Unobservable Skill Dispersion and Comparative Advantage

Matilde Bombardini[*], Giovanni Gallipoli[†] and Germán Pupato[‡]

First Draft: June 2009

This Draft: October 2013

Abstract

This paper investigates a theoretical mechanism linking comparative advantage to the distribution of skills in the working population. We develop a tractable multi-country, multi-industry model of trade with unobservable skills in the labor market and show that comparative advantage derives from (i) cross-industry differences in the substitutability of workers' skills and (ii) cross-country differences in the dispersion of skills. We establish the conditions under which higher skill dispersion leads to specialization in industries characterized by higher skill substitutability across tasks. The main results are robust when the model is extended to allow for partial observability of skills. Finally, we use distributions of literacy scores from the International Adult Literacy Survey to approximate cross-country productivity differences due to skill dispersion and we carry out a quantitative assessment of the impact of skill dispersion on the pattern of trade.

JEL Classification codes: F12, F16, J82.

We would like to thank Paul Beaudry, Pablo D'Erasmo, David Green, Patrick Francois, Keith Head, Thomas Lemieux, Giovanni Maggi, Vadim Marmer, Thomas Sampson, Francesco Trebbi, Jonathan Vogel and seminar participants at CIFAR, EIIT, Getulio Vargas Foundation, Queen's University, Ryerson University, University of Alberta, University of Auckland, UC Davis, UC San Diego, University of Chicago Booth, Universidad de Chile, University of Sydney and Yale University for helpful comments. Bingjing Li and Ruoying Wang provided excellent research assistance.

[*]University of British Columbia, CIFAR and NBER.

[†]University of British Columbia and RCEA.

[‡]Ryerson University.

1 Introduction

The theory of comparative advantage identifies factor endowments as a key determinant of the pattern of trade. In particular, the theoretical prediction that countries endowed with larger stocks of human capital export relatively more in skill-intensive industries has received support in the literature, see Romalis (2004). In previous work (Bombardini et al., 2012b, henceforth BGP) we build on this line of research and argue that the second moment of the distribution of skills also determines comparative advantage. In particular, we find that the degree of skill dispersion has a quantitative impact on trade flows similar to that of the aggregate endowment of human capital. This paper presents a multi-country multi-industry model that shows how skill dispersion generates comparative advantage and thus provides a theoretical underpinning to the empirical evidence in BGP.

Why would the skill distribution matter for specialization and trade? We argue that industries vary in the degree of substitutability of workers' skills in the production process. In particular, some industries, such as aerospace or engine manufacturing, require completing long sequences of tasks and poor performance at any single stage greatly reduces the value of output. These are industries with low skill substitutability, or O-ring as in Kremer (1993), where efficiency improves when workers of similar skills are employed in every stage of production. In other industries, such as apparel, teamwork is relatively less important, as skills are more easily substitutable and poor performance in some task can be mitigated by superior performance in others.

We investigate theoretically whether countries with greater skill dispersion specialize in industries characterized by higher substitutability of skills across tasks. We build a model with many countries and many industries. Countries only differ in the distribution of skills in the labor force, while industries only differ in the degree of skill substitutability in the production process. At the micro level, our framework features worker heterogeneity and non-linear returns to scale; however, at the industry level, it is isomorphic to a perfectly competitive model with CRS and technological differences across many countries and industries, as in Costinot et al. (2012).

We develop the central argument in a benchmark framework in which skills are not observable by firms, adapted from Akerlof (1970). Skill unobservability is not an unusual assumption in the literature, see for example Grossman and Maggi (2000), Grossman (2004) and Helpman et al. (2010).[1] This modelling choice is motivated by the facts, documented in BGP, that (i) observable characteristics of workers, including age and education, account for a minor share of total variation in work-related literacy scores within countries; and (ii) groups of observationally similar workers exhibit very different degrees of skill dispersion across countries.[2] One immediate advantage of this approach is tractability in a setting with many countries and many sectors. Essentially, under skill unobservability, every worker is indifferent between employment in any two firms because there is a single (type-independent) wage that clears the labor market. In equilibrium, workers are randomly matched to firms and thus the latter inherit the skill distribution prevailing in the economy -i.e. the distribution of workers' unobservable skills in every firm is identical to the distribution of unobservable skills in the country.

The results in the benchmark model carry over with minor qualifications when cross-country differences in observable skills are introduced in the analysis. When firms can observe some components of skills, the equilibrium features sorting on observable skills by industry. As in the benchmark model, however, we show that unobservable skill dispersion fully determines the pattern of trade in

[1]In Grossman and Maggi (2000) and Helpman et al. (2010) firms can observe a component of individual skills before hiring workers. There remains, however, another component which firms never observe. We introduce a similar extension in Section 6. In Grossman (2004), as in this paper, individual skills are unobservable by firms and the production process features imperfect skill substitutability.

[2]Naturally, firms are likely to be able to observe the skills of their employees more accurately than econometricians. Therefore, these facts should be interpreted as playing a motivational role for our modelling approach, and not as precise descriptions of the degree of uncertainty that firms face regarding the skills of their employees.

equilibria with wage equalization across-countries. In light of this result, the benchmark model is best interpreted as a mechanism illustrating how the dispersion of skills among workers with otherwise identical observable characteristics affects comparative advantage. In the rest of the paper, we thus sometimes refer to unobservable skills as 'residual' skills.

We show that the interaction of skill dispersion and skill substitutability generates differences in output per worker across industries and countries, driving the pattern of international trade. The central result of the paper establishes conditions under which countries with a high dispersion of skills in the labor force are relatively more productive, and thus export relatively more, in sectors where skills are more easily substitutable across tasks. Interestingly, we also show that the effects of skill dispersion on output and specialization are isomorphic to the effects of technological differences in Ricardian models, such as Costinot et al. (2012). In this sense, our work has implications for the quantitative assessment of Ricardian comparative advantage since cross-country differences in measured total factor productivity can arise as the by-product of differences in the distribution of factor endowments.[3]

The theoretical framework developed in this paper can also be used for quantitative analysis. Using distributions of literacy scores from the International Adult Literacy Survey (IALS), we construct productivity differences attributable to skill dispersion. Cross-country differences in trade costs are calibrated to match bilateral trade flows at the industry level. We use the numerical counterpart of the model to carry out a quantitative assessment of the impact of skill dispersion on trade patterns. In particular, we measure the general equilibrium responses in industry-level trade when cross-country differences in skill dispersion are eliminated. These experiments, which

[3] Costinot et al. (2012) argue that their exogenous productivity differences aim to capture factors such as climate, infrastructure, and institutions, which affect the productivity of all producers in a given country and industry. To the extent that the distribution of human capital endowments – above and beyond observable credentials – is the product of a country's social structure and norms, our explanation would pertain to the institutional view of comparative advantage.

we run under a variety of technology parametrizations, suggest a significant role of skill dispersion on trade flows, with different effects on different countries.

This paper is related to recent theoretical research studying how skill distributions influence the pattern of trade. The hypothesis that skill dispersion may lead to specialization was first put forth by Grossman and Maggi (2000) in the context of a two-country, two-sector model, with competitive labor markets and constant returns to skills. They show that, when skills are fully observable, cross-country differences in skill dispersion do not generate comparative advantage when production technologies display convex isoquants in worker skills.[4] Gains from trade do not exist because, in equilibrium, workers of identical abilities are paired together, i.e. self-matching prevails in every industry, making skill dispersion irrelevant.[5] Grossman and Maggi (2000) also consider the case in which a portion of individual skills is unobservable. As in our paper, workers (with identical observable skills) are randomly matched to firms. Although they show that skill dispersion generates gains from trade when both industries have convex isoquants, they cannot determine the resulting *pattern* of trade across countries, which is the main focus of this paper.[6] In order to make progress, we propose a specification of production technologies that gives us sufficient tractability to handle a wide range of substitution elasticities, i.e. varying degrees of concavity or convexity of the isoquants, in a parametric way. In combination with random matching, this assumption allows us to characterize the variation in productivity differentials arising from skill dispersion, across multiple countries and industries, which is the crucial step in pinning down the pattern of trade.

[4] Supermodularity implies that the marginal product of any worker is increasing in the ability of the co-worker. Submodularity of the production function implies the opposite.

[5] In the case of observable skill dispersion, gains from trade are conditional on the existence of a supermodular sector, where 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. In this case, the country with a more dispersed skill distribution specializes in the submodular sector.

[6] When one sector displays convex isoquants and the other displays concave isoquants, they show that unobservable skill dispersion reinforces the pattern of trade induced by observable skill dispersion.

Interest in the relevance of skill distributions for trade is relatively recent. Ohnsorge and Trefler (2007) propose a Roy-type model with two-dimensional worker heterogeneity to 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.

The next section describes consumer preferences, production technologies and the labor market. Section 3 discusses how unobservable skill dispersion generates productivity differences across countries and industries, driving comparative advantage. Section 4 studies the optimization problem of individual firms. Section 5 analyzes the implications of skill dispersion for the pattern of international trade. Section 6 introduces observable skills in the model. Section 7 presents the quantitative analysis and our counterfactual results. The paper ends with some concluding remarks.

2 Setup

2.1 Preferences

Countries are denoted by a subscript $c \in \{1, ..., C\}$, which is dropped when it creates no ambiguity. Each country c is populated by a measure L_c of individuals. Utility of the representative consumer depends on the consumption of differentiated goods Q_i with $i \in \{1, ..., I\}$. The utility function U

is Cobb-Douglas:

$$\log U = \sum_{i=1}^{I} \alpha_i \log Q_i,$$

where $0 < \alpha_i < 1$, $\sum_i \alpha_i = 1$ and Q_i is an aggregate consumption index over a fixed set Ω_i of varieties of i. Preferences exhibit constant elasticity of substitution σ_i across varieties of any good.[7] As a result, total expenditure on variety ω of good i is:

$$x_i(\omega) = \left[\frac{p_i(\omega)}{P_i}\right]^{1-\sigma_i} \alpha_i E, \tag{1}$$

where E is aggregate expenditure, $p_i(\omega)$ is the price of variety ω of i, and P_i is the ideal CES price index of Q_i.

2.2 Production

Each variety ω in the differentiated industry i is produced under perfect competition and free entry. The typical firm producing ω has to incur a fixed start-up cost of f units of the numeraire good, to be specified below. The amount of output produced $y(\omega)$ depends on the skill level of each worker hired $a > 0$, the measure of workers hired h, the distribution of skills across workers $\widetilde{G}(a)$ and a random productivity shock $\varepsilon > 0$. The distribution of skills matters for production because we assume that skills are imperfectly substitutable. In particular, the production function depends on the degree of substitutability λ among workers' skills in that industry and takes the form:

$$y = \varepsilon \left(\int h a^{\lambda} d\widetilde{G}(a) \right)^{\frac{\gamma}{\lambda}} \text{ with } 0 < \gamma < \lambda. \tag{2}$$

[7] More specifically:

$$Q_i = \left[\int_{\omega \in \Omega_i} q_i(\omega)^{\frac{\sigma_i - 1}{\sigma_i}} d\omega \right]^{\frac{\sigma_i}{\sigma_i - 1}} \text{ with } \sigma_i > 1$$

where $q_i(\omega)$ is the quantity consumed of variety ω of good i.

The parameter $\lambda \in \Lambda$ measures the degree of skill substitutability. The elasticity of substitution among skill levels, for a fixed mass of workers h, is given by $\frac{1}{1-\lambda}$. The set $\Lambda \subset (0, 1]$ determines the admissible range of substitutability. A key assumption is that each industry i is characterized by a different value of λ in production, and therefore by a different degree of substitutability among workers' skills. Since λ is the only characteristic that distinguishes technology across industries,[8] in the remainder of the paper we drop subscript i and index industries by their corresponding parameter λ.

Note that when skills are perfectly substitutable (that is, $\lambda = 1$) output depends on the average skills of workers but not on higher moments of the skill distribution and skill dispersion does not generate comparative advantage. In what follows, we exploit this natural property of the model to enhance the tractability of the analysis. In particular, we show that, if there is an industry with no transport costs, $\lambda = 1$ and no intra-industry goods' heterogeneity, equilibria with incomplete specialization display wage equalization across countries.[9] Throughout we assume that $1 \in \Lambda$ and use Λ_{-1} to denote the set of industries with imperfect skill substitutability; i.e. $\Lambda_{-1} = \Lambda \setminus \{1\}$.

In order to obtain a clear-cut characterization of the impact of skill dispersion on bilateral trade flows at the industry level, we introduce intra-industry heterogeneity in industries with imperfect skill substitutability, as in Costinot et al. (2012). In particular, for industries $\lambda \in \Lambda_{-1}$, the Ricardian productivity shock $\varepsilon = \varepsilon(\omega)$ is assumed to be an i.i.d. variety-specific draw from a Fréchet distribution $F(\varepsilon) = \exp\left[-\varepsilon^{-\theta}\right]$ for $\varepsilon > 0$, with $\theta > 1$. As mentioned, we assume that the distribution of ε is degenerate when skills are perfectly substitutable and set $\varepsilon(\omega) = 1$ for all varieties in industry $\lambda = 1$. Importantly, these Ricardian shocks are identical across countries and therefore do not generate comparative advantage in the model.

[8]The parameter γ is constant across industries and does not affect comparative advantage.

[9]The industry $\lambda = 1$ therefore plays a similar role to a homogeneous sector in many trade models.

Two properties of this production function are worth mentioning. First, every worker has a positive marginal product: in particular, the marginal product of adding worker h of skill level a_h is $\varepsilon h^{\frac{\gamma}{\lambda}-1}\left(\int a^{\lambda}d\widetilde{G}\left(a\right)\right)^{\frac{\gamma}{\lambda}-1}\frac{\gamma}{\lambda}a_h^{\lambda}$. Second, the parameter γ controls the return to the mass of workers, given λ and the skill distribution. In this setting with price taking and fixed costs, decreasing marginal returns in every industry are necessary for the existence of an equilibrium with free entry.[10] In what follows, we restrict Λ to ensure $\gamma < \lambda$.[11]

2.3 Labor Market

We model an environment with informational asymmetries, adapted from Akerlof (1970). Workers are characterized by different skill levels. Skill is a continuous variable distributed in the population of country c according to a distribution function $G_c\left(a\right)$ with support $[a_{\min}, a_{\max}]$, where $a_{\max} > a_{\min} > 0$. Importantly, skills are unobservable by the firms. A competitive equilibrium thus consists of a single wage rate w_c, independent of worker types, that clears the labor market.

