Abstract

Workers bring skillsets to the workplace with every skill geared towards a particular task. Firms on the other hand, are looking to select workers with the best possible skill-mix. So how do workers and firms match up given this multidimensional setting? I develop a multidimensional notion of assortative matching that helps to analytically solve this problem, using a framework based on Gaussian copulas. The key insights from this closed-form solution include: First, contrary to the one-dimensional setting, my model enables many ways to assortatively match workers with firms. This is why the assignment can be pinned-down not only by the sign, but also by the strength of worker-firm complementarities in each task. Second, I show how the multivariate distributions of firm and worker attributes influence both assignment and wages, in addition to how these supply and demand characteristics interact with technological change.

Keywords. Multidimensional Heterogeneity, Assortative Matching, P-matrix, Gaussian Copula, Task-Biased Technological Change.
1 Introduction

This paper develops a theory of pairwise matching with multidimensional types and transferable utility. I begin by generalizing the notions of positive assortative matching (PAM) and negative assortative matching (NAM) to the multidimensional setting. Assortative matching provides the equilibrium assignment with considerable structure. This makes it possible to analytically solve for assignment and supporting prices in a class of multivariate models that is based on Gaussian copulas. I then use this closed form solution to better understand both the technical and economic implications of multidimensional matching compared to the one-dimensional setting.

Understanding multidimensional matching from a theoretical point of view is important because in many markets, matching is based on multiple characteristics. In empirical economics, this is a widely accepted fact. A notable example is the study by Willis and Rosen (1979), which rejects the premise that talent is one-dimensional. It shows that college graduates would have performed worse than observationally similar high school dropouts, had the graduates dropped out of high school themselves. This suggests that worker performance depends on a bundle of different skills such as intellectual and manual skills. Some people are strong in both skills (e.g. mechanical engineers or surgeons) but this is rather exceptional since most agents specialize in only one of these skills.

This points to the main reason for requiring matching models with multidimensional heterogeneity: In the data, characteristics are not perfectly correlated, which is why agents can only be partially ordered. Hence, it is problematic to aggregate different attributes into a single one-dimensional index, according to which the agents are ranked and matched.

The mathematical literature on optimal transport has developed powerful tools to study the existence, uniqueness and purity (i.e. deterministic matching) of multidimensional assignments in the transferable utility context. However, the optimal transport approach provides little guidance on how to explicitly solve for the equilibrium allocation and its supporting price. This makes it difficult to derive economic implications from multivariate matching problems and to develop statistical models that can be applied to the data. This paper makes an attempt towards filling this gap.

Section 2 introduces the general theoretical framework. I develop a pairwise assignment model in which I interpret the two sides of the market as workers and firms. For the most part of the paper, both sides have two characteristics. Each worker possesses two skills and brings this skill bundle to the workplace. Let them be intellectual and manual skills. To produce output, two tasks must be performed – one requiring intellectual skills and the other requiring manual abilities. Firms, in turn,

---

1 There is a growing empirically-motivated literature that studies multidimensional matching on the marriage market. Chiappori et al. (2012a) and Chiappori et al. (2012b) find that American spouses trade education for beauty (and education for non-smoking) when selecting a partner. Choo and Siow (2006) also recognize the importance of multiple characteristics when estimating the gains from marriage. In a similar vein, Galichon and Salanié (2010) model the trade-off between different characteristics faced by spouses in order to estimate the marriage surplus function. Dupuy and Galichon (2012) show that not only education but also personality traits matter for marriage choices.

2 In non-technical terms, the optimal transport problem is to find a measure-preserving map that carries one distribution into another at minimal cost. This approach relies on linear programming. See Galichon and Salanié (2010), Chiappori et al. (2012b) and McCann et al. (2012) for economic applications as well as the purely theoretical work by Gretsky et al. (1992), Chiappori et al. (2010), Ekeland (2010), Dizdar (2012) and Dizdar and Moldovanu (2012). I will discuss the relation between the most related literature and my work in Section 5, after having presented the model.
differ in the complexity level of each task, which I will also refer to as productivities.

Within this task-based framework, I develop a generalization of PAM and NAM to the multi-dimensional setting. In non-technical terms, my definition of PAM (NAM) is that better (worse) workers match with better firms within each task dimension. Moreover, this pattern dominates the matching forces that relate skills and productivities across tasks. For instance, the best scientists usually work in the best universities (universities put a lot of weight on intellectual skills but little on manual dexterity) whereas the best technicians often work in professional motor sports (which require manual skills more than intellectual abilities), and not the other way around. This notion is captured by the \textit{P-matrix property} of the Jacobian of the matching function and is an intuitive generalization of one-dimensional assortative matching. P-matrix properties have so far not been exploited in the matching literature but have been used in other fields of economics to rule out multiple equilibria.\footnote{See, for instance, Simsek et al. (2005).}

I derive technological conditions, under which the equilibrium is not only pure but also assortative.\footnote{The presented condition is related to the \textit{twist condition} from optimal transport but is not equivalent. See Section 5.} Intuitively, if complementarities (substitutabilities) of skills and productivities within tasks dominate the ones across tasks, then the optimal assignment will be characterized by PAM (NAM).

In Section 3, assortative matching proves useful to analytically solve for the equilibrium assignment and wage function in the class of multidimensional models that is based on (a) Gaussian copulas and (b) linear-quadratic technology. Notice that this class is quite broad since I allow for arbitrary marginal skill and productivity distributions (possibly from different families of distributions) where dependence of the data is modeled via Gaussian copulas. Moreover, the model can deal with arbitrary skill and productivity correlations.

The main difficulty in solving this multidimensional model is that there are many ways to assortatively match workers with firms, which is due to the incomplete order of types. What matters for pinning down the assignment is not only the sign but especially the strength of skill-productivity complementarities in each task. This is in stark contrast to one-dimensional matching where there is only one possibility for matching workers to firms in an assortative way. A straightforward implication of this is that the sign of complementarities (i.e. super- or submodularity of the technology) determines the exact allocation, independent of the degree of super- or submodularity (Becker (1973)).

Much of Section 3 addresses this complication of how to explicitly pin down the \textit{vector-valued matching function}, which maps bivariate skills to bivariate productivities. The main idea is to make the two-dimensional problem similar to two separate one-dimensional problems since one-dimensional matching is much better understood. To this end, I use a measure-preserving transformation to (temporarily) un-correlate the Gaussian skills and productivities. This is the multidimensional equivalent to standardization. Along with assortative matching, this transformation gives rise to a tractable market clearing condition which is an important step towards computing the explicit assignment.

However, even if the technology is such that matching is positive assortative, there are still many ways to map workers to firms. One more step is required to pin down the exact assignment, which not only reflects the \textit{sign} of skill-productivity complementarities but particularly their \textit{relative strength}. The key objects in the described transformation are the \textit{square roots of skill and productivity co-
variance matrices, which appear in the market clearing condition. I parameterize these square roots such that they reflect the relative strength of skill-productivity complementarities across tasks. The continuum of square roots that achieves the de-correlation of skills and productivities ranges from the spectral square root (for fully symmetric complementarities across tasks) to the Cholesky square root (for fully asymmetric complementarities). To the best of my knowledge, this paper is the first to develop this technique for computing the assignment in closed form.

Section 4 turns to the economic implications of the Gaussian copula model. The model’s flexibility and analytical tractability allow me to investigate several interesting issues, such as the impact of (a) multivariate skill and productivity distributions and (b) various types of technological change on assignment, wages and wage inequality. Throughout these comparative statics, I emphasize the novel insights from the multidimensional model that go beyond those from the one-dimensional setting.

The first comparative statics regarding the distributions highlight the important role of skill and productivity correlation for assignment and wage inequality. These correlations are given by the dependence parameters of the Gaussian skill and productivity copulas. I interpret a low skill correlation as an abundance of specialists who are good in only one task (similarly, a low productivity correlation points to an abundance of firms that are highly productive in only one task). The opposite is true for high correlation, which points to generally educated workers (or general technologies).

The main results are as follows: The discrepancy between skill and productivity correlation (i.e. between skill supply and demand) causes misalignment between workers’ skills and firms’ skill requirements. This negatively affects the economy’s average income and is conducive to wage inequality. Interestingly, an economy with the highest skill correlation does not necessarily feature the highest wage inequality, which is contrary to Roy (1951) and Ohnsorge and Trefler (2007). Rather it depends on the interplay between skill supply and demand, that is, whether the supply of generalists is met by the firms’ demand. Issues regarding the degree of specialization of the workforce and its impact on labor market outcomes cannot be properly analyzed under one-dimensional heterogeneity since such models implicitly assume that all skills are perfectly correlated.

The second exercise analyzes two types of technological change and their effects on equilibrium outcomes: First, complementarity-enhancing technological change augments the productivity in both tasks symmetrically, which is why both the assignment and relative wages remain unchanged. Second, I analyze task-biased technological change (TBTC), which is captured by a change in the relative magnitude of skill-productivity complementarities across tasks.\footnote{This type of technological change has received much attention in the empirical literature. It is often referred to as routinization. See, e.g. Autor et al. (2003), Autor et al. (2006) and Autor and Dorn (2012). To fit their analysis, the two skills here can be interpreted as routine and non-routine skills.} Suppose that TBTC favors the intellectual task. The effect on the assignment is that matching along this task approaches the perfectly assortative allocation, meaning that the matched skills and productivities perfectly overlap. In turn, matching along the manual dimension diverges from the perfect assortative allocation. Thus, more important characteristics are matched up in a more assortative way, which comes at the cost of misalignment along other characteristics. As a direct consequence of this reallocation, workers who are specialized in the manual task experience a wage decline relative to both specialists in the
intellectual task and generalists. Moreover, the impact of TBTC on relative wages is most severe when labor supply is considerably more specialized than labor demand.

The exercise on technological change contains three main messages. First, there are important interactions between technological change and the underlying skill and productivity distributions. Second, this exercise reiterates that the multidimensional allocation not only depends on the sign but especially the relative strength of complementarities across tasks. Third, the trade-off between misalignment and assortative matching across tasks is a novel channel through which TBTC increases the return of intellectual skills relative to manual skills and thus wage inequality.

In Section 5, I compare this paper to both the economics literature on multidimensional matching under transferable utility and to the optimal transport literature. Section 6 provides a conclusion.

2 The General Model

2.1 Environment and Important Concepts

Agents: There are two types of agents, firms and workers. All are risk-neutral. There is a continuum of workers. Every worker is endowed with a skill bundle \( x = (x_A, x_B) \in X \subseteq \mathbb{R}^2_+ \), which he brings to the workplace. \( x_A \) (respectively \( x_B \)) corresponds to the skill necessary to perform task \( A \) (respectively \( B \)). I interpret task \( A \) as an intellectual task and task \( B \) as a task that involves manual dexterity. Points in \( X \) represent worker types. Denote the joint c.d.f. of \( (x_A, x_B) \) by \( H(x_A, x_B) \), which is assumed to be absolutely continuous with respect to the Lebesgue measure. There also is a continuum of firms. Each firm is endowed with a productivity bundle \( y = (y_A, y_B) \in Y \subseteq \mathbb{R}^2_+ \). \( y_A \) (respectively \( y_B \)) corresponds to the productivity or complexity level in task \( A \) (respectively \( B \)). Points in \( Y \) represent firm types. Denote the joint c.d.f. of \( (y_A, y_B) \) by \( G(y_A, y_B) \), which is also assumed to be absolutely continuous. Assume that the overall masses of firms and workers coincide.\(^6\)

Production: Every firm produces a single homogeneous final good by combining all inputs. Denote the technology by \( F(x_A, x_B, y_A, y_B) \). It is assumed that \( F(.) \) is twice continuously differentiable.

Labor market: Firms and workers match in pairs. The labor market is competitive (there are no search frictions) and the environment is static.

Pure Matching: The sorting between workers and firms is described by a map \( x^* = \nu(y) \), where \( \nu(y) \) is the worker type that firm \( y \) optimally chooses to hire (** indicates an equilibrium object). The focus here is on a bijective \( C^1 \) map \( \nu : \mathbb{R}^2_+ \to \mathbb{R}^2_+ \), which can be uniquely characterized by its inverse \( \mu \equiv \nu^{-1} \). I call \( y^* = \mu(x) \) the matching function, which describes the assignment of workers to firms. Throughout, the concept of pure matching will play an important role, which is defined as follows.

Definition 1 (Pure Matching) Matching is pure if \( y^* = \mu(x) \) is one-to-one almost surely.

Pure matching is closely related to the properties of the Jacobian of the matching function and particularly to the \( P \)-matrix property of the Jacobian, which is defined as follows.

\(^6\)Otherwise, there is equilibrium unemployment or idle firms, which unnecessarily complicates the model.
**Definition 2 (P-Matrix)** The Jacobian of the matching function, given by

\[
J_\mu(x_A, x_B) \equiv D_x y^* = \begin{bmatrix}
\frac{\partial y^*_A}{\partial x_A} & \frac{\partial y^*_A}{\partial x_B} \\
\frac{\partial y^*_B}{\partial x_A} & \frac{\partial y^*_B}{\partial x_B}
\end{bmatrix}
\]

is a P-matrix if

\[
[i] \quad \frac{\partial y^*_A}{\partial x_A} > 0 \quad [ii] \quad \frac{\partial y^*_B}{\partial x_B} > 0 \quad [iii] \quad \frac{\partial y^*_A}{\partial x_A} \frac{\partial y^*_B}{\partial x_B} - \frac{\partial y^*_A}{\partial x_B} \frac{\partial y^*_B}{\partial x_A} > 0.
\]

Generally, a matrix \(M\) is a P-matrix if all its principal minors are positive. Hence, every positive definite matrix is a P-matrix but the converse statement only holds when the matrix is symmetric. Gale and Nikaido (1965) link the P-matrix property of the Jacobian of a function to the function’s injectivity, which gives a sufficient condition for purity in the current setting: If \(D_x y^*\) is a P-matrix (or \(P^-\)-matrix), then the matching function is globally one-to-one.\(^7\) The P-matrix property of the Jacobian is also sufficient for global invertibility of the matching function, justifying my approach to consider \(\nu^{-1} \equiv \mu\) as the matching function instead of \(\nu\).\(^8\)

Notice the tight connection between the concepts of purity in the one and multidimensional setting. With one-dimensional traits, the equilibrium is pure if the real-valued matching function \(y = \mu(x)\) is strictly monotonic or one-to-one, i.e. \(\mu'(x) \leq 0\). When skills and productivities vary in more than one dimension, I replace this condition by the requirement that the Jacobian of \(\mu(x)\) (which is the matrix of first derivatives) is a P or \(P^-\)-matrix, guaranteeing that \(\mu(x)\) is one-to-one.

**Assortative Matching:** The P-matrix property of the Jacobian does not only ensure a pure assignment but also allows for a sensible definition of assortative matching with multidimensional types.

**Definition 3 (Assortative Matching with Multidimensional Types)**. The sorting pattern is PAM (NAM) if \(D_x y^*\) is a P-matrix (\(P^-\)-matrix).

To illustrate most arguments, I will focus on P-matrices and PAM. In economic terms, PAM means that intellectual types work in firms where workers need to perform complex intellectual tasks (part [i] in (1)). Similarly, workers with strong manual skills work in firms that attach considerable weight to the manual task (part [ii]). Moreover, condition [iii] says that these within-task matching forces dominate the between-task matching forces. Otherwise, strong scientists would work in the best garages whereas the best technicians would work at leading universities.

Definition 3 is a natural generalization of one-dimensional assortative matching where PAM is defined as a positive first derivative of the matching function. Just like in the one-dimensional case, Definition 3 captures two aspects: the direction of sorting in each task dimension (given by expressions [i] and [ii] in (1)) and purity of the assignment (for which the determinant condition [iii] is needed). In both the one-dimensional and multidimensional setting, PAM implies purity.

---

\(^7\)A matrix \(M\) is a \(P^-\)-matrix if \(-M\) is a P-matrix, i.e. [i] and [ii] have opposite inequalities in Definition 2.

\(^8\)See Theorem 1.1 in Chua and Lam (1972) and the references therein for the equivalence of the class of globally one-to-one and continuous functions from \(\mathbb{R}^n\) into \(\mathbb{R}^n\) and the class of globally homeomorphic functions from \(\mathbb{R}^n\) to \(\mathbb{R}^n\).
The figure below provides a graphical illustration of multidimensional PAM, using a discrete 2x2 example: Each side of the market has two attributes that can be high (H) or low (L). Hence, there are four worker and four firm types. In each subfigure, the left panel are worker types and the right panel are firm types. Dots indicate types. Assume that all dots carry the same mass of agents and suppose worker and firm types of the same color match. In subfigure (a), matching is characterized by PAM (which implies purity). In subfigure (b), matching is pure (i.e. every agent matches with a single preferred type) but PAM is violated along the A dimension. In subfigure (c), matching is neither positive assortative nor pure because agents are indifferent between several matches.

2.2 The Firm’s Problem

A firm with given productivity bundle \((y_A, y_B)\) chooses a worker with skill bundle \((x_A, x_B)\) in order to maximize profits. It takes the wage schedule as given, meaning that wages are not a function of productivities. In this section, I derive the firm’s problem and optimality conditions heuristically, taking as given that the wage function (denoted by \(w(x_A, x_B)\)) is twice continuously differentiable. Below, I show conditions under which \(w(x_A, x_B)\) satisfies this property. The firm’s problem is given by

\[
\max_{(x_A, x_B) \in X} F(x_A, x_B, y_A, y_B) - w(x_A, x_B)
\]  

(2)

The FOCs of this maximization problem read

\[
F_{x_A}(x_A, x_B, y_A, y_B) - w_{x_A}(x_A, x_B) = 0
\]  

(3)

\[
F_{x_B}(x_A, x_B, y_A, y_B) - w_{x_B}(x_A, x_B) = 0
\]  

(4)

where subscripts denote derivatives. Equations (3) and (4) hold only at the equilibrium assignment.

2.3 Equilibrium

I focus on a pure equilibrium, which is defined as follows.

**Definition 4 (Pure Equilibrium)** A pure equilibrium is a \(C^1\) assignment function \((y_A^*, y_B^*) = \mu(x_A, x_B)\) where \(\mu : X \to Y\), with \(X, Y \subseteq \mathbb{R}_+^2\), and a non-negative \(C^2\) wage schedule \(w : X \to \mathbb{R}_+\), s.t.

(i) Optimality: Price-taking firms maximize profits (2) by choosing \((x_A, x_B)\) for a given \(w(x_A, x_B)\).

(ii) Market Clearing: Feasibility of \((y_A^*, y_B^*) = \mu(x_A, x_B)\) requires that when \((x_A, x_B) \sim H\) then \((y_A^*, y_B^*) \sim G\).

(iii) Purity: \((y_A^*, y_B^*) = \mu(x_A, x_B)\) is a one-to-one function.

Optimality of the firm’s choice is a standard requirement of a competitive equilibrium. Market Clearing requires that the amount of workers of type \((x_A, x_B)\) demanded across all firm types cannot exceed the measure of such workers in the population. Purity implies that there exists a well-defined matching function so that each firm type prefers a single worker type. In a continuous setting, pure matching is equivalent to a deterministic mapping from workers’ skills to firms’ productivities.
(a) Purity and PAM

(b) Purity, No PAM

(c) No Purity, No PAM
A word on the existence of the equilibrium is in order. The problem presented can be captured by the Walrasian assignment problem by Gretsky et al. (1992). They prove existence of the Walrasian equilibrium with continuous seller and buyer types under transferable utility and milder assumptions on the surplus function. Since the focus of this paper is on the characterization of the equilibrium, the reader is referred to their proof of existence (Theorem 4 of their paper). For the class of multivariate models considered in Section 3, I prove existence by construction.

2.4 The Equilibrium Assignment

This section asks how properties of the production technology relate to the properties of the equilibrium assignment \((y_A^*, y_B^*) = \mu(x_A, x_B)\), which I will explicitly denote by \(y_A^* = y_A(x_A, x_B)\) and \(y_B^* = y_B(x_A, x_B)\). This assignment is only optimal if the second-order conditions of the firm’s problem, evaluated at \((y_A^*, y_B^*)\), are satisfied (i.e. negative semi-definite Hessian). Using these necessary second-order conditions for optimality along with an argument of local deviations, I show that if

\[
D_{xy}^2 F(x, y) = \begin{bmatrix}
F_{x_A y_A} & F_{x_A y_B} \\
F_{x_B y_A} & F_{x_B y_B}
\end{bmatrix}
\]

is a strictly diagonally dominant \(P\)-matrix (\(P^-\)-matrix), then \(D_{xy}^2 F(x, y)\) is a \(P\)-matrix (\(P^-\)-matrix). Moreover, under the same technological condition, \(y^* = \mu(x)\) constitutes a global maximum:

**Proposition 1 (Assortativeness and Global Maximum)** Let \(y = (y_A, y_B) \in Y \subseteq \mathbb{R}_+^2\) and \(x = (x_A, x_B) \in X \subseteq \mathbb{R}_+^2\). If \(D_{xy}^2 F(x, y)\) is a strictly diagonally dominant \(P\)-matrix (or \(P^-\)-matrix), then the assignment \((y_A^*, y_B^*) = \mu(x_A, x_B)\) \(i\) exhibits PAM (NAM), and \(ii\) is a global maximum.

The proof is in Appendix A.1. To gain intuition on the assortativeness result, focus on PAM:

If \(i\) \(F_{x_A y_A} > 0\) \(ii\) \(F_{x_B y_B} > 0\) \(iii\) \(F_{x_A y_A} > |F_{x_A y_B}|, F_{x_B y_A} > |F_{x_B y_B}|\)

then \(a\) \(\frac{\partial y_A^*}{\partial x_A} > 0\) \(b\) \(\frac{\partial y_B^*}{\partial x_B} > 0\) \(c\) \(\frac{\partial y_A^*}{\partial x_A} \frac{\partial y_B^*}{\partial x_B} - \frac{\partial y_A^*}{\partial x_B} \frac{\partial y_B^*}{\partial x_A} > 0\).

Intuitively, if there is complementarity between skills and productivities within both the intellectual task \(A\) and the manual task \(B\) (stated by \(i\) and \(ii\)), then it is optimal that workers with strong intellectual skills work in firms that value these skills (and similarly on the manual dimension), given by \(a\) and \(b\). However in general, sorting occurs along all skill and productivity dimensions, i.e. also between tasks (e.g. between the complexity level of the manual task and the intellectual skill). The diagonal dominance of the matrix of cross partials (given by \(iii\)) restricts the strength of these

---

9. They assume upper-semi continuity of the surplus function whereas I assume continuity. However, they additionally assume compactness of \(X, Y\).

10. Closely related existence results are given in Chiappori et al. (2010) and Ekeland (2010). Their work extends Gretsky’s existence results on the *endowment* economy where every seller is endowed with a given type of good to a *production* economy where sellers can choose the type of good they want to sell.

