The Pond Dilemma with Heterogeneous Relative Concerns*

Paweł Gola†

February 2025

Abstract

This paper explores team formation when workers differ in skills and their desire to out-earn co-workers. I cast this question as a two-dimensional assignment problem, characterise the equilibrium sorting and payoffs for a large class of specifications, and find that heterogeneity in status preferences drastically changes the distributional and organisational consequences of skill-biased technological change (SBTC). Strikingly, the benefits of SBTC trickle down to low-skill workers with weak relative concerns even when there are no complementarities in production. Moreover, SBTC incentivises domestic outsourcing, as firms seek to avoid detrimental social comparisons between high- and low-skill workers.

JEL Codes: C78, D91, J31.

Keywords: relative concerns, multidimensional sorting, trickle-down effects, skill-biased technological change, outsourcing

*The paper was first circulated in September 2024. I would like to thank Miguel Ballester, Stuart Breslin, Antonio Cabrales, Ed Hopkins, Tai-Wei Hu, Emir Kamenica, Alastair Langtry, Andriy Zapechelnyuk, Yuejun Zhao, and the seminar audiences in Edinburgh and Warsaw for helpful discussions and suggestions. I gratefully acknowledge the support of the UKRI Horizon Europe Guarantee (Grant Ref: EP/Y028074/1). All errors are mine.

†pawel.gola@ed.ac.uk, University of Edinburgh, 30 Buccleuch Place, Edinburgh, EH8 9JT.

1 Introduction

Which boats are lifted by a rising tide? When the tide is an improvement in technology, the answer appears obvious: The relative beneficiaries are those who become more productive, and possibly also those who produce the goods and services these more productive workers consume. Decades of empirical research on skill-biased technological change (SBTC) confirm this view: The main beneficiaries were high-skill workers (Bound and Johnson, 1992; Katz and Murphy, 1992; Juhn, Murphy, and Pierce, 1993) and those low-skill workers who work for high-earners (Mazzolari and Ragusa, 2013), typically in close geographical proximity (Manning, 2004). This, however, is at best half of the answer—after all, if not much skill is needed in their jobs, then why are these gains by a subset of low-skill workers not competed away over the medium-to-long-term?[1]

In this paper, I develop a theory of labour market sorting in the presence of heterogeneous relative concerns, which provides a simple answer to this question: The low-skill workers who benefit from SBTC are those with comparatively weak status concerns, that is, the workers who do not mind being surrounded by people who earn more than them. The point of departure for my theory is the observation that to form productive and durable teams, it is not sufficient to find workers with complementary skills. Preference and personality compatibility matter as well, and the relative concerns of the team-members are of particular importance: Many a famous sports team and music band have disintegrated because multiple members felt they deserved to be the biggest fish in that particular pond. Less anecdotally, there is strong empirical evidence that humans care about their relative position within the reference group and are willing to accept lower absolute wages to improve their relative earnings.[2] Moreover, the strength of these relative concerns—and their close cousin, competitiveness—differ across individuals, affecting their career and location choices.[3]

It is this heterogeneity in relative concerns which produces heterogeneous trickle-down effects of SBTC. In a nutshell, absent production complementarities, a high-skill worker with strong relative concerns would like to match a low-skill worker, as this would provide her with high within-firm status.[4] While a low-skill worker with strong relative concerns would require a large compensation for accepting low status, a low-skill worker without relative concerns requires no such compensation. High-skill workers with strong relative concerns therefore match with these status-indifferent low-skill workers and the avoided compensation is split between them.[5] Crucially, this wage premium received by status-indifferent low-skill workers is not a

[1]The literature on routine-biased technological change, started by Autor, Levy, and Murnane (2003), posits that low-skill workers in non-routine occupations also gained from technological change. Unless these beneficiaries are intrinsically different than other low-skill workers—in which case their gains stem from increased productivity—the puzzle remains: Why have these gains not been competed away? In other words, the puzzle is about gains in occupations with low barriers of entry.

[2]See, for example, Luttmer (2005); Card, Mas, Moretti, and Saez (2012); Perez-Truglia (2020); Bottan and Perez-Truglia (2020).

[3]See Buser, Niederle, and Oosterbeek (2014) and Bottan and Perez-Truglia (2020).

[4]I focus first on the case without production complementarities to present the core mechanism cleanly. Absent production complementarities, matching is driven by relative concerns only, with production complementarities output-maximisation matters as well. The interactions between production complementarities and status concerns generate additional insights about outsourcing and inequality that I discuss below.

[5]The exact split is determined endogenously and depends both on the production technology and on the distribution of preferences.

compensating differential for status disutility—they have none—but *reflects the high demand from high-skill workers for matches with low-skill workers.*

Skill-biased technological change increases the hypothetical compensation that low-skill workers with strong relative concerns would demand to match with a high-skill worker—but not their actual wages, because these matches never materialise. Since the increase in this hypothetical compensation worsens the outside option of high-skill workers, status-indifferent low-skill workers capture part of this increase, which raises their wage premium. Thus, SBTC triggers heterogeneous trickle-down effects. Note that this trickle-down can be very large: If most low-skill workers care strongly about relative concerns, then those who do not may reap a benefit larger than the *per capita* increase in output in the economy.

My theory connects SBTC to another puzzling empirical trend, the well-documented increase in domestic outsourcing (Goldschmidt and Schmieder, 2017; Bergeaud, Malgouyres, Mazet-Sonilhac, and Signorelli, 2024). The same heterogeneity in relative concerns that determines which low-skill workers benefit from SBTC also affects firm boundaries. Consider a world where, due to production complementarities, a match between a low-skill worker who cares greatly about status and a status-indifferent high-skill worker is output-maximising. Yet, it may not be mutually beneficial, because the worker with strong relative concerns would be very unhappy about their low status in such a match. However, if—as argued by Nickerson and Zenger (2008)—social comparisons are more salient within than across firm boundaries, a firm may salvage such output-maximising match by outsourcing the low-skill worker, thus avoiding the potentially detrimental social comparison altogether. The size of these potential distortions depends on the difference in productivity between the high- and low-skill workers, as this difference determines how much lower the low-skill worker's wage and status are. SBTC, by increasing this difference in productivity, raises the number of firms that choose to outsource.

The interaction between relative concerns and production complementarities yields another striking result: The presence of heterogeneous relative concerns can actually increase overall wage inequality in the economy. This occurs through its effect on equilibrium sorting. When low- and high-skill workers are complements, the equilibrium without relative concerns features negative assortative matching (NAM). However, with heterogeneous relative concerns, workers may prefer positive assortative matching (PAM), as this avoids detrimental social comparisons. While this shift from NAM to PAM always reduces within-firm inequality, it can increase between-firm inequality so dramatically that overall inequality rises. Importantly, this result emerges precisely when relative concerns are weak among high-skill workers but strong among low-skill workers—conditions that make the model isomorphic to one where all agents are averse to within-firm inequity. Thus, paradoxically, aversion to local wage inequality can amplify global inequality.

The rest of the paper is structured as follows. Section 2 discusses the related literature. Section 3 develops a one-sided, one-to-one assignment model in which workers have a specific keeping-up-with-Joneses (KUJ) utility function and differ in both skill and their status preferences. Section 4 characterises the equilibrium, and derives comparative statics results. Section 5 develops a social-comparisons-based theory of the firm, and uses it to explore the consequences of skill-biased technological changes for outsourcing. Section 6 concludes. Appendix A contains omitted proofs and derivations. Appendix B extends the analysis to the case of general KUJ utility.

2 Related Literature

Sorting with Relative Concerns This paper contributes to the literature on labour market sorting with heterogeneous relative concerns. Most importantly, this is the first paper to (a) provide analytical expressions for sorting and wages in settings with rich skill and preference heterogeneity, (b) allow a worker's productivity to depend on their co-worker's skill, (c) study how changes to the production function affect sorting, inequality and outsourcing in the presence of heterogeneous relative concerns, and (d) consider the impact that heterogeneous relative concerns have on firms' boundaries.

I am aware of three articles and one book that study the problem of how workers sort into teams/firms if they have heterogeneous relative concerns. The seminal work by Robert Frank (Frank, 1984b, 1985) posed this important problem and unearthed the fundamental insights that the presence of heterogeneous relative concerns means that within-firm wage inequality is lower than productivity inequality, and that workers with stronger relative concerns end up having less skilled co-workers. Fershtman, Hvide, and Weiss (2006) extended the problem posed by Frank to include effort provision, finding that firms consisting of workers with strong and weak relative concerns require workers with strong relative concerns to exert more effort.[6] Langtry (2023) differs from the other work on this topic (including mine), in that wages are set exogenously in his model, thus precluding high-skill workers from compensating low-skill workers for they lower status.[7] For that reason, relative concerns affect wage inequality through sorting only, and the trickle-down effect—critical for my results—is absent.[8]

In addition to that, Cabrales, Calvó-Armengol, and Pavoni (2008); Cabrales and Calvó-Armengol (2008) consider sorting in the presence of inequity aversion rather than relative concerns, and show that inequity aversion leads to more positive and assortative sorting in skill; a result that is also implied by my Proposition 2.[9] In contrast to my work, they only allow for additively separable production functions, and so the surprising result that inequality aversion may increase overall wage inequality does not occur in their settings.

Multidimensional Sorting This article is one of the very few to fully characterise the equilibrium of a two-dimensional assignment model. With the exception of Gola (2021), who assumes that each firm uses only one dimension of skill in production, the other characterizations all leverage bi-linear surplus functions and Gaussian distribution of traits to retain tractability

[6]There is a large literature that studies the impact of relative concerns on effort provision (see, for example, Hopkins and Kornienko, 2004, 2009); this literature, however, assumes homogeneous relative concerns and is not concerned with sorting.

[7]It is worth noting that the bulk of Langtry (2023) is concerned with the altogether different problem of optimal choice of consumption on a network, when agents care about their neighbours' consumption.

[8]In that sense, Langtry (2023) is actually closer to the literature focusing on workers who have homogeneous relative concerns and choose between two occupations (e.g., Fershtman and Weiss, 1993; Mani and Mullin, 2004; Gola, 2024) than it is to Frank (1984b, 1985); Fershtman et al. (2006) and the present paper. In both Langtry (2023) and the occupational choice literature, the firms/occupations do not internalise the externalities caused by relative concerns, and thus all of the impact of relative concerns happens through sorting, rather than wage setting. For that reason, relative concerns affect sorting even when all workers care about status equally.

[9]Cabrales et al. (2008) motivate their focus on inequity aversion by arguing that status concerns would produce counterfactual sorting in skills in the economy. While this is true when status concerns are homogeneous, one of the insights from the current work is that when status concerns are heterogeneous, then any degree of sorting in skills can be rationalised, irrespective of the properties of the production function.

(Tinbergen, 1956; Bojilov and Galichon, 2016; Lindenlaub, 2017). This article is the first one to (a) provide close form solutions for trait distributions that are not Gaussian, (b) allow for one of the dimensions of heterogeneity to be a social preference rather than skill and (c) consider a one-sided sorting model. I do this by leveraging a unique property of one-sided matching: It is always isomorphic to a two-sided model with a symmetric surplus function and identical distributions of traits on each side. This property is extremely useful, because under reasonable conditions on the surplus function (see, for example, Proposition 11(b) in Lindenlaub, 2017), equilibrium sorting in such two-sided problems involves positive and assortative matching within each of the skill/preference dimensions. In other words, one-sided sorting problems are more amenable to the introduction of multidimensional traits than two-sided problems are, because the assumption of identical trait distribution is much easier to satisfy.[10]

Theory of the Firm The theory presented in this paper provides a clear answer to a very old question posed by the the transaction costs theory of the firm literature (Coase, 1937; Williamson, 1971; Klein, Crawford, and Alchian, 1978): Given that transaction costs provide a rationale for concentrating economic activity within firms, why is it not the case that all economic activity takes place in one gigantic firm? In my model, firm size is limited by the need to weaken social comparisons between high-skill workers who have weak relative concerns and low-skill workers with strong relative concerns.[11] In that, my theory formalises and expands upon Nickerson and Zenger (2008), who propose an informal theory of the firm based on the need for a firm to manage the cost of social comparison. In addition to being the first to formalise the 'social comparison' theory of firm, I expand on Nickerson and Zenger (2008) by considering agents who differ in the strength of relative concerns, which allows me to explain why seemingly identical firms make different outsourcing decisions.[12] I also show that the 'social comparisons' theory of the firm provides a natural explanation for the rise in outsourcing in the recent decades: Simply put, by increasing the difference in productivity between high- and low-skill workers, skill-biased technological change has drastically increased the cost of within-firm social comparisons, thus increasing firms' incentives to outsource.

Technology and Outsourcing To the best of my knowledge, this is the first paper to provide a formal model linking (skill-biased) technological change with outsourcing. Bergeaud et al. (2024) show empirically that firms connected to broadband internet engage in domestic outsourcing more than firms without such connection. Bergeaud et al. use the informal theory developed in Abraham and Taylor (1996) to explain why technological change may cause outsourcing. One of the reasons for outsourcing put forward by Abraham and Taylor is that, in

[10]The solution method in Tinbergen (1956); Galichon (2016); Lindenlaub (2017) also leverages this property: Because Gaussian distributions are closed under linear transformations and surplus is bi-linear, there exist transformations of the workers' traits that have the same distributions on both sides of the market.

[11]The 'property right' (Grossman and Hart, 1986; Hart and Moore, 1990) and 'incentive systems' (Holmstrom and Milgrom, 1994; Holmström, 1999) provide complementary explanations for where firm boundaries are drawn.

[12]Interestingly, the seeds of the social comparison based theory of the firm are present already in Coase (1937), who dismisses that theory's importance on the grounds that it would imply that entrepreneurs earn less than their employees. This implication, however, is incorrect as soon as one allows for heterogeneity in skills: In my model, 'entrepreneurs' are paid more than 'employees', simply because they are more skilled. This is true even though 'entrepreneurs' are indeed taking a pay cut compared to the case where they same-match.

the absence of outsourcing, within-firm wage inequality may be constrained by workers' inequity aversion. I model inequity aversion/relative concerns explicitly and highlight that skill-biased technological change—by creating pressure for higher wage differentials between high- and low-skill workers—naturally leads to more outsourcing, which further amplifies SBTC's impact on wage inequality.

Trickle-Down Effects This paper proposes a novel channel through which productivity gains among high-skill workers trickle down to (some!) low-skill workers. Manning (2004) and Mazzolari and Ragusa (2013) proposed and tested the hypothesis that gains from skill-biased technological change trickle down to local low-skill workers who produce non-tradeable goods and services. I see my theory as complementary, as it explains why these gains are not competed away—the non-gaining low-skill workers have strong relative concerns—and which low-skill workers gain. While not explicitly about trickle-down effects, Aghion, Bergeaud, Blundell, and Griffith (2019) document a related empirical pattern: Innovative firms pay a wage premium to low-skill workers, but not to high-skill workers. Aghion et al. explain this by positing that high- and low-skill workers are complements in innovative firms and that low-skill workers differ in their soft skills. Borrowing the assumption that high- and low-skill workers are complements in innovative firms, the theory presented in this paper is consistent with this empirical puzzle: The complementarity between high- and low-skill workers forces innovative firms to hire low-skill workers, but to do so, they need to pay a compensating differential for low status. Indeed, my theory is consistent with another empirical fact documented by Aghion et al.: that innovative firms are more likely to outsource. In my model, the innovative firms that hire high-skill workers with weak relative concerns would outsource, and the non-outsourcing innovative firms would pay a wage premium to low-skill workers.

3 Model

There is a continuum of workers, who differ along two dimensions—skill $x_1 \in D_{x_1}$ and (the inverse of) the strength of their relative concerns $x_2 \in D_{x_2}$, where D_{x_1}, D_{x_2} are both closed subsets of $\mathbb{R}_{\geq 0}$. The joint distribution of $(x_1, x_2) = \mathbf{x}$ is denoted by $H : D_{\mathbf{x}} \to [0, 1]$, where $D_{\mathbf{x}} \equiv D_{x_1} \times D_{x_2}$. The marginal distribution of x_k will be denoted by H_{x_k}, and the maximum (minimum) of D_{x_i} by $\bar{x}_i$ ($\underline{x}_i$). Finally, I assume that $\Pr(X_2 \leq x_2 | x_1)$ is absolutely continuous in x_2 for any $x_1 \in D_{x_1}$.