Since the opportunity cost of accepting employment is the same for every worker, decisions to participate in the labor market are independent of the wage rate and a worker's type. Moreover, since every firm offers the same wage, it follows that every worker is indifferent between any two jobs in the economy. As a result, there is no sorting between workers and firms in equilibrium. This implies that every firm, in any industry λ and country c, inherits the residual skill distribution in

[10] Note that the return to the mass of workers also depends on λ and therefore varies across industries. However, as the analysis in section 5 shows, this effect plays no role in shaping comparative advantage in equilibria with wage equalization across countries.

[11] Relative to Grossman and Maggi (2000), here the emphasis is on the degree of substitutability across sectors and not on whether the production function is submodular or supermodular. In fact, because we introduce sufficiently strong decreasing returns, our production function is always submodular. We should stress though that this is inconsequential for the case of unobservable skills. The crucial factor is the degree of substitutability across sectors, which, in the case of only two tasks would be easily represented graphically by the curvature of the isoquants. Notice that, differently from Grossman and Maggi (2000), here isoquants may be convex even with a submodular production function (this is possible when we remove the crucial assumption of constant returns to talent present in Grossman and Maggi (2000)). This topic is explored in Bombardini et al. (2012a).

the labor force:

$$\widetilde{G}(a) = G_c(a).$$

There are a few considerations to make about random matching. First, as will become clear below, random matching allows us to study the essential implications of unobservable skills for comparative advantage within a tractable general equilibrium framework. More sophisticated forms of imperfect matching would result if screening or signalling technologies were introduced into the model. In general, these extensions would come at the cost of analytic solutions. However, our analysis in section 6 can be interpreted as an attempt to partially relax these assumptions while keeping sufficient tractability. When worker skills are modelled as the product of two components, one of which is observable by firms, we can think of either (i) firms having access to a costless screening technology that allows them to identify some specific dimension of worker skills or, alternatively, (ii) workers that can credibly and costlessly signal some of their skills. Second, our focus is mainly on unobservable skills because this case has received relatively less attention in the literature and, as we show below, it can generate patterns of comparative advantage that are consistent with the empirical evidence in BGP. In section 6, we show that the results of this benchmark setup are robust when cross-country differences in observable skills are introduced in the analysis.

3 Skill Dispersion as Comparative Advantage

This section shows how cross-country differences in the distribution of residual skills generate comparative advantage. To facilitate the discussion we write the production function in (2) as $y = \varepsilon A(\lambda, c)\, h^{\frac{\gamma}{\lambda}}$ where the factor $A(\lambda, c)$ is defined as:

$$A(\lambda, c) \equiv \left(\int a^{\lambda} dG_c(a) \right)^{\frac{\gamma}{\lambda}}. \tag{3}$$

We refer to $A(\lambda, c)$ as 'fundamental productivity', although it is not the result of countries having access to different technologies. The magnitude of $A(\lambda, c)$ depends on the combination of a country-specific skill distribution and an industry-specific level of skill substitutability in production. Therefore, unlike the i.i.d. technological shocks captured by ε, fundamental productivity varies systematically across countries and industries and, as section 5 shows, it is the sole determinant of the pattern of comparative advantage in this model. The goal of this section is to understand how variation in $A(\lambda, c)$ is affected by the distribution of skills.[12]

Motivated by the empirical evidence presented in BGP, we explore the conditions under which countries with higher skill dispersion have a comparative advantage in sectors with higher skill substitutability. Property 1, stated below, provides a general condition for this pattern of comparative advantage to emerge from differences in the distribution of skills.[13]

For simplicity we compare countries that have the same average skills, but different dispersion. We order countries so that, if $c < c'$, then country c' is characterized by a skill distribution $G_{c'}(a)$ which is a mean-preserving spread of the skill distribution $G_c(a)$ in country c.

Property 1 $A(\lambda, c)$ is *strictly log-supermodular in* λ *and* c, *i.e. for any* λ, $\lambda' \in \Lambda$ and c, $c' \in \{1, ..., C\}$ such that $\lambda < \lambda'$ *and* $c < c'$*:*

$$\frac{A(\lambda, c')}{A(\lambda, c)} < \frac{A(\lambda', c')}{A(\lambda', c)}. \tag{4}$$

Property 1 states that the fundamental productivity of firms in countries with high skill dispersion will be relatively larger in high substitutability sectors.

A general result of this type cannot be established for all skill distributions. We therefore

[12] Although in a different context, Benabou (2005) analyzes the related issue of endogenous technological choice as a function of the degree of skill dispersion in the population.

[13] It is important to notice that the patterns of comparative advantage and trade are always fully determined in this model once skill dispersion and substitutability have pinned down an arbitrary $A(\lambda, c)$ function, regardless of whether Property 1 holds or not.

provide three different approaches to studying this problem. First, we show that comparative advantage can be established for any distribution with an upper-bounded support by appropriately restricting the admissible range of substitutability. Second, we perform comparative statics for specific distributions of skills that do not satisfy the sufficient conditions of the first approach. Finally, we use the empirical distributions of IALS scores in 19 countries to check numerically whether Property 1 holds when λ is allowed to vary over a wider range, relaxing the condition $\lambda \leq 1$.

Our first approach yields a general result for skill distributions with bounded supports, that relies on restricting the admissible range of skill substitutability.[14]

Proposition 1 *Property 1 holds under the following sufficient conditions:*

(i) the skill distribution in any country is supported on a bounded interval $[a_{\min}, a_{\max}]$*, where* $a_{\max} > a_{\min} > 0$*;*

(ii) the set of admissible substitutability satisfies $\Lambda \subset [\lambda_{\min}, 1]$ *where* $0 < \lambda_{\min} < 1$*, implicitly defined by*

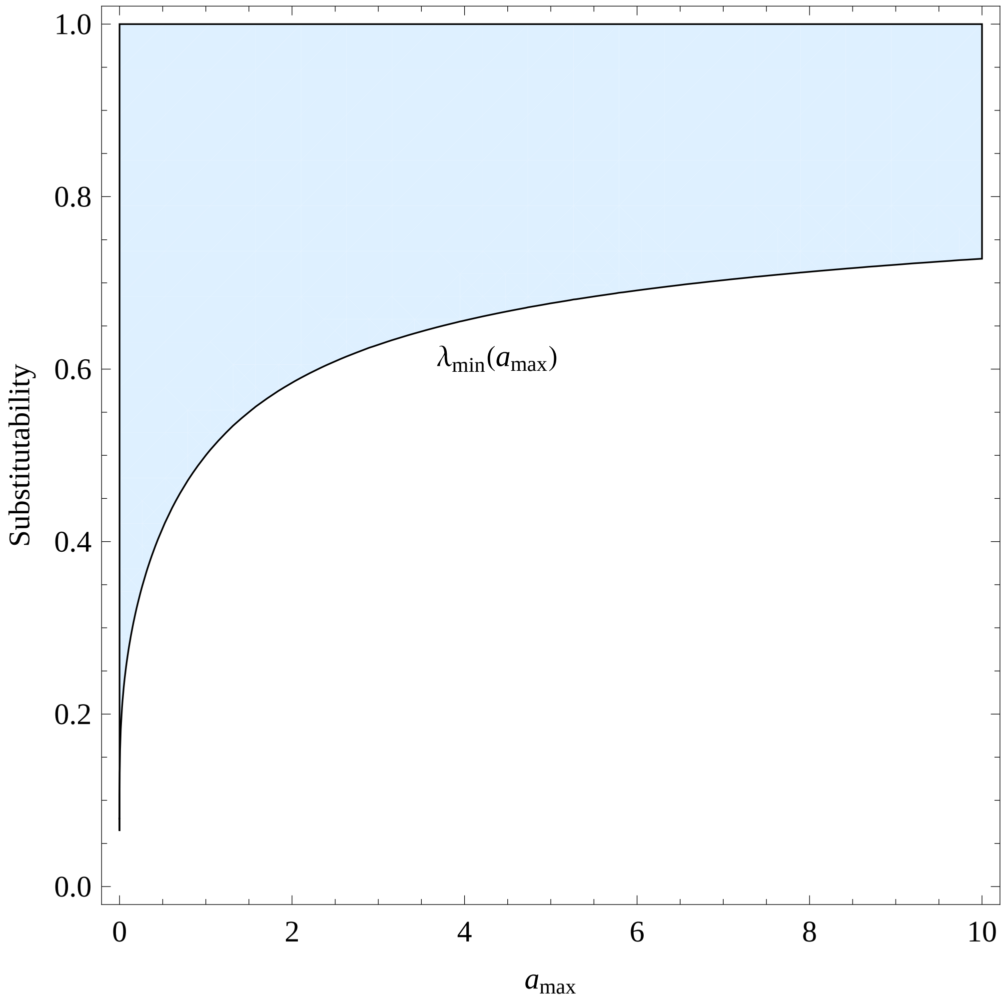
$$\log a_{\max} = \frac{2\lambda_{\min} - 1}{(1 - \lambda_{\min})\, \lambda_{\min}}.$$

Figure 1 displays the admissible range of substitutability for an arbitrary skill distribution $G_c(a)$ with bounded support. The lower bound on substitutability is a function of the upper bound of the skill distribution, $\lambda_{\min}(a_{\max})$, as defined in Proposition 1. As $a_{\max}$ increases, the lower bound on the admissible substitutability restriction becomes tighter, meaning that comparative advantage can be established for sets of industries with relatively higher skill substitutability in production. This means that Property 1 will not hold for some highly concave functions if the distribution of skills has a sufficiently large upper bound, a somewhat counterintuitive result which we further

[14] Imposing some upper bound on a is a reasonable restriction, as it only rules out the existence of infinitely productive workers.

discuss at the end of this section.

Figure 1: Admissible susbstitutability for bounded skill distributions



Our second approach to studying Property 1 relaxes the conditions on substitutability at the cost of concentrating on specific parametric distributions. We 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. In the appendix, we show that if skills are distributed according to a Log-normal distribution then, if country c and c' are characterized by skill distributions $G_c(a)$ and $G_{c'}(a)$ such that $G_{c'}(a)$ has equal mean and higher variance than $G_c(a)$ and if $\lambda < \lambda'$ then Property 1 holds. A similar result is established for the Pareto distribution. Besides these analytical results, we have also numerically computed the A's for several other distributions, including the uniform, triangular, gamma, beta and inverse Gaussian. For all these distributions, and for a wide range of parameters, it was not possible to find a numerical violation of the ranking in (4).

These analytical results provide sound support for (4) under the condition $\lambda \leq 1$. However, in anticipation of our quantitative exercises in section 7, where some sectors have a calibrated λ larger than one, we check numerically whether Property 1 is verified over a wider range of λ. In particular, we extend a numerical exercise originally presented in BGP, using the empirical distribution of IALS test scores in 19 countries to construct $A(\lambda, c)$ for an equally spaced grid of 50 different λ's defined over the $[0.1, 5]$ interval.[15] The data are described in Section 7. This grid covers the full range of λ's that we later employ in the quantitative section of the paper.[16] We obtain that, when averaging across country pairs, $\frac{A(\lambda,c')}{A(\lambda,c)}$ is increasing in λ at 96.9% of the grid points, where c' is a country with higher skill dispersion than c (as measured by the coefficient of variation of IALS scores). The corresponding exercise on the $[0.1, 1]$ interval yields a similar result, with relative productivity increasing at 97.3% of the grid points.

As a robustness check of the relevance of Property 1 we have investigated under what conditions a theoretical violation of the ranking could be engineered. One might expect that increasing differences in substitutability between production technologies would result in more, rather than less, relative advantage. Instead, Proposition 1 implies that Property 1 will not hold if some industries have low enough substitutability in production. To make sense of this result we rely on findings pertaining to choice theory and risk aversion, due to Ross (1981). He shows that, if one adopts the Arrow-Pratt definition of risk aversion (essentially a measure of concavity), there exist

[15] In contrast, BGP uses a grid of λ's defined over the $[0.1, 1]$ interval.

[16] The only exception is one large outlying λ in the parametrization where $\gamma = 0.7$.

lotteries such that a more risk-averse individual may be willing to pay less than a less risk-averse individual in order to avoid an increase in risk. The specific lottery employed as an example by Ross can be explained in the context of our model if one reinterprets $A(\lambda, c)^{\gamma}$ as certainty equivalent, i.e. the constant skill level that would make a firm as productive, and $f(a) = a^{\lambda}$ as utility function. Starting from a skill distribution where most of the workers have low skills, with the exception of few very talented individuals, consider adding a small amount of dispersion at the high skill level. A sector with lower λ may counterintuitively see its certainty equivalent drop by relatively less with such increase in dispersion. The intuition is that the increase in dispersion at very high skills happens in a range where $f(a)$ has relatively little curvature. Moreover, in order to avoid an increase in dispersion, a firm in a low λ sector would have to consider lowering its certainty equivalent in a relatively steep portion of $f(a)$ and could be less willing to do so (relative to a higher λ sector). Using the example by Ross it is possible to find a small set of parameters which generate such a violation of Property 1. Proposition 1 spells out a condition which rules out this phenomenon, by lowering the upper bound of skills to the point at which higher skill dispersion always results in large enough output losses.

