Supply and Demand in a Two-Sector Matching Model*

Paweł Gola†

This version: November 17, 2017

Job Market Paper

[Please download the most recent version at pawelgola.com/GolaJMP.pdf]

Abstract

This paper merges a model of workers' self-selection (across manufacturing and services) with an assignment model (within sectors). I provide comparative statics results for changes in the production function, the distribution of skill and the distribution of firms' productivity. Any manufacturing-specific technological improvement that favors high-skilled workers results in better sorting into manufacturing. If unemployment is positive, this raises wage inequality in both industries. Perverse output effects are possible: manufacturing might contract if the technological improvement favors low-skilled workers. Finally, in the symmetric case, wages become more polarized if sectors start using more similar bundles of skills.

Keywords: two-sector matching, self-selection, imperfect substitution, demand for skill, supply of skill, wage inequality, wage polarization

JEL codes: C78, D31, J24, J31

*I am very grateful to my doctoral advisor, Ian Jewitt, for his support and guidance. I would like to thank Peter Eso for feedback and encouragement, as well as Matej Bajgar, Leonie Baumann, Juan Block, River Chen, Aytek Erdil, Edoardo Gallo, Axel Gottfries, Claudia Herresthal, Ed Hopkins, Jashua Lanier, Karan Nagpal, Bartosz Redlicki, Mikhail Safronov, Federico Torracchi and Łukasz Woźny for detailed comments on this and earlier drafts. I would also like to thank Michał Burzyński, Vasco Carvalho, Paweł Dziewulski, Meg Meyer, Climent Quintana-Domeque, Larry Samuelson, Coen Teulings and seminar participants at Oxford, Cambridge, the University of Luxembourg, the 2014 Ce^2 Workshop, the 2014 WIEM Conference, the 2015 RES Conference, the 2016 SSES Congress and the 2016 North American Summer Meeting of the Econometric Society for discussions and suggestions. The paper benefited greatly from insightful comments by James J. Heckman and four anonymous referees. All errors are mine.

†Emmanuel College, University of Cambridge. E-mail: pg454@cam.ac.uk

1 Introduction

Industries differ in the technologies they use and, consequently, in the bundles of skill they demand. Within sectors, workers differ in the ability to perform their jobs, whereas firms differ in their organizational structures and production processes. In this paper, I study how this inherent heterogeneity in technology and skills affects outcomes, wages in particular, when workers are free to choose the sector they work in. Specifically, I answer the following questions: what is the impact of i) a sector-specific improvement in technology or training and ii) a change in the skills used by sectors on a) sectoral distributions of skills, b) sectoral distributions of wages and c) the output of each sector.

To address these questions, I propose a new model that introduces assignment of workers to firms (within sectors) in the vein of Becker (1973) and Sattinger (1979) into a model of self-selection (across sectors) in the vein of Roy (1951). Workers' self-selection implies that sectoral distributions of skill are determined endogenously and depend, among other factors, on the technology used in each sector. The assignment of workers to firms introduces imperfect substitution of workers with different skill levels.[1] This is in contrast to existing models of self-selection which, starting with Heckman and Sedlacek (1985), assume that workers are perfect substitutes. However, perfect substitution is inconsistent with empirical evidence (Katz and Murphy, 1992).[2]

I study a model with two sectors, manufacturing and services, and derive sharp monotone comparative statics results. Firstly, I show that if the output produced by all manufacturing workers rises but the increase is greater for workers of higher skill, then a greater number of high-skill workers joins manufacturing. This increases wage inequality not just in manufacturing, but also in services.[3] Hence, wage inequality gets transmitted across sectors. If, however, the increase in output favors workers of lower skill, then some high-skill workers might leave for services, possibly decreasing the average skill in manufacturing. As a consequence of that, manufacturing might even contract.

Secondly, I show that if sectors start using more similar bundles of skill, the supply of skill *de facto* declines. This raises both wage polarization and wage inequality, but decreases output, as long as sectors are symmetric and jobs abundant, i.e. there are more

[1] To see why, note that in Sattinger's (1979) model firms and workers match in a fixed proportion: i.e. every firm hires the same, exogenously given number of workers. Hence, a firm cannot costlessly substitute a skilled worker with a greater number of less skilled ones. In aggregate, neither can sectors, because highly productive firms are scarce. For example, suppose that two high-skilled workers leave manufacturing. The first is replaced by $x > 1$ workers of low skill. To replace the second high skilled worker more than x low-skilled workers are needed, as the second worker's replacements match with less productive firms than the workers replacing the first worker.

[2] Perfect substitution of workers implies that if some high-skill workers leave manufacturing, wages increase by the same proportion for low and high-skill workers (the 'proportionality hypothesis'). Katz and Murphy (1992) provide evidence that changes in the relative supply of high and low-skilled workers significantly affect their relative wages.

[3] This is guaranteed to happen only if jobs are scarce, i.e. there are fewer firms than workers. Otherwise, the increase in output levels itself could – in some case – decrease inequality.

jobs than workers.

These results differ qualitatively from comparative statics that arise in existing sorting models, which further underscores the importance of properly accounting for the imperfect substitution of workers. It is worth noting that I am able to derive those results in the absence of functional form assumptions and despite the fact that, as a consequence of workers' imperfect substitution, wages in my model depend on the entire distribution of skills in a sector. I accomplish this by leveraging the well-known relationship between the distribution of skill and wages that arises under positive and assortative matching.

Overview. The paper is organized in three main sections.

Section 2 sets up the model and characterizes the unique equilibrium. In the model, there is a continuum of heterogeneous workers and two sectors, manufacturing (M) and services (S), each populated by a continuum of heterogeneous firms.[4] Each worker is endowed with a vector of fundamental skills $\mathbf{x}$ and each firm in sector $i \in \{M, S\}$ is endowed with a scalar productivity h_i. A match between a firm and a worker produces some surplus, which – in the absence of other inputs – can be interpreted as its output expressed in monetary terms.[5] The surplus produced by a match is determined by a surplus function that depends on the sector, the firm's productivity and the worker's skill. In particular, the two sectors use workers' skills in different proportions. In the competitive equilibrium, both workers and firms take wages as given, each worker sorts into the sector that pays a higher wage for her skill endowment and each firm hires at most one worker to maximize profits. Wages are set to clear the market.

To make the model tractable, I assume that in each sector the vector of fundamental skills $\mathbf{x}$ can be aggregated into a univariate index v_i in such a way that the vector (v_M, v_S) contains all the relevant information about $\mathbf{x}$.[6] The indices v_M, v_S are normalized to have standard uniform marginal distributions and, hence, are referred to as *relative skills*.[7] I assume, as is standard in the matching literature, that surplus is increasing in productivity, strictly increasing in relative skill and supermodular in productivity and relative skill.

Section 3 focuses on changes in technology and skill distribution that increase (decrease) the *difference* in surplus produced by manufacturing workers of any two relative skill levels. I will call this an increase (decrease) in the *vertical differentiation* of man-

[4]This means that the distribution of firms' productivity in each sector is exogenously determined, although the productivity distribution among firms that end up matched is not.

[5]Alternatively, this model can be seen as a reduced form of a model in which there are further inputs and there exists an explicit output function. In such a case, the surplus is the revenue minus the cost of non-labor inputs for an optimal choice of non-labor inputs.

[6]This is a two-sector version of the separability assumption from Chiappori, Oreffice, and Quintana-Domeque (2011) and it makes the relative skills an analogue of the tasks from Heckman and Sedlacek (1985).

[7]For example, consider a worker with relative skills vector $(0.25, 0.5)$. This means that 25% of the population is more skilled than this worker in manufacturing and that half of the population is more skilled than her in services.

ufacturing workers. Let me first illustrate the importance of vertical differentiation for sorting and its implications for output by the means of an example. Subsequently, I will discuss the implications for wage inequality.

Suppose there exist fewer firms than workers (jobs are *scarce*) and the vector of fundamental skills has three components: x_1 is manufacturing-specific, x_2 is services-specific and x_3 is a general purpose skill. Surplus in manufacturing increases strictly in the manufacturing-specific and general purpose skills, but does not depend on the services-specific skill; analogously for services. Suppose that the government wants to boost the total output produced in manufacturing and, to this end, decides to invest in population-wide training of x_1 in a way that shifts its distribution up. This increases the supply of a fundamental skill that is used *only* in manufacturing, whilst leaving the supply of fundamental skills used in services unchanged. In consequence, the output produced by a worker of any relative skill (v_M, v_S) increases for matches with manufacturing firms, but remains the same for matches with firms in services. In a model with perfect substitutes this would necessarily lead to an expansion of manufacturing. With imperfect substitution, however, it can easily backfire.

This can happen if the government's investment makes manufacturing workers less vertically differentiated.[8] How does this change firms' hiring choices in the short-term, before wages have the time to adjust? As the *difference* in surplus produced by different workers decreases, but the *difference* in wages does not, workers of high skill become relatively overpaid. In consequence, all firms want to hire a worker of lower relative skill than in the old equilibrium. Further, because I assume that the surplus produced in any match always covers the reservation payoffs of both parties, scarcity of jobs implies that all firms want to hire some worker. In consequence, more workers of high skill and fewer workers of low skill want to work in manufacturing than are demanded by manufacturing firms. Therefore, in equilibrium, wages rise for workers of low manufacturing skill and fall for workers with high manufacturing skill, forcing some marginal high-skill workers out of manufacturing, but drawing in additional low skill workers from services. The distribution of relative skill improves in manufacturing and worsens in services, both in first order stochastic dominance sense. Overall, services expand, whereas the impact on the total output produced in manufacturing is ambiguous.

Further, I show that a shock that directly affects manufacturing only, might well increase both absolute and relative wage inequality in services.[9] Several attempts have been made in recent years to infer the causes of the increases in wage inequality from cross-sector comparisons (e.g. Bakija, Cole, and Heim, 2010; Kaplan and Rauh, 2010,

[8]This would be the case if e.g. fundamental surplus was multiplicatively separable in the fundamental skills and concave in x_1, and the distribution of x_1 shifted up by a constant. Intuitively, think of a situation in which the government's policy provides more resources for teaching of all students, but focuses on the less able ones.

[9]It is worth noting, however, that both increases would be stronger in the directly affected sector.

2013). Ideally, such analyses should take into account wage inequality transmission. For example, Kaplan and Rauh (2013) concluded that the fact that "the increase in pay at the highest income levels is broad based" is more consistent with the superstar and scale effects (Rosen, 1981) explanations of raising wage inequality, than with an increase in managerial power or a weakening of social norms. If, however, an increase in wage inequality that originates in a narrow subset of sectors can spread across the economy, such conclusions would seem premature. This is less of a concern in models in which workers are perfect substitutes, where any change in the sectoral supply of skills will have the same relative impact on the wages of all workers. With imperfect substitution of skills, however, this is not the case. If jobs are scarce, then an increase in vertical differentiation in manufacturing draws in high-skill workers from services. The latter increases the wages of highly skilled services workers, as compared to workers of low skill. In consequence, wage inequality increases in both sectors.[10]

Although there are no strategic interactions in my model, my results do suggest that there might be strategic reasons for increases in wage inequality. I demonstrate this in Section 3.1.4, where I provide an example in which jobs are scarce and the two sectors (or regions in which they are concentrated) can make an investment in infrastructure, which increases the vertical differentiation of their workers. Rising vertical differentiation increases wage inequality and creates a negative sorting externality for the other industry. Therefore, each sector is willing to make the investment even in cases when this is not socially optimal.[11] In equilibrium, both sectors over-invest, which cancels out the sorting effect. Thus, the sectors end up with lower output (net of the investment's cost) and higher wage inequality.

Section 4 provides comparative statics results for changes in the interdependence of relative skills.[12] This analysis is restricted to the symmetric case of the model, in which the two sectors differ in the relative skill they use, but are identical otherwise. The main result is that an increase in the interdependence of relative skills constitutes a negative supply shock and, if jobs are abundant, increases *wage polarization* and wage inequality at the top. In particular, it increases the very top wages by more than the lowest wage, which themselves increase by more than wages in the middle of the distribution. I then go on to demonstrate that the interdependence of relative skills depends, among other things, on the skill contents of occupations, i.e. the degree to which manufacturing and services use the same fundamental skills in the production process. Therefore, the results in this section suggest a new channel through which routinization, i.e. the automation

[10]In particular, wage range increases in both sectors and top wages in manufacturing rise proportionately more than the lowest wages. Wages in services increase proportionately more for workers of high skill than of low skill as long as reservation wages are not too small.

[11]That is, decrease the sum of surplus produced in the economy, net of the investment's cost.

[12]I use concordance Scarsini (1984) as the notion of interdependence. Keeping marginal distributions unchanged, two random variables become more concordant if their joint distribution shifts up. For jointly normally distributed variables, concordance increases if and only if correlation increases.

driven change in task and skill content of occupations (Autor, Levy, and Murnane, 2003; Spitz-Oener, 2006), could have caused the well documented increase in wage polarization and inequality in the 1990s in the US (Acemoglu and Autor, 2011).[13]

The intuition for this result is simple: as each worker can work in at most one sector, higher interdependence means that more high skills will remain unused. Therefore, the *de facto* supply of relative skill in the economy declines. Together with symmetry, this implies that the distribution of relative skills deteriorates in each sector in the sense of first order stochastic dominance. To see why this implies a polarization of wages, it is easiest to start with the original model of Roy (1951), in which the mapping from relative skills to wages is given exogenously.[14] Nevertheless, as relative skills become more interdependent, wage distributions do change, because now most ranks in the wage distribution are occupied by workers of lower relative skill than previously. This is the *distribution effect*. The top and the bottom of the wage distribution are unaffected by the distribution effect, as they are still occupied by the highest and least skilled workers, respectively. Thus, the distribution effect increases wage polarization: wages fall for all interior ranks, but remain unchanged at the top and bottom. In my model, there is, further, the *wage effect*, as the deterioration in the sectoral distributions of relative skill increases wages for all but the least skilled workers. This creates, additionally, an increase in wage inequality at the top.[15]

Section 5 reviews further related literature and places my main contributions into it. Section 6 concludes. All proofs can be found in the Appendix.

2 The Model

In this section I set the model up, characterize the equilibrium and prove its uniqueness.

2.1 The Setup

There are two sectors – manufacturing and services – and two populations: workers and firms.

Workers There is a unit measure of workers, each endowed with a vector of fundamental skills $\mathbf{x} = (x_1, x_2 \ldots x_N) \in I_{\mathbf{x}} \subset \mathbf{R}^N$. Denote the distribution of $\mathbf{x}$ as F. Workers can either work for a firm and receive a wage or remain unemployed, in which case their payoff is normalized to 0.

[13]To the best of my knowledge, the idea that skill contents can constitute a negative supply shock and *through that* affect wage polarization has not been discussed in the literature so far.

[14]The relative skills of my model are a close analogue of the sector specific skills assumed by Roy (1951).

[15]The increase in wages counteracts the increase in wage polarization, but it can be shown that wages will fall for at least some ranks in the interior of the wage distribution.

Firms There is a measure R of firms, each endowed with vector $(z,i) \in I_z \times \{M,S\}$, where z denotes the firm's productivity, $i \in \{M,S\}$ denotes the sector the firm operates in and $I_z \subset \mathbf{R}$. The (exogenous) distribution of (z,i) is denoted as H_Z, whereas the measure of firms in sector i is denoted as $R_i > 0$. Each firm hires at most one worker. A worker-firm pair produces surplus according to the fundamental surplus function $\Pi : I_{\mathbf{x}} \times I_z \times \{M,S\} \to \mathbf{R}_{\geq \mathbf{0}}$. For example, if a manufacturing firm with productivity h hires a worker with skill $\mathbf{x}$, they produce surplus $\Pi(\mathbf{x}, z, M)$. The fact that the surplus function depends on the sector means that workers' skill and firms' productivity might be used differently in each industry. If a firm does not hire a worker, it receives a reservation profit normalized to 0.

2.1.1 Assumptions

Following Heckman and Sedlacek (1985) I assume that the fundamental surplus functions in each sector are *separable* in fundamental skills and productivity.[16]

Assumption 1 (Properties of Surplus)**.** In both sectors, fundamental surplus Π is separable in skills and productivity i.e. for each sector there exist mappings $v_i : I_{\mathbf{x}} \to [0,1]$ (*relative skill*), $h_i : I_z \to [0,1]$ (*relative productivity*) and $\pi_i : [0,1]^2 \to \mathbf{R}_{\geq \mathbf{0}}$ (*the reduced surplus*) such that:

A1.1 *Separability*: $\pi_i(v_i(\mathbf{x}), h_i) = \Pi(\mathbf{x}, z, i)$;

A1.2 *Differentiability*: π_i is twice continuously differentiable;

A1.3 *Increasing surplus*: $\frac{\partial}{\partial v_i}\pi_i > 0$, $\frac{\partial}{\partial h_i}\pi_i \geq 0$;

A1.4 *Supermodular surplus*: $\frac{\partial^2}{\partial v_i \partial h_i}\pi_i \geq 0$.

Separability means that the impact of fundamental skills $\mathbf{x}$ on surplus in sector i is fully captured by the one-dimensional index $v_i = v_i(\mathbf{x})$. Together with separability, the increasingness assumption (A1.3) implies that workers and firms can be totally ordered within each sector with respect to the surplus they produce.[17] Supermodularity (A1.4) implies that highly productive firms benefit more from hiring high-skill workers. This ensures that within-sector matching is positive and assortative and makes the model tractable. The comparative statics results would be unchanged if surplus was submodular.[18]

---

[16] In matching context, Chiappori, Oreffice, and Quintana-Domeque (2012) assume that there exists a single one-dimensional index that summarizes agents' preferences. My assumption is weaker, in that I only require firms' 'preferences' over workers to be the same within a sector, but allow them to differ across sectors.

[17] That is, if one manufacturing firm produces more surplus by hiring worker $\mathbf{x}'$ than worker $\mathbf{x}$, then all firms do; and analogously for firms.

[18] Whether my results would hold also for surpluses that are neither super- not submodular is an open question, but seems likely, given that the two extreme cases yield the same results.

Assumption 2 (Properties of the Copula)**.** Distributions F and H_Z and surplus Π are such that the joint distribution C of relative skill $(v_M, v_S) \in [0,1]^2$ as well as the distributions H_i of relative productivity $h_i \in [0,1]$:

A2.1 *Differentiability*: are twice continuously differentiable and

A2.2 *Full support*: have strictly positive, finite density on their respective supports.[19]

The full support assumption allows me to normalize the indices v_M, v_S and h_M, h_S in such a way that their marginal distributions are standard uniform, using the fact that Assumption 1 defines them only up to a monotone transformation.[20] This is why I refer to them as relative skills and productivities, respectively. Note that because the marginal distributions of v_M, v_S are standard uniform, C is a copula (Sklar, 1959). Further, the full support assumption precludes perfect (positive or negative) correlation, but otherwise allows for very general dependence structures. For example, for C belonging to the family of Gaussian Copulas, Assumption 2 allows for any correlation parameter $\rho \in (-1, 1)$.

The formulation of the model in terms of uniformly distributed relative skills will be referred to as *the canonical formulation*.[21] Apart from examples and applications, I will be working with the canonical formulation exclusively. For that reason, most of the time I will refer to v_i, h_i and π_i as, respectively, 'skill', 'productivity' and 'surplus', dropping the 'relative' and 'reduced'.

Assumption 3 (Non-Degenerate Solutions)**.** For any $i, j \in \{M, S\}$ with $j \neq i$ it is the case that $R_i < 1$ or $\pi_i(0, 1 - \frac{1}{R^i}) < \pi_j(1, 1)$.

This assumption is necessary and sufficient for all equilibria of this model to be non-degenerate, so that a positive measure of workers works in each sector.

2.1.2 Supply, Demand and Equilibrium

Supply of Relative Skills A worker with skill (v_M, v_S) who joins sector i receives wage $w_i(v_i)$, where $w_i : [0,1] \to \mathbf{R}$. Workers sort into the sector that maximizes their wages. A worker with skill (v_M, v_S) joins manufacturing if and only if:

$$w_M(v_M) \geq \max\{w_S(v_S), 0\}, \tag{1}$$

[19]It suffices if both conditions hold just on $(0,1)^2$ for distribution C. In particular, all results hold for Gaussian Copula.

[20]To see that this is a normalization, consider any v_i' and π_i' that meet Assumptions 1 and 2. Define the marginal distribution of V_i' as F_i. Then take $V_i = F_i(V_i')$ which ensures standard uniform marginal distribution; this gives $\pi_i(v_i, h_i) = \pi_i'(F_i^{-1}(v_i), h_i)$.

[21]The canonical formulation defines equivalence classes: any two models with the same canonical formulation will give raise to the same outcomes (i.e. wage and output distributions).

joins services if and only if:

$$w_S(v_S) > \max\{w_M(v_M), 0\}, \tag{2}$$

and remains unemployed otherwise.[22]

Sectoral *supply of relative skill* of level t, $S_i(t)$, is defined cumulatively, as the measure of workers with sector specific skill of at least t who join sector i, for given wage functions w_M, w_S:

$$S_M(t) = \Pr\Big(V_M \geq t,\ w_M(V_M) \geq w_S(V_S),\ w_M(V_M) \geq 0\Big), \tag{3}$$

$$S_S(t) = \Pr\Big(V_S \geq t,\ w_M(V_M) < w_S(V_S),\ w_S(V_S) \geq 0\Big). \tag{4}$$

Note that $S_i(0)$ gives us the total measure of workers who joined sector i. Further, together with the joint distribution C, either of S_M, S_S determines the other. [23]

Demand for Relative Skills The demand for skills in each sector is determined by the firms' hiring decisions, which in turn are driven by profit maximization, with firms taking the wage function as given. Firm h_i earns profit $r_i(h_i)$ and hires worker $v_i^*(h_i)$, where $r_i : [0,1] \to \mathrm{R}$ and $v_i^* : [0,1] \to [0,1]$, with:

$$r_i(h_i) = \max_{v\in[0,1]} \ \pi_i(v, h_i) - w_i(v), \tag{5}$$

$$v_i^*(h_i) \in \arg\max_{v\in[0,1]} \ \pi_i(v, h_i) - w_i(v). \tag{6}$$

Demand for skills is defined analogously to skill supply. The sectoral *demand for relative skill* of level t, $D_i(t)$, is equal to the measure of sector i firms that hire workers with sector specific skill of at least t, for a given wage function w_i:

$$D_i(t) = R_i \Pr\Big(v_i^*(H_i) \geq t,\ r_i(H_i) \geq 0\Big). \tag{7}$$

This definition assumes that profits are strictly increasing in productivity, which is the case as long as $\frac{\partial}{\partial h_i}\pi_i > 0$. A more general definition, holding also when surplus does not depend on productivity, is provided in Appendix A.[24]

[22] Of course, a worker for whom $w_S(v_S) = w_M(v_M)$ is indifferent and could join either sector; however, because such workers will be of zero measure in equilibrium, we can assign all of them to manufacturing without loss of generality.

[23] The exact relation between them is not trivial and is left for later. See proof of Theorem 2 in Appendix C for details.

[24] Additionally, my definition of v_i^* implies it is a function, which excludes the possibility of impure matchings. This greatly simplifies notation and is without loss of generality, because all matchings – even impure ones – will result in the same wage functions.

The Competitive Equilibrium I focus on the competitive equilibrium, which is defined as follows.

Definition 1 (Equilibrium)**.** An equilibrium is characterized by:
i) two *sectoral relative skill supply functions* $S_i : [0, 1] \to [0, 1]$, consistent with workers' sorting decisions and given by Equations (3) and (4);
ii) two *sectoral relative skill demand functions* $D_i : [0, 1] \to [0, 1]$, consistent with firms' profit maximization and given by Equation (7);
iii) two increasing *sectoral wage functions* $w_i : [0, 1] \to \mathbf{R}$, which clear the markets: $S_i(u) = D_i(u)$ for $i \in \{M, S\}$ and all $u \in [0, 1]$.

It is worth noting that, because this model is an assignment game, the competitive equilibrium coincides with the core (Gretsky, Ostroy, and Zame, 1992).[25] I require wage functions to be increasing for expositional ease: this can be ensured by assuming that a worker of higher skill can always pretend to have lower skill, but not the other way round.[26]

2.2 Characterization Strategy

To characterize the competitive equilibrium I deploy a two-step strategy. In the first step, I treat sectoral supply functions S_i as given and derive the wage functions that equate supply with demand in each sector. This is very similar to the problem first solved by Sattinger (1979). In the second step, I use those wages to find the sectoral supply functions in a manner somewhat similar to Roy's model.

2.2.1 First Step

In this part, I treat the sectoral supply functions as given and find the wage functions for which demand will equal supply. Let the *critical skill* v_i^c be the relative skill of the least skilled worker who joins sector i:

$$v_i^c = \sup\{v \in [0, 1] : S_i(v) = S_i(0)\}. \tag{8}$$

In equilibrium, $S_i(0)$ cannot be greater than $R_i = D_i(0)$, as otherwise the market would never clear; I will restrict attention to supply functions that meet this condition.