Workers sort into teams of size two, which makes this a one-sided, one-to-one assignment model. A match between a worker with skill x_1^k and a worker with skill x_1^j produces output according to a symmetric, increasing and twice-continuously differentiable function $F : D_{x_1}^2 \to \mathbb{R}$. The output $F(x_1^k, x_1^j)$ is then endogenously split into the wages of the two workers.

In contrast to standard assignment models, a worker's utility depends not only on their own wage, w^k, but also on the average wage within k's team $\bar{w}^{k,j} \equiv F(\mathbf{x}^k, \mathbf{x}^j)/2$. In other words, agents have a 'Keeping-Up with Joneses' (KUJ, see Gali, 1994) utility function, with

$$U(w^k, \bar{w}^{k,j}; x_2^k) \equiv x_2^k w^k + (1 - x_2^k)(w^k - \bar{w}^{k,j}). \tag{1}$$

The closer x_2 is to 0 the stronger are the worker's *relative concerns*. In particular, for workers with $x_1 = 1$, (1) reduces to standard neo-classical preferences, whereas workers with $x_2^k > 1$ enjoy working in teams in which the average wage is high, which can be interpreted as a preference for *global status* (see Section 3.4.3). Section 3.4.1 discusses the place of this utility function within the larger class of social preferences.

Each agent's outside option is strictly lower than $F(x_1^k, x_1^k)/2$, the utility the agent would receive in a 'same-match' (i.e., a match with a worker of the same type). Finally, note that if $x_2 = 1$ for all workers, then the model reduces to a standard Sattinger (1979) sorting model. I will refer to this case the *benchmark* and denote it by the subscript B.

3.1 The Surplus Function

The main advantage of the specific form of a KUJ utility function posited in (1), is that it renders utility *perfectly transferable*: Irrespective of the value x_2^k, x_2^j take, any increase in the worker's own wage w^k increases their utility by the same amount as it reduces the utility of their co-worker.

Given that, we can define the *surplus function* $\Pi : D_{\mathbf{x}}^2 \to \mathbb{R}$ as the sum of the utilities of the two teammates. Using the fact that $w^k + w^j = 2\bar{w}^{k,j} = F(x_1^k, x_1^j)$, one can easily show that the surplus function depends only on the teammates types, with

$$\Pi(\mathbf{x}^j, \mathbf{x}^k) \equiv 0.5F(x_1^k, x_1^j)\left(x_2^k + x_2^j\right). \tag{2}$$

Thus, the surplus of the match depends not only on the output produced, but also (negatively) on the strength of the teammates' relative concerns.

3.2 The Competitive Equilibrium

A function $\mu : D_{\mathbf{x}} \to D_{\mathbf{x}}$ is a *sorting* if it satisfies $\mu(\mu(\mathbf{x})) = \mathbf{x}$: That is, if the co-worker of $\mathbf{x}$'s co-worker is $\mathbf{x}$ themselves. A sorting μ is *feasible* if $\mu(\mathbf{X}) \sim H$, that is, if the distribution of traits implied by the sorting is the same as the actual distribution of traits; the set of all feasible sortings is denoted by $S(H)$.

All worker's take the payoff function $u(\mathbf{x}^j)$ as given. A sorting μ is *individually rational* given a payoff function u if

$$\mu(\mathbf{x}^k) = \mathbf{x}^j \Rightarrow \mathbf{x}^j \in \arg\max_{\mathbf{x}} \Pi(\mathbf{x}^j, \mathbf{x}^k) - u(\mathbf{x}^j).$$

Finally, in equilibrium it must be the case that

$$u(\mathbf{x}^k) = \max_{\mathbf{x}^j} \Pi(\mathbf{x}^j, \mathbf{x}^k) - u(\mathbf{x}^j). \tag{3}$$

Definition 1. A competitive equilibrium consists of a sorting $\mu^* : D_{\mathbf{x}}^2 \to [0, 1]$ and a payoff function $u^* : D_{\mathbf{x}} \to \mathbb{R}$, such that μ^* is feasible and individually rational given u^*, and u^* satisfies (3).

3.3 Assumptions

In order to characterise the equilibrium I will need two further assumptions. To state the first one, let me define the functions

$$L(\mathbf{x}^k; x_1', x_1) \equiv 0.5(F(x_1^k, x_1') - F(x_1^k, x_1))x_2^k, \quad V(l; x_1', x_1) = \Pr(L(\mathbf{x}; x_1', x_1) \leq l).$$

Assumption 1 (Common Rankings)**.** For any $x_1''', x_1'', x_1', x_1 \in D_{x_1}$ such that $x_1''' > x_1''$ and $x_1' > x_1$, and any $\mathbf{x}^k \in D_{\mathbf{x}}$ we have that

$$V(L(\mathbf{x}^k; x_1''', x_1''); x_1''', x_1'') = V(L(\mathbf{x}^k; x_1', x_1); x_1', x_1) \equiv v_1(\mathbf{x}^k).$$

I will refer to $v_1(\mathbf{x}^k)$ as the *skill-preference index.*

As we will see more explicitly in the proof of Theorem 2, Assumption 1 implies that there exists an index of skill and preference, such that the surplus function is supermodular only in $(x_1^k, v_1(\mathbf{x}^j))$ and $(x_1^j, v_1(\mathbf{x}^k))$. In other words, Assumption 1, implies that the strength of relative concerns will affect sorting only through the skill-preference index.[13]

To state the second assumption, I will require the concept of a bi-variate copula.

Definition 2 (Copula)**.** A *bivariate copula* is a supermodular function $C : [0,1]^2 \to [0,1]$ such that $C(0, v) = C(u, 0) = 0$, $C(1, v) = v$ and $C(u, 1) = u$.

A function $C_{\mathbf{Y}} : [0,1]^2 \to [0,1]$ is the copula of the bivariate random vector $\mathbf{Y}$ if

$$C(\Pr(Y_1 \leq y_1), \Pr(Y_2 \leq y_2)) = \Pr(Y_1 \leq y_1, Y_2 \leq y_2).$$

Sklar's Theorem (Sklar, 1959) ensures there exists a copula for every random vector, and that this copula is unique if the random vector is continuously distributed. For discrete random vectors, a continuum of copulas exists.

Assumption 2 (Properties of the Copula)**.** The distribution H and the production function F are such that there exists an *exchangeable* copula C of (X_1, v_1), that is, such a copula that $C(u, v) = C(v, u)$ for all $u, v \in [0, 1]$.

Assumptions 1 and 2 are undoubtedly strong; and yet, they are satisfied for many natural combinations of assumptions about the production function F and the traits distribution H. The following assumption describes three large classes of specifications for which Assumptions 1 and 2 are both satisfied.

Assumption 3 (Sufficient Conditions)**.** The distribution H and the production function F satisfy at least one of the following conditions:

A3.1 *Additive Production with Exchangeable Distribution:* $F(x_1^k, x_1^j) = K(x_1^k) + K(x_1^j)$ and the copula C_H of (x_1, x_2) is exchangeable, that is, $C_H(u, v) = C_H(v, u)$ for all $(u, v) \in [0, 1]$; or

[13] Assumption 1 is inspired by Assumption 1 in Gola (2021).

A3.2 *Binary Skills:* $x_1 \in \{l, h\}$, with $h > l$ and $\Pr(X_1 = h) = 0.5$; [14] or

A3.3 *Multiplicative Production with Log-Elliptical Distribution* $F(x_1^k, x_1^j) = A + ((x_1^k x_1^j)^c - 1)/c$, where $c \neq 0$ and $(\ln x_1, \ln x_2) \sim EC_2(\Delta, \Omega; \phi)$, that is, the characteristic function of the joint distribution of $(\ln x_1, \ln x_2)$ is of the form $c_X(t) \equiv \exp(it^T\Delta)\phi(t^T\Omega t)$, where $i = \sqrt{-1}$.

Each of these specifications is permissive in some dimensions, and restrictive in others. A detailed discussion of the specifications covered in Assumption 3 is provided in Section 3.4.4.

Proposition 1. Assumption 3 implies Assumptions 1 and 2.

In Section 4, I will derive general results that hold for, at least, all of the cases covered by Assumption 3. In Section 5, I will focus on the binary skills case (A3.2), as this case allows for both super- and submodular production functions while retaining tractability.

3.4 Discussion

3.4.1 The Utility Function

This paper employs the 'Keeping Up with the Joneses' (KUJ) utility function, introduced by Gali (1994), which is the standard framework for modelling cardinal status in economics (Hopkins, 2024).[15] While most applications consider consumption rather than wages, the static nature of my model makes these equivalent. The specific functional form in (1) follows Fershtman et al. (2006), who use an isomorphic specification (adding effort provision) to study sorting under relative concerns. As explained in Langtry (2023), this utility function can be also interpreted as a special case of Kőszegi and Rabin (2006)'s reference-dependent utility, where the reference point is the average within-team consumption. Finally, my utility function is also a special case of the ERC (equity, reciprocity and competition) motivation function, which makes (1) consistent with a wide range of puzzling lab results, such as the non-appropriation of full rents in ultimatum games (Bolton and Ockenfels, 2000).

The simple KUJ specification used here is needed to fully characterise sorting in settings with continuously distributed skills. As discussed in Sections 2 and 3.4.4, even with perfectly transferable utility, characterizing solutions to multidimensional assignment problems is notoriously difficult. More general utility specifications would introduce imperfectly transferable utility, making the continuous skill distribution case intractable. However, in the binary skill case, the solution can be characterised under more general KUJ specifications under a condition similar to the generalised increasing differences (GID) condition introduced by Legros and Newman (2007); this is done in Appendix B. The solution under general KUJ utility is qualitatively the same as the one to the baseline model, suggesting that my main insights are unlikely to be driven by the specific functional form chosen.

[14] If $p \equiv \Pr(X_1 = h) \neq 0.5$ then the economy consists of two completely separate sub-economies: one in which workers always same-match, as there are too few workers of the other skill for them to possibly match, and one consisting of an equal measure of low- and high-skill workers who choose whether to same- or cross-match. The first economy is trivial, and the second is isomorphic to an economy with $p = 0.5$. Thus, the assumption that $p = 0.5$ is without loss for the characterisation of equilibrium.

[15] Apart from the articles discussed in the main body, KUJ utility functions has been adopted, for example, in Clark and Oswald (1998); Ghiglino and Goyal (2010); Barnett, Bhattacharya, and Bunzel (2019) and Langtry and Ghinglino (2023).

3.4.2 Inequity Aversion

My model can also be used to study the impact that inequity aversion has on labour market sorting. Consider a simple Fehr and Schmidt (1999) utility function with heterogeneous inequity aversion:

$$U(w^k, w^s) = w^k - \alpha \max\{\bar{w}^{k,j} - w^k, 0\} - \beta \max\{w^k - \bar{w}^{k,j}, 0\}$$

Here, $\alpha \geq 0$, $\beta \in [0, 1)$—to ensure that, keeping the team's output constant, utility increases in own wage—and both α and β differ across individuals. In general, a sorting model in which workers have this utility function is different from mine. However, if skill are binary, then the inequity aversion model is isomorphic to my model if $x_2 = 1/(1 - \beta)$ for high-skill workers and $x_2 = 1/(1+\alpha)$ for low-skill workers. To understand why, note that because utility increases in own wage and output increases in skill, in the relative concerns model, the high-skill workers will always earn more than the low-skill worker in any cross-match. Thus, having $x_2 > (<)1$ is equivalent to having inequity aversion for high- (low-)skill workers. I will call a trait distribution H *inequity aversion equivalent* if $\Pr(X_2 \geq 1|x_1 = H) = \Pr(X_2 \leq 1|x_1 = L) = 1$, in which case the binary skill model is isomorphic to the inequity aversion model. This class of preferences will play an important role in the discussions about sorting (Section 4.1.5) and wage inequality (Section 4.2.3).

3.4.3 Global Status

The model features only within-firm social comparisons (local status) but no between-firm comparisons (global status). That is unrealistic: people compare themselves not only to their co-workers, but also to friends, family and acquaintances. Fortunately, restricting attention to local status is without loss of generality. To see this, let us add a global status term into each worker's utility function:

$$U(w^k, \bar{w}^{k,j}, \bar{w}; x_2^k) \equiv x_2^k w^k + (1 - x_2^k)(w^k - \bar{w}^{k,j}) + x_3^k(\bar{w}^{k,j} - \bar{w}),$$

where x_3 is worker-specific preference for global status, and $\bar{w}$ is the average economy-wide wage. The idea here is simple: Social comparisons outside of the workplace are likely based on within-firm average wage, which is more easily observable for an outsider than the worker's individual wage. Thus, people who work for a firm that pays high average wages enjoy high global status.

As any pair of workers is of measure zero, the sorting decision of any individual worker has no impact on $x_3\bar{w}$, and thus this term can be dropped. It then follows by the same logic as in Section 3.1 that

$$\Pi(\mathbf{x}^j, \mathbf{x}^k) \equiv 0.5F(x_1^k, x_1^j)\left(x_2^k + x_3^k + x_2^j + x_3^j\right)$$

In other words, we can define a new random variable $\tilde{x}_2 \equiv x_2 + x_3$, and then the model with global status becomes isomorphic to the baseline model in which workers' type is $(x_1, \tilde{x}_2)$. Thus, global status provides a compelling interpretation for $\tilde{x}_2 > 1$: Workers with $\tilde{x}_2 > 1$ are simply workers with weak relative concerns and strong global status concerns.[16]

[16] Some of the results regarding wage inequality assume that $\bar{x}_2 \leq 1$: In the light of this discussion, this assumption requires that workers have sufficiently weak global status concerns.

3.4.4 The Production Function and the Distribution of Traits

The special cases satisfying Assumptions 1 and 2 are natural and arguably larger than those in other studies that offer characterisations of two-dimensional assignment problems. As we will see in Section 4, two forces determine equilibrium sorting in this model: output maximisation and the need to match high-skill workers to co-workers with weak relative concerns. The three cases of Assumption 3 reflect these forces by either isolating them or allowing them to interact in tractable ways. While the cases do not nest each other, they are complementary: the first isolates the role of relative concerns, the second allows for general production and utility functions, and the third provides a tractable framework with continuous skills and production complementarities.

Assumption A3.1 The additively separable case isolates the effect of relative concerns on sorting by removing production complementarities. Since without complementarities aggregate production is independent of the matching pattern, this case is particularly useful for understanding how relative concerns shape labor market outcomes.

The only distributional requirement is that H's copula is exchangeable ($C_H(u,v) = C_H(v,u)$). This is not restrictive: Most commonly used bivariate copulas, including all Archimedean and elliptical copulas, are exchangeable.[17] Since this assumption is about the copula, marginal distributions remain unrestricted. Both Frank (1984a) and Fershtman et al. (2006) restrict attention to the additively separable case.

Assumption A3.2 The binary-skill case, while very restrictive in the skill dimension, allows for fully flexible production functions and trait dependence. This makes it ideal for studying how production complementarities interact with relative concerns. Its tractability enables extensions to asymmetric production functions, outsourcing decisions and general utility functions (Sections 3.4.5, 5 and Appendix B), while the economic insights from this case generalize to continuous skills.

Assumption A3.3 This case assumes a multiplicative production function and log-elliptical skill distribution. The production function $F(x_1^k, x_1^j) = A + ((x_1^k x_1^j)^c - 1)/c$ is the standard multiplicative production function commonly used in assignment models (e.g., Costrell and Loury, 2004; Tervio, 2008) but extended to accommodate not only super- ($c > 0$) but also submodularity ($c < 0$).

The log-elliptical distribution assumption is both natural and flexible. The class of log-elliptical distributions is at least as large as the class of non-negative one-dimensional random variables (Theorem 2.12 in Fang, 1990), and includes many standard distributions like log-normal, log-t-student, and (log) scale-mixtures of normals.[18] These distributions are widely used in asset pricing (e.g., Kim, 1998), portfolio choice (e.g., Owen and Rabinovitch, 1983),

[17] These are the two most commonly used families of copulas. Archimedean copulas include, among others, Clayton, Frank, Plackett, Gumbel. Elliptical copulas include the Gaussian and t-student copulas.

[18] Note that while all these examples have full support on $\mathbb{R}_{\geq 0}$, there exist bi-variate log-elliptical distributions that restrict the support to some subset of $\mathbb{R}_{\geq 0}$ with a log-elliptical boundary. Thus, one can easily create cases where $x_2 < 1$ for all workers.

and machine learning (e.g., Louizos, Ullrich, and Welling, 2017), as they combine the appealing properties of log-normal distributions with the ability to model heavy tails.

