Unobservable Skill Dispersion and Comparative Advantage

Matilde Bombardini*, Giovanni Gallipoli† and Germán Pupato‡

First Draft: June 2009

This Draft: December 2011

Abstract

This paper investigates a theoretical mechanism linking comparative advantage to the skill distribution in the working population. We develop a tractable multi-country, multi-industry model of trade with unobservable skills and search frictions 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 triggers specialization in industries characterized by higher substitutability of skills across tasks.

JEL Classification codes: F12, F16, J82.

We would like to thank Paul Beaudry, David Green, Patrick Francois, Keith Head, Thomas Lemieux, Giovanni Maggi, Vadim Marmer, 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.

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

†University of British Columbia and RCEA.

‡Graduate School of Economics, Getulio Vargas Foundation.

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., forthcoming, 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, 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 differ in the degree of skill substitutability in the production process. At the micro level, our framework features search frictions, 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. (forthcoming).

We introduce search frictions in the labor market as in Helpman et al. (2010) and focus the analysis on skills that are not observable ex-ante (i.e. before workers are matched). The latter modelling choice reflects 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 skill dispersions across countries. Ex-ante unobservability is also consistent with evidence suggesting that firms only slowly learn the skills of their workers (see for example Altonji and Pierret, 2001 and Altonji, 2005). Therefore the 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 sometimes refer to such skills as 'residual' skills.

One immediate advantage of our assumptions is that the model remains tractable in a setting with many countries and many sectors. Essentially, ex-ante unobservability and search frictions imply that firms randomly sample workers from the country's residual skill distribution. As a result, firms inherit the 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.

We show that this mechanism 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 firms located in countries with a high dispersion of skills in the labor force are relatively more productive, and 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 identical to the effects of technological differences in Ricardian models, such as Costinot et al. (forthcoming). 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 productive endowments.[1]

Our framework can be used to derive a theoretically consistent econometric specification to test its main predictions, a task pursued in BGP. In addition, it establishes a direct link between the degree of substitutability and the dispersion of residual wages within industries. In particular, we show that, as a consequence of random matching on residual skills, residual wage distributions at the industry level primarily reflect the degree of substitutability among skills in production. Industries with lower substitutability are characterized by a more compressed wage distribution because, for example, workers with higher than average skills contribute relatively less to surplus, a fact reflected in their wage.

Our analysis 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) -henceforth GM- whose modelling assumptions complement those of this paper. In particular, GM focus on the case of fully observable skills and competitive labor markets. In a two-country, two-sector model, they show that, for the case of observable skill dispersion, trade is conditional on the existence of a supermodular sector, where workers of identical abilities are paired together, i.e. self-matching prevails, and of a submodular sector, where

[1] Costinot et al. (forthcoming) 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.

the most skilled workers are paired with the least skilled co-workers, i.e. cross-matching prevails.[2] In GM, the country with more dispersed skill distribution specializes in the submodular sector. In this paper we study the effects of ex-ante unobservability of skills, extending the model to many countries and industries to make it empirically testable. Although GM also consider the case in which a portion of skills is unobservable, they do so in a stylized 2x2 setting where skills are complementary in one sector and substitutable in the other. Therefore their predictions do not easily generalize to more than two industries. Our analysis, instead, provides general results about comparative advantage for many industries which vary in the *degree* of substitutability.[3]

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 different skill distributions generate productivity differences across countries and industries, driving comparative advantage. Section 4 studies the optimization problem of indi-

[2] 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.

[3] We can explain the difference between GM and our setup in terms of isoquants in the workers' skills space. In GM one industry has convex isoquants, while the other industry has concave isoquants. We focus on the more usual case in which all industries feature convex isoquants, but differ in the *degree* of convexity.

vidual firms. Section 5 analyzes the implications of skill dispersion for the pattern of international trade. Section 6 characterizes the link between skill complementarity and wage inequality, which is extensively used in BGP. Section 7 explains how search costs, the mass of firms, labor allocations and expenditure decisions across countries and industries are determined in general equilibrium. 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 a continuum of differentiated goods $Q(i)$ with $i \in I$. The utility function U is Cobb-Douglas:

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

where $0 < \alpha(i) < 1$, $\int_{i \in I} \alpha(i) \, di = 1$ and $Q(i)$ is an aggregate consumption index over a fixed set Ω of varieties of i. Preferences exhibit constant elasticity of substitution $\sigma(i)$ across varieties of any good.[4] As a result, total expenditure on variety ω of good i is:

$$x(\omega, i) = \left[\frac{p(\omega, i)}{P(i)} \right]^{1-\sigma(i)} \alpha(i) E \tag{1}$$

[4]More specifically:

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

where $q(\omega, i)$ is the quantity consumed of variety ω of good i. We note that functions $\sigma(i)$ and $\alpha(i)$ can be allowed to vary arbitrarily across countries with very minor modifications to the analysis.

where E is aggregate expenditure, $p(\omega, i)$ 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 f. 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 $\tilde{g}(a)$ 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:

$$y = \varepsilon \left(\int a^{\lambda} h \tilde{g}(a)\, da \right)^{\frac{\gamma}{\lambda}} \text{ with } \gamma < \lambda < 1 \tag{2}$$

The parameter λ 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 larger λ, the more substitutable workers of different skill levels are. 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' skill levels. Since λ is the only characteristic that distinguishes technology across industries,[5] in the remainder of the paper we drop subscript i and index industries by their corresponding parameter λ.

The random Ricardian productivity shock $\varepsilon = \varepsilon(\omega)$ introduces intra-industry heterogeneity in labor productivity across varieties. In particular, we assume that ε is an i.i.d. draw from a Fréchet

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

distribution $F(\varepsilon) = \exp\left[-\varepsilon^{-\theta}\right]$ for all $\varepsilon > 0$, with $\theta > 1$. Importantly, F is identical across countries and industries. Therefore, these technology shocks do not generate comparative advantage in the model, unlike in Costinot et al. (forthcoming).

Two properties of this production function are worth mentioning. First, every worker has positive marginal product: in particular, the marginal product of adding worker h of skill level a_h is $\varepsilon h^{\frac{\gamma}{\lambda}-1}\left(\int a^{\lambda}\tilde{g}(a)\,da\right)^{\frac{\gamma}{\lambda}-1}\frac{\gamma}{\lambda}a_h^{\lambda}$. 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.[6] In what follows, we restrict the lower bound of λ by assuming $\gamma < \lambda$.[7]

2.3 Labor Market

We introduce labor market frictions in the spirit of the standard Diamond-Mortensen-Pissarides approach, following the specification in Helpman et al. (2010). Every worker looks for a job in one of the differentiated industries.[8] Workers are characterized by different skill levels and skill is a continuous variable distributed in the population of country c according to a density function $g(a,c)$. Firms pay a cost bh to randomly sample and match with a measure h of workers. The search cost b is endogenously determined by the labor market tightness, as shown in section 7.2.

We make the simplifying assumption that workers ignore the full distribution of wages in each industry, but know the expected wage and the probability of sectoral unemployment.[9] As a result,

[6]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.

[7]Relative to GM, 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 GM, here isoquants are always convex, even though we have a submodular production function (this is possible when we remove the crucial assumption of constant returns to scale present in GM).

[8]Workers do not have incentives to sort across different varieties within a sector because, as shown in section A.3, the wage does not depend on the variety produced by the employer.

[9]An alternative assumption, with identical implications, is that workers learn their skills only after production

the prevailing distribution of residual skills in the labor force is also inherited by the measure of workers searching for a job in each industry. Moreover, residual skills are not observable to the firm when hiring.[10] The combination of these assumptions yields no sorting in residual skills between workers and firms and implies that every firm, in any industry λ and country c, inherits the residual skill distribution in the general population:

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

There are a few considerations to make about random matching. Although we do not explicitly model observable skills, we do not rule out that sorting on observables may play a role in an extension of the current framework. More simply, our focus is on skills that are costly or time-consuming to observe before matching takes place because, as we show below, this may generate patterns of comparative advantage that are consistent with the empirical evidence in BGP. Notice that, in this model, nothing prevents firms from firing workers that turn out to be 'bad' ex-post. However, their marginal product is still positive and, as section 6 shows, this simply results in lower equilibrium wages. Moreover, random matching makes our theoretical framework tractable in a context with many countries, many industries and arbitrary skill distributions.

takes place.