11. A matrix \(M\) is strictly diagonally dominant if \(|m_{ii}| > \sum_{j \neq i}|m_{ij}|, i = 1, 2, \ldots, n\).
between-task sorting patterns: If within-task complementarities dominate between-task complementarities, then it is optimal that the overall strength of sorting within tasks dominates the strength of sorting between tasks. This is captured by the positive determinant of \( D_{xy}^2 \) in \([c]\), which is crucial for purity. Intuitively, diagonal dominance ensures that the scientist works at the university and not in a firm looking for an expert technician.\(^{12}\)

The stated sufficient condition for PAM is \textit{distribution free}. Notice that the \( P \)-matrix property of \( D_{xy}^2 \) alone (and in particular, its positive determinant) is \textit{not} sufficient to ensure assortative matching. In addition, the stronger notion of diagonal dominance is needed. To illustrate this, I construct an example where PAM fails despite the \( P \)-matrix property of \( D_{xy}^2 \) (see Appendix D).\(^{13}\)

This section closes with a comparison to the one-dimensional setting. With one-dimensional traits, the requirement of a negative definite Hessian collapses to the requirement on the second-order condition, given by \(-F_{xy} \frac{\partial \mu(x)}{\partial x} < 0\). If \( F_{xy} \) is positive then matching is pure and positive assortative. Purity is given by the strict monotonicity of the matching function \( \mu(x) \) and PAM by its positive slope. Similarly in this model, I impose conditions on the matrix of cross-partial \( D_{xy}^2 \) to obtain PAM. The difference is that with multiple dimensions not only the signs but also the relative magnitude of different complementarities need to be restricted in order to ensure assortative matching.

2.5 The Equilibrium Wage Function

This section derives conditions for the existence of a unique wage schedule that supports the equilibrium assignment. The equilibrium wage is the solution of a system of \textit{partial differential equations} (PDEs), which are given by the first-order conditions of the firm, (3) and (4), evaluated at the equilibrium assignment. To solve a system of PDEs, integrability conditions of the system need to be specified in order to make the system involutive (i.e. formally integrable). For the linear system of first-order PDEs given above, there is only one integrability condition. It is given by the commutativity of mixed partial derivatives and obtained by cross-differentiating (3) and (4), when evaluated at \((y_A^*, y_B^*)\):

\[
\frac{\partial^2 w(x_A, x_B)}{\partial x_A \partial x_B} = \frac{\partial^2 w(x_A, x_B)}{\partial x_B \partial x_A} \iff F_{x_A y_A} \frac{\partial y_A^*}{\partial x_B} + F_{x_A y_B} \frac{\partial y_B^*}{\partial x_B} = F_{x_B y_A} \frac{\partial y_A^*}{\partial x_A} + F_{x_B y_B} \frac{\partial y_B^*}{\partial x_A}.
\]

(5)

This condition is equivalent to the requirement that the Hessian of the firm’s problem is symmetric. The next proposition states the result on existence and uniqueness of the equilibrium wage function.

**Proposition 2 (Existence and Uniqueness of the Wage Function)** Given a continuously differentiable assignment \( y^* = \mu(x) \), condition (5) is necessary and sufficient for the existence of a unique solution to the system (3) and (4), given by \( w(x) \), such that \( w(x) = w_0 \).\(^{14}\)

\(^{12}\)Similarly, in the case of NAM, assortative matching within tasks dominates the assortativeness across tasks, only that in this case high productivity workers are matched with low productive firms.

\(^{13}\)The reader is advised to turn to this example after having read Section 3.

\(^{14}\)\(w_0\) is the reservation wage of the least productive worker \( x \), set s.t. he is indifferent between working and not working.
The proof relies on Frobenius’ Theorem, which is stated in the Appendix A.2. Integrability condition (5) has both technical and economic implications. Technically, given (5), there exists a $C^2$ wage function $w(x_A, x_B)$, justifying the differentiation-based approach above.

Condition (5) also carries an important economic message. It highlights the stronger link between technology and assignment in the multidimensional setting compared to the one-dimensional setting. It implies that the Jacobian of the matching function (and hence the assignment) does not only depend on the signs of the cross partial derivatives, $F_{x_iy_j}, i, j \in \{A, B\}$, but also on their strength. Changing the strength (but not the signs) of $F_{x_iy_j}$ will induce worker reallocation without necessarily violating PAM or NAM. Matching multidimensional types thus generates something similar to an intensive margin even though firms and workers match in pairs.

This is different from the one-dimensional setting, where there is no such integrability condition because the wage is the solution to a single ordinary differential equation. There, the assignment only depends on the sign of $F_{x_iy_i}$ but not on its strength: Supermodularity (submodularity) of the technology implies PAM (NAM). Given the direction of sorting, there exists a unique measure-preserving increasing (decreasing) map of skills to productivities, which can be pinned down by labor market clearing alone. With multiple dimensions, there is no complete order of types. Hence, there is no unique measure-preserving positive (or negative) assortative map of skills to productivities. The optimal assignment must be jointly determined by labor market clearing and the firm’s problem. This is central to the next section, where I analytically solve this model for a class of examples.

3 The Gaussian Copula Model for Multidimensional Assignment

A main goal of this paper is to derive economic implications from the multidimensional assignment model that provide new insights compared to one-dimensional matching. This section presents an important step towards this objective. It develops a technique to explicitly compute the multidimensional assignment and corresponding wage, based on Gaussian copulas. The idea is to transform skills and productivities from arbitrary marginal distributions into variables that are marginally Gaussian and then use a copula to bind them together. This approach enables me to apply a second transformation that un-correlates Gaussian random variables and transforms them into independent standard normal ones. This transformation is crucial for the tractability of the model. The occurrence when the original variables are marginally Gaussian is captured as a special case.

The advantage of copulas is their flexibility. Copulas are multivariate distributions where the margins cannot only come from different families of distributions but can also include continuous and discrete, or positive and negative variables at the same time. It is in this sense that this copula model is more flexible than the general model discussed above. The construction of copulas allows for separating considerations about the marginal distributions on the one hand and dependence of the data on the other. The Gaussian copula is particularly tractable and has been widely used for applications when the data exhibit little or no tail dependence. For an extensive discussion on

---

15(a) The only constraint on the correlation matrix is that it has to be positive definite. (b) One can specify different levels of correlation between the margins.
copulas in general and on the Gaussian copula in particular, see Cherubini et al. (2004).

3.1 The Gaussian Copula

In what follows, I formally define copulas with focus on the Gaussian copula and show why they prove useful in this model. I focus on the case of $N = 2$ although this section can be generalized to arbitrary $N$. A two-dimensional copula is a c.d.f. whose support is contained in $[0, 1]^2$ and whose one-dimensional margins are uniform distributions $U(0, 1)$. Consider a bivariate distribution function $Q(x_1, x_2)$ with univariate marginal distributions $Q_1(x_1), Q_2(x_2)$ and quantile functions $Q_1^{-1}, Q_2^{-1}$. The copula associated with $Q(x_1, x_2)$ is a distribution function $C : [0, 1]^2 \rightarrow [0, 1]$ such that

$$Q(x_1, x_2) = C(Q_1(x_1), Q_2(x_2); \theta)$$

(6)

where $\theta$ is the dependence parameter of the copula: It measures the dependence between the marginals. Equation (6) captures the relation between distribution functions and copulas. Denote $u_1 = Q_1(x_1), u_2 = Q_2(x_2)$, where $u_1, u_2$ are independent uniformly distributed on $[0, 1]$. Then, $x_1 = Q_1^{-1}(u_1), x_2 = Q_2^{-1}(u_2)$ and (6) can be expressed as

$$C(u_1, u_2; \theta) = Q(Q_1^{-1}(u_1), Q_2^{-1}(u_2)).$$

(7)

To construct the Gaussian copula, skills and productivities from arbitrary marginal distributions are converted into Gaussian variables via the inverse transform method and then bound together via (7). To illustrate this, consider the case of bivariate skills. Assume that the two skills $x_A$ and $x_B$ have marginal distributions $H_A(x_A)$ and $H_B(x_B)$, respectively. I generate standard normally distributed skills from the original skills, denoted by $\tilde{x}_A$ and $\tilde{x}_B$, where $\tilde{x}_A = \Phi^{-1}(H_A(x_A))$ and $\tilde{x}_B = \Phi^{-1}(H_B(x_B))$. The dependence of $(\tilde{x}_A, \tilde{x}_B)$ is modeled using the Gaussian copula

$$C(x_A, x_B) = \Phi_2(\Phi^{-1}(H_A(x_A)), \Phi^{-1}(H_B(x_B))).$$

Then, $(\tilde{x}_A, \tilde{x}_B)$ are standard bivariate normal with correlation $\rho_{\tilde{x}}$.

3.2 Environment

Let skills $(x_A, x_B)$ and productivities $(y_A, y_B)$ follow arbitrary absolutely continuous marginal distributions.\textsuperscript{16} Transform them into standard bivariate Gaussian variables, denoted by tilde, using the discussed copula transformation:

$$\begin{bmatrix} \tilde{x}_A \\ \tilde{x}_B \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{\tilde{x}} \\ \rho_{\tilde{x}} & 1 \end{bmatrix} \right), \quad \begin{bmatrix} \tilde{y}_A \\ \tilde{y}_B \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{\tilde{y}} \\ \rho_{\tilde{y}} & 1 \end{bmatrix} \right)$$

Denote the bivariate distribution functions of transformed skills and productivities by $\Phi_{\tilde{x}}(\tilde{x}_A, \tilde{x}_B)$ and $\Phi_{\tilde{y}}(\tilde{y}_A, \tilde{y}_B)$, respectively. Assume, $\rho_{\tilde{x}}, \rho_{\tilde{y}} \in (-1, 1)$. The production technology is specified in

\textsuperscript{16}The Gaussian copula model can also deal with discrete variables. Only redefine PAM without requiring derivatives.
terms of the transformed variables and given by:

\[ F(\tilde{x}_A, \tilde{x}_B, \tilde{y}_A, \tilde{y}_B) = \alpha \tilde{x}_A \tilde{y}_A + \beta \tilde{x}_A \tilde{y}_B + \gamma \tilde{x}_B \tilde{y}_A + \delta \tilde{x}_B \tilde{y}_B \] (8)

The model can be solved in closed form under (8). However, for the sake of illustration here I focus on:

\[ F(\tilde{x}_A, \tilde{x}_B, \tilde{y}_A, \tilde{y}_B) = \alpha \tilde{x}_A \tilde{y}_A + \delta \tilde{x}_B \tilde{y}_B \] (9)

The analysis under the bilinear technology (9) conveys the full intuition. See Appendix D for how to apply the solution technique to the general technology (8). Without loss of generality, set \( \alpha = 1 \) and \( \delta \in [0, 1] \) where \( \delta \) indicates the relative strength of skill-productivity complementarities across tasks. (9) captures that there is within-task complementarity but between-task complementarity is shut down. Based on the general model from Section 2, certain properties of the equilibrium assignment are already known at this point without having to check second-order conditions of the firm’s problem.\(^{17}\) Under (9), \( D_{xy}^2 F(\tilde{x}, \tilde{y}) \) is a strictly diagonally dominant \( P \)-matrix. Consequently, the assignment in transformed variables will be unique and characterized by purity and PAM. These properties will prove useful in the construction of the equilibrium below.

I solve this assignment problem in two steps. First, I construct the equilibrium assignment and then the wage schedule that supports it. Computing the assignment and wages in the transformed variables will prove very tractable. Going back from the transformed to the original space is straightforward. The text gives the main steps and Appendix B provides the details.

### 3.3 The Equilibrium Assignment Functions \( \tilde{y}_A^* \) and \( \tilde{y}_B^* \)

The objective is to compute the equilibrium assignment functions \( \tilde{y}_A^* = \tilde{y}_A(\tilde{x}_A, \tilde{x}_B) \) and \( \tilde{y}_B^* = \tilde{y}_B(\tilde{x}_A, \tilde{x}_B) \). They must be consistent with both labor market clearing and the firm’s optimality. As pointed out in the general model, there are many possibilities of how to match workers with firms in a positive assortative way. This is the main difficulty in solving for the multidimensional assignment. What matters for pinning it down is not only the sign but also the relative strength of skill-productivity complementarities across tasks, captured by parameter \( \delta \). The solution to this problem is based on the idea of making the two-dimensional problem similar to two separate one-dimensional problems. To this end, I apply a measure-preserving transformation that un-correlates the Gaussian variables. In particular, let \( x \) be a \( p \)-variate random vector with mean \( \mu \) and nonsingular covariance matrix \( \Sigma \). Then,

\[ z = \Sigma^{-\frac{1}{2}}(x - \mu) \] (10)

has mean \( 0 \) and covariance matrix \( I_p \). The matrix \( \Sigma^{-\frac{1}{2}} \) is the inverse of any square root of the covariance matrix, i.e. \( \Sigma^{\frac{1}{2}}(\Sigma^{\frac{1}{2}})^T = \Sigma \). There exists an infinite number of such square roots. I use this\(^{17}\)

\[ \max_{(\tilde{x}_A, \tilde{x}_B) \in \tilde{X}} \tilde{x}_A \tilde{y}_A + \delta \tilde{x}_B \tilde{y}_B - w(\tilde{x}_A, \tilde{x}_B). \]

\(^{17}\)Analogously to the general model, the firm’s problem is given by
degree of freedom to parameterize the square roots by the technology parameter \( \delta \). This uniquely pins down the assignment in form of a vector-valued matching function. Here, I sketch the main steps.

Denote by \( \Sigma \hat{x} \) (resp. \( \Sigma \hat{y} \)) the covariance matrix of skills (resp. productivities). Apply (10) to the standard bivariate normal skills and productivities

\[
\begin{align*}
\hat{z}_x &= \begin{bmatrix} \hat{z}_{\hat{A}} \\ \hat{z}_{\hat{B}} \end{bmatrix} = \Sigma^{-\frac{1}{2}}_\hat{x} \begin{bmatrix} \hat{x}_A \\ \hat{x}_B \end{bmatrix} \quad \text{and} \quad \hat{z}_y &= \begin{bmatrix} \hat{z}_{\hat{A}} \\ \hat{z}_{\hat{B}} \end{bmatrix} = \Sigma^{-\frac{1}{2}}_\hat{y} \begin{bmatrix} \hat{y}_A \\ \hat{y}_B \end{bmatrix}
\end{align*}
\] (11)

where \( \hat{z}_x \) and \( \hat{z}_y \) are the vectors of uncorrelated skills and productivities, respectively. The labor market clearing condition can now be specified in terms of the uncorrelated variables, which is consistent with labor market clearing of the original variables because the applied transformation is measure-preserving. Because of the properties of the equilibrium assignment (PAM and purity), skills are mapped to productivities in an increasing way as follows

\[
(1 - \Phi(z_{\hat{y}_A}))(1 - \Phi(z_{\hat{y}_B})) = (1 - \Phi(z_{\hat{x}_A}))(1 - \Phi(z_{\hat{x}_B}))
\] (12)

where \( \Phi(.) \) again denotes the standard normal c.d.f. The interpretation of (12) is that if the firm \((z_{\hat{y}_A}, z_{\hat{y}_B})\) matches with worker \((z_{\hat{x}_A}, z_{\hat{x}_B})\), then the mass of workers with better skills than \((z_{\hat{x}_A}, z_{\hat{x}_B})\) must be equal to the mass of firms that are more productive than \((z_{\hat{y}_A}, z_{\hat{y}_B})\) (due to PAM). Because of purity, no firm that is better than \((z_{\hat{y}_A}, z_{\hat{y}_B})\) would also be willing to match with a worker that is worse than \((z_{\hat{x}_A}, z_{\hat{x}_B})\) in one or even both skills. There are no such indifferences.

The market clearing condition (12) implicitly defines the vector-valued matching function of transformed variables, denoted by \( \hat{\mu}^x : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \). The objective is to back out two real-valued assignment functions of this vector-valued matching function. To do so, set equal the quantiles of the marginal skill and productivity distributions within the A and within the B dimension

\[
\Phi(z_{\hat{y}_i}) = \Phi(z_{\hat{x}_i}) \quad \forall \ i \in \{A, B\}
\] (13)

which gives a system of two equations. In principle, there are many possible ways to match up the marginals in (12) but due to PAM (i.e. positive diagonal elements of \( D_x \hat{y}^* \)), this is the only sensible way. System (13) can be explicitly solved for productivities \( \hat{y}_A \) and \( \hat{y}_B \) as functions of skills \( \hat{x}_A \) and \( \hat{x}_B \), which constitutes the candidate equilibrium assignment

\[
\begin{bmatrix} \hat{y}_A^* \\ \hat{y}_B^* \end{bmatrix} = \Sigma^{-\frac{1}{2}}_\hat{y} \Sigma^{-\frac{1}{2}}_\hat{x} \begin{bmatrix} \hat{x}_A \\ \hat{x}_B \end{bmatrix}
\] (14)

where \( D_x \hat{y}^* = \Sigma^{-\frac{1}{2}}_\hat{y} \Sigma^{-\frac{1}{2}}_\hat{x} \) is the Jacobian of the matching function. System (14) is the candidate for the vector-valued matching function, mapping bivariate skills into bivariate productivities. By (12), it is measure-preserving (i.e. in line with labor market clearing). For (14) to also be consistent with firms’ optimality, \( D_x \hat{y}^* \) must be a P-matrix and, moreover, must satisfy the integrability condition

\[^{18}\text{I will verify below that the market clearing in the transformed variables } (\hat{z}_x, \hat{z}_y), \text{ which is based on PAM and purity, gives rise to a measure-preserving assignment in } (\hat{x}, \hat{y}) \text{ that also admits PAM and purity.}\]
(5), which collapses under the bilinear technology to:

$$\frac{\partial \tilde{y}_A^*}{\partial \tilde{x}_B} = \delta \frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_A}. \quad (15)$$

Integrability condition (15) clearly shows that $D_A y^*$ (and hence the assignment) is not only determined by the sign of skill-productivity complementarities but also by their relative magnitude, captured by $\delta \in [0,1]$. To ensure that (15) holds, the matrix square roots $\Sigma_{\tilde{x}}^1$ and $\Sigma_{\tilde{y}}^1$ play the key role because they can incorporate features of the technology. Notice that a covariance matrix has an infinite number of square roots because it is a symmetric positive definite matrix. I compute them by applying an orthonormal transformation $R$ to a given square root $\Sigma_{\tilde{x}}^1$

$$\Sigma_{\tilde{x}}^{1/2} R = \Sigma_{\tilde{x}}^{1/2} \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \quad (16)$$

where the (orthogonal) columns of $R$ have unit length: $\sqrt{\alpha^2 + \beta^2} = 1$. Then, $\Sigma_{\tilde{x}}^{1/2} R (\Sigma_{\tilde{x}}^{1/2} R)^T = \Sigma_{\tilde{x}}^{1/2} R R^T (\Sigma_{\tilde{x}}^{1/2})^T = \Sigma_{\tilde{x}}^{1/2} (\Sigma_{\tilde{x}}^{1/2})^T = \Sigma$. It follows that $\Sigma_{\tilde{x}}^{1/2} R$ is also a square root of $\Sigma$. I show in Appendix B.2 how the rotation matrix $R$ can be parameterized by the technology parameter $\delta \in [0,1]$, such that the resulting assignment is consistent with the firm’s optimality for any level of complementarities across tasks. The next proposition states the main results on the equilibrium assignment.

**Proposition 3 (Equilibrium Assignment)**

(i) For $\delta = 1$, $\tilde{y}_A^* = \tilde{y}_A(\tilde{x}_A, \tilde{x}_B)$ and $\tilde{y}_B^* = \tilde{y}_B(\tilde{x}_A, \tilde{x}_B)$ is given by

$$\begin{bmatrix} \tilde{y}_A^* \\ \tilde{y}_B^* \end{bmatrix} = \Sigma_{\tilde{x}}^{1/2} \Sigma_{\tilde{y}}^{-1/2} D_{\tilde{x}} \begin{bmatrix} \tilde{x}_A \\ \tilde{x}_B \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \left( \frac{1 + \rho_y}{\sqrt{1 + \rho_y}} + \frac{1 - \rho_y}{\sqrt{1 - \rho_y}} \right) \\ \frac{1}{2} \left( \frac{1 + \rho_y}{\sqrt{1 + \rho_y}} - \frac{1 - \rho_y}{\sqrt{1 - \rho_y}} \right) \end{bmatrix} \frac{1}{2} \left( \frac{\sqrt{1 + \rho_y}}{\sqrt{1 - \rho_y}} - \frac{\sqrt{1 - \rho_y}}{\sqrt{1 + \rho_y}} \right) \begin{bmatrix} \tilde{x}_A \\ \tilde{x}_B \end{bmatrix} \quad (17)$$

where $\Sigma_{\tilde{x}}^{1/2}$ ($\Sigma_{\tilde{x}}^{1/2}$) is the spectral square root of the productivity (skill) covariance matrix.

(ii) For $\delta = 0$, $\tilde{y}_A = \tilde{y}_A(\tilde{x}_A, \tilde{x}_B)$ and $\tilde{y}_B = \tilde{y}_B(\tilde{x}_A, \tilde{x}_B)$ is given by

$$\begin{bmatrix} \tilde{y}_A \\ \tilde{y}_B \end{bmatrix} = L_{\tilde{y}} (L_{\tilde{x}})^{-1} D_{\tilde{x}} \begin{bmatrix} \tilde{x}_A \\ \tilde{x}_B \end{bmatrix} = \frac{1}{\rho_y - \rho_z} \sqrt{\frac{1 - \rho_y^2}{1 - \rho_z^2}} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} \tilde{x}_A \\ \tilde{x}_B \end{bmatrix} \quad (18)$$

where $L_{\tilde{y}} (L_{\tilde{x}})$ is the Cholesky square root of the productivity (skill) covariance matrix.

(iii) For $\delta \in (0, 1)$, $\tilde{y}_A^* = \tilde{y}_A(\tilde{x}_A, \tilde{x}_B)$ and $\tilde{y}_B^* = \tilde{y}_B(\tilde{x}_A, \tilde{x}_B)$ are computed from (14), using appropriate rotations of the spectral square roots of skill and productivity covariance matrices.

(iv) If $\rho_z = \rho_y$, then $\frac{\partial \tilde{y}_A}{\partial \tilde{x}_A} = \frac{\partial \tilde{y}_B}{\partial \tilde{x}_B} = 1$ and $\frac{\partial \tilde{y}_A}{\partial \tilde{x}_B} = \frac{\partial \tilde{y}_B}{\partial \tilde{x}_A} = 0$, independent of $\delta$.

See Appendix B.2 for the proof. Under the symmetric technology of part (i), the equilibrium assignment is fully symmetric across the two tasks. The spectral square root, which is the unique symmetric
positive definite square root of the covariance matrix is used to compute this assignment. Part (ii) captures the completely asymmetric case, where only task $A$ matters for production. The Cholesky square root is the unique lower triangular square root, and hence asymmetric. It delivers an asymmetric assignment, which is optimal when the technology exhibits extreme asymmetries: Matching along the $A$ dimension is perfectly assortative, which I formally define as follows.