[25]An earlier version of this paper has used the core – and the related concept of a stable matching – to define the equilibrium. This older version is available at www.pawelgola.com/research.

[26]Even without this assumption, wages will be increasing on any relative skill interval for which demand is strictly increasing. However, for skill levels not demanded by any firm, wages can in principle be of any level that ensures no firm wants to hire workers with such skill. This does not matter for equilibrium supply and demand, but makes for cluttered exposition.

Proposition 1. In equilibrium, wage functions are such that:

$$w_i(v_i) = \int_{v_i^c}^{v_i} \frac{\partial}{\partial v_i}\pi_i\Big(v, 1 - \frac{S_i(v)}{R_i}\Big)\mathrm{d}v + w_i(v_i^c) \qquad \text{for } v_i \geq v_i^c \tag{9}$$

$$w_i(v_i) \geq w_i(v_i^c) + \pi_i\Big(v_i, 1 - \frac{S_i(0)}{R_i}\Big) - \pi_i\Big(v_i^c, 1 - \frac{S_i(0)}{R_i}\Big) \qquad \text{for } v_i < v_i^c. \tag{10}$$

where $w_i(v_i^c) \in [0, \pi_i(v_i^c, 1 - \frac{S_i(0)}{R_i})]$. If, however, $S_i(0) < R_i$ then $w_i(v_i^c) = \pi_i(v_i^c, 1 - \frac{S_i(0)}{R_i})$.

It is well-known that if the surplus function is supermodular, workers and firms match positively and assortatively: that is, the most productive firm matches with the worker of highest relative skill, the second most productive firm matches with the second most skilled worker, and so on. Accordingly, my Proposition 1 is essentially a restatement of Sattinger's (1979) famous result on wage functions under positive and assortative matching.[27] The reason for which the relationship between wages and the supply of skills needs to be of the form specified in Proposition 1 can be easily understood from the first order condition of the firm's hiring decision:

$$w_i'(v_i^*(h_i)) = \frac{\partial}{\partial v_i}\pi_i\Big(v_i^*(h_i), 1 - \frac{S_i(v_i^*(h_i))}{R_i}\Big).$$

Thus, the difference in wages paid to workers of marginally different skill is equal to the difference in the surplus they produce. The value of this marginal surplus, however, depends on the firm the worker is matched with. This, in turn, depends on the supply of relative skills in that sector: the fewer high-skill workers are available, the better match can be secured by any worker. The wage paid to the worker with critical skill v_i^c depends on whether workers are in short supply in that sector. If this is the case, then competition drives the profits of the least productive matched firm to 0.

2.2.2 Second Step

In the second step, I treat the sectoral wage functions w_M, w_S as given and derive the corresponding sectoral supply functions.[28] Note that by Proposition 1 wages are strictly increasing in each sector for $v_i \geq v_i^c$. This has two important implications for sorting. Firstly, any worker with relative services skill $v_S > v_S^c$ can earn a strictly positive wage and therefore will never choose to remain unemployed. Secondly, for any such worker

[27]The only difference is that my result explicitly allows for surplus to be weakly supermodular, i.e. allows for $\frac{\partial^2}{\partial v_i \partial h_i}\pi_i \geq 0$. Legros and Newman (2002) call 'famous' the result that under weakly supermodular surpluses any stable matching can be supported only by payoff schemes that support PAM. However, they don't provide any references and their Proposition 3 holds only for one-sided matching markets. Sattinger (1979) shows that the cross-derivative of the surplus function needs to be positive, but his argument holds only for strictly positive cross-derivatives.

[28]Technically, Equations 3 and 4 already do that. The challenge, however, is to express S_M, S_S as a function of wages in a way that will allow me to characterize the equilibrium.

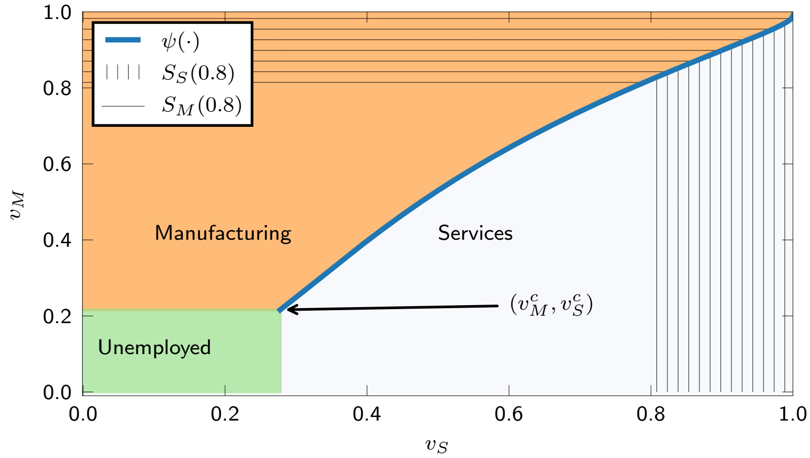

Figure 1: Graphical representation of the relation between the separating function and sorting. The hatched areas represent the space of workers with skill $V_i \geq 0.8$ who join occupation i; their supply $S_i(0.8)$ depends on how many workers reside in this space (which depends on the copula).

there will exist a cut-off value $\psi(v_S)$ of the relative manufacturing skill such that she will strictly prefer to join services if $v_M > \psi(v_S)$ and strictly prefer to join manufacturing if $v_M < \psi(v_S)$. Therefore, the sorting of workers to sectors can be easily expressed by the means of the critical skills v_M^c, v_S^c and the *separation function* $\psi : [v_S^c, 1] \to [v_M^c, 1]$, which is a mapping from manufacturing skill into the corresponding cut-off values of the services skill.

Lemma 1. In equilibrium, the critical skills in manufacturing and services are, respectively:

$$v_M^c = \sup\{v_M \in [0,1] : w_M(v_M) \leq \max\{w_S(0), 0\} \text{ or } v_M = 0\}, \tag{11}$$

$$v_S^c = \sup\{v_S \in [0,1] : w_S(v_S) \leq \max\{w_M(0), 0\} \text{ or } v_S = 0\}. \tag{12}$$

Provided that $v_M^c, v_S^c < 1$, it is the case that $w_M(v_M^c) = w_S(v_S^c)$.

The separation function depends on sectoral wage functions as follows:

$$\psi(v_S) = \max\{v_M \in [v_M^c, 1] : w_M(v_M) \leq w_S(v_S)\} \text{ for } v_S \geq v_S^S. \tag{13}$$

Note that for v_S's such that $w_S(v_S) \leq w_M(1)$ this implies:

$$w_S(v_S) = w_M(\psi(v_S)). \tag{14}$$

The critical skills and the separation function are sufficient to characterize the sorting of workers to sectors. This is depicted in Figure 1. By the definitions of v_M^c and v_S^c all but a zero measure of workers with $(v_M, v_S) < (v_M^c, v_S^c)$ remain unmatched. Manufacturing is populated by workers with $v_M > v_M^c$ and $v_M > \psi(v_S)$. Services are populated by workers with $v_S > v_S^c$ and $v_M < \psi(v_S)$. The set of workers unassigned by these rules is of measure zero. Thus, v_M^c, v_S^c and ψ fully determine the sectoral supply functions.

Lemma 2. Given the critical skills v_M, v_S and the separation function ψ, the supply of relative skill in manufacturing and services is, respectively:

$$S_M(v) = \begin{cases} \int_v^1 C_{v_M}(r, \phi(r))\mathrm{d}r, & v \geq v_M^c \\ S_M(v_M^c) & v < v_M^c, \end{cases} \quad S_S(v) = \begin{cases} \int_v^1 C_{v_S}(\psi(r), r)\mathrm{d}r, & v \geq v_S^c \\ S_S(v_S^c) & v < v_S^c, \end{cases}$$

where $\phi : [v_S^c, 1] \to [v_M^c, 1]$ depends on ψ as follows:

$$\phi(v_M) = \sup\{v_S \in [v_S^c, 1] : \psi(v_S) < v_M\}.$$

In equilibrium, the separation function determines sectoral supply of skill, sectoral supply of skill determines wages and wages determine the separation function. Any separation function that corresponds to supply and wage functions that hold in some equilibrium will be called an *equilibrium separation function*. Equilibrium separation functions can be found by substituting Equation (2) into the results in Proposition 1 and then substituting those into Equations (11)-(13).

Theorem 1. An equilibrium exists and is essentially unique. In particular, equilibrium supply, demand and separation functions are unique.

The proof relies on constructing a map, the fixed point of which is equivalent to the solution of (13) and finding a norm for which this map is a *contraction mapping*.[29] This proves that $\psi(\cdot)$ is unique *given* (v_M^c, v_S^c) – and also continuous in them. Then showing existence and uniqueness is merely a matter of proving that Equations (11)-(12) have a unique solution given the function $\psi(\cdot, v_M^c, v_S^c)$.

The existence of a unique equilibrium separation function implies trivially that there exists an equilibrium. It means, further, that the equilibrium is essentially unique, in that equilibrium supply and demand functions are unique. Equilibrium wage functions, however, are uniquely determined only for $v_i \geq v_i^C$ by Proposition 1 and even then possibly only up to the lowest wage $w_i(v_i^c)$.[30]

2.2.3 Sattinger and Roy

The first step in my characterization strategy is very similar to Sattinger (1979), the second to Roy (1951). This is not a coincidence: in fact, both one-sector matching and

[29] The norm I use is Bielecki's norm for a high-enough parameter λ.
[30] This is the case if $R_M + R_S = 1$, otherwise $w_i(v_i^c)$ is uniquely determined.

Roy-like models are nested within this framework.[31] In the case of Sattinger's model, that's fairly obvious: if one of the sectors is sufficiently more productive than the other and there is more than a unit measure of firms in it (so if Assumption 3 does not hold), then all workers will work in that industry and the model collapses to just one sector. The same would be the case if R_M or R_S was equal to 0.

As for Roy-like models, suppose that firms in both sectors are identical, in which case the surplus produced by any match depends on the worker's skill only.[32] If, on top of that, there is an abundance of firms in each sector ($R_i > 1$), then firms have no market power. Hence, workers receive the entire surplus ($w_i(v_i) = \pi_i(v_i)$) and their wage does not depend on the sectoral supply of skill, exactly as in Roy's model. In other words, Roy-like models can be seen as two-sector matching models in which all firms from the same sector are homogeneous.[33]

3 Changes to the Reduced Surplus Functions

In this section I study the effects of changes to the reduced surplus functions. I focus on changes that are sector-specific, as I am interested in the way in which shocks and policy interventions spread through the economy.

In all comparative statics exercises in this paper I compare the equilibria of two specifications of the model – the *old* and the *new* one.[34] The old specification is denoted by c_1 and the new one by c_2. For example, $\pi_M(\cdot, c_1)$ is the old reduced surplus function in manufacturing, whereas $\pi_M(\cdot, c_2)$ is the new one. I start the analysis by defining two key concepts, vertical differentiation of workers and an increase in surplus levels.

Definition 2 (Vertical Differentiation)**.** Workers in manufacturing become (strictly) more *vertically differentiated* if, for any $h_M \in [0, 1]$ and any $0 \leq v'_M < v''_M \leq 1$:

$$\pi_M(v''_M, h_M, c_2) - \pi_M(v'_M, h_M, c_2)(>) \geq \pi_M(v''_M, h_M, c_1) - \pi_M(v'_M, h_M, c_1).$$

Workers become more vertically differentiated in manufacturing if the difference in the surplus they produce increases for all levels of relative skill and all firms.[35] This is

[31]The actual models by Sattinger and Roy are not, strictly speaking. In the former case, the reason is that Sattinger allows for cases in which both firms and workers are unemployed, which is ruled out here by the assumption of positive surpluses – so only certain special cases of his model are nested. In the latter, the reason is that Roy uses bi-variate log-normal distribution of skills, which is not defined over an rectangle – however, we can get an arbitrarily good approximation of Roy's model, by using bi-variate log-normal distribution, truncated arbitrarily high and arbitrarily close to zero. This is done in Section 3.1.1 and Appendix D.

[32]This implies $\Pi_z(\bullet) = 0$, which is allowed by my assumptions.

[33]An example would be the Gaussian-Exponential specification, which will be introduced in Section 3.1.1, with $R_M, R_S > 1$ and $\gamma_M = \gamma_S = 0$.

[34]With both specifications meeting all conditions from Section 2, including Assumption 3.

[35]This is equivalent to an increase in the marginal surplus of relative skill for all matches in manufac-

equivalent to an increase in the *spread* of the distribution of $\Pi(\mathbf{X}, h, M)$ in the sense of Bickel and Lehmann (1979) (for all $h \in [0, 1]$).

Definition 3 (Increase in Levels)**.** The level of surplus produced in manufacturing *increases universally* if, for all $(v_M, h_M) \in [0, 1]^2$, $\pi_M(v_M, h_M, c_2) \geq \pi_M(v_M, h_M, c_1)$.

The level of surplus produced in manufacturing increases universally if any match in manufacturing produces more surplus than before. This is a very strong condition. Yet, I will demonstrate in this section that, despite its strength, a universal increase in manufacturing surplus level is not sufficient to generate an increase in the equilibrium supply of skill in manufacturing. Further, I will provide an example showing that this condition does not even guarantee – on its own – that the total surplus (so the sum of surpluses produced by all firms) produced in manufacturing increases.

In Section 3.1 I first show that if jobs are scarce ($R_M + R_S \leq 1$) an increase in vertical differentiation of manufacturing workers is sufficient for an increase in the equilibrium supply of relative skill in that sector, regardless of what happens to the levels of surplus. Then I consider a series of structured examples, which a) illustrate that vertical differentiation of workers can increase due to changes in either the distribution of fundamental skills $\mathbf{x}$, the distribution of fundamental productivities $(z|i)$ or changes in the fundamental surplus function Π; b) demonstrate that a universal increase in manufacturing surplus levels accompanied by a fall in vertical differentiation can result in lower total surplus in manufacturing and c) explore the implications of my results. In Section 3.2 I then consider the abundant jobs case ($R_M + R_S > 1$) and show that in this case an increase in both differentiation and levels is needed to ensure higher equilibrium supply of relative skill in manufacturing.

3.1 Scarce Jobs

Scarcity of jobs implies that all firms are matched, the measure of firms in each sector is fixed and the positive and assortative matching function in each sector is equal to the distribution of the relevant relative skill in that sector, $G_i : [v_i^c, 1] \to [0, 1]$:

$$G_i(v) = \Pr\Big(V_i \leq t | w_M(V_M) \geq w_S(V_S), w_M(V_M) \geq 0\Big).$$

Lemma 3. If $R_M + R_S \leq 1$ then $S_i(0) = R_i$ which implies that $G_i(v_i) = 1 - \frac{S_i(v_i)}{R_i}$. If jobs are strictly scarce ($R_M + R_S < 1$), then additionally $w_M(v_M^c) = w_S(v_S^c) = 0$.

All firms must be matched if jobs are scarce, because otherwise the unmatched firms would hire the unmatched workers. With strictly scarce jobs, the competition from the

turing: $\frac{\partial}{\partial v_M}\pi_M(v_M, h_M, c_2) \geq (>) \frac{\partial}{\partial v_M}\pi_M(v_M, h_M, c_1)$ for all $(v_M, h_M) \in [0, 1]^2$.

unemployed workers drives the wages of the least skilled employed workers down to their reservation payoff.

Proposition 2. If jobs are scarce and workers in manufacturing become more vertically differentiated, then (i) the distribution of relative skill in manufacturing improves in first order stochastic dominance sense; (ii) the distribution of relative skill in services deteriorates in first order stochastic dominance sense. If, further, the increase in differentiation is strict, then strictly more relative skill is supplied to manufacturing and strictly less to services.

This result can be best understood by focusing on its impact on the demand for relative skills. How do manufacturing firms' hiring decisions change after an increase in vertical differentiation, but before the wage functions had time to adjust?[36] Because the *difference* in surplus produced by different workers has increased, but the *difference* in wages has not, workers of high skill become relatively underpaid. In consequence, all firms want to hire a worker of higher relative skill than in the old equilibrium: the demand for relative skill shifts up.[37] This is depicted in the left panel of Figure 2. Note that with scarce jobs, all firms want to hire some worker and, hence, their hiring decisions depend on the differences in surplus only: the levels play no role at all. The upward shift in skill demand draws in additional marginal high-skill workers into manufacturing and causes some marginal low skill workers leave for services, which is depicted in the right panel of Figure 2.

In addition to its impact on sorting, the upward shift in demand for relative skill affects wages as well. To bring the most interesting results into focus, in the discussion on wages I focus on strict increases in vertical differentiation; all results extend easily to the more general case. It is also worth noting that with strict supermodularity results (i) and (iii) below hold strictly.

Proposition 3. If jobs are scarce and workers in manufacturing become strictly more vertically differentiated, then (i) in services, wages increase for all workers and the higher the relative skill the greater the increase; (ii) in manufacturing, wages increase strictly for a positive mass of the most relative skilled workers; and (iii) in both sectors the range of the wage distribution increases, strictly in manufacturing. If jobs are strictly scarce, then,

[36] That is, if firms still have to pay wage $w_M(v_M)$ for relative skill v_M.

[37] The reasoning here is the same as in the monotone comparative statics results in Milgrom and Shannon (1994), with vertical differentiation being an analogous condition to increasing differences. Accordingly, an increase in vertical differentiation is a stronger condition that is needed for Proposition 2 to hold. What is sufficient is that the marginal surplus of relative skill increases for all existing matches, rather than globally (see Theorem 2 and proof thereof in Appendix C). For small changes in the surplus function this weaker condition is equivalent to an upward shift in demand. In general, however, these two are not equivalent: as an example, consider a change in surplus from $v_M h_M + 1$ to 1 for all $v_M < 1$ and 2 for $v_M = 1$. Under old wages, all firms would like to hire $v_M = 1$. However, the measure of such workers is zero, so in equilibrium essentially all manufacturing firms would be hiring only those workers whose services skill is to low to enable them employment in services.

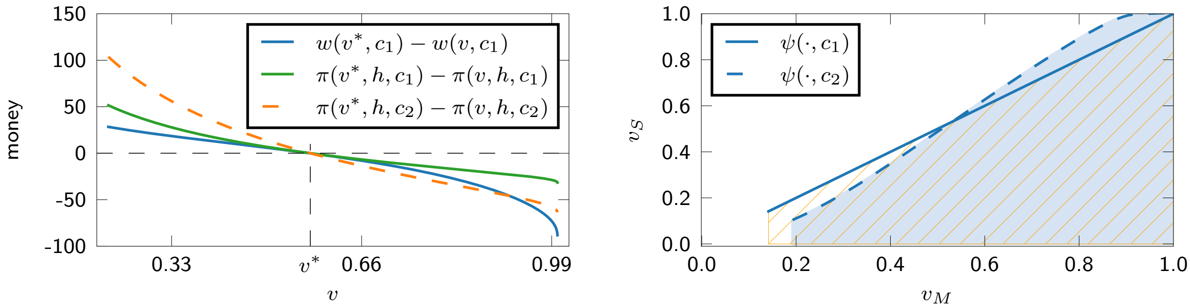

Figure 2: Left panel: v^* is the skill of the worker hired by firm h in the old equilibrium. All workers with skill $v \in (v^*, v')$ are strictly preferred by the firm after the increase in differentiation, but before wage functions had time to adjust.
Right panel: change in equilibrium separating function.

further, the increases in wages and wage range in services are strict and manufacturing wages fall strictly for a positive mass of the least relative skilled workers.

Wages increase for manufacturing workers of high relative skill, as those workers are now in higher demand; similarly, wages fall for low-skilled workers in manufacturing, as those workers are now less in demand.[38] It is not true, in general, that the wage of a worker with higher skill increases by more than the wage of a worker with lower skill. This is because there are two effects in play in manufacturing. On the one hand, the increase in vertical differentiation increases the wages of high-skill workers by more. On the other hand, however, the increase in the supply of skill means that workers with higher skill face tougher competition than previously, which lowers their wages in comparison to manufacturing workers of lower skill. In the services sector, the fall in the supply of skill is the only force affecting wages. Thus, wages become more spread out in the Bickel and Lehmann (1979) sense: i.e. the difference in wages earned by workers of any two levels of relative skill increases.

In addition to the results in Proposition 3, in the case of services we can say a little bit more about measures of inequality other than range. Keeping relative skill distribution constant, the increase in the spread of services wages raises its variance.[39] This means that variance increases for the *incumbents*, i.e. the workers who work in services both in the old and the new equilibrium.[40] It does not, however, necessarily imply an overall increase in wage variance in services, as the distribution of relative skill does change.[41] Keeping the wage function constant, the fall in the supply of relative skill has an ambiguous

[38]If $R_M + R_S = 1$ then it is possible that $v_M^c = 0$ both in the old and new equilibrium, in which case $w_M(0) = w_S(0) = 0$ and there is no change in wage for the least skilled worker.

[39]Follows from 3.B.25 in Shaked and Shanthikumar (2007).

[40]Formally, the set of incumbents is defined as $I = \{(v_M, v_S) : w_S(v_S, c_1) > w_M(v_M, rho_1) \text{ and } w_S(v_S, c_2) > w_M(v_M, rho_2)\}$.

[41]In fact, it is the change in relative skill distribution that spreads out sector two wage in the first place.

effect on variance. Hence, the overall effect of an increase in vertical differentiation of manufacturing workers on wage variance in services is ambiguous.

Both wage range and variance are measures of absolute wage inequality. In many contexts, however, relative inequality might be of more interest. In manufacturing, the ratio of wages earned by highest and lowest skilled workers increases strictly, because high-skill workers earn strictly more and workers of lowest skill strictly less than previously. In services, however, the direction of the change in relative inequality depends on the workers' reservation payoff. So far, the reservation payoff has been normalized to 0, as it is of no import for the equilibrium, changes in sorting patterns and changes in wage levels.[42] It does, however, matter for the change in relative inequality in services: the same increase in the difference in wages can result in higher or lower ratio of wages, depending on what the original wage level was. In particular, if jobs are strictly scarce and the reservation payoff is high enough, relative wage inequality increases in services, in that the ratio of wages earned by high and low-skilled workers goes up.[43] Note that this could never happen in a model where workers are perfect substitutes.

Proposition 4. If jobs are scarce and manufacturing workers become more vertically differentiated, then both total surplus and profits fall in services. If, further, the level of surplus in manufacturing increases universally, then total surplus produced in manufacturing increases as well, as does total surplus produced in the economy.

As was the case with wages, profits and total surplus in services are affected only by the fall in skill supply. Lower supply of skills means that firms are matched with less skilled workers, which decreases both their profits and the surplus they produce. In manufacturing, total surplus and profits are affected both by the increase in skill supply and the change in the surplus function. If surplus levels increase universally, then the surplus produced by any manufacturing firm increases and, in consequence, total surplus in manufacturing rises. If vertical differentiation increases, but surplus levels fall universally, the effect on production in manufacturing is ambiguous. An example is provided in Section 3.1.2.

[42]As longs as the assumption that every match produces more than the sum of the workers' and firms' reservation payoffs.

[43]We can write the ratio of wages earned by two different workers in services as:

$$\frac{w_S(v''_S) + p_w}{w_S(v'_S) + p_w} = 1 + \frac{w_S(v''_S) - w_S(v'_S)}{p_w + w_S(v'_S)},$$

where p_W denotes the workers' reservation payoff and $w_S(\cdot)$ is the wage function under $p_W = 0$. This increases if and only if:

$$p_w > \frac{(w_S(v''_S, c_1) - w_S(v'_S, c_1))w_S(v'_S, c_2) + (w_S(v''_S, c_2) - w_S(v'_S, c_2))w_S(v'_S, c_1)}{(w_S(v''_S, c_2) - w_S(v''_S, c_1)) - (w_S(v'_S, c_2) - w_S(v'_S, c_1))}.$$

If jobs are scarce and for v''_M close to 1 and v'_M close to $v^c_S(c_1)$ both the denominator and numerator of the RHS have to be strictly positive by Proposition 3 and thus there exists a high enough p_w for which relative inequality increases.

The change in profits of manufacturing firms depends, besides the level of surplus and the supply of skill, also on the vertical differentiation of firms, i.e. the change in $\frac{\partial}{\partial h}\pi_M$. This is explored in more detail in Section 3.2.

3.1.1 Gaussian-Exponential Specification

In the rest of my analysis of the scarce jobs case I provide three examples, each of them illustrating a different type of change that could cause an increase (or decrease) in vertical differentiation of workers. I do this using the following, Gaussian-Exponential specification of the model.