[10] Note that, contrary to the case described by Helpman et al. (2010), with our production function firms would not want to screen workers even if a screening technology were available, because the marginal product of an additional worker is always positive. This holds always true in the current static problem. In a dynamic framework firms might have incentives to lay off unproductive workers and replace them with potentially more productive ones. This is beyond the scope of this paper and left for future research. However we notice that re-matching will not be empirically relevant if learning about workers' skills is slow enough and if industry-specific human capital accumulation is relatively fast.

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:

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

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.

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.[11]

We focus on comparing fundamental productivity across 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 $g(a, c')$ which is a mean-preserving spread of the skill distribution $g(a, c)$ in country c.

[11] 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.

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

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

Property 1 states that 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 provide two different approaches to studying this problem. First, we show that comparative advantage can be established for any upper-bounded distribution by restricting the range of substitutability. Second, we perform comparative statics for specific distributions of skills that do not satisfy the sufficient conditions of the first approach. These analytical results buttress existing numerical evidence which is consistent with (3): in BGP we construct $A(\lambda, c)$ for a grid of λ in the $[0, 1]$ interval, using the empirical distribution of test scores for 19 countries that participated in the International Adult Literacy Survey (IALS) and verify that, when averaging across country pairs, $\frac{A(\lambda, c')}{A(\lambda, c)}$ is increasing in λ for 97% of the grid points.

Our first approach yields a general result for skill distributions with bounded supports, that relies on restricting the degree of substitutability of skill in technology.[12]

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

(i) the support of the skill distribution in every country is bounded from above by $a_{\max}$*;*

(ii) the degree of substitutability in each industry is high enough: that is, $\lambda > \overline{\lambda}$ *where* $\overline{\lambda}$ *is defined by the following condition*

$$\log a_{\max} = \frac{2\overline{\lambda} - 1}{\left(1 - \overline{\lambda}\right) \overline{\lambda}}.$$

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

Notice that as $a_{\max}$ increases, the 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 discuss at the end of this section.

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.

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

Proposition 2 establishes an analytical result; however, we have also numerically computed the A's for several other distributions, such as 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 (3).

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 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. We follow Stole and Zwiebel (1996) in allowing workers to renegotiate wage contracts after being matched, therefore the firm also engages in a bargaining game with its employees over the surplus generated. In the appendix section 6, we show that the firm's equilibrium share of revenue, s, equals $\lambda/(\lambda + \gamma)$.

We drop the industry index to simplify notation. Under perfect competition, the price $p_j(\omega)$

that consumers in country j pay for variety ω is

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

where $\mu_{cj}(\omega)$ is the unit cost of producing and delivering variety ω from country c to country j. Therefore, firms in country c find it profitable to start production of variety ω 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.[13] Otherwise, the firm could reallocate output across destinations and increase profits. Let us denote this FOB price as p. The firm's problem is to maximize profits:

$$\max_h \ sp\varepsilon A(\lambda, c) h^{\frac{\gamma}{\lambda}} - bh - f.$$

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 $u(y)$. The minimum of this curve pins down the firm's efficient scale y^* and employment h^*. The minimum average production cost is:

$$u_c(y^*) = \kappa \frac{b^{\frac{\gamma}{\lambda}}}{\varepsilon A(\lambda, c)},$$

where $\kappa > 0$ under decreasing marginal returns.[14]

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

[14] The expressions are: $h^* = \frac{f\gamma}{b(\lambda-\gamma)}$, $y^* = \varepsilon A(\lambda, c) h^{*\frac{\gamma}{\lambda}}$ and $\kappa \equiv f^{\frac{\lambda-\gamma}{\lambda}} (\lambda-\gamma)^{\frac{\gamma}{\lambda}-1} \gamma^{-\frac{\gamma}{\lambda}} \lambda$.