This specification provides a tractable framework that is both larger and more natural than the Gaussian-bilinear specification of Tinbergen (1956); Bojilov and Galichon (2016) and Lindenlaub (2017), which is the only specification admitting closed-form solutions in two-sided multidimensional assignment models.[19] Moreover, this specification has all the ingredients needed to bring the model to the data in future work—the log-elliptical distribution allows for enough skewness to match empirical wage distributions, and the production function allows for both super- and submodularity.

3.4.5 Asymmetric Production

On its own, the assumption that F is symmetric is without loss. To see why, let us follow Kremer and Maskin (1996) in assuming that production requires that two tasks—the key- the support-task—need to be performed. A key worker of skill x_1^k and support worker of skill x_1^s produce $\phi(x_1^k, x_1^s)$, where the function ϕ could be asymmetric. Because it is always optimal to maximise production within the team, we can then recover the production function as $F(x_1^k, x_1^s) = \max\{\phi(x_1^k, x_1^s), \phi(x_1^s, x_1^k)\}$, which is clearly symmetric.

Having said that, Assumption 1 is difficult to satisfy for production functions constructed from asymmetric ϕ functions. The reason is that if ϕ is asymmetric then the max operator necessarily makes $F(x_1^k, x_1') - F(x_1^k, x_1)$ dependent on (x_1', x_1) and I conjecture that the only continuous function that allows for such a dependence and satisfies Assumption 1 is the multiplicative production function from Assumption A3.3.

Importantly, however, the binary-skill case allows for completely arbitrary production functions including ones that are constructed from asymmetric ϕ. Indeed, the submodularity induced by the max operator in the construction of F is my main justification for focusing on submodular production functions in Section 5 (see Section 5.2 for a discussion). While I suspect that allowing for production functions constructed from asymmetric ϕ's in the other two cases would produce additional insights, much like it does in the model without relative concerns, the fact that the main results are the same in the binary and continuos cases serves as a reassurance that the symmetry of ϕ is not driving the main results.

3.4.6 Relative Concerns as Private Information

My model, in which there is complete information about workers' types, is isomorphic to a model in which only skills are public information, but preferences are private information.[20] I show this formally in Appendix A, but the intuition is simple: A worker's strength of relative concerns does not affect their co-worker's payoff; only the wage offered to the co-worker and the worker's skill do. As a result, the co-worker is indifferent between all workers of the same skill

[19] My production function is the one-dimensional equivalent of the bi-linear production function. The Gaussian distribution is a special case of the elliptical distribution—and working with logs is more natural than with levels here, as logs restricts skills to be non-negative.

[20] The question of what happens if both traits are private information is outside of the scope of this paper. While there is a growing literature concerned with this problem in the context of two-sided markets (Liu, Mailath, Postlewaite, and Samuelson, 2014; Liu, 2020, 2024), this literature remains silent on one-sided problems.

who offer them the same wage, and workers have no incentive to lie about the strength of their relative concerns.

4 Characterising the Equilibrium

In this Section, I will characterise the equilibrium sorting, payoff and wage functions, and use this characterisation to explore how the presence of relative concerns affects sorting patterns, welfare and wage inequality in comparison to the benchmark.

4.1 Equilibrium Sorting

It is well-established that in two-sided sorting problems with transferable utility the solution to the planner's problem coincides with the competitive equilibrium (Gretsky, Ostroy, and Zame, 1992). McCann and Trokhimtchouk (2010) show that the Monge-Kantorovich duality holds also for one-sided problems with transferable utility.

Theorem 1 (McCann and Trokhimtchouk (2010))**.** In a sorting model with transferable utility, sorting μ^* and a payoff function u^* constitute a competitive equilibrium if and only if μ^* solves the planner's problem:

$$\mu^* \in \arg\max_{\mu \in \mathcal{S}(H)} V_{P\mathbf{x}}(\mu), \text{ where } V_P(\mu) \equiv \int_{D_{\mathbf{x}}} \Pi(\mathbf{x}, \mu(\mathbf{x})) \mathrm{d}H(\mathbf{x}),$$

and u^* solves its dual problem, that is

$$u^* \in \arg\min_u \left(\int_{D_{\mathbf{x}}} u(\mathbf{x}) \mathrm{d}H(\mathbf{x}) \quad \text{s.t. } \forall_{(\mathbf{x}^j, \mathbf{x}^k) \in D^2_{\mathbf{x}}} u(\mathbf{x}^k) + u(\mathbf{x}^j) \geq \Pi(\mathbf{x}^j, \mathbf{x}^k) \right).$$

Thus, I can characterise the equilibrium sorting of my model by solving the planner's problem. Under Assumptions 1 and 2, the solution to the planner's problem exists and is easy to characterise.

Theorem 2 (Equilibrium Sorting)**.** Under Assumptions 1 and 2, a sorting μ^* is an equilibrium sorting if and only if induces positive and assortative matching between workers' x_1 and co-workers' v_1, that is, iff μ^* satisfies (a) $\Pr(\mu^*_1(\mathbf{x}) \leq x) = H_{x_1}(x)$ and $\Pr(v_1(\mu^*(\mathbf{x})) \leq v) = v$, and (b) $x'_1 > x_1 \Rightarrow v_1(\mu^*(\mathbf{x}')) > v_1(\mu^*(\mathbf{x}))$.

In particular, if H_{x_1} is strictly increasing, then the equilibrium sorting is given by

$$\mu^*(x_1, x_2) = [H^{-1}_{x_1}(v_1(x_1, x_2)), z(H^{-1}_{x_1}(v_1(x_1, x_2)), H_{x_1}(x_1)]^T, \tag{4}$$

where $z(x_1, \cdot)$ is the inverse of v_1 with respect to x_2, so that $z(x_1, v_1(x_1, x_2)) \equiv x_2$.

Proof. Define $\mathbf{v}(\mathbf{x}^k, \mathbf{x}^j) \equiv (v_1(\mathbf{x}^k), x^j_1)$, that is, a vector of the worker's skill-preference index and the co-worker's skill. The idea of the proof is to rewrite the planner's problem in terms of the vectors $\mathbf{v}(\mathbf{x}^k, \mathbf{x}^j), \mathbf{v}(\mathbf{x}^j, \mathbf{x}^k)$ and show that the resulting surplus function implies that x^j_1 is a complement to $v_1(\mathbf{x}^k)$ and x^k_1 is a complement to $v_1(\mathbf{x}^j)$, and that there are no other relevant complementarities or substitutabilities between the elements of $\mathbf{v}(\mathbf{x}^k,\mathbf{x}^j),\mathbf{v}(\mathbf{x}^j,\mathbf{x}^k)$. Thus, the planner wants to match workers with high x_1^k to co-workers with high v_1^j.

Formally, select an arbitrary $\tilde{x}_1 \in D_{x_1}$, and define

$$\pi(\mathbf{v}(\mathbf{x}^k,\mathbf{x}^j),\mathbf{v}(\mathbf{x}^j,\mathbf{x}^k)) \equiv V^{-1}(v_1(\mathbf{x}^k);x_1^j,\tilde{x}_1) + V^{-1}(v_1(\mathbf{x}^j);x_1^k,\tilde{x}_1)$$
$$V_{P\pi}(\mu) \equiv \int_{D_{\mathbf{x}}} \pi(\mathbf{v}(\mu(\mathbf{x}),\mathbf{x}),\mathbf{v}(\mathbf{x},\mu(\mathbf{x})))\mathrm{d}H(\mathbf{x}).$$

Note that

$$\Pi(\mathbf{x}^k,\mathbf{x}^j) = \pi(\mathbf{v}(\mathbf{x}^k,\mathbf{x}^j),\mathbf{v}(\mathbf{x}^j,\mathbf{x}^k)) + 0.5F(x_1^k,\tilde{x}_1)x_2^k + 0.5F(\tilde{x}_1,x_1^j)x_2^j.$$

Because the last two terms are additively separable in $\mathbf{x}^j,\mathbf{x}^k$, they can be added or subtracted from the surplus function with no impact on the maximiser of the planner's problem, so that

$$\max_{\mu\in\mathcal{S}(H)} V_{P\mathbf{x}}(\mu) = E_H\left((F(x_1,\tilde{x}_1))x_2\right) + \max_{\mu\in\mathcal{S}(H)} V_{P\pi}(\mu).$$

Because $\mathbf{v}(\mathbf{x}^k,\mathbf{x}^j)$ depends only on x_1^j and $v_1(\mathbf{x}^k)$, the strength of relative concerns x_2 affects a worker's match *only through its impact on the rank* $v_1(\mathbf{x}^k)$.

By construction, the mapping π is additively separable in $\mathbf{v}(\mathbf{x}^k,\mathbf{x}^j)$ and $\mathbf{v}(\mathbf{x}^j,\mathbf{x}^k)$. Therefore, the only aspects of μ that affect $V_{P\mathbf{v}}(\mu_g)$ are the bi-variate distributions of $\mathbf{v}(\mathbf{x}^k,\mathbf{x}^j)$ and $\mathbf{v}(\mathbf{x}^j,\mathbf{x}^k)$ it induces. Note that $V^{-1}(v_1;v_2,\bar{x}_1) = 0.5(F(x_1,v_2) - F(x_1,\bar{x}_1))z(x_1,v_1)$ for all x_1 in the domain of $z(\cdot,v_2)$. It follows by Assumption 1 that

$$\pi((v_1,v_2'),\mathbf{v}'') - \pi((v_1,v_2),\mathbf{v}'') = V^{-1}(v_1;v_2',\bar{x}_1) - V^{-1}(v_1;v_2,\bar{x}_1) = V^{-1}(v_1;v_2',v_1)$$

increases in v_1 and thus $\pi(\mathbf{v}(\mathbf{x}^k,\mathbf{x}^j),\mathbf{v}(\mathbf{x}^j,\mathbf{x}^k))$ is supermodular in $\mathbf{v}(\mathbf{x}^k,\mathbf{x}^j)$. Therefore, by standard results (see, for example, Theorem 4.3 in Galichon, 2016), $V_{P\mathbf{v}}(\mu_g)$ cannot reach a value higher than that achieved for sortings that satisfy (b), that is, sorting which ensure that v_1^j increases deterministically in x_1^k. Because $v_1(\mu^*(\mathbf{x}))$ strictly increases in x_1, $\mu_1^*(\mathbf{x})$ increases in $v_1(\mathbf{x})$; together with (a)—which must be satisfied for any feasible sorting—this implies that $\mu_1^*(\mathbf{x})$ depends on x_1 only through $v_1(\mathbf{x})$. It thus follows that the copula of $(\mu_1^*(\mathbf{x}),v_1(\mu^*(\mathbf{x}))$ is the same as the copula of $(v_1(\mathbf{x}),x_1)$, and thus μ^* is feasible by Assumption 2; it follows that $\mu^* \in \arg\max_{\mu\in\mathcal{S}(H)} V_{P\mathbf{x}}(\mu)$. Finally, with strictly increasing H_{x_1} only sorting μ^* satisfies (a) and (b). $\square$

In equilibrium, high-skill workers match workers with high skill-preference index. To better understand what this implies for sorting in the (x_1,x_2) space, let us work through the case of homogeneous x_2, as well as the three specifications satisfying Assumption 3.

4.1.1 Homogeneous Relative Concerns

If x_2 is the same for all workers, then $L(\mathbf{x}^k,x_1',x_1)$ depends on x_1^k only. Assumption 1 is then satisfied if and only if the production function is either strictly supermodular or strictly submodular. Under strictly supermodular production $L(\mathbf{x}^k,x_1',x_1)$ strictly decreases in x_1^k, and thus

$v_1(\mathbf{x}) = H_{x_1}(x_1)$ and we obtain positive and assortative matching (PAM) in skills; under strictly submodular production $L(\mathbf{x}^k, x_1', x_1)$ strictly decreases in x_1^k, and thus $v_1(\mathbf{x}) = 1 - H_{x_1}(x_1)$ and negative assortative matching (NAM) obtains. Therefore, the sorting patterns are exactly the same as those derived by Becker (1973) and Sattinger (1979) for the model without relative concerns. This is our first insight: sorting depends on the strength relative concerns only if workers' preferences are heterogeneous.

4.1.2 Additive Production

Under Assumption A3.1, $L(\mathbf{x}^k, x_1', x_1)$ becomes $(K(x_1') - K(x_1))x_2^k$ and thus depends only on x_2^k (positively) but not on x_1^k. Accordingly, $v_1(\mathbf{x}) = H_{x_2}(x_2)$ and

$$\mu^*(\mathbf{x}) = [H_{x_1}^{-1}(H_{x_2}(x_2)), H_{x_2}^{-1}(H_{x_1}(x_2)). \tag{5}$$

When production is additive, any sorting pattern produces the same total output. Therefore, the only aspect of sorting that matters for the social planner (and thus also in equilibrium) is the social comparisons it induces. The welfare-maximising sorting induces then negative and assortative matching between a worker's skill and the strength of their co-worker's relative concerns: A match between a high-skill worker with strong relative concerns and a low-skill worker with weak relative concerns allows the high-skill worker to enjoy his high status, without imposing much of a loss on the low-skill worker. High-skill workers with weak relative concerns and low-skill workers with strong relative concerns same-match.

4.1.3 Binary Skills

Under Assumption A3.2, $x_1 \in \{l, h\}$ where $h > l$, and thus $L(\mathbf{x}^k; h, l) = (F(x_1^k, h) - F(x_1^k, l))x_2^k$. Denote the distribution of x_2 conditional on x_1 by G_{x_1}; it follows directly from the definition of $v_1(\mathbf{x})$ that

$$v_1(\mathbf{x}) = \sum_{j \in \{l,h\}} 0.5 \left[G_j\left(x_2(F(x_1, h) - F(x_1, l))/(F(j, h) - F(j, l))\right)\right]. \tag{6}$$

By Theorem 1 any worker with $v_1(\mathbf{x}^k) > (<) 0.5$ matches a co-worker of high (low) skill. Define $\bar{y}$ such that $\bar{y} = 1$ if $a_F < G_l^{-1}(0)/G_h^{-1}(1)$, $\bar{y} = 0$ if $a_F > G_l^{-1}(1)/G_h^{-1}(0)$, and $\bar{y}$ solves $a_F = G_l^{-1}(1 - \bar{y})/G_h^{-1}(\bar{y})$ otherwise, where $a_F \equiv (F(h, h) - F(h, l))/(F(h, l) - F(l, l))$. A rearrangement of (6) yields then that in equilibrium high-skill workers with $x_2 \leq G_h^{-1}(\bar{y})$ match low-skill workers with $x_2 \geq G_l^{-1}(1 - \bar{y})$ and all remaining workers same-match.

With supermodular (submodular) production output maximisation requires positive (negative) and assortative sorting in skills. However, the need to maximise production needs to be traded-off against the desire to match high-skill workers with workers that care little about social comparisons. The outcome of this trade-off depends in general on (a) how strong the supermodularity (submodularity) of the production function is and (b) how strong are the relative concerns of high-skill workers compared to low-skill workers. High-skill workers with very strong relative concerns always match low-skill workers with very weak relative concerns. However, the definition of "strong" or "weak" relative concerns depends very much on the complementarity between workers of the same skill, as captured by a_F. In general, the stronger the complementarity, the larger the difference in the strength of relative concerns needs to be to warrant a match between a high- and a low-skill worker.