The vector $\mathbf{x}$ of fundamental skills is jointly normally distributed with mean $\boldsymbol{\mu}_x$ and covariance matrix $\boldsymbol{\Sigma}_x$. Productivity z conditional on the firm operating in sector i is distributed uniformly on $[\underline{\beta}_i, \bar{\beta}_i]$, with $\bar{\beta}_i > \underline{\beta}_i \geq 0$. The fundamental surplus function in sector i is given by $\Pi(\mathbf{x}, z, i) = \left(A_i + \frac{1-e^{-\delta\boldsymbol{\alpha}_i\mathbf{x}^T}}{\delta}\right)z^{\gamma_i}$, where $\boldsymbol{\alpha}_i = [\alpha_{i1}, \alpha_{i2}\ldots\alpha_{iN}]^T$ is a N-dimensional vector of (fundamental) skill requirements, $A_i > 0$ determines the extent to which productivity influences surplus *irrespective* of skill, δ_i determines the curvature of surplus as a function of skill and $\gamma_i \geq 0$ determines the extent to which skills and productivity are supermodular. Note that with $\delta < 0$ this model is equivalent to a model in which $\ln(\mathbf{x})$ is jointly normally distributed and surplus is multiplicative: $(A_i + \boldsymbol{\alpha}_i\mathbf{x}^T)z_i$.[44]

In the Gaussian-Exponential specification, each sector uses some linear combination of the fundamental skill components in its production process. Denote the linear combination of skills required by sector i as $v_i'(\mathbf{x}) = \boldsymbol{\alpha}_i\mathbf{x}^T$, which gives us a vector of indices (v_M', v_S') that are $\mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distributed, with

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_M \\ \mu_S, \end{bmatrix} = \boldsymbol{\alpha}\boldsymbol{\mu}_x,\ \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_M^2 & \rho\sigma_M\sigma_S \\ \rho\sigma_M\sigma_S & \sigma_S^2 \end{bmatrix} = \boldsymbol{\alpha}\boldsymbol{\Sigma}_x\boldsymbol{\alpha}^T \text{ and } \boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_M \\ \boldsymbol{\alpha}_S \end{bmatrix}.$$

Normalising v_i' and $z|i$ in such a way that their marginal distributions are standard normal, yields the following canonical formulation of the model:

$$\pi_i(v_i, h_i) = \left(A_i + \frac{1}{\delta}\left(1 - e^{-\delta\left(\Phi^{-1}(v_i)\sigma_i + \mu_i\right)}\right)\right)\left((\bar{\beta}_i - \underline{\beta}_i)h_i + \underline{\beta}_i)\right)^{\gamma_i},$$
$$C(v_M, v_S) = \boldsymbol{\Phi}_\rho\left(\Phi^{-1}(v_M), \Phi^{-1}(v_S)\right),$$

where $\boldsymbol{\Phi}_\rho$ is the cdf of a standardised bivariate normal distribution with correlation ρ and Φ is the cdf of univariate standard normal distribution.[45] In the three following

[44] Note that if jobs were abundant in both sectors ($R_i > 1$) and $\delta = -1$, $A_i = 1$ and $\gamma_i = 0$, then this specification reduces to the model in Roy (1951): the logarithm of skills is joint normally distributed and the surplus function does not depend on firm's productivity.

[45] The Gaussian-Exponential specification does not meet Assumption 1, as it is not defined for $v_i = 1$ and not differentiable for $v_i \in \{0, 1\}$. Formally, I solve this problem by working with a surplus function

examples, the assumption that jobs are scarce is maintained; in general, however, the GE specification is applicable also if jobs are abundant.[46]

3.1.2 Public Investment in Education

Suppose the government would like to boost the total surplus produced in manufacturing, by investing in the education of a skill that is used more intensively in manufacturing than services.[47] To make things really simple, I focus on the extreme case in which one of the fundamental skills (x_1) is (nearly) manufacturing-specific, with $\alpha_{S1} \approx 0$.[48] Investment in the quality of x_1 training increases its mean μ_1. Table 1 provides a numerical example with three fundamental skills: x_1 is manufacturing-specific, x_2 is services-specific, whereas x_3 is a general purpose skill, used with equal weight in both sectors.

The investment in skill x_1 will have two effects: the *direct* effect and the *sorting* effect. The direct effect is the change that would happen if there was no re-sorting of workers; the sorting effect captures the impact of re-sorting. The direct effect is positive, because an improvement in μ_{M1} increases π_M for any possible match. The direction of the sorting effect, however, depends on whether the investment in x_1 increases or decreases the vertical differentiation of manufacturing workers. As a benchmark, note that for univariate normally distributed variables, a change in μ has no effect on the spread of their distribution. Therefore, for surplus that is linear in skill ($\delta \to 0$) an investment in μ_1 does not affect workers' vertical differentiation; hence, there is no sorting effect and manufacturing expands purely due to the positive direct effect.

The situation is quite different if δ is negative. In this case the fundamental surplus function is convex in skill and, therefore, any improvement in the distribution of x_1 in FOSD sense increases the differences in the surplus produced by workers of different relative skill. In other words, manufacturing workers become more vertically differentiated. This causes a positive sorting effect in manufacturing, in line with Proposition 2. At the same time, the supply of skills decreases in services, causing its contraction.

with an additional truncation parameter a_i, which approaches π_i as $a_i \to 0$ (see Appendix D for details). Then, in simulations, I just set a_i close to 0. This procedure is equivalent to first aggregating $\mathbf{x}$ and then truncating the aggregate v_i' at values far removed from the mean. For expositional simplicity I will ignore this problem in the main body.

[46]With no changes if $\delta < 0$; otherwise, v_i' needs to be truncated to ensure that $\pi_i > 0$, because – otherwise – π_M would admit negative values for $v_i \approx 0$, regardless of the value of other parameters. See Appendix D for details.

[47]There exists a plethora of reasons why this might be the governments goal, some of them economically justifiable, others less so. The government might be worried about deindustrialization for strategic reasons, or because manufacturing sector is a powerful engine for productivity growth and experiences unconditional productivity convergence, unlike services (Rodrik, 2013). On the other hand, the government might be focused on manufacturing for purely political reasons: thinking manufacturing growth is more salient or having electorate and donors that are concentrated in manufacturing.

[48]If the model was initially symmetric, this assumption would not be needed, $\alpha_{M1} > \alpha_{S1}$ being sufficient. In the asymmetric case, however, the exact shape of the equilibrium separation function matters for how much more important x_1 needs to be in manufacturing than services for Proposition 2 to apply.

Sector	Type of effect	% change in output for $\delta =$ −1 (convex)	0 (linear)	1 (concave)
Manufacturing	direct	1.12	0.48	0.64
	sorting	1.34	0	-0.68
	overall	**2.04**	**0.48**	**-0.04**
Services	**overall**	**-0.82**	**0**	**0.47**

Table 1: The effects of an increase in μ_1 from 0 to 1.
Computed for Gaussian-Exponential specification: $N = 3$; $\mathbf{x} \approx \mathbf{N}(\mathbf{0}, \mathbf{I})$ ($\mathbf{I}$ is the identity matrix);$\boldsymbol{\alpha}_S = [0.2, 0, \sqrt{0.96}]^T$, $\boldsymbol{\alpha}_S = [0, 0.2, \sqrt{0.96}]^T$; $A_i = 41$; $\underline{\beta}_i = 2$, $\bar{\beta}_i = 3$, $\gamma_i = 2$, $R_i = 0.49$.

If δ is positive, then surplus in services is concave in skills. By analogous reasoning, an increase in μ_1 makes manufacturing workers less vertically differentiated. This is plausible in the context of education, where often the biggest gains in productivity come from closing the gap between students that are already high achieving, and those who are underperforming. However, if this is the case, then the sorting effect is negative: the investment in the manufacturing-specific skill makes firms less willing to hire workers of high relative skill and worsens sorting into that sector. In extreme cases, the negative sorting effect can dominate the positive direct effect and cause a decline in manufacturing (Table 1 provides an example). Thus, this seemingly straightforward policy can easily backfire in this model. Note that this could never happen in a selection model with perfect substitution of workers.[49]

The results for the concave case have further counterintuitive implications. For example, if the government had to choose whether to invest in x_1 and x_2 then, wanting to boost manufacturing, its best option might be to invest in the *services-specific* skill. Furthermore, such an investment could lead to a contraction of services, implying that manufacturing expands by more than the whole economy. This puts in doubt the conclusions from Justman and Thisse (1997) and Poutvaara (2008), who argue that if workers can migrate between regions (sectors) then governments will necessarily under-invest in the training of skills, foreign (here: services) specific skills in particular.[50]

As an aside, this example is a good showcase for the canonical formulation's ad-

[49]In such models, the change in skill supply has two effects: it changes sorting and the relative price of the manufacturing task (which is analogous to the relative manufacturing skill in this model). Suppose manufacturing contracted following an increase in the supply of the manufacturing-specific skill. This is possible only if the supply of the manufacturing task decreased, which – in turn – implies that the supply of the services task increased. But this would raise the relative price of the manufacturing task, which – together with the increase in the supply of x_1 – would lead to an expansion of manufacturing; contradiction.

[50] This literature models the strategic interactions much more carefully, which is beyond the scope of this paper; nevertheless, this example should make it clear that in my model it is not certain that there will be underinvestment in foreign-specific skills.

vantages. In terms of the fundamental formulation, an increase in μ_1 affects only the (multivariate) sectoral supply of (fundamental) skills. In equilibrium, sectoral supply of fundamental skills is driven partly by the increase in μ_1 and partly by sorting, generally with an ambiguous end effect. In the canonical formulation, this is decomposed into an improvement in the reduced surplus function, which captures the direct effect of the increase in μ_1; and a change in the sectoral supply of relative skill, capturing the effect on sorting. In particular, this allows for a meaningful comparison of the quality of workers who sort into manufacturing, even in cases when the distribution of fundamental skill has changed.

3.1.3 Productivity Distribution and Inequality Transmission

I will now study the impact of an improvement in the distribution of firms' productivity: the most natural interpretation of such an improvement is the introduction of a more efficient technology, although trade liberalization could have a similar effect (see Melitz, 2003; Sampson, 2014).[51] In the Gaussian-Exponential specification, the distribution of firms' productivity z in sector i depends on two parameters: $\underline{\beta}_i$ and $\bar{\beta}_i$. An increase in either of those improves the distribution of productivity in first order stochastic dominance sense, but their effect on the spread of productivity distribution differs. Specifically, an increase in $\underline{\beta}_i$ makes firms more similar, in the sense that $\frac{\partial}{\partial h}\pi_i$ falls, whereas $\bar{\beta}_i$ has the opposite effect. This difference matters for how firms' profits change, but – in the baseline model discussed here – has the same effect on sorting, qualitatively.[52] This is because any improvement in the distribution of $z|i$ increases the fundamental productivity of a firm with relative productivity h_i. As fundamental surplus is strictly supermodular in the Gaussian-Exponential specification (for $\rho_i > 0$), this means that for any relative productivity h_i the difference in the surplus produced by workers of high and low relative skill increases. Thus, surplus levels increase and manufacturing workers become more vertically differentiated. This reasoning applies more generally than just for the Gaussian-Exponential specification.

Lemma 4. If $\Pi(\mathbf{x}, z, i)$ is strictly increasing in productivity then, under Assumptions 1 and 3, an improvement in the distribution of $(Z|M)$ in the first order stochastic dominance sense implies that manufacturing workers become more vertically differentiated.

Therefore, if jobs are scarce, Propositions 2 to 4 hold if either $\underline{\beta}_i$ and $\bar{\beta}_i$ increase. Consequently, the adoption of more efficient technologies makes the least skilled workers in that sector worse off (Proposition 3).[53] Crucially, the relation between the distribution

[51] Note, however, that in Melitz model trade liberalization changes also the mass of firms, which is not the case here.

[52] With endogenous entry of firms, the effect on sorting would be different.

[53] This is reminiscent of the effects in the task-based model of Acemoglu and Autor (2011), where e.g. high skill augmenting technology could result in a fall in wages for medium skilled workers.

		% change in		
Sector	Type of effect	surplus	wage variance	wage range
Manufacturing	direct	19.37	82.39	39.42
	overall	20.28	-13.52	14.31
Services	overall	-0.77	58.44	13.54

Table 2: The effects of an increase in $\underline{\beta}_M$ from 2 to 2.5.
Computed for Gaussian-Exponential specification: $N = 3$; $\mathbf{x} \approx \mathbf{N}(\mathbf{0}, \mathbf{I})$ ($\mathbf{I}$ is the identity matrix); $\boldsymbol{\alpha}_S = [0.2, 0, \sqrt{0.96}]^T$, $\boldsymbol{\alpha}_S = [0, 0.2, \sqrt{0.96}]^T$; $\delta = 1$; $A_i = 41$; $\underline{\beta}_i = 2$, $\bar{\beta}_i = 3$, $\gamma_i = 2$, $R_i = 0.49$.

of productivity and workers' differentiation implies that technological change which is restricted to just one sector can increase wage inequality in many industries. A numerical example using the concave case of the Gaussian-Exponential specification is provided in Table 2. In this example, $\underline{\beta}_M$ increases and all manufacturing firms become more productive, but also more similar. This affects wage inequality in manufacturing directly, increasing drastically both wage range and wage variance. However, the sorting effect works in the opposite direction: wage inequality falls in manufacturing and increases in finance. Overall, and in line with Proposition 2 wage range increases in both sectors. The increase is stronger in manufacturing (which is true in general, not just in this example), but the magnitude is very similar in both sectors. The results for variance are even more striking, albeit less general: variance *decreases* in manufacturing overall, but increases in services. This suggests that, for example, Rosen's (1981) superstar effect could be driving the increases in wage inequality also in sectors that are not directly affected by the improvements in communication technology.

The effect of an improvement in $\underline{\beta}_i$ on total surplus is as implied by Proposition 4 and Lemma 4: there is an increase in manufacturing and fall in services.

3.1.4 Inter-regional Competition for Skills as a Force for Inequality

Suppose that the two sectors are concentrated in two distinct regions: the manufacturing region and the services region. In the English context this could be thought of as the North of England and London. The manufacturing region considers an investment in regional broadband infrastructure. I will assume that such an investment would improve the surplus produced by all matches in manufacturing, but particularly so for matches involving high-skill workers: for suggestive evidence that broadband internet is indeed a complement with skill see Akerman, Gaarder, and Mogstad (2015). Formally, this will be captured by an increase in the exponent γ_M.[54]

[54] I will assume here that $\underline{\beta}_M \geq 1$, to ensure that an increase in γ_M raises surplus.

		% change in		
Sector	Type of effect	surplus	wage variance	wage range
Manufacturing	direct	9.87	23.56	11.49
	overall	10.30	38.15	11.21
Services	overall	-0.41	-14.33	0.24

Table 3: The effects of an increase in γ_M from 2 to 2.1.
Computed for Gaussian-Exponential specification: $N = 3$; $\mathbf{x} \approx \mathbf{N}(\mathbf{0}, \mathbf{I})$ ($\mathbf{I}$ is the identity matrix); $\boldsymbol{\alpha}_S = [0.2, 0, \sqrt{0.96}]^T$, $\boldsymbol{\alpha}_S = [0, 0.2, \sqrt{0.96}]^T$; $\delta = -1$; $A_i = 41$; $\underline{\beta}_i = 2$, $\bar{\beta}_i = 3$, $\gamma_i = 2$, $R_i = 0.49$.

An increase in γ_M increases both the level of surplus and workers' vertical differentiation in manufacturing. Thus, its direct effect is an increase in total surplus and an increase in inequality. After workers re-sort, total surplus increases further, whilst wage range decreases. In the numerical example provided in Table 3 resorting increases wage variance in manufacturing and decreases it in services, but this effect is not general (as we have seen in Section 3.1.3). In services, total surplus falls and wage range increases.

Suppose that the two sectors are symmetric and both regions need to decide whether to invest in broadband. In particular, suppose that the direct increase in total surplus is lower than the investment's cost, but the overall increase is well worth the cost, regardless of what the other region decided. The two regional governments have to simultaneously choose whether to invest in broadband infrastructure and their payoff is the percentage change in average surplus net of the investment cost per capita (if any). The strategic form of this game is presented below, for the specification from Table 3 and investment cost per-capita equal to 10% of the pre-game average surplus.

		Manufacturing	
		$\gamma_M = 1$	$\gamma_M = 1.1$
Services	$\gamma_S = 1$	0%, 0%	-0.41%, 0.3%
	$\gamma_S = 1.1$	0.3%, -0.41%	-0.13%, -0.13%

This is clearly a Prisoner's Dilemma, with both governments investing in broadband in equilibrium. This means that neither of them improves the supply of skills in their region and their gains are limited to the investment's direct effect, which is not worth its cost. As a consequence, wage inequality increases in both regions as well (see Table 3).[55] In general, investments that increase workers' vertical differentiation impose a negative externality on other regions, whilst also increasing overall wage inequality. This suggests that interregional, or in fact international, competition for high-skill workers is a force for greater inequality.

[55] Note that because there is no resorting in the equilibrium of the investment game, both sectors will experience just the direct effects of the investment.

3.2 Abundant Jobs

If jobs are abundant ($R^1 + R^2 > 1$), the level of surplus plays a role in determining the extensive margin of firm's hiring decision: that is, whether to hire any worker at all or to exit the market. In particular, if there was no change in workers' vertical differentiation in manufacturing but the level of surplus fell, then some low productivity manufacturing firms would likely decide to leave the market, which would shift the demand for relative manufacturing skill down. To address this, in the section I focus on changes in reduced surplus that both increase workers' vertical differentiation and increase the levels of surplus.

Proposition 5. If a) jobs are abundant, b) surplus levels increase universally in manufacturing and c) manufacturing workers become more vertically differentiated, then more relative skill is supplied to manufacturing and less to services in equilibrium ($S_M(v_M)$ increases and $S_S(v_S)$ falls for all v_i). If the increase in vertical differentiation is strict, then the changes in relative skill supply are strict for some v_i.

Let us again consider the change in manufacturing firms' hiring decisions after the reduced surplus function has changed, but before wage functions adjusted. By the same logic as outlined in Section 3.1 every firm will want to hire a more skilled worker than previously. Additionally, some firms that did not find it profitable to hire anyone previously, will now decide to hire a low skilled worker, because of the increase in surplus levels. Thus, again, the demand for relative skill in manufacturing shifts up, which draws in additional workers from services, so that employment rises in manufacturing and falls in services.[56] Note that this time some of those additional workers could be of relatively low skill, as the increase in surplus levels implies that manufacturing generally became more productive, compared to services.

The difference between the scarce and abundant jobs cases is that in the former only the intensive margin of demand matters (which worker is hired by firm h_M), whereas in the abundant jobs case the extensive margin matters as well (does firm h_M hire any worker). Furthermore, if jobs are abundant in both sectors ($R_i \geq 1$), then in the important special case of Roy-like models the intensive margin does not matter at all for the equilibrium supply of relative skill.In general, demand shifts at the intensive margin increase the relative market power of high-skill workers, allowing them to receive a greater share of surplus. However, in Roy-like models, firms are not sufficiently heterogeneous to have any market power at all, as each firm can always be replaced by an identical, unmatched company. Therefore, workers always receive the entire surplus, regardless of how differentiated they are. More generally, vertical differentiation of workers has an

[56] The increase (fall) in manufacturing (services) employment follows immediately from the increase (fall) in skill supply.

impact on equilibrium sorting only if firms heterogeneity is substantial enough to allow at least some of them a degree of market power.

Proposition 6. If a) jobs are abundant, b) surplus levels increase universally in manufacturing and c) manufacturing workers become more vertically differentiated, then all wages increase in services, as well as wages of the most skilled manufacturing workers.

Less skill is supplied to services and thus, again, wages rise.[57] In manufacturing, the wages of high-skill workers rise, as they are in higher demand. However, the change in wages of least skilled workers is ambiguous: this is because the demand for those workers might not have fallen. Whilst it is true that the firms that employed them previously will now demand better workers, manufacturing firms that previously were not employing anyone, might now want to hire those low-skilled workers.

The impact on wage range is ambiguous in both sectors, as the increase in surplus levels may, in certain cases, be equality-enhancing.

Proposition 7. If a) jobs are abundant, b) surplus levels increase universally in manufacturing and c) manufacturing workers become more vertically differentiated, then both total surplus and profits fall in services, whereas the total surplus produced in manufacturing rises.

The results for profits and total surplus are exactly the same as in the scarce jobs case, as is the intuition for them. In the abundant jobs case, however, it is particularly easy to see why an increase in workers' differentiation and surplus levels is not enough to ensure an increase in manufacturing profits. Suppose that the old surplus function $\pi_M(v_M, h_M, c_1) = v_M h_M$ changes to $\pi_M(v_M, h_M, c_2) = v_M$ and that $R_M > 1$. This increases both the surplus produced in any match and increases the marginal surplus of skill, in both cases because $h_M \leq 1$. Under the new surplus function, however, firms are identical and abundant and, thus, receive no profit (regardless of the supply of skill in manufacturing).

4 Changes in Interdependence

In this section, I study the effects of an increase in the interdependence of relative skills. To make this problem tractable, I restrict attention to the symmetric case. I show that, if jobs are abundant, an increase in skill interdependence results in wage polarization, that is wages decrease more for workers in the interior of the wage distribution than for the lowest and highest earners. Later, I use the Gaussian-Exponential specification

[57] It is worth noting, that although wages spread in the Bickel and Lehmann (1979) sense, this does not necessarily result in a widening of the pay gap between top and bottom workers. The reasons is that critical skill in sector two may rise – as the measure of active sector two firms falls.

from Section 3.1.1 to demonstrate that a change in the proportion in which sectors use fundamental skills will result in a change in relative skill interdependence. For example, if cognitive skills become more important in manufacturing, then – assuming that originally manufacturing was less cognitive skill intensive – relative skill interdependence increases. This matches empirical evidence: skill contents of occupations appear changed between the 1970s and late 1990s (Autor et al., 2003), which was followed by an increase in wage polarization in the 1990s and early 2000s (Acemoglu and Autor, 2011).[58]

Definition 4. The model is *symmetric* iff: (i) $C(v_M, v_S) = C(v_S, v_M)$ for all $(v_M, v_S) \in [0,1]^2$; (ii) $\pi_M(v,h) = \pi_S(v,h)$ for all $(v,h) \in [0,1]^2$ and (iii) $R_M = R_S$.

In the symmetric case, workers simply choose the sector in which their relative skill is higher. This results in identical sectoral relative skill distributions, wages and sector sizes.

Definition 5 (Scarsini, 1984)**.** The copula $C(\bullet, c_2)$ is more concordant than copula $C(\bullet, c_1)$ iff:

$$C(v_M, v_S, c_2) \geq C(v_M, v_S, c_1) \tag{15}$$

for all $(v_M, v_S) \in [0,1]^2$.

The concordance ordering formalizes the idea of greater interdependence, as higher concordance implies that large values of V_M are more likely to go with large values of V_S. For example, for bivariate random variables with a Gaussian copula, an increase in concordance is equivalent to an increase in the correlation parameter ρ (Joe, 1997).[59] Note that an increase in concordance is equivalent to a fall in overall supply of skill. This is because an increase in concordance increases the measure of workers with relative skill lower than (v,v) for any $v \in [0,1]$; accordingly, there are fewer workers whose skill *in at least one sector* is greater than v. And because each worker can work in at most one sector, what truly matters is how many workers are highly skilled in at least one dimension. I will call the concordance ordering strict whenever Equation 15 holds strictly for all $(v_M, v_S) \in (0,1)^2$.

Proposition 8. Suppose the model is symmetric and the distribution of relative skills becomes more concordant. Then, in each sector: (i) the distribution of relative skill deteriorates in first order stochastic dominance sense; (ii) wages increase for any relative skill level; (iii) wage range increases; (iv) profits fall for all firms and (v) the total surplus decreases.

[58]It is worth noting that my results do not imply that *any* change in skill contents will result in greater polarization of wages, only changes that make the two sectors use more *similar* fundamental skills.

[59]I call ρ the correlation parameter, because for variables with Gaussian copula and normal marginal distributions it is equal to linear correlation.

A more concordant relative skill distribution implies a downward shift in relative skill supply in both sectors, which directly translates into less skill being supplied in equilibrium in each sector. This decreases profits and output, but increases wages. The lowest wage, however, remains unchanged and thus wage range goes up.

Proposition 8 specifies changes in wages as a function of relative skill: empirically, however, it is much easier to observe *the distribution of wages.* Fortunately, if jobs are abundant and the model is symmetric, then the distribution of wages can be studied easily. Denote the distribution of wages in the economy as D_W.[60] In the symmetric case, the overall wage distribution is identical to either of the sectoral wage distributions, with:

$$D_W(t) = G_i(w_i^{-1}(t)),$$

which results in an inverse wage distribution function $W : [0, 1] \to \mathbf{R}_{\geq 0}$:

$$W(t) = w_i\big(G_i^{-1}(t)\big). \tag{16}$$

The inverse wage distribution function gives us the wage earned by a worker whose rank in the wage distribution is t. For example, $W(0.5)$ is the median wage. As in Section 3.1 in the discussion on wages are will focus on the case of strict supermodularity ($\frac{\partial^2}{\partial v_i \partial h_i}\pi_i > 0$).

Proposition 9 (Wage Polarization)**.** Suppose that jobs are abundant, the model is symmetric and the distribution of relative skills becomes strictly more concordant. Then the distribution of wages becomes more *polarized*, i.e. i) the lowest wage remains unchanged; ii) highest wage increases strictly and iii) there exists a rank $\bar{t} \in (0, 1)$ such that the inverse distribution function takes strictly lower values for $t \in (0, \bar{t})$.