Profit maximization under perfect competition implies that marginal revenue, sp, equals marginal costs. In turn, free entry implies that every firm produces at the efficient scale, where marginal cost equals average cost, u_c. Therefore, assuming that there is a continuum of small potential entrants, the industry's average cost function is perfectly elastic at the FOB price p that satisfies the zero-profit condition $sp = u_c$. As a result, the industry's unit cost of producing and delivering variety ω from country c to country j is:

$$\mu_{cj}(\omega) = \tau_{cj} \cdot \frac{u_c(y^*(\omega))}{s}$$

where $\tau_{cj} \geq 1$ is the iceberg transport cost from c to j, with strict inequality whenever $c \neq j$. 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

Geographical specialization is determined by the location of the minimum cost suppliers to each destination, which is a function 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 a continuum of industries and multiple countries. This section shows how comparative advantage is in turn fully determined by fundamental productivity.

At the micro level our framework features search frictions, worker heterogeneity and non-linear returns to scale; however, at the industry level, it is isomorphic to the perfectly competitive model with CRS and technological differences across many countries and sectors in Costinot et al. (forthcoming). We therefore follow their derivation of the pattern of trade.

As a first step, we provide an expression for bilateral exports at the industry level. 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 $\Omega_{cj}(\lambda) \equiv \left\{\omega \in \Omega \mid \mu_{cj}(\omega, \lambda) = \min_{1 \leq c' \leq C} \mu_{c'j}(\omega)\right\}$ is the subset of varieties exported by country c to country j in industry λ. Then,

Lemma 3

$$x_{cj}(\lambda) = \frac{\left[b(\lambda, c)^{\frac{\gamma}{\lambda}} \tau_{cj}(\lambda)/A(\lambda, c)\right]^{-\theta}}{\sum_{c'=1}^{C} \left[b(\lambda, c')^{\frac{\gamma}{\lambda}} \tau_{c'j}(\lambda)/A(\lambda, c')\right]^{-\theta}} \alpha(\lambda) E_j \tag{4}$$

Proof. The result follows from the proof of Lemma 1 in the online appendix of Costinot et al. (forthcoming), 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. $F(z) = \exp\left[-\left(z/A(\lambda, c)\right)^{-\theta}\right]$, for all $z > 0$. Notice that κ and s vary by industry, but not by country, and therefore cancel out in 4. ■

Intuitively, 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 capture a higher proportion of the industry's varieties imported by consumers in any destination.

Assume $\tau_{cj}(\lambda) = \tau_{cj} \cdot \tau_j(\lambda)$ for all λ and $c \neq j$. Then, Lemma 3 can be used to establish that for any importer j and any pair of exporters c, $c' \neq j$, the ranking of relative fundamental productivities fully determines the ranking of relative exports. That is, for any pair of industries λ and λ':[15]

$$\frac{A(\lambda, c')}{A(\lambda, c)} \leq \frac{A(\lambda', c')}{A(\lambda', c)} \Leftrightarrow \frac{x_{c'j}(\lambda)}{x_{cj}(\lambda)} \leq \frac{x_{c'j}(\lambda')}{x_{cj}(\lambda')} \tag{5}$$

Notice that variation in search costs does not affect the ranking of relative exports. This is because in equilibrium $b(\lambda, c)$ is both log-supermodular and log-submodular as shown in Lemma

[15] This result uses the fact -stated in Lemma 6 below- that variation in equilibrium search costs across countries and industries does not change the ranking of relative exports.

6 below. 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 4 *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* λ.

Proof. Follows immediately from Property 1 and the ranking of relative exports in 5. ■

6 Wage Distribution and Complementarity

This section discusses how wages are determined. In addition, it shows that the model can provide a proxy for the degree of substitutability, which is not directly observable and for which, to the best of our knowledge, there are no direct estimates available. In particular, this section establishes the existence of a one-to-one link between the degree of substitutability and the dispersion of wages in each industry, which can then be used to construct an empirical test of the main predictions of the model.