4.1.4 Multiplicative Production and Log-Elliptically Distributed Traits

Under Assumption A3.3, $L(\mathbf{x}^k, x_1', x_1) = (x_1^c - (x_1')^c)(x_1^k)^c x_2^k$; thus $v_1(\mathbf{x}^k)$ is simply equal to worker's $\mathbf{x}^k$ rank in the distribution of $\bar{v}_2 \equiv c \ln x_1 + \ln x_2$. As any linear transformation of an elliptically distributed random variable remains elliptically distributed with the same generator function ϕ (e.g. Theorem 2.16 in Fang, 1990), it follows that $(x_1, \bar{v}_2) \sim EC_2(\mathbf{A}\Delta, \mathbf{A}\Omega\mathbf{A}^T; \phi)$, where $A \equiv \left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right)$. Denote the square root of the ratio of variances of $\bar{v}_2$ and x_1 by r: it follows then from Theorem 1 and some linear algebra that:

$$\mu_1^*(\mathbf{x}) = (x_1^c x_2)^{1/r} \, e^{\delta_1(1-\frac{c}{r})+\frac{\delta_2}{r}}, \quad \mu_2^*(\mathbf{x}) = e^{(1+\frac{c}{r})[\delta_1(\frac{c}{r}-1)-\delta_2]} \left((x_1^{\frac{r^2}{c}-c})/x_2 \right)^{c/r}. \tag{7}$$

As in the binary case, if $c > 0$ ($c < 0$), and thus if production is supermodular (submodular), then the equilibrium sorting optimally trades-off the need to match high-skill workers to high- (low-) skill co-workers, with the need to match high-skill workers to workers who have weak relative concerns. There is, however, one noteworthy change compared to the binary skills case: When skills are elliptically distributed, then the measure of same-matching workers is 0 as long as Ω is of full rank. This is because, instead of same-matching, a worker with very high skill and weak (but not very weak) relative concerns can now match a worker with high (but not very high) skill and very weak relative concerns.

4.1.5 Sorting in Skills

The preceding discussion made it very clear that the direction and the strength of sorting in the skill dimension depends not only on the properties of the production function, but also on the distribution of relative concerns. Indeed, in the additive case sorting in skills is solely determined by the interdependence between skill and the strength of relative concerns—for example, sorting is perfectly positive and assortative in skills if and only if x_2 increases deterministically in x_1. More generally, whether workers of high skill are complements or substitutes matters for the direction and the strength of sorting in skill as well, just as it does in the standard sorting model—however, even then the distribution of relative concerns continues to play a (possibly dominant) role.

Proposition 2. Consider an economy (F, H) which satisfies Assumption 3. For any $\rho \in [-1, 1]$ there exists a distribution of traits $\tilde{H}$, such that (a) economy $(F, \tilde{H})$ satisfies Assumption 3, (b) the marginal distribution of skill is the same under H and $\tilde{H}$ ($H_{x_1} = \tilde{H}_{x_1}$), and (c) $\text{Corr}(X_1, \mu_1^*(X_1, X_2)) = \rho$ for every equilibrium sorting function μ^* of economy $(F, \tilde{H})$.

Proposition 2 implies that we can fix the production side of the economy—that is, the production function and the marginal distribution of skill—and yet produce any degree of sorting in skill in equilibrium, simply by altering the preference structure of the economy. This means that the observed strong positive empirical correlation between co-workers' skills (see, e.g. Freund, 2022) is consistent with a strictly submodular production function. The intuition is straightforward. Suppose that workers' skill and relative concerns are perfectly negatively correlated, which is the case, for example, if preferences are inequity aversion equivalent. A high-skill worker faces then a trade-off between maximizing their own wage by matching a low-skill worker, and minimising the within-firm wage differential by same-matching. If the relative concerns of high-skill workers are very weak compared to low-skill workers, then the welfare gain from same-matching outweighs the welfare loss stemming from the loss of output, and positive assortative matching in the skill dimension prevails.

4.2 Equilibrium Payoffs and Wages

Equation (3) and the Envelope Theorem imply that $u^*_{x_2}(\mathbf{x}) = 0.5F(x_1, \mu_1^*(\mathbf{x}))$. Therefore, the equilibrium payoff function must satisfy

$$u^*(\mathbf{x}) = 0.5\left(u^*(x_1, x_2^*) + \int_{x_2^*}^{x_2} F(x_1, \mu_1^*(x_1, s))\,\mathrm{d}s\right). \tag{8}$$

Thus, as long as for every x_1 there exists some x_2^* for which $u^*(x_1, x_2^*)$ can be determined, we would be able to derive $u^*(\mathbf{x})$. The obvious candidate for such (x_1, x_2^*) are workers who same-match in equilibrium, that is match with a co-worker of the same skill. As $\mu^*(\mu^*(\mathbf{x})) = \mathbf{x}$, co-workers of the same skill must split output equally—otherwise one of them could do strictly better by matching a worker of identical type, rather than just skill—and thus the utility of the same-matching worker $(x_1, x_2^*(x_1)$ equals $0.5F(x_1, x_1)x_2^*(x_1)$. Therefore, the equilibrium utility function can be readily derived from (8) as long as for every x_1 there exists some $x_2^*(x_1)$ such that the worker $(x_1, x_2^*(x_2))$ same-matches. While I will shy away from fully characterising the necessary and sufficient conditions for this to occur, it follows from 4.1.2-4.1.4 that this is generically the case as long as Assumption 3 is satisfied.[21]

Proposition 3 (Equilibrium Payoffs and Wages)**.** If Assumptions 1 and 2 are satisfied, and μ^* is such that for every x_1 there exists a $x_2^*(x_1)$ for which $\mu_1(x_1, x_2^*(x_1)) = x_1$, then the equilibrium payoff function and wage functions u^*, w^* are as follows:

$$u^*(\mathbf{x}) = \underbrace{0.5F(\mathbf{x_1})x_2}_{\equiv u_S(x_1) \text{ (same-match payoff)}} + \underbrace{0.5\int_{x_2^*(x_1)}^{x_2} x_2 - s\,\mathrm{d}F(x_1, \mu_1^*(x_1, s)),}_{\text{cross-matching benefit}} \tag{9}$$

$$w^*(\mathbf{x}) = \underbrace{0.5F(\mathbf{x_1})}_{\equiv w_S(x_1) \text{ (same-match wage)}} + \underbrace{0.5\int_{x_2^*(x_1)}^{x_2} 1 - s\,\mathrm{d}F(x_1, \mu_1^*(x_1, s))\,.}_{\text{cross-matching payment}} \tag{10}$$

Here, $\mathbf{x_1}$ denotes (x_1, x_1).

Proof. (9) follows from substituting $u^*(x_1, x_2^*) = 0.5F(x_1, x_1)x_2^*(x_1)$ and (4) into (8), integration by parts and some rearranging. (10) follows then from substituting $w^*(\mathbf{x}) = u^*(\mathbf{x}) + 0.5(1 - x_2)F(x_1, \mu_1^*(\mathbf{x}))$—which itself follows from (1)—into (9). □

[21] This is the case as long as the sorting function is not perfectly negative and assortative, that is, as long as $\bar{y} < 1$ under Assumption A3.2 and $x_1, \bar{v}_2$ are not perfectly negatively correlated under Assumptions A3.1 and A3.3.

The payoff of every worker can the decomposed into the same-match payoff—that is, the utility they would receive if matched with a co-worker of identical skill—and the benefit of cross-matching. As same-matching is always an option, the benefit of cross-matching must be positive. Similarly, the wage a worker receives consists of a same-match component and a cross-matching payment. Assuming $\bar{x}_2 \leq 1$, this cross-matching payment is positive for workers matched to co-workers of higher skill, as they have to receive an additional payment to endure the low within-team status. Conversely, for workers matched to less skilled co-workers the cross-matching payment is negative.

4.2.1 Wages and Relative Concerns

Keeping skill constant, workers with stronger relative concerns earn lower wages: as $v_1(\mathbf{x})$ increases in x_2, so does $\mu_1^*(\mathbf{x})$ and thus

$$w(x_1, x_2') - w(x_1, x_2) = \int_{x_2}^{x_2'} 1 - s\,\mathrm{d}F(x_1, \mu_1^*(x_1, s)) < 0,$$

as long as $1 > x_2 > x_2'$. This implies that workers with strong relative concerns earn lower wages than workers with weak relative concerns. The reason for this, perhaps, slightly counter-intuitive result is that *there is no effort provision in this model* and hence ones *relative* wage can be increased only by matching a less-skilled (and thus lower earning!) co-worker. Alas, as production increases in skill, matching a less skilled co-worker comes at the cost of decreasing the worker's absolute wage.

4.2.2 The Trickle-Down Effect of Technological Change

Changes to the production function affect payoffs and wages through two channels: directly, through changes to $F(x_1', y) - F(x_1, y)$, and indirectly, through their impact on the matching function, specifically its skill-component μ_1^*.[22] Therefore, the effect that technological change has on payoffs will depend precisely on whether it is skill-biased and on the direction in which it affects sorting.

Definition 3. A technological change is a change in the production function from some $F(\cdot, \cdot; \theta_1)$ to some $F(\cdot, \cdot; \theta_2)$. A technological change is

D3.1 *skill-biased* if, for all $x_1' \geq x_1$ and all y,

$$F(x_1', y; \theta_2) - F(x_1, y; \theta_2) \geq F(x_1', y; \theta_1) - F(x_1, y; \theta_1); \tag{11}$$

D3.2 *NAM-biased* (*PAM-biased*) if, for all $x' > x, y' > y$,

$$\Delta a_F(x', x, y', y) \equiv \frac{F(x', y'; \theta_2) - F(x, y'; \theta_2)}{F(x', y; \theta_2) - F(x, y; \theta_2)} - \frac{F(x', y'; \theta_1) - F(x, y'; \theta_1)}{F(x', y; \theta_1) - F(x, y; \theta_1)} < (>)0.$$

My definition of skill-biased technological change (SBTC) is very general and requires only that the difference in output produced by workers of higher skill increases compared

[22] $x_2^*(x_1)$ is also affected by changes in μ_1^*.

to workers of lower skill. A change in technology that decreases (increases) $a_F(x', x, y', y) \equiv (F(x', y') - F(x, y'))/(F(x', y) - F(x, y))$ is called NAM- (PAM-) biased because it makes workers of similar skill stronger substitutes (complements) and thus, as shown in Lemma 1 in Appendix A, makes sorting more negative (positive) and assortative.

Skill-biased technological change affects not only high-skill workers' but also low-skill workers' payoffs and wages. As the following Proposition shows, the gains from SBTC trickle-down to low-skill workers as long the change in technology is also NAM-biased, because in that case the direct productivity and indirect sorting channels work in unison.

Proposition 4. Under Assumption 3 and the premise of Proposition 3, any technological change that is both NAM- and skill-biased increases $u^*(\mathbf{x}) - u_S(x_1)$ (and $w^*(\mathbf{x}) - w_S(x_1)$) for all $\mathbf{x}$ such that $x_1 < H_{x_1}(0.5)$, $\mu_1^*(\mathbf{x}) > x_1$ (and $x_2 < 1$).[23]

To better understand these results, let's first examine the direct productivity channel in isolation. I will do this by focusing on the additive case (i.e., $F(x_1^k, x_1^j) = K(x_1^k) + K(x_1^j)$) where sorting patterns remain fixed since μ^* depends only on the marginal distributions of traits. Furthermore, to bring the difference between sorting with and without relative concerns into contrast, I will consider a change in K that increases $K'(x_1)$ for all x_1 above some cutoff $\hat{x}$, but leaves it unchanged otherwise.

In the presence of relative concerns, such a change in K raises the payoffs of low-skill workers with weak relative concerns, as those workers have co-workers with skill above the cut-off. Their co-workers have to, essentially, pass a part of the increase in the average wage within the team onto the low-skill worker in order to keep the match mutually beneficial. In the benchmark model, in contrast, wages and payoffs of workers with $x_1 < \hat{x}$ would be unaffected. Importantly, low-skill workers with strong relative concerns are matched to co-workers with skill below the cutoff, and thus do not see any increase in payoff. Even worse, if the cutoff $\bar{x}$ is very low, then workers with skill just above the cutoff will have low skill and yet, if they have strong relative concerns, then their wages and payoffs will increase less than in the benchmark. Thus, the welfare impact of SBTC is starkly different for low-skilled workers with weak relative concerns than for those with strong ones.

More generally, of course, the trickle-down effect of SBTC depends also on its impact on sorting. If the change in technology is NAM-biased then low-skill workers are matched with more skilled co-workers than before, which amplifies the trickle-down effect. If, however, technological change is PAM-biased, then low-skill workers end up with less skilled partners, which dampens the trickle-down effect. If the bias towards PAM is sufficiently strong, the sorting effect may well dominate the direct effect and all low-skill workers may gain less from SBTC than if they were same-matched.[24]

If present, the trickle-down effect triggered by SBTC can be very substantial indeed. For comparison, consider a policy where the government taxes the entire increase in output caused

[23] The proposition generalises naturally to cases, such as an increase in σ under Assumption A3.3 with log-normal skills, where technological change is skill-biased only in some regions of the skill-space. Specifically, consider a TC for which (11) holds only for $x_1, x_1' \in D_S \subset D_{x_1}$. In that case, $u^*(\mathbf{x}) - u_S(x_1)$ (and $w^*(\mathbf{x}) - w_S(x_1)$) increase for all $\mathbf{x}$ such that $x_1 < H_{x_1}(0.5)$, $\mu_1^*(\mathbf{x}) > x_1$, $[x_1, \mu_1^*(\mathbf{x})] \subset D_S$ (and $x_2 < 1$).

[24] By revealed preference this would be the case, for example, in a binary-skill model in which technological kept $F(\mathbf{l})$ unchanged but increased a_F above $G_l^{-1}(1)/G_l^{-1}(0)$, thus inducing perfect PAM.

by SBTC and redistributes it equally. Such a *full redistribution of gain from technological change* would increase each worker's wage by $\Delta K \equiv \int_{D_{x_1}} \Delta K(y) \mathrm{d}H_{x_1}(y)$, where $\Delta K(x_1) = K(x_1;\theta_2) - K(x_1,\theta_1)$. The following result is meant as an illustration of the scale of the trickle-down effect; to keep this illustration simple, I will restrict attention to the additive case.

Proposition 5. Under Assumption A3.1 and the premise of Proposition 3, if H_{x_2} is first-order stochastically dominated by $U[0,\bar{x}_2]$ then the worker with lowest skill and weakest relative concerns prefers no redistribution over full redistribution of technological gains, that is, $u^*(\underline{x}_1,\bar{x}_2;\theta_2) \geq u^*(\underline{x}_1,\bar{x}_2;\theta_1) + x_2\Delta K$. Similarly, if H_{x_2} is first-order stochastically dominated by $U[0,1]$, then worker receives a greater increase in income with no redistribution than under full distribution of technological gains, that is, $w^*(\underline{x}_1,\bar{x}_2;\theta_2) - w^*(\underline{x}_1,\bar{x}_2;\theta_1) \geq \Delta K$.

If most workers have much stronger relative concerns (so lower values of x_2) than the worker with lowest skill and weakest relative concerns, then the outside options of high-skill workers with strong relative concerns are weak and there is a lot of demand to match with $(\underline{x}_1,\bar{x}_2)$. Jointly, this implies that this worker can appropriate a substantial portion of the gains from SBTC—and because she has weak relative concerns, she does not mind the increase in the gap between her own and her co-workers' wage all that much.

It is worth noting that even though $(\underline{x}_1,\bar{x}_2)$ has the lowest skill, she does not earn particularly low wages, at least as long as $\bar{x}_2 \leq 1$. It is very easy to see that if $\bar{x}_2 \leq 1$, then $(\underline{x}_1,\bar{x}_2)$ prefers no redistribution only if she earns an above average wage—after all, full redistribution would give her an average wage without having to suffer the ignomity of earning less than the average wage in her firm. If $\bar{x}_2 > 1$ instead—that is, if $(\underline{x}_1,\bar{x}_2)$ cares more about global than within-firm status—then she could be earning arbitrarily low wages and nevertheless oppose redistributive policies.

4.2.3 Wage Inequality

Intuitively, as long as all workers dislike earning less than their co-worker (i.e., $\bar{x}_2 \leq 1$), wage inequality should be lower in the presence of relative concerns than if all workers received the same-match wage, simply because any low-skill workers who are not same-matched must earn higher—and any high-skill workers lower—wages than the same-match wage. Proposition 6 confirms this intuition for any specification of the model which satisfies Assumption 3.[25]

Proposition 6. Under Assumption 3, if $\bar{x}_2 \leq 1$, then $\mathrm{Var}(w_S) \geq \mathrm{Var}(w^*)$.[26]

It follows immediately, that—under the conditions imposed by Proposition 6—wage inequality in my model must be lower than in the benchmark as long as the benchmark wages are more unequal than the same-match wages, that is, as long as $\mathrm{Var}(w_B) \geq \mathrm{Var}(w_S)$. By standard results, this is always the case if production is either supermodular or additive, as then $w_B(\mathbf{x}) = w_S(\mathbf{x})$.