**Definition 5 (Perfect Assortativeness)** An assignment along dimension $i \in \{A, B\}$ is perfectly assortative if $\frac{\partial \tilde{y}^*_i}{\partial \tilde{x}_i} = 1$ and $\frac{\partial \tilde{y}^*_i}{\partial \tilde{x}_j} = 0$ for $i \neq j$.

Perfect assortativeness along dimension $i$ means that the marginal skill and productivity distributions fully overlap in this dimension. The opposite of perfect assortativeness is misalignment, which I define as the dissimilarity between a worker’s skills and a firm’s productivities in a given match.\(^{19}\) Misalignment (I will sometimes refer to it as structural mismatch) can be interpreted as a discrepancy between a worker’s skills and a firm’s skill requirements. When task $A$ receives all the weight in production, matching in this dimension is perfectly assortative, whereas matching along the $B$ task is convoluted by mismatch. This is intuitive because task $B$ does not matter for output.\(^{20}\)

Part (iii) states that when there are some asymmetries in the production technology, then the assignment is in-between the two polar cases (i) and (ii). Part (iv) captures the special case, in which skill and productivity distributions coincide ($\rho_{\tilde{x}} = \rho_{\tilde{y}}$). Then, matching along both dimensions is perfectly assortative, which is independent of possible asymmetries in the technology captured by $\delta < 1$. Interestingly, this holds for any level of correlation: As long as skill and productivity distributions are the same, the resulting assignment is equivalent to the one under independence of skills and productivities. In contrast, if skill supply and demand differ ($\rho_{\tilde{x}} \neq \rho_{\tilde{y}}$ by part (iv)), then labor market clearing requires that there is some structural mismatch.

It follows from this discussion that there are two sources of structural mismatch in the economy, technology and distributions, which hints at a certain degree of indeterminacy between the two. The first source stems from asymmetries in the production technology (parts (ii) and (iii)). The second source is due to the discrepancy between skill and productivity distributions or, in other words, between skill supply and demand ($\rho_{\tilde{x}} \neq \rho_{\tilde{y}}$ by part (iv)). It arises because the frictionless labor market must clear no matter how different skill and productivity distributions are.

### 3.4 The Equilibrium Wage Function

I close the model by computing the wage function that supports the assignment found above. In the Gaussian copula model, the wage function admits a closed-form solution. I proceed by guess and verify.

\(^{19}\)The degree of misalignment can be captured by the Kullback-Leibler divergence, $D_{KL}(g,h)$, which is a measure of dissimilarity of two distributions $h$ and $g$. It is zero when the two distributions are the same and increases in their dissimilarity (see Appendix B.4). In this example, $D_{KL}(g,h) = \frac{1}{2} \left( 2 \frac{1-\rho_{\tilde{x}}^2}{1-\rho_{\tilde{y}}^2} + \log \left( \frac{1-\rho_{\tilde{y}}^2}{1-\rho_{\tilde{x}}^2} \right) - 2 \right)$, which is zero when $\rho_{\tilde{x}} = \rho_{\tilde{y}}$.

\(^{20}\)Notice that assignment (18) is consistent with the equilibrium (i.e. with market clearing and optimality of the firm’s decision) but is not unique. It is a pure assignment. However, with completely asymmetric technology, agents do not care about task $B$. Firms are indifferent between workers that are equally capable of doing task $A$ but differ in their ability to complete task $B$. This can also be seen from $D_{\tilde{x}}F$, which is not a $F$-matrix for $\delta = 0$, which is why purity is not guaranteed. Hence, there are also non-pure equilibria. Characterizing them goes beyond the scope of this paper.
Proposition 4 (Equilibrium Wage Schedule) The equilibrium wage function is given by the sum of marginal products integrated along the assignment paths

$$w(\hat{x}_A, \hat{x}_B) = \int_0^{\hat{x}_A} \frac{\partial \hat{y}^*_A}{\partial \hat{x}_A} \hat{x}_A d\hat{x}_A + \frac{1}{2} \hat{x}_B \int_0^{\hat{x}_A} \frac{\partial \hat{y}^*_A}{\partial \hat{x}_B} d\hat{x}_A + \frac{1}{2} \hat{x}_B \int_0^{\hat{x}_B} \frac{\partial \hat{y}^*_B}{\partial \hat{x}_A} d\hat{x}_B + \delta \int_0^{\hat{x}_B} \frac{\partial \hat{y}^*_B}{\partial \hat{x}_B} d\hat{x}_B$$

$$= \frac{1}{2} \frac{\partial \hat{y}^*_A}{\partial \hat{x}_A} \hat{x}_A^2 + \frac{\partial \hat{y}^*_A}{\partial \hat{x}_B} \hat{x}_A \hat{x}_B + \frac{1}{2} \frac{\partial \hat{y}^*_B}{\partial \hat{x}_B} \hat{x}_B^2 + w_0$$

(19)

where the derivatives can be computed from Proposition 3 and where $w_0$ is the constant of integration.$^{21}$

Appendix B.3 verifies that the constructed wage schedule supports the equilibrium assignment $\hat{y}^*_A$ and $\hat{y}^*_B$ from Proposition 3. It inherits its additive structure from the additive separability of the production function. Since integrability condition (15) is satisfied by construction of the assignment, it is also unique. Section 4 extensively discusses the properties of the wage function and how they depend on the equilibrium assignment. Notice that in the presented Gaussian copula model, the equilibrium (i.e. an assignment and a supporting wage schedule) has been found by construction and hence, exists.

3.5 Assignment and Wage in Terms of Original Variables

Important properties of equilibrium assignment and wages in transformed variables $(\hat{x}, \hat{y})$ also hold for the original variables $(x, y)$. Notice that the assignment in terms of original variables is given by

$$\hat{y}^*_A = \hat{y}^*_A(x_A) = J_{\tilde{\mu}_{11}} x_A + J_{\tilde{\mu}_{12}} x_B \Leftrightarrow \Phi^{-1}(G_A(y_A)) = J_{\tilde{\mu}_{11}} \Phi^{-1}(H_A(x_A)) + J_{\tilde{\mu}_{12}} \Phi^{-1}(H_B(x_B))$$

$$\hat{y}^*_B = \hat{y}^*_B(x_A) = J_{\tilde{\mu}_{21}} x_A + J_{\tilde{\mu}_{22}} x_B \Leftrightarrow \Phi^{-1}(G_B(y_B)) = J_{\tilde{\mu}_{21}} \Phi^{-1}(H_A(x_A)) + J_{\tilde{\mu}_{22}} \Phi^{-1}(H_B(x_B))$$

where $J_{\tilde{\mu}_{ij}}, i, j \in \{1, 2\}$ are the elements of the matching function’s Jacobian from Proposition 3. $(\hat{x}, \hat{y})$ were obtained by monotone transformations of $(x, y)$. Since the sign of the first partial derivative is invariant under monotone transformations of the variables, it holds that $\frac{\partial \hat{y}^*_i}{\partial x_i} > 0, i \in \{A, B\}$. Moreover, the determinant of the matching function’s Jacobian in original variables is positive if and only if the determinant of the Jacobian in transformed variables is positive since

$$Det(D_y^*) = R(\Phi, H_A, H_B, G_A, G_B)Det(D_y^*)$$

where $R(.)$ is a function that takes positive values (see Appendix B.5). Hence, matching in original variables satisfies PAM.$^{22}$ Moreover, the results of Proposition 3 qualitatively apply to $(x, y)$.

To find the wage function in terms of the original variables, $w(x_A, x_B)$, substitute $\Phi^{-1}(H_i(x_i))$, $i \in \{A, B\}$ into (19). Notice that (19) is a positive-definite quadratic form in standard normal variables for all original skills $(x_A, x_B)$. Hence, even though original skills and productivities are allowed

$^{21}$w_0 is the reservation wage of the least productive worker, $(x_A, x_B) = (0, 0)$, which makes him indifferent between working and not working. Since this worker produces nothing, he earns zero wage on the job and hence I can set $w_0 = 0$.

$^{22}$If some of the original variables were discrete, one may interpret the transformed variables as latent variables of the original variables. In that case, PAM in original variables would hold up to a non-decreasing transformation.
to be negative, wages and output are non-negative for all matches that form in equilibrium (see Appendix B.6). The intuition is that supermodularity of skills and productivities within tasks induces PAM, which is a force towards matching firms and workers of similar types.

In sum, there are two main messages from the Gaussian copula model: First, it highlights a key difference between one- and multidimensional matching. With multiple characteristics there are many ways to assortatively match workers to firms, which is why not only the signs of complementarities matter for the assignment but also their relative strength.\footnote{This is similar to Galichon and Salanié (2010) and Dupuy and Galichon (2012) who point out a trade-off between matching patterns along different characteristics. See Section 5 on the relation between my paper and the literature.}

Second, the Gaussian copula approach provides the model with much flexibility and tractability: It allows for (a) arbitrary marginal skill and productivity distributions and (b) for any correlation between skills and between productivities (e.g. positive skill and negative productivity correlation). Due to this flexibility stemming from the distributions, the assumed technology is less restrictive since there is a certain degree of indeterminacy between distributions and technology. The same assignment can be achieved by different combinations of the two. Moreover, this framework can be generalized to \( N \) heterogeneity dimensions (Appendix B.7). Lastly, the transformation used to compute the assignment is highly tractable, giving rise to an analytical solution in the transformed variables and – depending on the original distributions – also in the original variables. Most importantly, key properties of assignment and wages in transformed variables can be verified for the original variables.

This section closes with an illustrative example of the Copula approach. Consider standard normal marginal distributions of skills and productivities, \( H_i = \Phi, G_i = \Phi \), so that \( \tilde{x}_i = x_i \) and \( \tilde{y}_i = y_i, \forall i \in \{A, B\} \). Then, the assignment is linear in the original skills \((x_A, x_B)\) and productivities \((y_A, y_B)\)

\[
\begin{align*}
\tilde{y}_A &= J_{\mu_11} \tilde{x}_A + J_{\mu_12} \tilde{x}_B \quad \iff \quad y_A^* = J_{\mu_11} x_A + J_{\mu_12} x_B \\
\tilde{y}_B &= J_{\mu_21} \tilde{x}_A + J_{\mu_22} \tilde{x}_B \quad \iff \quad y_B^* = J_{\mu_21} x_A + J_{\mu_22} x_B
\end{align*}
\]

where \( J_{\mu_{ij}} = J_{\mu_{ij}}, i, j \in \{1, 2\} \) are given by Proposition 3. Hence, the example with (standard) normally distributed skills and productivities is captured as a special case of the Gaussian copula model.

### 4 Comparative Statics

The central idea of assignment models is that the allocation of workers to firms shapes wages, and hence, wage inequality. To see this assignment-wage link, recall the wage function from Proposition 4\footnote{Spelled out, this is}

\[
w(\tilde{x}) = \frac{1}{2} \tilde{x}^T J_\mu \tilde{x} \tag{20}
\]
where $J_{\tilde{\mu}} = D_{\tilde{x}} \tilde{y}^*$ is the Jacobian of the matching function. At this point, I focus on $\delta = 1$.\textsuperscript{25} From (20), the Hessian of the wage function equals the Jacobian of the matching function, $H(w(\tilde{x})) = J_{\tilde{\mu}}$, showing the tight link between sorting and wages: Since sorting is positive assortative (implying that $J_{\tilde{\mu}}$ is a symmetric $P$-matrix or positive definite) wages are convex. Convex wages mean that workers with large (absolute) amounts of skills earn over-proportionally more than workers with small (absolute) amounts of skills. However, skills are not the only force behind high earnings. Due to PAM, skill differences are magnified because skilled workers are matched to more productive firms, convexifying the wage schedule. One the other hand, if sorting was negative assortative with $J_{\tilde{\mu}}$ being a symmetric $P^-$-matrix (i.e. negative definite), the wage function would be concave.\textsuperscript{26}

A major focus of this section is on wage inequality. To understand its determinants, one needs to analyze the determinants of the assignment, namely distributions and technology. This motivates two types of comparative statics: First, how does sorting depend on the skill and productivity distributions and how does this feed into wages and inequality? Moreover, how do various types of technological change affect sorting and wage inequality? I also discuss their interaction, i.e. how the response of wages to technological change depends on skill and productivity distributions.

I address these questions in the simplest possible setting that conveys the full intuition: I assume standard normally distributed skills and productivities, which is a special case of the Gaussian copula model. (See Appendix C.1 for the extension to non-standard normality of skills and productivities). Propositions 3 and 4 apply with $\tilde{x}_i = x_i$ and $\tilde{y}_i = y_i, \forall i \in \{A, B\}$, so I will discard the tilde from now on. I begin with a symmetric technology ($\delta = 1$). Moreover, I assume $\rho_x, \rho_y \in [0, 1)$. This is not necessary but allows me to focus on predictions regarding generalists (i.e. workers with high skill correlation) versus specialists (i.e. workers with low skill correlation), which are insightful.

Notice that because of standard normality, skills and productivities are allowed to be negative. However, all matches give rise to positive output and hence, positive wages (see above). Standard normality has the advantage that both the wage schedule and distribution take on clean and tractable forms.

### 4.1 Comparative Statics I: Distributions

The distributions of skills and productivities reflect supply and demand for skill. The shortage or abundance of skills relative to their demand crucially determine the assignment of workers to firms and earnings. This section analyzes how the wage distribution depends on the underlying distributions through sorting. This will also help to understand how the impact of technological change (TC) depend on the interplay between skill supply and demand (next section). The moments of the wage distribution can be computed in closed form because the wage function is a quadratic form in standard normal variables (see (20)). I focus on mean, variance and skewness. It can be shown that

\textsuperscript{25}For $\delta \in [0, 1)$, the wage is still a quadratic form but $J_{\tilde{\mu}}$ is replaced by a slightly different (symmetric) matrix with $\frac{\partial^2 y_i}{\partial x_j \partial x_k}$ and $\delta \frac{\partial^2 y_i}{\partial x_j \partial x_k}$ on the off-diagonal.

\textsuperscript{26}Even though this is not the focus here, it is worth mentioning that in this model, wage data is sufficient to determine the direction of sorting. In several one-dimensional models, this has been refuted (see e.g. Eeckhout and Kircher (2011)).
the average wage is (see Appendix C.2 for a formal statement of this section’s results):

\[ E(w(x_A, x_B)) = \frac{1}{2} \left( \sqrt{(1 + \rho_x)(1 + \rho_y)} + \sqrt{(1 - \rho_x)(1 - \rho_y)} \right) \]  

(21)

It is maximized when skill and productivity distributions perfectly overlap \((\rho_x = \rho_y)\). The same holds for output. At that point, skill supply and demand are completely aligned, giving rise to a perfectly assortative assignment in both tasks (see Proposition 3 (iv)). In turn, the economy performs poorest on average when the misalignment between skills and skill requirements is largest \((\rho_i = 0, \rho_j \to 1 \text{ or } \rho_y = 0, \rho_x \to 1)\).

Besides studying the average income, much economic research is concerned with analyzing wage inequality often measured by the variance of the wage distribution. Larger wage dispersion is a proxy for larger wage inequality. In this model, it is given by:

\[ Var(w(x_A, x_B)) = 1 + \rho_x \rho_y \]  

(22)

The maximum variance of the wage distribution is reached when all agents have perfectly balanced skill and productivity bundles \((\rho_x, \rho_y \to 1)\). That is, there are exclusively **generalists** in the economy with identical amounts of the two skills and, moreover, technology is general in the sense that firms are equally productive in both tasks. The intuition is that there are many workers with either very low or high amounts of both skills (in absolute terms) who match with firms who require these extreme skill bundles. This leads to a dispersed set of wages, with a large mass of both low and high income earners.

An alternative measure of wage inequality is the skewness of the wage distribution, here given by:

\[ E \left( \left( \frac{w - E(w)}{\sqrt{Var(w)}} \right)^3 \right) = \frac{\sqrt{(1+\rho_y)(1+\rho_x)} + \sqrt{(1-\rho_y)(1-\rho_x)}}{1+\rho_x \rho_y} \left( \frac{69 + 99 \rho_x \rho_y - 33 \sqrt{(1-\rho_y^2)(1-\rho_x^2)}}{(1+\rho_x \rho_y)^2} \right) \]  

(23)

It is common for empirical wage distributions to be positively skewed, indicating that a large fraction of workers earns little while a small fraction earns over-proportionally much. It can be shown that in this model positive skewness is maximized when the misalignment between skill supply and demand is considerable. This happens either when firms demand **specialists** with unbalanced skill bundles \((\rho_y = 0)\) but with most workers being **generalists** \((\rho_x \to 1)\) or when the opposite is true \((\rho_y \to 1, \rho_x = 0)\). In both cases, the scarce workers with highly demanded skills are **underqualified** in their jobs and earn impressive wages.\(^{27}\) On the other hand, the large mass of workers with undesired skill bundles is **overqualified** for jobs and earns little, pushing up wage inequality (see Appendix C.2). In the first case, it is the specialists who earn the highest wage. In the second case, the high-income earners are the generalists. Similar to the variance, wage inequality measured by the skewness is minimized when firms want specialists and there are many of them.

In sum, structural mismatch between skill supply and demand negatively affects the economy’s average income and is conducive to inequality. The latter has found support in the empirical liter-

\(^{27}\) *Underqualified* means that the firm’s productivities exceed their skills; the opposite is true for **overqualified** workers.
Moreover, in my model, when both skills and productivities are strongly correlated wage inequality is considerable, as measured by the variance. Interestingly, the economy with the highest skill correlation does not necessarily have the largest wage inequality. This is in contrast to the assignment models by Roy (1951) and Ohnsorge and Trefler (2007) where a higher skill correlation unambiguously leads to more inequality. My model demonstrates that their assumption of one-dimensional firm types is not innocuous. Instead, the interplay between skill supply and demand, captured by $\rho_x$ and $\rho_y$, is crucial for understanding wage inequality.

4.2 Comparative Statics II: Technological Change

In this section, I analyze two types of technological change. First, I study how assignment and wages respond to complementarity-enhancing technological change (CETC), meaning that skill-productivity complementarities are equally increased in both tasks. Second, I analyze the effect of task-biased technological change (TBTC) on assignment and wages, where skill-productivity complementarities are increased in one task relative to the other.

There are three main messages: First, while CETC is symmetric across tasks, leaving the assignment and relative wages unchanged, TBTC asymmetrically impacts the two tasks, which changes both assignment and relative prices. The comparison of the two exercises highlights a key insight from the multidimensional model: what matters most for assignment and relative wages is the relative strength of complementarities across tasks. Second, both exercises emphasize that the impact of technological change on assignment and wages depends on the (mis)alignment of skill and productivity distributions. Third, relative to the one-dimensional assignment literature, I highlight a new channel through which TBTC affects wage inequality.

4.2.1 Complementarity-Enhancing Technological Change

Consider the production technology, $F(x_A, x_B, y_A, y_B) = Z(x_Ay_A + \delta x_By_B)$, where $Z$ is an economy-wide productivity parameter. As before, symmetric tasks are considered, i.e. $\delta = 1$. Complementarity-enhancing technological change can be thought of as an increase in the skill-productivity complementarities in both tasks, given by $F_{x_i y_i} = Z$. Consider a change from $Z$ to $Z'$ such that $Z' > Z$. I informally discuss the effects here. See Appendix C.3 for formal statements and proofs.

CETC does not affect the assignment since there is no change in the underlying distributions and task weights have been kept symmetric. Since CETC is factor neutral, relative wages remain unchanged as well. Yet the economy’s overall output increases. All workers benefit from it, pushing up the average wage. It is noteworthy that when skill supply and demand are well-aligned ($\rho_x = \rho_y$), output and wages increase most strongly. Furthermore, CETC increases wage dispersion. The reason is that skilled workers sort into the best firms (due to PAM), which interacts with the now stronger skill-productivity complementarities, leading to more dispersed wages. The increase in wage dispersion due to CETC is stronger when wage dispersion was large to start with (i.e. when $\rho_x, \rho_y \to 1$).

Slonimczyk (2011) confirms for the U.S. that misalignment of workers’ skills and firms’ skill requirements leads to wage inequality.
4.2.2 Task-Biased Technological Change

Task-biased technological change (TBTC) can be thought of as a change in the relative strength of complementarities across tasks, i.e. a change in the parameter $\delta$. I interpret this type of technological change as routinization which has received much attention in the empirical literature.\(^{29}\) According to the routinization hypothesis, workers performing routine tasks are increasingly substituted by computers and machines. This does not imply that the prevalence of those tasks in the production process has diminished over time – quite the opposite (Acemoglu and Autor (2011)).

In the presented model, this phenomenon can be captured by a decrease in the skill-productivity complementarities in the manual task. Recall the technology $F(x_A, x_B, y_A, y_B) = x_A y_A + \delta x_B y_B$ and consider a change from $\delta$ to $\delta'$ such that $\delta' < \delta$. Then, $\delta'$ is called task-biased relative to $\delta$, with the bias favoring the intellectual task $A$. Indicate variables that depend on $\delta'$ by prime.

**Proposition 5 (Task-Biased Technological Change)** Suppose $\delta' < \delta = 1$. Then:

(i) Assignment:

For $\rho_x < \rho_y$, $\frac{\partial y_B'}{\partial x_B} < \frac{\partial y_A'}{\partial x_A} < 1$ and $\frac{\partial y_B'}{\partial x_B} < \frac{\partial y_A'}{\partial x_A}$

For $\rho_x > \rho_y$, $\frac{\partial y_B'}{\partial x_B} > \frac{\partial y_A'}{\partial x_A} > 1$ and $\frac{\partial y_B'}{\partial x_B} < \frac{\partial y_A'}{\partial x_A}$

For $\rho_x = \rho_y$, $\frac{\partial y_A'}{\partial x_A} = \frac{\partial y_B'}{\partial x_B} = 1$ and $\frac{\partial y_A'}{\partial x_A} = \frac{\partial y_B'}{\partial x_B} = 0$.

(ii) Relative wages: Let a worker with $(x_A, x_B) = (|x|, 0)$ be a specialist in task $A$, with $(x_A, x_B) = (0, |x|)$ be a specialist in $B$ and with $(x_A, x_B) = (|x|, |x|)$ be a generalist where $|x| < \infty$. Then:

$\frac{w'(|x|, 0)}{w'(0, |x|)} > \frac{w(|x|, 0)}{w(0, |x|)}$ and more so when $\rho_y > \rho_x$

$\frac{w'(|x|, |x|)}{w'(0, |x|)} > \frac{w(|x|, |x|)}{w(0, |x|)}$ and more so when $\rho_y > \rho_x$.

(iii) Moments: $\forall \rho_x, \rho_y$, $E(w(x_A, x_B))$ decreases, $Var(w(x_A, x_B))$ decreases, $E \left[ \left( \frac{w-\mu_w}{\sigma_w} \right)^3 \right]$ increases.

As TBTC increases the relative level of complementarities in task $A$, matching along this dimension approaches the perfect assortative allocation $\left( \frac{\partial y_A'}{\partial x_A} \rightarrow 1, \frac{\partial y_B'}{\partial x_B} \rightarrow 0 \right)$ whereas matches in task $B$ are increasingly characterized by mismatch (part (i)). This is displayed below. Figure 1 plots $y_A$ as a function of $x_A$ and $x_B$ before (first panel) and after TBTC (second panel) to see how $y_A$ co-varies with skills in each dimension. Figure 2 plots $y_B$ as a function of $x_B$ and $x_A$ and has the same structure.