Because of search frictions, matched workers are not costlessly interchangeable with outside workers. We assume that the firm and its workers engage in bargaining to share the surplus as in Stole and Zwiebel (1996). Workers' outside option is unemployment, which yields a payoff of zero. Stole and Zwiebel show that the bargaining solution yields payoffs that correspond to the Shapley value. As discussed above, we assume that at the bargaining and production stage workers' skills are revealed, and workers of different skills receive different wages as a result of intra-firm bargaining. This stark assumption captures some realistic features of the hiring process and subsequent employment, because workers' skills are difficult to assess until they start working (and even after that, as documented by Altonji and Pierret, 2001).

Section A.3 in the appendix provides the derivation of the Shapley value for a worker of skill a. Since the average wage also differs across industries, we normalize the wage of a worker of skill a in industry λ by the average wage in the industry. The normalized wage is $\widetilde{w}(a,\lambda) = \frac{a^{\lambda}}{E\left(a^{\lambda}\right)}$, which depends on λ and reflects the marginal product of a worker of skill a when added to the production team. The higher the substitutability across workers the larger the marginal product of a worker with high skills. By contrast, if λ is low, teamwork becomes more important and a highly skilled worker has a relatively lower marginal product because her skills are very different from the average skills of her co-workers. An implication of this wage structure is that workers with identical skills, but employed in different sectors, generally receive different wages, as returns to skills vary across industries.[16]

As the distribution of residual skills is the same in each industry, it follows that, in our framework, the distribution of wages within an industry depends exclusively on technological factors, which affect the marginal product of workers with different skills.[17] The following proposition establishes the existence of a one-to-one correspondence between the degree of skill substitutability and several common measures of wage dispersion.

Proposition 5 *For any non-degenerate distribution of skills $g(a,c)$, the following three measures of sectoral wages' dispersion are strictly increasing in the degree of substitutability of workers' skills, λ: (i) the Coefficient of Variation; (ii) the Gini Coefficient and (iii) the Inter-Percentile Ratio*[18]

[16] The point is made by Heckman and Scheinkman (1987), who show that returns to unobservable characteristics are different across sectors. This simple fact is also a recurring object of attention for modern search theory.

[17] This is equivalent to saying that within-industry wage dispersion does not reflect compositional differences in unobserved ability across industries.

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

Proposition 5 establishes that the more substitutable workers are, the less compressed the wage distribution is. The intuition follows from our discussion of normalized wages.

7 General Equilibrium

This section explains how to solve for the remaining endogenous variables of the model.

7.1 Trade Balance

Imposing a balanced trade condition allows us to pin down expected income per capita in each country, denoted W_c. As in Costinot et al. (forthcoming), we denote by $\pi_{cc'}(\lambda) \equiv x_{cc'}(\lambda)/\sum_j x_{jc'}$ the share of exports from c to c' in industry λ. Then, for any country c, balanced trade requires

$$\sum_{c'} \int \pi_{cc'}(\lambda)\, \alpha(\lambda)\, \beta_{c'} d\lambda = \beta_c$$

where $\beta_c \equiv W_c L_c / \sum_{c'} W_{c'} L_{c'}$ is the share of country c in world income. Having determined exports in section 5, this system provides $C-1$ independent equations which, together with a choice of numeraire, pin down W_c in every country. Aggregate income follows immediately using $E_c = W_c L_c$.

7.2 Search Costs

We follow Helpman et al. (2010) and Blanchard and Gali (2010) in specifying the search cost as an increasing function of the labor market tightness (χ),

$$b = \delta_0 \chi^{\delta_1}, \qquad \delta_0 > 0,\, \delta_1 > 0 \tag{6}$$

where the labor market tightness is defined as the ratio of matched workers to those searching for a job in the industry. In an equilibrium with incomplete specialization, workers are indifferent between searching for employment in any industry. This requires expected wages to be equal across industries and, therefore, equal to expected income per-capita in the economy W_c. In turn, the expected wage in a given industry equals the expected wage conditional on being matched times the probability of being matched, χ. Since matched workers earn a fraction $1-s$ of total revenues, the indifference condition can be written as

$$(1-s)\frac{bh^{*}+f}{sh^{*}}\chi = W_c. \tag{7}$$

Equations 6 and 7 allow us to solve for the equilibrium search cost and labor market tightness in each industry. In particular, using the solutions for h^{*} and s, we obtain

$$b(\lambda, c) = \delta_0 \left[\frac{W_c}{\delta_0}\frac{\lambda^2}{\gamma(2\lambda-\gamma)}\right]^{\frac{\delta_1}{1+\delta_1}} \tag{8}$$

The following result summarizes how the equilibrium search cost varies across countries and industries.