If production is submodular, however, sorting is negative and assortative in the benchmark and, by revealed preference, workers' wages are higher than the same-match wage. The sign of

[25]This intuition could fail if skills were binary and $p > 0.5$, because the presence of relative concerns would increase the inequality between the two sub-economies discussed in footnote 14.

[26]In the case of log-elliptical distributions in which $x_1 > 1$ with a positive probability, such as the log-normal, the same conclusion holds as long as $H_{x_2}(1)$ is sufficiently close to 1.

$\mathrm{Var}(w_B) - \mathrm{Var}(w_S)$ is then ambiguous and depends (only) on the production function and the distribution of workers' skill. This is easiest to see when skills are binary, in which case wages are not unique in the benchmark and any wage structure of the form

$$w_B(x_1) = 0.5(F(\mathbf{x_1}) + \alpha_{x_1}(2F(h,l) - F(\mathbf{h}) - F(\mathbf{l})),$$

is sustainable in equilibrium. Here, $\alpha_{x_1} \in [0,1]$ denotes the bargaining power of workers with skill x_1, and $\alpha_l + \alpha_h = 1$.[27] Thus, in the binary skills case $\mathrm{Var}(w_B) - \mathrm{Var}(w_S) = 0.25(2\alpha_l - 1)(2F(h,l) - F(\mathbf{h}) - F(\mathbf{l}))$, which is positive if and only if $\alpha_l \geq 0.5$. In other words, the benchmark wage distribution is less unequal than the same-match one if and only if low-skill workers have a stronger bargaining position in the benchmark.[28]

Recall that Proposition 2 implies that if low-skill workers have sufficiently stronger relative concerns than high-skill workers then same-matching obtains in equilibrium. Thus, if the benchmark wage distribution is indeed less unequal than the same-match one, then the presence of relative concerns may well increase wage inequality in comparison to the benchmark.

Corollary 1. Consider an economy (F, H) which satisfies Assumption 3.
(i) If $\mathrm{Var}(w_B) < \mathrm{Var}(w_S)$, then there exists a distribution of traits $\tilde{H}$, such that (a) economy $(F, \tilde{H})$ satisfies Assumption 3, (b) the marginal distribution of skill is the same under H and $\tilde{H}$ ($H_{x_1} = \tilde{H}_{x_1}$), and (c) $\mathrm{Var}(w_B) < \mathrm{Var}(w^*)$.
(ii) Suppose that, in addition, (F, H) satisfies the premise of Proposition 6. If $\mathrm{Var}(w_B) \geq \mathrm{Var}(w_S)$, which is always satisfied for supermodular F, then $\mathrm{Var}(w_B) \geq \mathrm{Var}(w^*)$.

Corollary 1(i) follows from Proposition 2 and the properties of the same-match and benchmark wages, neither of which depend on the fact that all workers have $x_2 \leq 1$. Corollary 1(ii), in contrast, follows from Proposition 6 and thus requires this additional assumption.

Remark 1. In the binary skill case, inequity aversion equivalent distributions of preferences (see 3.4.2) are exactly of the form that pushes sorting to be more positive and assortative in skills. In particular, it follows immediately from the discussions in 3.4.2 and 4.1.3 that if skills are binary and, for example, all workers have Fehr and Schmidt (1999) utility with $\beta \in [1 - a_F, 1]$, then everyone will same-match in the equilibrium of the model with inequity aversion.[29] This implies that if $a_F < 1$ and low-skill workers have higher bargaining power than high-skill workers ($\alpha \geq 0.5$)—and thus the benchmark distribution of wages is less unequal than the same-match one—then *sufficiently strong inequity aversion increases wage inequality in the economy.*

[27] Workers' bargaining power does not matter in this model as long as some workers same-match, because competition reduces the bargaining set to a singleton. If, however, production is submodular and skills are binary, then there is no same-matching in the benchmark and the bargaining set is not a singleton anymore.

[28] The sign of $\mathrm{Var}(w_B) - \mathrm{Var}(w_S)$ remains ambiguous also when benchmark wages are uniquely determined. To see this, consider the case of multiplicative production and log-normally distributed skills. In that case, the benchmark and same-match wages become $w_B(x_1) = (Ac + \exp(2c\delta_1) - 1)/(2c) + \exp(2c\delta_1)(\ln x_1 - 2\delta_1)$ and $w_S(x_1) = A + (x_1^{2c} - 1)/c$ and thus $\mathrm{Var}(w_B) - \mathrm{Var}(w_S) = \exp(4\delta_1 c)\left(\omega_{11} - \exp(2\omega_{11}) + \exp(\omega_{11})\right)$. Clearly, this expression (a) equals to 0 and has a positive derivative for $\omega_{11} = 0$ and (b) is concave in ω_{11}. It follows, therefore, that there exists some $\bar{\omega}_{11}$ such that if $\bar{\omega}_{11} < \omega_{11}$ then $\mathrm{Var}(w_B) > \mathrm{Var}(w_S)$ and if $\bar{\omega}_{11} > \omega_{11}$ then $\mathrm{Var}(w_B) < \mathrm{Var}(w_S)$.

[29] This is a sufficient rather than a necessary condition. More generally, same-matching obtains if for all workers $a_F > \frac{1-\underline{\beta}_h}{1+\underline{\alpha}_l}$, where $\underline{\beta}$ is the lowest value of β among high-skill workers and $\underline{\alpha}_l$ is the lowest value of α among low-skill workers.

This rather striking result seems to capture a mechanism that may hold much more broadly than just in the labour market: The desire to minimise within-group (here, within-firm) inequality, may push agents to sort with agents who are similar to them. While this indeed eliminates inequity within-groups it *maximises* inequality between-groups—and if the structure of the economy is such that between-group inequality is a greater concern than within-group inequality, it may well increase overall inequality.

4.2.4 Welfare

In order to assess the welfare effects of social comparisons, suppose that the reference point to which each worker compares themselves depends on the intensity of social interactions p, with $\bar{w}^{k,j} = (1-p)w^k + 0.5pF(x_1^k, x_1^j)$. To fix ideas, suppose that $p = 0$ corresponds to a case where the production process is fully remote and anonymous, so that the workers do not know who their co-workers are, and thus in calculating the reference point take only their own wage into account. In other words, with $p = 0$ workers have no-one else to compare themselves to. In that case the utility function of worker $\mathbf{x}^k$ reduces to $x_2^k w^k$. Given the absence of social comparisons, sorting and wages must be as in the benchmark, and thus the payoff of worker $\mathbf{x}^k$ equals $x_2^k w_B(x_1^k)$ in equilibrium. The case of $p = 1$ corresponds to fully in-office, team-work based production, where the reference point is the within-firm average wage. This setup allows us to whether an increase in the intensity of interactions from $p = 0$ to $p = 1$ is welfare-improving.

Proposition 7. Suppose that Assumptions 1 and 2 are satisfied.
(i) If F is supermodular, then $\Pr(u^*(X_1, X_2) \geq X_2 w_B(X_1)) = 1$. If, in addition, Assumption A3.1 or A3.3 is satisfied and H has full support, then $\Pr(u^*(X_1, X_2) > X_2 w_B(X_1)) = 1$.
(ii) If F is strictly submodular and there exists some $\mathbf{x}'$ such that $\mu_1(\mathbf{x}') = x_1'$, then $\Pr(u^*(X_1, X_2) \geq X_2 w_B(X_1)) < 1$.

The welfare impact of social comparisons hinges on the properties of the production function. Under supermodular production, same matching always guarantees the benchmark wage, regardless of how intense social interactions are—by revealed preference workers must therefore benefit from social comparisons. Under submodular production, the negative and assortative sorting pattern that prevails in the benchmark allows workers to earn higher wages than the same-match payoff: If, therefore, under $p = 1$ some workers decide to same-match in order to avoid negative social comparisons, then they must be worse off than they would be under $p = 0$. This indicates that if each team could choose p, then all teams would choose $p = 1$ under supermodular production, but some would choose $p = 0$ under submodular production. This observation forms the basis of the 'theory of the firm' outlined below.

5 Outsourcing

In this Section, I show that the interaction between skill-biased technological change and relative concerns explains the marked increase in domestic outsourcing (Goldschmidt and Schmieder, 2017; Bergeaud et al., 2024). To do this, I first need to allow the teams to choose where to draw the boundary of the firm. As this extension makes the model significantly less tractable than the baseline (e.g. Assumptions 1 and 2 become difficult to satisfy), I restrict attention to the binary-skill case—which remains simple enough to solve—throughout.

The basic premise is very simple: Wage comparisons weigh lighter in agents' utility when they happen across firm boundaries. In other words, the co-worker's high wage bothers the worker less if the worker is a subcontractor rather than a subordinate. More specifically, suppose that each matched team has the option of *outsourcing*, that is, forming two separate firms instead of one. As is standard in the theory of the firm literature, outsourcing comes at a cost $c \geq 0$: Contracts need to be written, there is additional accounting, etc. The possible advantage of outsourcing, however, is that each of the new firms consists of a single worker, so that $\bar{w}^{k,j} = w^k$ and thus $u(w^k, \bar{w}^{k,j}; x_2^k) = x_2^k w^k$. If the team decides not to outsource, then each co-worker's utility is as in the baseline model.

Proposition 8. Suppose that Assumption A3.2 is satisfied, and denote by s_F the loss of output resulting from same-matching, with $s_F \equiv F(h,l) - 0.5(F(\mathbf{h}) + F(\mathbf{l}))$.
(i) If $c > s_F$, then there is no outsourcing and the equilibrium is as described in 4.1.3.
(ii) If $c \in [0, s_F]$, then all teams formed in any equilibrium are between a high- and a low-skill worker. Define y^o as the solution to

$$\frac{G_l^{-1}(1-y^o)}{G_h^{-1}(y^o)} = \min\{\frac{G_l^{-1}(1)}{G_h^{-1}(0)}, \max\{\frac{G_l^{-1}(0)}{G_h^{-1}(1)}, b_F\}\}. \tag{12}$$

where $\alpha_l \in [0,1]$ denotes the bargaining power of low-skill workers and

$$b_F \equiv 1 - \frac{2c}{(1-\alpha_l)(F(h,l) - F(\mathbf{l})) + \alpha_l(F(\mathbf{h}) - F(h,l)) + 2\alpha_l c}.$$

Low-skill workers are outsourced iff their $x_2 < G_l(1-y^o)$. Outsourced low-skill workers match with high-skill workers of $x_2 > G_h^{-1}(y^o)$, the non-outsourced low-skill workers match with high-skill workers of $x_2 \leq G_h^{-1}(y^o)$.

In the baseline model, welfare-reducing social comparisons can be avoided only through same-matching. Outsourcing provides an alternative way of opting-out from these comparisons. For that alternative to be used, it must be cheaper than same-matching. As same-matching is output-maximising under supermodular production, outsourcing can happen only if production is submodular ($s_F > 0$), and the cost of outsourcing is low compared to the loss of production stemming from same-matching ($c < s_F$).

To understand which teams outsource, consider the *marginal team*, that is a team which is indifferent between forming one or two firms. If outsourcing comes at a cost, their indifference implies that social comparisons are costly within that team, and thus the high-skill worker has weaker relative concerns than the low-skill worker. In the extreme case of cost-less outsourcing, the high- and low-skill workers forming the marginal team have equally strong relative concerns. Of course, the high-skill workers forming (non-)outsourcing teams have weaker (stronger) relative concerns than the high-skill worker in the marginal team; and *vice versa* for the low-skill workers.

5.1 The Impact of SBTC on Outsourcing and Sorting

Let as focus on the case in which outsourcing and non-outsourcing teams co-exist; that is, I assume $c < s_F$ and H is such that $y^o \in (0, 1)$. It follows directly from (12) in Proposition 8 that SBTC decreases y^o and thus raises the number of outsourcing teams. In other words, as long as any jobs were outsourced initially, skill-biased technological change will cause more outsourcing. To understand the intuition, recall that in the marginal team the high-skill worker has stronger relative concerns than the low-skill worker. SBTC further increases the inequality within that team—and with that the welfare loss from social comparisons. As a result, the marginal team now strictly prefers to outsource, and the number of outsourcing teams increases. Furthermore, the wages of the newly outsourced low-skill workers fall in comparison to non-outsourced low-skill workers, which is consistent with the empirical findings from Goldschmidt and Schmieder (2017) and Bergeaud et al. (2024).

If $c < s_F$ then all production teams consist of one high- and one low-skill worker, and thus SBTC has no effect on how workers sort into *production teams*. Crucially, however, SBTC does affect how workers sort into *firms*, because every outsourcing team consists of two single-worker firms! Thus, for an econometrician who observes the composition of firms but not teams, workers from outsourcing teams are sorted positively and assortatively, whereas workers in non-outsourcing firms are negatively sorted. It follows, therefore, that increase in outsourcing caused by SBTC results in workers sorting more positively into firms.

5.2 Discussion

Social Comparisons Across Firm Boundaries I assumed that social comparisons within a production team are much weaker if that team is split into two firms. Nickerson and Zenger (2008) attribute this weakening of social comparisons across firm boundaries to the salience of within-firm comparisons, and to within-firm competition for resources. They also provide a number of persuasive case studies in which the firm boundary mattered critically for the strength of social comparisons. Another justification of this assumption can be derived from Coase (1937), who hypothesised that some people like to direct others, and some like to be directed, and differentiated between 'employees' and 'subcontractors' precisely by the degree to which their work is directed. Under this interpretation, co-workers in outsourcing teams work together, but—in contrast to a non-outsourcing firm—none of them is directed by the other, and thus social comparisons matter less.

Theory of the Plant vs. Firm A compelling feature of this extended model is that it provides both a 'theory of plant' (sorting into production teams) and 'theory of firm' (a team's decision whether to form one or two firms), and that production equivalent 'plants' draw their firm boundaries differently. Furthermore, the 'plant'- and firm-formation decisions interact in this model. As outsourcing becomes viable, a high-skill worker who would have previously same-matched, switches to having an outsourced low-skill co-worker. This implies that having the option of cheaply redrawing the boundary of the firm affects what "plants" are formed.

Submodular Production The condition $c \in (0, s_F)$ is satisfied only when the production function is submodular. This is potentially problematic, because supermodular production functions are more commonly assumed in the sorting literature. However, this is largely because the empirical correlation in the level of co-workers skill is large and positive (see Figure 1(b) in Freund, 2022, for example), a fact which in standard models can be reproduced only with supermodular production. In my model, however, submodular production is perfectly consistent with positive and assortative matching in skills, as long as low-skill workers have stronger relative concerns than high-skill workers (see Section 4.1.5). Furthermore, (locally) submodular production has been convincingly microfounded by Kremer and Maskin (1996) as the by-product of workers' self-selection into roles within the firm, and more recently by Boerma, Tsyvinski, and Zimin (2021) as the outcome of within-team problem solving.

Lower Cost of Outsourcing An obvious alternative explanation for the trends in outsourcing, sorting and inequality is a decrease in the cost of outsourcing. It follows immediately from Proposition 8 that this would have the same qualitative impact on outsourcing and sorting as SBTC.

Inequity Aversion Inequity aversion equivalent preferences satisfy the premise of Proposition 8. Thus, if the cost of outsourcing is neither too high nor too low (so that both outsourcing and non-outsourcing teams co-exist), then the impact of SBTC on outsourcing and sorting under inequity aversion is the same as under relative concerns.

Remote Work Instead of outsourcing, teams could escape detrimental social comparisons by re-organising production. For example, teams could choose to work remotely, which would decrease the intensity of social interactions. The results about outsourcing can be freely reinterpreted as results about the prevalence of remote work: That is, SBTC would make remote work more common. The results about sorting, however, would change under this reinterpretation: When firms avoid social comparisons by re-organising production rather than by outsourcing, then measured and real sorting coincide, and thus SBTC has no impact on measured sorting.

6 Concluding Remarks

In this paper, I develop a one-sided assignment model in which workers differ in skill and the strength of their relative concerns. While this heterogeneity makes the problem naturally two-dimensional, I am able to fully characterise the equilibrium for a large class of cases by leveraging the fact that the distribution of traits of 'workers' must be the same as that of 'co-workers' in equilibrium.

As utility is transferable, equilibrium sorting optimally trades off output maximisation with the need to maximise the welfare gain stemming from within-team social comparisons. This produces several key results: sorting can be positive (negative) assortative in skill even when production is submodular (supermodular), and the benefits of skill-biased technological change trickle down specifically to low-skill workers who care little about relative status. Indeed, when the overall level of relative concerns in the population is sufficiently high, these low-skill workers may earn above average wages despite their low skill level.