I interpret this as an increase in wage polarization because it implies that wages for the lowest and highest ranked workers increase more (decrease less) then for some workers of intermediate rank. This is depicted in Figure 3. To understand the intuition for this result note that an increase in the interdependence of relative skills has two effects on the inverse wage function. Firstly, because the distribution of relative skill deteriorates in each sector, $G_i^{-1}(t)$ falls, which means that any rank $t \in (0, 1)$ is now occupied by a less skilled worker. This is the *distribution effect.* It is the distribution effect that provides the force for an increase in wage polarization, by decreasing wages in the middle of the distribution, whilst keeping wage at the very top and bottom of the distribution unchanged.

Secondly, $w_i(v_i)$ increases, which means that the wage paid for any relative skill level $v_i \in (0, 1]$ increases strictly. This is the *wage effect.* The wage effect counteracts the increase in wage polarization somewhat, because it increases the wages earned by workers of higher relative skill by more than those earned by workers of lower skill. On the other

[60] $D_W(w) = \Pr\big(\max\{w_M(V_M), w_S(V_S)\} \geq w \big| \max\{w_M(V_M), w_S(V_S)\} > 0\big)$.

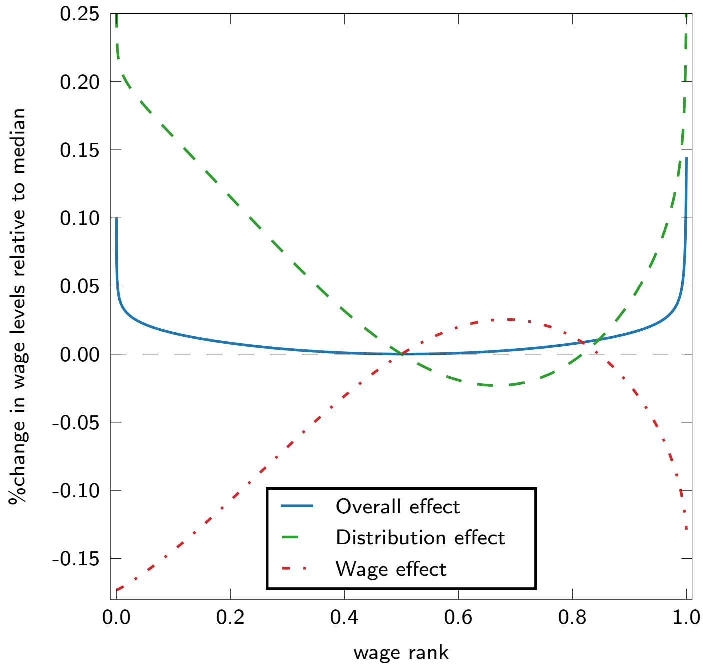

Figure 3: The effect of an increase in concordance on wages. The Gaussian-Exponential specification, increase from $\rho = 0.66$ to $\rho = 0.99$.
Specification: $\sigma_i = 1$, $\mu_i = 0$, $\delta = -\frac{1}{3}$, $\beta_i = 4$, $A_i =$ and $\gamma =$, for $i \in \{1, 2\}$.

hand, it is the wage effect that increases wage inequality, measured as the difference (or ratio) of wages earned by highest and least earning workers: although the fall in the overall supply of skill increases wages for all relative skill levels, the increase is greater for high-skill workers.[61] Overall, the distribution effect dominates for some small enough ranks t, whereas the wage effect dominates for high enough ranks.

It is worth noting that in a standard Roy's model only the distribution effect is present, because the wage paid to a worker with vector (v_M, v_S) is given exogenously. In consequence, an increase in skill concordance would increase wage polarization also in Roy's model, but there would be no increase in wage inequality. In the model of Heckman and Sedlacek (1985) both wage polarization and wage inequality would increase, but the latter only in absolute terms: the ratio of the top to lowest wage would remain unchanged. The relationship between the correlation of skills and wage inequality in the model of Heckman and Sedlacek (1985) has been previously pointed out by Gould (2002), but not its effect on wage polarization.[62] In my model, wage polarization and wage inequality increase both in absolute and relative terms.[63] This is consistent with empirical trends:

[61] The change in levels is the greater the higher the skill; as we can see in Figure 3 this is not quite the case for relative changes, but still the increase for the highest skilled worker will be strictly greater than that for the least skilled worker.

[62] Gould (2002) explains this result in terms of the relative importance of comparative advantage, rather than in terms of a decrease in the supply of skills.

[63] The increase of relative inequality and polarization follows immediately from the fact that the lowest

wages have become polarized in the 90's and 2000's, but nevertheless the strongest growth was observed at the top (Acemoglu and Autor, 2011).

Finally, a quick comment on the asymmetric case. The results in the symmetric case are easy to derive, because the fall in the supply of skills affects both sectors equally: and, thus, it can be expressed as a downward shift in the supply of skills in both sectors (evaluated for the old wage functions). If the model is asymmetric, then the change in skill supply might affect one of the sectors disproportionately strongly (or even exclusively) and, thus, the simple logic of the symmetric case does not go through. However, as an increase in concordance does imply a fall in the overall supply of skill, total surplus produced in the economy falls also in the asymmetric case.

4.1 Changes in Skill Contents

There exists strong empirical evidence that the skill and task content of occupations has changed in recent decades (Autor et al., 2003; Spitz-Oener, 2006), which has been linked to the increase in wage polarization (Acemoglu and Autor, 2011). The standard explanation for these changes is the "routinization hypothesis", so the idea that *routine* jobs are being replaced by machines. In this subsection, I will use the Gaussian-Exponential specification of my model to show that changes to the mix of skills required in each occupation affect the interdependence of relative skills. Hence, changes in interdependence of relative skills are a further channel through which changes in skill contents of occupations might have affected wage polarization.

Recall the Gaussian-Exponential specification from Section 3.1.1. For simplicity, suppose there are just two fundamental skills: cognitive (x_1) and manual (x_2). To isolate the impact of changes in skill requirements on task interdependence I assume that sector i uses the cognitive skill with weight α_i and the manual skill with weight $\sqrt{1-\alpha_i^2}$, and that $\mathbf{x} = (x_C, x_M)$ is $\mathbf{N}(\mathbf{0}, \mathbf{I})$ distributed, where $\alpha_i \in [0, 1]$ and $\mathbf{I}$ is the identity matrix. This ensures that $\sigma_M = \sigma_S = 1$, $\mu_M = \mu_S = 0$ and, hence, implies that the reduced surplus functions do not depend on the vector of skill requirements $\boldsymbol{\alpha}_i$.

A change in α_i represents technological change, but not necessarily a technological improvement.[64] In practice, changes in skill contents are likely caused by actual technological improvements, resulting in an increase in the surplus produced by most, if not all, matches. Here, however, I just want to make a simple point: that a change in skill contents could cause wage polarization *even if* it did not affect surplus functions, through its impact on the overall supply of relative skills.

Because x_1 and x_2 are jointly normally distributed, the interdependence between the

wage does not change.

[64]Intuitively, an increase in α_i can be thought of as a replacement of a worker who was performing a manual task, with a machine that does not require any manual skill to operate, but needs to be occasionally reprogrammed and, thus, requires a degree of cognitive skill.

relative skills v_M, v_S is captured by the parameter ρ of the Gaussian copula. Specifically, an increase in ρ is equivalent to an increase in concordance. Under the assumptions above $\rho = \alpha_S\alpha_M + \sqrt{(1-\alpha_M^2)(1-\alpha_S^2)}$: clearly, relative skill interdependence depends on the skill contents of each sector. Differentiating ρ with respect to α_M reveals that relative skill interdependence will increase in response to a small increase in the weight of cognitive skills in manufacturing if and only if $\alpha_M < \alpha_S$. Thus, if the importance of cognitive skills increases in a sector that is relatively manual skill intensive, the interdependence of relative skills increases. If, on the other hand, cognitive skills become more important in the relatively cognitive skill intensive sector, interdependence falls. Empirical evidence suggests that the substitution towards less routine, and therefore cognitive skill intensive tasks has been stronger in sectors that were previously routine task intensive (Autor et al., 2003), which is consistent with an increase in the interdependence between relative manufacturing and services skills. Furthermore, Gould (2002) provides evidence that skill correlation has been indeed on the rise in the US.

This explanation for increases in wage polarization is different from the mechanisms discussed so far in the literature. The existing explanations are somewhat mechanical, in that they usually assume that the jobs in the middle of the wage distribution are the ones being replaced by machines(Acemoglu and Autor, 2011; Costinot and Vogel, 2010). The explanation presented here does not make this assumption: in fact, as the analysis is restricted to the symmetric case of the model, some of the workers whose skills are now less needed must have been earning very high wages before. Instead, polarization increases, because all industries start demanding a more similar set of skills.

5 Related Literature

This paper builds on the work of Becker (1973), Sattinger (1979) and Roy (1951), combining their approaches towards matching and self-selection, respectively. My model nests Sattinger-like, one sector matching models and Roy-like, two sector comparative advantage models within one framework. The sectors in Roy (1951) can be interpreted as Becker-like matching markets with homogenous and abundant firms, implying that companies have no market power and workers earn the entire surplus. In Roy's model, therefore, sorting depends only on surplus' levels, but not on vertical differentiation. The introduction of firm heterogeneity gives firms some market power and is the reason why the vertical differentiation of workers matters for skill supply. Compared to Becker (1973) and Sattinger (1979), the addition of another sector allows the study of interactions between two matching markets, as well as the determinants of sectoral skill supply.

There are a number of papers that provide comparative statics results for the standard, one-sector differential rents model (e.g. Costrell and Loury, 2004; Gabaix and Landier, 2008; Tervio, 2008).[65] These results capture only the direct effect of exogenous changes, as the within-sector distribution of skill is fixed. As noted in Costrell and Loury (2004), this is a serious limitation – which is addressed directly in this paper. In particular, I show that the two types of shocks which are of particular importance in this literature (i.e. multiplicative surplus shocks and first order stochastic dominance improvements in the distribution of firms' productivity) result in a greater supply of skill in the affected sector, increasing, as they do, both the levels of surplus and worker's vertical differentiation (Sections 3.2). This undoes at least part of their positive direct effect on wage levels and inequality. In fact, with scarce jobs, wages will certainly fall for the least skilled workers in the industry in which the change took place.

My model is a multivariate matching problem. Most of the literature in this area is focused on marriage markets (Anderson, 2003; Chiappori, Oreffice, and Quintana-Domeque, 2011, 2012); however, a recent paper by Lindenlaub (2017) does investigate multivariate matching in labor markets. Lindenlaub defines positive and assortative matching in a general setting with bivariate skills and skill-demands, and provides sufficient conditions for its existence. However, the model is solved and comparative statics are provided only for the very special quadratic-Gaussian case.[66] The comparative statics results focus on technological change, modelled as a multiplicative surplus shock, and only the knife-edge case of no unmatched firms and workers is considered. My paper studies the determinants of skill supply more generally, including changes to the distribution of skills, and all my results hold for general surplus and distribution functions. On the other hand, my model does not allow for jobs that require two types of skills – however, given Lindenlaub's finding that the demand for skills is strongly negatively correlated in the US, this does not seem restrictive.

There exists a small, but quickly growing literature on multi-sector matching. The models in McCann, Shi, Siow, and Wolthoff (2015) and Grossman, Helpman, and Kircher (2013), however, differ substantially from the one presented here, both focusing on one-to-many rather than one-to-one matching. McCann et al. (2015) have a complicated model, with three markets and schooling. This comes at the cost of using specific functional forms and not providing comparative statics results. Meanwhile, Grossman et al. (2013) focus on the impact of trade liberalization, rather than changes in skill and productivity distributions. Their skills are one-dimensional and they restrict attention to cases where re-sorting happens at the extensive margin only. The model in Dupuy (2015) is quite similar to the one presented here (albeit less general), as it is a differential rents match-

[65]See Sattinger (1993) for an overview of the different types of assignment models.

[66]The quadratic surplus function coupled with normal skill distribution implies a surplus function that is not monotonic in workers' skills, so that workers with extremely high and extremely low skills produce the same surplus and earn the same wages. Adding non-interaction skill terms does not resolve this problem in a satisfactory manner, as evidenced by the fact that the surplus function estimated in Lindenlaub (2017) (Table 1) is non-monotonic in manual skill.

ing model with two-dimensional skills.[67] Dupuy (2015) proves the existence (but not uniqueness) of an equilibrium and then proceeds to study the impact of multiplicative shocks on self-selection and inequality. However, unlike this paper, Dupuy (2015) does not show the equilibrium effect of such shocks, providing only a first order result.[68] Mak and Siow (2017) also propose a similar model, with the difference that workers self-select into different sides of the market, rather than into separate sectors. They then calibrate the model to Brazilian data to explain changes in within and across-firm wage inequality, but do not provide any comparative statics results.

My model extends Roy (1951) in a different direction than the strand of 'Roy-like' assignment models (Sattinger, 1975; Teulings, 1995, 2005), in which comparative advantage drives the matching of workers to tasks within a *single* sector and skills are one-dimensional. In this paper, comparative advantage drives between-sectors assignment, whereas within-sector matching is determined by the scale of operation effect. There are, further, a number of papers in the trade literature that build on such 'Roy-like' assignment models (Costinot and Vogel, 2010; Sampson, 2014). Those models feature multiple matching markets – countries – but, as labor is assumed internationally immobile, there is no sorting between markets, which is the main focus here. The comparative statics results in this literature focus mostly on the impact of trade, but Costinot and Vogel (2010) consider also a multiplicative shock in the foreign country and find that this increases inequality in both countries. Hence, although the transmission channel is different, the inequality effect of a multiplicative shock is similar to that in my model.

This paper is also related to the literature on wage polarization. In the one-dimensional models of Costinot and Vogel (2010), Acemoglu and Autor (2011) workers are totally ordered, workers performing routine tasks are ranked in the middle of the wage distribution and, hence, a decrease in the relative demand for routine tasks causes wage polarization. As pointed out by Boehm (2015), the relation between routinization and wage polarization is not immediate if skills are multi-dimensional. This is because of rank switchings: even if workers performing routine tasks were originally ranked in the middle of the wage distribution, their rank might fall as an outcome of routinization. Multi-dimensional models, such as this, allow for rank switching and, thus, seem more appropriate for the study of wage polarization. Compared to Boehm (2015) I propose a different channel through which routinization causes wage polarization. Boehm (2015) allows only for changes in

[67] Generally, the type of model considered in Dupuy differs from mine – as firms are one-dimensional, the size of any sector can never be exogenously fixed. However, as Dupuy further assumes that the masses of workers and firms are equal, his specification is equivalent to the special case of my model for which the measure of firms is equal to 1 in each sector.

[68] The equilibrium adjustment process can be thought of as a following chain of events: the shock changes wages, which impacts sorting, which impacts wages, which impacts sorting etc. until a new equilibrium is reached. The results in Dupuy (2015) consider only the first change in sorting, not the subsequent ones. Proposition 5 in this paper does derive the full equilibrium effect for the type of shock considered by Dupuy.

the prices of tasks performed by different occupations, whilst keeping the task and skill content of each occupation constant.[69] There exists, however, empirical evidence that the task and skill content of occupations has changed in recent decades (Autor et al., 2003; Spitz-Oener, 2006). I show that changes in skill contents of occupations can, on their own, produce wage polarization through their effect on the interdependence of relative skills.[70]

6 Conclusions

This article developed a novel model of workers' self-selection, one that allows for imperfect substitution of skills within sectors. This was accomplished by merging a standard model of self-selection (across sectors) in the vein of Roy (1951) with an assignment model (within sectors) in the vein of Becker (1973) and Sattinger (1979). The within-sector assignment created imperfect substitution of skills but also caused wage functions in each sector to depend on the entire distribution of skill in that sector. Despite this difficulty, I was able to derive a series of sharp monotone comparative statics results without making functional form assumptions.

Firstly, I have shown that if the level of surplus produced by manufacturing matches increases universally and manufacturing workers become more vertically differentiated, then the supply of skills is guaranteed to increase in manufacturing. If jobs are scarce, this increases wage inequality in both sectors. On the other hand, if the improvement in surplus levels decreases vertical differentiation of manufacturing workers, then the supply of skill might fall in that sector. In some cases, this will lead to a contraction of the manufacturing sector.

Secondly, I studied what happens if sectors start requiring different bundles of skills than previously. This changes the interdependence between the sector-specific skill indices (*relative skills*). In particular, if sectors start using more similar skill bundles, then the interdependence of relative skills increases. In consequence, the overall supply of skill declines: there are fewer workers available that have a high relative skill in at least one sector. In the symmetric case, this decreases the sectoral supply of skill in both sectors and increases both wage inequality and wage polarization.

The model developed in this paper can contribute to a variety of areas in economics.

[69] In symmetric Roy's model with perfect substitution of workers (in the vein of Heckman and Sedlacek (1985)), which is the model used by Boehm (2015), a change in task prices that maintains symmetry could never cause wage polarization. Thus, because all my results on polarization are derived for the symmetric case, it follows that these two channels are, in fact, different.

[70] Lindenlaub (2017) studies the effects of changes in skill interdependence on the mismatch between firms and workers, but does not derive the effect this has on wage polarization. In another set of results she studies the impact of an increase in the relative importance of cognitive tasks on polarization, but as her model has no occupations, this change happens across the board. In my model, such a universal change in skill contents has an ambiguous effect on polarization, even in the symmetric case.

For example, in another paper (Burzyński and Gola, 2017), we use an extension of this framework to study migration decisions and the distributional impact of migration on wages. As workers are imperfect substitutes in the model, we are able to shed light on who gains and who loses from migration. Calibrating the model to US and Mexican wage data, we find that although migration increases the average wage in the US, the majority of the native population (62 %) end up earning lower wages.

A Demand: Formal Definition and Shifts

The definition of sectoral demand for relative skill provided in Section 2 holds for a given hiring function and under the assumption that profit is strictly increasing. However, if e.g. surplus does not depend on firm productivity, then a) firms will be indifferent between many different workers and *de facto* there will exist many different hiring functions and b) all firms will make the same profits. Here, I amend the definition of sectoral demand to allow for this possibilities. Accordingly, the economy will be in equilibrium if there exists at least one demand function consistent with firms maximisation problem for which the market clears.

Definition 6. A mapping $v_i^* : [0,1] \to [0,1] \cup \{-1\}$ is a *hiring function* in sector i for wage function w_i, if a) for $v^*(h) \in [0,1]$, $v_i^*(h) \in \arg\max_{v_i} \pi_i(v_i, h) - w_i(v_i)$ and $\pi_i(v_i^*(h), h) - w_i(v_i^*) \geq 0$ and b) for $v_i^*(h) = -1$, $\pi_i(v_i^*, h) - w_i(v_i) \geq 0$ for all $v_i \in [0,1]$.

Given a talent level v_i and an input function v_i^*, define the set $B(v_M, v_i^*) = \{h \in [0,1], v_i^*(h) \geq v_M\}$.

Definition 7. A mapping $D_i : [0,1] \to [0,R]$ is a *sector i demand function for relative skill* for wage function w_i, if there exists a hiring function such that $R_M \int_{B(v_i,v_i^*)} 1 \mathrm{d}v_i = D_i(v_i)$, for all $v_i \in [0,1]$.

For any matching problem, I will denote as $DC(c)$ the set of all possible cumulative demand functions and as $DC(v_M, c)$ the set of their values for talent v_M.

Definition 8. Demand for skill *shifts up* if $\inf DC(v_M, c_2) \geq \sup DC(v_M, c_1)$, for the old equilibrium wage function $w_M(\cdot, c_1)$ and all $v_M \in [0,1]$.

Proposition 10. If both manufacturing workers become more vertically differentiated and surplus levels increase universally in manufacturing, the demand for relative skill shifts up in manufacturing. If jobs are scarce, an increase in workers' vertical differentiation alone suffices for an upward shift of skill demand.

Proof. The partial order $([0,1], \geq)$ is clearly a lattice and the function $\pi_i(v,h) - w_i(v)$ is supermodular in v (for any h). Thus, as an increase in vertical differentiation implies that $\pi_i(v,h) - w_i(v)$ has increasing differences in c it follows from the results in

Topkis (1978) and Milgrom and Shannon (1994) that the set $V^*(c_i) = \{v \in [0,1] : v \in \arg\max \pi_i(v,h,c) - w_i(v)\}$ increases in the strong set order sense with a change from c_1 to c_2. This proves the second statement, as $v_i^*(h) \in [0,1]$ for all firms in that case. As for the first claim, note that the increase in surplus levels means that each firm's profit increases for the old choice of inputs, and hence, by profit maximisation, also for the new choice. Thus, no firms leave the market and the result follows. □

B Stable Matchings and Assignments

Proof of Proposition 1. This proof will refer to the formal definition of a sectoral demand function introduced in Appendix A rather than the simplified definition from Section 2.1.2. First, I propose the following hiring function $v^* : [0,1] \to [v_i^c, 1] \cup \{-1\}$:

$$v^*(h) = \begin{cases} -1 & \text{for } h \in [0, 1 - \frac{S_i(0)}{R_i}) \\ \max\{v \in [0,1] : S(v) = R_i(1-h)\} & \text{otherwise.} \end{cases}$$

Clearly, the set $B(v, v^*) = \{h \in [0,1] : h \geq 1 - \frac{S_i(v)}{R_i}\}$ for this v^*. Thus, the corresponding demand schedule is simply $D_i(v, v^*) = R_i \int_{1-\frac{S_i(v)}{R_i}}^{1} 1\mathrm{d}t = S_i(v)$, as required. In other words, hiring function v^* ensures that demand equals supply.

Secondly, I show that the hiring function is consistent with firms' maximisation problem for the wage functions proposed. Consider any firm with $h \in [1 - \frac{S_i(0)}{R_i}, 1]$; for v^* to be consistent with that firm's profit maximisation it must be the case that:

$$\underset{v\in[0,1]}{\forall} \quad \pi_M(v_M^*(h), h) - \pi_M(v, h) \geq w_M(v_M^*(h)) - w_M(v).$$

This is met as long as:

$$\int_v^{v^*(h)} \int_{1-\frac{S_i(s)}{R_i}}^{h} \frac{\partial^2}{\partial v \partial h} \pi_i(s,t)\mathrm{d}t.$$

Note that $h = 1 - \frac{S_i(v^*(h))}{R_i}$ and that $1 - \frac{S_i(v)}{R_i}$ is an increasing function of v. Thus, if $v^*(h) > v$ then $1 - \frac{S_i(s)}{R_i} \leq h$ for all $s \in [v, v^*(h)]$. Similarly, if $v^*(h) \leq v$ then $1 - \frac{S_i(s)}{R_i} \geq h$ $s \in [v, v^*(h)]$. This and the supermodularity of reduced surplus ensure that the above condition is always met.

Note that the set $[0, 1 - \frac{S_i(0)}{R_i})$ is non-empty only if $S_i(0) < R_i$. If that's the case, for any firm with $h \in [0, 1 - \frac{S_i(0)}{R_i})$ it must be the case that:

$$\underset{v\in[0,1]}{\forall} \quad \pi_M(v, h) - w_M(v) \leq 0,$$

which for $v \in [0, v_i^c)$ gives:

$$\pi_i(v,h) - \pi_i(v, 1-\frac{S_i(0)}{R_i}) \leq w_i(v_i^c) - \pi_i(v_i^c, 1-\frac{S_i(0)}{R_i}) \geq 0.$$

The LHS is greatest for $h \approx 1-\frac{S_i(0)}{R_i}$, in which case the LHS is arbitrarily close to 0. Thus, this condition is ensured to be met if and only if $w_i(v_i^c) = \pi_i(v_i^c, 1-\frac{S_i(0)}{R_i})$, as required. For $v \in [v_i^c, 1]$ we need:

$$\pi_i(v,h) - \pi_i(v, 1-\frac{S_i(0)}{R_i}) + \pi_i(v, 1-\frac{S_i(0)}{R_i}) - w_i(v) \leq 0. \tag{17}$$

Note that $\pi_i(v, 1-\frac{S_i(0)}{R_i}) - w_i(v)$ is the profit firm $1-\frac{S_i(0)}{R_i}$ would make from hiring worker v. This is weakly smaller than $1-\frac{S_i(0)}{R_i}$ makes by hiring $v^*(1-\frac{S_i(0)}{R_i}) = v_i^c$, which is equal to 0, because $w_i(v_i^c) = \pi_i(v_i^c, 1-\frac{S_i(0)}{R_i})$. Thus, the LHS of Equation (17) is negative, as required.

