_i$ = $Teamwork_i$ (1)	$O{*}NET_i$ = $Impact_i$ (2)	$O{*}NET_i$ = $Communication_i$ (3)	$O{*}NET_i$ = $Contact_i$ (4)	*Aggregate* $O{*}NET_i$ (5)
$O{*}NET_i \times$ *Skill dispersion*$_H$	−0.029**	−0.027**	−0.028**	−0.023**	−0.032**
	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)
Trade barriers	Yes	Yes	Yes	Yes	Yes
Exporter FE	Yes	Yes	Yes	Yes	Yes
Importer-industry FE	Yes	Yes	Yes	Yes	Yes
Observations	58,124	58,124	58,124	58,124	58,124
R^2	0.70	0.70	0.70	0.70	0.70

Notes: The dependent variable is the natural logarithm of exports from country H to country F in industry i. Standardized beta coefficients are reported. Bootstrap standard errors clustered by importer-exporter pair in parenthesis (50 replications).
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

dispersion in country H (measured by the standard deviation of residual scores) and the corresponding O*NET ranking: $Teamwork_i$, $Impact_i$, $Communication_i$, 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 5 at the 1 percent significance level. The estimates of the effect of skill dispersion are of similar magnitude to the ones generated using the wage dispersion rankings.[30]

Since we find consistent results across all four correlated survey-based measures of complementarity, and in order to provide a concise robustness analysis section, we create an O*NET ranking based on the four rankings above. Column 5 of Table 5 reports similar results using this Aggregate $O{*}NET_i$ ranking.[31]

D. *Identification and Robustness*

In this section we discuss some issues related to the identification of the effects quantified in Tables 4 and 5. Table 6 below reports results with both wage dispersion (columns 1, 3, and 5) and aggregate O*NET rankings (columns 2, 4, and 6), although we only include coefficient estimates using the standard deviation of residual skills. Results are unchanged if we employ the 95-5 and Gini skill dispersion measures.

The Extensive Margin of Trade: Selection.—Tables 4 and 5 report estimation results which do not take into account the fact that a substantial fraction of bilateral

[30] In unreported regressions we check that these results are qualitatively unchanged if: (i) skill dispersion 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.

[31] *Aggregate* $O{*}NET_i$ is a ranking variable based on the median and average of the four $O{*}NET_i$ rankings (as we did for each individual *O*NET* ranking, the average is employed to break ties in rankings based on the median).

Table 6—Robustness

	HMR selection		Controls		Predicted skills	
Sustitutability$_i$	Wage dispersion$_i$ (1)	O*NET$_i$ (2)	Wage dispersion$_i$ (3)	O*NET$_i$ (4)	Wage dispersion$_i$ (5)	O*NET$_i$ (6)
Substitutability$_i$ × Residual skill dispersion$_H$	0.016**	−0.033**	0.032**	−0.066**	0.035**	−0.050**
	(0.004)	(0.010)	(0.009)	(0.012)	(0.009)	(0.011)
Substitutability$_i$ × Predicted skill dispersion$_H$					−0.004	−0.019*
					(0.004)	(0.008)
Capital intensity$_i$ × Capital endowment$_H$			0.029**	0.030**	0.029**	0.030**
			(0.008)	(0.008)	(0.008)	(0.008)
Skill Intensity$_i$ × Skill endowment$_H$			0.033**	0.018**	0.033**	0.023**
			(0.006)	(0.006)	(0.006)	(0.006)
Differentiated$_i$ × Judicial quality$_H$			0.021	0.020	0.021	0.018
			(0.011)	(0.012)	(0.011)	(0.011)
Substitutability$_i$ × Labor rigidity$_H$			0.008*	−0.036**	0.007*	−0.034**
			(0.004)	(0.006)	(0.004)	(0.006)
Share top code$_i$ × Residual skill dispersion$_H$			−0.006	0.029**	−0.006	0.029**
			(0.007)	(0.005)	(0.007)	(0.005)
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	52,455	52,455	41,301	41,301	41,301	41,301
R^2	0.69	0.70	0.73	0.73	0.73	0.73

Notes: The dependent variable is the natural logarithm of exports from country H to country F in industry i. All columns employ the standard deviation of IALS log-scores as a measure of skill dispersion. As proxy for skill substitutability: columns 1, 3, and 5 employ a ranking based on the standard deviation of residual wages; columns 2, 4, and 6 employ *Aggregate O*NET$_i$* ranking. Standardized beta coefficients are reported. Bootstrap standard errors clustered by importer-exporter pair in parenthesis (50 replications). The regression includes an unreported polynomial in the probability to export, obtained from the first stage.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

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 (4) requires excluding observations for countries which do not trade in specific industries. These amount to 66.5 percent of the sample. As discussed in Helpman, Melitz, and Rubinstein (2008), 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 (4), including β. In order to correct for selection bias, we implement the two-step estimation procedure proposed by Helpman, Melitz, and Rubinstein (2008) (details in the online Appendix). Table 6 reports second stage results obtained using the selection correction. Columns 1 and 2 document the robustness of the skill dispersion effect.