4 The Firm's Problem

This section analyzes the optimal entry and employment choices of a typical firm producing variety ω of good λ in country c. Under perfect competition, the price $p_j(\omega, \lambda)$ that consumers in country j pay for variety ω of good λ is

$$p_j(\omega, \lambda) = \min_{1 \leq c \leq C} \{\mu_{cj}(\omega, \lambda)\},$$

where $\mu_{cj}(\omega, \lambda)$ is the unit cost of producing and delivering variety ω of good λ from country c to country j. Therefore, firms in country c find it profitable to start production of this variety only if they are the minimum cost suppliers in at least one of the C potential destinations. Conditioning on this, the representative firm maximizes profits by choosing total output and allocating it across the markets it decides to serve; that is, quantities to sell in the domestic and export markets. However, since firms are price takers, in equilibrium prices will be such that every destination *served* is equally profitable for firms producing a given variety.[17] Otherwise, the firm could reallocate output across destinations and increase profits. Let us denote this FOB price as $\widehat{p}_c(\omega, \lambda)$. The profit maximization problem of a producer of variety ω of good λ in country c can then be written as:

$$\max_h \ \widehat{p}_c(\omega, \lambda)\varepsilon(\omega)A\left(\lambda, c\right) h^{\frac{\gamma}{\lambda}} - w_c h - f.$$

Alternatively we can view the firm's problem as one of cost minimization. The combination of a fixed cost and decreasing marginal returns to h implies that firms face a U-shaped average cost curve. The minimum of this curve pins down the firm's employment $h_c^*(\lambda) = f\gamma / \left[w_c\left(\lambda - \gamma\right)\right]$ and efficient scale $y_c^*(\omega, \lambda) \equiv \varepsilon(\omega)A\left(\lambda, c\right) h_c^*(\lambda)^{\frac{\gamma}{\lambda}}$. The minimum average production cost, denoted $m_c^*(\omega, \lambda)$ is:

$$m_c^*(\omega, \lambda) = \kappa(\lambda)\frac{(w_c)^{\frac{\gamma}{\lambda}}}{\varepsilon(\omega)A\left(\lambda, c\right)},$$

where $\kappa(\lambda) > 0$ under decreasing marginal returns.[18]

Under perfect competition, profit maximizing firms equate output price, $\widehat{p}_c(\omega, \lambda)$, to marginal cost. In turn, free entry implies that every firm produces at the efficient scale, where marginal cost equals average cost, $m_c^*(\omega, \lambda)$. Therefore, assuming that there is a continuum of small potential

[17]In particular, producer prices faced by every firm producing a given variety in the same location will be equalized across destinations.

[18]Specifically, $\kappa(\lambda) \equiv f^{\frac{\lambda-\gamma}{\lambda}} \left(\lambda - \gamma\right)^{\frac{\gamma}{\lambda}-1} \gamma^{-\frac{\gamma}{\lambda}} \lambda$.

entrants, the industry's average cost function is perfectly elastic at the FOB price $\widehat{p}_c(\omega, \lambda)$ that satisfies the zero-profit condition $\widehat{p}_c(\omega, \lambda) = m_c^*(\omega, \lambda)$. As a result, the industry's unit cost of producing and delivering variety ω of good λ from country c to country j is:

$$\mu_{cj}(\omega, \lambda) = \tau_{cj} \cdot m_c^*(\omega, \lambda),$$

where $\tau_{cj} \geq 1$ is the iceberg trade cost from c to j. Two remarks are in order. First, we assume that the industry with perfect skill substitutability features no producer heterogeneity ($\varepsilon = 1$) and no transport costs. If there is positive production in this industry in all countries, then these assumptions imply that unit costs in this industry are equalized across countries and so are wage rates. In summary, these assumptions yield wage equalization in equilibria with incomplete specialization. Second, note that unit costs are inversely related to fundamental productivity. Therefore, under Property 1, countries with higher skill dispersion will, ceteris paribus, have relatively lower unit costs of producing varieties in industries with higher substitutability.

5 The Pattern of Trade

The analysis focuses on equilibria with incomplete specialization in production, in which every country produces positive output in every industry. As explained above, this leads to wage equalization across countries, thus we choose labor as the numeraire and let $w_c = 1$ for all c. As a result, the pattern of bilateral trade is indeterminate in the industry that exhibits perfect skill substitutability. Whenever skills are imperfectly substitutable, however, geographical specialization is determined by the locations of the minimum cost suppliers to each destination, which are functions of trade costs and both fundamental and random productivity shocks. We have seen how the interaction of skill dispersion and skill substitutability generates fundamental productivity differences across industries and countries. This section shows how the pattern of trade is in turn fully determined by fundamental productivity.

At the micro level our framework features worker heterogeneity and non-linear returns to scale; however, at the industry level, it is isomorphic to the perfectly competitive model with homogeneous workers, CRS and cross-country technological differences in Costinot et al. (2012).[19] We therefore follow their derivation of the pattern of trade.

As a first step, we provide an expression for bilateral exports in industries with imperfect skill substitutability. For $\lambda \in \Lambda_{-1}$, let $x_{cj}(\lambda) \equiv \sum_{\omega \in \Omega_{cj}(\lambda)} x_j(\omega, \lambda)$ denote the value of total exports from country c to country j in industry λ, where $\Omega_{cj}(\lambda) \equiv \left\{\omega \in \Omega \mid \mu_{cj}(\omega, \lambda) = \min_{1 \leq c' \leq C} \mu_{c'j}(\omega, \lambda)\right\}$ is the subset of varieties exported by country c to country j in industry λ. Then,

Lemma 2 *For any $\lambda \in \Lambda_{-1}$,*

$$x_{cj}(\lambda) = \frac{\left[(w_c)^{\frac{\gamma}{\lambda}} \tau_{cj}(\lambda)/A(\lambda, c)\right]^{-\theta}}{\sum_{c'=1}^{C} \left[(w_{c'})^{\frac{\gamma}{\lambda}} \tau_{c'j}(\lambda)/A(\lambda, c')\right]^{-\theta}} \alpha(\lambda) E_j. \tag{5}$$

Proof. The result follows from the proof of Lemma 1 in the online appendix of Costinot et al. (2012), using the fact that $\varepsilon(\omega) \overset{i.i.d.}{\sim}$ *Fréchet* implies that $z(\omega, \lambda, c) \equiv \varepsilon(\omega) A(\lambda, c) \overset{i.i.d.}{\sim}$ *Fréchet* with scale parameter $A(\lambda, c)$; i.e. $F(z) = \exp\left[-(z/A(\lambda, c))^{-\theta}\right]$, for all $z > 0$. Note that $\kappa(\lambda)$ varies by industry, but not by country, and therefore cancels out in (5). ■

Intuitively, in industries with imperfect skill substitutability, the location of the minimum cost supplier of any single variety is indeterminate because it depends on the random component of productivity. However, because these shocks are purely idiosyncratic, a country with a higher fundamental productivity will, ceteris paribus, capture a higher proportion of the industry's varieties

[19] In particular, production technologies in our model can be equivalently interpreted as depending on h homogeneous workers, with fundamental productivity equal to $A(\lambda, c)$ in industry λ and country c. Furthermore, as shown in the previous section, under free entry, industry unit costs are constant, as with CRS technologies.

imported by consumers in any destination.

In order to isolate the role of skill dispersion in shaping comparative advantage, assume that there are no other exogenous sources of relative cost differences across exporters, i.e. $\tau_{cj}(\lambda) = \tau_{cj} \cdot \tau_j(\lambda)$ for $\lambda \in \Lambda_{-1}$. Then, under wage equalization, Lemma 2 implies that for any importer j and any pair of exporters c and c', the ranking of relative fundamental productivities fully determines the ranking of relative exports to a given market j. That is, for any two industries λ, $\lambda' \in \Lambda_{-1}$:

$$\frac{A(\lambda, c')}{A(\lambda, c)} \leq \frac{A(\lambda', c')}{A(\lambda', c)} \Leftrightarrow \frac{x_{c'j}(\lambda)}{x_{cj}(\lambda)} \leq \frac{x_{c'j}(\lambda')}{x_{cj}(\lambda')}. \tag{6}$$

In our framework, $A(\lambda, c)$ is log-supermodular in λ and c, when countries are ordered according to increasing skill dispersion. This yields the main result of the paper linking skill dispersion, comparative advantage and trade flows.

Proposition 3 *Consider an equilibrium with incomplete specialization in production, in which skill dispersion is the only source of relative cost differences across countries, i.e.* $\tau_{cj}(\lambda) = \tau_{cj} \cdot \tau_j(\lambda)$ *for* $\lambda \in \Lambda_{-1}$. *Then, 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 higher substitutability* $\lambda \in \Lambda_{-1}$.

Proof. It follows immediately from Property 1 and the ranking of relative exports in (6). ■

We now briefly explain how to solve for the remaining endogenous variables of the model. The mass of firms in industries with imperfect skill substitutability, denoted $M_c(\lambda)$, can be determined from the market clearing condition that total expenditure on a given industry's varieties equals the sum of the revenues of domestic producers of these varieties. Under wage equalization, equilibrium firm revenues are $h^*(\lambda) + f$, obtained from section 4. Therefore, the market clearing condition in industry $\lambda \in \Lambda_{-1}$ can be written as $\sum_j x_{cj}(\lambda) = M_c(\lambda)[h^*(\lambda) + f]$, which, together with the solutions for trade flows (5), can be used to solve for $M_c(\lambda)$ as a function of expenditure levels E_c. The allocation of labor in this industry, denoted $L_c(\lambda)$, satisfies $L_c(\lambda) = M_c(\lambda)[h^*(\lambda) + f]$. In turn, the allocation of labor in the industry with perfect skill substitutability is $L_c(1) = L_c - \sum_{\lambda \in \Lambda_{-1}} L_c(\lambda)$, ensuring full employment.[20] From this, we can compute the mass of firms and total exports to the rest of the world, $M_c(1)$ and $\sum_j x_{cj}(1)$, using the solution for $h^*(1)$ and imposing market clearing. Finally, aggregate expenditure levels are obtained by imposing the equality of income and expenditure in every country, so that $E_c = \sum_{\lambda \in \Lambda} \sum_j x_{cj}(\lambda)$ for all $c = 1, ..., C$.

6 Observable skills

This section extends the model to account for observable skills. The equilibrium features sorting by industry on observable skills. As in the benchmark model, however, unobservable skill dispersion fully determines the pattern of trade in equilibria with wage equalization.

Suppose there are N types of workers, indexed by $n = \{1, ..., N\}$. The input of a worker i of type n is now the product of observable and unobservable components of ability, denoted q_n and a_{ni} respectively. By definition, every worker of type n is endowed with the same skill q_n, which is observable to both firms and workers.[21] Component a_{ni}, however, can vary across workers of the same type. As before, we assume a_{ni} is unobservable to firms. The structure of the labor market is therefore unchanged, except that the equilibrium now requires determining N type-specific market clearing wages in each country, denoted $w_{n,c}$.

[20] Incomplete specialization requires $L_c(1) > 0$ for all c, which can be ensured by appropriate choice of labor endowments and expenditure shares across countries.

[21] One possible interpretation is that n indexes education-experience cells and q_n is the expected value of the (log) ability of workers belonging to cell n. This interpretation would require the normalization $E_n[\log(a_{ni})] = 0$ for all n.

In this context, the technology (2) of a firm in industry λ naturally extends to:

$$y = \varepsilon \left[\sum_{n=1}^{N} \int_0^{h_n} (q_n a_{ni})^{\lambda} \, di \right]^{\frac{\gamma}{\lambda}}, \tag{7}$$

where h_n is the number of workers of type n employed.[22] Note that if $N = q_N = 1$, we obtain (2).

Let $G_{n,c}(a)$ be the distribution of unobserved skills of type n in country c. To keep matters as close to the benchmark model as possible, we assume that unobservable skills are identically distributed across skill types, i.e. $G_{n,c}(a) = G_c(a)$ for every n, as in Grossman and Maggi (2000).[23] Without loss of generality, we further restrict the analysis to two observable types $n = \{1, 2\}$, with $q_1 < q_2$ and refer to type-1 and type-2 workers as unskilled and skilled, respectively. Together with random matching on unobservable skills, these assumptions allow us to express output as $y = \varepsilon A(\lambda, c) \left[h_1 q_1^{\lambda} + h_2 q_2^{\lambda} \right]^{\gamma/\lambda}$, where $A(\lambda, c) \equiv \left(\int a^{\lambda} dG_c(a) \right)^{\gamma/\lambda}$ is defined exactly as in section 3.

Dropping the country and industry indices momentarily in order to lighten notation, the firm chooses h_1 and h_2 in order to maximize $\widehat{p}y - w_1 h_1 - w_2 h_2 - f$. The first-order condition for h_n, $n = \{1, 2\}$, is $\widehat{p}\varepsilon A \left[h_1 q_1^{\lambda} + h_2 q_2^{\lambda} \right]^{\gamma/\lambda - 1} q_n^{\lambda} (\gamma/\lambda) \leq w_n$, with equality if $h_n > 0$.[24] Note that both first-order conditions hold with equality only in a threshold industry λ_c^*, satisfying

$$\left(\frac{q_2}{q_1} \right)^{\lambda_c^*} = \frac{w_{2,c}}{w_{1,c}}. \tag{8}$$

Producers in industry λ_c^* are indifferent between both types of workers so, for expositional purposes,

[22] Recall that in the benchmark model without observable skills, there are no skill-specific tasks, i.e. total output is simply the integral of individual inputs a_i^{λ} across employees, regardless of a_i. By applying the same concept in the present context, we obtain (7).

[23] Allowing for 'heteroskedasticity' in unobservable skills across countries implies that unobserved skill dispersion for each type of worker operates as an independent source of comparative advantage. Although potentially relevant in an empirical analysis, this extension would make the analysis cumbersome, without adding further insight.

[24] The production function is quasiconcave in (h_1, h_2), since we still assume $0 < \gamma < \lambda$. The first-order conditions are thus necessary and sufficient for a maximum.

we focus on equilibria such that industry λ_c^* hires skilled workers.[25] Since $q_1 < q_2$, firms in industries $\lambda \geq \lambda_c^*$ exclusively employ skilled workers, while firms in industries $\lambda < \lambda_c^*$ exclusively employ unskilled workers. Therefore, the equilibrium displays sorting on observable skills and random matching on unobservable skills.[26] Although the present model considers only two factors, the result is reminiscent of the sorting pattern obtained in Costinot and Vogel (2010) with a continuum of sectors and factors. The key similarity lies in the assumptions of log-supermodularity of the sector-skill specific productivity in their paper and of q^λ in this paper.