Given a supply function S_i sector i of my model is trivially a special case of the assignment model specified in Chiappori, McCann, and Nesheim (2010). Label the firms as buyers and the workers as sellers. Then this model meets the conditions of the semi-convex buyer setting from Chiappori et al. (2010).[71] Hence, their Proposition 3 holds. This implies that for any stable matching the marginal profits are equal to $\frac{\partial}{\partial h} r_M(h)$ almost everywhere for $h \in [1-\frac{S_i(0)}{R_i}, 1]$, where $r_M(h) = \pi_M(v^*(h), h) - w_M(v^*(h))$. This in turn implies that the wage function for any stable matching is of the proposed form almost everywhere.[72] □

Proof of Lemma 1. I will first show that either $w_M(v_M^c) = \max\{w_S(0), 0\}$ or $v_M^c \in \{0, 1\}$. First, suppose that $w_M(v_M^c) < \max\{w_S(0), 0\}$. This is possible only if $v_M^c = 1$. Otherwise, as w_M is continuous for $v_M \geq v_M^c$, there must exist some $\epsilon > 0$ such that $w_M(v_M) < \max\{w_S(0), 0\}$ for all workers with $v_M \in [v_M^c, v_M^c + \epsilon]$. But then all such workers strictly prefer to either remain unemployed or join services (as w_S is increasing by assumption) and $S_M(v_M^c) = S_M(v_M')$, which contradicts the definition of v_M^c.

Suppose that $w_M(v_M^c) > \max\{w_S(0), 0\}$. Firstly, consider $v_S^c > 0$. In such a case, all (but possibly a positive measure of) workers with $(v_M, v_S) \in [0, v_M^c] \times [0, v_S^c]$ would join manufacturing. As C has full support, if $v_M^c > 0$ then a strictly positive measure of workers lives in this rectangle, which contradicts the definition of v_M^c. Secondly, v_S^c could be equal to 0. By continuity of the reduced surplus function, there exists some $v_M'' < v_M^c$,

[71] Surplus function is twice differentiable, $[0,1]$ and $[v_M^c, 1]$ are smooth manifolds and standard uniform distribution puts zero mass on any $h \in [0,1]$. See Definition 4 in Chiappori et al. (2010) and its discussion.

[72] If for any $v_i \in [0, v_i^c]$ we had $w_i(v_i) < w_i(v_i^c) + pi_i\left(v_i, 1-\frac{S_i(0)}{R_i}\right) - \pi_i\left(v_i^c, 1-\frac{S_i(0)}{R_i}\right)$, then this worker would be demanded by firm $h = 1-\frac{S_i(0)}{R_i}$. If this was the case for a set of points of positive Lebesgue measure, then $D_i(0) \neq S_i(0)$.

such that

$$w_M(v_M^c) + \pi_M\Big(v_M'', 1 - \frac{S_M(0)}{R_M}\Big) - \pi_M\Big(v_M^c, 1 - \frac{S_M(0)}{R_i}\Big) > w_S(0).$$

Further, w_S is continuous for all $v_S \geq v_S^c$: thus, there must exist some $v_S'' > 0$, such that

$$w_S(v_S) < w_M(v_M^c) + \pi_M\Big(v_M'', 1 - \frac{S_M(0)}{R_M}\Big) - \pi_M\Big(v_M^c, 1 - \frac{S_M(0)}{R_i}\Big) \leq w_M(v_M'')$$

for all $v_S \in [0, v_S'']$. Thus, all workers with $(v_M, v_S) \in (v_M'', v_M^c) \times [0, v_S'']$ would join manufacturing, which contradicts the definition of v_M^c. The proof for v_S^c is analogous.

It follows that $v_M^c = v_M' = \sup\{v_M \in [0,1] : w_M(v_M) \leq \max\{w_S(0), 0\}$ or $v_M = 0\}$. First, suppose that $v_M' < v_M^c$ (so $w_M^c > 0$) then by the definition of v_M' follows that $w_M(v_M) > \max\{w_S(0), 0\}$; contradiction. Now suppose that $v_M' > v_M^c$, which is possible only if $v_M^c < 1$. This implies that there exists some $v_M \in (v_M^c, v_M')$ such that $w_M(v_M) \leq \max\{w_S(0), 0\}$. By Proposition1 w_M is strictly increasing for $v_M \geq v_M^c$ and hence $w_M(v_M^c) < \max\{w_S(0), 0\}$ which was shown to be impossible.

Finally, let me prove the last statement. First, I will consider the case of $v_M^c, v_S^c \in (0,1)$. This implies that a) some workers are unemployed (because workers with $(v_M, v_S) < (v_M^c, v_S^c)$ cannot join either sector by definition of critical skills and b) that $w_M(v_M^c) = \max\{w_S(0), 0\}$ and $w_S(v_S^c) = \max\{w_M(0), 0\}$. Suppose $w_S(0) > 0$; then all workers would prefer to join services than remain unemployed, which contradicts point a) above; thus $w_M(v_M^c) = 0$. An analogous reasoning holds for $w_S(v_S^c)$.

Now suppose that $v_M^c = 0$. Clearly, $w_M(0) \geq 0$ as otherwise workers with manufacturing skill close to 0 would prefer to remain unemployed, contradicting $v_M^c = 0$. Thus, if $v_S^c > 0$ then $w_S(v_S^c) = w_M(v_M^c)$, as required. This leaves the possibility that $v_S^c = 0$. Suppose $w_M(0) \neq w_S(0)$. Without loss, suppose that $w_S(0) > w_M(0)$. Then by continuity of w_M follows that there exists some $v_M'' > 0$ such that $w_M(v_M) < w_S(0)$ for all workers with $v_M \in [0, v_M'']$; and, thus, all such workers strictly prefer to join services, contradicting the definition of v_M^c.

□

Proof of Lemma 2. I start with manufacturing. The probability that a worker with relative skill $v_M \geq v_M^c$ chooses services is $\Pr(\psi(V_S) < v_M | V_M = v_M)$. Note that because ψ is weakly increasing, it follows that if $\psi(v_S') < v_M$ then $\psi(v_S'') < v_M$ for any $v_S' \geq v_S'' \geq v_S^C$. Thus:

$$\Pr(\psi(V_S) < v_M | V_M = v_M) = C_{v_M}(v_M, \phi(v_M)) \qquad \text{for } v_M \geq v_M^c,$$

where $\phi(v_M) = \sup\{v_S \in [v_S^c, 1] : \psi(v_S) < v_M\}$. Because $S_M(1) = 0$, this gives us the required expression for $S(v_M)$ if $v_M \geq v_M^c$. And, of course, for any $v_M < v_M^c$, $S_M(v_M) = S_M(0)$ by the definition of critical relative skill.

The proof for $S_S(\cdot)$ is analogous. $\square$

Proof of Theorem 1. Define the *extended separating function* $\psi^e : [v_S^c, 1] \to [v_M^c, 1+B]$ as:

$$\psi^e(v_S) = {v_M}^c + \int_{{v_S}^c}^{v_S} \frac{\frac{\partial}{\partial v_S}\pi_S^e\Big(t, 1 - \frac{\int_t^1 C^e_{v_S}(\psi(r),r)\mathrm{d}r}{R_S}\Big)}{\frac{\partial}{\partial v_M}\pi_M^e\Big(\psi(t), 1 - \frac{\int_t^1 C^e_{v_M}(r,\phi(r))\mathrm{d}r}{R_M}\Big)}\mathrm{d}t, \tag{18}$$

where he extended functions $C^e(\bullet)$, $\pi_M^e(\bullet)$ and $\pi_S^e(\bullet)$ are defined as follows: (1) C^e : $[0, 1+B] \times [0,1] \to [0,1]$

$$C^e(v_M, v_S) = \begin{cases} C(v_M, v_S) & \text{for } (v_M, v_S) \in [0,1] \times [0,1] \\ v_S & \text{for } (v_M, v_S) \in (1, 1+B] \times [0,1], \end{cases}$$

(2): $\pi_M^e(v_M, h) : [0, 1+B] \times [0, \frac{1+R_M}{R_M}] \to \mathbf{R}^+$:

$$\pi_M^e(v_M, h) = \begin{cases} \pi_M(v_M, h) & \text{for } (v_M, h) \in [0,1]^2 \\ \pi_M(1,h) + (v_M - 1)\frac{\partial}{\partial v_M}\pi_M(1,h) & \text{for } (v_M, h) \in (1, B] \times [0,1], \\ \pi_M(v_M, 1) & \text{for } (v_M, h) \in [0,1] \times (1, \frac{1+R_M}{R_M}], \\ \pi_M(1,1) + (v_M - 1)\frac{\partial}{\partial v_M}\pi_M(1,1) & \text{for } (v_M, h) \in (1, B] \times (1, \frac{1+R_M}{R_M}], \end{cases}$$

(3): $\pi_S^e(v_S, h) : [0,1] \times [0, 1 + \frac{1}{R_S}] \to \mathbf{R}^+$:

$$\pi_S^e(v_S, h) = \begin{cases} \pi_S(v_S, h) & \text{for } (v_M, h) \in [0,1]^2 \\ \pi_S(v_S, 1) & \text{for } (v_M, h) \in [0,1] \times (1, 1 + \frac{1}{R_S}], \end{cases}$$

and $B = \frac{\max \frac{\partial}{\partial v_S}\pi_S}{\min \frac{\partial}{\partial v_M}\pi_M}$. Note that $C^e(\cdot, v_S)$, $C^e_{v_S}(\cdot, v_S)$, $\frac{\partial}{\partial v_M}\pi_M^e(\cdot,\cdot)$ and $\frac{\partial}{\partial v_S}\pi_S^e(v_S, \cdot)$ are Lipschitz continuous[73]; denote their Lipschitz-constants as L^1, L^2 ,L^3, L^4 and L^5 respectively.

Clearly, given v_M^c and v_S^c the separating function ψ uniquely determines the extended

[73] I will do this in detail for $C^e_{v_S}(v_M, v_S)$ – the reasoning for the other two is analogous. $C^e_{v_S}(v_M, v_S)$: $[0, 1+B] \times [0,1] \to [0,1]$:

$$C^e_{v_S}(v_M, v_S) = \begin{cases} C_{v_S}(v_M, v_S) & \text{for } (v_M, v_S) \in [0,1] \times [0,1] \\ 1 & \text{for } (v_M, v_S) \in (1, 1+B] \times [0,1], \end{cases}$$

is clearly continuous in u. It is equally easy to see that the function $C^e_{v_S}(\cdot, v_S)$ is differentiable almost everywhere and its derivative is Lebesque integrable. It is also the case that for any $(v_M, v_S) \in (1, 1+B] \times [0,1]$ we have:

$$C^e_{v_S}(a, v_S) + \int_a^1 C^e_{uv}(r, v_S)\mathrm{d}r + \int_1^{v_M} 0\mathrm{d}r = 1,$$

which means that $C^e_{v_S}(\cdot, v_S)$ is absolutely continuous. Moreover, as $C^e(\bullet)$ is twice continuously differentiable and any continuous function defined on a compact set is bounded it follows that $C^e_{v_S}(\cdot, v_S)$ is essentially bounded; and a differentiable almost everywhere, absolutely continuous function with an essentially bounded derivative is Lipschitz-continuous.

separation function ψ^e. Similarly, it should be clear that

$$\psi(v_S) = \begin{cases} \psi^e(v_S) & \text{if } \psi^e(v_S) \leq 1, \\ 1 & \text{otherwise.} \end{cases}$$

The result for $\psi^e(v_S) \leq 1$ follows from noting that ψ^e is strictly increasing and then substituting Equation (9) into Equation (14), differentiating wrt v_S, dividing both sides by $\frac{\partial}{\partial v_M}\pi_M\Big(\psi(v_S), \frac{1}{R_M}\int_{v_M{}^c}^{\psi(v_S)} C_{v_M}(r, \psi^{-1}(r))\mathrm{d}t\Big)$ and then integrating from $v_S{}^c$ to v_S (and remembering that $\psi(v_S{}^c) = v_M{}^c$).[74] The other part follows from the fact that for v_S's such that $w_S(v_S) \leq w_M(1)$ we have $\psi(v_S) = 1$ and $\psi^e(v_S) > 1$ (because ψ^e is strictly increasing).

Thus, it is suffient to prove that ψ^e, v^c_M, v^c_S exist and are unique. Let me make a few observations that will prove useful.

Relation Between Supply Functions By differentiating $C(\psi(r), r)$ rearranging and integrating from $v_S{}^c$ to v_S, we arrive at

$$S_M(0) - S_M(\psi(v_S)) + S_S(0) - S_S(v_S) = C(\psi(v_S), v_S) - C(v_M{}^c, v_S{}^c). \tag{19}$$

Determining the Critical Skills As the critical skills v^c_M, v^c_S are also unknown, we need to find conditions that will pin them down. Let me start by denoting the measure of employed workers as $M = S_M(0) + S_S(0)$. Clearly, $M = \min\{R_M + R_S, 1\}$ in equilibrium: otherwise we have $S_i(0) < R_i$ in some sector i, implying that a positive measure of workers with relative skill below (v^c_M, v^c_S) would strictly prefer to join sector i than remain unemployed. By Equation (19) this gives $1 - M = C^e(v^c_M, v^c_S)$, determining one of the critical skills as a function of the other. Furthermore, note that Assumption 3 implies that $v^c_M, v^c_S < 1$ and thus $S_M(0), S_S(0) > 0$.[75] Therefore, from Proposition 1 and Lemma 1 it follows that if $S_M(0) < R_M$ then:

$$\pi_M(v^c_M, 1 - \frac{S_M(0)}{R_M}) = w_M(v^c_M) = w_S(v^c_S) \leq \pi_S(v^c_S, 1 - \frac{M - S_M(0)}{R_S}),$$

and alogously for services. This determines the other critical skill if $R_M + R_S > 1$. Finally, recall that market clearing implies that $S_i(0) \leq R_i$, implying that if $R_M + R_S \leq 1$ we have $S_M(0) = R_M$ and $S_S(0) = R_S$.

[74] This gives us Equation (18), but with ψ rather than ψ^e on the right hand side.

[75] If $R_i < 1$ this follows immediately from $1 - M = C^e(v^c_M, v^c_S)$. Otherwise, suppose that $v^c_M = 1$; then $S_M(0) = 0 < R_M$ and $w_M(1) = \pi_M(1, 1) > \pi_S(0, 1 - \frac{1}{R_S}) \geq w_S(0)$. But then, by continuity of π_M and Proposition 1 follows that there must exist some $\epsilon > 0$ such that all workers with $(v_M, v_S) \in [0, \epsilon] \times [1-\epsilon, 1]$ would prefer to join manufaturing, contradicting $v^c_M = 1$.

The Set of Equations and Inequalities By substituting $S_i(v_i) = S_i(0) - S_i(v_i)$ and Equation (19) into Equation (18) we arrive at:

$$\psi^e(v_S) = {v_M}^c + \int_{{v_S}^c}^{v_S} \frac{\frac{\partial}{\partial v_S}\pi_S^e\left(t, \frac{R_S - S_S(0) + \int_{{v_S}^c}^t C^e_{v_S}(\psi(r),r)\mathrm{d}r}{R_S}\right)}{\frac{\partial}{\partial v_M}\pi_M^e\left(\psi^e(t), \frac{R_M - 1 + S_S(0) + C^e(\psi^e(t),t) - \int_{{v_S}^c}^t C^e_{v_S}(\psi^e(r),r)\mathrm{d}r}{R_M}\right)}\mathrm{d}t. \tag{20}$$

This, together with:

$$M = \min\{R_M + R_S, 1\} \tag{21}$$

$$1 - M = C^e({v_M}^c, {v_S}^c), \tag{22}$$

$$S_S(0) = \int_{{v_S}^c}^1 C^e_{v_S}(\psi(r), r)\mathrm{d}r, \tag{23}$$

$$S_S(0) \in \Theta(M) = [\max\{0, M - R_M\}, \min\{1, R_S\}] \tag{24}$$

$$S_M(0) < R_M \Rightarrow \pi^e_M(v^c_M, 1 - \frac{S_M(0)}{R_M}) \leq \pi^e_S(v^c_S, 1 - \frac{M - S_M(0)}{R_S}), \tag{25}$$

$$S_S(0) < R_S \Rightarrow \pi^e_S(v^c_S, 1 - \frac{M - S_M(0)}{R_S}) \leq \pi^e_M(v^c_M, 1 - \frac{S_M(0)}{R_M}). \tag{26}$$

constitutes the set of Equations and Inequalities that determines ψ^e, v^c_M, v^c_S.

The remainder of the proof shows that there exists a unique solution to Equations (20)-(26). Define the set:

$$K = \{d \in C[0,1] : |d(v_S) - 1| \leq 1 + B\},$$

where $C[0,1]$ is the set of all continuous functions that map from $[0,1]$. The constant function $d(v_S) = 1$ lies in K and hence the set is non-empty. Define a norm, $||\cdot||_\lambda$ on $C[0,1]$:

$$||h||_\lambda = \sup_{[0,1]} e^{-\lambda v_S}|h(v_S)|,$$

where λ is some weakly positive number. K is a complete metric space for this norm.[76]

Endow the sets $[0,1]^2$ and $\Theta(M)$ with the Euclidean norm and define a mapping $T : K \times [0,1]^2 \times \Theta(M) \to K$:

$$(Td)(v_S, {v_S}^c, {v_M}^c, S_S(0)) = \begin{cases} {v_M}^c & \text{for } v_S < {v_S}^c \\ {v_M}^c + \int_{{v_S}^c}^{v_S} \frac{\frac{\partial}{\partial v_S}\pi^e_S(t, \frac{R_S - S_S(0) + \int_{{v_S}^c}^t C^e_{v_S}(d(r),r)}{R_S}\mathrm{d}r)}{\frac{\partial}{\partial v_M}\pi^e_M(d(t), \frac{R_M - 1 + S_S(0) + C^e(d(t),t) - \int_{{v_S}^c}^t C^e_{v_S}(d(r),r)\mathrm{d}r}{R_M}\mathrm{d}r)}\mathrm{d}t & \text{for } v_S \geq {v_S}^c. \end{cases}$$

[76] If we endowed K with the sup-norm, then K would be a closed subspace of $C[0,1]$; since $C[0,1]$ is complete in the sup-norm, so is K. And it was shown by Bielecki (1956) that the $||\cdot||_\lambda$ norm is equivalent to the sup-norm for any $C[a,b]$ – and thus if K is a complete metric space for the sup-norm it is also a complete metric space for $||\cdot||_\lambda$.

Note that this map is well-defined, as for any ${v_S}^c \in [0,1]$ and $d \in K$:

$$\frac{R_S - S_S(0) + \int_{{v_S}^c}^t C^e_{v_S}(d(r),r)}{R_S}\mathrm{d}r \leq 1 + \int_{{v_S}^c}^t \frac{1}{R_S}\mathrm{d}r \leq \frac{1}{R_S} + 1$$
$$\frac{R_M - 1 + S_S(0) + C(d(t),t) - \int_{{v_S}^c}^t C^e_{v_S}(d(r),r)\mathrm{d}r}{R_M} \leq \frac{R_M + C(d(t),t)}{R_M} \leq \frac{1}{R_M} + 1;$$

and that it is continuous in v, ${v_S}^c$, ${v_M}^c$ and $S_S(0)$. It is also the case that for $v_S \geq {v_S}^c$:

$$||[(Td)(v_S, {v_S}^c, {v_M}^c, S_S(0)) - 1| \leq \int_{{v_S}^c}^{v_S} B\mathrm{d}t + |{v_M}^c - 1| \leq 1 + B,$$

and for $v_S < {v_S}^c$:

$$||[(Td)(v_S, {v_S}^c, {v_M}^c, S_S(0)) - 1| \leq |{v_M}^c - 1| \leq 1 + B,$$

so indeed $T(K) \subset K$. Finally, it should be clear that for any $({v_S}^c, {v_M}^c, S_S(0))$ the restriction of any fixed point of $(Td)(\bullet)$ to $[{v_S}^c, 1]$ gives us the solution to (20) and that any solution to (20) can be easily extended into a fixed point of $(Td)(\bullet)$. Therefore, it suffices to show that there exists such a λ that for any $({v_S}^c, {v_M}^c, S_S(0)) \in [0,1]^2 \times \Theta(M)$, $Td(\bullet)$ is a contraction wrt to the norm $||\cdot||_\lambda$ to show that (20) has a unique solution for any feasible $({v_M}^c, {v_S}^c, S_S(0))$.

Let us drop $({v_S}^c, {v_M}^c, S_S(0))$ from the definition of the map (remembering that we are keeping them constant) and enhance our notation by new maps: $S_S : [{v_S}^c, 1] \times K \to [0,1]$, $P_S : [{v_S}^c, 1] \times K \to [0, 1 + \frac{1}{R_S}]$ and $P_M : [0, B] \times K \to [0, 1 + \frac{1}{R_M}]$:

$$\begin{aligned}
(S_S d)(v_S) &= S_S(0) - \int_{{v_S}^c}^{v_S} C^e_{v_S}(d(r),r)\mathrm{d}r,\\
(P_S d)(v_S) &= \frac{R_S - (S_S d)(v_S)}{R_S},\\
(P_M d)(d(v_S)) &= \frac{R_M - 1 + C^e(d(v_S), v_S) + (S_S d)(v_S)}{R_M}.
\end{aligned}$$

Take any any $t \geq {v_S}^c$ and any $d_1, d_2 \in S$ and for any map $(fd)(t)$ denote $(fd_1)(t) - (fd_2)(t)$ as $\Delta_d(fd)(t)$. Then we have:

$$\begin{aligned}
|\Delta_d(S_S(0)d)(t)| &= |\int_{v^c}^t C^e_v(d_1(r),r) - C^e_v(d_2(r),r)\mathrm{d}r| \qquad (27)\\
&\leq \int_{v^c}^t |C^e_v(d_1(r),r) - C^e_v(d_2(r),r)|\mathrm{d}r \leq \int_{v^c}^t L_2|d_1(r) - d_2(r)|\mathrm{d}r|\\
&= L_2 \int_{v^c}^t e^{\lambda r}e^{-\lambda r}|d_1(r) - d_2(r)|\mathrm{d}r \leq L_2||d_1 - d_2||_\lambda \int_{v^c}^t e^{\lambda r}\mathrm{d}r
\end{aligned}$$

$$= \frac{L_2}{\lambda}||d_1 - d_2||_\lambda(e^{\lambda t} - e^{\lambda v^c}) \leq \frac{L_2}{\lambda}||d_1 - d_2||_\lambda e^{\lambda t},$$

which can be used to establish:

$$|\Delta_d(P_S d)(t)| \leq \frac{L_2}{\lambda R_S}||d_1 - d_2||_\lambda e^{\lambda t} \tag{28}$$

$$|(P_M d_1)(d_1(t)) - (P_M d_2)(d_2(t))| = |\frac{C^e(d_1(v), v) - C^e(d_2(v), v) - \Delta_d(S_S(0)d)(v)}{R_M}| \tag{29}$$

$$\leq \frac{1}{R_M}(|C^e(d_1(v), v) - C^e(d_2(v), v)| + |\Delta_d(S_S(0)d)(v)|)$$

$$\leq \frac{L_2}{\lambda_M}||d_1 - d_2||_\lambda e^{\lambda t} + \frac{L^1}{R_M}|d_1(t) - d_2(t)|.$$