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. Columns 3 and 4 of Table 6 show that the estimated effect of the interaction of substitutability ranking and residual skill dispersion is robust to a number of controls for other potential determinants of comparative advantage. We introduce controls for standard Heckscher-Ohlin sources of comparative advantage: the interaction of factor endowment of a country (in particular human capital and physical capital) and factor intensity of the sector (human capital and physical capital), in the spirit of Romalis (2004). Since 95 percent confidence intervals overlap, the impact on trade flows of our interaction of interest is quantitatively similar to the Heckscher-Ohlin effects of the human and physical capital interactions, $\textit{skill intensity}_i \times \textit{skill endowment}_H$ and $\textit{capital intensity}_i \times \textit{capital endowment}_H$. We also control for institutional characteristics of exporters. In particular, we interact $\textit{Differentiated}_i$ (a measure of sector *i* contract intensity) with $\textit{Judicial quality}_H$ (a measure of judicial quality) as in Nunn (2007) and our skill substitutability proxies with $\textit{Labor rigidity}_H$, a measure of labor law rigidity in country *H*, from Tang (2008). Including these alternative controls does not substantially affect the magnitude of our variable of interest and indicates that institutional quality has an impact on trade flows that is quantitatively similar to that of skill dispersion. We also introduce the share of individual wages that are top-coded within an industry, $\textit{Share top code}_i$, interacted with skill dispersion 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).[32]

Finally, we expand our analysis by including a measure of observable skill dispersion. Although not a formal test of GM's theory,[33] columns 5 and 6 add an interaction of skill substitutability with the coefficient of variation of the predicted component of skills as estimated in (6). The coefficient on our interaction of interest is unchanged, while the effect of observable skill dispersion is broadly in line with the intuition suggested by the GM model, although not always statistically significant.

Reverse Causality.—Wage dispersion rankings and skill dispersion might be partly influenced by the pattern of international trade, potentially resulting in reverse causality.[34] The orthogonality condition needed for consistent estimation of β in equation (4) is

$$E(\textit{WageDisp}_s \times \textit{SkillDisp}_c \times \varepsilon_{HFi}) = 0 \quad \forall s, c. \tag{8}$$

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

$$E(\textit{WageDisp}_s \times \varepsilon_{HFi} | \textit{SkillDisp}_c) = 0 \quad \forall s, c, \tag{9}$$

[32] For briefness we include all controls at once. The working paper version reports estimates with controls included one at a time.

[33] A difficulty in testing GM is that it is unclear how their predictions can be extrapolated in order to carry out a multi-country and multi-sector empirical analysis of the impact of skill dispersion on trade flows. Moreover, our substitutability proxies only identify a *ranking* of industries according to the degree of skill substitutability, but not whether any given sector's technology is submodular or supermodular in skills. When skills are observable, GM find that skill dispersion has an ambiguous effect on the pattern of trade across industries that are ranked in terms of skill substitutability. For example, in a two-country two-sector setting, skill dispersion will not generate comparative advantage if both production technologies have different degrees of supermodularity in skills. Conversely, trade will emerge if one of the sectors has a submodular production function. As a result, the same ranking can yield different trade patterns.

[34] It is less obvious how international trade may affect the survey based rankings O^*NET_i.

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 (9) would arise if ε_{HFi} 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, Itskhoki, and Redding (2010) 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.[35] However, as shown in Helpman, Melitz, and Rubinstein (2008), the correction for self-selection into the export market discussed in Section IIID effectively removes this potential bias.

Furthermore, since we measure wage dispersion at the industry level using US data, we can check the robustness of our estimates by removing the United States from our set of exporters. To the extent that the US 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 wage dispersion. This procedure yields a coefficient on our interaction of interest of 0.035 (with standard error 0.01), effectively unchanged when compared to the specification in column 3 of Table 6.

An alternative sufficient condition, $E(SkillDisp_c \times \varepsilon_{HFi} | WageDisp_s) = 0$ for all s, c, which guarantees (8), is discussed in online Appendix E.