The slope of the straight lines indicates how strong the sorting forces are within tasks. On the other hand, the slope of the dotted lines is an indicator of how strong sorting forces are between tasks. Recall that, due to the bilinear technology, the within-force is desirable whereas the between-force is not. Before TBTC, tasks $A$ and $B$ receive identical weights in production ($\delta = 1$), hence, the first panels in both figures are identical. Comparing the second panel across figures shows how TBTC

\(^{29}\)See, for instance, Autor et al. (2003), Autor et al. (2006), Acemoglu and Autor (2011) and Autor and Dorn (2012).
triggers asymmetries in the assignment across tasks. While matching along the $A$ dimension becomes almost perfectly assortative with $y_A \approx x_A$ (Figure 1), matching on the $B$ dimension is characterized by significant misalignment with $y_B$ responding similarly to changes in $x_B$ and $x_A$ (Figure 2).

Figure 1: $y_A$ as a Function of $x_A$ and $x_B$ before ($\delta = 1$) and after ($\delta = 0.1$) TBTC under $\rho_y > \rho_x$

Figure 2: $y_B$ as a Function of $x_B$ and $x_A$ before ($\delta = 1$) and after ($\delta = 0.1$) TBTC under $\rho_y > \rho_x$

This trade-off between assortativeness and misalignment across tasks arises in a multidimensional environment with asymmetric technology. Matching in tasks with stronger complementarities is more assortative. This comes at the cost of mismatch along other dimensions. Notice that there is one situation in which there is no reallocation despite TBTC: This is when skill supply and demand overlap perfectly ($\rho_x = \rho_y$). Sorting in both tasks is perfectly assortative to start with, which is why the assignment of workers to firms along the A-dimension cannot further improve as $\delta$ decreases.

TBTC also triggers changes in relative wages (part(ii)). Initially, specialists in A and specialists in B earn the same. Due to TBTC, workers who are specialized in task A experience a wage increase
relative to workers specialized in task $B$. Moreover, generalists gain ground on workers specialized in task $B$. These relative price changes are not only due to changes in parameter $\delta$ (direct effect) but also due to the reallocation process of workers to firms (part (i)). Also, they particularly depend on whether new matches are characterized by a better or worse fit. Here, matches along the $A$ dimension improve their fit, while the opposite is true for matches along the $B$ dimension.

Interestingly, the wage responses to TBTC are influenced by the economy’s skill supply and demand. The relative wage changes are larger when firms demand generalists with balanced skill bundles and when, at the same time, the supply of these workers is relatively scarce ($\rho_y > \rho_x$). Intuitively, a larger degree of specialization (which is not met by the demand) makes it harder for workers to adjust to adverse technological shocks. The demand for generalists cannot be fully met by their supply, driving up the generalist wage. Moreover, specialists in task $A$ fill this gap (because they are more productive than their counterparts who are specialized in $B$) and see their relative wage increase strongly. This is depicted in Figure 3 below, where the wage function is plotted for $[0, 1] \times [0, 1]$ before TBTC ($\delta = 1$), and after ($\delta = 0.1$).

![Figure 3: Wage Schedule as a Function of $x_A$ and $x_B$ before ($\delta = 1$) and after ($\delta = 0.1$) TBTC](image)

Part (iii) analyzes the effect of TBTC on the moments of the wage distribution. TBTC decreases the average income. This is intuitive because it decreases the productive capacity of the economy, suppressing output and wages. As far as wage inequality is concerned, the variance of the wage distribution decreases in response to TBTC. To gain intuition into this result, recall that the variance has a supremum at $\rho_x = \rho_y = 1$, where the wage distribution is characterized by many low and high income earners. TBTC decreases the productivity of matches in the $B$-task not only in relative but also in absolute terms. This hits high-income earning generalists more than their low salaried counterparts, which is why the variance of the wage distribution decreases. On the other hand, the wage distribution becomes more positively skewed. The reason is that TBTC takes away some productive capacity of the agents, which is why the proportion of high-income earners shrinks whereas the mass of low-income earners increases even further.
4.3 Discussion

What can be learned from the presented comparative statics of the multidimensional model? First, I address technological change. Costinot and Vogel (2010) and Acemoglu and Autor (2011) analyze task-biased technological change in one-dimensional assignment models. In these frameworks, an adverse technology shock reduces firms’ demand for medium-skilled workers and hence decreases their relative wages. Analyzing TBTC in the presented multidimensional model generates novel predictions that are complementary to their theories: First, besides the results on aggregate wage inequality changes, this model yields detailed predictions on the winners and losers of TBTC, where I distinguish between generalists and different types of specialists. More importantly, there is an interaction between TBTC and the underlying distributions. The impact of TBTC on relative wages is most severe when labor supply is considerably more specialized than demand. Moreover, the multidimensional model identifies a new channel of how TBTC affects wage inequality. TBTC endogenously changes the allocation of workers to jobs, improving the fit of worker-firm pairs along the advantaged task relative to those pairs on the adversely affected task. This is what I call the assortativeness-mismatch trade-off across competing tasks. It is noteworthy that this reallocation takes place despite pairwise matching and the fact that the P-matrix property of $D^2_{xy}F$ (i.e. PAM) is unaffected by TBTC.

Notice that the multidimensional setting is well-suited to analyze TBTC because two different skills are involved: routine and non-routine. Contrary to the one-dimensional models, this model does not assume that routine skills are only used by medium-skilled workers. Instead, I make the more natural assumption that both types of skills are used on every job yet in different proportions.

The preceding comparative statics also highlight the role of skill and productivity distributions (parameterized by skill correlation $\rho_x$ and productivity correlation $\rho_y$) for assignment and wages and their interaction with various types of technological change. This is not the first theory to stress the importance of workers’ skill correlation (i.e. the degree of specialization) for labor market outcomes. Roy (1951) and more recently Ohnsorge and Trefler (2007) propose that a higher skill correlation leads to more wage inequality. In models by Wasmer (2003) and Wasmer (2006), a more specialized workforce is less mobile and has more difficulty adjusting to shocks. In a similar vein, Lazear (2009) predicts that wage losses upon replacement increase in the degree of specialization.

My model adds an important consideration regarding the demand side (captured by $\rho_y$), enriching the existing predictions. It shows how the interplay between skill and productivity correlation shapes wage inequality and its response to technological change. A specialized workforce as such does not necessarily imply less wage inequality or a disadvantage in the presence of technological shocks. It all depends on the (mis)alignment between skill supply and demand as well as the type of shock.

Finally, my model suggests new ways to analyze the data. Differences in the underlying skill and productivity distributions might be an alternative explanation as to why labor market outcomes differ across countries that are under the same technological regime. So far, the interaction between technological change and country-specific labor market institutions has figured prominently in explaining

---

30 Contrary to Costinot and Vogel (2010) and Acemoglu and Autor (2011), there is no intensive margin here.
31 This is similar to the Skill Weights approach by Lazear (2009).
these differences. Investigating the role of skill and productivity distributions (with focus on general versus specialized) in the divergence of labor market outcomes is a promising area for future research.  

5 The Literature

This work primarily relates to two strands of literature: multidimensional matching under transferable utility and optimal transport.

**Multidimensional Matching.** Choo and Siow (2006) propose a transferable utility model of the marriage market to estimate the marriage matching function from observed matches in the US. Their model allows for multidimensional observed and unobserved heterogeneity under the assumption that there is no interaction between unobservable characteristics of partners (separability assumption). More recently, Galichon and Salanié (2010) study optimal matching in a model with multidimensional observed and unobserved heterogeneity. Under the same separability assumption, they show that optimal matching on observable characteristics is *non-pure*.

Both studies differ from my research in terms of objective and modeling choices. Choo and Siow (2006) estimate the gains from marriage and investigate the effect of US policy changes on these gains, i.e. the focus is empirical. In turn, Galichon and Salanié (2010) develop a technique to estimate the complementarities in the surplus function from the observed matches. They pursue this objective without providing a closed form solution of their model. Conversely, my objectives are to (a) develop multidimensional notions of PAM (NAM) that provide guidance on how to solve multivariate matching problems in closed form and (b) use the analytical solution of such a model to understand its economic implications. In the above-mentioned papers, modeling devices are (un)observed heterogeneity, discrete types and extreme value distributions of the unobserved heterogeneity. I rely on observed heterogeneity, Gaussian copulas and continuous types. Notice, however, that there is an important conclusion common to both the paper by Galichon and Salanié (2010) and my own: With multidimensional matching, there is a trade-off between matching along competing characteristics which depends on the complementarity weights in the surplus function.

In related work, Dupuy and Galichon (2012) extend the set-up of Galichon and Salanié (2010) to continuous types. Using a linear-quadratic surplus function, they develop a technique to estimate the affinity matrix (i.e. the matrix of complementarity/substitutability weights in the surplus function) and indices of mutual attractiveness. Applying this model to data on Dutch marriages, they confirm that matching in this market is multidimensional (e.g. both education and personality traits matter) and, in line with my work, there are trade-offs between matching along competing characteristics.

Chiappori et al. (2012b) develop a transferable utility model of multidimensional matching and test predictions of how spouses trade off education and non-smoking. They characterize the assignment in closed form, assuming: (i) Agents have a continuous uniformly distributed trait (education) and

---

32 Wasmer (2003) and Wasmer (2006) make a related point but focus on how different country-specific labor market institutions encourage different degrees of labor specialization. Based on differences in specialization, he aims to explain US-European differences regarding productivity, unemployment and mobility.

33 I will not discuss papers with non-transferable utility because there is little relation.

34 These indices are useful to identify the dimensions on which sorting occurs.
a binary trait (smoking-status). (ii) These traits are independent. (iii) The surplus function is such that the utility loss due to smoking is proportional to the surplus generated by the spouses’ skills. While their modeling choices suit their empirical application well, my paper provides a technique to derive closed forms for more general environments (with arbitrary marginal type distributions, correlation between different traits and possibly N-dimensional heterogeneity).

McCann et al. (2012) develop a model to explain partner, educational and occupational choices when agents have two types of skills: cognitive skills $k$ and social skills $n$. This discussion focusses on their marriage market. Under the assumption that the sex ratio between man and women of each type is equal to one and the specified technology $Y^M = \sqrt{k_i k_j 2^{n_i n_j - n_i - n_j}}$, $n_i, n_j \geq 2$, the authors prove that matching between agents $i$ and $j$ is such that equal types match, i.e. $n_i = n_j$ and $k_i = k_j$. When looking at this marriage market through the lens of my framework, it can be shown that $D_{kn}^2 F(k, n)$ is a $P$-matrix but not diagonally dominant. This shows that my previously mentioned distribution-free condition is sufficient but not necessary for PAM. What is needed to ensure PAM significantly depends on the underlying distributions. The marriage market example by McCann et al. (2012) illustrates that under the assumption of completely symmetric distributions, less than diagonal dominance is needed to obtain assortativeness.

**Mathematical Literature/Optimal Transport.** In non-technical terms, the optimal transport problem is to find a measure-preserving map that carries one distribution into another at minimal cost, using linear programming. A tight link has been established between the following two formulations of the assignment problem: a hedonic pricing problem with transferable utility (like the problem in this paper) and an optimal transport problem. Shapley and Shubik (1971) show this equivalence in a discrete and Gretsky et al. (1992) in a continuous setting.

Different from Gretsky et al. (1992), in the multidimensional assignment problems of Chiappori et al. (2010) and Ekeland (2010), sellers can also choose the characteristics of the good they sell. In this setting, Chiappori et al. (2010) establish a similar equivalence, namely between the hedonic pricing, the stable matching and the optimal transport problems. Apart from providing existence and uniqueness results both Chiappori et al. (2010) and Ekeland (2010) establish purity of the assignment: Their sufficient condition for purity is the twist condition, which states that $D_x F(x, y)$ is injective with respect to $y$. Notice that the $P$-matrix property of $D_{xy}^2 F(x, y)$ from my paper is sufficient for the twist condition to hold. Since $D_{xy}^2 F(x, y)$ is the Jacobian of $D_x F(x, y)$, the $P$-matrix property is sufficient for $D_x F(x, y)$ to be injective. While this literature developed powerful general tools to study multidimensional matching problems, it provides little guidance on how to solve them explicitly. This is what my paper seeks to address.

---

35 In their paper, the marriage market is the only one, in which choices are based on two characteristics on both sides of the market and, hence, where a comparison to my multidimensional set-up is appropriate.

36 Optimal transport has a long tradition in mathematical theory. See Villani (2009) for a recent reference book.

37 Additionally, both show the equivalence to a third formulation, namely the market game. Notice that the provided examples in Gretsky et al. (1992) are restricted to one-dimensional types.

38 Notice that Chiappori et al. (2010) relies on linear programming while Ekeland (2010) applies convex analysis.
6 Concluding Remarks

In many markets, matching is based on multiple attributes. While the optimal transport literature has developed powerful tools to study the existence and uniqueness of the equilibrium in multidimensional matching problems, there has been little progress in solving and characterizing multidimensional examples, as noted by Ekeland (2010). This paper makes an attempt towards filling this gap.

I show how to analytically solve for the equilibrium assignment and wage function in the class of multivariate models that is based on Gaussian copulas, transferable utility and pairwise matching. In solving this model, I rely on (a) my multidimensional notion of assortative matching and (b) a transformation that un-correlates Gaussian variables. I then use this closed-form solution to generate novel predictions of how assignment and wages depend on multivariate skill and productivity distributions (i.e. skill supply and demand) and their interaction with technological change. This analysis highlights a key feature of multidimensional matching: There are many ways to assortatively match workers with firms, contrary to the one-dimensional setting. What matters for pinning down the assignment is not only the sign but also the strength of skill-productivity complementarities in each task.

This analysis may be extended in various ways: Empirically, one can investigate the model’s predictions on sorting and wage inequality using multidimensional data on worker skills and firm characteristics (currently being developed in a parallel research). When working with data, one has to bear in mind that matching is not always pure since observationally similar agents choose matches that are observationally different. The focus of this paper, however, is on pure (i.e. one-to-one) assignments. Developing techniques to explicitly construct non-pure assignments is an important issue and is complementary to this paper.

As far as the theory is concerned, one extension is to introduce search frictions into the environment. Some progress was achieved on developing sufficient conditions for multidimensional positive and negative assortative matching under search frictions and directed search (see Appendix E). However, the aforementioned techniques to analytically solve for the equilibrium cannot be applied to the environment with frictions and directed search. The complication stems from the queue length of workers for each job posting, which is an endogenous object. It appears in the labor market clearing condition and is endogenously correlated with firm types. But the applied transformation only helps to deal with complications stemming from the exogenous correlation of skills and productivities. This considerably complicates the construction of the vector-valued matching function and no progress has been made in this respect. A second theoretical question is how explicit assignments can be computed for a broader class of production technologies. Lastly, there is an open issue as to how far the presented results are useful for environments with non-transferable utility. These questions could be considered in future research.
A Proofs General Model (Section 2)

A.1 The Equilibrium Assignment

In order to prove Proposition 1, the following Lemma and Corollary are useful.

Lemma 1 (P-Matrix Property) Let \( y = (y_A, y_B) \in Y \subseteq \mathbb{R}_+^2 \) and \( x = (x_A, x_B) \in X \subseteq \mathbb{R}_+^2 \). If

\[
D_{xy}^2 F(x, y) = \begin{bmatrix}
F_{xAyA} & F_{xyB} \\
F_{xAyB} & F_{xyB}
\end{bmatrix}
\]

is a strictly diagonally dominant \( P \)-matrix (\( P^- \)-matrix), then \( J \hat{\mu}(x) \equiv D_{xy}^* \) is a \( P \)-matrix (\( P^- \)-matrix).

Throughout the proof I will make the following assumption:

Assumption 1 \( D_{xy}^2 F \) is a strictly diagonally dominant \( P \)-matrix.

Proof.

It will be shown that under Assumption 1, optimality of the firm’s choice requires that the Jacobian of the matching function, \( D_{xy}^* \), is a \( P \)-matrix. The proof for the case when \( D_{xy}^2 F \) is a \( P^- \)-matrix is analogous and therefore omitted. I proceed in several steps.

1. The Hessian evaluated at the equilibrium assignment, given by

\[
H^* = D_{xx}^2 F(x, y^*) - D_{xx}^2 w(x),
\]

is negative semi-definite. These are the necessary second order conditions for optimality.

2. \( Det(H^*) > 0 \). Differentiate the first order conditions, evaluated at the optimal assignment \( y^* = \mu(x) \), with respect to the skill vector \( x \), which gives

\[
H^* = D_{xx}^2 F(x, y^*) - D_{xx}^2 w(x) = -(D_{xy}^2 F(x, y^*)) (D_{xy}^*).
\]

where \( D_{xy}^* \) is the Jacobian of the matching function. Since \( D_{xy}^2 F \) is a \( P \)-matrix everywhere (and, hence, also along the equilibrium allocation \( y^* \)), it is non-singular and hence the inverse \( (D_{xy}^2 F(x, y^*))^{-1} \) exists. From (24), it is given by

\[
(D_{xy}^2 F(x, y^*))^{-1} = -(D_{xy}^*) (D_{xx}^2 F(x, y^*) - D_{xx}^2 w(x))^{-1}.
\]

It follows that \( (D_{xx}^2 F(x, y^*) - D_{xx}^2 w(x))^{-1} \) exists, and thus \( Det(H^*) \neq 0 \). Then, by Step 1, it must be \( Det(H^*) > 0 \). Hence, \(-H^*\) is a \( P \)-matrix. However, the matrix product of two \( P \)-matrices is not necessarily a \( P \)-matrix. This is why more steps are required to prove the result.

3a. If \( D_{xy}^* \) is sign-symmetric then it is a \( P \)-matrix. Suppose that \( D_{xy}^* \) is sign symmetric, i.e. \( \frac{\partial y_j^*}{\partial x_i} \frac{\partial y_i^*}{\partial x_j} > 0, i, j = A, B, i \neq j \). For sign-symmetric matrices, positivity of principal minors and stability are equivalent (see Theorem 2.6. in Hershkowitz and Keller (2005)). In the following, I
show that \(D_x y^*\) has positive eigenvalues, i.e. is stable. From (24) \(-H^* = (D^2_{xy} F(x, y^*)) (D_x y^*)\), which has all positive eigenvalues (Step 2). Denote \(M = D^2_{xy} F(x, y^*)\), \(N = D_x y^*\). Denote the eigenvalues of \(-H^*\) by \(\lambda^H\). They must obey the characteristic equation \(\text{det}(MN - \lambda^H I) = 0\). Since \(M\) is a P-matrix (Assumption 1), it is invertible and the characteristic equation can be reformulated as \(\text{det}(N - \lambda^H M^{-1}) = 0\), where \(\lambda^H\) is the generalized eigenvalue of the square matrices \((N, M^{-1})\). Given \((N, M^{-1})\), the generalized Schur decomposition factorizes both matrices \(N = Q S T'\) and \(M^{-1} = QTZ'\), where \((Q, Z)\) are orthogonal matrices and \((S, T)\) are upper triangular matrices with the eigenvalues of \((N, M^{-1})\) on their diagonals. The (real) generalized eigenvalues can be computed as \(\lambda^H = \frac{S_{ii}}{T_{ii}}\). Notice that \(T_{ii} > 0, \forall i\) because \(M\) is a diagonally dominant P-matrix, which implies stability (i.e. positive real part of eigenvalues) and \(\lambda^{M^{-1}} = \frac{1}{\lambda^H}\). For \(\lambda^H > 0\), it must be that \(S_{ii} > 0\), i.e. \(N = D_x y^*\) has positive eigenvalues. It follows that \(D_x y^*\) has positive principal minors and, thus, is a P-matrix.

3b. If \(D_x y^*\) is anti sign-symmetric then it is a P-matrix. First, notice that \(\text{det}(D_x y^*) > 0\). This follows from \(\text{det}(D^2_{xy} F) > 0\) (Assumption 1) and \(\text{det}(H^*) > 0\) (Step 2) where

\[
\text{det}(H^*) = (F_{xAyA} F_{yB} - F_{xAyB} F_{xB} yA) \left( \frac{\partial y_A^*}{\partial x_A} \frac{\partial y_B^*}{\partial x_B} - \frac{\partial y_A^*}{\partial x_B} \frac{\partial y_B^*}{\partial x_A} \right) = \text{det}(D^2_{xy} F) \text{det}(D_x y^*). \tag{26}
\]

It remains to prove that \(D_x y^*\) has positive diagonal elements. Suppose that \(D_x y^*\) is anti-sign symmetric, \(\frac{\partial y_A^*}{\partial x_i} \frac{\partial y_B^*}{\partial x_j} < 0, i, j = A, B, i \neq j\). Under this assumption and given steps 1-2, \(D_x y^*\) can have the following sign patterns:

\[
D_x y^* = \begin{bmatrix} - & - \\ + & - \end{bmatrix}, D_x y^* = \begin{bmatrix} + & - \\ + & + \end{bmatrix} \tag{27}
\]

\[
D_x y^* = \begin{bmatrix} - & - \\ + & + \end{bmatrix}, D_x y^* = \begin{bmatrix} + & + \\ - & + \end{bmatrix} \tag{28}
\]

\[
D_x y^* = \begin{bmatrix} + & + \\ + & + \end{bmatrix}, D_x y^* = \begin{bmatrix} + & + \\ + & - \end{bmatrix} \tag{29}
\]

Notice that the Hessian of the firm’s problem, evaluated at \(y^*\), is given by:

\[
H^* = \begin{bmatrix}
\left( F_{xAyA} \frac{\partial y_A}{\partial x_A} + F_{xAyB} \frac{\partial y_B}{\partial x_A} \right) & - \left( F_{xAyA} \frac{\partial y_A}{\partial x_B} + F_{xAyB} \frac{\partial y_B}{\partial x_B} \right) \\
- \left( F_{xByB} \frac{\partial y_A}{\partial x_A} + F_{xByA} \frac{\partial y_A}{\partial x_B} \right) & \left( F_{xByB} \frac{\partial y_A}{\partial x_A} + F_{xByA} \frac{\partial y_A}{\partial x_B} \right)
\end{bmatrix} \tag{30}
\]

It will be shown that those sign patterns in (27)-(29) that are not in line with the P-matrix property of \(D_x y^*\) either violate the negative definiteness of (30) or allow for a deviation that is output increasing.

First, suppose that \(F_{xAyB}, F_{xByA} \geq 0\). The sign patterns in (27) cannot be contradicted but are in line with P-maticies. The first two in (28) are contradicted by the optimality requirements \(H_{22} < 0\) and \(H_{11} < 0\), respectively. The last two sign patterns in (28) violate the symmetry requirement of

\[39\text{If N has complex eigenvalue, S is quasi-upper triangular.}\]
the Hessian $H_{12} = H_{21}$. I will now use an argument of feasible local deviations to show that the sign patterns in (29) cannot constitute an equilibrium.