Denote $\sup \frac{\partial}{\partial v_S}\pi_S(v_S, h) = L_6$, $\inf \frac{\partial}{\partial v_M}\pi_M(v_M, h) = L_7$ and note that continuity of $\frac{\partial}{\partial v_M}\pi_M$ and $\frac{\partial}{\partial v_S}\pi_S$ and the fact that $\frac{\partial}{\partial v_M}\pi_M > 0$ imply that both L^6 and L^7 are finite. Using all this, we can write, for any $v_S \geq {v_S}^c$ and any $d_1, d_2 \in S$:

$$|\Delta_d(Td)(v)| = |\int_{v^c}^{v} \frac{\frac{\partial}{\partial v_S}\pi_S^e(t, (P_S d_1)(t))}{\frac{\partial}{\partial v_M}\pi_M^e(d_1(r), (P_M d_1)(d_1(t)))} - \frac{\frac{\partial}{\partial v_S}\pi_S^e(t, (P_S d_2)(t))}{\frac{\partial}{\partial v_M}\pi_M^e(d_2(r), (P_M d_2)(d_2(t)))}\mathrm{d}t|$$

$$\leq \int_{v^c}^{v} |\frac{\frac{\partial}{\partial v_S}\pi_S^e(t, (P_S d_1)(t))}{\frac{\partial}{\partial v_M}\pi_M^e(d_1(r), (P_M d_1)(d_1(t)))} - \frac{\frac{\partial}{\partial v_S}\pi_S^e(t, (P_S d_2)(v,))}{\frac{\partial}{\partial v_M}\pi_M^e(d_1(r), (P_M d_1)(d_1(t)))}$$

$$+ \frac{\frac{\partial}{\partial v_S}\pi_S^e(t, (P_S d_2)(t))}{\frac{\partial}{\partial v_M}\pi_M^e(d_1(r), (P_M d_1)(d_1(t)))} - \frac{\frac{\partial}{\partial v_S}\pi_S^e(t, (P_S d_2)(t))}{\frac{\partial}{\partial v_M}\pi_M^e(d_2(r), (P_M d_2)(d_2(t)))}|\mathrm{d}t$$

$$\leq \int_{v^c}^{v} \frac{|\frac{\partial}{\partial v_S}\pi_S^e(t, (P_S d_1)(t)) - \frac{\partial}{\partial v_S}\pi_S^e(t, (P_S d_2)(t))|}{L_7}$$

$$+ L_6(|\frac{\frac{\partial}{\partial v_M}\pi_M^e(d_1(r), (P_M d_1)(d_1(t))) - \frac{\partial}{\partial v_M}\pi_M^e(d_2(r), (P_M d_2)(d_2(t)))}{\frac{\partial}{\partial v_M}\pi_M^e(d_1(r), (P_M d_1)(d_2(t)))\frac{\partial}{\partial v_M}\pi_M^e(d_2(r), (P_M d_2)(d_2(t)))}|\mathrm{d}t$$

$$\leq \int_{v^c}^{v} \frac{L_5}{L_7}|\Delta_d(P_S d)(t)|$$

$$+ \frac{L_6}{L_7^2}[|\frac{\partial}{\partial v_M}\pi_M^e(d_1(r), (P_M d_1)(d_1(t))) - \frac{\partial}{\partial v_M}\pi_M^e(d_2(t), (P_M d_1)(d_1(t))|]$$

$$+ \frac{L_6}{L_7^2}[|\frac{\partial}{\partial v_M}\pi_M^e(d_2(t), (P_M d_1)(d_1(t)) - \frac{\partial}{\partial v_M}\pi_M^e(d_2(r), (P_M d_2)(d_2(t)))|]\mathrm{d}t$$

$$\leq \int_{v^c}^{v} \frac{L_5 L_2}{\lambda L_7 R_S}||d_1 - d_2||_\lambda e^{\lambda(t - v^c)} + \frac{L_3 L_6}{L_7^2}|d_1(t) - d_2(t)|$$

$$+ \frac{L_4 L_6}{L_7^2}|(P_M d_1)(d_1(t)) - P_M d_2)(d_2(t))|\mathrm{d}t$$

$$\leq \frac{L_5 L_2}{\lambda^2 L_7 R_S}||d_1 - d_2||_\lambda e^{\lambda v} + \frac{L_3 L_6}{\lambda L_7^2}||d_1 - d_2||_\lambda e^{\lambda v}$$

$$+ \int_{v^c}^{v} \frac{L_4 L_6}{L_7^2}(\frac{L_2}{\lambda_M}||d_1 - d_2||_\lambda e^{\lambda(t - v^c)} + \frac{L_1}{R_M}|d_1(t) - d_2(t)|)\mathrm{d}t$$

$$\leq \frac{1}{\lambda}||d_1 - d_2||_\lambda e^{\lambda v}\Big[\frac{L_5 L_2}{\lambda L_7 R_S} + \frac{L_3 L_6}{L_7^2} + \frac{L_4 L_6}{L_7^2}\Big(\frac{L_2}{\lambda_M} + \frac{L_1}{R_M}\Big)\Big]$$

Now, for $v_S < {v_S}^c$ this has to hold as well, as then $|(Td_1)(v_S) - T(d_2)(v_S)| = 0$; therefore, for any $v_S \in [0,1]$ we have that:

$$\Delta_d(Td)(v_S)| \leq \frac{1}{\lambda}||d_1 - d_2||_\lambda e^{\lambda v_S}\Big[\frac{L_5 L_2}{\lambda L_7 R_S} + \frac{L_3 L_6}{L_7^2} + \frac{L_4 L_6}{L_7^2}\Big(\frac{L_2}{\lambda_M} + \frac{L_1}{R_M}\Big)\Big].$$

Dividing both sides of that by $e^{\lambda v_S}$ and then taking sup on both sides we get:

$$||(Td_1)(t) - T(d_2)(t)||_\lambda \leq \frac{1}{\lambda}||d_1 - d_2||_\lambda \Big[\frac{L_5 L_2}{\lambda L_7 R_S} + \frac{L_3 L_6}{L_7^2} + \frac{L_4 L_6}{L_7^2}\Big(\frac{L_2}{\lambda_M} + \frac{L_1}{R_M}\Big)\Big]. \quad (30)$$

Therefore, there has to exist a high enough λ for which our map $(Td)(v_S)$ is a contraction in the metric space $(S, ||\cdot||_\lambda)$ – which, by Banach's Fixed-Point Theorem means that $(Td)(v_S)$ has a unique fixed point, which in turn means that Equation (20) has a single solution for any given $({v_S}^c, {v_M}^c, S_S(0)) \in [0,1]^2 \times \Theta(M)$. Note that Equation (30) does not depend on $({v_S}^c, {v_M}^c, S_S(0))$ – and thus, by standard results (see e.g. Hasselblatt and Katok, 2003, p. 68) it follows that as $(Td)(v_S, {v_S}^c, {v_M}^c, S_S(0))$ is continuous in ${v_S}^c$, ${v_M}^c$ and $S_S(0)$ the fixed point – and thus the solution of (20) – is continuous in them as well.

Denote the fixed point of $(Td)(\cdot, {v_S}^c, {v_M}^c, S_S(0))$ as $d^*(\cdot, {v_S}^c, {v_M}^c, S_S(0))$ – then the following result holds:

Lemma 5. The function $d^*(\cdot, {v_S}^c, {v_M}^c, S_S(0))$ is weakly decreasing in ${v_S}^c$ and $S_S(0)$ and weakly increasing in ${v_M}^c$ for all v_S's. Moreover, for some v_S's, $d^*(\cdot, {v_S}^c, {v_M}^c, S_S(0))$ is strictly decreasing in ${v_S}^c$ and $S_S(0)$ (strictly increasing in ${v_M}^c$).

Proof. I start with the claims regarding $d(v_S, \cdot, {v_M}^c, S_S(0))$ and suppress ${v_M}^c$ and $S_S(0)$ from notation for that part of the proof. Take any ${v_S}_2^c > {v_S}_1^c \in [0,1]$, denote $d^*(v_S, {v_S}_2^c) - d^*(v_S, {v_S}_1^c)$ as $\Delta_{{v_S}^c} d^*(v_S, {v_S}^c)$ and define:

$$S_S(v_S, {v_S}^c) = S_S(0) - \int_{{v_S}^c}^{v_S} C_{v_S}(d^*(r, {v_S}^c), r)\mathrm{d}r,$$
$$P_S(v_S, {v_S}^c) = \frac{R_S - S_S(v_S, {v_S}^c)}{R_S},$$
$$P_M(d^*(v_S, {v_S}^c), {v_S}^c) = \frac{R_M - 1 + C(d^*(v_S, {v_S}^c), r) + S_S(v_S, {v_S}^c)}{R_M}.$$

Then for any $v_S \geq {v_S}_2^c$ we have:

$$\Delta_{v^c} d^*(v, v^c) = \Delta_{v^c} d^*(v_2^c, v^c) + \int_{v_2^c}^{v} \frac{\frac{\partial}{\partial v_S}\pi_S^e(t, P_S(v_2^c, t))}{\frac{\partial}{\partial v_M}\pi_M^e(d^*(t, v_2^c), P_M(v_2^c, d^*(t, v_2^c)))} - \frac{\frac{\partial}{\partial v_S}\pi_S^e(t, P_S(v_1^c, t))}{\frac{\partial}{\partial v_M}\pi_M^e(d^*(t, v_1^c), P_M(v_1^c, d^*(t, v_1^c)))}\mathrm{d}t.$$

It is trivial that for any $v_S \in [{v_S}_1^c, {v_S}_2^c)$ we have $\Delta_{{v_S}^c} d^*(v_S, {v_S}^c) < 0$, which proves the second (strict) part of this claim. Thus, we only need to show now that $\Delta_{{v_S}^c} d^*(v_S, {v_S}^c) \leq$

0 for all $v_S \in [v_{S2}^c, 1]$. Suppose not. Then the set $\Omega^{gen} = \{v_S \in [v_{S2}^c, 1] : \Delta_{v_S{}^c} d^*(v_S, v_S{}^c) > 0\}$ has to be non-empty. Then we have that for $v_S{}^g = \inf \Omega^{gen}$, $\Delta_{v_S{}^c} d^*(v_S{}^g, v_S{}^c) = 0$ and $\Delta_{v_S{}^c} d^*_{v_S}(v_S{}^g, v_S{}^c) > 0$. The sign of $\Delta_{v_S{}^c} d^*_{v_S}(v_S{}^g, v_S{}^c)$ depends only on the signs of[77]:

$$\frac{\partial}{\partial v_S}\pi_S^e(v_S{}^g, P_S(v_{S2}^c, v_S{}^g)) - \frac{\partial}{\partial v_S}\pi_S^e(v_S{}^g, P_S(v_{S1}^c, v_S{}^g))$$

and

$$\frac{\partial}{\partial v_M}\pi_M^e(d^*(v_S{}^g, v_{S1}^c), P_M(v_{S1}^c, d^*(v_S{}^g, v_{S1}^c))) - \frac{\partial}{\partial v_M}\pi_M^e(d^*(v_S{}^g, v_{S2}^c), P_M(v_{S2}^c, d^*(v_S{}^g, v_{S2}^c))).$$

However, as $\Delta_{v_S{}^c} d^*(v_S{}^g, v_S{}^c) = 0$ and both surplus functions are weakly supermodular, these in turn depend only on the sign of $S_S(0)(v_{S2}^c, v_S{}^g) - S_S(0)(v_{S1}^c, v_S{}^g)$. As for any $v_S \leq v_S{}^g$ it was the case that $\Delta_{v_S{}^c} d^*(v_S{}^g, v_S{}^c) \leq 0$ and $v_{S2}^c \geq v_{S1}^c$, it follows that: $S_S(0)(v_{S2}^c, v_S{}^g) - S_S(0)(v_{S1}^c, v_S{}^g) \leq 0$ and thus:

$$\Delta_{v_S{}^c} d^*_{v_S}(v_S, v_S{}^c) \leq 0,$$

which means that Ω^{gen} has to be empty and proves our first claim.

The proof for $S_S(0)$ is analogous[78]. For $v_M{}^c$, note that for a change in $v_M{}^c$, $\Delta_{v_M{}^c} d^*(v_S{}^c, v_M{}^c)$ is positive. The subsequent reasoning is analogous, but with opposite signs (the strict decreasingness follows from $\Delta_{v_M{}^c} d^*(v_S{}^c, v_M{}^c) < 0$ and continuity). □

Everything I derived so far applies both for cases with abundant and scarce jobs. From now on, however, I will consider those cases separately.

Scarce jobs If $R_M + R_S < 1$, then $M = R_M + R_S$, which reduces (26) to $S_S(0) = R_S$ and gives $C(v_M{}^c, v_S{}^c) = 1 - R_M - R_S > 0$. For $(v_M, v_S) > 0$, $C(\bullet)$ is strictly increasing in both parameters, which allows us to define $v_M{}^c$ as a strictly decreasing, continuous function of $v_S{}^c$. Define $\underline{v_S}$ as $v_M{}^c(\underline{v_S}) = 1$ and note that, as $v_M{}^c \in [0, 1]$, Equation (22) shrinks the range of feasible $v_S{}^c$'s to $[\underline{v_S}, 1]$. Hence, $d^*(v_S, v_S{}^c, v_M{}^c, S_S(0))$ depends only on v_S and $v_S{}^c$ and is decreasing and continuous in $v_S{}^c$ – I will denote it as $d^*(v_S, v_S{}^c)$

[77] To see this, note that:

$$\Delta_{v^c} d^*_v(v^g, v^c) = \frac{\frac{\partial}{\partial v_S}\pi_S^e(v^g, P_S(v_2^c, v^g))}{\frac{\partial}{\partial v_M}\pi_M^e(d^*(v^g, v_2^c), P_M(v_2^c, d^*(v^g, v_2^c)))} - \frac{\frac{\partial}{\partial v_S}\pi_S^e(v^g, P_S(v_1^c, v^g))}{\frac{\partial}{\partial v_M}\pi_M^e(d^*(v^g, v_1^c), P_M(v_1^c, d^*(v^g, v_1^c)))}$$

$$= \frac{\frac{\partial}{\partial v_S}\pi_S^e(v^g, P_S(v_2^c, v^g)) - \frac{\partial}{\partial v_S}\pi_S^e(v^g, P_S(v_1^c, v^g))}{\frac{\partial}{\partial v_M}\pi_M^e(d^*(v^g, v_2^c), P_M(v_2^c, d^*(v^g, v_2^c)))}$$

$$+ \frac{\partial}{\partial v_S}\pi_S^e(v^g, P_S(v_1^c, v^g)) \frac{\frac{\partial}{\partial v_M}\pi_M^e(d^*(v^g, v_1^c), P_M(v_1^c, d^*(v^g, v_1^c))) - \frac{\partial}{\partial v_M}\pi_M^e(d^*(v^g, v_2^c), P_M(v_2^c, d^*(v^g, v_2^c)))}{\frac{\partial}{\partial v_M}\pi_M^e(d^*(v^g, v_1^c), P_M(v_1^c, d^*(v^g, v_1^c)))\frac{\partial}{\partial v_M}\pi_M^e(d^*(v^g, v_2^c), P_M(v_2^c, d^*(v^g, v_2^c)))}.$$

[78] For $v_S = v_S{}^c$ we have $\Delta_M d^*(v_S, S_S(0)) = 0$ and $\Delta_M d^*_{v_S}(v_S, S_S(0)) < 0$. The sign of $\Delta_M d^*_{v_S}(v_S{}^g, S_S(0))$ depends on $S_S(0)_1 - S_S(0)_2 < 0$ and the difference in $S_S(0)(v_S, S_S(0))$, which is weakly negative for the same reasons as above. Thus, $\Delta_M d^*_{v_S}(v_S{}^e, S_S(0)) \leq 0$, which implies that $d^*(v_S, a, \cdot)$ will never strictly increase.

from now on. Thus, the modified system of equations reduces to:

$$R_S = \int_{v_S{}^c}^{1} C^e_{v_S}(d^*(r, v_S{}^c), r)\mathrm{d}r.$$

The RHS is continuous in $v_S{}^c$, as $d^*(v_S, v_S{}^c))$ is continuous in $v_S{}^c$. For $v_S{}^c = \underline{v_S}$, we have $d^*(v_S, v_S{}^c) \geq 1$ regardless of v_S and therefore $\int_0^1 C^e_{v_S}(d^*(r, v_S{}^c), r)\mathrm{d}r = 1$, whereas for $v_S{}^c = 1$, $\int_1^1 C^e_{v_S}(d^*(r, v_S{}^c), r)\mathrm{d}r = 0$; thus, a solution to (23) (given $R_S \in (0, 1)$) exists. It is unique, as $d^*(v_S, \cdot)$ is weakly decreasing for all and strictly decreasing for some v_S and thus the RHS crosses R_S only once from above.

Abundant jobs If $_M + R_S \geq 1$, then $M = 1$ and thus $C(v_M{}^c, v_S{}^c) = 0$. Hence, $\min\{v_M{}^c, v_S{}^c\} = 0$ and I cannot define $v_M{}^c$ as a function of $v_S{}^c$, as there is a continuum of $v_S{}^c$'s for which $C(0, v_S{}^c) = 0$. I address this by defining the set $\Gamma^c = \{(v_M{}^c, v_S{}^c) : \min\{v_M{}^c, v_S{}^c\} = 0\}$, a new variable $a \in [-1, 1]$ and writing $v_M{}^c$ and $v_S{}^c$ as:

$$v_M{}^c(a) = \begin{cases} -a & \text{for } a \leq 0, \\ 0 & \text{for } a > 0, \end{cases} \qquad v_S{}^c(a) = \begin{cases} 0 & \text{for } a \leq 0, \\ a & \text{for } a > 0. \end{cases}$$

For any a, $(v_M{}^c(a), v_S{}^c(a)) \in \Gamma^c$ and for any $(v_M{}^c, v_S{}^c) \in \Gamma^c$ there exists a unique a, such that $(v_M{}^c(a), v_S{}^c(a)) = (v_M{}^c, v_S{}^c)$. Thus, if there exists a unique a that solves Equation (23), there also exists a unique $(v_M{}^c, v_S{}^c)$ that solves it. Moreover, $v_S{}^c(a)$ is continuous and increasing, and $v_M{}^c(a)$ is continuous and decreasing. Therefore the function $d^*(v_S, a, S_S(0)) = d^*(v_S, v_S{}^c(a), v_M{}^c(a), S_S(0))$ is continuous and decreasing (strictly for some v_S's) in a. Thus, I can write Equation (23) as:

$$S_S(0) = \begin{cases} \int_0^1 C^e_{v_S}(d^*(r, a, S_S(0)), r)\mathrm{d}r & \text{for } a < 0, \\ \int_a^1 C^e_{v_S}(d^*(r, a, S_S(0)), r)\mathrm{d}r & \text{for } a \geq 0. \end{cases}$$

The RHS is continuous in a, as $d^*(v_S, a, S_S(0))$ is continuous in a. For $a = -1$, we have $\int_0^1 C^e_{v_S}(d^*(r, a, S_S(0)), r)\mathrm{d}r = 1$; for $a = 1$, we have $\int_a^1 C^e_{v_S}(d^*(r, a, S_S(0)), r)\mathrm{d}r = 0$; thus, a solution to (23) (given $S_S(0) \in \Theta(1)$) exists. It is unique, as $d^*(v_S, \cdot, S_S(0))$ is weakly decreasing for all and strictly decreasing for some v_S and thus the RHS crosses $S_S(0)$ only once from above.

As $d^*(v_S, \cdot, \cdot)$ is continuous, $a(S_S(0))$ is continuous as well. It is strictly decreasing in $S_S(0)$, as the LHS is strictly increasing in $S_S(0)$ and the RHS is weakly decreasing in $S_S(0)$ and strictly decreasing in a; thus, if $S_S(0)$ increases, Equation (23) is met only if a decreases. As $a(S_S(0))$ is unique and a defines uniquely $(v_M{}^c, v_S{}^c)$, there exist unique $v_M{}^c(S_S(0))$ and $v_S{}^c(S_S(0))$; the former is non-decreasing and the latter non-increasing; and for any $S_S(0)_2 > S_S(0)_1$ we have that $v_M{}^c(S_S(0)_2) > v_M{}^c(S_S(0)_1)$ or $v_S{}^c(S_S(0)_2) < v_S{}^c(S_S(0)_1)$.

The modified set reduces to:

$$S_S(0) > 1-_M \Rightarrow \pi_M\Big(u^c(S_S(0)), \frac{R_M - 1 + S_S(0)}{R_M}\Big) \le \pi_S\Big(v^c(S_S(0)), \frac{R_S - S_S(0)}{R_S}\Big) \quad (31)$$

$$S_S(0) < R_S \Rightarrow \pi_M\Big(u^c(S_S(0)), \frac{R_M - 1 + S_S(0)}{R_M}\Big) \ge \pi_S\Big(v^c(S_S(0)), \frac{R_S - S_S(0)}{R_S}\Big) \quad (32)$$

$$S_S(0) \in \Theta(1). \quad (33)$$

Note that ${v_M}^c(0) = 0$, ${v_S}^c(0) = 1$, ${v_M}^c(1) = 1$ and ${v_S}^c(1) = 0$. Condition (31) -(32) will be trivially met if there exists some $S_S(0) \in \Theta(1)$ such that:

$$\pi_M\Big({v_M}^c(S_S(0)), \frac{R_M - 1 + S_S(0)}{R_M}\Big) = \pi_S\Big({v_S}^c(S_S(0)), \frac{R_S - S_S(0)}{R_S}\Big).$$

If there is no such $S_S(0)$, then it has to be the case that either (a) $LHS > RHS$ for all $S_S(0) \in \Theta(1)$ or (b) $RHS > LHS$ for all $S_S(0) \in \Theta(1)$. However, (a) is possible only if $\max\{0, 1-_M\} = 1-_M$, as $LHS > RHS$ for $S_S(0) = 0$ violates condition (d). And for $S_S(0) = 1-_M$, $LHS > RHS$ meets (31) -(32), as the first inequality doesn't have to hold. For similar reasons, (b) is possible only if $\min\{1, R_S\} = R_S$, in which case $RHS > LHS$ meets (31) -(32). Thus, existence of a solution to (31) -(32) follows. Hence, there exists a solution to the modified and original sets.

For uniqueness, remember that $d^*(v_S, a(S_S(0)), S_S(0))$ is unique and, thus, it suffices to show that the solution to (31) -(32) is unique. Denote the set of all $S_S(0) \in \Theta(1)$ that meet (31) -(32) as Ω^M. Consider $\min \Omega^M = S_S(0)_1$. Note that $S_S(0)_1$ exists as Ω^M is non-empty and $\pi_M(\cdot,\cdot)$, $\pi_S(\cdot,\cdot)$, ${v_S}^c(\cdot)$ and ${v_M}^c(\cdot)$ are continuous. Suppose $S_S(0)_1 = \min\{1, R_S\}$ – then the solution is unique. Now suppose that $S_S(0)_1 < \min\{1, R_S\}$, which implies that for any $S_S(0)_2 \in \Omega^M$ such that $S_S(0)_2 > S_S(0)_1$ we need to have:

$$\pi_M\Big({v_M}^c(S_S(0)_2), \frac{R_M - 1 + S_S(0)_2}{R_M}\Big) \le \pi_S\Big({v_S}^c(S_S(0)_2), \frac{R_S - S_S(0)_2}{R_S}\Big)$$

and for $S_S(0)_1$ we have:

$$\pi_M\Big({v_M}^c(S_S(0)_1), \frac{R_M - 1 + S_S(0)_1}{R_M}\Big) \ge \pi_S\Big({v_S}^c(S_S(0)_1), \frac{R_S - S_S(0)_1}{R_S}\Big).$$

This is a contradiction, as $\frac{\partial}{\partial v_M}\pi_M > 0$, $\frac{\partial}{\partial h}\pi_M \ge 0$, $\frac{\partial}{\partial v_S}\pi_S > 0$, $\frac{\partial}{\partial h}\pi_S \ge 0$, ${v_M}^c(\cdot)$ is weakly increasing, ${v_S}^c(\cdot)$ is weakly decreasing and ${v_M}^c(S_S(0)_2) > {v_M}^c(S_S(0)_1) \vee {v_S}^c(S_S(0)_2) < {v_S}^c(S_S(0)_1)$. Thus $S_S(0)_2$ does not exist and $S_S(0)_1$ is the only element in Ω^M, which completes the proof. □

Proof of Lemma 3. First note that if $S_M(0) + S_S(0) < 1$ then $v_M^c, v_S^c > 0$, which – in turn – by Assumption **??** the reasoning from the proof of Lemma **??** implies that $w_i(v_i^c) = 0$.

Because $S_i(0) \le R_i$ this proves the last statement immediately. The first claim follows then by Proposition 1, because $S_i(0) < R_i$ implies then that $w_i(v_i^c) > 0$ and $S_M(0) + S_S(0) < 1$, which is a contradiction. Finally, $G_i(v_i) = 1 - \frac{S_i(v_i)}{R_i}$ follows from $S_i(0) = R_i$ and the definitions of the distribution of relative skills in sector i and the supply of relative skills in sector i. □

C Comparative Statics

To simplify what follows, I first introduce new notation. The difference between the old and new values of any object O is denoted as $\Delta_c O$. The greater of the old and new values of O is denoted as $\max O$. Thus, for instance, the change in manufacturing size is $\Delta_c S_M(0)$ and the greater critical relative skill in services is $\max {v_S}^c$.

Additionaly, I define the *star relative skill in services*:

$$\bar{v}_S = \sup\{v_S \in [0,1] : \psi(v_S) < 1.\}$$

This definition implies that all workers with $v_S > \bar{v}_S$ join services. The manufacturing analogue can be defined as $\bar{v}_M = \psi(\bar{v}_S)$. Finally, the positive assortative matching function is defined as $P_i(v_i) = 1 - \frac{S_i(v_i)}{R_i}$.

Definition 9. *Vertical differentiation* in manufacturing increases by (strictly) more than in services if, for all (v_M, h):

$$\psi_{v_S}(v_S, c_1)\Delta_c \frac{\partial}{\partial v_M}\pi_M(\psi(v_M, c_1), P_M(\psi(v_s), c_1)) \geq (>)\Delta_c \frac{\partial}{\partial v_S}\pi_s(v_s, P_S(v_S, c_1)).$$

Note that a manufacturing-specific increase in vertical differentiation (Definition **??**) implies trivially that vertical differentiation increased by more in manufacturing than in services.

Definition 10. The matching problems $(Q(c_1), Q(c_2))$ have *(strong) impossibility property* if it is impossible that ${v_S}^c(c_2) < (\leq){v_S}^c(c_1)$ and $\Delta_c S_S(0) > (\geq)0$.

Theorem 2. Suppose $(Q(c_1)$, $Q(c_2))$ exhibit the impossibility property and sector sizes are unchanged. If vertical differentiation increases by more in manufacturing than services, then (i) $P_S(v_S, c_2) \geq P_S(v_S, c_1)$ for all v_S and (ii) $P_M(v_M, c_2) \leq P_M(v_M, c_1)$ for all v_M. If the property is strong, then (i) holds strictly for a positive measure of v_S and (ii) for a positive measure of v_M. If $\Delta_c \frac{\partial}{\partial v_M}\pi_M(\bullet) > 0$ for all (v_M, h), then (iii) $P_S(v_S, c_2) > P_S(v_S, c_1)$ for all $v_S \in [\max {v_S}^c, \max \bar{v}_S)$ and (iv) $P_M(v_M, c_2) < P_M(v_M, c_1)$ for all $v_M \in [\max {v_M}^c, \max \bar{v}_M)$.