Since firms hire a single type of workers, the solutions to the profit maximization problem are very similar to the benchmark model. In particular, under free-entry, it is straightforward to show that the industry's unit cost of producing and delivering variety ω from country c to country j, is $\mu_{cj}(\omega, \lambda) = \tau_{cj}(\lambda)\kappa_{n_\lambda,c}\left(w_{n_\lambda,c}\right)^{\gamma/\lambda} / \left[\varepsilon(\omega) A\left(\lambda, c\right)\right]$, where subscript n_λ indicates the type of workers (optimally) employed in industry λ; i.e. $n_\lambda = 1$ if $\lambda < \lambda_c^*$ and $n_\lambda = 2$ otherwise. In turn, $\kappa_{n_\lambda,c} \equiv f^{\frac{\lambda-\gamma}{\lambda}}\left(\lambda-\gamma\right)^{\frac{\gamma}{\lambda}-1}\gamma^{-\frac{\gamma}{\lambda}}\lambda\left(q_{n_\lambda}\right)^{-\gamma}$ is a function of parameters that vary across industries and, possibly, countries.[27] Having determined unit costs, total bilateral exports at the industry level are derived as in Lemma 2:

$$x_{cj}(\lambda) = \frac{\left[\kappa_{n_\lambda,c}\left(w_{n_\lambda,c}\right)^{\frac{\gamma}{\lambda}}\tau_{cj}(\lambda)/A\left(\lambda,c\right)\right]^{-\theta}}{\sum_{c'=1}^{C}\left[\kappa_{n_\lambda,c'}\left(w_{n_\lambda,c'}\right)^{\frac{\gamma}{\lambda}}\tau_{c'j}(\lambda)/A\left(\lambda,c'\right)\right]^{-\theta}}\alpha(\lambda)E_j. \tag{9}$$

As in the benchmark model, we can now establish the pattern of trade for the class of equilibria featuring wage equalization across countries. In this case, $w_{n,c} = w_n$ for all c together with (8) imply that the assignment of observable skills to industries is the same in every country; i.e. $\lambda_c^* = \lambda^*$

[25] We assume $\lambda_c^* \in \Lambda$. This will be the case when, for example, Λ is an interval in the real line.

[26] Intuitively, sorting is driven by the assumption of no type-specific tasks embedded in (7) together with the fact that returns to individual skills increase in λ.

[27] $\kappa_{n_\lambda,c}$ will vary across countries as long as q_{n_λ} does. This can occur when the threshold industry λ^* differs across countries so that, in a given industry λ, countries employ workers of different types.

for all c. In the absence of other sources of comparative advantage, it is easy to verify that equation (6) holds and thus relative trade flows are fully determined by unobservable skill dispersion. The next result immediately follows from these observations, along the lines of Proposition 3.

Proposition 4 *Consider an equilibrium with wage equalization across countries in the model with multiple observable skills. Assume that transport costs are not a source of relative cost differences across countries, i.e.* $\tau_{cj}(\lambda) = \tau_{cj} \cdot \tau_j(\lambda)$ *for* $\lambda \in \Lambda$*. Then, under Property 1, a country with relatively higher dispersion of unobservable skills has a comparative advantage, and therefore exports relatively more to any destination, in sectors with higher substitutability.*

We note that wage equalization can be ensured in a similar way as in the benchmark model and, more generally, as in standard Heckscher-Ohlin theory. That is, by an appropriate choice of endowments and expenditure shares such that every country is incompletely specialized in a set of N, freely traded, homogeneous goods, each employing a different type of workers.

Finally, we would like to remark that in our model, because of the assumption of wage equalization across countries, observable skill dispersion is not a source of comparative advantage among differentiated goods.[28] However, the analysis in this section is not intended to provide definitive conclusions regarding the effect of observable skill dispersion on trade flows, but rather to show that observable skills can be incorporated into our model in a natural way, without necessarily changing our core results regarding unobservable skill dispersion.

7 Quantitative analysis

Next, we return to the model outlined in Sections 2-5 to perform a quantitative assessment of the effect of skill dispersion on trade flows. For this exercise we consider a set of 63 manufacturing

[28]Standard Heckscher-Ohlin comparative advantage may emerge in the homogeneous good sectors, in which skills are perfectly substitutable. We do not focus on this well established result.

industries and 18 countries for which we can observe both bilateral trade flows and the distribution of workers' skills.[29]

7.1 Eliminating cross-country differences in skill dispersion

We undertake a set of counterfactual experiments aimed at inferring how changes in comparative advantage due to skill dispersion alter trade flows. To do so we fix a reference country and study the general equilibrium responses of its trade flows with the rest of the world (ROW) after removing all comparative advantage due to skill dispersion. More specifically, after choosing a reference country 0, we set each industry's $A(\lambda, c) = A(\lambda, 0)$ for all the other countries c; then we solve for the counterfactual equilibrium and we compare both exports and imports of country 0 to the ROW, industry by industry, against those in the benchmark equilibrium for which the $A(\lambda, c)$ are constructed from data.[30] Consistently with the model, and to minimize confounding effects, we perform the quantitative analysis in a setting in which (i) wages are equalized in equilibrium, (ii) consumers have identical preferences, i.e. $\alpha(\lambda, c) = \alpha(\lambda)$ for all c, and (iii) skill dispersion is the only source of comparative advantage. The third assumption requires $\tau_{cj}(\lambda) = \tau_{cj} \cdot \tau_j(\lambda)$ for every c, j and λ. Below we describe all the experiments, then we discuss their implementation.

Change in exports, by industry. Under the assumptions (i)-(iii) above, we use equation (5) to obtain an expression for the exports of country 0 to the ROW in industry λ. Counterfactual

[29] Our sample of industries and countries is identical to the one studied in BGP, with the exclusion of Chile. We are unable to measure Chilean domestic absorption rates from OECD-STAN data and thus we drop this country from the sample (see Appendix).

[30] These counterfactual experiments are similar to those presented in Costinot et al. (2012) to study Ricardian comparative advantage, with two important differences: (i) we allow for counterfactual changes in aggregate expenditures and, (ii) we also report results at a more disaggregated (industry) level.

equilibrium exports, relative to those in the benchmark equilibrium, are

$$\frac{\sum_j x'_{0j}(\lambda)}{\sum_j x_{0j}(\lambda)} = \frac{\sum_j \frac{\tau_{0j}^{-\theta}}{\sum_{c'} \tau_{c'j}^{-\theta}} E'_j}{\sum_j \frac{[\tau_{0j}/A(\lambda,0)]^{-\theta}}{\sum_{c'} \left[\tau_{c'j}/A(\lambda,c')\right]^{-\theta}} E_j}, \tag{10}$$

where primes denote counterfactual variables. For example, $x_{0j}(\lambda)$ and $x'_{0j}(\lambda)$ are exports from 0 to j in industry λ in, respectively, the benchmark and counterfactual equilibria. Further below we explain how we compute equation (10) using data and, then, we use it to assess the impact of skill dispersion on each industry's exports.

It is also interesting to study how this ratio changes across industries. Consider any two industries λ_1 and λ_2; then it is easy to show that

$$\frac{\sum_j x'_{0j}(\lambda_1)}{\sum_j x_{0j}(\lambda_1)} \leq \frac{\sum_j x'_{0j}(\lambda_2)}{\sum_j x_{0j}(\lambda_2)} \tag{11}$$

if and only if

$$\sum_j \frac{[\tau_{0j}/A(\lambda_2,0)]^{-\theta}}{\sum_{c'} \left[\tau_{c'j}/A(\lambda_2,c')\right]^{-\theta}} E_j \leq \sum_j \frac{[\tau_{0j}/A(\lambda_1,0)]^{-\theta}}{\sum_{c'} \left[\tau_{c'j}/A(\lambda_1,c')\right]^{-\theta}} E_j. \tag{12}$$

The last expression suggests that, following a loss of comparative advantage, the decline of country 0's relative exports is larger in industries in which this country had a relatively higher fundamental productivity advantage with respect to the ROW. This can be seen most clearly when trade costs are symmetric, i.e. $\tau_{c'j} = \tau_{0j}$ for all c' and j. In this case (12) simplifies to

$$\frac{A(\lambda_2,0)^{\theta}}{\sum_{c'} A(\lambda_2,c')^{\theta}} \leq \frac{A(\lambda_1,0)^{\theta}}{\sum_{c'} A(\lambda_1,c')^{\theta}}. \tag{13}$$

If the above inequality holds, country 0 has a comparative advantage in industry λ_1 relative to the ROW and, using equation 11, its counterfactual exports to the ROW should drop relatively more in industry λ_1. For example, if we take a high skill dispersion country (e.g. the US) as country 0, its fundamental productivities in the initial equilibrium would be relatively larger (relative to the ROW) in industries with higher λ. It follows that, when plotting $\sum_j x'_{0j}(\lambda)/\sum_j x_{0j}(\lambda)$ as a function of λ, we expect to observe, on average, a negatively sloping curve. The steepness of this curve indicates the strength of CA generated by skill dispersion. In contrast, if the reference country has low skill dispersion (e.g. Germany), that slope is expected to be positive. In the numerical experiments below we do not assume symmetric trade costs and, therefore, we also allow for some confounding effects.

Change in imports, by industry. Similarly, the change in imports to country 0 from the ROW in the counterfactual equilibrium can be expressed as

$$\frac{\sum_{c\neq 0} x'_{c0}(\lambda)}{\sum_{c\neq 0} x_{c0}(\lambda)} = \frac{1+\frac{[\tau_{00}(\lambda)/A(\lambda,0)]^{-\theta}}{\sum_{c\neq 0}[\tau_{c0}(\lambda)/A(\lambda,c)]^{-\theta}}}{1+\frac{\tau_{00}(\lambda)^{-\theta}}{\sum_{c\neq 0}\tau_{c0}(\lambda)^{-\theta}}}\frac{E'_0}{E_0} \tag{14}$$

Comparing such changes in any two industries λ_1 and λ_2 under the hypothesis that trade costs do not generate comparative advantage, one obtains

$$\frac{\sum_c x'_{c0}(\lambda_1)}{\sum_c x_{c0}(\lambda_1)} \leq \frac{\sum_c x'_{c0}(\lambda_2)}{\sum_c x_{c0}(\lambda_2)}$$

if and only if

$$\frac{A\left(\lambda_1,0\right)^{\theta}}{\sum_{c\neq 0}\left[\tau_{c0}/A\left(\lambda_1,c\right)\right]^{-\theta}} \leq \frac{A\left(\lambda_2,0\right)^{\theta}}{\sum_{c\neq 0}\left[\tau_{c0}/A\left(\lambda_2,c\right)\right]^{-\theta}}. \tag{15}$$

This condition is similar to (13), with the difference due to the presence of trade costs. Simply put, the condition suggests that if we take a high skill dispersion country (US) as country 0, we expect its fundamental productivities to be relatively larger (compared to the ROW) in industries with higher λ. Equation (15) suggests that, when plotting $\sum_c x'_{c0}(\lambda)/\sum_c x_{c0}(\lambda)$ as a function of λ, we should observe a positive slope. This is quite intuitive since a positive slope means that, when the US loses its comparative advantage in high substitutability industries, the ROW exports relatively more to the US in those industries.[31] Like before, condition (15) is informative about the slope of relative changes in imports across industries, while the relative change within any industry is summarized by (14).

7.2 Calibration

To perform the quantitative exercises described above one needs to assign values to a subset of model parameters. Here we overview these parameters. Estimates and further details are deferred to the Appendix. Once parameter values are obtained, we solve for bilateral trade flows and aggregate expenditures in both the benchmark and counterfactual equilibria, as outlined at the end of Section 5. This allows us to compute the numerical counterparts of expressions (10) and (14) for our sample of countries and industries.

Technology. The crucial technology parameters in equation (2) are the industry-specific λ's, which determine the elasticity of substitution among workers' skills, and γ, which controls the degree of diminishing returns to labor, for a given λ.[32] From the analysis in section 4, it is straightforward to check that the wage bill as a fraction of firm revenue is equal to γ/λ, in industry λ.[33] We therefore use industry-level data to compute the share of labor in value added to approximate γ/λ. The source of the data is the 2002 Economic Census and the variables employed are total compensation[34] and value added at the NAICS 6-digit level, aggregated to the 63 industries that

[31] In (15) the transport costs τ generate a different weighting of the A productivities for ROW and this may, in principle, weaken somewhat the result.

[32] The empirical literature has so far paid little attention to the estimation of individual skills' substitutability in production. Most studies of heterogeneous labor demand have instead focused on coarser classifications, such young vs old, male vs female, college vs no college. A discussion of relevant issues can be found in the original survey by Hamermesh (1986).

[33] More specifically, for any variety ω of good λ, $(w_c h_c^*(\lambda)) / (\widehat{p}_c(\omega,\lambda) y_c^*(\omega,\lambda)) = \gamma/\lambda$.

[34] Results are unaffected when employing total payroll as a measure of wage bill.

are constructed in BGP.

For a given γ, the procedure described above allows us to pin down λ in each of the 63 industries considered. However, implementing our quantitative exercises still requires assigning a value for γ. To gauge the robustness of our results, we parametrize the technology under alternative values of γ and then simply invert the estimated γ/λ ratios to derive the associated sets of λ's. In the context of our model, it is reasonable to impose a value of 1 as an upper-bound on the range of γ. Otherwise, the calibrated lambdas would be larger than 1 in every industry, a highly implausible scenario in which production technologies in every manufacturing industry have concave isoquants.[35] Therefore, we posit that $\gamma \in (0, 1)$ and experiment with three values. Namely, we consider a benchmark parametrization of $\gamma = 0.5$ (the middle value of the unit range) but we also report results for γ equal to 0.3 and 0.7.[36] More details about the technology parameterization are presented in the Appendix.