Proceed by contradiction and suppose that the first sign pattern in (29) is an equilibrium. Consider two types of workers $(x_A, x_B)$ and $(x_A + \Delta_x, x_B)$, where $\Delta_x > 0$, and suppose they are matched with firm types $(y_A + \Delta_y, y_B)$ and $(y_A, y_B + \Delta_y)$, $\Delta_y > 0$, respectively (which is in line with the sign pattern under consideration). Notice that in this frictionless environment with quasi-linear utility, any equilibrium maximizes output. There is a profitable deviation to the proposed assignment if swapping the firm types is output increasing, i.e. if

$$F(x_A + \Delta_x, x_B, y_A + \Delta_y, y_B) + F(x_A, x_B, y_A, y_B + \Delta_y) > F(x_A + \Delta_x, x_B, y_A, y_B + \Delta_y) + F(x_A, x_B, y_A + \Delta_y, y_B)$$

$$\Leftrightarrow F(x_A + \Delta_x, x_B, y_A + \Delta_y, y_B) - F(x_A, x_B, y_A + \Delta_y, y_B) > F(x_A + \Delta_x, x_B, y_A, y_B) - F(x_A, x_B, y_A, y_B + \Delta_y).$$

Divide both sides by $\Delta_x > 0$ and take the limit $\Delta_x \to 0$:

$$\lim_{\Delta_x \to 0} \frac{F(x_A + \Delta_x, x_B, y_A + \Delta_y, y_B) - F(x_A, x_B, y_A + \Delta_y, y_B)}{\Delta_x} \geq \lim_{\Delta_x \to 0} \frac{F(x_A + \Delta_x, x_B, y_A, y_B + \Delta_y) - F(x_A, x_B, y_A, y_B + \Delta_y)}{\Delta_x}$$

$$\Leftrightarrow \lim_{\Delta_x \to 0} F_{x_A}(x_A, x_B, y_A + \Delta_y, y_B) - F_{x_A}(x_A, x_B, y_A, y_B) > \lim_{\Delta_x \to 0} F_{x_A}(x_A, x_B, y_A, y_B + \Delta_y) - F_{x_A}(x_A, x_B, y_A, y_B)$$

(31)

Again divide both sides by $\Delta_y > 0$ and take the limit $\Delta_y \to 0$:

$$\lim_{\Delta_y \to 0} \frac{F_{x_A}(x_A, x_B, y_A + \Delta_y, y_B) - F_{x_A}(x_A, x_B, y_A, y_B)}{\Delta_y} \geq \lim_{\Delta_y \to 0} \frac{F_{x_A}(x_A, x_B, y_A, y_B + \Delta_y) - F_{x_A}(x_A, x_B, y_A, y_B)}{\Delta_y}$$

$$\Leftrightarrow \lim_{\Delta_y \to 0} F_{x_A y_A}(x_A, x_B, y_A, y_B) > \lim_{\Delta_y \to 0} F_{x_A y_B}(x_A, x_B, y_A, y_B)$$

(32)

which is true under diagonal dominance of $D^2_{xy} F$ (Assumption 1). Hence, the first sign pattern in (29) cannot be an equilibrium. An analogous argument can be made against the second sign pattern in (29). Second, suppose that $F_{x_A y_B}, F_{x_B y_A} < 0$. In this case, the first sign pattern in (28) is contradicted by $H_{11} < 0$, the second one is contradicted by $H_{22} < 0$. The third and the fourth sign patterns in (28) are contradicted by diagonal dominance of $D^2_{xy} F$, namely by $|F_{x_B y_B}| > |F_{x_B y_A}|$ and $|F_{x_A y_A}| > |F_{x_A y_B}|$, respectively. Last, the sign patterns in (29) are contradicted by the symmetry of the Hessian $H_{12} = H_{21}$. Third, suppose that $F_{x_A y_B} > 0, F_{x_B y_A} < 0$. The first two sign patterns in (28) contradict $H_{11} < 0$ and $H_{22} > 0$, respectively. The last sign pattern in (28) contradicts $H_{22} < 0$. The remaining sign pattern in (28) and the two sign patterns in (29) are not optimal because, again, there are feasible and profitable deviations from such assignments (analogous argument as above). The argument for the case when $F_{x_A y_B} < 0, F_{x_B y_A} > 0$ is analogous and therefore omitted. ■

**Corollary 1 (Assortativeness and Local Maximum)** Let $y = (y_A, y_B) \in Y \subseteq \mathbb{R}^2_+$ and $x = (x_A, x_B) \in X \subseteq \mathbb{R}^2_+$. If $D^2_{xy} F(x, y)$ is a strictly diagonally dominant P-matrix (P(-)-matrix), then (i) the assignment $(y^*_A, y^*_B) = \mu(x_A, x_B)$ exhibits PAM (NAM) and (ii) it constitutes a local maximum.
Proof.
(i) Assortativeness: Follows from the definition of assortativeness (Definition 3) and Lemma 1.
(ii) Local Maximum: If the Jacobian of a function is a P-matrix (or a \( P^{(-)} \)-matrix), then the function is injective (one-to-one) on any rectangular region of \( \mathbb{R}^n \) (Gale and Nikaido (1965), Theorem 4). It follows from Lemma 1 and the Gale-Nikaido theorem that the solution to the firm’s problem a local maximum.

Proof of Proposition 1.
(i) Assortativeness: Follows directly from Corollary 1.
(ii) Global Maximum: It will be shown that the solution to the firm’s problem is a global maximum. I proceed by contradiction. Consider a firm \( y = (y_A, y_B) \) which optimally chooses worker \( x = (x_A, x_B) \), i.e. \( y = \mu(x) \).

Consider another firm \( y' = (y'_A, y'_B) \), \( y' \neq y \), for which worker \( x' = (x'_A, x'_B) \), \( x \neq x' \), is an optimal choice, and hence \( y' = \mu(x') \). Let \( y = \mu(x) \) and \( y' = \mu(x') \) be the local optima from Corollary 1. Now suppose that worker \( x' \) is also an optimal choice for firm \( y \), that is \( x' \) satisfies the optimality (first-order) conditions of both firms:

\[
F_x(x', y) = w_x(x') \tag{34}
\]

\[
F_x(x', y') = w_x(x'). \tag{35}
\]

I will show that, under Assumption 1, (34) and (35) cannot hold simultaneously. It suffices to show that the function \( F_x = (F_{xA}, F_{xB}) \) is one-to-one, i.e. \( F_x(x, y) = F_x(x, y') \) implies \( y = y' \). By Assumption 1, \( D^2_{xy}F(x, y') \) is a P-matrix. Moreover, \( F_x \) is defined over a rectangular region on \( \mathbb{R}^4 \).

It follows from the Gale-Nikaido Theorem (Gale and Nikaido (1965)) that \( F_x \) is injective with respect to \( y \). Thus, (34) and (35) cannot hold simultaneously because

\[
F_x(x', y) = F_x(x', y') \tag{36}
\]

only if \( y = y' \), contradicting the assumption that \( y \neq y' \). It follows that the singleton solution to the firm’s problem found in Corollary 1 is not only a local but also a global maximum.

A.2 The Wage Function

Proof of Proposition 2.
The proof is based on Frobenius Theorem. Consider a system of linear first-order partial differential equations

\[
\frac{\partial u^\rho}{\partial x^i} = \psi^\rho_i(x, u) \quad i = 1, ..., n; \rho = 1, ..., N \tag{37}
\]

where \( u : \mathbb{R}^n \rightarrow \mathbb{R}^N \). Consider the following theorem.

\[40\] More precisely, this is \( x = \nu(y) \). But recall that \( \nu^{-1} = \mu \) is the unique inverse and hence the assignment can be completely characterized by the inverse \( \mu \).
Theorem 1 (Frobenius Theorem) The necessary and sufficient conditions for the unique solution \( u^\alpha = u^\alpha(x) \) to the system (37) such that \( u(x_0) = u_0 \) to exist for any initial data \((x_0, u_0) \in \mathbb{R}^{n+N}\) is that the relations
\[
\frac{\partial \psi_i^\alpha}{\partial x^j} - \frac{\partial \psi_j^\alpha}{\partial x^i} + \sum_\beta \left( \frac{\partial \psi_i^\alpha}{\partial u^\beta} \psi_j^\beta - \frac{\partial \psi_j^\alpha}{\partial u^\beta} \psi_i^\beta \right) = 0 \quad \forall i, j = 1, ..., n, \quad \alpha, \beta = 1, ..., N. \tag{38}
\]
hold where \( \psi_i^\beta = \frac{\partial u^\beta}{\partial x^i}, \psi_j^\beta = \frac{\partial u^\beta}{\partial x^j} \).

Applying Frobenius’ Theorem to this model implies: \( u = w, \ x = (x_A, x_B) \) and \( \psi_i(x, u) = F_x(x_A, x_B, y_A(x_A, x_B), y_B(x_A, x_B)) \). Notice that \( N = 1 \) because \( w(.) \) is a real-valued function. Then, (38) reduces to
\[
\frac{\partial \psi_i}{\partial x^j} - \frac{\partial \psi_j}{\partial x^i} = 0 \tag{39}
\]
which is condition (5) from the main text since \( F_{x_A x_B} = F_{x_B x_A} \). Hence, given (5), the involutivity condition from Frobenius theorem is satisfied. A unique (local) solution to the system of linear partial differential equations (3) and (4) exists.

B Proofs of Gaussian Copula Model (Section 3)

B.1 Labor Market Clearing under PAM (or NAM)

Having applied the measure-preserving transformation (10) to skills and productivities, the labor market clearing of transformed variables under PAM reads
\[
\int_{\tilde{z}_{A}}^{\infty} \int_{\tilde{z}_{B}}^{\infty} g(\tilde{z}_{A}, \tilde{z}_{B}) d\tilde{z}_{B} d\tilde{z}_{A} = \int_{\hat{z}_{A}}^{\infty} \int_{\hat{z}_{B}}^{\infty} h(\hat{z}_{A}, \hat{z}_{B}) d\hat{z}_{B} d\hat{z}_{A} \tag{40}
\]
where \( h \) and \( g \) denote the standard normal p.d.f.’s of the uncorrelated skills and productivities, respectively. Equation (12) follows immediately, taking into account that the \( z' \)s are independent and standard normally distributed. Similarly, under NAM, the market clearing would read
\[
\int_{\tilde{z}_{A}}^{\infty} \int_{\tilde{z}_{B}}^{\infty} g(\tilde{z}_{A}, \tilde{z}_{B}) d\tilde{z}_{B} d\tilde{z}_{A} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\tilde{z}_{A}, \tilde{z}_{B}) d\tilde{z}_{B} d\tilde{z}_{A}.
\]
B.2 The Equilibrium Assignment

The following two lemmas are building blocks for the proof of Proposition 3.

Lemma 2 (Continuum of Square Roots) (i) There exists a continuum of square roots of the covariance matrix $\Sigma$, denoted by $S$. Denote its elements by $\Sigma^{1/2} \in S$, where $\Sigma^{1/2}(\Sigma^{1/2})^T = \Sigma$.

(ii) The elements of $S$ can be computed by applying an orthonormal transformation to any given square root. In particular, let $R$ be an orthogonal matrix, i.e. its columns are mutually orthogonal unit vectors. Hence, $R^{-1} = R^T$. Then, $\Sigma^{1/2}R(\Sigma^{1/2}R)^T = \Sigma^{1/2}RR^T(\Sigma^{1/2})^T = \Sigma^{1/2}(\Sigma^{1/2})^T = \Sigma$.

Proof.

(i) The existence of an infinite number of square roots of the covariance matrix follows from its symmetry. The following non-linear system

$$\Sigma^{1/2}(\Sigma^{1/2})^T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \Sigma $$

is underdetermined. Thus, it either has none or an infinite number of solutions. Since $\Sigma$ is positive-definite, one square root can be computed using the spectral square root decomposition

$$\Sigma = CDC'$$

or,

$$\Sigma^{1/2} = CD^{1/2}C'$$

where $D$ is a diagonal matrix with the eigenvalues of $\Sigma$ as diagonal entries and $C$ is a matrix of orthonormal eigenvectors of $\Sigma$. Since the spectral square root is one solution to (41), it follows that the system has an infinite number of solutions.

(ii) follows directly from orthonormality of $R$, as stated in the Lemma.

The next lemma states how the orthogonal transformation matrices $R_i, i \in \{\tilde{x}, \tilde{y}\}$ can be parameterized by $\delta$.

Lemma 3 (Orthogonal Transformation Matrices) The system of equations to be solved is given by:

$$\alpha^2 + \beta^2 = 1$$

$$\frac{\partial y^*_A}{\partial x_B} = \frac{\delta}{\partial x_A}$$

$$\frac{\partial y^*_B}{\partial x_A}$$

33
where \( \alpha_\tilde{z}, \beta_\tilde{z}, \alpha_\tilde{y}, \beta_\tilde{y} \) are the elements of the orthogonal transformation matrices:

\[
R_\tilde{z} = \begin{bmatrix} \alpha_\tilde{z} & -\beta_\tilde{z} \\ \beta_\tilde{z} & \alpha_\tilde{z} \end{bmatrix}, \quad R_\tilde{y} = \begin{bmatrix} \alpha_\tilde{y} & -\beta_\tilde{y} \\ \beta_\tilde{y} & \alpha_\tilde{y} \end{bmatrix}
\]

(i) For all \( \delta \in [0, 1] \), the solution to system (43)-(45) is given by \( \alpha_\tilde{z} = \pm 1, \beta_\tilde{z} = 0 \) and

\[
\alpha_\tilde{y} = \pm \frac{(1 + \delta) \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} + \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right)}{\sqrt{(1 - \delta)^2 \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} - \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right)^2 + (1 + \delta)^2 \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} + \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right)^2}}
\]

\[
\beta_\tilde{y} = \sqrt{1 - \alpha_\tilde{y}^2}.
\]

(ii) For \( \rho_\tilde{z} \leq \rho_\tilde{y} \), set \( \alpha_i > 0 \). For \( \rho_\tilde{z} > \rho_\tilde{y} \), set \( \alpha_i < 0 \), where \( i \in \{ \tilde{x}, \tilde{y} \} \).

**Proof.**

To solve (43)-(45), first express the off-diagonal elements of \( D_\tilde{z}\tilde{y}^* \), \( \frac{\partial \tilde{y}^*_A}{\partial x_A} \) and \( \frac{\partial \tilde{y}^*_B}{\partial x_B} \) as functions of the unknowns. To this end, I compute a candidate equilibrium assignment from (14) where I use rotations of the spectral square root (given by (42)) to uncorrelate skills and productivities. They are given by:

\[
\Sigma^\frac{1}{2}_i = \begin{bmatrix} \frac{1}{2} (\sqrt{1 + \rho_i} + \sqrt{1 - \rho_i}) & \frac{1}{2} (\sqrt{1 + \rho_i} - \sqrt{1 - \rho_i}) \\ \frac{1}{2} (\sqrt{1 + \rho_i} - \sqrt{1 - \rho_i}) & \frac{1}{2} (\sqrt{1 + \rho_i} + \sqrt{1 - \rho_i}) \end{bmatrix} \begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}, \forall \ i \in \{ \tilde{x}, \tilde{y} \}.
\]

Using (48), the candidate equilibrium assignment can be computed from (14) as:

\[
\begin{bmatrix} \tilde{y}_\tilde{x} \\ \tilde{y}_\tilde{y} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \alpha_\tilde{x} (\alpha_\tilde{x} + \beta_\tilde{x}) \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} + \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right) + (\beta_\tilde{x} \alpha_\tilde{x} - \alpha_\tilde{x} \beta_\tilde{x}) \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} - \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right) \\ \frac{1}{2} \alpha_\tilde{y} (\alpha_\tilde{y} + \beta_\tilde{y}) \left( \sqrt{\frac{1 + \rho_\tilde{x}}{1 - \rho_\tilde{x}}} + \sqrt{\frac{1 - \rho_\tilde{x}}{1 + \rho_\tilde{x}}} \right) + (\beta_\tilde{y} \alpha_\tilde{y} - \alpha_\tilde{y} \beta_\tilde{y}) \left( \sqrt{\frac{1 + \rho_\tilde{x}}{1 - \rho_\tilde{x}}} - \sqrt{\frac{1 - \rho_\tilde{x}}{1 + \rho_\tilde{x}}} \right) \end{bmatrix} \begin{bmatrix} \alpha_\tilde{x} & -\beta_\tilde{x} \\ \beta_\tilde{x} & \alpha_\tilde{x} \end{bmatrix} \begin{bmatrix} \tilde{x}_\tilde{x} \\ \tilde{x}_\tilde{y} \end{bmatrix}, \forall \ i \in \{ \tilde{x}, \tilde{y} \}.
\]

(i) The underdetermined system (43)-(45) has one degree of freedom. I exploit it by setting \( \beta_\tilde{z} = 0 \), which immediately gives \( \alpha_\tilde{z} = \pm 1 \) from equation (43). It remains to determine two unknowns, \( \alpha_\tilde{y}, \beta_\tilde{y} \), from two equations (44) and (45). From (44), \( \beta_\tilde{y} = \mp \sqrt{1 - \alpha_\tilde{y}^2} \). Using this relation along with \( \alpha_\tilde{z} = \pm 1, \beta_\tilde{z} = 0 \) and candidate assignment (49), integrability condition (45) reads:

\[
\left( \alpha_\tilde{y} \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 + \rho_\tilde{z}}} - \sqrt{\frac{1 - \rho_\tilde{y}}{1 - \rho_\tilde{z}}} \right) - \sqrt{1 - \alpha_\tilde{y}^2} \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} + \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right) \right) = \delta \left( \alpha_\tilde{y} \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 + \rho_\tilde{z}}} - \sqrt{\frac{1 - \rho_\tilde{y}}{1 - \rho_\tilde{z}}} \right) + \sqrt{1 - \alpha_\tilde{y}^2} \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} + \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right) \right)
\]

Reorganizing terms and solving for \( \alpha_\tilde{y} \) yields:

\[
\alpha_\tilde{y} = \pm \frac{(1 + \delta) \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} + \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right)}{\sqrt{(1 - \delta)^2 \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} - \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right)^2 + (1 + \delta)^2 \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{y}}} + \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{y}}} \right)^2}}
\]
Using (51), \( \beta_\tilde{y} \) can be backed out from (44)\textsuperscript{41}

\[
\beta_\tilde{y} = \sqrt{1 - \alpha_\tilde{y}^2}.
\] (52)

(ii) Rearranging (50) yields:

\[
\alpha_\tilde{y}(1 - \delta) \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 + \rho_\tilde{x}}} - \sqrt{\frac{1 - \rho_\tilde{y}}{1 - \rho_\tilde{x}}} \right) = \sqrt{1 - \alpha_\tilde{y}^2}(\delta + 1) \left( \sqrt{\frac{1 + \rho_\tilde{y}}{1 - \rho_\tilde{x}}} + \sqrt{\frac{1 - \rho_\tilde{y}}{1 + \rho_\tilde{x}}} \right)
\] (53)

While RHS \( \geq 0, \forall \rho_\tilde{x}, \rho_\tilde{y}, LHS \lesssim 0 \) for \( \rho_\tilde{y} \lesssim \rho_\tilde{x} \). It follows that \( \alpha_\tilde{y} \geq 0 \) for \( \rho_\tilde{x} \leq \rho_\tilde{y} \) and \( \alpha_\tilde{y} < 0 \) for \( \rho_\tilde{x} > \rho_\tilde{y} \).

**Proof of Proposition 3.**

*Computing the Assignment:*

(i) For \( \delta = 1 \), from (46)

\[
\alpha_\tilde{y} = \pm 1.
\] (54)

The orthogonal transformation (48) delivers:

\[
\begin{bmatrix}
\frac{1}{2} \left( \sqrt{1 + \rho_\tilde{x}} + \sqrt{1 - \rho_\tilde{x}} \right) & \frac{1}{2} \left( \sqrt{1 + \rho_\tilde{x}} - \sqrt{1 - \rho_\tilde{x}} \right) \\
\frac{1}{2} \left( \sqrt{1 + \rho_\tilde{y}} - \sqrt{1 - \rho_\tilde{y}} \right) & \frac{1}{2} \left( \sqrt{1 + \rho_\tilde{y}} + \sqrt{1 - \rho_\tilde{y}} \right)
\end{bmatrix}
\begin{bmatrix}
\pm 1 & 0 \\
0 & \pm 1
\end{bmatrix}
= \pm \begin{bmatrix}
\frac{1}{2} \left( \sqrt{1 + \rho_\tilde{x}} + \sqrt{1 - \rho_\tilde{x}} \right) & \frac{1}{2} \left( \sqrt{1 + \rho_\tilde{x}} - \sqrt{1 - \rho_\tilde{x}} \right) \\
\frac{1}{2} \left( \sqrt{1 + \rho_\tilde{y}} - \sqrt{1 - \rho_\tilde{y}} \right) & \frac{1}{2} \left( \sqrt{1 + \rho_\tilde{y}} + \sqrt{1 - \rho_\tilde{y}} \right)
\end{bmatrix}
\text{for } i \in \{ \tilde{x}, \tilde{y} \}.
\] (55)

To see that these are the *spectral square roots* of the covariance matrix (or minus one times them), I derive them below using the spectral square root decomposition, which is given by

\[
\Sigma = CDC' \quad \iff \quad \Sigma^{\frac{1}{2}} = CD^{\frac{1}{2}}C'
\] (56)

where \( D \) is a diagonal matrix with the eigenvalues of \( \Sigma \) as diagonal entries and \( C \) is a matrix of orthonormal eigenvectors of \( \Sigma \). The matrix \( \Sigma^{\frac{1}{2}} \) in (56) is called the *spectral square root* of \( \Sigma \). Notice that for \( \Sigma^{\frac{1}{2}} \) to be positive-definite, the positive square roots of the diagonal entries of \( D \) are used.