Proof of Theorem 2. The results for services are proved in a series of lemmas and the result for manufacturing follow easily (details at the end of the proof). But first, I define the following three sets of services talent levels:

$$\Xi^0 = \{v_S \in [\max {v_S}^c, \min \bar{v}_S] :\ \psi(v_S, c_2) = \psi(v_S, c_1) \wedge P_S(v_S, c_2) \leq P_S(v_S, c_1)\}$$
$$\Xi^1 = \{v_S \in [\max {v_S}^c, \min \bar{v}_S] :\ \psi(v_S, c_2) \leq \psi(v_S, c_1) \wedge P_S(v_S, c_2) < P_S(v_S, c_1)\}$$
$$\Xi^2 = \{v_S \in [\max {v_S}^c, \min \bar{v}_S) :\ \psi(v_S, c_2) \leq \psi(v_S, c_1) \wedge P_S(v_S, c_2) \leq P_S(v_S, c_1)\}.$$

Lemma 6. If vertical differentiation increases by more in manufacturing than services then Ξ^1 (Ξ^2) is empty.

Proof of Lemma 6. Remember that $\frac{\partial}{\partial v_S} P_S(v_S) = \frac{\psi_{v_S}(v_S) C_{uv}(\psi(v_S), v_S)}{R_S}$. Take any $v_{S0} \in \Xi^0$ Note that by Equation (19) we have $\Delta_c P_M(\psi(v_{S0}, c_1)) \geq 0$. Then we have:

$$\Delta_c \frac{\partial}{\partial v_M} w_M(\psi(v_0, c_1)) = \Delta_c \frac{\partial}{\partial v_M} \pi_M(\psi(v_0, c_1), P_M(\psi(v_0, c_1), c_2)) + \int_{P_M(\psi(v_0,c_1),c_1)}^{P_M(\psi(v_0,c_1),c_2)} \frac{\partial^2}{\partial v_M \partial h} \pi_M(\psi(v_0, c_1), r, c_1) \mathrm{d}r \geq (>) 0,$$

as $\Delta_c \frac{\partial}{\partial v_M} \pi_M(v_M, h) \geq (>) 0$ for any (v_M, h), $\pi_M(\bullet)$ is supermodular and $\Delta_c P_M(\psi(v_{S0}, c_1)) \geq 0$. Whereas for v_{S0} we have:

$$\Delta_c \frac{\partial}{\partial v_S} w_S(v_{S0}) = \int_{P_S(v_{S0},c_1)}^{P_S(v_{S0},c_2)} \frac{\partial^2}{\partial v_S \partial h} \pi_S(v_{S0}, r) \mathrm{d}r \leq 0,$$

as $\pi_S(\bullet)$ is supermodular and $\Delta_c P_S(v_{S0}) \leq 0$. By differentiating Equation (**??**) wrt to v for both c_2 and c_1, taking differences and rearranging, we arrive at:

$$\Delta_c \psi_{v_S}(v_{S0}) = \frac{1}{\frac{\partial}{\partial v_M} w_M(\psi(v_{S0}, c_1), c_2)} \big[\Delta_c \frac{\partial}{\partial v_S} w_S(v_{S0}) - \psi_{v_S}(v_{S0}, c_1) \Delta_c \frac{\partial}{\partial v_M} w_M(\psi(v_{S0}, c_1)\big],$$

from which follows trivially that $\Delta_c \psi_{v_S}(v_{S0}) \leq (<) 0$.

Take any $v_{S1} \in \Xi^1$. Suppose that $\Delta_c \psi(v_S) \leq 0$ for all $v_S \in [v_{S1}, \min \bar{v}_S]$, which implies that $\bar{v}_S(c_2) > \bar{v}_S(c_1)$. Remember that both for c_1 and c_2 we need to have $P_S(1) = 1$ and thus $\Delta_c P_S(1) = 0$. Therefore:

$$0 = \Delta_c P_S(v_{S1}) + \frac{\int_{v_{S1}}^{\bar{v}_S(c_2)} C_{v_S}(\psi(v_S, c_2), v_S) \mathrm{d}v_S - \int_{v_{S1}}^{\bar{v}_S(c_1)} C_{v_S}(\psi(v_S, c_1), v_S) \mathrm{d}v_S - \Delta_c \bar{v}_S}{R_S}$$
$$= \Delta_c P_S(v_{S1}, c_1) - \int_{v_{S1}}^{\bar{v}_S(c_1)} \int_{\psi(v_S,c_2)}^{\psi(v_S,c_1)} \frac{C_{uv}(s, v_S)}{R_S} \mathrm{d}s \mathrm{d}v_S - \int_{\bar{v}_S(c_2)}^{\bar{v}_S(c_1)} \frac{1 - C_{v_S}(\psi(v_S, c_2), v_S)}{R_S} \mathrm{d}v_S.$$

Note that as $\psi(v_S, c_2) \leq 1$ it follows that $C_{v_S}(\psi(v_S, c_2), v_S) \leq 1$; hence we have that the two latter terms on the RHS are weakly and the first is strictly negative – contradiction. Therefore there needs to exist some $v_S \in (v_{S1}, \min \bar{v}_S]$ such that $\Delta_c \psi(v_S) > 0$ for Ξ^1 to be non-empty. Denote the set of all such v's as Ξ^3; then it follows that $\inf\Xi^3 \in \Xi^0$.[79] But under $\Delta_c \frac{\partial}{\partial v_M}\pi_M(v_M,h) \geq (>)0$ for all (v_M,h) for any $v_S \in \Xi^0$, $\Delta_c\psi_{v_S}(v_S) \leq (<)0$, which contradicts $v_S = \inf\Xi^3$. Thus Ξ^1 has to be empty and the result for a weak increase in vertical differentiation holds.

Now consider any $v_{S2} \in \Xi^2$. Note that under $\Delta_c \frac{\partial}{\partial v_M}\pi_M(v_M,h) > 0$ for all (v_M,h) there has to exist some arbitrarily small $\epsilon > 0$ such that $v_{S2} + \epsilon < \min \bar{v}_S$ and for any $v_S \in (v_{S2}; v_{S2}+\epsilon]$ we have $\Delta_c\psi(v_S) < 0$. This follows from continuity if $\Delta_c\psi(v_{S2}) < 0$ and from the fact that if $\Delta_c\psi(v_{S2}) = 0$ then $v_{S2} \in \Xi^0$ and $\Delta_c\psi_{v_S}(v_{S2}) < 0$. Therefore, trivially, $\Delta_c P_S(v_{S2}+\epsilon) < 0$ and thus $v_{S2}+\epsilon \in \Xi^1$, which means that Ξ^2 has to be empty and concludes the proof. □

Lemma 7. Suppose Ξ^1 is empty. Consider some $v_{Se} \in [\max {v_S}^c, \min \bar{v}_S]$. Then $\Delta_c P_S(v_{Se}) \geq 0$ implies $\Delta_c P_S(v_{Se}) \geq 0$ for all $v_S \in [v_{Se}, \min \bar{v}_S]$. If Ξ^2 is empty, then additionally $\Delta_c P_S(v_{Se}) > 0$ implies $\Delta_c P_S(v_{Se}) > 0$ for all $v_S \in [v_{Se}, \min \bar{v}_S)$.

Proof. I will start with the first claim. Suppose it is false. Then the set $\Upsilon^1 = \{v_S \in [v_{Se}, \min \bar{v}_S] : \Delta_c P_S(v_S) < 0\}$ has to be non-empty. Take some ${v_S}^1 \in \Upsilon^1$ and define $\Upsilon^2 = \{v_S \in [v_{Se}, v_{S1}] : \Delta_c P_S(v_S) \geq 0\}$. By continuity of $\Delta_c P_S(v_S)$ the point ${v_S}^2 = \max \Upsilon^2$ exists and is $< {v_S}^1$. Therefore, for any $v_S \in ({v_S}^2, {v_S}^1]$ we have $\Delta_c P_S(v_S) < 0$. However, as:

$$\Delta_c P_S({v_S}^1) = \Delta_c P_S({v_S}^2) + \frac{1}{R_S}\int_{{v_S}^2}^{{v_S}^1}\int_{\psi(r,c_1)}^{\psi(r,c_2)} C_{uv}(s,r)\mathrm{d}s\mathrm{d}r,$$

this implies that there exists some $v_{S1} \in ({v_S}^2, {v_S}^1]$ such that $\Delta_c\psi(v_{S1}) < 0$ and thus $v_{S1} \in \Xi^1$ – contradiction.

Let us move to the second claim. Again, suppose it is false. Then the set $\Upsilon^3 = \{v_S \in [v_{Se}, \min \bar{v}_S] : \Delta_c P_S(v_S) \leq 0\}$ has to be non-empty; but as $\Delta_c P_S(v_S)$ is continuous in v, the non-emptiness implies that ${v_S}^3 = \min \Upsilon^3$ exists. Additionally, ${v_S}^3 > v_{Se}$, as $\Delta_c P_S(v_{Se}) > 0$. Define a new set $\Upsilon^4 = \{v_S \in [v_{Se}, {v_S}^3] : \Delta_c\psi(v_S) \leq 0\}$ and ${v_S}^4 = \max \Upsilon^4$; by definition of ${v_S}^3$, for any $v_S < {v_S}^3 \wedge \in \Upsilon^4$ we have that $\Delta_c P_S(v_S) > 0$. As $[v_{Se}, {v_S}^3]$ is a compact set and $\Delta_c\psi(v_S)$ is continuous ${v_S}^4$ won't exist only if Υ^4 is empty; but an empty Υ^4 implies that $\Delta_c\psi(v_S) > 0$ for any $v_S \in [v_{Se}, {v_S}^3]$, which in turn means that $\Delta_c P_S({v_S}^3) > 0$, which contradicts the definition of ${v_S}^3$. Therefore ${v_S}^4$ needs to exist. Now suppose that ${v_S}^4 < {v_S}^3$; then we have $\Delta_c P_S({v_S}^4) > 0$ and for any $v_S \in ({v_S}^4, {v_S}^3], \Delta_c\psi(v_S) > 0$, which implies that $\Delta_c P_S({v_S}^3) > 0$ and also contradicts the definition of ${v_S}^3$. Therefore it has to be the case that ${v_S}^3 = {v_S}^4$; but this implies that $\Delta_c(\psi({v_S}^3)) \leq 0$ and $\Delta_c P_S({v_S}^3) \leq 0$, which contradicts emptiness of Ξ^2 □

Lemma 8. $\Delta_c P_S(\min \bar{v}_S) \geq 0$ implies that (i) for any $v_S > \min \bar{v}_S$ we have $\Delta_c P_S(v_S) \geq 0$ and (ii) for all $v_S \in [\min \bar{v}_S, \max \bar{v}_S)$ we have $\Delta_c P_S(v_S) > 0$.

[79] By continuity of $\Delta_c\psi(v_S)$, which follows from continuity of $\psi(v_S)$.

Proof. Note that $\Delta_c P_S(\min \bar{v}_S) > (\geq) 0$ implies that $\bar{v}_S(c_2) > (\geq)\bar{v}_S(c_1)$[80]. Thus, if $\Delta_c P_S(\min \bar{v}_S) = 0$ then $\min \bar{v}_S = \max {v_S}^c$ and the second claim follows trivially. Whereas if $\Delta_c P_S(\min \bar{v}_S) > 0$ then $\bar{v}_S(c_2) > \bar{v}_S(c_1)$ and by the fact that all agents with $v_S \in (\bar{v}_S, 1]$ join services for sure it follows that for $v_S \in (\bar{v}_S(c_1), \bar{v}_S(c_2))$ we also have $\Delta_c P_S(v_S) > 0$. Claim (i) for $v_S > \max \bar{v}_S$ follows easily from the aforementioned property of $\bar{v}_S$. □

Lemma 9. The (strong) impossibility property implies that if ${v_S}^c(c_2) < (\leq){v_S}^c(c_1)$ then $\Delta_c P_S({v_S}^c(c_1)) > 0$.

Proof. This follows from the fact that $\Delta_c {v_S}^c < (\leq) 0$ implies that $G_S({v_S}^c(c_1)) > (\geq) 0$, the fact that:

$$\Delta_c P_S(v_S) = \frac{1}{R_S}\big((G_S(v_S, c_2) - 1)\Delta_c S_S(0) + S_S(0)(c_1)\Delta_c G_S(v_S)\big) \tag{34}$$

and the fact that ${v_S}^c(c_1) < 1$ and thus $G_S(v_S, c_2) - 1 < 0$. □

Lemma 10. Empty Ξ^1 and impossibility property jointly imply $\Delta_c P_S(\max {v_S}^c) \geq 0$. If either the increase in vertical differentiation is strict or the property is strong then this inequality holds strictly.

Proof. Suppose (strong) impossibility property holds. Define a set $\Xi^5 = \{v_S \in [\max {v_S}^c, \max \bar{v}_S) : \Delta_c \psi(v_{S5}) < 0 \wedge \Delta_c P_S(v_{S5}) \leq 0\}$. By continuity, there has to exist some some arbitrarily small $\epsilon > 0$ such that $v_{S5} + \epsilon \in \Xi^1$; thus, by Lemma 6, an increase in vertical differentiation implies that Ξ^5 has to be empty.

If ${v_S}^c(c_2) < (\leq){v_S}^c(c_1)$ – then by Lemma 9 we have $\Delta_c P_S(\max {v_S}^c) > 0$. If ${v_S}^c(c_2) \geq {v_S}^c(c_1)$ and $\max {v_S}^c \geq \min \bar{v}_S$, then – as $\bar{v}_S > {v_S}^c$ – it has to be that $\bar{v}_S(c_2) > {v_S}^c(c_2) > \bar{v}_S(c_1)$. But as all agents with $v_S > \bar{v}_S$ join services, this implies $\Delta_c P_S({v_S}^c(c_2)) > 0$.

Thus, we only need to show the result for $\max {v_S}^c < \min \bar{v}_S$ and ${v_S}^c(c_2) \geq (>){v_S}^c(c_1)$. As $\Delta_c M^3 = 0$ we have $C({v_M}^c(c_1), {v_S}^c(c_1)) = C({v_M}^c(c_2), {v_S}^c(c_2))$ and thus $\Delta_c {v_S}^c \geq (>) 0$ implies $\Delta_c {v_M}^c \leq 0$. As $\psi({v_S}^c) = {v_M}^c$ and $\psi(v_S)$ is strictly increasing for any c we have: $\psi({v_S}^c(c_2), c_1) \geq (>){v_M}^c(c_1)$, ${v_M}^c(c_1) \geq {v_M}^c(c_2)$ and ${v_M}^c(c_2) = \psi({v_S}^c(c_2), c_2)$, which trivially implies that:

$$\Delta_c \psi({v_S}^c(c_2)) \leq (<) 0.$$

If second property holds, then this inequality holds weakly, which together with empty Ξ^1 implies $\Delta_c P_S({v_S}^c(c_1)) \geq 0$. If the second property is strong, then $\Delta_c \psi({v_S}^c(c_2)) < 0$,

[80] To see this, denote the c_i for which $\bar{v}_S(c_i) = \max \bar{v}_S$ as c_m; then, as $\Delta_c P_S(1) = 0$, we have:

$$0 = \Delta_c P_S(\min \bar{v}_S, c_1) + \frac{1}{R_S}\int_{\bar{v}_S(c_2)}^{\bar{v}_S(c_1)} \frac{1 - C_{v_S}(\psi(v_S, c_m), v_S)}{R_S} \mathrm{d}v_S.$$

As $1 - C_{v_S}(\psi(v_S, c_m), v_S) \geq 0$, the fact that $\Delta_c P_S(\min \bar{v}_S) > (\geq) 0$ implies that for this to hold we need $\bar{v}_S(c_2) > (\geq)\bar{v}_S(c_1)$.

which – as Ξ^5 is empty – implies $\Delta_c P_S(\max v_S{}^c) > 0$. If Ξ^2 is empty, then we have that $\Delta_c\psi(v_S{}^c(c_2)) \le 0$ implies $\Delta_c P_S(\max v_S{}^c) > 0$, which concludes the proof. $\square$

Lemma 11. Empty Ξ^1 and impossibility properties imply jointly that for any $v_S < \max v_S{}^c$, $\Delta_c P_S(v_S) \ge 0$.

Proof. Suppose $\Delta_c v_S{}^c < 0$ – then for all $v_S < \max v_S{}^c$ we have that $G_S(v_S{}^c(c_1)) \ge 0$ and by impossibility property that $\Delta_c S_S(0) \le 0$. Thus, the claim follows from Equation (34). Now suppose that $\Delta_c v_S{}^c \ge 0$. This implies that for any $v_S \le v_S{}^c(c_2)$ it is the case that $\Delta_c G_S(v_S{}^c(c_2)) = 0 - G_S(v_S, c_2) \le 0$ and this expression is decreasing in v. As by Lemma 10 $\Delta_c P_S(v_S{}^c(c_2)) \ge 0$ it follows from Equation (34) that $\Delta_c P_S(v_S) \ge 0$ for all $v_S < \max v_S{}^c$, as required. Note that this implies also that $\Delta_c P_S(0, c) = -\Delta S_S(0) \ge 0$. $\square$

Lemma 12. For all $v_S \in [\max v_S{}^c, \min \bar{v}_S]$, if $\Delta_c P_S(v_S) \ge (>)0$ then $\Delta_c P_M(\psi(v_S, c_2)) \le (<)0$.

Proof. From Equation (19), Lemma 2 and the fact that $P_i(v_i) = 1 - \frac{S_i(v_i)}{R_i}$ follows that:

$$\Delta_c P_S(\psi(v, c_2)) = -\frac{1}{R_M} R_S \Delta_c P_S(v) + \frac{1}{R_M}\Big[\int_{\psi(v,c_1)}^{\psi(v,c_2)} C_u\big(r, v\big)\mathrm{d}r - \int_{\psi(v,c_1)}^{\psi(v,c_2)} C_u\big(r, \phi(r, c_1)\big)\mathrm{d}r\Big].$$

If $\psi(v_S, c_2) \ge \psi(v_S, c_1)$ then for any $r \in [\psi(v_S, c_1), \psi(v_S, c_2)]$, $\phi(r, c_1) \ge v_S$ and my claim follows. If $\psi(v_S, c_2) < \psi(v_S, c_1)$ then for any $r \in [\psi(v_S, c_2), \psi(v_S, c_1)]$, $\phi(r, c_1) < v_S$ and my claim follows as well. $\square$

All results for services follow easily from Lemmas 6, 7, 8, 10 and 11 as well as continuity of $\Delta_c P_S(\cdot)$. As Lemma 8 has an exact manufacturing analogue, the manufacturing results for $v_M \ge \max v_M{}^c$ follow from services results and Lemma 12. The results for $v_M < \max v_M{}^c$ follow from reasoning analogous to that in proof of Lemma 11 once we note that $\Delta_c S_S(0) \le 0$ implies $\Delta_c S_M(0) \ge 0$. $\square$

To prove Propositions 2 and 5, it suffices to show that the impossibility property holds, as a (strict) increase in the spread of surplus implies a (strict) increase in vertical differentiation and the results follow from Theorem 2.

Proof of Proposition 2. The impossibility property is met, as $\Delta_c S_S(0) = 0$. $\square$

Proof of Proposition 5. Suppose the impossibility property does not hold, then $\Delta_c S_S(0) > 0$ and $\Delta_c v_S{}^c < 0$, which implies that $\Delta_c S_M(0) < 0$, $\Delta_c v_M{}^c \ge 0$ and trivially $\Delta_c P_S < 0$

and $\Delta_c P_M > 0$. Sector two expansion implies $S_S(0)(c_1) < R_S$; manufacturing shrinkage implies $R_M > S_M(0)(c_2)$, and thus from (25)-(26) in the proof of Theorem 1 follows that:

$$\pi_M(u^c(c_2), P_S(c_2), c_2) \leq \pi_S(v^c(c_2), P_S(c_2)) \tag{35}$$

$$\pi_S(v^c(c_1), P_S(c_1)) \leq \pi_M(u^c(c_1), P_S(c_1), c_1). \tag{36}$$

Given that $\frac{\partial}{\partial v_S}\pi_S > 0$ and $\frac{\partial}{\partial h}\pi_S \geq 0$, we have that RHS of (35) is strictly less than the LHS of (36) and therefore $\pi_M({v_M}^c(c_2), P_S(c_2), c_2) < \pi_M({v_M}^c(c_1), P_S(c_1), c_1)$. However, as $\Delta_c \pi_M({v_M}^c(c_1), P_M(c_1)) \geq 0$, $\frac{\partial}{\partial v_M}\pi_M > 0$ and $\frac{\partial}{\partial h}\pi_M \geq 0$ this is impossible and impossibility property holds. □

Lemma 13. Scarce jobs and a strict increase in vertical differentiation imply that $\Delta \bar{v}_M \leq 0$, $\Delta \bar{v}_S \geq 0$, with at least one of these holding strictly.

Proof. The first part follows trivially from Lemmas 8 and 12. Suppose $\Delta \bar{v}_S = 0$; consider the set $\Omega^T = \{v_S \in [\max {v_S}^c, \min \bar{v}_S] : \Delta\psi(v_S) < 0\}$ and its minimum ${v_S}^5$. Suppose ${v_S}^5 \neq \min \bar{v}_S$, then, by Theorem 2, $\Delta P_S(\min \bar{v}_S) > 0$, which implies $\Delta \bar{v}_S > 0$, contradiction. Therefore, if $\Delta \bar{v}_S = 0$, then $\Delta\psi(\min \bar{v}_S) < 0$, which implies $\Delta \bar{v}_M < 0$ and concludes the proof. □

Proof of Proposition 3. (i) From Theorem 2 in Appendix C and Lemma 3 follows that for (strictly) scarce jobs a strict increase in vertical differentiation results in a (strict) decrease in ${v_S}^c$. As $w_S(v_S^c)$ remains unchanged and surplus function is supermodular, the (strict) increase in lowest wage follows from inspection of Equation (**??**). Note that for any ${v_S}'' > {v_S}' \geq {v_S}^c$ we have:

$$w_S({v_S}'') = \int_{{v_S}'}^{{v_S}''} \frac{\partial}{\partial v_S}\pi_S(r, P_S(r))\mathrm{d}r + w_S({v_S}'). \tag{37}$$

As $P_S(r)$ increases and surplus is supermodular, it follows that $w_S({v_S}'')$ increases by more than $w_S({v_S}')$.

(ii) Proposition 2 and of Lemma 13 (in Appendix C) imply that $\bar{v}_M(c_2) \leq \bar{v}_M(c_1)$ and $\bar{v}_S(c_2) \geq \bar{v}_S(c_1)$ and at least one of these inequalities is strict. This gives:

$$w_M(\bar{v}_M(c_1), c_2) \geq w_M(\bar{v}_M(c_2), c_2) = w_S(\bar{v}_S(c_2), c_2) \geq w_S(\bar{v}_S(c_1), c_2)$$
$$w_S(\bar{v}_S(c_2), c_1) \geq w_S(\bar{v}_S(c_1), c_1) = w_M(\bar{v}_M(c_1), c_1) \geq w_M(\bar{v}_M(c_2), c_1)$$

with at least one inequality holding strictly, which trivially implies:

$$w_M(\bar{v}_M(c_1), c_2) - w_M(\bar{v}_M(c_1), c_1) > w_S(\bar{v}_S(c_2), c_2) - w_S(\bar{v}_S(c_2), c_1). \tag{38}$$

Thus, $w_M(\bar{v}_M(c_1))$ increases strictly. For any $v_M > \bar{v}_M$ we have that:

$$w_M(v_M) = \int_{\bar{v}_M}^{v_M} \frac{\partial}{\partial v_M} \pi_M(r, G_M(r))\mathrm{d}r + w(\bar{v}_M(c_1)). \tag{39}$$

For $v_M > \bar{v}_M(c_1)$, $G_M(v_M)$ does not change; and as surplus' spread implies that $\frac{\partial}{\partial v_M}\pi_M(v_M, h)$ strictly increases, it follows that $w_M(v_M, c_2) > w_M(v_M, c_1)$ for any $v_M \in [\bar{v}_M(c_1), 1]$.

(iii) Follows from (i), (ii) and the fact that with scarce jobs $C^i(c_2) = C^i(c_1)$.