The welfare implications depend critically on the production technology. Under supermodular production, compared to a model without social comparisons, all workers are better off and wage inequality is lower. Under submodular production, however, the presence of heterogeneous relative concerns may increase wage inequality and harm both high-skill workers with low relative concerns and low-skill workers with strong relative concerns.

Building on that final insight, I argue that skill-biased technological change may have driven the observed increase in domestic outsourcing. Following Nickerson and Zenger (2008), I assume that the salience of social comparisons weakens if one of the team-members is outsourced. If that is the case, then teams consisting of high-skill workers with low relative concerns and low-skill workers with strong relative concerns would like to outsource the low-skill worker, even though outsourcing is costly. Skill-biased technological change increases within-team inequality and thus increases the cost of keeping the low-skill worker in-house for such teams; as a result, the number of outsourcing teams increases.

A Omitted Proofs and Derivations

Proof of Proposition 1 I will show that each of A3.1-A3.2 implies Assumptions 1 and 2.

A3.1: Assumption 1 is satisfied trivially, with $v_1(x_1^j, x_2^j; x_1, x_1') = H_{x_2}(x_2)$. The copula of x_1, v_1 and x_1, x_2 are the same, and thus Assumption 2 is satisfied as well. A3.2: Assumption 3 $\Rightarrow$ Assumption 1 immediately, because if the distribution of skill is binary, then there are only two levels of skill.

I will show that 3 $\Rightarrow$ Assumption 2 by constructing copula C_h of x_1, v_1 that satisfies Assumption 2. First, define the functions

$$
\begin{aligned}
&G(v) \equiv \Pr(v_1 \le v | X_2 = L), \quad Z(v) \equiv \Pr(v_1 \le v | X_2 = H) = 2v - G(v),\\
&C^1(u,v) = \frac{G(\min\{u,0.5\})G(\min\{v,0.5\})}{2G(0.5)},\\
&C^2(u,v) = \frac{G(\max\{u,0.5\}) - G(0.5)}{2}\frac{Z(\min\{v,0.5\})}{Z(0.5)},\\
&C^3(u,v) = \frac{(Z(\max\{u,0.5\}) - Z(0.5))(Z(\max\{v,0.5\}) - Z(0.5)}{2(1-Z(0.5))}.
\end{aligned}
$$

The candidate copula is then

$$
C_h(x_1, v_1) \equiv C^1(x_1, v_1) + (C^2(x_1, v_1) + C^2(v_1, x_1)) + C^3(x_1, v_1). \tag{13}
$$

C_h is symmetric because $C^1(x_1, v_1), C^3(x_1, v_1)$ and $C^2(x_1, v_1) + C^2(v_1, x_1)$ are all symmetric. It is a copula (a) because $\frac{\partial}{\partial v_1}\frac{\partial}{\partial v_1}C_h(x_1, v_1)$ exists almost everywhere, and is strictly positive wherever it exists, and (b) by symmetry and the facts that $C_h(x_1, 0) = 0$, $C_h(x_1, 1) = x_1$. What remains to be shown is that $C_h(H_{x_1}(x_1), v_1) = \Pr(X_1 \le x_1, v_1 \le v_1)$. This is true for $x_2 = h$ by the definition of a copula, and follows for $x_2 = l$ by inspection of (13) and the fact that $\Pr(X_1 \le l, v_1 \le v_1) = 0.5G(v)$.

A3.3: An important property of the elliptical family of distributions, is that any linear transformation of an elliptically distributed random variable is elliptically distributed and has the same generator function ϕ (Theorem 2.16 in Fang, 1990). Formally, if $\mathbf{z} \sim EC_n(\Delta_z, \Omega_z, \phi)$ with $\text{rank}(\Omega_z) = k$, $\mathbf{B}$ is an $m \times n$ matrix and $\boldsymbol{p}$ is an $m \times 1$ vector, then $\boldsymbol{p} + \mathbf{B}\mathbf{z} \sim EC_m\left(\boldsymbol{p} + \mathbf{B}\Delta_z, \mathbf{B}\Omega_z\mathbf{B}^T, \phi\right)$. Define the random variable $\bar{\mathbf{V}} = \mathbf{A}(\ln X_1, \ln X_2)^T$, where $A \equiv \left(\begin{smallmatrix}1 & 0\\ c & 1\end{smallmatrix}\right)$. Clearly, $v_1(x_1^j, x_2^j; x_1, x_1') = \Pr(\bar{v}_1 \le c\ln x_1 + \ln x_2)$. Assumption 1 is thus satisfied; Assumption 2 is satisfied as well, because the copula of (X_1, v_1) is also the copula of $\bar{\mathbf{V}}$, and copulas of log-elliptical random variables are exchangeable.

The Argument from 3.4.6 Truth-telling about ones preference is incentive compatible under the sorting and payoff functions that hold in the competitive equilibrium. To see this, suppose that x_1 is perfectly observable, but x_2 is not. Workers first announce some $\hat{x}_2$, and after that sorting commences with $(x_1, \hat{x}_2)$ treated as each workers true type. Finally, on top of the requirements specified in Definition 1, we impose the *truth-telling condition*:

$$\begin{aligned} u^*(x_1, x_2) &= \max_{\hat{x_2}} w^*(x_1, \hat{x_2}) - 0.5(1 - x_2)F(x_1, \mu_1^*(x_1, \hat{x_2})) \\ &= \max_{\hat{x_2}} 0.5(1 + x_2)F(x_1, \mu_1^*(x_1, \hat{x_2})) - w^*(\mu(x_1, \hat{x_2})), \end{aligned} \tag{14}$$

where $w^*(\mathbf{x}) = u^*(\mathbf{x}) + 0.5(1 - x_2)F(x_1, \mu_1^*(\mathbf{x}))$. In other words, worker $\mathbf{x}$ is free to match with any co-worker of a worker with skill x_1 as long as they pay them the equilibrium wage; and, of course, truth-telling requires that the utility maximising choice is the one that corresponds to their true x_2. Critically, however, every worker was equally free to do so under complete information! Formally, we have that

$$\Pi(\mathbf{x}^j, \mathbf{x}^k) - u(\mathbf{x}^j) = 0.5F(x_1^k, x_1^j)(1 + x_2^k) - w^*(\mathbf{x}^j)$$

and thus individual rationality implies (14). Therefore, the requirement of truth-telling does not change the equilibrium conditions.

Proof of Proposition 2 It suffices to show that there exists a parameterised family of joint distributions $\tilde{H}^\rho$ that (i) satisfies conditions (a) and (b) and (ii) induces a family C^ρ of copulas of the joint distributions of $(X_1, \mu_1^*(X_1, X_2))$ which depends continuously on the parameter ρ and nests both Fréchet–Hoeffding upper and lower bounds.[30] This is because X_1 and $\mu_1^*(X_1, X_2)$ have the same marginal distributions, and thus the Fréchet–Hoeffding upper (lower) bound copula produces $\mathrm{Corr}(X_1, \mu_1^*(X_1, X_2))$ equal to 1 (-1).

If (F, H) satisfies Assumption A3.1, then $\mathrm{Corr}(X_1, \mu_1^*(X_1, X_2)) = \mathrm{Corr}(X_1, H_{x_1}^{-1}(H_{x_2}(X_2)))$, and the result follows by setting $\tilde{H}^\rho$ to be the family of Gaussian copulas. If (F, H) satisfies Assumption A3.2, let us set $\tilde{H}^\rho$ so that $G_h^{-1}(y) = 1/(2 - y + (1 - \rho)a_F)$ and $G_l^{-1}(y) = a_F/((2 - y)a_F + (1 + \rho))$, which is clearly continuous in ρ and, by the discussion in 4.1.3, attains the Fréchet–Hoeffding lower (upper) bound for $\rho = -1$ $(\rho = 1)$.

Finally, if (F, H) satisfies Assumption A3.3, then set $\tilde{H}^\rho$ to be $EC_2(\Delta, \Omega(\rho); \phi)$ distributed, where $\omega(\rho)_{11} = \omega_{11}$, $\omega(\rho)_{22} = 4c^2\omega_{11}$, $\omega(\rho)_{12} = \omega(\rho)_{21} = -\rho\sqrt{\omega(\rho)_{11}\omega(\rho)_{22}}$, and $\rho \in [-1, 1]$. It follows from the proof of Proposition 1 that $\ln X_1, \ln \mu_1^*(X_1, X_2)$ is EC_2 distributed, with both marginals equal to H_{x_1} and

$$\mathrm{Corr}(\ln X_1, \ln \mu_1^*(X_1, X_2)) = \mathrm{sgn}(c)\frac{1 - 2\rho\mathrm{sgn}(c)}{\sqrt{(1 - 2\mathrm{sgn}(c))^2 + 4\mathrm{sgn}(c)(1 - \rho)}}.$$

Therefore, the copula of $X_1, \mu_1^*(X_1, X_2)$ depends continuously on ρ, and for $\rho = 1$ $(\rho = -1)$ $\mathrm{Corr}(\ln X_1, \ln \mu_1^*(X_1, X_2)) = -1 (= 1)$ and thus the copula of $X_1, \mu_1^*(X_1, X_2)$ reaches the Fréchet–Hoeffding lower (upper) bound.

Proof of Proposition 4 Let me start by showing the impact of NAM-biased TC on μ^*.

[30] In the binary skills case, this suffices because the correlation between workers' and their co-workers' skill depends on $\bar{y}$ only, which is the same for all possible equilibrium sortings.

Lemma 1. Under Assumption 3, PAM-biased technological change increases (decreases) $\mu^*(\mathbf{x})$ for all $x_1 > (<) H_{x_1}(0.5)$. Conversely, NAM-biased technological change increases (decreases) $\mu^*(\mathbf{x})$ for all $x_1 < (>) H_{x_1}(0.5)$.

Proof. In the additively separable case, $\mu^*(\mathbf{x})$ is a function of marginal distributions only, and thus any technological change that respect Assumption A3.1 leaves $\mu^*(\mathbf{x})$ unchanged and thus weakly increases it.

In the binary case, sorting depends on production through a_F only. Clearly, PAM-biased technological change increases a_F and NAM-biased technological change decreases it, so that the result follows directly from 4.1.3.

In the multiplicative case, the only parameter affecting production is c. It is easy to verify that any change in technology that decreases (increases) c is NAM-biased (PAM-biased). Elementary algebra reveals that $\frac{\partial}{\partial c}\ln(\mu_1(\mathbf{x})) = c(\ln(x_1) - \delta_1)/r$, and the result follows because the (elliptical) marginal distribution of $\ln x_1$ is symmetric, so that δ_1 is the median skill.

$\square$

Define $\tilde{z}(\cdot; x_1)$ as the left-inverse of $\mu_1^*(x_1; \cdot)$, with $\tilde{z}(y; x_1) \equiv \inf\{x_2 \in D_{x_2} : \mu_1^*(x_1, x_2) \geq y\}$.[31] Integrating by substitution transforms (9) and (10) into:

$$u^*(\mathbf{x}) = 0.5F(\mathbf{x_1})x_2 + 0.5\int_{x_1}^{\mu_1^*(x_1,x_2)} x_2 - \tilde{z}(y; x_1)\,\mathrm{d}F(x_1, y), \tag{15}$$

$$w^*(\mathbf{x}) = 0.5F(\mathbf{x_1}) + 0.5\int_{x_1}^{\mu_1^*(x_1,x_2)} 1 - \tilde{z}(y; x_1)\,\mathrm{d}F(x_1, y). \tag{16}$$

As $\mu_1^*(x_1, \cdot)$ is increasing (possibly weakly), so is $\tilde{z}(y; x_1)$. It follows from the definition of $\tilde{z}$ that $\tilde{z}(\mu_1^*(x_1, x_2), x_1) \leq x_2$ and that $\tilde{z}(y; x_1)$ weakly decreases if $\mu_1^*(x_1, x_2)$ has increased. Thus, by inspection of (15) that if $\mu_1^*(\mathbf{x}) > x_1$ then skill-biased technological change increases $u^*(\mathbf{x}) - u_S(x_1)$ as long as it increases $\mu_1^*(\mathbf{x})$; then the result for $u^*(\mathbf{x}) - u_S(x_1)$ follows from Lemma 1. The result for $w^*(\mathbf{x}) - w_S(x_1))$ follows from an analogous reasoning and the fact that if $x_2 < 1$ then $\tilde{z}(y; x_1) < 1$ for all $y \in [x_1, \mu_1^*(\mathbf{x})]$.

Proof of Proposition 5 Under Assumption A3.1, (15) and (16) simplify to

$$u^*(\mathbf{x}) = 0.5F(\mathbf{x_1})x_2 + \int_{x_1}^{\mu_1^*(x_1,x_2)} x_2 - H_{x_2}^{-1}(H_{x_1}(y))\,\mathrm{d}K(y), \tag{17}$$

$$w^*(\mathbf{x}) = 0.5F(\mathbf{x_1}) + \int_{x_1}^{\mu_1^*(x_1,x_2)} 1 - H_{x_2}^{-1}(H_{x_1}(y))\,\mathrm{d}K(y). \tag{18}$$

Integrating ΔK by parts and rearranging yields

$$u^*(\underline{x}_1, \bar{x}_2; \theta_2) - u^*(\underline{x}_1, \bar{x}_2; \theta_1) - x_2\Delta K = \int_0^1 \frac{\partial}{\partial x_1}\Delta K(H_{x_1}^{-1}(s))(\bar{x}_2 s - H_{x_2}^{-1}(s))\mathrm{d}H_{x_1}^{-1}(s)$$

$$w^*(\underline{x}_1, \bar{x}_2; \theta_2) - w^*(\underline{x}_1, \bar{x}_2; \theta_1) - \Delta K = \int_0^1 \frac{\partial}{\partial x_1}\Delta K(H_{x_1}^{-1}(s))(s - H_{x_2}^{-1}(s))\mathrm{d}H_{x_1}^{-1}(s).$$

[31] Of course, if x_1 is continuous, then $\tilde{z} = z$.

The results follow then from the definition of SBTC and the fact that $U[0,x]$ first-order stochastically dominates H_{x_2} if and only if $xs \geq H_{x_2}^{-1}(s)$ for all $s \in [0,1]$.

Proof of Proposition 6 **Assumption A3.1**.

Notice that, by (18), (a) the wage function and the average wage in the economy depend only on the marginals of the traits distribution, but not its copula and (b) if the copula of H is the Fréchet–Hoeffding upper bound, then each worker receives exactly half of the production of a same-matched team, which is the benchmark wage. Thus, wage variance is lower in my model than in the baseline as long as $\int_{D_{\mathbf{x}}} w(\mathbf{x})^2 \mathrm{d}H(\mathbf{x}) < \int_{D_{\mathbf{x}}} w(\mathbf{x})^2 \mathrm{d}\bar{C}(H_{x_1}(x_1), H_{x_2}(x_2))$, where $\bar{C}(\mathbf{v}) \equiv \min\{v_1, v_2\}$ is the Fréchet–Hoeffding upper bound copula. This is clearly true—by the definitions of the Fréchet–Hoeffding upper bound and the supermodular order—because the square function is convex and $w(\mathbf{x})$ is additevely separable and (under the assumption that $\underline{x}_2 \leq 1$) increases in both variables, and thus $w(\mathbf{x})^2$ is supermodular.

Assumption A3.2. If $\bar{y} = 0$ then $w_S = w^*$ and the result is immediate. If $\bar{y} = 1$ and $\bar{x}_2 \leq 1$ then all low- (high) skill workers must earn more (less) than $w_S(l)$ ($w_S(h)$) and the result follows as well. For $\tilde{y} \in (0,1)$, denote the wage received by a worker of skill x_1 from cross-matching by $w^c(x_1)$. Of course, a high- (low-) skill worker with $x_2^h \equiv G_h^{-1}(\bar{y})$ ($x_2^l \equiv G_l^{-1}(1-\bar{y})$ is indifferent between same- and cross-matching, so that $w^c(x_1) - 0.5(1 - x_2^{x_1})F(h,l) = x_2^{x_1} w^S(x_1)$. After some rearranging, one can show that this implies that $(x_2^h + x_2^l)\Delta w^c = 2x_2^h x_2^l \Delta w^S$, and $F(h,l) - (F(\mathbf{h}) + F(\mathbf{l}))/2 = 0.5\Delta w^S (x_2^h - x_2^l)/(x_2^h + x_2^l)$, where $\Delta w^S = 0.5(F(h) - F(l))$ is the difference between the high- and low-skill same-matched wages, and $\Delta w^c \equiv w^c(h) - w^c(l)$ is the difference between the high- and low-skill wages of workers who cross-match. Clearly then, $\mathrm{Var}(W_B) = 0.25(\Delta w^S)^2$ and

$$\mathrm{Var}(W) = \mathrm{Var}(W_B)(1 + \frac{4\bar{y}(x_2^h x_2^l)^2 + (1-\bar{y})\bar{y}(x_2^h - x_2^l)^2 - \bar{y}(x_2^h + x_2^l)^2}{(x_2^h + x_2^l)^2},$$

Elementary algebra reveals that if $x_2^h x_2^l \leq 1$ then $\mathrm{Var}(W) - \mathrm{Var}(W_B) \leq 0$.