From (56) it follows that \( \Sigma^{\frac{3}{2}}_\tilde{y} \) is given by:

\[
\Sigma^{\frac{3}{2}}_\tilde{y} = \begin{bmatrix}
\sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \\
\frac{\sqrt{1 + \rho_\tilde{y}}}{\sqrt{2}} & \frac{\sqrt{1 - \rho_\tilde{y}}}{\sqrt{2}}
\end{bmatrix}
\begin{bmatrix}
\sqrt{\frac{1 + \rho_\tilde{y}}{2}} & 0 \\
0 & \sqrt{\frac{1 - \rho_\tilde{y}}{2}}
\end{bmatrix}
\begin{bmatrix}
\sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \\
\frac{\sqrt{1 + \rho_\tilde{y}}}{\sqrt{2}} & \frac{\sqrt{1 - \rho_\tilde{y}}}{\sqrt{2}}
\end{bmatrix}
= \begin{bmatrix}
\frac{1}{2} \left( \sqrt{1 + \rho_\tilde{y}} + \sqrt{1 - \rho_\tilde{y}} \right) & \frac{1}{2} \left( \sqrt{1 + \rho_\tilde{y}} - \sqrt{1 - \rho_\tilde{y}} \right) \\
\frac{1}{2} \left( \sqrt{1 + \rho_\tilde{y}} - \sqrt{1 - \rho_\tilde{y}} \right) & \frac{1}{2} \left( \sqrt{1 + \rho_\tilde{y}} + \sqrt{1 - \rho_\tilde{y}} \right)
\end{bmatrix}
\] (57)

\textsuperscript{41}Notice that \( \beta_\tilde{y} = -\sqrt{1 - \alpha_\tilde{y}^2} \) is also possible but does not affect the result, which is why I focus on the positive square root.
Moreover, since

\[ \Sigma^{-\frac{1}{2}} = CD^{-\frac{1}{2}}C = C \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} \end{bmatrix} C'. \]  

(58)

where \( \lambda_1, \lambda_2 \) are the eigenvalues of \( \Sigma \) and \( C \) is a matrix of the corresponding orthonormal eigenvectors, the matrix \( \Sigma_{\tilde{x}}^{-\frac{1}{2}} \) is given by

\[ \Sigma_{\tilde{x}}^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & \frac{1}{\sqrt{\lambda_2}} \\ \frac{1}{\sqrt{\lambda_2}} & -\frac{1}{\sqrt{\lambda_1}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & \frac{1}{\sqrt{\lambda_2}} \\ \frac{1}{\sqrt{\lambda_2}} & -\frac{1}{\sqrt{\lambda_1}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & \frac{1}{\sqrt{\lambda_2}} \\ \frac{1}{\sqrt{\lambda_2}} & -\frac{1}{\sqrt{\lambda_1}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & \frac{1}{\sqrt{\lambda_2}} \\ \frac{1}{\sqrt{\lambda_2}} & -\frac{1}{\sqrt{\lambda_1}} \end{bmatrix} \]  

(59)

It follows that the Jacobian of the matching function is given by:

\[ D_{\tilde{x}} \mathbf{y}^* = (\Sigma_{\tilde{y}}^{\frac{1}{2}} R_{\tilde{y}})(\Sigma_{\tilde{x}}^{\frac{1}{2}} R_{\tilde{x}})^{-1} = \Sigma_{\tilde{x}}^{\frac{1}{2}} \Sigma_{\tilde{x}}^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{2} \left( \frac{1}{\sqrt{\lambda_1}} + \frac{1}{\sqrt{\lambda_2}} \right) & \frac{1}{2} \left( \frac{1}{\sqrt{\lambda_1}} - \frac{1}{\sqrt{\lambda_2}} \right) \\ \frac{1}{2} \left( \frac{1}{\sqrt{\lambda_2}} + \frac{1}{\sqrt{\lambda_1}} \right) & \frac{1}{2} \left( \frac{1}{\sqrt{\lambda_2}} - \frac{1}{\sqrt{\lambda_1}} \right) \end{bmatrix} \]  

(60)

The assignment is then computed using (14).

(ii) For \( \delta = 0 \), it follows from Lemma 3 that \( R_{\tilde{y}} \) and \( R_{\tilde{x}} \) are respectively given by

\[ R_{\tilde{y}} = \begin{bmatrix} \alpha_{\tilde{y}} & \sqrt{1 - \alpha_{\tilde{y}}^2} \\ \sqrt{1 - \alpha_{\tilde{y}}^2} & \alpha_{\tilde{y}} \end{bmatrix} = \begin{bmatrix} \pm \frac{1}{2} \left( \frac{\sqrt{1 + \rho_y}(1 + \rho_x) + \sqrt{1 - \rho_y}(1 - \rho_x)}{\sqrt{1 + \rho_y}(1 + \rho_x) - \sqrt{1 - \rho_y}(1 - \rho_x)} \right) \\ \frac{1}{2} \left( \frac{\sqrt{1 + \rho_y}(1 + \rho_x) - \sqrt{1 - \rho_y}(1 - \rho_x)}{\sqrt{1 + \rho_y}(1 + \rho_x) + \sqrt{1 - \rho_y}(1 - \rho_x)} \right) \end{bmatrix} \]  

(61)

\[ R_{\tilde{x}} = \begin{bmatrix} \alpha_{\tilde{x}} & -\sqrt{1 - \alpha_{\tilde{x}}^2} \\ \sqrt{1 - \alpha_{\tilde{x}}^2} & \alpha_{\tilde{x}} \end{bmatrix} = \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}. \]  

(62)

Let \( \Sigma_{\tilde{y}}^{\frac{1}{2}} \) and \( \Sigma_{\tilde{x}}^{\frac{1}{2}} \) be the spectral square roots of skill and productivity covariance matrices, given by (57) and by the inverse of (59), respectively. Then,

\[ \Sigma_{\tilde{y}}^{\frac{1}{2}} R_{\tilde{y}} = \begin{bmatrix} \frac{1}{2} \left( \frac{\sqrt{1 + \rho_y} + \sqrt{1 - \rho_x}}{\sqrt{1 + \rho_y} + \sqrt{1 - \rho_x}} \right) \\ \frac{1}{2} \left( \frac{\sqrt{1 + \rho_y} - \sqrt{1 - \rho_x}}{\sqrt{1 + \rho_y} + \sqrt{1 - \rho_x}} \right) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \left( \frac{\sqrt{1 + \rho_y} + \sqrt{1 - \rho_x}}{\sqrt{1 + \rho_y} + \sqrt{1 - \rho_x}} \right) \\ \frac{1}{2} \left( \frac{\sqrt{1 + \rho_y} - \sqrt{1 - \rho_x}}{\sqrt{1 + \rho_y} + \sqrt{1 - \rho_x}} \right) \end{bmatrix} \]  

(63)

\[ (\Sigma_{\tilde{x}}^{\frac{1}{2}} R_{\tilde{x}})^{-1} = \begin{bmatrix} \frac{1}{2} \left( \frac{\sqrt{1 + \rho_x} + \sqrt{1 - \rho_x}}{\sqrt{1 + \rho_x} - \sqrt{1 - \rho_x}} \right) \\ \frac{1}{2} \left( \frac{\sqrt{1 + \rho_x} - \sqrt{1 - \rho_x}}{\sqrt{1 + \rho_x} - \sqrt{1 - \rho_x}} \right) \end{bmatrix}. \]
It can be shown that the Jacobian is then given by:

\[
D_{\tilde{y}^*} = \Sigma_{\tilde{y}}^2 R_{\tilde{y}} (\Sigma_{\tilde{x}}^2 R_{\tilde{x}})^{-1} = \begin{bmatrix}
1 & 0 \\
\rho_{\tilde{y}} - \rho_{\tilde{x}} \sqrt{1 - \rho_{\tilde{y}}^2} & \sqrt{1 - \rho_{\tilde{x}}^2} \\
\rho_{\tilde{x}} \sqrt{1 - \rho_{\tilde{y}}^2} & \sqrt{1 - \rho_{\tilde{x}}^2}
\end{bmatrix}
\]  \hspace{1cm} (64)

In the following, it is shown that (64) is equivalent to \( L_{\tilde{y}} (L_{\tilde{x}})^{-1} \) where \( L_i, i \in \{\tilde{x}, \tilde{y}\} \), is the Cholesky square root of skill and productivity covariance matrices, which is the unique lower triangular matrix \( L_i \) such that \( L_i(L_i)^T = \Sigma_i, i \in \{\tilde{x}, \tilde{y}\} \). By definition, \( L_i \) is a square root of \( \Sigma_i \). Under the assumption of standard normality, \( L_i \) is given by:

\[
L_i = \begin{bmatrix}
1 & 0 \\
\rho_i \sqrt{1 - \rho_i^2} & \sqrt{1 - \rho_i^2}
\end{bmatrix} \forall \ i \in \{\tilde{x}, \tilde{y}\}
\]  \hspace{1cm} (65)

Hence,

\[
L_{\tilde{y}} (L_{\tilde{x}})^{-1} = \begin{bmatrix}
1 & 0 \\
\rho_{\tilde{y}} \sqrt{1 - \rho_{\tilde{y}}^2} & \sqrt{1 - \rho_{\tilde{y}}^2}
\end{bmatrix} \begin{bmatrix}
1 & 0 \\
\rho_{\tilde{x}} \sqrt{1 - \rho_{\tilde{x}}^2} & \sqrt{1 - \rho_{\tilde{x}}^2}
\end{bmatrix}^{-1} = \begin{bmatrix}
1 & 0 \\
\rho_{\tilde{y}} - \rho_{\tilde{x}} \sqrt{1 - \rho_{\tilde{y}}^2} & \sqrt{1 - \rho_{\tilde{x}}^2} \\
\rho_{\tilde{x}} \sqrt{1 - \rho_{\tilde{y}}^2} & \sqrt{1 - \rho_{\tilde{x}}^2}
\end{bmatrix}
\]  \hspace{1cm} (66)

which coincides with (64). The equilibrium assignment is then given by (14)

\[
\begin{bmatrix}
\tilde{y}_A^* \\
\tilde{y}_B^*
\end{bmatrix} = (\Sigma_{\tilde{y}}^2 R_{\tilde{y}} (\Sigma_{\tilde{x}}^2 R_{\tilde{x}})^{-1} \begin{bmatrix}
\tilde{x}_A \\
\tilde{x}_B
\end{bmatrix}
\]

\[
= (L_{\tilde{y}})(L_{\tilde{x}})^{-1} \begin{bmatrix}
\tilde{x}_A \\
\tilde{x}_B
\end{bmatrix}
\]

\[
= \begin{bmatrix}
1 & 0 \\
\rho_{\tilde{y}} \sqrt{1 - \rho_{\tilde{y}}^2} & \sqrt{1 - \rho_{\tilde{y}}^2}
\end{bmatrix} \begin{bmatrix}
1 & 0 \\
\rho_{\tilde{x}} \sqrt{1 - \rho_{\tilde{x}}^2} & \sqrt{1 - \rho_{\tilde{x}}^2}
\end{bmatrix}^{-1} \begin{bmatrix}
\tilde{x}_A \\
\tilde{x}_B
\end{bmatrix}
\]

\[
= \begin{bmatrix}
1 & 0 \\
\rho_{\tilde{y}} - \rho_{\tilde{x}} \sqrt{1 - \rho_{\tilde{y}}^2} & \sqrt{1 - \rho_{\tilde{x}}^2}
\end{bmatrix} \begin{bmatrix}
\tilde{x}_A \\
\tilde{x}_B
\end{bmatrix}
\]  \hspace{1cm} (67)

(iii) The equilibrium assignment is obtained by using the suitable rotation matrices from Lemma 3 for each value of \( \delta \in (0, 1) \) together with the candidate equilibrium assignment (14).

(vi) For \( \rho_{\tilde{y}} = \rho_{\tilde{x}} \), (46) yields

\[
\alpha_{\tilde{y}} = \pm 1
\]  \hspace{1cm} (68)

and hence, \( \beta_{\tilde{y}} = 0 \). Substituting this (along with \( \alpha_{\tilde{x}} = \pm 1, \beta_{\tilde{x}} = 0 \)) into the candidate assignment given by (14) yields \( y_A^* = \tilde{x}_A \) and \( y_B^* = \tilde{x}_B \) and, hence, \( \frac{\partial y_A^*}{\partial x_A} = \frac{\partial y_B^*}{\partial x_B} = 1 \) and \( \frac{\partial y_A^*}{\partial x_B} = \frac{\partial y_B^*}{\partial x_A} = 0 \).
Consistency of the Assignment Functions from (i)-(iv) with the Equilibrium. Three properties have to be verified: (a) Consistency with market clearing; (b) the assignment satisfies PAM; (c) the integrability condition is satisfied. (a) Market clearing is satisfied by (12) and because the transformation (10) is measure-preserving, (b) Verifying the PAM-property amounts to checking that \( D_x \tilde{y}^* \) is a P-matrix. Using Lemmas 2 and 3 equilibrium assignment (49) can be simplified by substituting in the expressions for \( \alpha \)'s and \( \beta \)'s:

\[
\begin{bmatrix}
\tilde{y}_A^* \\
\tilde{y}_B^*
\end{bmatrix} = \begin{bmatrix}
\frac{1 + \delta}{\sqrt{1 - \rho_y^2}} \\
\frac{\rho_y^2 - \rho_x^2}{\rho_y^2 - \rho_x^2 + \sqrt{1 - \rho_y^2 \sqrt{1 - \rho_x^2}}}
\end{bmatrix} \begin{bmatrix}
\tilde{x}_A \\
\tilde{x}_B
\end{bmatrix}
\]

Taking derivatives yields:

\[
\frac{\partial \tilde{y}_A^*}{\partial \tilde{x}_A} > 0 \\
\frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_B} > 0
\]

\[
\frac{\partial \tilde{y}_A^*}{\partial \tilde{x}_A} \frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_B} - \frac{\partial \tilde{y}_A^*}{\partial \tilde{x}_B} \frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_A} = \sqrt{\frac{1 - \rho_y^2}{1 - \rho_x^2}} > 0 \text{ for } \rho_y < < 1
\]

where (70) and (71) follow immediately from (69). Hence, \( D_x \tilde{y}^* \) is a P-matrix. (c) The assignment was derived under the integrability condition (45), i.e. it is satisfied. ■

B.3 The Equilibrium Wage Function

Proof of Proposition 4.

The guess of the equilibrium wage (19), given by

\[
w(\tilde{x}_A, \tilde{x}_B) = \frac{1}{2} \left[ \tilde{x}_A + \tilde{x}_B \right] \left[ \begin{array}{c}
\delta \\
\delta
\end{array} \right] \begin{bmatrix}
\frac{\partial \tilde{y}_A^*}{\partial \tilde{x}_A} \\
\frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_B}
\end{bmatrix}
\]

needs to be verified. Notice that the constant of integration is set to zero, \( w_0 = 0 \), making the least productive worker, \((\tilde{x}_A, \tilde{x}_B) = (0, 0)\), indifferent between working and not working. Given (73), the partial derivatives of the wage with respect to skills \( \tilde{x}_A, \tilde{x}_B \) are given by

\[
\frac{\partial w(\tilde{x}_A, \tilde{x}_B)}{\partial \tilde{x}_A} = \frac{\partial \tilde{y}_A^*}{\partial \tilde{x}_A} \tilde{x}_A + \frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_A} \tilde{x}_B
\]

\[
\frac{\partial w(\tilde{x}_A, \tilde{x}_B)}{\partial \tilde{x}_B} = \delta \left( \frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_B} \tilde{x}_B + \frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_A} \tilde{x}_A \right)
\]
which coincide with the first-order conditions of the firm,

\[
\frac{\partial w(\tilde{x}_A, \tilde{x}_B)}{\partial \tilde{x}_A} = \tilde{y}_A^* \tag{76}
\]

\[
\frac{\partial w(\tilde{x}_A, \tilde{x}_B)}{\partial \tilde{x}_A} = \delta \tilde{y}_B^* \tag{77}
\]

evaluated at the equilibrium assignment (69). Moreover, the integrability condition

\[
\frac{\partial^2 w(\tilde{x}_A, \tilde{x}_B)}{\partial \tilde{x}_A \partial \tilde{x}_B} = \frac{\partial^2 w(\tilde{x}_A, \tilde{x}_B)}{\partial \tilde{x}_B \partial \tilde{x}_A} \Leftrightarrow \frac{\partial \tilde{y}_A^*}{\partial \tilde{x}_B} = \delta \frac{\partial \tilde{y}_B^*}{\partial \tilde{x}_A} \tag{78}
\]
is satisfied by construction of the equilibrium assignment, which gives uniqueness of (73) by Proposition 2.

\[\square\]

B.4 Kullback-Leibler Divergence

**Definition 6 (Kullback-Leibler Divergence)** The Kullback-Leibler divergence between probability distributions \(p_1\) and \(p_2\) is defined as

\[
D_{KL}(p_1, p_2) = \int p_1(x) \log \left( \frac{p_1(x)}{p_2(x)} \right) dx.
\]

Let \(h\) be the p.d.f. of skill distribution and \(g\) be the p.d.f. of the productivity distribution defined as above. It can be shown that the KL-divergence for two normal distributions is given by

\[
D_{KL}(g, h) = \frac{1}{2} \left( \text{tr}(\Sigma_x^{-1} \Sigma_y) + (\mu_x - \mu_y)^T \Sigma_x^{-1} (\mu_x - \mu_y) + \log \left( \frac{\det(\Sigma_x)}{\det(\Sigma_y)} \right) - n \right). \tag{79}
\]

where \(n = 2\) because bivariate Gaussian variables are considered. With bivariate standard normal skill and productivity distributions, (79) becomes

\[
D_{KL}(g, h) = \frac{1}{2} \left( 2 \frac{1 - \rho_x \rho_y}{1 - \rho_x^2} + \log \left( \frac{1 - \rho_x^2}{1 - \rho_y^2} \right) - 2 \right)
\]

which is zero for \(\rho_x = \rho_y\).

B.5 PAM in Original Variables

**Lemma 4 (PAM in Original Variables)** \((y_A^*, y_B^*) = \mu(x_A, x_B)\) satisfies PAM.
Proof. \((y_A^*, y_B^*) = \mu(x_A, x_B)\) is implicitly given by

\[
\Phi^{-1}(G_A(y_A^*)) = J_{\mu_1} \Phi^{-1}(H_A(x_A)) + J_{\mu_2} \Phi^{-1}(H_B(x_B)) \tag{80}
\]

\[
\Phi^{-1}(G_B(y_B^*)) = J_{\mu_1} \Phi^{-1}(H_A(x_A)) + J_{\mu_2} \Phi^{-1}(H_B(x_B)) \tag{81}
\]

where \(J_{\mu_{ij}}, i, j \in \{1, 2\}\) are the elements of the Jacobian of the matching function in transformed variables (Proposition 3). It has to be shown that \(J_\mu \equiv D_\mu y^*\) is a \(P\)-matrix. \((\tilde{x}, \tilde{y})\) are obtained by monotone transformations of \((x, y)\). Since the sign of the first partial derivative is invariant under monotone transformations of the variables and \(\frac{\partial y_i^*}{\partial x_i} > 0, i \in \{A, B\}\), it holds that \(\frac{\partial y_i^*}{\partial x_i} > 0, \forall i \in \{A, B\}\). It remains to show that \(\text{Det}(D_\mu y^*) > 0\). Implicit differentiate (80) and (81) with respect to \(x_A\) and \(x_B\), which gives four derivatives. For instance:

\[
\frac{\partial y_A^*}{\partial x_A} = \frac{\partial \Phi^{-1} \partial H_A}{\partial H_A \partial x_A} J_{\mu_1}
\]

Then compute \(\text{Det}(D_\mu y^*)\) as:

\[
\text{Det}(D_\mu y^*) = \frac{\partial y_A^*}{\partial x_A} \frac{\partial y_B^*}{\partial x_B} - \frac{\partial y_A^*}{\partial x_A} \frac{\partial y_B^*}{\partial x_B} = \frac{\partial \Phi^{-1} \partial H_A}{\partial H_A \partial x_A} \frac{\partial \Phi^{-1} \partial H_B}{\partial H_B \partial x_B} (J_{\mu_1} J_{\mu_2} - J_{\mu_1} J_{\mu_2})
\]

where \(R(\Phi, H_A, H_B, G_A, G_B)\) takes only positive values because it involves derivatives of strictly increasing c.d.f.'s and where \(\text{Det}(D_\mu \tilde{y}^*)\) is positive from the proof of Proposition 3. Hence, \(\text{Det}(D_\mu y^*) > 0\) if and only if \(\text{Det}(D_\mu \tilde{y}^*) > 0\).

**B.6 Non-Negative Output and Wages**

**Lemma 5 (Non-Negative Output and Wages)** Both equilibrium output \(F(x_A, x_B, y_A^*, y_B^*)\) and equilibrium wages \(w(x_A, x_B)\) are non-negative for all \(x_A, x_B \in X\).

**Proof of Lemma 5.**

Both equilibrium output \(F(\tilde{x}_A, \tilde{x}_B, y_A^*, y_B^*) = \tilde{x}_A \tilde{y}_A^* + \tilde{x}_B \tilde{y}_B^* = \frac{\partial y_A^*}{\partial x_A} \tilde{x}_A^2 + 2 \frac{\partial y_A^*}{\partial x_B} \tilde{x}_B \tilde{x}_A + \delta \frac{\partial y_B^*}{\partial x_B} \tilde{x}_B^2 = \tilde{x}^T J_\mu \tilde{x} > 0\) and equilibrium wages \(w(\tilde{x}_A, \tilde{x}_B) = \frac{1}{2} \frac{\partial y_A^*}{\partial x_A} \tilde{x}_A^2 + \frac{\partial y_A^*}{\partial x_B} \tilde{x}_B \tilde{x}_A + \frac{1}{2} \delta \frac{\partial y_B^*}{\partial x_B} \tilde{x}_B^2 = \frac{1}{2} \tilde{x}^T J_\mu \tilde{x} > 0\) are positive-definite quadratic forms in standard normal variables and hence positive for all non-zero column vectors \(\tilde{x} = (\tilde{x}_A, \tilde{x}_B)\). Notice that I used \(\frac{\partial y_A^*}{\partial x_B} = \frac{\partial y_B^*}{\partial x_A}\) (by (15)). Hence, for all \((x_A, x_B)\) such that \(\tilde{x}_A = \Phi^{-1}(H_A(x_A)) \neq 0\) or \(\tilde{x}_B = \Phi^{-1}(H_B(x_B)) \neq 0\), \(F(x_A, x_B, y_A^*, y_B^*)\) and \(w(x_A, x_B)\) are positive.
B.7 Assortative Matching with N-Dimensional Heterogeneity

**Proposition 6 (Pure and Assortative Equilibrium with N Dimensional Types)** Let $y = (y_1, ..., y_N) \in Y \subseteq \mathbb{R}^N$ and $x = (x_1, ..., x_N) \in X \subseteq \mathbb{R}^N$. If $D^2_{xy} F(x, y)$ is a diagonal $P$-matrix ($P^-$-matrix), then

(i) $J_\mu(x)$ is a $P$-matrix ($P^-$-matrix), i.e. the equilibrium assignment $\mu(x) = y^*$ is characterized by PAM (NAM), (ii) and $\mu(x) = y^*$ is globally unique.

**Proof.**

I focus on the case where $D^2_{xy} F(x, y)$ is a diagonal $P$-matrix. First notice that under this assumption, $D_{xy} y^*$ is sign-symmetric. To see this, notice that by the symmetry of the Hessian and $F_{xiyj} = 0, i, j \in \{1, 2, ..., N\}, i \neq j$,

$$H^*_i = H^*_j \Leftrightarrow F_{xiy_i} \frac{\partial y^*_i}{\partial x_j} = F_{x_jy_i} \frac{\partial y^*_j}{\partial x_i} \forall i, j \in \{1, 2, ..., N\}, i \neq j,$$

and hence $D_{xy} y^*$ is sign-symmetric, i.e. $\frac{\partial y^*_i}{\partial x_j} \frac{\partial y^*_j}{\partial x_i} > 0 \forall i, j \in \{1, 2, ..., N\}, i \neq j$. Moreover, $D_{xy} y^*$ is stable (see Step 3a in proof of Lemma 1). A sign-symmetric and stable matrix is a $P$-matrix (Theorem 2.6. in Hershkowitz and Keller (2005)), which proves (i). The proof for part (ii) works analogous to part (ii) in the proof of Proposition 1. 