I will turn now to wages of the least talented agents with strictly scarce jobs. By Theorem 2 we have that ${v_M}^c(c_2) > {v_M}^c(c_1)$. As wages strictly increase in talent, it follows from definition of critical relative skill that $w_M({v_M}^c(c_2), c_1) > w_M({v_M}^c(c_1), c_1) = 0$. Note that existence of a positive mass of agents for whom wages decrease (increase) follows from continuity of wage functions. □

Proof of Proposition 4. The first claim follows from inspection of Equation (40): each firm is matched with a less productive agent, so $(P_S)^{-1}$ decreases for all h and the profit constant falls as well, as it is equal to $\pi_S({v_S}^c, 0)$ and ${v_S}^c$ falls by Proposition 2. The second follows trivially from the fact that the pool of services firms is unchanged, the supply of talent falls and surplus is increasing in talent.

As for the last two claims, note that, fixing sectoral supply functions, an increase in surplus' levels increases the total surplus in the economy. As the stable assignment is surplus maximising in this model, the change from the old to new stable assignment has to further improve total surplus[81]. Finally, as services's total surplus falls, it has to increase in manufacturing. □

Proof of Lemma 4. If $\Pi(\mathbf{x}, z, i)$ is strictly increasing in productivity, then $h_i(\cdot)$ must be strictly increasing in z.[82] Denoting the distribution of Z conditonal on $i = M$ as H_{Z_M} it follows from the fact that h_M has standard uniform distribution that $H_{Z_M}(z) = h_M(z)$. Thus, a FOSD improvement in the distribution of $Z|M$ implies that $h_M(z, c_2) \leq h_M(z, c_1)$. Consider such z and z' that $h_M(z', c_2) = h_M(z, c_1)$; it follows that $z' \geq z$. Thus:

$$\frac{\partial}{\partial v_M}\pi_M(v_M, h_M(z', c_2), c_2) = \frac{\partial}{\partial v_M}\pi_M(v_M, h_M(z', c_1), c_1) \geq \frac{\partial}{\partial v_M}\pi_M(v_M, h_M(z, c_1), c_1),$$

as required, where the final inequality follows from the supermodularity of reduced surplus function. □

Lemma 14. If jobs are abundant and surplus both increases and becomes strictly more spread out, the lowest profit rises in manufacturing and a decreases in services

[81] This is the case, as my model can be rewritten as a special case of the assignment model described in Gretsky et al. (1992) and thus the equivalence of stable and efficient matching showed by them holds for my model as well.

[82] For any $z' > z$ we have $\pi_i(v_i, h_i(z')) > \pi_i(v_i, h_i(z))$. Suppose that $h_i(z') \leq h_i(z)$. Then by Assumption A1.3 $\pi_i(v_i, h_i(z')) \leq \pi_i(v_i, h_i(z))$; contradiction.

$(r_M(P_M(c_2), c_2) \geq r_M(P_M(c_1), c_1)$ and $r_S(P_S(c_2), c_2) \leq r_S(P_S(c_1), c_1))$, where $P_i = 1 - \frac{S_i(0)}{R_i}$.

Proof. In manufacturing, there are two possibilities: $S_M(0)(c_1) <_M$ and $S_M(0)(c_1) =_M$. If the former is the case, then $r_M(P_M(c_1), c_1) = 0$ and the result follows trivially. If the latter is true, then $S_M(0)(c_2) =_M$ and by Proposition 5 we have $v_M{}^c(c_2) \geq v_M{}^c(c_1)$, $v_S{}^c(c_2) \leq v_S{}^c(c_1)$, $P_M(c_1) = P_S(c_2)$ and $P_S(c_2) = P_S(c_1)$.[83] This implies that $r_M(P_M(c_2), c_2) \geq r_M(P_M(c_1), c_1)$, as by measure consistency and Equations (25)-(26):

$$r_M(P_M(c_i), c_i) = \pi_M(v_M{}^c(c_i), P_M(c_i)) - \pi_S(v_S{}^c(c_i), P_S(c_i)),$$

for $i = 1, 2$. The result for services follows from analogous reasoning, but the two cases are $S_S(0)(c_2) < R_S$ and $S_S(0)(c_2) = R_S$. □

Proof of Proposition 6. Consider $T = A_A^1(c_2) \cap A_A^1(c_1)$, the set of agents who work in services in both matching problems. Denote the least talented of those agents –$\inf_y T$ – as $\max_c v_S{}^c$. Her wage depends on two factors: positively on the surplus she produces and negatively on its share received by the firm she is matched with. The first factor always increases, as she is matched with a more productive firms. The change in the second factor can be both positive (for $\max_c v_S{}^c = v_S{}^c(c_1)$) and negative (for $\max_c v_S{}^c = v_S{}^c(c_2)$). If the former is the case, however, then the increase in surplus received by her firm:

$$\Delta_c r_S(P_S(v_S{}^c(c_1)) = \int_{v_S{}^c(c_2)}^{v_S{}^c(c_1)} \frac{\partial}{\partial v_S} P_S(r, c_2) \frac{\partial}{\partial h} \pi_S(r, P_S(r, c_2)) \mathrm{d}r + \Delta_c C_R^2,$$

is always less than the increase in the surplus she produces (by Lemma 14):

$$\pi_S(v_S{}^c(c_1), P_S(v_S{}^c(c_1), c_2)) - \pi_S(v_S{}^c(c_1), P_S(c_1)) = \int_{v_S{}^c(c_2)}^{v_S{}^c(c_1)} \frac{\partial}{\partial v_S} P_S(r, c_2) \frac{\partial}{\partial h} \pi_S(v_S{}^c(c_1), P_S(r, c_2)) \mathrm{d}r.$$

Thus, $w_S(\max_c v_S{}^c, c_2) - w_S(\max_c v_S{}^c, c_1) \geq 0$ and by inspection of Equation (37) we have that $w_S(v_S'', c_2) - w_S(v_S'', c_1) \geq w_S(v_S', c_2) - w_S(v_S', c_1)$, for any $v_S'' > v_S' \geq \max_c v_S{}^c$. It follows that wages increase for all $v_S \in T$. This and revealed preference imply that all agents who used to work in services are better off[84]. As the top wages increase by more in manufacturing (by the same reasoning as in the proof of Proposition 3) it follows that wages increase for most talented manufacturing workers. □

[83] The change in critical levels follows from the fact that with fixed sector sizes an improvement in skill supply is equivalent to a FOSD improvement in the distribution of skill.

[84] This is trivial if $v_S{}^c(c_2) \leq v_S{}^c(c_1)$. If $v_S{}^c(c_2) \geq v_S{}^c(c_1)$ then the agents with $v_S \in [v_S{}^c(c_1), v_S{}^c(c_2))$ will move to manufacturing; but as the lowest wages are the same in both sectors, they earn more than $w_S(v_S{}^c(c_2), c_2)$, which in turn is greater than their old wage.

Proof of Proposition 7. The result wrt manufacturing (two) profits follows from Proposition 5, the definition of an increase (decrease) in supply, the definition of PAM ($P^i(\cdot)$), Lemma 14 in Appendix C and inspection of Equation (40). The increases in total output follow from analogous reasoning as in the proof of Proposition 4. □

Proof of Proposition 8. First, notice that symmetry and Theorem 1 imply trivially that $v_M^c = v_S^c$, $v_M^* = v_S^*$ and $\psi(v_S) = v_M$ characterise the unique separating function. Therefore, v_i^c solves $C(v_i^c, v_i^c) = \max\{1-2R_M, 0\}$, $S_i(v_i) = \frac{1-C(v_i,v_i)}{2}$ and $G_i(v_i) = \frac{C(v_i,v_i)-1}{2\min\{R_i,0.5\}}+1$ for $v_i \geq v_i^c$.

I will prove the proposition for manufacturing, results for services follow from symmetry. (i) This follows immediately from the definition of an increase in concordance $C(v_i^c, v_i^c) = \max\{1-2R_M, 0\}$ and $G_i(v_i) = \frac{C(v_i,v_i)-1}{2R_i} + 1$.

(ii) Recall Equation (9). The lowest wage, $w_i(v_i^c)$, remains unchanged.[85] Since surplus function, marginals and sector sizes do not change either, by Definition **??** wages depend only on within-sector relative skill distributions and hence, they increase for any $v_M > {v_M}^c(c_2)$; for $v_M \in [0, {v_M}^c(c^2)]$ they remain constant (and equal to 0).

(iii) Trivial, as $w_i(v_i^c, c_2) = w_i(v_i^c, c_1)$ (see footnote 85) and $w_M(1, c_2) \geq w_M(1, c_1)$.

(iv) Denote the worst matched firm in sector i as $P_i = 1 - \frac{S_i(0)}{R_i}$. As the sector size does not change, neither do P_i. and A_F^i (the set of matched firms)[86]. All unmatched firms make zero profit and the profit function for all matched firms in sector i is given by (see Sattinger, 1979):

$$r^i(h_i) = \int_{P_i}^{h_i} \pi_h^i((P^i)^{-1}(r), r)\mathrm{d}r + r_i(P_i). \tag{40}$$

The lowest profit $r_i(P_i)$ will either remain unchanged (for $R_M \geq \frac{1}{2}$) or will fall (for $R_M < \frac{1}{2}$, as then $r_i(P_i) = \pi_M(v_{Mc}, 0)$ and ${v_M}^c$ decreased). Thus, (iv) follows from (i).

(v) By (i), every firm produces lower surplus. □

Proof of Proposition 9. The wage received by agent of rank t is paid by firm $h_i(t) = \frac{R_i - \frac{1}{2} + t}{R_i}$ (by symmetry, job abundance and PAM). Thus:

$$W(t) = \pi_i(G_i^{-1}(t), h_i(t)) - r_i(h_i(t)),$$

where $G_i(v_i) = C(v_i, v_i)$. By Proposition 8 an increase in concordance will decrease both $G_i^{-1}(t)$ and $r_i(h_i(t))$, which means that the change in $W(t)$ is of ambiguous sign in general.

[85] If jobs are strictly scarce ($R_M \leq \frac{1}{2}$), then $w_i(v_i^c) = 0$. If jobs are abundant ($R_M > \frac{1}{2}$), then $w_i(v_i^c) = \pi_M({v_M}^c, S_M(0))$ and by Corollary **??** we have that ${v_M}^c = 0$. Finally, for $R_M + R_S = 1$ the lowest wage is equal to some constant $\in (0, 1)$ – and there is no reason why this constant would change.

[86] For non-strictly supermodular surplus functions and the abundant jobs case ($R_M + R_S > 1$), the set of matched firms is not unique. Hereafter, I will assume that in such cases firms with $h \geq P_i$ become matched. This simplifies notation and is without any loss in generality, as the profits of the relevant firms will be 0.

However, symmetry and job abundance imply that $v_i^c = 0$ and $W(0) = \pi_i(0, 1 - \frac{R_i}{2})$ irrespective of concordance, which proves i). Similarly, $G_i^{-1}(1) = 1$ in any specification of this model, which proves ii).

To prove iii) consider this auxiliary surplus function $\pi_i^a(v_i, h_i) = 1 - e^{-\delta v_i}$, with $\delta > \max\{-\min_{(v,h)\in[0,1]^2}(\frac{\frac{\partial^2}{\partial^2 v_i}\pi(v_i,h_i)}{\frac{\partial}{\partial v_i}\pi(v_i,h_i)}), 0\}$.[87] π_i^a does not depend on firms' productivity and hence $r_i^a(h_i) = 0$. Thus, a strict increase in concordance implies a strict fall in $W^a(t)$ for all $t \in (0,1)$. As $W^a(0) = 0$ irrespective of copula, it must be the case that there exists some $\bar{t} \in (0,1)$ such that $\Delta_c \frac{\partial}{\partial t} W^a(t) < 0$ for all $t \in (0, \bar{t})$.[88] This gives:

$$\frac{\delta e^{-\delta G_i^{-1}(t,c_2)}}{g(G^{-1}(t,c_2),c_2)} - \frac{\delta e^{-\delta G_i^{-1}(t,c_1)}}{g(G^{-1}(t,c_1),c_1)} < 0,$$

which after some rearrangement yields:

$$\ln\Big(\frac{g(G^{-1}(t,c_2),c_2))}{g(G^{-1}(t,c_1),c_1)}\Big) > -\delta \Delta_c G_i^{-1}(t),$$

for all $t \in (0, \bar{t})$.

Recall that $G_i(v_i) = C(v_i, v_i)$ and thus does not depend on the surplus function. Suppose that $\Delta_c \frac{\partial}{\partial t} W(t) \geq 0$ for some $t \in (0, \bar{t})$. This implies

$$\frac{\frac{\partial d}{\partial v_i}\pi_i(G_i^{-1}(t,c_2),h_i(t))}{g(G^{-1}(t,c_2),c_2)} - \frac{\frac{\partial d}{\partial v_i}\pi_i(G_i^{-1}(t,c_1),h_i(t))}{g(G^{-1}(t,c_1),c_1)} \geq 0,$$

which yields:

$$-\int_{G_i^{-1}(t,c_2)}^{G_i^{-1}(t,c_1)} \frac{\frac{\partial^2}{\partial^2 v_i}\pi(s,h_i)}{\frac{\partial}{\partial v_i}\pi(s,h_i)} \mathrm{d}s \geq \ln\Big(\frac{g(G^{-1}(t,c_2),c_2))}{g(G^{-1}(t,c_1),c_1)}\Big) > -\delta \Delta_c G_i^{-1}(t).$$

This gives:

$$-\int_{G_i^{-1}(t,c_2)}^{G_i^{-1}(t,c_1)} \min_{(v,h)\in[0,1]^2} \frac{\frac{\partial^2}{\partial^2 v_i}\pi(v_i,h_i)}{\frac{\partial}{\partial v_i}\pi(v_i,h_i)} \mathrm{d}s \geq -\int_{G_i^{-1}(t,c_2)}^{G_i^{-1}(t,c_1)} \frac{\frac{\partial^2}{\partial^2 v_i}\pi(s,h_i)}{\frac{\partial}{\partial v_i}\pi(s,h_i)} \mathrm{d}s \geq -\delta \Delta_c G_i^{-1}(t),$$

$$-\Big(\min_{(v,h)\in[0,1]^2} \frac{\frac{\partial^2}{\partial^2 v_i}\pi(v_i,h_i)}{\frac{\partial}{\partial v_i}\pi(v_i,h_i)}\Big) \geq \frac{1}{\Delta_c G_i^{-1}(t)} \int_{G_i^{-1}(t,c_2)}^{G_i^{-1}(t,c_1)} \frac{\frac{\partial^2}{\partial^2 v_i}\pi(s,h_i)}{\frac{\partial}{\partial v_i}\pi(s,h_i)} \mathrm{d}s \geq \delta,$$

which contradicts the fact that $\delta > -\min_{(v,h)\in[0,1]^2}(\frac{\frac{\partial^2}{\partial^2 v_i}\pi(v_i,h_i)}{\frac{\partial}{\partial v_i}\pi(v_i,h_i)})$. Thus, $\Delta_c \frac{\partial}{\partial t} W(t) < 0$ for all $t \in (0, \bar{t})$, which implies that $\Delta_c W(t) = \Delta_c W(0) + \int_0^t \Delta_c \frac{\partial}{\partial t} W(s) \mathrm{d}s < 0$ for all $t \in (0, \bar{t})$,

[87] $\max_{(v,h)\in[0,1]^2} \frac{\partial^2}{\partial^2 v_i}\pi_i$ is finite and $\min_{(v,h)\in[0,1]^2} \frac{\partial}{\partial v_i}\pi(v_i,h_i) > 0$, both by Assumption 1.

[88] $\Delta_c W(t) = \Delta_c W(0) + \int_0^t \Delta_c \frac{\partial}{\partial t} W^a(s)\mathrm{d}s$. Suppose $\bar{t}$ does not exist. Then there exists some t_1 such that $\Delta_c \frac{\partial}{\partial t} W^a(t) \geq 0$, which implies $\Delta_c W(t) \geq 0$ for $t \in (0, t_1)$; contradiction.

which completes the proof. □

D Approximating Gaussian-Exponential

Consider the following specification in the canonical form:

$$\pi_i(v_i, h_i) = \Big(A_i - \frac{1}{\delta} e^{-\delta\big(\Phi^{-1}((1-2a)v_i + a)\sigma_i + \mu_i\big)}\Big)\Big((\bar{\beta}_i - \underline{\beta}_i)h_i + \underline{\beta}_i)\Big)^{\gamma_i},$$
$$C(v_M, v_S) = \mathbf{\Phi}_\rho\Big(\Phi^{-1}(v_M), \Phi^{-1}(v_S)\Big),$$

where $a \in [0, \frac{1}{2})$. This specification i) reduces to the Gaussian-Exponential specification for $a = 0$ and ii) results in π_i that is twice continuously differentiable on $[0,1]^2$ for $a \in (0, \frac{1}{2})$.[89]

First, suppose that $\delta < 0$; then $A_i - \frac{1}{r} e^{-r\big(\Phi^{-1}((1-2a)v_i + a)\sigma_i + \mu_i\big)} > 0$ for all $v_i \in [0,1]$ and $a \in (0, \frac{1}{2})$; thus, the specification above meets all of the model assumptions for $a \in (0, \frac{1}{2})$.

If δ is positive, then the situation is slightly more complex, as two further assumptions are violated in the GE specification. Define $\bar{v}_i = \Phi(-\frac{\ln(r_i A_i) - \mu_i}{\sigma_i}) \in (0, 1)$; then, firstly, surplus will not increase in productivity for $v_i < \bar{v}_i$. Secondly, surplus will be negative for small enough $v_i < \bar{v}_i$. These two assumptions will also be a problem for very small but positive a and therefore the argument above cannot be directly applied. Suppose, however, that $R_1 + R_2 < 1 - \bar{v}_i$. Then those two violations are inconsequential, because workers with $v_i < \bar{v}_i$ would never work in sector i: even if sector $j \neq i$ hired a measure of R_j best sector i workers, there would be still more than R_i measure of workers with $v_i \geq \bar{v}_i$ available, which is enough to fill all jobs in sector i. Thus, the equilibrium characterisation specified in Section 2.2 yields the correct equilibrium and all my comparative statics results hold if jobs are sufficiently scarce.

References

Acemoglu, D. and D. Autor (2011). *Skills, Tasks and Technologies: Implications for Employment and Earnings*, Volume 4 of *Handbook of Labor Economics*, Chapter 12, pp. 1043–1171. Elsevier.

Akerman, A., I. Gaarder, and M. Mogstad (2015). The Skill Complementarity of Broadband Internet. *The Quarterly Journal of Economics 130*(4), 1781–1824.

[89] It also corresponds to a fundamental specification with the same distributions are Gaussian-Exponential, but following surplus function:

$$\Pi(\mathbf{x}, z, i) = \Big(A_i + \frac{1 - e^{-\delta\Phi^{-1}((1-2a)\Phi(\frac{1}{\sigma_i}(\boldsymbol{\alpha}_i \mathbf{x}^T - \mu_i)) + a)\sigma_i + \mu_i}}{\delta}\Big) z^{\gamma_i}.$$

Anderson, S. (2003). Why Dowry Payments Declined with Modernization in Europe but Are Rising in India. *Journal of Political Economy 111*(2), pp. 269–310.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics 118*(4), 1279–1333.

Bakija, J., A. Cole, and B. Heim (2010, March). Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data. Department of Economics Working Papers 2010-22, Department of Economics, Williams College.

Becker, G. S. (1973, July-Aug.). A Theory of Marriage: Part I. *Journal of Political Economy 81*(4), 813–46.

Bickel, P. and E. Lehmann (1979). Descriptive Statistics for Nonparametric Models IV. Spread. In J. Jureckova (Ed.), *Contributions to Statistics : Jaroslav Hajek Memorial Volume*. D. Reidel Pub. Co Dordrecht, Holland ; Boston.

Bielecki, A. (1956). Une remarque sur la methode de Banach-Cacciopoli-Tikhonov dans la theorie des equations differentielles ordinaires. *Bull. Acad. Polon. Sci. Cl. III. 4*, 261–264.

Boehm, M. (2015, 03). The Price of Polarization: Estimating Task Prices Under Routine-Biased Technical Change.

Burzyński, M. and P. Gola (2017). Mexican Immigration to the US: Selection, Sorting and Matching.

Chiappori, P.-A., R. McCann, and L. Nesheim (2010, February). Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. *Economic Theory 42*(2), 317–354.

Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque (2011). Matching with a Handicap: The Case of Smoking in the Marriage Market. (1011-07).

Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque (2012). Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market. *Journal of Political Economy 120*(4), 659 – 695.

Costinot, A. and J. Vogel (2010, 08). Matching and Inequality in the World Economy. *Journal of Political Economy 118*(4), 747–786.

Costrell, R. M. and G. C. Loury (2004, December). Distribution of Ability and Earnings in a Hierarchical Job Assignment Model. *Journal of Political Economy 112*(6), 1322–1363.

Dupuy, A. (2015, January). The Assignment of Workers to Tasks with Endogenous Supply of Skills. *Economica*.

Gabaix, X. and A. Landier (2008, February). Why Has CEO Pay Increased so Much? *The Quarterly Journal of Economics 123*(1), 49–100.

Gould, E. D. (2002). Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States. *Journal of Labor Economics 20*(1), 105–147.

Gretsky, N. E., J. M. Ostroy, and W. R. Zame (1992, January). The Nonatomic Assignment Model. *Economic Theory 2*(1), 103–27.

Grossman, G. M., E. Helpman, and P. Kircher (2013, October). Matching and Sorting in a Global Economy. NBER Working Papers 19513, National Bureau of Economic Research, Inc.

Hasselblatt, B. and A. Katok (2003). *A First Course in Dynamics: With a Panorama of Recent Developments*. Cambridge University Press.

Heckman, J. and G. Sedlacek (1985). Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-selection in the Labor Market. *Journal of Political Economy 93*(6), 1077–1125.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Monographs on Statistics and Applied Probability. Chapman & Hall.

Justman, M. and J.-F. Thisse (1997, September). Implications of the mobility of skilled labor for local public funding of higher education. *Economics Letters 55*(3), 409–412.

Kaplan, S. and J. Rauh (2010). Wall Street and Main Street: What Contributes to the Rise in the Highest Incomes? *Review of Financial Studies 23*(3), 1004–1050.

Kaplan, S. N. and J. Rauh (2013). It's the Market: The Broad-Based Rise in the Return to Top Talent. *Journal of Economic Perspectives 27*(3), 35–56.

Katz, L. and K. M. Murphy (1992). Changes in relative wages, 19631987: Supply and demand factors. *The Quarterly Journal of Economics 107*(1), 35–78.

Legros, P. and A. F. Newman (2002). Monotone Matching in Perfect and Imperfect Worlds. *Review of Economic Studies 69*, 925–942.

Lindenlaub, I. (2017). Sorting Multidimensional Types: Theory and Application. *Review of Economic Studies*.

Mak, E. and A. Siow (2017). Occupational Choice and Matching in the Labor Market. *IZA Discussion Paper No. 10584.*.

McCann, R. J., X. Shi, A. Siow, and R. Wolthoff (2015). Becker meets ricardo: Multisector matching with communication and cognitive skills. *Journal of Law, Economics, and Organization 31*(4), 690–720.

Melitz, M. J. (2003, November). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Milgrom, P. and C. Shannon (1994, January). Monotone Comparative Statics. *Econometrica 62*.

Poutvaara, P. (2008, 09). Public and Private Education in an Integrated Europe: Studying to Migrate and Teaching to Stay? *Scandinavian Journal of Economics 110*(3), 591–608.

Rodrik, D. (2013). Unconditional Convergence in Manufacturing. *Quarterly Journal of Economics 128*(1), 165–204. This is a substantially revised version of "Unconditional Convergence" below.

Rosen, S. (1981). The Economics of Superstars. *American Economic Review 71*(5), 845–58.

Roy, A. D. (1951). Some Thoughts On The Distribution Of Earnings. *Oxford Economic Papers 3*(2), 135–146.

Sampson, T. (2014). Selection into Trade and Wage Inequality. *American Economic Journal: Microeconomics 6*(3), 157–202.

Sattinger, M. (1975). Comparative Advantage and the Distributions of Earnings and Abilities. *Econometrica 43*(3), 455–468.

Sattinger, M. (1979, March). Differential Rents and the Distribution of Earnings. *Oxford Economic Papers 31*(1), 60–71.

Sattinger, M. (1993). Assignment Models of the Distribution of Earnings. *Journal of Economic Literature 31*(2), pp. 831–880.

Scarsini, M. (1984). On measures of concordance. *Stochastica 8*(3), 201–218.

Shaked, M. and G. Shanthikumar (2007). *Stochastic Orders* (1 ed.). Springer Series in Statistics. Springer-Verlag New York.

Sklar, A. (1959). *Fonctions de repartition a n dimensions et leurs marges.*

Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics 24*(2), 235–270.

Tervio, M. (2008, June). The Difference That CEOs Make: An Assignment Model Approach. *The American Economic Review 98*(3), 642–668.

Teulings, C. (2005). Comparative Advantage, Relative Wages, and the Accumulation of Human Capital. *Journal of Political Economy 113*(2), 425–461.

Teulings, C. N. (1995, April). The Wage Distribution in a Model of the Assignment of Skills to Jobs. *Journal of Political Economy 103*(2), 280–315.

Topkis, D. M. (1978). Minimizing a Submodular Function on a Lattice. *Operations Research 26*(2), 305–321.