Skill distributions. To approximate a country's skill distribution we use the distribution of IALS scores (see BGP for a discussion). We do this for 18 countries.[37] Together with the estimated lambdas, we use (3) to compute fundamental productivity measures $A(\lambda, c)$ for each country c in industry λ.

Transportation costs. The transportation costs $\tau_{cj}(\lambda)$ have an important role in the numerical analysis, capturing all 'residual' sources of unmodeled heterogeneity that affect trade between each country pair (c, j) in industry λ. We estimate $\tau_{cj}(\lambda)$ in two steps: first, we use the OECD-

[35]The ONET database studied in BGP reveals that scores of indicators of low skill substitutability, such as the importance of teamwork, are high in many US manufacturing industries.

[36]In a different context, when aggregating across 67 different occupations, Hsieh et al. (2013) acknowledge the lack of information about the elasticity of substitution and simply choose a value of 3, doing robustness checks around this benchmark value. We take an approach similar to theirs. While their setup is different from ours, their assumed elasticity is well within the range we allow for in our application.

[37]The results presented here are produced using 'residual' IALS obtained by regressing IALS scores on individual-level observable characteristics, as documented in BGP. We also perform the same experiments using raw IALS scores, and the results are qualitatively and quantitatively similar. In line with the fact that most (roughly 2/3) of IALS variation comes from the residual component, the experiments with raw scores exhibit effects on trade flows which are between 15 and 30 percent stronger. More details are available from the authors.

STAN bilateral trade data at the industry level. The $\tau_{cj}(\lambda)$ are chosen to replicate bilateral trade flows by industry; this results in a set of $18 \times 17 \times 63$ elements corresponding to all country-industry triplets. In the Appendix we describe how these trade costs are derived as a function of bilateral trade flows in each industry. Next, we decompose (through an OLS projection) each $\tau_{cj}(\lambda)$ into two multiplicative components: (1) a country-pair effect; (2) a country-industry component. These are the two terms $\hat{\tau}_{cj}$ and $\hat{\tau}_j(\lambda)$, which we then use in the simulations. The decomposition $\hat{\tau}_{cj}(\lambda) = \hat{\tau}_{cj}\hat{\tau}_j(\lambda)$ guarantees that transport costs are not a source of comparative advantage.[38]

Idiosyncratic variety shocks. Parameter θ, characterizing the Fréchet distribution of the Ricardian productivity shocks $\varepsilon(\omega)$, is set to 6.53, following Costinot et al. (2012).

Preferences. To solve for an equilibrium we only need to set the expenditure share parameters $\alpha(\lambda)$, which we assume identical in all countries. Details of how we calibrate these consumption shares are in the Appendix.

7.3 Results

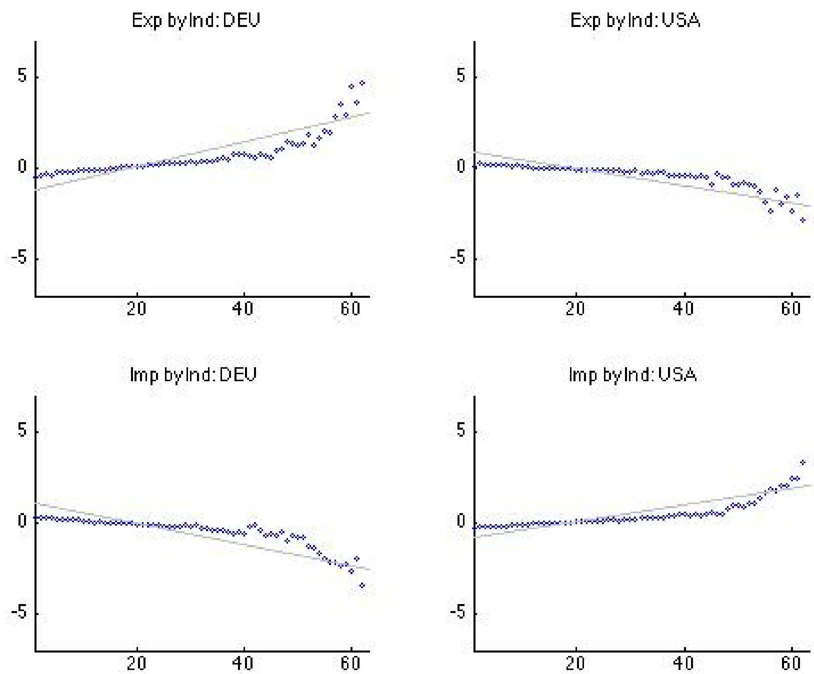
Shutting down differences in relative productivities guarantees the absence of comparative advantage due to skill dispersion. However, the choice of a reference country is important in these experiments because it affects the level of fundamental productivities' in different industries. For this reason we compute a battery of 18 counterfactual experiments, allowing for a different reference country in each one of them. In what follows we plot graphs based on technology estimates obtained using the wage shares measured through total compensation and the benchmark parametrization of $\gamma = 0.5$. In the appendix we report results for alternative values of γ. Figure 2 plots the percentage changes in imports and exports, setting Germany and the US as reference countries.[39]

[38] Using this decomposition of the transport costs means that the simulations do not exactly match bilateral trade flows in each industry.

[39] Changes are computed like in equations 10 and 14. Exports of a country to itself are not included.

While both countries are at similar stages of development, they are at opposite ends of the skill dispersion range. Each square reports the proportional (percentage) change in exports or imports for the reference country, plotted over the 63 differentiated goods' sectors: industries on the horizontal axis are ordered in terms of increasing λ (i.e. skills' substitutability in production). In each graph we overlay a linear fit, to emphasize the direction and steepness of the changes as the skill substitutability parameter λ becomes larger.

Figure 2: Changes in trade flows for Germany and US: $\gamma = 0.5$

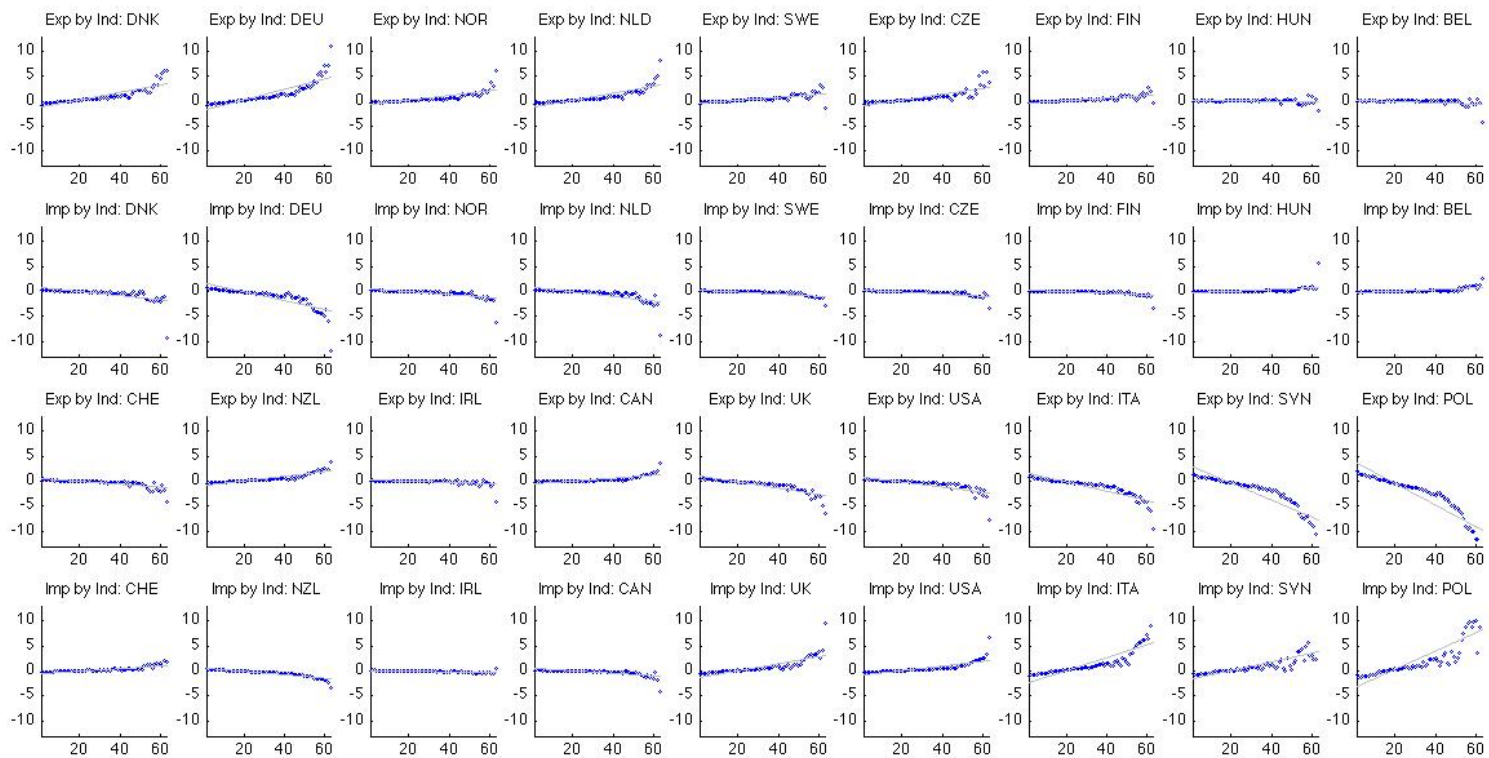


Note: the horizontal axis reports the 63 industries in order of increasing λ. The vertical axis reports the percentage change in trade.

These plots reveal that the expected patterns of change in exports and imports are borne out by the experiments. Export changes for the highest skill dispersion country (the US) show a decreasing pattern as we move towards higher λ industries, while the slope of the same curve for Germany exhibits a positive slope. The opposite is true for imports, confirming that the loss of comparative advantage is reflected more strongly in the sectors which most benefit from a country's skill dispersion in the benchmark equilibrium.

This quantitative result is further illustrated by Figure 3 which reports trade changes for the full set of 18 reference countries. The first two rows show, respectively, proportional changes in exports and imports by industry for the nine countries with the lowest skill dispersion. The last two rows do the same for the nine countries with the highest skill dispersion. In each row skill dispersion is increasing as countries change from left to right. From these plots it is apparent that, when we equalize the $A(\lambda, c)$ using reference countries at the lower end of the skill dispersion range (like Denmark and other Nordic countries) export changes become larger as we move right towards industries with higher λ, while the opposite is true for import changes. In contrast, when the reference country in the counterfactual experiments have high skill dispersion like the US, the UK or Italy, changes in trade flows by industry have slopes which are the reversed, as these countries lose their comparative advantage in high λ industries. Interestingly, for countries in the middle of the skill dispersion range (e.g., Hungary, Belgium, Switzerland) there is little or no slope in the curves plotting export and import changes by industry: these are countries where comparative advantage due to skill dispersion is generally weaker in the benchmark equilibrium. Beyond the evolution of trade flow changes across industries, these experiments can also shed some light on the relative intensity of the changes. To this purpose we compute for each country the average of the absolute percentage changes in imports and exports. Table 1 reports these average absolute changes. The table also reports the (raw, not absolute) percentage change experienced in the industries corresponding to the 10th and 90th percentiles in the distribution of estimated λ's. The latter values clearly illustrate the range of variation in percentage changes.

Figure 3: Changes in trade flows across industries : $\gamma = 0.5$



Note: the horizontal axis reports the 63 industries in order of increasing λ. The vertical axis reports the percentage change in exports or imports.

Table 1: Changes in trade flows ($\gamma = 0.5$)

country	exports			imports		
	(1)	(2)	(3)	(4)	(5)	(6)
	average	10%	90%	average	10%	90%
DNK	0.93	-0.25	3.81	0.50	0.12	-1.08
NLD	0.92	-0.23	2.71	0.60	0.16	-1.92
NOR	0.59	-0.17	1.30	0.37	0.10	-1.21
FIN	0.48	-0.13	1.70	0.37	0.09	-1.04
DEU	1.03	-0.24	3.54	0.82	0.22	-2.34
SWE	0.46	-0.09	1.18	0.20	0.07	-0.60
CZE	0.56	-0.04	3.32	0.17	0.04	-0.44
BEL	0.15	-0.01	0.18	0.09	-0.02	0.32
HUN	0.22	0.01	1.54	0.04	-0.01	0.10
CHE	0.20	0.04	-0.54	0.15	-0.06	0.59
CAN	0.55	-0.18	1.85	0.40	0.18	-1.84
NZL	0.31	-0.04	0.48	0.26	0.02	-0.47
IRL	0.28	0.01	-0.41	0.08	0.02	-0.06
UK	0.54	0.22	-1.82	0.57	-0.20	2.05
USA	0.67	0.16	-1.96	0.68	-0.18	2.08
ITA	0.97	0.26	-2.00	1.44	-0.31	3.68
SVN	1.67	0.42	-4.15	0.84	-0.35	3.29
POL	2.71	0.75	-7.45	2.34	-0.73	6.46

Notes:

– Columns (1) and (3) report the average of absolute percentage changes in trade; columns (2), (4), (5) and (6) report raw (not absolute) percentage changes in trade for the industries corresponding to the 10^{th} and 90^{th} percentile in the distribution of estimated λ's (skill substitutability).

– The countries are ranked by skill dispersion.

– 10^{th} percentile and 90^{th} percentile are the 6th and 58th industries ranked by λ's.