Extending technology (9) to $N$ dimensions satisfies the assumption of Proposition 6. The techniques for solving the Gaussian copula model do not hinge on two-dimensional heterogeneity and, thus, can be applied to $N$ dimensions as well.

C Proofs of Comparative Statics (Section 4)

C.1 Non-Standard Normally Distributed Skills and Productivities

The continuum of square roots used to un-correlate skills and productivities can not only deal with asymmetries in the technology (i.e. $\delta < 1$) but also with asymmetries in the distributions (i.e. different means and variances of skill and productivity distributions). This section solves the model in closed form for non-standard normally distributed skills and productivities.

**Proposition 7 (Equilibrium Assignment and Wages under Normality)** Denote the non-standard normally distributed skills by $(\hat{x}_A, \hat{x}_B)$ and the productivities by $(\hat{y}_A, \hat{y}_B)$, where:

$$\begin{bmatrix} \hat{x}_A \\ \hat{x}_B \end{bmatrix} \sim N \left( \begin{bmatrix} \mu_{\hat{x}_A} \\ \mu_{\hat{x}_B} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\hat{x}_A} & \sigma_{\hat{x}_A \hat{x}_B} \rho_{\hat{x}} \\ \sigma_{\hat{x}_A \hat{x}_B} \rho_{\hat{x}} & \sigma^2_{\hat{x}_B} \end{bmatrix} \right), \quad \begin{bmatrix} \hat{y}_A \\ \hat{y}_B \end{bmatrix} \sim N \left( \begin{bmatrix} \mu_{\hat{y}_A} \\ \mu_{\hat{y}_B} \end{bmatrix}, \begin{bmatrix} \sigma^2_{\hat{y}_A} & \sigma_{\hat{y}_A \hat{y}_B} \rho_{\hat{y}} \\ \sigma_{\hat{y}_A \hat{y}_B} \rho_{\hat{y}} & \sigma^2_{\hat{y}_B} \end{bmatrix} \right)$$

(i) Assignment: Parts (i)-(iii) of Proposition 3 apply with minor modifications (see Proof). For part (vi) of Proposition 3 to hold, set $\sigma_{\hat{x}_i} = \sigma_{\hat{y}_i}, \mu_{\hat{x}_i} = \mu_{\hat{y}_i}, \forall i \in \{A, B\}$ in addition to $\rho_{\hat{y}} = \rho_{\hat{x}}$. 

41
(ii) Wage Schedule: If skills and productivities are normally distributed, the wage is given by

$$w(\hat{x}_A, \hat{x}_B) = \frac{1}{2} \frac{\partial y^*_A}{\partial x_A} s^2 + \frac{\partial y^*_A}{\partial x_A} s^2 \hat{x}_A \hat{x}_B + \frac{1}{2} \frac{\partial y^*_B}{\partial x_B} s^2 \hat{x}_B \hat{x}_B$$

$$+ \left( \mu_{g_A} - \frac{\sigma_{g_A}}{\sigma_{x_A}} \frac{\partial y^*_A}{\partial x_A} \frac{\partial y^*_A}{\partial x_B} \right) \hat{x}_A + \delta \left( \mu_{g_B} - \frac{\sigma_{g_B}}{\sigma_{x_B}} \frac{\partial y^*_B}{\partial x_A} \frac{\partial y^*_B}{\partial x_B} \right) \hat{x}_B + w_0$$

where $\frac{\partial y^*_i}{\partial x_i}, i, j \in \{A, B\}$ are the elements of the Jacobian of the matching function under standard normality (see Proposition 3).

Proof. .

(i) Recall the equilibrium assignment under standard normality:

$$\begin{bmatrix} \tilde{y}^*_A \\ \tilde{y}^*_B \end{bmatrix} = \begin{bmatrix} \frac{1 + \sqrt{1 - \rho^2}}{\sqrt{1 - \delta^2}} \\ \frac{1 + 2 \delta (\rho \rho_{y_B} + \sqrt{1 - \rho^2})^2}{} \end{bmatrix} \begin{bmatrix} \sqrt{1 + 2 \delta (\rho \rho_{y_B} + \sqrt{1 - \rho^2})^2} \\ \sqrt{1 + 2 \delta (\rho \rho_{y_B} + \sqrt{1 - \rho^2})^2} \end{bmatrix} \begin{bmatrix} \hat{x}_A \\ \hat{x}_B \end{bmatrix}$$

Denote $J_\mu(\hat{x}_A, \hat{x}_B) = D_y \tilde{y}^*$ the Jacobian of the matching function with standardized variables, where:

$$J_\mu(\hat{x}_A, \hat{x}_B) = \begin{bmatrix} \frac{\partial y^*_A}{\partial x_A} & \frac{\partial y^*_A}{\partial x_B} \\ \frac{\partial y^*_B}{\partial x_A} & \frac{\partial y^*_B}{\partial x_B} \end{bmatrix}$$

Non-standardized skills, denoted by $\hat{x}_i$, and productivities $y_i$ need to be standardized in order for (84) to hold: $\tilde{x}_i = \frac{\hat{x}_i - \mu_{\hat{x}_i}}{\sigma_{\hat{x}_i}}$ and $\tilde{y}_i = \frac{y_i - \mu_{y_i}}{\sigma_{y_i}}$. This transformation is reversed below to express the assignment in terms of the non-standardized variables $\hat{x}_i, \tilde{y}_i$. The equilibrium assignment with normally distributed skills and productivities is given by

$$\Leftrightarrow \begin{bmatrix} \tilde{y}^*_A \\ \tilde{y}^*_B \end{bmatrix} = J_\mu \begin{bmatrix} \hat{x}_A \\ \hat{x}_B \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} \frac{\tilde{y}^*_A - \mu_{\tilde{y}_A}}{\sigma_{y_B}} \\ \frac{\tilde{y}^*_B - \mu_{\tilde{y}_B}}{\sigma_{y_B}} \end{bmatrix} = J_\mu \begin{bmatrix} \frac{\hat{x}_A - \mu_{x_A}}{\sigma_{x_A}} \\ \frac{\hat{x}_B - \mu_{x_B}}{\sigma_{x_B}} \end{bmatrix}$$

where $J_\mu$ is given in (85). (86) can be solved explicitly for $(\tilde{y}^*_A, \tilde{y}^*_B)$ as a function of $(\hat{x}_A, \hat{x}_B)$. Parts (i)-(iii) of Proposition 3 readily apply. Moreover, when setting $\rho_{\hat{x}} = \rho_{\tilde{y}}, \sigma_{\hat{x}} = \sigma_{\tilde{y}}, \mu_{\hat{x}} = \mu_{y}, \forall i \in \{A, B\}$ in (86), one obtains, $\tilde{y}^*_A = \hat{x}_A, \tilde{y}^*_B = \hat{x}_B$ (part (iv) of Proposition 3 applies).

(ii) The wage guess needs to be verified. Denote the parameter in (84) by $\hat{\delta}$ instead of $\delta$ and set

$$\hat{\delta} = \delta \frac{\sigma_{g_B}}{\sigma_{x_B}}$$

(87)
where $\delta \in [0,1]$ is again the relative task weight in the production function. Given (83), one obtains:

$$\frac{\partial w(\hat{x}_A, \hat{x}_B)}{\partial \hat{x}_A} = \frac{\partial \delta y_A^s}{\partial \hat{x}_A} \sigma_{\hat{x}_A} (\hat{x}_A - \mu_{\hat{x}_A}) + \frac{\partial \delta y_A^B}{\partial \hat{x}_A} \sigma_{\hat{x}_A} (\hat{x}_B - \mu_{\hat{x}_B}) + \mu_{\delta y_A}$$

$$\frac{\partial w(\hat{x}_A, \hat{x}_B)}{\partial \hat{x}_B} = \delta \left( \frac{\partial y_A^B}{\partial \hat{x}_A} \sigma_{\hat{x}_A} (\hat{x}_A - \mu_{\hat{x}_A}) + \frac{\partial y_A^B}{\partial \hat{x}_B} \sigma_{\hat{x}_B} (\hat{x}_B - \mu_{\hat{x}_B}) + \mu_{\delta y_B} \right).$$

which coincide with the first-order conditions of the firm, (76) and (77), evaluated at the equilibrium assignment (86). To derive (88) and (89), I made use of the integrability condition

$$\frac{\partial^2 w(\hat{x}_A, \hat{x}_B)}{\partial \hat{x}_A \partial \hat{x}_B} = \frac{\partial^2 w(\hat{x}_A, \hat{x}_B)}{\partial \hat{x}_B \partial \hat{x}_A}$$

$$\Leftrightarrow \frac{\partial \delta y_A^s}{\partial \hat{x}_A} \sigma_{\hat{x}_A} = \delta \frac{\partial \delta y_B^s}{\partial \hat{x}_B} \sigma_{\hat{x}_A}$$

$$\Leftrightarrow \frac{\partial \delta y_A^s}{\partial \hat{x}_B} = \delta \frac{\partial \delta y_B^s}{\partial \hat{x}_B}$$

which is satisfied by construction of the equilibrium assignment. Moreover, (90) implies uniqueness of $w(\hat{x}_A, \hat{x}_B)$ by Proposition 2.

### C.2 Distributions

**Lemma 6 (The Effect of Distributions on Assignment and Wages).**

(i) **Assignment:** Let a worker with $(x_A, x_B) = (|x|, 0)$ be a specialist in task $A$, with $(x_A, x_B) = (0, |x|)$ be a specialist in $B$ and with $(x_A, x_B) = (|x|, |x|)$ be a generalist where $|x| < \infty$. Then:

(a) **Generalist:** $|y_i| \leq |x|$ for $\rho_x \leq \rho_y \quad \forall \quad i \in \{A, B\}$

(b) **Specialist in A:** $|y_A| \leq |x|$ for $\rho_x \leq \rho_y$

(c) **Specialist in B:** $|y_B| \leq |x|$ for $\rho_x \leq \rho_y$

Define over and under-qualification from a worker’s perspective as follows: A worker $(x_A, x_B)$ is over(under)-qualified if $x_A \geq (\leq) y_A, x_B \geq (\leq) y_B$ with at least one inequality being strict. Hence, generalists are over(under)qualified for $\rho_x < (>\rho_y$, and specialists are over(under)qualified for $\rho_x > (\rho_y$.

(ii) **Moments of the Wage Distribution:**

(a) $E(w(x_A, x_B))$ is maximized when $\rho_x = \rho_y$ and has two infima at $\rho_x = 0, \rho_y = 1$ and $\rho_y = 0, \rho_x = 1$.

(b) $Var(w(x_A, x_B))$ has a supremum at $\rho_x = 1, \rho_y = 1$. It is minimized for $\rho_i = 0, i \in \{x, y\}$.

(c) The wage distribution is positively skewed. It achieves its maximum positive skewness for $\rho_x \to 1, \rho_y = 0$ and $\rho_x = 0, \rho_y \to 1$ and its minimum at $\rho_x = \rho_y = 0$. 

43
Proof.
(i) Assignment:
(a) The generalist has $|x_A| = |x_B| = |x|$. From $J_\mu$ (notice that $J_\mu = J_\mu$ and $J_\mu$ is given by (17)), it follows that

$$y_i = \sqrt{\frac{1 + \rho_y}{1 + \rho_x}} |x| \quad \forall \ i \in \{A, B\}$$

Hence, $\forall \ i \in \{A, B\} \ |y_i| \geq |x|$ for $\rho_x \leq \rho_y$.
(b) Specialist A has $|x_A| = |x|$ and $|x_B| = 0$. Only the firm match along the A dimension matters for output and hence for the wage. From $J_\mu$, it follows that

$$y_A = \frac{1}{2} \left( \sqrt{\frac{1 + \rho_y}{1 + \rho_x} + \sqrt{\frac{1 - \rho_y}{1 - \rho_x}}} \right) |x|$$

where

$$\frac{1}{2} \left( \sqrt{\frac{1 + \rho_y}{1 + \rho_x} + \sqrt{\frac{1 - \rho_y}{1 - \rho_x}}} \right) \geq 1$$

$$\Leftrightarrow \sqrt{1 + \rho_y} \sqrt{1 - \rho_x} + \sqrt{1 - \rho_y} \sqrt{1 + \rho_x} - 2 \sqrt{1 - \rho_x^2} \geq 0$$

$$\Leftrightarrow \rho_x \geq \rho_y.$$

Hence, $|y_A| \geq |x|$ for $\rho_x \geq \rho_y$.
(c) The same argument establishes that $|y_B| \geq |x|$ for $\rho_x \geq \rho_y$.

(ii) Moments: The mean, variance, third moment and skewness of the wage distribution are given by

$$E(w(x_A, x_B)) = tr(J_\mu \Sigma_x) = \frac{1}{2} \left( \sqrt{(1 + \rho_x)(1 + \rho_y)} + \sqrt{(1 - \rho_x)(1 - \rho_y)} \right)$$

$$Var(w(x_A, x_B)) = 2tr(J_\mu \Sigma_x J_\mu \Sigma_x) = 1 + \rho_x \rho_y$$

$$E(w^3) = 3tr(J_\mu \Sigma_x) + 6tr(J_\mu \Sigma_x)tr(J_\mu \Sigma_x J_\mu \Sigma_x) + 8tr(J_\mu \Sigma_x J_\mu \Sigma_x J_\mu)$$

$$E \left[ \left( \frac{w - E(w)}{\sqrt{Var(w)}} \right)^3 \right] = \frac{E(w^3) - 3E(w)Var(w) - E(w)^3}{Var(w)^{\frac{3}{2}}}$$

$$= \frac{\left( \sqrt{(1 + \rho_y)(1 + \rho_x)} + \sqrt{(1 - \rho_y)(1 - \rho_x)} \right) \left( \frac{69}{4} + \frac{63}{4} \rho_x \rho_y - \frac{33}{4} \sqrt{(1 - \rho_y^2)(1 - \rho_x^2)} \right)}{(1 + \rho_x \rho_y)^{\frac{3}{2}}}$$

where, as before, $J_\mu$ denotes the Jacobian of the matching function and $\Sigma_x$ the covariance matrix of skills. See e.g. Magnus (1978) for the derivation of moments of quadratic forms in normal variables.
(a) Maximizing \( E(w(x_A, x_B)) \) with respect to \( \rho_y \) and \( \rho_x \) yields:

\[
\max_{\rho_x, \rho_y} \frac{1}{2} \left( \sqrt{(1 + \rho_x)(1 + \rho_y)} + \sqrt{(1 - \rho_x)(1 - \rho_y)} \right)
\]

- \( \rho_x : \rho_x = \rho_y \)
- \( \rho_y : \rho_x = \rho_y \)

The Hessian of this maximization problem is given by:

\[
H(E(w)) = \begin{bmatrix}
-\frac{1}{8} \left( \frac{\sqrt{1 - \rho_y}}{(1 - \rho_x)^2} + \frac{\sqrt{1 + \rho_y}}{(1 + \rho_x)^2} \right) & \frac{1}{8} \left( \frac{1}{\sqrt{1 - \rho_x} \sqrt{1 - \rho_y} + \frac{1}{\sqrt{1 + \rho_x} \sqrt{1 + \rho_y}}} \right) \\
\frac{1}{8} \left( \frac{\sqrt{1 - \rho_x}}{(1 - \rho_y)^2} + \frac{\sqrt{1 + \rho_x}}{(1 + \rho_y)^2} \right) & -\frac{1}{8} \left( \frac{\sqrt{1 - \rho_x}}{(1 - \rho_y)^2} + \frac{\sqrt{1 + \rho_x}}{(1 + \rho_y)^2} \right)
\end{bmatrix}
\]

It can be shown that the Hessian is negative semi-definite \( \forall \rho_x, \rho_y \). Hence, \( E(w(x_A, x_B)) \) is concave. At \( \rho_x = \rho_y \), \( \text{Det}(H(E(w))) = 0 \), indicating that \( \rho_x = \rho_y \) are (degenerate) maxima. Due to concavity of \( E(w) \), its infimum is achieved at the boundary of the domain \( 0 \leq \rho_x < 1, 0 \leq \rho_y < 1 \). \( E(w) \) has two infima at \( \rho_x = 0, \rho_y = 1 \) and \( \rho_y = 0, \rho_x = 1 \).

(b) \( \text{Var}(w(x_A, x_B)) = 1 + \rho_x \rho_y \) has no interior maximum. It is immediate that the supremum of \( \text{Var}(w(x_A, x_B)) \) is at \( \rho_x = 1, \rho_y = 1 \) (and that the minimum is reached at \( \rho_i = 0 \)).

(c) Follows from simulations. The code is available upon request. ■

C.3 Technological Change

**Lemma 7 (Complementarity-Enhancing Technological Change)** Suppose \( Z' > Z \). Then:

(i) The assignment of workers to firms remains unchanged.

(ii) The relative wage of any two workers remains unchanged.

(iii) Mean output and mean wage: \( \forall \rho_x, \rho_y, \frac{\partial E(F(x_A, x_B, y_A, y_B))}{\partial Z} > 0 \) and \( \frac{\partial E(w(x_A, x_B))}{\partial Z} > 0 \) and more so for \( \rho_x = \rho_y \). Variance of the wage distribution: \( \frac{\partial \text{Var}(w(x_A, x_B))}{\partial Z} > 0, \frac{\partial^2 \text{Var}(w(x_A, x_B))}{\partial Z \partial \rho_x} > 0, \frac{\partial^2 \text{Var}(w(x_A, x_B))}{\partial Z \partial \rho_y} > 0 \).

**Proof.**

Assume the production technology is \( F = Z(x_A y_A + \delta x_B y_B) \) with \( \delta = 1 \) and there is TETC, i.e. \( Z' > Z \).

(i) \( Z \) affects both tasks symmetrically. Part (ii) of Proposition (3) applies.

(ii) \( w(x) = \frac{1}{2} Z x^T J_\mu x \) and, hence, \( \frac{w'(x)}{w(x)} = \frac{w'(\bar{x})}{w(\bar{x})} \forall \ x \neq \bar{x} \).

(iii) \( E(w(x_A, x_B)) = Z \frac{1}{2} \left( \sqrt{(1 + \rho_x)(1 + \rho_y)} + \sqrt{(1 - \rho_x)(1 - \rho_y)} \right), \\
E(F(x_A, x_B, y_A, y_B)) = Z \left( \sqrt{(1 + \rho_x)(1 + \rho_y)} + \sqrt{(1 - \rho_x)(1 - \rho_y)} \right) \right) \right) \right) \right). \) Both are increasing in \( Z \).

\( \text{Var}(w(x_A, x_B)) = Z^2(1 + \rho_x \rho_y) \), which is increasing in \( Z \). The results on the cross-partial follow immediately. ■

**Proof of Proposition 5.**

Assume the production technology is \( F = x_A y_A + \delta' x_B y_B \) and \( \delta' < \delta = 1 \) (there is TBTC). Denote the variables that depend on \( \delta' \) by prime.
From (69), the elements of the Jacobian are as follows:

\[
\begin{align*}
\frac{\partial y^*_A}{\partial x_A} &= \frac{1 + \delta \sqrt{1-\rho_y^2}}{\sqrt{1 + 2\delta(\rho_x \rho_y + \sqrt{1 - \rho_y^2 \sqrt{1 - \rho_x^2}}) + \delta^2}} \\
\frac{\partial y^*_B}{\partial x_B} &= \frac{\delta + \sqrt{1-\rho_y^2}}{\sqrt{1 + 2\delta(\rho_x \rho_y + \sqrt{1 - \rho_y^2 \sqrt{1 - \rho_x^2}}) + \delta^2}} \\
\frac{\partial y^*_A}{\partial x_B} &= \frac{\delta \left( \rho_y - \rho_x \frac{\sqrt{1-\rho_y^2}}{\sqrt{1-\rho_x^2}} \right)}{\sqrt{1 + 2\delta(\rho_x \rho_y + \sqrt{1 - \rho_y^2 \sqrt{1 - \rho_x^2}}) + \delta^2}} \\
\frac{\partial y^*_B}{\partial x_A} &= \frac{\rho_y - \rho_x \frac{\sqrt{1-\rho_y^2}}{\sqrt{1-\rho_x^2}}}{\sqrt{1 + 2\delta(\rho_x \rho_y + \sqrt{1 - \rho_y^2 \sqrt{1 - \rho_x^2}}) + \delta^2}}
\end{align*}
\]

1. Step: Show that \( \left| \frac{\partial y^*_A}{\partial x_A} - 1 \right| < \left| \frac{\partial y^*_B}{\partial x_B} - 1 \right| \).

First notice that \( \frac{\partial y^*_A}{\partial x_A} \leq 1 \) for \( \rho_x \leq \rho_y \):

\[
\frac{\partial y^*_A}{\partial x_A} = \frac{1 + \delta' \sqrt{1-\rho_y^2}}{\sqrt{1 + 2\delta'(\rho_x \rho_y + \sqrt{1 - \rho_y^2 \sqrt{1 - \rho_x^2}}) + \delta'^2}} \leq 1
\]

\[
\Leftrightarrow \left( 1 + \delta' \frac{\sqrt{1-\rho_y^2}}{\sqrt{1-\rho_x^2}} \right) ^2 \leq 1 + 2\delta'(\rho_x \rho_y + \sqrt{1 - \rho_y^2 \sqrt{1 - \rho_x^2}}) + \delta'^2
\]

\[
\Leftrightarrow 2\delta' \left( \frac{1 - \rho_y^2 \rho_x^2 - \rho_x \rho_y \sqrt{1 - \rho_x^2}}{1 - \rho_x^2} + \delta'^2 \left( \frac{1 - \rho_y^2}{1 - \rho_x^2} - 1 \right) \right) \leq 0 \quad \text{for} \quad \rho_x \leq \rho_y.
\]

Moreover, from (91) and (92), \( \frac{\partial y^*_A}{\partial x_A} \geq \frac{\partial y^*_B}{\partial x_B} \) for \( \rho_x \leq \rho_y \) since

\[
1 + \delta \frac{\sqrt{1-\rho_y^2}}{\sqrt{1-\rho_x^2}} \geq \delta + \frac{\sqrt{1-\rho_y^2}}{\sqrt{1-\rho_x^2}}
\]

\[
\Leftrightarrow (1 - \delta) \left( \frac{\sqrt{1-\rho_x^2} - \sqrt{1-\rho_y^2}}{\sqrt{1-\rho_x^2}} \right) \leq 0 \quad \text{for} \quad \rho_x \leq \rho_y
\]
which completes the first step.

2. Step: Show that \[ \left| \frac{\partial y_A'}{\partial x_B} \right| < \left| \frac{\partial y_B'}{\partial x_A} \right|. \] This is immediate from (93) and (94) for \( \delta' < 1 \).