IV. Magnitudes

Although regression coefficients are standardized and therefore readily comparable, in this section we interpret their magnitude in terms of trade flows. For ease of comparison with other control variables we focus on the full specification in column 3 of Table 6 and take 0.032 as the estimated effect of the interaction of country skill dispersion and industry substitutability measures. Consider two countries, the United Kingdom and Canada, and two industries, "computers" and "plastics." These countries and industries are chosen because residual skill dispersion in the United Kingdom is (approximately) one standard deviation higher than in Canada and the residual wage dispersion rank in computers is one standard deviation higher than in plastics. Since the standard deviation of log exports is 2.204 (Table A-5 in the online Appendix), the expected ratio of relative exports of computers in the United Kingdom and Canada, i.e., $E\left[\frac{X_{UK,F}(computers)}{X_{UK,F}(plastics)} / \frac{X_{CAN,F}(computers)}{X_{CAN,F}(plastics)}\right]$, is given by $e^{0.032 \times 2.204}$. This implies that, all else constant, skill dispersion induces exports of computers (relative to plastics) in the United Kingdom to be 7.3 percent higher than in Canada. To put this result in perspective, the estimates from column 3 of Table 6 imply that similar exercises yield a figure of 7.5 percent due to cross-country differences in human capital abundance (the Heckscher-Ohlin channel) and 4.7 percent due to institutional quality as in Nunn (2007).

One could also adopt the standard "Rajan-Zingales" (Rajan and Zingales 1998) approach of comparing industries and countries at the 25th and 75th percentiles

[35] Exporters do pay higher wages. See, for example, Bernard, Jensen, and Lawrence (1995) and Bernard and Jensen (1997).

of their respective distributions. Implementing this exercise for the skill dispersion channel requires similar calculations as before, except that now we consider the countries at the 25th and 75th percentiles of the skill dispersion distribution and the industries at the 25th and 75th percentiles of the residual wage dispersion rankings. As a result, the relative exports of the 75th percentile country in the 75th percentile sector are 24.5 percent higher due to the skill dispersion channel, 10.9 percent due to the skill endowment channel and 11.9 percent due to the institutional quality channel.

V. Conclusions

Relative differences in aggregate factor endowments are central to the classical theory of international trade. In this paper we push this idea further and argue that the entire distribution of a factor, and not just its aggregate endowment, can help rationalize observed trade flows. The analysis presents evidence that skill dispersion in the labor force has a quantitatively comparable effect to skill abundance in shaping comparative advantage. In particular we explore the prediction, developed in Bombardini, Gallipoli, and Pupato (2009), that if (i) workers and firms randomly match along the unobservable component of skills, and (ii) industries vary in the degree to which they can substitute workers of different skills across production tasks, then firms in sectors with higher complementarity are relatively more productive in countries with lower skill dispersion.

The empirical finding that countries with higher residual skill dispersion specialize in low complementarity sectors is robust to alternative measures of skill substitutability and skill dispersion, as well as to controls for alternative sources of comparative advantage. Importantly, the magnitude of the effect of skill dispersion is comparable to that of the aggregate skill endowment and institutional quality.

Finally, we notice that the analysis in the paper has implications for the impact of trade on residual wage inequality, which are beyond the scope of this study. Our results, taken at face value, imply that a more disperse skill distribution might have an indirect effect on a country's earnings distribution, as higher skill dispersion induces specialization in sectors that generate high residual wage dispersion.

REFERENCES

Altonji, Joseph G., and Charles R. Pierret. 2001. "Employer Learning and Statistical Discrimination." *Quarterly Journal of Economics* 116 (1): 313–50.

Benabou, Roland. 1996. "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance." *American Economic Review* 86 (3): 584–609.

Bernard, Andrew B., and J. Bradford Jensen. 1997. "Exporters, Skill Upgrading, and the Wage Gap." *Journal of International Economics* 42 (1–2): 3–31.

Bernard, Andrew B., and J. Bradford Jensen, and Robert Z. Lawrence. 1995. "Exporters, Jobs, and Wages in US Manufacturing: 1976–1987." *Brookings Papers on Economic Activity* 67–112.

Blackburn, McKinley, and David Neumark. 1992. "Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials." *Quarterly Journal of Economics* 107 (4): 1421–36.

Bombardini, Matilde, Giovanni Gallipoli, and Germán Pupato. 2009. "Skill Dispersion and Trade Flows." National Bureau of Economic Research Working Paper 15097.

Bombardini, Matilde, Giovanni Gallipoli, and Germán Pupato. 2012a. "Skill Dispersion and Trade Flows: Dataset." *American Economic Review*. http://dx.doi.org/10.1257/aer.102.5.2327.

Bombardini, Matilde, Giovanni Gallipoli, and Germán Pupato. 2012b. "Unobservable Skill Dispersion and Comparative Advantage." http://faculty.arts.ubc.ca/mbombardini/skilldisp-theory.pdf.

Bougheas, Spiros, and Raymond Riezman. 2007. "Trade and the Distribution of Human Capital." *Journal of International Economics* 73 (2): 421–33.

Bowen, Harry P., Edward E. Leamer, and Leo Sveikauskas. 1987. "Multicountry, Multifactor Tests of the Factor Abundance Theory." *American Economic Review* 77 (5): 791–809.