Assumption A3.3. Consider an arbitrary wage function w and a feasible sorting function μ. Variance decomposition yields $\mathrm{Var}(w(x_1, x_2)) = \mathrm{BWI}(w,\mu) + \mathrm{WWI}(w,\mu)$, where

$$\mathrm{BWI}(w,\mu) = \mathrm{Var}\left(w(\mathbf{x}) + w(\mu(\mathbf{x})\right), \qquad \mathrm{WWI}(w,\mu) = 0.25\mathrm{E}\left(w(\mathbf{x}) - w(\mu(\mathbf{x}), \mu_2(\mathbf{x}))\right)^2;$$

thus it suffices to show that $\mathrm{WWI}(w_B, \mu^*) \geq \mathrm{WWI}(w^*, \mu^*)$ and $\mathrm{BWI}(w_B, \mu^*) \geq \mathrm{BWI}(w^*, \mu^*)$.

The first part is easy. It follows directly from (10) that worker $\mathbf{x}$ earns more (less) than their benchmark wage if $\mu_1^*(\mathbf{x}) \geq (\leq) x_1$ and $z(x_1, H_{x_1}(s)) \leq 1$ for all $s \in [x_1, \mu_1(\mathbf{x}]$. Note that $z(x_1, H_{x_1}(\mu^*(\mathbf{x})) = z(x_1, v_1(\mathbf{x})) = x_2$, and $x_1 = \mu_1^*(\mu_1^*(\mathbf{x}), \mu_2^*(\mathbf{x}))$, and thus $z(x_1, H_{x_1}(x_1)) = \mu_2(x_1, x_2)$. It follows, therefore, that if $\max\{x_2, \mu_2^*(x_1, x_2)\} \leq 1$ and $\mu_1^*(\mathbf{x}) \geq (\leq)\mathbf{x}$ then $w(\mu^*(\mathbf{x})) \leq w_B(\mu^*(\mathbf{x}))$ and $w(\mathbf{x}) \geq w_B(\mathbf{x})$. The assumption that $\bar{x}_2 \leq 1$ ensures that $\max\{x_2, \mu_2^*(x_1, x_2)\} \leq 1$, and it thus follows that $\mathrm{WWI}(w_B, \mu^*) \geq \mathrm{WWI}(w^*, \mu^*)$.[32]

[32]If $\bar{x} > 1$ but $H(1)$ is arbitrarily close to 0, the conclsion is the same, because $\max\{x_2, \mu_2^*(x_1, x_2)\} > 1$ for arbitrarily few workers.

Moving on to BWI, we have that

$$c^2\mathrm{BWI}(w_B,\mu^*) = 0.25\mathrm{Var}(x_1^c+\mu_1^*(\mathbf{x})^c) = 0.5(\mathrm{Var}(x_1^c)+\mathrm{Cov}(x_1^c,\mu_1^*(\mathbf{x})^c)),$$
$$c^2\mathrm{BWI}(w^*,\mu^*) = \mathrm{Var}((x_1)^c\mu_1^*(\mathbf{x})^c) = E((x_1^c\mu_1^*(\mathbf{x})^c)^2) - E(x_1^c\mu_1^*(\mathbf{x})^c)^2.$$

Define the random variables $s \equiv c(\ln x_1 + \ln \mu_1^*(\mathbf{x}))$ and $z \equiv (s-2\delta_1)/\alpha$; where $\alpha^2 \equiv \mathrm{Corr}(\ln x_1, \ln \mu_1^*(\mathbf{x}))+1$. By Theorem 2.16 in Fang (1990), $z \sim EC_1(0, c^2\omega_{11};\phi)$; denote the cdf of z by G. Next, let us write

$$c^2(\mathrm{BWI}(w^*,\mu^*)-\mathrm{BWI}(w_S,\mu^*)) = e^{-4\delta_1}\underbrace{(0.5E(z^2)-E(z)^2)}_{\equiv T(\alpha)} - A.$$

Here, $A = 0.5\mathrm{Var}(x_1^c)+0.5\mathrm{E}(x_1^c)\mathrm{E}(\mu_1^*(\mathbf{x})^c))$. It is easy to see that if $\alpha=2$, then $\mathrm{BWI}(w^*,\mu^*) = \mathrm{BWI}(w_S,\mu^*)$, and if $\alpha = 0$, then $\mathrm{BWI}(w^*,\mu^*) \le \mathrm{BWI}(w_S,\mu^*)$.[33] Thus, by standard arguments, if $T(\cdot)$ is convex then $\mathrm{BWI}(w^*,\mu^*) - \mathrm{BWI}(w_S,\mu^*) < 0$.

Proof that $T(\cdot)$ is convex. Let us start by rewriting $T(\alpha)$

$$\begin{aligned}
T(\alpha) &= 0.5\int_{-\infty}^{\infty} e^{2(\alpha z)}\mathrm{d}G(z) - \left(\int_{-\infty}^{\infty} e^{\alpha z}\mathrm{d}z\right)^2 \\
&= \int_{-\infty}^{\infty} e^{\alpha z}\left(\int_{-\infty}^{r} 0.5e^{\alpha z} - e^{\alpha r}\,\mathrm{d}G(r) + \int_{r}^{\infty} 0.5e^{\alpha z} - e^{\alpha r}\,\mathrm{d}G(r)\right)\mathrm{d}G(z) \\
&= \int_{-\infty}^{\infty}\int_{r}^{\infty} e^{\alpha r}(0.5e^{\alpha r}-e^{\alpha z}) + e^{\alpha z}(0.5e^{\alpha z}-e^{\alpha r})\,\mathrm{d}G(r)\,\mathrm{d}G(z) \\
&= 0.5\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 0.5e^{2\alpha z} - 2e^{\alpha(r+z)} + 0.5e^{2\alpha r}\,\mathrm{d}G(r)\,\mathrm{d}G(z).
\end{aligned}$$

Denote $0.5e^{2\alpha z} - 2e^{\alpha(r+z)} + 0.5e^{2\alpha r}$ by $p(z,r;\alpha)$. Note that (a) $G(z) = 1-G(z)$, because any elliptical distribution is symmetric and (b) $p(z,r;\alpha) = p(r,z;\alpha)$; we can thus write:

$$T(\alpha) = \int_0^{\infty}\int_0^{\infty} \underbrace{p(z,r;\alpha)+p(-z,r;\alpha)+p(z,-r;\alpha)+p(-z,-r;\alpha)}_{\equiv P(z,r;\alpha)}\,\mathrm{d}G(r)\,\mathrm{d}G(z).$$

Clearly, it suffices thus to show that $\frac{\partial^2}{\partial\alpha^2}P(z,r;\cdot) \ge 0$ for any $(z,r)\in\mathbb{R}_+^2$; as $P(z,r;\alpha)$ is symmetric in z,r, we can assume, wlog, that $z>r$. First, note that

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha^2}p(z,r;\alpha) &= 2(z^2e^{2\alpha z} + r^2e^{2\alpha r} - (r+z)^2e^{\alpha(r+z)}), \\
\frac{\partial}{\partial z}\frac{\partial^2}{\partial\alpha^2}p(z,r;\alpha) &= 4(\underbrace{ze^{2\alpha z} - (r+z)e^{\alpha(r+z)}}_{l(z,r;\alpha)}) \\
&\quad + 2\alpha(\underbrace{2z^2e^{2\alpha z} - (r+z)^2e^{\alpha(r+z)}}_{k(z,r;\alpha)}, \\
\frac{\partial^2}{\partial z^2}\frac{\partial^2}{\partial\alpha^2}p(z,r;\alpha) &= 4(e^{2\alpha z} - e^{\alpha(r+z)}) + 2\alpha\frac{\partial}{\partial z}\frac{\partial^2}{\partial\alpha^2}p(z,r;\alpha) \\
&\quad + 2\alpha(4ze^{2\alpha z} + \alpha(r+z)^2e^{\alpha(r+z)}).
\end{aligned}$$

[33] If $\alpha = 0$, then $\mathrm{Corr}(\ln x_1, \ln \mu_1^*(\mathbf{x}))$ and the variance of s is 0.

Next, note that

$$\frac{\partial}{\partial z}\frac{\partial^2}{\partial \alpha^2}P(z,r;\alpha) \geq 0 \Rightarrow \frac{\partial^2}{\partial z^2}\frac{\partial^2}{\partial \alpha^2}P(z,r;\alpha) > 0,$$
$$K(z,r;\alpha) \equiv k(z,r;\alpha) + k(-z,r;\alpha) + k(z,-r;\alpha) + k(-z,-r;\alpha) \geq \frac{\partial^2}{\partial \alpha^2}P(z,r;\alpha),$$

because $z(e^{2\alpha z} - e^{-2\alpha z} > 0$ and

$$e^{2\alpha z} - e^{\alpha(r+z)} + e^{-2\alpha z} - e^{-z\alpha(z+r)} = \alpha \int_{r+z}^{z} e^{\alpha s} - e^{-\alpha s}\mathrm{d}s > 0$$

for any $r \in \mathbb{R}$ and $z \geq \max\{0, r\}$. Finally, it is immediate that $\frac{\partial^2}{\partial \alpha^2}P(z,z;\alpha) = 0$ and that $L(z,z;\alpha) = 0$, where $L(z,r;\alpha) \equiv l(z,r;\alpha) + l(-z,r;\alpha) + l(z,-r;\alpha) + l(-z,-r;\alpha)$. Jointly these facts imply that $\frac{\partial^2}{\partial \alpha^2}P(z,r;\alpha) \geq 0$ for all $(z,r) \in \mathbb{R}^2_+$. Suppose not. Then there must exist some $(z^*, r*) \in \mathbb{R}^2_+$ such that $\frac{\partial^2}{\partial \alpha^2}P(z^*, r*; \cdot) < 0$ and (by symmetry) $z^* > r*$. This is only possible if the set $\Omega \equiv \{z \in [r*, z^*] : \frac{\partial}{\partial z}\frac{\partial^2}{\partial \alpha^2}P(z,r;\alpha) < 0\}$ is non-empty; denote its infimum by z'. For all $z \in [r^*, z']$ it must be the case that $\frac{\partial^2}{\partial \alpha^2}P(z,r^*;\alpha), \frac{\partial}{\partial z}\frac{\partial^2}{\partial \alpha^2}P(z,r^*;\alpha) \geq 0$. This implies that $\frac{\partial}{\partial z}\frac{\partial^2}{\partial \alpha^2}P(z',r^*;\alpha) > 0$; contradiction!

Proof of Proposition 7 The first part of (i) is immediate, because a worker of an arbitrary type $\mathbf{x}$ can guarantee themselves the payoff of $0.5x_2F(x_1,x_1) = x_2w_B(x_1)$ by same-matching. The second part follows, because if either Assumption A3.1 or A3.3 is satisfied and H has full support, then—by (5) and (7)—only a measure zero of workers same-matches and conditional on x_1 utility is minimised for the same-matching workers by Proposition 3 and footnote 21.

If surplus is strictly submodular, then the above logic breaks down, because $u^*(x_1, x_2^*(x_1)) = 0.5x_2F(x_1,x_1) < x_2w_B(x_1)$; that is, while same-matching is still any workers' outside option, this outside option is strictly worse than what they would receive with $p = 0$. Clearly then, $u^*(\mathbf{x}') < x_2w_B(x_1')$, and the result follows from the absolute continuity of $\Pr(x_2|X_1 = x_1)$ and continuity of $u^*(\bullet)$.

Proof of Proposition 8 (i) In this case, outsourcing is dominated by same-matching: In choosing between these options relative concerns do not matter, and $c > s_F$ implies that the output from same-matching is higher than from outsourcing.

(ii) By the same logic as in (i), $c < s_F$ implies that outsourcing dominates same-matching for all teams, and thus all teams consist of one high- and one low-skill worker. Denote by $w^o(\mathbf{x}), u^o(\mathbf{x})$ ($w^n(\mathbf{x}), u^n(\mathbf{x})$) the wage and payoff received by a worker of type $\mathbf{x}$ if their team is (non-)outsourcing. In equilibrium, $w^n(\mathbf{x}), w^o(\mathbf{x})$ are constant in x_2; otherwise, no-one would match with the workers earning $\max_{x_2} w^i(x_1,x_2)$ for $i \in \{n,o\}$ and $x_1 \in \{h,l\}$. The outsourcing teams split the benefit of outsourcing $s_F - c$ according to their bargaining power, so that

$$w^o_{x_1} = 0.5F(\mathbf{x}_1) + \alpha_{x_1}(s_F - c), \tag{19}$$

where $\alpha_l + \alpha_h = 1$. Note that $\frac{\partial}{\partial x_2}(u^n(h,x_2) - u^o(h,x_2)) = 0.5F(h,l) - w^o_h < 0$ ($\frac{\partial}{\partial x_2}(u^n(l,x_2) - u^o(l,x_2)) = 0.5F(h,l) - w^o_l > 0$), so that there exists a cutoff level of x_2 such all high (low) skill workers with x_2 greater (lower) that the cutoff are in outsourcing teams. By feasibility, the measure of (non-)outsourcing high-skill workers must be equal to the measure of (non-)outsourcing low-skill workers; thus, there exists some y^o such that a high- (low-)skill worker outsources (gets outsourced) iff $x_2 > G_h^{-1}(y^o)$ ($x_2 < G_l^{-1}(1-y^o)$). Of course, if $y^o \in (0,1)$, then workers $(h, G_h^{-1}(y^o))$ and $(l, G_l^{-1}(1-y^o))$ are indifferent between outsourcing and not, so that

$$w_h^n = G_h^{-1}(y^o)w_h^o + 0.5(1 - G_h^{-1}(y^o))F(h,l),\ w_l^n = G_l^{-1}(1-y^o)w_l^o + 0.5(1 - G_l^{-1}(1-y^o))F(h,l). \tag{20}$$

(12) follows then from adding up (20) for the two skill types, using $w_h^n + w_l^n = F(h,l)$, $w_h^o + w_l^o = F(h,l) - c$, (19) and some algebra.[34]

B General Utility, Production and Distribution Functions

Consider the model from Section 3 but replace the special case of the KUJ utility function from (1) with a general KUJ utility function:

$$U(w^k, \bar{w}^{k,j}; x_2), \tag{21}$$

which is twice continuously differentiable and increasing in w^k.

Imperfectly Transferable Utility Under general KUJ utility, utility is imperfectly transferable in this model. Given that, and adapting from Legros and Newman (2007), in order to define the equilibrium we need to first specify the *utility possibility frontier* $\psi : D_{\mathbf{x}}^2 \times \mathbb{R} \to \mathbb{R}$, such that

$$\psi(\mathbf{x}^j, \mathbf{x}^k, u) \equiv \max_{w^j} U\left(F(x_1^k, x_1^j) - w^j, 0.5F(x_1^k, x_1^j), x_2^k\right)$$
$$\text{subject to } U(w^j, 0.5F(x_1^k, x_1^j), x_2^j) \geq u.$$

In other words, the utility possibility frontier $\psi(\mathbf{x}^j, \mathbf{x}^k, u)$ is equal to the highest utility worker $\mathbf{x}^k$ can achieve with a co-worker $\mathbf{x}^j$ if the co-worker receives utility of at least u. Define $g : \mathbb{R}^3$, as the inverse of $U(\cdot, \bar{w}^{k,j}; x_2)$ so that $g(U(w^j; \bar{w}^{k,j}, x_2); \bar{w}^{k,j}, x_2) = w^j$. It is easy to see that ψ becomes then

$$\psi(\mathbf{x}^j, \mathbf{x}^k, u) = U\left(F(x_1^k, x_1^j) - g(u; 0.5F(x_1^k, x_1^j), x_2^j), 0.5F(x_1^k, x_1^j), x_2^k\right).$$

With the inverse g and the utility possibility frontier define, we can now state the assumption that ensures that x_2 captures the inverse of the strength of relative concerns.