These results confirm that the largest changes are experienced by countries which lie at the opposite ends of the skill dispersion range. Changes are also more pronounced for sectors at the extremes of the λ range, in particular for the highest values of λ. Moreover, looking at the changes in industries at the extremes of the substitutability scale it is apparent that the signs of changes are inverted for countries which lie at different ends of the skill dispersion range. Finally, to summarize the individual country results, we take a weighted average of the absolute changes in exports (imports) across countries, where weights correspond to exports (imports) shares in the benchmark. This results in an average change of 1.63% in exports and 1.35% in imports. These findings are in line with our theoretical analysis and suggest a non-trivial role for skill dispersion in determining trade flows. In the Appendix we also report results for simulations in which we solve and perturb the benchmark under different technology parameters, in particular γ. Setting γ to, respectively, 0.3 and 0.7 , the weighted average absolute change in exports are, respectively, 0.85% and 4.50%. To put these results in perspective, one would obtain a 5.68% average change in absolute value of exports when using counterfactual results presented by Costinot et al. (2012) from a similar experiment designed to quantitatively assess Ricardian comparative advantage.[40]

Finally, to gain a sense of the welfare implications of these counterfactuals we also computed changes in the price index for the set of differentiated products. Under the assumption that $\gamma = 0.5$, we focused on counterfactuals involving only the highest and lowest skill dispersion countries (respectively, Denmark and Poland). This exercise suggests that shutting down cross-country differences in skill-dispersion would increase the price index by 0.18% when taking Denmark as the reference country. The same exercise results in an average decrease in the price index of 0.35% when Poland is the reference country. These results combine the effect of removing differences in skill dispersion on comparative advantage and on the absolute level of productivity. When $\gamma = 0.5$, many industries feature calibrated λ's above 1, which explains why raising skill dispersion to the level of Poland has a beneficial effect on all countries' price indices. This effect more than offsets the increase in the price index due to the elimination of comparative advantage when a high fraction of λ's are above one. The opposite happens when $\gamma = 0.3$ and most λ's are below one (implying convex isoquants): the Denmark experiment yields an average differentiated product price index decrease of 0.03% for the Denmark case and an increase of 0.065% for the Poland case.

[40] This number is obtained by taking an un-weighted average of the absolute value of export changes presented in column (1) of Table 7 of Costinot et al. (2012). Some caution should be exercised in comparing results since several details of the counterfactual experiments differ across papers, including the samples of countries considered.

8 Conclusions

Relative differences in the distribution of factors of production are central to the classical theory of international trade. The Heckscher-Ohlin-Samuelson factor proportion model stresses the idea that cross-country differences in aggregate factor endowments play a major role in predicting trade flows. This paper belongs to a line of research that emphasizes that the entire distribution of a productive factor can enhance our understanding of trade patterns.

We develop a theoretical framework where, because skills are unobservable by firms, workers and firms are randomly matched. Skill dispersion affects industries differently because some technologies are more capable of substituting skills across production tasks than others. All industries in each country inherit the population distribution of unobservable skills and, as a result, firms in sectors with lower substitutability are relatively more productive in countries with lower skill dispersion. The paper therefore shows how differences in the dispersion of human capital inputs may lead to 'Ricardian-looking' differences in measured labor productivity at the country-industry level. In Ricardian models such productivity differences are often assumed to be the result of access to different (broadly defined) technologies. Our findings suggest that productivity wedges may also arise as the by-product of cross-country differences in the distribution of skills, when the latter are not properly accounted for. Higher productivity implies that countries with low unobservable skill dispersion specialize in (and export) goods produced under low skill substitutability. This result survives the introduction of cross-country differences in observable skills and provides a theoretical underpinning to the empirical analysis in BGP.

Relative to the seminal contribution of Grossman and Maggi (2000) we adopt a specific production function and we are therefore able to specify theoretical conditions under which, even in the presence of convex isoquants in all industries, comparative advantage and the pattern of trade are well defined.

In a calibration exercise we show that productivity wedges associated to skill dispersion are quantitatively important determinants of trade flows. We find that eliminating all skill dispersion differences produces a reallocation of production and an average absolute change in exports of almost 2%. Moreover these changes exhibit substantial heterogeneity across countries and sectors.

References

Akerlof, G. A., 1970. The market for "lemons": Quality uncertainty and the market mechanism. The Quarterly Journal of Economics, 488–500.

Benabou, R., 2005. Inequality, technology and the social contract. Handbook of economic growth 1, 1595–1638.

Bombardini, M., Gallipoli, G., Pupato, G., 2012a. How robust is the skill-dispersion-complementarity hypothesis? Mimeo, The University of British Columbia.

Bombardini, M., Gallipoli, G., Pupato, G., 2012b. Skill dispersion and trade flows. American Economic Review 102 (5), 2327–2348.

Bougheas, S., Riezman, R., 2007. Trade and the distribution of human capital. Journal of International Economics 73 (2), 421–433.

Costinot, A., Donaldson, D., Komunjer, I., 2012. What goods do countries trade? a quantitative exploration of ricardo's ideas. The Review of Economic Studies 79 (2), 581–608.

Costinot, A., Vogel, J., 2010. Matching and Inequality in the World Economy. Journal of Political Economy 118 (4), 747–786.

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.

Hamermesh, D. S., 1986. The demand for labor in the long run. Handbook of Labor Economics 1, 429–471.

Helpman, E., Itskhoki, O., Redding, S., 2010. Inequality and unemployment in a global economy. Econometrica 78 (4), 1239–1283.

Hsieh, C.-T., Hurst, E., Jones, C. I., Klenow, P. J., 2013. The allocation of talent and us economic growth. Tech. rep., National Bureau of Economic Research.

Kremer, M., 1993. The O-ring theory of economic development. The Quarterly Journal of Economics, 551–575.

Ohnsorge, F., Trefler, D., 2007. Sorting It Out: International Trade with Heterogeneous Workers. Journal of Political Economy 115 (5), 868–892.

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.

A Appendix

To simplify notation we define $B(\lambda, c)$ such that $B^{\gamma}(\lambda, c) = A(\lambda, c)$. It is immediate to show that $B(\lambda, c)$ is log-supermodular if and only if $A(\lambda, c)$ is log-supermodular. Proof are conducted in terms of $B(\lambda, c)$ for notational simplicity.

A.1 Proof of proposition 1

By definition of log-supermodularity we need to prove that, if $G_{c'}(a)$ is a mean-preserving spread of $G_c(a)$, then:

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

The partial derivative has the following expression:

$$\frac{\partial \log B(\lambda, c)}{\partial \lambda} = \frac{1}{\lambda} \frac{\int a^{\lambda} \log(a) dG_c(a)}{\int a^{\lambda} dG_c(a)} - \frac{1}{\lambda^2} \log \left(\int a^{\lambda} dG_c(a) \right) \tag{A-1}$$

A mean-preserving spread of $g(a, c)$ increases the second term of the right-hand side of (A-1) by definition, since a^{λ} 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. $\log a < \frac{2\lambda - 1}{(1-\lambda)\lambda}$. Note that the right-hand side of this inequality is continuous and increasing in λ, and $\lim_{\lambda \to 1} \frac{2\lambda-1}{(1-\lambda)\lambda} = \infty$. Then, if a is bounded above by $a_{\max}$, there exists a value $\lambda_{\min} < 1$ such that $\log a_{\max} = \frac{2\lambda_{\min} - 1}{(1-\lambda_{\min})\lambda_{\min}}$. If $\lambda > \lambda_{\min}$ then $\frac{\partial \log B(\lambda, c)}{\partial \lambda}$ increases with a mean preserving spread of $g(a, c)$.

We conclude that given an upper bound on the support of the skill distribution $a_{\max}$, we can find a value $\lambda_{\min}$ satisfying $0 < \lambda_{\min} < 1$ such that $A(\lambda, c)$ is log-supermodular in λ and c, for $\lambda_{\min} \leq \lambda \leq 1$. This completes the proof of Proposition 1.

A.2 Property 1 when Skill Distributions are Log-Normal or Pareto

(i) Pareto Distribution - Under the assumption that skills follow a Pareto distribution with mean μ and standard deviation σ, B takes the following expression:[41]

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

[41] 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 B as a function of those parameters: $B = a_{\min} \left(\frac{k}{k-\lambda}\right)^{\frac{1}{\lambda}}$. Since we are interested in a mean-preserving increase in variance, we express the B as a function of μ and σ, which are related to shape and location parameters according to the following equations:

$$a_{\min} = \frac{\mu^2+\sigma^2-\sigma\sqrt{\mu^2+\sigma^2}}{\mu}, \quad k = \frac{\sigma+\sqrt{\mu^2+\sigma^2}}{\sigma}$$

Since B is twice differentiable in σ and λ, the result in Proposition 4 is equivalent to B being log-supermodular in λ and σ, that is $\frac{\partial^2 \log B}{\partial\sigma\partial\lambda} > 0$. The expression for the cross partial derivative is the following:

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

$\lambda < 1$ ensures B is log-supermodular in λ and σ.

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

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

It is easy to show that under this distribution, B is log-supermodular for every $\lambda > 0$ since the following expression is always positive:

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

A.3 Numerical Implementation

In what follows we describe how we parameterize the model and how we solve for an equilibrium.

Preference shares. We assume that preferences are homogeneous across countries. To measure these shares we proceed in three steps: first, we use OECD-STAN data to approximate, for each country, the relative share of consumption in each of the 63 goods (industries). Second, we average each industry's shares across countries and obtain a common set of 63 values for $\alpha(\lambda)$. To simulate trade flows we need the preference share for the non-differentiated good (i.e., the good produced by the industry in which $\lambda = 1$ and there are no transport costs). This latter sector is only added for convenience in the model and cannot be related to a direct data counterpart. The (unobserved) consumption share of the non-differentiated good (the one produced by the industry in which $\lambda = 1$) can be obtained as a by-product of the general equilibrium solution and we set it at the lowest value which guarantees: (1) trade equilibrium, with no deficits or surpluses; (2) incomplete specialization.[42]

In what follows we provide some details about the individual steps. As we mentioned, the first step is to use OECD-STAN data to derive country-specific preference shares in each industry λ. We denote the expenditure share spent on each differentiated good λ by country j as $\alpha_j(\lambda)$ and compute it as follows:

$$\alpha_j(\lambda) = \frac{x_{jj}(\lambda) + \sum_{c\neq j} x_{cj}(\lambda)}{\sum_{\lambda}\left[x_{jj}(\lambda) + \sum_{c\neq j} x_{cj}(\lambda)\right]}$$

[42]Note that we set this share to the lowest possible value such that there is imperfect specialization in both benchmark calibration and experiments.

where $x_{jj}(\lambda)$ is the consumption of good λ produced in j. In turn $x_{jj}(\lambda)$ is calculated as:

$$x_{jj}(\lambda) = \left(\frac{1}{IPR_j(\lambda)} - 1\right) \sum_{c \neq j} x_{cj}(\lambda)$$

where $IPR_j(\lambda) = \frac{IMP_j(\lambda)}{PROD_j(\lambda)+IMP_j(\lambda)-EXP_j(\lambda)}$. Data on production $PROD_j(\lambda)$, imports $IMP_j(\lambda)$ and exports $EXP_j(\lambda)$ is from the OECD-STAN database. Whenever import penetration ratios (IPR) are not between zero and one, we replace them with their average across sectors for the same country.

Then, for each industry j we average across countries, weighing by population. This results in 63 different values (a set $\{\alpha(\lambda)\}$ for $\lambda = 1, 2, ..., 63$) which we use as the common preference shares in all countries.

Estimating the preference share parameter for the non-differentiated (residual) good (the industry with no trade costs and $\lambda = 1$) requires an extra step because there is no data which can be used to measure the share of consumption in this industry. In what follows we superimpose the symbol "~" to a model variable if it indicates its observed data value. We use the fact that, by definition,

$$\alpha^c_{\lambda=1} = \frac{\sum_j x_{jc}(\lambda = 1)}{\tilde{E}^c}$$

Clearly we cannot observe $\sum_j x_{jc}(1)$ from data. However we can derive it using the following equilibrium conditions,

$$W\tilde{L}^c = \sum_j x_{cj}(\lambda = 1) + \sum_{\lambda \neq 1} \sum_j \tilde{x}_{cj}(\lambda) \tag{A-3}$$

$$\sum_j x_{cj}(\lambda = 1) + \sum_{\lambda \neq 1} \sum_j \tilde{x}_{cj}(\lambda) = \sum_j x_{jc}(\lambda = 1) + \sum_{\lambda \neq 1} \sum_j \tilde{x}_{jc}(\lambda) \tag{A-4}$$

From equation A-3 it is possible to see that, given some W_1, we can identify $\sum_j x_{cj}(\lambda = 1)$; this can be used, in turn, in equation A-4 to obtain an estimate of $\sum_j x_{jc}(\lambda = 1)$. As discussed in the paper, W is constant across countries in equilibrium. Moreover there are infinitely many triplets of W, $\sum_j x_{cj}(\lambda = 1)$ and $\sum_j x_{jc}(\lambda = 1)$ which satisfy equations $(A-3, A-4)$. We pick the W which guarantees incomplete specialization in the benchmark equilibrium. More specifically, suppose that country c^* has the largest observed exports per worker in the differentiated goods sector (that is, the highest value $\sum_{\lambda \neq 1} \sum_j \tilde{x}_{c^*j}(\lambda)/\tilde{L}^{c^*}$). Then, we assume that $W = \sum_{\lambda \neq 1} \sum_j \tilde{x}_{c^*j}(\lambda)/\tilde{L}^{c^*}$, meaning that country c^* only produces an infinitely small amount of output in industry $\lambda = 1$. Since W is constant across countries, we can recover $\sum_j x_{cj}(\lambda = 1)$ for all the other countries using equation (A-3) and then verify that none of them is less or equal zero. In this way we also derive, using (A-4) for each country, the value of $\sum_j x_{jc}(\lambda = 1)$ and, given $\tilde{E}^c$, we can estimate $\alpha^c_{\lambda=1}$. Finally we average out these estimates to compute the common, economy-wide $\alpha_{\lambda=1}$, which is used to simulate the economy. Note that the 63 α_j for the differentiated goods' sectors are rescaled by $(1-\alpha_{\lambda=1})$, so that $\sum_{j=1}^{63} \frac{\alpha_j}{(1-\alpha_{\lambda=1})} + \alpha_{\lambda=1} = 1$.