(ii) Let a worker with \((x_A, x_B) = (|x|, 0), |x| < \infty\) be a specialist in task \(A\) and the worker \((x_A, x_B) = (0, |x|), |x| < \infty\) be a specialist in task \(B\). Then, since \(\delta' < \delta = 1\), using (19), the following inequality holds:

\[
\frac{w'(|x|, 0)}{w'(0, |x|)} = 1 + \frac{1}{2} x^2 \frac{\partial y_A'}{\partial x_A} + \frac{1}{2} x^2 \frac{\partial y_B'}{\partial x_B} = \frac{1}{4} \left( \frac{\sqrt{1+\rho_y}}{\sqrt{1-\rho_x}} + \frac{\sqrt{1-\rho_y}}{\sqrt{1-\rho_x}} \right) = \frac{1}{2} x^2 \frac{\partial y_A'}{\partial x_A} = w(|x|, 0).
\]

Moreover, \(\frac{w'(|x|, 0)}{w'(0, |x|)}\) is decreasing in \(\sqrt{1-\rho_y} \sqrt{1-\rho_x}\). Since \(\sqrt{1-\rho_y} / \sqrt{1-\rho_x}\) is larger for \(\rho_x > \rho_y\), the result follows.

Let a worker with \((x_A, x_B) = (|x|, |x|), |x| < \infty\) be a generalist. Then, using (19):

\[
\frac{w(|x|, |x|)}{w(0, |x|)} = 1 + \frac{1}{2} \left( 1 + \delta \sqrt{1-\rho_y} \sqrt{1-\rho_x} \right) + \delta \left( \rho_y - \rho_x \sqrt{1-\rho_y} \sqrt{1-\rho_x} \right) + \frac{1}{2} \delta \left( \delta + \sqrt{1-\rho_y} \sqrt{1-\rho_x} \right) \equiv Z.
\]

Since,

\[
\frac{\partial Z}{\partial \delta} = \sqrt{1-\rho_y} \left( -\frac{1}{4} \delta^2 - \frac{1}{2} \rho_x + \frac{1}{2} \delta^2 \rho_x \right) - \frac{1}{2} \delta - \frac{1}{2} \delta^2 \rho_y < 0 \quad \text{for} \quad \rho_x < \rho_y < 1
\]

it follows from \(\delta' < \delta = 1\) that \(w'(|x|, |x|) / w'(0, |x|) > w(|x|, |x|) / w(0, |x|)\).
Furthermore,

$$
\frac{\partial^2 Z}{\partial \delta \partial \rho_x} = \left( \delta \rho_x (B + 1) - \frac{3}{2} AB \rho_x + \delta \left( \frac{\delta^2}{4} (1 - \rho_x^2) + \rho_x + \delta \rho_y \rho_y \right) + A \delta^2 \right) \left( \frac{A \delta}{(1 - \rho_x^2) (\frac{1}{2} \delta (\delta + A))^3} \right) > 0
$$

(97)

$$
\frac{\partial^2 Z}{\partial \delta \partial \rho_y} = \left( \delta \rho_y (-B - 1) + \frac{3}{2} AB \rho_y - \rho_y \delta (1 + \delta) - \frac{1}{4} \delta (\delta^2 + A) \right) \left( \frac{\delta}{\sqrt{(1 - \rho_x^2)(1 - \rho_y^2) (\frac{1}{2} \delta (\delta + A))^3}} \right) < 0
$$

(98)

where \( A \equiv \sqrt{\frac{1 - \rho_y^2}{1 - \rho_x^2}} \) and \( B \equiv (-\frac{1}{4} \delta^2 - \frac{1}{4} + \frac{1}{2} \delta^2 \rho_x) < 0 \). It follows from (95)-(98) that \( \frac{w'(|x|,|x|)}{w'(0,|x|)} > \frac{w(|x|,|x|)}{w(0,|x|)} \) is more pronounced for \( \rho_x < \rho_y \) and, in fact, the change in relative wages is maximized for \( \rho_y \to 1 \) ad \( \rho_x \to 0 \).

(iii) Using the results from (i), the equilibrium wage can be expressed as:

$$
w(x_A, x_B) = \frac{1}{2} [x_A \ x_B] \left[ \begin{array}{c}
\frac{1 + \delta \sqrt{1 - \rho_x^2}}{\sqrt{1 + 2 \sqrt{(\rho_x^2 + \sqrt{1 - \rho_x^2}) + \delta^2}}} \\
\frac{\delta (\rho_y - \rho_x \sqrt{1 - \rho_x^2})}{\sqrt{1 + 2 \sqrt{(\rho_x^2 + \sqrt{1 - \rho_x^2}) + \delta^2}}}
\end{array} \right] [x_A \ x_B] = \frac{1}{2} [x_A \ x_B] \left[ \begin{array}{c}
a \ b \\
c \ d
\end{array} \right]
$$

(99)

The expressions for the moments of the wage distribution are given in the Proof of Lemma 6. Simply replace the Jacobian \( J_\mu \) with the matrix \( M \) from (99)

$$
E(w(x_A, x_B)) = tr(M \Sigma_x) = \frac{1}{2} a + b \rho_x + \frac{1}{2} d
$$

$$
Var(w(x_A, x_B)) = 2tr(M \Sigma_x M \Sigma_x) = \frac{1}{2} ((a + b \rho_x)^2 + 2(a \rho_x + b)(b + d \rho_x) + (b \rho_x + d)^2)
$$

$$
E \left[ \left( \frac{w - E(w)}{\sqrt{Var(w)}} \right)^3 \right] = \frac{E(w^3) - 3E(w)Var(w) - E(w)^3}{Var(w)^{\frac{3}{2}}} = \frac{3E(w) + 8tr(M \Sigma_x M \Sigma_x M \Sigma_x) - E(w)^3}{Var(w)^{\frac{3}{2}}}
$$

where \( a, b, c, d \) are defined in (99) and where

$$
\text{tr}(M \Sigma_x M \Sigma_x M \Sigma_x) = (a + b \rho_x)((a + b \rho_x)^2 + (a \rho_x + b)(b + d \rho_x)) + 2(b + d \rho_x)(a \rho_x + b)(a + 2b \rho_x + d)
$$

$$
+ (b \rho_x + d)((b + d \rho_x)(a \rho_x + b) + (b \rho_x + d)^2)
$$

The results on how these moments respond to TBTC are based on simulations and available upon request. ■
D P-matrix Property of $D^2_{xy}F$ is Not Sufficient for P-matrix Property of $D_{x,y^*}$

To show this, assume the following production technology:

$$F(x_A, x_B, y_A, y_B) = \alpha x_A y_A + \beta x_A y_B + \gamma x_B y_A + \delta x_B y_B$$

For the sake of the argument, set $\alpha = \delta = 1$. Hence, the firm’s problem is given by:

$$\max_{x_A, x_B \in X} x_A y_A + \beta x_A y_B + \gamma x_B y_A + x_B y_B - w(x_A, x_B)$$

From before, I know that for $\beta = \gamma = 0$, the spectral square root achieves the de-correlation of the variables. So, I will again use rotations of the spectral square root for the transformation, where the rotation matrices are parameterized by $\beta, \gamma$. The orthonormal rotation matrices are given by:

$$R_x = \begin{bmatrix} \alpha_x & -\beta_x \\ \beta_x & \alpha_x \end{bmatrix}, \quad R_y = \begin{bmatrix} \alpha_y & -\beta_y \\ \beta_y & \alpha_y \end{bmatrix}$$

They have to satisfy the following conditions:

$$\alpha_x^2 + \beta_x^2 = 1 \quad (100)$$
$$\alpha_y^2 + \beta_y^2 = 1 \quad (101)$$
$$\frac{\partial y_A^*}{\partial x_B} + \beta \frac{\partial y_B^*}{\partial x_B} = \gamma \frac{\partial y_A^*}{\partial x_A} + \frac{\partial y_A^*}{\partial x_A} \quad (102)$$

where (102) is the integrability condition, which can be expressed as:

$$\alpha_y \left( \sqrt{\frac{1 + \rho_y}{1 + \rho_x}} + \sqrt{\frac{1 - \rho_y}{1 - \rho_x}} \right) (\beta - \gamma) = \sqrt{1 - \alpha_y^2} \left( 2 \left( \sqrt{\frac{1 + \rho_y}{1 - \rho_x}} + \sqrt{\frac{1 - \rho_y}{1 + \rho_x}} \right) + \left( \frac{1 + \rho_y}{1 - \rho_x} - \sqrt{\frac{1 - \rho_y}{1 + \rho_x}} \right) (\beta + \gamma) \right) \quad (103)$$

(103) can be solved for $\alpha_y$:

$$\alpha_y = \pm \frac{\left( 2 \left( \sqrt{\frac{1 + \rho_y}{1 + \rho_x}} + \sqrt{\frac{1 - \rho_y}{1 - \rho_x}} \right) + \left( \frac{1 + \rho_y}{1 - \rho_x} - \sqrt{\frac{1 - \rho_y}{1 + \rho_x}} \right) (\beta + \gamma) \right)}{\sqrt{\left( \sqrt{\frac{1 + \rho_y}{1 + \rho_x}} + \sqrt{\frac{1 - \rho_y}{1 - \rho_x}} \right)^2 (\beta - \gamma)^2 + \left( 2 \left( \sqrt{\frac{1 + \rho_y}{1 + \rho_x}} + \sqrt{\frac{1 - \rho_y}{1 + \rho_x}} \right) + \left( \frac{1 + \rho_y}{1 - \rho_x} - \sqrt{\frac{1 - \rho_y}{1 + \rho_x}} \right) (\beta + \gamma) \right)^2}} \quad (104)$$

From (101) $\beta_y = \sqrt{1 - \alpha_y^2}$. Also, set $\alpha_x = \pm 1$ (there is one degree of freedom in system (100)-(102)), which implies $\beta_x = 0$ by (100). Notice that for $\beta > (<) \gamma$, the parameters $\alpha_y, \alpha_x > (<) 0$ satisfy (103).
Under this parameterization of the rotation matrices, the Jacobian of $\mu(x)$ is given by:

$$\frac{\partial y_A^*}{\partial x_A} = \frac{1}{2} \left( \alpha y \left( \frac{\sqrt{1 + \rho y}}{\sqrt{1 + \rho x}} + \frac{\sqrt{1 - \rho y}}{\sqrt{1 - \rho x}} \right) + \sqrt{1 - \alpha y^2} \left( \frac{\sqrt{1 + \rho y}}{\sqrt{1 - \rho x}} - \frac{\sqrt{1 - \rho y}}{\sqrt{1 + \rho x}} \right) \right)$$

(105)

$$\frac{\partial y_A^*}{\partial x_B} = \frac{1}{2} \left( \alpha y \left( \frac{\sqrt{1 + \rho y}}{\sqrt{1 + \rho x}} - \frac{\sqrt{1 - \rho y}}{\sqrt{1 - \rho x}} \right) - \sqrt{1 - \alpha y^2} \left( \frac{\sqrt{1 + \rho y}}{\sqrt{1 - \rho x}} + \frac{\sqrt{1 - \rho y}}{\sqrt{1 + \rho x}} \right) \right)$$

(106)

$$\frac{\partial y_B^*}{\partial x_A} = \frac{1}{2} \left( \alpha y \left( \frac{\sqrt{1 + \rho y}}{\sqrt{1 + \rho x}} - \frac{\sqrt{1 - \rho y}}{\sqrt{1 - \rho x}} \right) + \sqrt{1 - \alpha y^2} \left( \frac{\sqrt{1 + \rho y}}{\sqrt{1 - \rho x}} + \frac{\sqrt{1 - \rho y}}{\sqrt{1 + \rho x}} \right) \right)$$

(107)

$$\frac{\partial y_B^*}{\partial x_B} = \frac{1}{2} \left( \alpha y \left( \frac{\sqrt{1 + \rho y}}{\sqrt{1 + \rho x}} + \frac{\sqrt{1 - \rho y}}{\sqrt{1 - \rho x}} \right) - \sqrt{1 - \alpha y^2} \left( \frac{\sqrt{1 + \rho y}}{\sqrt{1 - \rho x}} - \frac{\sqrt{1 - \rho y}}{\sqrt{1 + \rho x}} \right) \right)$$

(108)

Now, for the sake of the argument, let $\beta \to \infty$ and $\gamma = 0$. Notice that this parameterization does not violate the P-matrix property of $D_{2xy}F$. Then,

$$\lim_{\beta \to \infty} \alpha y = \frac{1}{2} \left( \sqrt{1 + \rho y} \sqrt{1 + \rho x} - \sqrt{1 - \rho y} \sqrt{1 - \rho x} \right).$$

(109)

As $\beta$ becomes large, (105) reads

$$\frac{\partial y_A^*}{\partial x_A} \bigg|_{\beta \to \infty} = \frac{1}{2} \sqrt{1 + \rho y} \sqrt{1 + \rho x} - \sqrt{1 - \rho y} \sqrt{1 - \rho x} \geq 0$$

(110)

where I used (109). Notice that $\frac{\partial y_A^*}{\partial x_A} \bigg|_{\beta \to \infty} < 0$ for

$$\sqrt{1 + \rho y} \sqrt{1 + \rho x} - \sqrt{1 - \rho y} \sqrt{1 - \rho x} < 0$$

$$\iff \rho y + \rho x < 0.$$

(111)

violating the $P$-matrix property of $D_{2xy}$. Similarly, it can be shown that

$$\frac{\partial y_A^*}{\partial x_B} \bigg|_{\beta \to \infty} < 0$$

$$\frac{\partial y_B^*}{\partial x_A} \bigg|_{\beta \to \infty} = 1$$

$$\frac{\partial y_B^*}{\partial x_B} \bigg|_{\beta \to \infty} = 0$$

Hence, the $P$-matrix property of $D_{2xy}F$ is not sufficient for the $P$-matrix property of $D_{2y}$. 

50
In this section, the model of multidimensional heterogeneity is embedded into a model with search frictions on the labor market and directed search. The search frictions stem from the lack of coordination of a large number of agents when applying for jobs. As in the baseline model with competitive labor market, I derive a sufficient condition on the production technology for purity of the equilibrium and assortative matching. This condition is a straightforward generalization of root-supermodularity from the one-dimensional setting (Eeckhout and Kircher (2010)) to the setting with multidimensional heterogeneity. I briefly outline the model, which is identical to theirs except that it allows for two-dimensional skills and productivities.

The frictional hiring process of firms can be described by a static game with three stages: In the first stage, every firm, characterized by some productivity bundle \( y = (y_A, y_B) \), posts a wage \( w(y) \). In a second stage, unemployed workers observe these wages. They anticipate that different wages \( w(y) \) are associated with different applicant-vacancy ratios \( q(y) \in [0, \infty] \), which I will refer to as the queue length. They choose to apply at firm \( y \), characterized by a pair \( (w(y), q(y)) \), to maximize their expected income. In the last stage, firms that receive at least one application hire one worker and production takes place. If a firm receives more than one applicant, then it hires one of them at random. Unmatched workers remain unemployed and unmatched firms will end up with a vacant job. Unmatched agents produce nothing and have zero payoff.

Denote the probability that firm \( y \) fills a vacancy by \( m(q(y)) \) and the probability that a worker is hired by that firm by \( m(q(y)) q(y) \) where the matching technology has the following properties: \( m(.) \) is twice differentiable and satisfies \( m_q(.) > 0, m_{qq}(.) < 0, \frac{\partial m(q(y))}{\partial q(y)} < 0, m(0) = 0 \) and \( m(\infty) = 1 \). In words, the vacancy-filling probability is strictly increasing and strictly concave in the queue length whereas the worker’s hiring probability is strictly decreasing in the number of other workers queueing up for the same job.

In order for a firm to attract a certain worker type \( x = (x_A, x_B) \) it needs to offer him an expected payoff that is at least as high as his expected equilibrium market utility \( U(x) \). \( U(x) = \frac{m(q(y))}{q(y)} w(y) \) is what he would get at his best alternative job, where \( w(y) \) is the worker’s actual wage when hired. The expected market wage \( U(x) \) implicitly defines the queue length \( q(y) \) as a function of the actual wage \( w(y) \), which can be shown to be an increasing function. Notice that the firms (and workers) take \( U(x) \) and hence the relationship between \( q(y) \) and \( w(y) \) as given, which is justified by the large number of both workers and firms. Contrary to the competitive labor market, however, here the firm chooses the wage. The firm’s problem is:

\[
\begin{align*}
\max_{w, q} & \quad m(q)(F(x, y) - w) \\
\text{s.t.} & \quad m(q) q w \geq U(x)
\end{align*}
\]  

Before analyzing the equilibrium properties, it is useful to define a pure equilibrium in this setting.
Definition 7 (Pure Equilibrium with Search) A pure symmetric competitive search equilibrium consists of wages, queue length, and market utility \((w(y), q(y), U(x))\) as well as a mapping \(y^* = \mu(x)\) s.t.:  
(i) Firm optimality: Given \(U(x)\) and other firms’ strategies, each firm \(y\) solves (112) s.t. (113);  
(ii) Worker optimality: A worker \(x\) applies to a job at \(y\) only if \(\frac{m(q(y))}{q(y)} w(y) \geq U(x)\);  
(iii) Purity: \(y^* = \mu(x)\) is a one-to-one function;  
(iv) Market clearing: The queue length \(q\) satisfies the market clearing for applications.

In equilibrium, the firm will not offer more utility to the worker than necessary to attract him. Hence, the constraint (113) holds with equality. The maximization problem can be reformulated as:

\[
\max_{x,q,w} \quad m(q)(F(x,y) - w) \\
\text{s.t.} \quad \frac{m(q)}{q} w = U(x) \tag{114}
\]

In words, the firm can choose the wage and the trading probabilities in order to attract a certain worker type (to whom it needs to offer his market utility). The constraint is similar to a participation constraint. When substituting (114) into the firm’s objective function, the maximization problem reads:

\[
\max_{x,q} \quad m(q)F(x,y) - qU(x)
\]

The first order conditions with respect to \(q, x\) are respectively given by

\[
\begin{align*}
q_y(q(x,y)) - U(x) = 0 \\
m(q)F_{x,y}(x,y) - qU_{x,y}(x) = 0 \\
m(q)F_{x,x}(x,y) - qU_{x,x}(x) = 0
\end{align*}
\]

where subscripts denote derivatives. An assignment is consistent with the equilibrium only if the second-order conditions are satisfied, i.e. if the Hessian is negative semi-definite. Denote the decision variables of the firm by the vector \(z = (x, q)\), where \(x = (x_A, x_B)\). Hence, the Hessian is a 3x3 matrix. Differentiating the FOCs (evaluated at \(y^*\)) with respect to \(z = (x, q)\) yields the Hessian evaluated at the equilibrium assignment \(y^*\)

\[
D^2_{zz} \left( m(q(y^*))F(x,y^*) - q(y^*)U(x) \right) = D^2_{qq} \left( m(q(y^*))F(x,y^*) \right) \left( \frac{F(x,y^*)F_{yy}(x,y^*)}{F_{x}(x,y^*)F_{y}(x,y^*)} - \epsilon_M(q(y^*)) \right) D_x y^* \tag{115}
\]

where \(\epsilon_M(q) \equiv \frac{m(q)(m(q)y - m(q))}{m(q)m_{qq}(q)y} = \frac{M(q)M_{qq}(q)}{M_{qq}(q)M(q)}\) is the elasticity of the aggregate matching function, as in Eeckhout and Kircher (2010).\(^{42}\) A necessary condition for optimality is that (115) is negative semi-definite. Notice that by assumption \(m_{qq}(q) < 0\) and hence \(D^2_{qq} (m(q)F(x,y^*))\) is a negative scalar. (115) is the multidimensional extension of expression (12) in Eeckhout and Kircher (2010).

\(^{42}\) The aggregate matching function is defined as the total number of matches that form when \(u\) workers and \(v\) vacancies are in the market \(M(u,v) = vm(u/v) = vm(q)\).
Optimality requires that
\[
\left( \frac{F(x, y^*)F_{xy}(x, y^*)}{F_x(x, y^*)F_y(x, y^*)} - \epsilon_M(q(y^*)) \right) (D_x y^*)
\] (116)

is positive semi-definite. The following proposition states a sufficient condition for purity and assortativeness of the equilibrium assignment in the multidimensional setting.

**Proposition 8 (Pure and Assortative Equilibrium under Search)** Let \( y = (y_A, y_B) \in Y \subseteq \mathbb{R}_+^2 \) and \( x = (x_A, x_B) \in X \subseteq \mathbb{R}_+^2 \). If
\[
\begin{bmatrix}
FF_{xy} - \epsilon_M(q) \\
FF_y x - \epsilon_M(q)
\end{bmatrix} = \begin{bmatrix}
\frac{FF_{xy}}{F_y F_x} - \epsilon_M(q) & 0 \\
0 & \frac{FF_{xy}}{F_y F_x} - \epsilon_M(q)
\end{bmatrix}
\] (117)
is a P-matrix (P\(^-\)-matrix), then \( D_x y^* \) is a P-matrix (P\(^-\)-matrix). The assignment \( \mu(x) = y^* \) is positive (negative) assortative. The equilibrium (if it exists) is globally unique.\(^{43}\)

**Proof of Proposition 8.** Since \( \begin{bmatrix}
FF_{xy} - \epsilon_M(q)
\end{bmatrix} \) is assumed to be diagonal, the symmetry of the Hessian, given by (116), requires that \( D_x y^* \) is sign-symmetric. Steps 1-3a of the proof of Lemma 1 apply, which proves the result that \( D_x y^* \) is a P-matrix. The result on assortativeness follows. The proof for global uniqueness is stated by Eeckhout and Kircher (2010) (p. 569) and applies here as well. \(\blacksquare\)

In the setting with one-dimensional heterogeneity, the matrix condition of Proposition 8 reduces to the technological condition of root-supermodularity, \( \frac{FF_{xy}}{F_y F_x} - \epsilon_M(q) \geq 0 \), which ensures both assortativeness and uniqueness. Root-supermodularity is a stronger notion of complementarity than supermodularity. This concept is extensively discussed in Eeckhout and Kircher. When there are search frictions, high skilled workers and high productivity firms have strong incentives to secure a match during search. The types who provide this trading insurance are the low types because matching is less important to them. Hence, search frictions are a force towards negative assortative matching. In order for PAM to obtain, the complementarities between skills and productivities must be stronger than supermodularity (which is required under the competitive labor market). If the production function is root-submodular, NAM obtains in equilibrium.

In the multidimensional setting, the interpretation of root-supermodularity is identical. However, similar to the frictionless model, one has to distinguish between production complementarities within and between tasks. If complementarities (given by root-supermodularity) within tasks are considerably stronger than complementarities between tasks, then the equilibrium is pure and the assignment is assortative along each task dimension. It is an open issue whether this condition can be relaxed to allow for between-task complementarities and whether there is a role for diagonal dominance as in the frictionless environment.

\(^{43}\)The focus here is on the characterization of the equilibrium. Existence is dealt with in Eeckhout and Kircher (2010).
References


54