Brown, Giorgina, John Micklewright, Sylke V. Schnepf, and Robert Waldmann. 2007. "International Surveys of Educational Achievement: How Robust Are the Findings?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 170 (3): 623–46.

Costinot, Arnaud, and Jonathan Vogel. 2010. "Matching and Inequality in the World Economy." *Journal of Political Economy* 118 (4): 747–86.

Cuñat, Alejandro, and Marc J. Melitz. 2010. "A Many-Country, Many-Good Model of Labor Market Rigidities as a Source of Comparative Advantage." *Journal of the European Economic Association* 8 (2–3): 434–41.

Dahl, Gordon B. 2002. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70 (6): 2367–2420.

Davis, Donald R., and David E. Weinstein. 2001. "An Account of Global Factor Trade." *American Economic Review* 91 (5): 1423–53.

Friesen, Jane, and Brian Krauth. 2007. "Sorting and Inequality in Canadian Schools." *Journal of Public Economics* 91 (11–12): 2185–2212.

Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9 (3): 271–303.

Grossman, Gene M. 2004. "The Distribution of Talent and the Pattern and Consequences of International Trade." *Journal of Political Economy* 112 (1): 209–39.

Grossman, Gene M., and Giovanni Maggi. 2000. "Diversity and Trade." *American Economic Review* 90 (5): 1255–75.

Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivken. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18 (5): 527–44.

Hanushek, Eric A., and Ludger Wößmann. 2006. "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries." *Economic Journal* 116 (510): C63–76.

Helpman, Elhanan, and Oleg Itskhoki. 2010. "Labour Market Rigidities, Trade and Unemployment." *Review of Economic Studies* 77 (3): 1100–1137.

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding. 2010. "Inequality and Unemployment in a Global Economy." *Econometrica* 78 (4): 1239–83.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein. 2008. "Estimating Trade Flows: Trading Partners and Trading Volumes." *Quarterly Journal of Economics* 123 (2): 441–87.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." National Bureau of Economic Research Working Paper 7867.

Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti. 2008. "Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data." *Journal of Labor Economics* 26 (2): 247–85.

James, Estelle. 1993. "Why Do Different Countries Choose a Different Public-Private Mix of Educational Services?" *Journal of Human Resources* 28 (3): 571–92.

Kremer, Michael. 1993. "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics* 108 (3): 551–75.

Lazear, Edward P., and Kathryn L. Shaw, eds. 2008. *The Structure of Wages: An International Comparison*. A National Bureau of Economic Research Conference Report. Chicago: University of Chicago Press.

Levchenko, Andrei A. 2007. "Institutional Quality and International Trade." *Review of Economic Studies* 74 (3): 791–819.

Manova, Kalina. 2008a. "Credit Constraints, Equity Market Liberalizations and International Trade." *Journal of International Economics* 76 (1): 33–47.

Manova, Kalina. 2008b. "Credit Constraints, Heterogeneous Firms, and International Trade." National Bureau of Economic Research Working Paper 14531.

Nunn, Nathan. 2007. "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade." *Quarterly Journal of Economics* 122 (2): 569–600.

Ohnsorge, Franziska, and Daniel Trefler. 2007. "Sorting It Out: International Trade with Heterogeneous Workers." *Journal of Political Economy* 115 (5): 868–92.

Oi, Walter Y., and Todd L. Idson. 1999. "Firm Size and Wages." In *Handbook of Labor Economics Volume 3B*, edited by O. Ashenfelter and D. Card, 2165–2214. Handbooks in Economics, vol. 5. New York: Elsevier Science, North-Holland.

Rajan, **Raghuram G.**, **and Luigi Zingales.** 1998. "Financial Dependence and Growth." *American Economic Review* 88 (3): 559–86.

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

Stevenson, **David L.**, **and David P. Baker.** 1991. "State Control of the Curriculum and Classroom Instruction." *Sociology of Education* 64 (1): 1–10.

Takii, **Katsuya**, **and Ryuichi Tanaka.** 2009. "Does the Diversity of Human Capital Increase GDP? A Comparison of Education Systems." *Journal of Public Economics* 93 (7–8): 998–1007.

Tang, **Heiwai.** Forthcoming. "Labor Market Institutions, Firm-specific Skills, and Trade Patterns." *Journal of International Economics*.

Trefler, **Daniel.** 1993. "International Factor Price Differences: Leontief Was Right!" *Journal of Political Economy* 101 (6): 961–87.

Trefler, **Daniel.** 1995. "The Case of the Missing Trade and Other Mysteries." *American Economic Review* 85 (5): 1029–46.

US Census Bureau. 2003. Census 2000, Public Use Microdata Sample (PUMS), United States, http://www.census.gov/main/www/pums.html.