Transport (residual) costs. The transport cost capture all residual heterogeneity underlying observed trade flows. We obtain estimates of the $\tau_{cj}(\lambda)$ values in two steps: first, we use the fact that any ratio of transport costs in the differentiated goods' sectors can be expressed as a function of observed bilateral trade flows; that is,

$$\frac{\tilde{x}_{jj}(\lambda)}{\tilde{x}_{cj}(\lambda)} = \frac{\tau_{jj}(\lambda)^{-\theta}}{\tau_{cj}(\lambda)^{-\theta}} = \frac{1}{\tau_{cj}(\lambda)^{-\theta}}$$

where the transportation cost of any country j to itself, $\tau_{jj}(\lambda)$, is normalized to one. Therefore, one can recover the transportation costs of any country pair in a given industry by solving $\tau_{cj}(\lambda) = \left(\frac{\tilde{x}_{jj}(\lambda)}{\tilde{x}_{cj}(\lambda)}\right)^{\frac{1}{\theta}}$, where the right hand side values are taken from data.[43] The estimated transport costs serve the purpose of a model residual, capturing any remaining source of trade flow variations which cannot be explained by the explicitly modeled sources of comparative advantage.

Once we have recovered transport costs for all country-pairs and industries, we decompose them into: (i) a common country pair component τ_{cj} (equal for all industries) and (ii) a country-industry specific term $\tau_j(\lambda)$. We do this by estimating the non-parametric regression,

$$\log \tau_{cj}(\lambda) = \phi_{cj} + \varphi_{j\lambda} + u_{cj\lambda}$$

This decomposition guarantees that transport costs do not generate comparative advantage. The τ's we employ in the counterfactual are the fitted values of the above regression, namely. $\widehat{\tau}_{cj}(\lambda) = \exp(\widehat{\phi}_{cj} + \widehat{\varphi}_{j\lambda})$.

Fundamental Productivities. The value of $A(\lambda, c)$ for each country-industry pair is approximated using the equation $A(\lambda, c) = \left(\int a^{\lambda} g(a,c) da\right)^{\frac{\gamma}{\lambda}}$ where a denotes the skill level of a worker, measured by the IALS scores. The mean skill level is normalized to one in all countries to eliminate comparative advantage due to average skills, and $g(a, c)$ denotes the skill distribution of country c. The discrete version of this proxy is $A(\lambda, c) = (\sum_{a \in \Lambda} a^{\lambda} w(a,c))^{\frac{\gamma}{\lambda}}$, where Λ denotes the set of possible values a can take in the data and $w(a, c)$ denotes the weight of workers with skill level a in country c.

Computing equilibria and counterfactual analysis. Once we have set values for all $\tau_{cj}(\lambda)$, $A(\lambda, c)$ and preference shares α_λ, we solve for the benchmark economy equilibrium. Assuming we know the total expenditure of country j, E_j, the simulated bilateral trade flows in the 63 differentiated goods' sectors $x_{cj}(\lambda)$ are given by,

$$x_{cj}(\lambda) = \frac{[\tau_{cj}/A(\lambda,c)]^{-\theta}}{\sum_{c'}[\tau_{c'j}/A(\lambda,c')]^{-\theta}} \alpha(\lambda)(1-\alpha_{\lambda=1})E_j \tag{A-5}$$

Similarly. imports and exports in the residual industry (non-differentiated goods' sector with

[43] Whenever observed trade flows are zero, we set them to be a tiny positive value.

$\lambda = 1$) can be calculated using, respectively, the condition $\sum_j x_{jc}(\lambda = 1) = \alpha_{\lambda=1} E_c$ and the trade balance condition in equation (A-4). With these simulated trade flows in hand, the equilibrium wage W must be such that, for each country, the following holds

$$W\tilde{L}^c = \sum_j x_{cj}(\lambda = 1) + \sum_{\lambda \neq 1} \sum_j x_{cj}(\lambda)$$

The question is: how can we derive a set of country-specific expenditures E_c such that the above conditions are not violated? It turns out that in our simple model the relative expenditures are trivially pinned down by

$$\frac{\tilde{L}_1}{\tilde{L}_c} = \frac{\sum_j x_{1j}(\lambda = 1) + \sum_{\lambda \neq 1} \sum_j x_{1j}(\lambda)}{\sum_j x_{cj}(\lambda = 1) + \sum_{\lambda \neq 1} \sum_j x_{cj}(\lambda)} = \frac{E_1}{E_c}$$

In our implementation we proxy the total employment ratio between two countries by the countries' population ratio.[44] Normalizing $E_1 = 1$, we have that $E_c = \frac{\tilde{L}_c}{\tilde{L}_1}$.

In the counterfactual analysis, we simply change $A(\lambda, c)$ while keeping transport costs and preference parameters unchanged. Then, counterfactual bilateral trade flows $x_{cj}(\lambda)$ can be computed, as long as E_j is known, using equation (A-5) and a new equilibrium can be computed again, exactly in the way described above. This gives a counterfactual wage W_{cf}. However, to compare the change in bilateral trade flows we assume that the wage in the benchmark and counterfactual economies are the same. This can be done easily: since trade flows and wages are linear in the expenditure E, we renormalize the countries expenditures E by a factor of $\frac{W}{W_{cf}}$ so that wages are equalized in the benchmark and counterfactual equilibria.

A.4 Quantitative Analysis: Additional Results

As we mentioned in Section 7, we experiment with alternative parameterizations of the production technologies. Every time we re-parameterize production technologies we solve for a new equilibrium, and obtain a new set of trade costs which match trade flows. In what follows we report technology estimates as well as results of counterfactual experiments under alternative parametrizations.

To pinpoint the values of γ and the λs we use the model restriction $\frac{w_\lambda h}{py} = \frac{\gamma}{\lambda}$. For each industry we use data on total compensation (or, alternatively, annual payroll) from the 2002 US Economic Census and divide them by value added to approximate $\frac{w_\lambda h}{py}$ and, hence, the ratio $\frac{\gamma}{\lambda}$. However these ratios do not allow us to estimate the value of γ and of all 63 λ's. To obtain these parameter estimates we proceed by setting γ to alternative values and then invert the $\frac{\gamma}{\lambda}$ ratios to get the associated λs. We experiment with three different values of γ: 0.3, 0.5 and 0.7. These values guarantee that all industries have convex isoquants. Estimates of the associated λs are presented in Table A.1, using the total compensation proxy for the wage bill. When we set $\gamma = 0.3$ we get substantially lower values for the λs, with a value larger than one only in 14 out of 63 industries; when $\gamma = 0.7$ all the industries have $\lambda > 1$, an admittedly extreme assumption which however helps us provide an upper bound for the changes in trade flows.

[44] Results do not change if we use total workforces.

Table A.1: Estimates of industry-specific λ values under alternative normalizations of γ

industry	$\gamma = 0.3$	$\gamma = 0.5$	$\gamma = 0.7$
Agricultural chemical mfr.	1.43	2.38	3.33
Agricultural implement, construction, mining and oil field machinery mfr.	0.68	1.13	1.58
Aircraft, aerospace prod. and parts mfr.	0.62	1.03	1.44
Aluminum production and processing	0.66	1.09	1.53
Animal food, grain and oilseed milling	1.58	2.64	3.69
Animal slaughtering and processing	0.77	1.28	1.80
Apparel accessories and other apparel mfr.	0.55	0.91	1.28
Bakeries	0.80	1.33	1.87
Beverage mfr.	1.37	2.28	3.20
Cement, concrete, lime, and gypsum product mfr.	0.75	1.24	1.74
Commercial and service industry machinery mfr.	0.67	1.11	1.55
Communications, audio, and video equipment mfr.	0.79	1.31	1.83
Computer and peripheral equipment mfr.	1.04	1.74	2.43
Cut and sew apparel mfr.	0.78	1.30	1.83
Cutlery and hand tool mfr.	0.77	1.28	1.79
Dairy product mfr.	1.05	1.76	2.46
Electrical lighting, equipment, and supplies mfr., n.e.c.	0.69	1.15	1.60
Electronic component and product mfr., n.e.c.	0.87	1.45	2.04
Engines, turbines, and power transmission equipment mfr.	0.97	1.62	2.27
Fabric Mills	0.59	0.98	1.37
Fiber, yarn, and thread mills	0.51	0.85	1.19
Footwear mfr.	0.56	0.93	1.30
Foundries	0.52	0.87	1.22
Fruit and vegetable preserving and specialty food mfr.	1.26	2.11	2.95
Furniture and related product mfr.	0.62	1.03	1.44
Glass and glass product mfr.	0.66	1.11	1.55
Household appliance mfr.	0.74	1.24	1.73
Industrial and miscellaneous chemicals	0.94	1.57	2.20
Iron and steel mills and steel product mfr.	0.61	1.02	1.43
Leather tanning and prod., except footwear mfr.	0.69	1.15	1.61
Machine shops; turned product; screw, nut and bolt mfr.	0.53	0.88	1.23
Machinery mfr., n.e.c.	0.58	0.97	1.35
Medical equipment and supplies mfr.	0.80	1.33	1.86
Metalworking machinery mfr.	0.48	0.79	1.11
Miscellaneous nonmetallic mineral product mfr.	0.78	1.30	1.82
Motor vehicles and motor vehicle equipment mfr.	0.73	1.22	1.71
Navigational, measuring, electromedical, and control instruments mfr.	0.58	0.98	1.37
Nonferrous metal, except aluminum, production and processing	0.63	1.05	1.47
Ordnance and miscellaneous fabricated metal prod. mfr.	0.65	1.08	1.51
Other transportation equipment mfr.	0.84	1.40	1.96

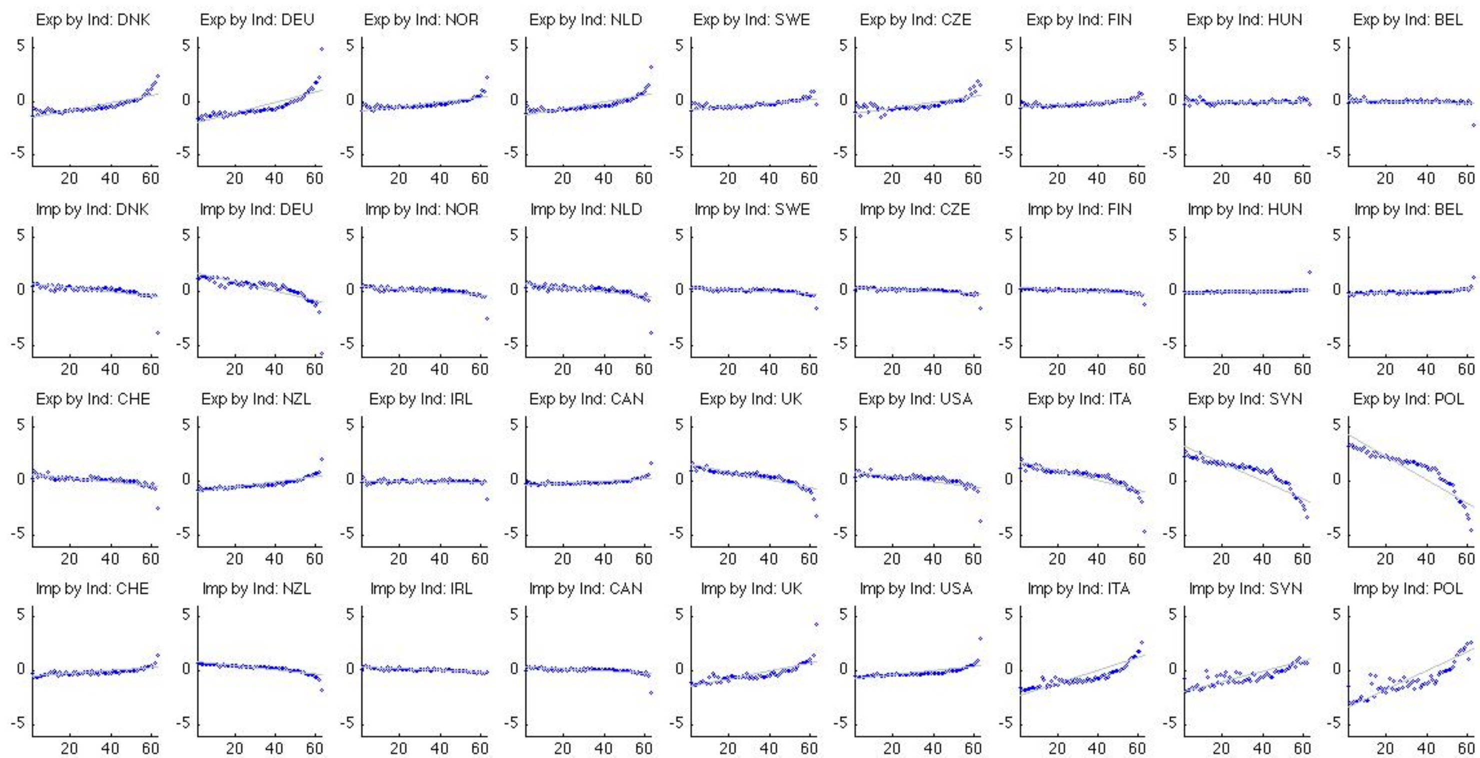
Paint, coating, and adhesive mfr. B46	1.04	1.74	2.43
Paperboard containers, boxes misc. paper and pulp prod.	0.70	1.17	1.64
Petroleum and Coal prod. mfr.	1.35	2.24	3.14
Pharmaceutical and medicine mfr.	1.78	2.97	4.16
Plastics product mfr.	0.70	1.17	1.64
Pottery, ceramics, structural clay and related prod. mfr.	0.63	1.05	1.47
Prefabricated wood buildings, mobile homes and miscellaneous wood prod.	0.54	0.90	1.25
Printing and related support activities	0.58	0.97	1.36
Pulp, paper, and paperboard mills	0.99	1.65	2.31
Railroad rolling stock mfr.	0.72	1.20	1.68
Resin, synthetic rubber and fibers, and filaments mfr.	1.02	1.71	2.39
Rubber prod.	0.55	0.91	1.28
Sawmills and wood preservation	0.53	0.88	1.24
Seafood and other miscellaneous foods, n.e.c.	1.44	2.40	3.36
Ship and boat building	0.56	0.94	1.31
Soap, cleaning compound, and cosmetics mfr.	1.96	3.27	4.58
Structural metals, and tank and shipping container mfr.	0.60	0.99	1.39
Sugar and confectionery prod.	1.15	1.92	2.69
Textile and fabric finishing and coating mills	0.69	1.15	1.61
Textile product mills	0.69	1.15	1.60
Tobacco mfr.	5.27	8.78	12.30
Toys, amusement, sporting goods and miscellaneous mfr., n.e.c.	0.65	1.09	1.52
Veneer, plywood, and engineered wood prod.	0.53	0.88	1.24

We have solved for alternative benchmark equilibria and studied the counterfactuals for all the technology parameterizations described above. We report the results for technology estimates where we set γ to either 0.3 or 0.7.[45] Figure A.1 reports changes in trade flows by industry (ordered in terms of increasing substitutability) using all 18 reference countries. The patterns confirm the findings discussed in the main body of the paper, although the magnitude of the changes tends to be smaller, with most of them being in the 1% range. When we instead simulate the model using technology estimates associated to the normalization $\gamma = 0.7$ we obtain larger changes in trade flows, as shown in Figure A.2.