Lemma 6 *The equilibrium search cost $b(\lambda, c)$ is both log-supermodular and log-submodular in λ and c. That is, for any pair of countries c, c' and any two industries λ, λ':*

$$\frac{b(\lambda, c')}{b(\lambda, c)} = \frac{b(\lambda', c')}{b(\lambda', c)}$$

Proof. Follows immediately by applying equation 8 ■

Lemma 6 establishes that relative search costs across industries will be the same in any two countries. Although unit costs and industry exports are functions of b, this result implies that skill dispersion does not generate comparative advantage through an effect on search costs in the model.

7.3 Mass of Firms and Labor Allocations

The mass of firms in each industry and country, 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 firms that supply these varieties. Since equilibrium firm revenues are $bh^* + f$, the market clearing condition in any given industry can be written as

$$\sum_j x_{cj} = M_c \left[\frac{bh^* + f}{s} \right]$$

which, together with the solutions for trade flows and search cost, can be used to solve for $M_c(\lambda)$.

In turn, the mass of workers searching for employment in any given industry, $L_c(\lambda)$, can be determined by noting that, from the solution to the bargaining game, expected labor payments are a fraction $1 - s$ of total revenue:

$$W_c L_c = (1 - s) \, M_c \left[\frac{bh^* + f}{s} \right].$$

8 Conclusions

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

We develop a theoretical framework where, because of frictions in the labor market and an ex-ante unobservable component in skills, 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 ex-ante unobservable skills and, as a result, firms in sectors with lower substitutability are relatively more productive in countries with lower skill dispersion. The model provides an observable proxy for the otherwise unobservable degree of substitutability among workers' skills, that is the dispersion of residual wages at the industry level. This result provides a foundation for some of the empirical evidence in BGP.

It is worth noting that this paper also 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.

References

Acemoglu, D., Antràs, P., Helpman, E., 2007. Contracts and technology adoption. American Economic Review 97 (3), 916–943.

Altonji, J., 2005. Employer learning, statistical discrimination and occupational attainment. American Economic Review, 112–117.

Altonji, J., Pierret, C., 2001. Employer learning and statistical discrimination. Quarterly Journal of Economics 313.

Aumann, R. J., Shapley, L. S., 1974. Values of non-atomic games. Princeton University Press Princeton.

Blanchard, O., Gali, J., 2010. Labor markets and monetary policy: a new keynesian model with unemployment. American Economic Journal: Macroeconomics 2 (2), 1–30.

Bombardini, M., Gallipoli, G., Pupato, G., forthcoming. Skill Dispersion and Trade Flows. American Economic Review.

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., forthcoming. What goods do countries trade? a quantitative exploration of ricardo's ideas. Review of Economic Studies.

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.

Heckman, J., Scheinkman, J., 1987. The importance of bundling in a Gorman-Lancaster model of earnings. The Review of Economic Studies, 243–255.

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

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.

Stole, L. A., Zwiebel, J., 1996. Intra-firm bargaining under non-binding contracts. The Review of Economic Studies, 375–410.

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 iff $A(\lambda, c)$ is log-supermodular. The proofs of Proposition 1 and 2 are conducted in terms of $B(\lambda, c)$ for simplicity.

A.1 Proof of proposition 1

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

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

The partial derivative has the following expression:

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

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. $\log a < \frac{2\lambda - 1}{(1-\lambda)\lambda}$. Since the right-hand side of this inequality is continuous and increasing in λ, it is equal to zero for $\lambda = \frac{1}{2}$ and $\lim_{\lambda \to 1} \frac{2\lambda - 1}{(1-\lambda)\lambda} = \infty$ then, if a is bounded above by $a_{\max}$, then there exists a value $\overline{\lambda} < 1$ such that $\log a_{\max} = \frac{2\overline{\lambda} - 1}{(1-\overline{\lambda})\overline{\lambda}}$. If $\lambda > \overline{\lambda}$ then $\frac{\partial \log B(\lambda, c)}{\partial \lambda}$ increases with a mean preserving spread of $g(a, c)$.