Finally, for comparison, in Figures A.3 and A.4 we separately report the estimated changes in trade flows using as reference country only Germany and the US, under technology parameterizations with, respectively, $\gamma = 0.3$ or $\gamma = 0.7$. It is apparent that the patterns of change are identical, and the main difference is that $\gamma = 0.7$ implies larger effects, suggesting that the estimates of the effects of skill dispersion on trade patterns presented in the main body of the paper are fairly conservative.

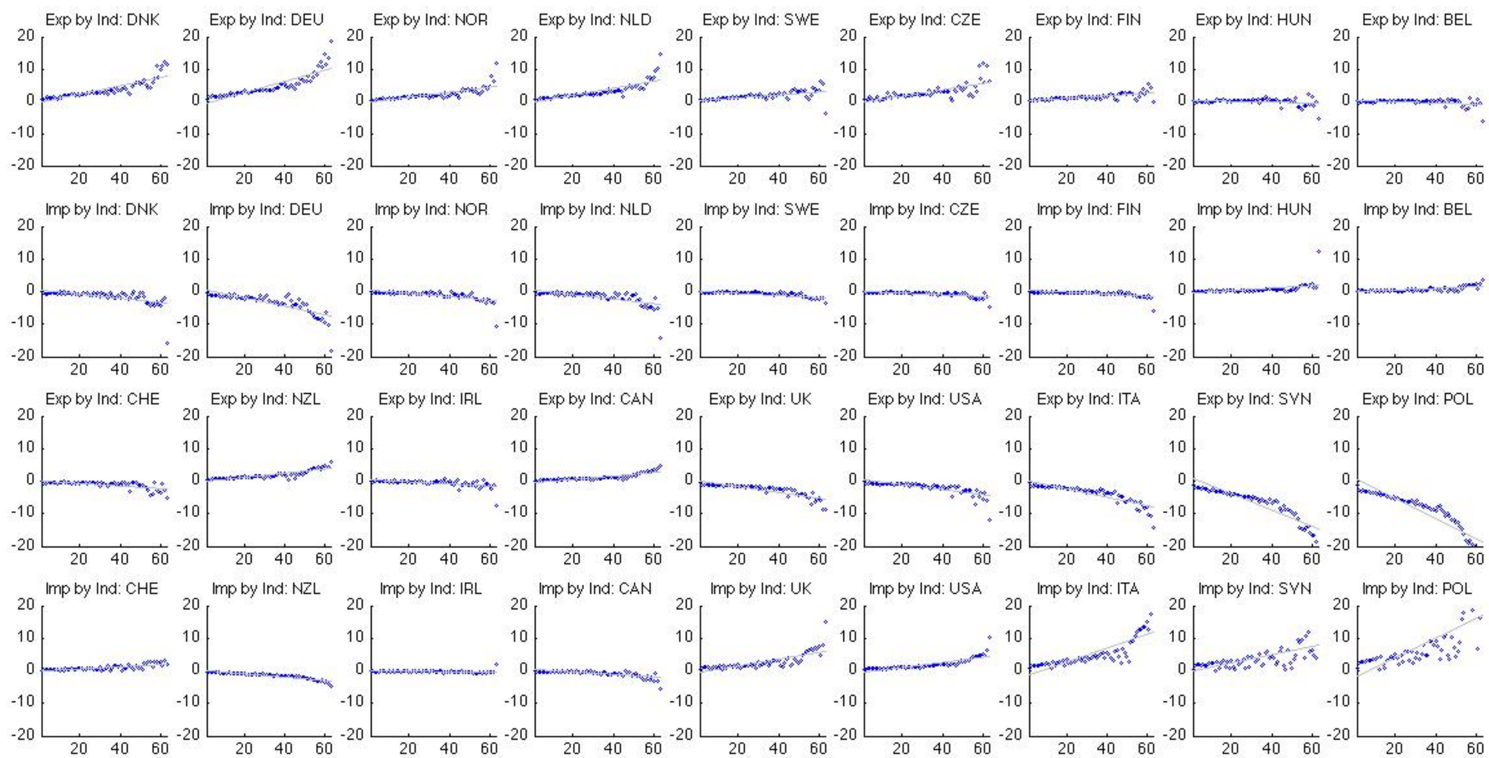
[45] We present results based on wage bills measured through total compensation. Available from the authors are several additional results based on wage bills measured from payroll data.

Figure A.1: Changes in trade flows across industries: $\gamma = 0.3$



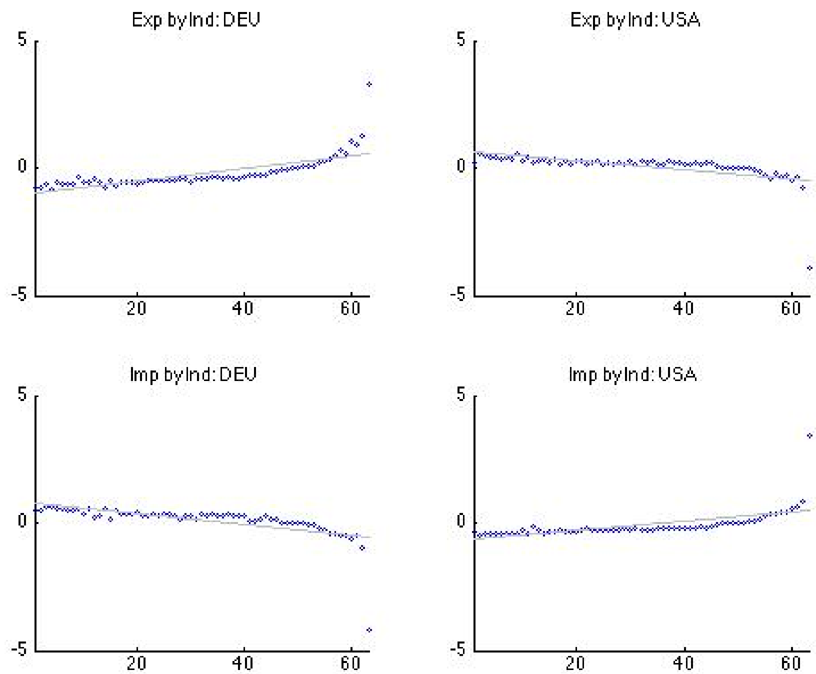
Note: the horizontal axis reports the 63 industries in order of increasing λ. The vertical axis reports the percentage change in trade.

Figure A.2: Changes in trade flows across industries: $\gamma = 0.7$



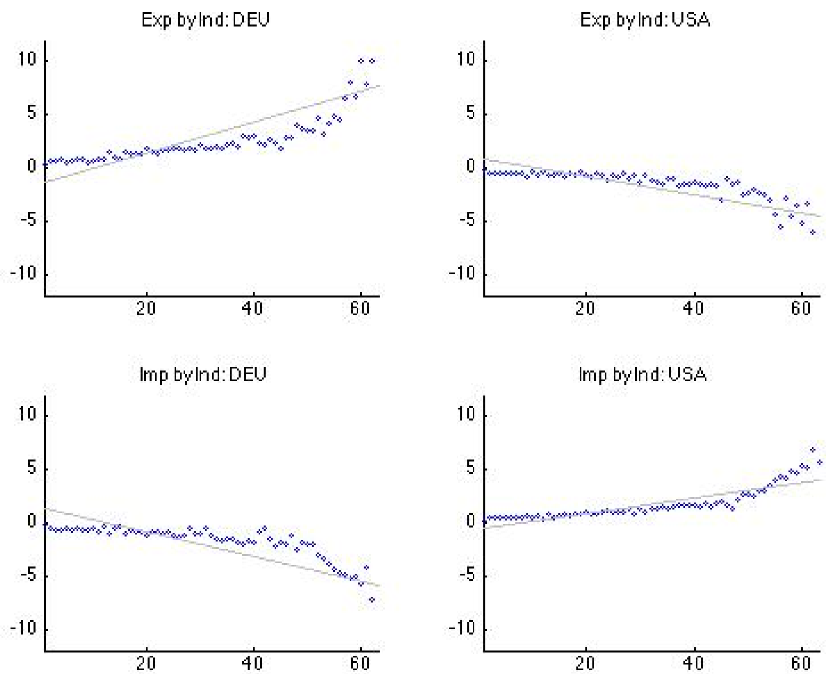
Note: the horizontal axis reports the 63 industries in order of increasing λ. The vertical axis reports the percentage change in trade.

Figure A.3: Changes in trade flows for Germany and US: $\gamma = 0.3$



Note: the horizontal axis reports the 63 industries in order of increasing λ. The vertical axis reports the percentage change in trade.

Figure A.4: Changes in trade flows for Germany and US: $\gamma = 0.7$



Note: the horizontal axis reports the 63 industries in order of increasing λ. The vertical axis reports the percentage change in trade.

Table A.2: Changes in trade flows (wage: $\gamma = 0.3$)

country	exports			imports		
	(1)	(2)	(3)	(4)	(5)	(6)
	average	10%	90%	average	10%	90%
DNK	0.47	-0.63	0.74	0.23	0.31	-0.21
NLD	0.46	-0.59	0.54	0.28	0.42	-0.39
NOR	0.27	-0.42	0.25	0.15	0.25	-0.23
FIN	0.24	-0.31	0.33	0.15	0.23	-0.20
DEU	0.51	-0.60	0.70	0.38	0.56	-0.48
SWE	0.23	-0.23	0.25	0.09	0.18	-0.14
CZE	0.28	-0.10	0.65	0.07	0.10	-0.08
BEL	0.07	-0.03	0.01	0.06	-0.04	0.08
HUN	0.11	0.02	0.31	0.01	-0.02	0.02
CHE	0.09	0.09	-0.11	0.08	-0.14	0.12
CAN	0.32	-0.44	0.37	0.24	0.45	-0.37
NZL	0.08	-0.09	0.08	0.06	0.03	-0.06
IRL	0.10	0.00	-0.06	0.04	0.06	-0.04
UK	0.35	0.59	-0.41	0.35	-0.52	0.45
USA	0.31	0.38	-0.37	0.33	-0.44	0.40
ITA	0.43	0.65	-0.38	0.60	-0.76	0.70
SVN	0.85	1.06	-0.84	0.49	-0.87	0.65
POL	1.44	1.91	-1.56	1.12	-1.84	1.29

Notes:

- Columns (1) and (3) report the average of absolute percentage changes in trade; columns (2), (4), (5) and (6) report raw (not absolute) percentage changes in trade for the industries corresponding to the 10^{th} and 90^{th} percentile in the distribution of estimated λ's (skill substitutability).
- The countries are ranked by skill dispersion.
- 10^{th} percentile and 90^{th} percentile are the 6th and 58th industries ranked by λ's.

Table A.3: Changes in trade flows (wage: $\gamma = 0.7$)

country	exports			imports		
	(1) average	(2) 10%	(3) 90%	(4) average	(5) 10%	(6) 90%
DNK	2.85	0.70	8.88	1.41	-0.34	-2.45
NLD	2.78	0.64	6.17	1.65	-0.45	-4.25
NOR	2.00	0.46	3.08	1.09	-0.28	-2.78
FIN	1.58	0.35	3.99	1.10	-0.26	-2.42
DEU	3.17	0.65	8.13	2.29	-0.59	-5.15
SWE	1.52	0.24	2.48	0.78	-0.17	-1.10
CZE	1.77	0.11	7.68	0.54	-0.10	-1.03
BEL	0.41	0.05	0.65	0.31	0.04	0.57
HUN	0.57	-0.04	3.51	0.15	0.04	0.23
CHE	0.40	-0.11	-1.18	0.53	0.17	1.33
CAN	1.84	0.50	4.01	1.30	-0.51	-3.95
NZL	1.13	0.10	1.33	0.85	-0.07	-1.41
IRL	0.82	-0.04	-1.17	0.23	-0.04	0.17
UK	1.51	-0.58	-3.64	1.63	0.53	4.27
USA	1.82	-0.47	-4.47	1.81	0.52	4.81
ITA	2.87	-0.74	-4.65	4.92	0.88	8.77
SVN	4.44	-1.16	-9.06	2.13	0.98	7.49
POL	7.30	-2.03	-15.75	6.58	2.03	14.74

Notes:

- Columns (1) and (3) report the average of absolute percentage changes in trade; columns (2), (4), (5) and (6) report raw (not absolute) percentage changes in trade for the industries corresponding to the 10^{th} and 90^{th} percentile in the distribution of estimated λ's (skill substitutability).
- The countries are ranked by skill dispersion.
- 10^{th} percentile and 90^{th} percentile are the 6th and 58th industries ranked by λ's.