A.2 Proof of proposition 2

(i) *Pareto Distribution* - Under the assumption that skills follow a Pareto distribution with mean μ and standard deviation σ, B takes the following expression:[19]

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

[19] 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:

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

$$k = \frac{\sigma + \sqrt{\mu^2 + \sigma^2}}{\sigma}$$

Since B is twice differentiable in σ and λ, the result in Proposition 3 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:

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

and $\lambda < 1$ so 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 since the following expression is always positive:

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

A.3 Derivation of the Shapley Value

In this section we provide details on how to derive the share of revenues accruing to the firm and the wages paid to workers. Stole and Zwiebel (1996) have proved the equivalence of their bargaining solution to the Shapley value of the corresponding cooperative game not only for the case of identical workers, but also for the case of heterogeneous workers,[20] therefore we calculate the Shapley value directly.[21] The Shapley value of the firm is heuristically derived as its marginal contribution averaged over all possible orderings of employees and the firm itself. The case of heterogeneous employees is easy to handle under our assumption of a continuum of workers because no matter how the firm is ordered, it is preceded by a mass of workers whose skill distribution mirrors the overall skill distribution in the workers population, so the only variable we have to keep track of is the mass of workers preceding the firm, define it n, which varies from zero to h. As discussed in Acemoglu et al. (2007), since the firm is an essential input its marginal contribution is equal to revenues when n workers are employed in production. The Shapley value of the firm S_{firm} is therefore:

$$\begin{aligned} S_{firm} &= \int_0^h \frac{1}{h} p\varepsilon A n^{\frac{\gamma}{\lambda}} dn = \\ &= sp\varepsilon h^{\frac{\gamma}{\lambda}} A \end{aligned}$$

where p is the producer price (discussed below), s is defined as

$$s = \frac{\lambda}{\lambda+\gamma}.$$

[20] See their Theorems 8 and 9, p. 393.

[21] The analogous of the Shapley value for a continuum of players is derived in Aumann and Shapley (1974).

Intuitively, the share of revenues accruing to the firm depends on the curvature of the revenue function, which in turn depend on γ and λ.

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

$$r\left(n\right) = p\varepsilon \left[\int_a a^{\lambda} n(a) da\right]^{\frac{\gamma}{\lambda}}$$

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

$$\frac{\partial r\left(n\right)}{\partial n(a)} = p\varepsilon\frac{\gamma}{\lambda}\left[\int_a n\widetilde{g}(a)a^{\lambda}da\right]^{\frac{\gamma}{\lambda}-1} a^{\lambda} = p\varepsilon\frac{\gamma}{\lambda}A\left(\lambda,c\right)^{\frac{\gamma-\lambda}{\gamma}} a^{\lambda}n^{\frac{\gamma}{\lambda}-1}$$

The Shapley value and wage of worker of skill a in industry λ is:

$$\begin{aligned} w(a,\lambda) &= \frac{1}{h}\int_o^h \frac{n}{h}\frac{\partial r\left(n\right)}{\partial n(a)}dn = \\ &= p\varepsilon A\left(\lambda,c\right)^{\frac{\gamma-\lambda}{\gamma}}\frac{\gamma}{\gamma+\lambda}a^{\lambda}h^{\frac{\gamma}{\lambda}-1} \end{aligned} \quad \text{(A-3)}$$

Because p is the FOB price, it is constant across destinations (does not depend on transportation costs) and is inversely proportional to ε :

$$p = \frac{\kappa_{\lambda} b^{\frac{\gamma}{\lambda}}}{s\varepsilon A\left(\lambda,c\right)}$$

In equilibrium $h^{*} = \frac{f\gamma}{b(\lambda-\gamma)}$ and does not depend on ε. We can therefore rewrite (A-3) as:

$$w(a,\lambda) = \frac{\kappa_{\lambda} b^{\frac{\gamma}{\lambda}}}{s}A\left(\lambda,c\right)^{-\frac{\lambda}{\gamma}}\frac{\gamma}{\gamma+\lambda}a^{\lambda}\left(\frac{f\gamma}{b\left(\lambda-\gamma\right)}\right)^{\frac{\gamma}{\lambda}-1} \quad \text{(A-4)}$$

Notice that (A-4) does not vary across varieties within the same industry for a given a so that an individual with skill a receives the same wage regardless of the variety in which she works. Since we are going to be interested in comparing dispersion across sectors and the average wage also differs across sectors, we normalize wages by the average wage in the industry $E\left[w(a,\lambda)\right]$. The normalized wage is denoted by $\widetilde{w}(a,\lambda) = \frac{w(a,\lambda)}{E[w(a,\lambda)]}$ and takes the following form:

$$\widetilde{w}(a,\lambda) = \frac{a^{\lambda}}{E\left(a^{\lambda}\right)}$$

A.4 Proof of Proposition 5

We consider three measures of wage dispersion:

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

$$Var\left(\widetilde{w}(a,\lambda)\right) = \frac{E\left(a^{2\lambda}\right)}{E\left(a^{\lambda}\right)^2} - 1, \tag{A-5}$$

(ii) the Gini Coefficient, defined with respect to the Lorenz Curve for normalized wages at the industry level $\Lambda\left(w,\lambda\right)$,

(iii) the Inter-Percentile Ratio IPR_{kj} defined as:

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

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

(i) *Coefficient of Variation*

Since the variance of normalized wages is equal to the square of the coefficient of variation we prove the result for the former. We start by rewriting (A-5) in an explicit form, dropping the country index c to simplify notation:

$$Var\left(\widetilde{w}(a,\lambda)\right) = \frac{\int a^{2\lambda}\tilde{g}\left(a\right)da}{\left(\int a^{\lambda}\tilde{g}\left(a\right)da\right)^2} - 1 \tag{A-6}$$

The derivative of (A-6) with respect to λ is non-negative if an only if the following inequality is satisfied:

$$\left(\int_0^{\infty} a^{2\lambda}\log a\ \tilde{g}\left(a\right)da\right)\left(\int_0^{\infty} a^{\lambda}\ \tilde{g}\left(a\right)da\right) \geq \left(\int_0^{\infty} a^{\lambda}\log a\ \tilde{g}\left(a\right)da\right)\left(\int_0^{\infty} a^{2\lambda}\ \tilde{g}\left(a\right)da\right) \tag{A-7}$$

The left-hand side of (A-7), which we denote by Φ_L can be rewritten as:

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

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

$$\Phi_L = \int_0^{\infty}\left(\int_0^{a} b^{\lambda}\ \tilde{g}\left(b\right)db\right)a^{2\lambda}\log a\ \tilde{g}\left(a\right)da + \int_0^{\infty}\left(\int_a^{\infty} b^{\lambda}\ \tilde{g}\left(b\right)db\right)a^{2\lambda}\log a\ \tilde{g}\left(a\right)da \tag{A-8}$$

We change the order of integration in the second component of Φ_L so that we can rewrite (A-8) it as:

$$\Phi_L = \int_0^{\infty}\left(\int_0^{a} b^{\lambda}\ \tilde{g}\left(b\right)db\right)a^{2\lambda}\log a\ \tilde{g}\left(a\right)da + \int_0^{\infty}\left(\int_0^{b} a^{2\lambda}\log a\ \tilde{g}\left(a\right)da\right)b^{\lambda}\ \tilde{g}\left(b\right)db \tag{A-9}$$

Finally, a change of variable in the second component of (A-9) allows us to express Φ_L as:

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

If the same decomposition is performed on the right-hand side of (A-7) we can rewrite the inequality as follows:

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

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

(ii) *Gini Coefficient*

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

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

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

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

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

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

Simplifying and factorizing leads to the following inequality:

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

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

(iii) *Inter-Percentile Ratio*

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

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

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