Assumption 4. The utility possibility frontier satisfies the generalised increasing differences (GID) condition (Legros and Newman, 2007) with respect to x_1^k, x_2^j, that is, the function

$$t(x_1^k; x_2^k, \mathbf{x}^k, u) \equiv -\frac{\frac{\partial}{\partial x_2^j}\psi(\mathbf{x}^j, \mathbf{x}^k, u)}{\frac{\partial}{\partial u}\psi(\mathbf{x}^j, \mathbf{x}^k, u)} = -\frac{\frac{\partial}{\partial x_2^j} g(u; 0.5F(x_1^k, x_1^j), x_2^j)}{\frac{\partial}{\partial u} g(u; 0.5F(x_1^k, x_1^j), x_2^j)}$$

[34]The results for the cases where $y^o \in \{0,1\}$ follow naturally from the fact that $G_l^{-1}(1-y^o)/G_h^{-1}(y^o)$ is decreasing in y^o.

is increasing.

Note that the original definition of GID from Legros and Newman (2007) differs from the one, but the equivalence between these two definitions has been shown by Chade, Eeckhout, and Smith (2017). Moreover, in previous work this condition has been defined only for one-dimensional assignment problem: However, my assumption specifies the relationship between the skill of the worker and the relative concerns of the co-worker only, and hence the definition is the same.

Finally, note that Assumption 4 is satisfied (a) for the utility specified in (1) because $\frac{\partial^2}{\partial x_1^k \partial x_2^j}\psi > 0$ and $\frac{\partial^2}{\partial x_1^k \partial u}\psi = 0$ and (b) for the workhorse KUJ utility used in the analysis in Gali (1994): $U(w^k, \bar{w}^{k,j}; x_2) = (w^k)^\alpha (w^{j,k})^{x_2(1-\alpha)}$.

The Competitive Equilibrium The competitive equilibrium is still as in Definition 1, with two exceptions. First, in the general model sorting μ is individually rational if

$$\mu(\mathbf{x}^k) = \mathbf{x}^j \Rightarrow \mathbf{x}^j \in \arg\max_{\mathbf{x}} \psi(\mathbf{x}^k, \mathbf{x}, u(\mathbf{x})).$$

Second, (3) is amended to

$$u(\mathbf{x}^k) = \max_{\mathbf{x}^j} \psi(\mathbf{x}^j, \mathbf{x}^k, u(\mathbf{x}^j)). \tag{22}$$

The Binary Skills Case In the binary skills case equilibrium sorting can be fully characterised even for general KUJ utility.[35]

Proposition 9. Suppose that Assumption A3.2 is satisfied and workers' utility function is given by (21). Define the function $b : [0, 1] \to \mathbb{R}$, such that

$$\begin{aligned} b(y) =& g(U(0.5F(\mathbf{l}), 0.5F(\mathbf{l}); G_l^{-1}(1-y)); 0.5F(h,l), G_l^{-1}(1-y)) \\ &+ g(U(0.5F(\mathbf{h}), 0.5F(\mathbf{h}); G_h^{-1}(y)); 0.5F(h,l), G_h^{-1}(y)) \end{aligned}$$

as well as $\tilde{y}$ such that $\tilde{y} = 1$ if $b(1) < F(h,l)$, $\tilde{y} = 0$ if $b(0) > F(h,l)$, and $\tilde{y}$ solves $b(y) = F(h,l)$ otherwise. In the unique equilibrium high-skill workers with $x_2 \leq G_h^{-1}(\tilde{y})$ match low-skill workers with $x_2 \geq G_l^{-1}(1-\tilde{y})$ and all remaining workers same-match.

Proof. By the same argument as in the proof of Proposition 8, all low- (high-) skill workers matched to a high- (low-) skill co-worker earn the same wage. Denote this wage by $w^h(l)$ $(w^l(h))$. Given that, we can define the utility received by worker $\mathbf{x}^k$ when matched to a worker of skill: x_1^k:

$$u^{x_1^j}(\mathbf{x}^k) \equiv U(w^{x_1^j}(x_1^k), 0.5F(x_1^k, x_1^j), x_2^k).$$

A high-skill worker will chose to same-match if $u^h(h, x_2) > u^l(h, x_2)$ and cross-match if the inequality flips. Thus, as long as $b_h(x_2) \equiv g(u^h(h, x_2); 0.5F(h,l), x_2)$ is strictly increasing, there will exist a unique cutoff value of x_2, such that all high-skill workers with x_2 above the cutoff will same-match, and all those with x_2 below the cutoff will cross-match. Clearly, $b_h(x_2)$ is strictly

[35] If skills are continuously distributed, then one can show that Assumption 4 implies that $\mu_1^*(\mathbf{x})$ is increasing in x_2 by adapting the classic argument from Sattinger (1979) (proof available on request).

increasing if and only if

$$\frac{\partial}{\partial x_2^k} u^h(h, x_2^k) \frac{\partial}{\partial u} g(u^h(h, x_2); 0.5F(h,l), x_2) + \frac{\partial}{\partial x_2^k} g(u^h(h, x_2); 0.5F(h,l), x_2) \geq 0,$$

which is equivalent to

$$\frac{\frac{\partial}{\partial x_2^k} g(u^h(h, x_2); 0.5F(h,l), x_2)}{\frac{\partial}{\partial u} g(u^h(h, x_2); 0.5F(h,l), x_2)} \geq \frac{\frac{\partial}{\partial x_2^k} g(u^h(h, x_2); 0.5F(h,h), x_2)}{\frac{\partial}{\partial u} g(u^h(h, x_2); 0.5F(h,h), x_2)}$$

which is satisfied by Assumption 4.

By similar logic, one can show that $b_l(x_2) \equiv g(u^l(l, x_2); 0.5F(h,l), x_2)$ is decreasing, and thus all low-skill workers with x_2 above (below) some cutoff will cross- (same-)match. Thus, by feasibility of equilibrium sorting, there must exist such a $\tilde{y}$ that high-skill workers with $x_2 \leq G_h^{-1}(\tilde{y})$ match low-skill workers with $x_2 \geq G_l^{-1}(\tilde{y})$ and all remaining workers same-match. Naturally, if $\tilde{y} \in (0,1)$, then workers $(h, G_h^{-1}(\tilde{y}))$ and $(l, G_l^{-1}(\tilde{y}))$ are indifferent between matching with each other and same-matching, implying that

$$\begin{aligned}&U(0.5F(\mathbf{h}), 0.5F(\mathbf{h}); G_h^{-1}(y)) = \\ &U(F(h,l) - g(U(0.5F(\mathbf{l}), 0.5F(\mathbf{l}); G_l^{-1}(1-y)); 0.5F(h,l), G_l^{-1}(1-\tilde{y})), 0.5F(h,l); G_h^{-1}(\tilde{y})).\end{aligned}$$

Taking the inverse g on both sides yields $b(\tilde{y}) = F(h,l)$, as required. If $b(1) > F(h,l)$ then all high-skill workers prefer to same-match than to pay the low-skill worker with weakest relative concerns enough to provide them with their same-match utility; similarly if $b(1) < F(h,l)$, then all high-skill workers find it more beneficial to match with the low-skill worker with strongest relative concerns than to same-match. $\square$

Thus, the basic structure of the equilibrium is qualitatively the same in the binary skill case under the general KUJ utility function as under the special one considered in the baseline. The only major difference is the condition determining the cutoff $\tilde{y}$: In the baseline model, this was a function of a simple sufficient statistics of F, a_F; in the general case, the cutoff still depends on all possible production plans $(F(\mathbf{l}, F(h,l), F(\mathbf{h}))$ but the relationship is much more complicated.

References

Abraham, K. G. and S. K. Taylor (1996). Firms' use of outside contractors: Theory and evidence. *Journal of Labor Economics 14*(3), 394–424.

Aghion, P., A. Bergeaud, R. W. Blundell, and R. Griffith (2019). The innovation premium to soft skills in low-skilled occupations. *Available at SSRN 3489777*.

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

Barnett, R. C., J. Bhattacharya, and H. Bunzel (2019). The fight-or-flight response to the joneses and inequality. *Journal of Economic Dynamics and Control 101*, 187–210.

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

Bergeaud, A., C. Malgouyres, C. Mazet-Sonilhac, and S. Signorelli (2024). Technological change and domestic outsourcing. *Journal of Labor Economics 0*(ja), null.

Boerma, J., A. Tsyvinski, and A. P. Zimin (2021, September). Sorting with team formation. Working Paper 29290, National Bureau of Economic Research.

Bojilov, R. and A. Galichon (2016). Matching in closed-form: equilibrium, identification, and comparative statics. *Economic Theory 61*(4), 587–609.

Bolton, G. E. and A. Ockenfels (2000). Erc: A theory of equity, reciprocity, and competition. *American economic review 91*(1), 166–193.

Bottan, N. L. and R. Perez-Truglia (2020). Choosing your pond: Location choices and relative income. *Review of Economics and Statistics, forthcoming*.

Bound, J. and G. Johnson (1992). Changes in the structure of wages in the 1980's: An evaluation of alternative explanations. *The American Economic Review 82*(3).

Buser, T., M. Niederle, and H. Oosterbeek (2014, 05). Gender, Competitiveness, and Career Choices *. *The Quarterly Journal of Economics 129*(3), 1409–1447.

Cabrales, A. and A. Calvó-Armengol (2008). Interdependent preferences and segregating equilibria. *Journal of Economic Theory 139*(1), 99–113.

Cabrales, A., A. Calvó-Armengol, and N. Pavoni (2008). Social preferences, skill segregation, and wage dynamics. *The Review of Economic Studies 75*(1), 65–98.

Card, D., A. Mas, E. Moretti, and E. Saez (2012, October). Inequality at Work: The Effect of Peer Salaries on Job Satisfaction. *American Economic Review 102*(6), 2981–3003.

Chade, H., J. Eeckhout, and L. Smith (2017). Sorting through search and matching models in economics. *Journal of Economic Literature 55*(2), 493–544.

Clark, A. E. and A. J. Oswald (1998). Comparison-concave utility and following behaviour in social and economic settings. *Journal of Public Economics 70*(1), 133–155.

Coase, R. H. (1937). The nature of the firm. *Economica 4*(16), 386–405.

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

Fang, K. (1990). *Symmetric Multivariate and Related Distributions (1st ed.).* Chapman and Hall/CRC.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics 114*(3), 817–868.

Fershtman, C., H. Hvide, and Y. Weiss (2006). Cultural Diversity, Status Concerns and the Organization of Work. In *The Economics of Immigration and Social Diversity*. Emerald group Publishing Limited.

Fershtman, C. and Y. Weiss (1993, July). Social Status, Culture and Economic Performance. *Economic Journal 103*(419), 946–59.

Frank, R. (1984a, September). Are Workers Paid their Marginal Products? *The American Economic Review 74*(4), 549–571.

Frank, R. (1984b). Interdependent Preferences and the Competitive Wage Structure. *The RAND Journal of Economics 15*(4), 510–520.

Frank, R. H. (1985). *Choosing the Right Pond : Human Behavior and the Quest for Status*. Oxford University Press New York ; Oxford [Oxfordshire].

Freund, L. (2022). Superstar teams: The micro origins and macro implications of coworker complementarities. *Available at SSRN 4312245*.

Gali, J. (1994). Keeping up with the joneses: Consumption externalities, portfolio choice, and asset prices. *Journal of Money, Credit and Banking 26*(1), 1–8.

Galichon, A. (2016). *Optimal Transport Methods in Economics*. Princeton University Press.

Ghiglino, C. and S. Goyal (2010). Keeping up with the neighbors: social interaction in a market economy. *Journal of the European Economic Association 8*(1), 90–119.

Gola, P. (2021). Supply and Demand in a Two-Sector Matching Model. *Journal of Political Economy 129*(3), 940–978.

Gola, P. (2024). On the importance of social status for occupational sorting. *The Economic Journal 134*(661), 2009–2040.

Goldschmidt, D. and J. F. Schmieder (2017). The rise of domestic outsourcing and the evolution of the german wage structure. *The Quarterly Journal of Economics 132*(3), pp. 1165–1217.

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

Grossman, S. J. and O. D. Hart (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of political economy 94*(4), 691–719.

Hart, O. and J. Moore (1990). Property rights and the nature of the firm. *Journal of political economy 98*(6), 1119–1158.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The review of Economic studies 66*(1), 169–182.

Holmstrom, B. and P. Milgrom (1994). The firm as an incentive system. *The American economic review*, 972–991.

Hopkins, E. (2024). Cardinal sins? conspicuous consumption, cardinal status and inequality. *Journal of the European Economic Association*, jvae025.

Hopkins, E. and T. Kornienko (2004, September). Running to Keep in the Same Place: Consumer Choice as a Game of Status. *American Economic Review 94*(4), 1085–1107.

Hopkins, E. and T. Kornienko (2009, November). Status, affluence, and inequality: Rank-based comparisons in games of status. *Games and Economic Behavior 67*(2), 552–568.

Juhn, C., K. M. Murphy, and B. Pierce (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy 101*(3), 410–442.

Katz, L. and K. M. Murphy (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics 107*(1), 35–78.

Kim, C. (1998). Stochastic dominance, pareto optimality, and equilibrium asset pricing. *The Review of Economic Studies 65*(2), 341–356.

Klein, B., R. G. Crawford, and A. A. Alchian (1978). Vertical integration, appropriable rents, and the competitive contracting process. *The journal of Law and Economics 21*(2), 297–326.

Kőszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Kremer, M. and E. Maskin (1996). Wage inequality and segregation by skill.

Langtry, A. (2023). Keeping up with "the joneses": reference-dependent choice with social comparisons. *American Economic Journal: Microeconomics 15*(3), 474–500.

Langtry, A. and C. Ghinglino (2023). Status substitution and conspicuous consumption. *arXiv preprint arXiv:2303.07008*.

Legros, P. and A. F. Newman (2007). Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities. *Econometrica 75*(4), 1073–1102.

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

Liu, Q. (2020, August). Stability and bayesian consistency in two-sided markets. *American Economic Review 110*(8), 2625–66.

Liu, Q. (2024). Chapter 5 - matching with incomplete information. Volume 1 of *Handbook of the Economics of Matching*, pp. 237–274. North-Holland.

Liu, Q., G. J. Mailath, A. Postlewaite, and L. Samuelson (2014). Stable matching with incomplete information. *Econometrica 82*(2), 541–587.

Louizos, C., K. Ullrich, and M. Welling (2017). Bayesian compression for deep learning. *Advances in neural information processing systems 30*.

Luttmer, E. F. P. (2005). Neighbors as Negatives: Relative Earnings and Well-Being. *The Quarterly Journal of Economics 120*(3), pp. 963–1002.

Mani, A. and C. H. Mullin (2004, October). Choosing the Right Pond: Social Approval and Occupational Choice. *Journal of Labor Economics 22*(4), 835–862.

Manning, A. (2004). We can work it out: the impact of technological change on the demand for low-skill workers. *Scottish Journal of Political Economy 51*(5), 581–608.

Mazzolari, F. and G. Ragusa (2013). Spillovers from high-skill consumption to low-skill labor markets. *Review of Economics and Statistics 95*(1), 74–86.

McCann, R. J. and M. Trokhimtchouk (2010). Optimal partition of a large labor force into working pairs. *Economic Theory 42*(2), 375–395.

Nickerson, J. A. and T. R. Zenger (2008). Envy, comparison costs, and the economic theory of the firm. *Strategic Management Journal 29*(13), 1429–1449.

Owen, J. and R. Rabinovitch (1983). On the class of elliptical distributions and their applications to the theory of portfolio choice. *The Journal of Finance 38*(3), 745–752.

Perez-Truglia, R. (2020). The effects of income transparency on well-being: Evidence from a natural experiment. *American Economic Review 110*(4), 1019–54.

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

Sklar, A. (1959). *Fonctions de Repartition a n Dimensions et Leurs Marges.*

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

Tinbergen, J. (1956). On the theory of income distribution. *Weltwirtschaftliches Archiv*.

Williamson, O. E. (1971). The vertical integration of production: Market failure considerations. *The American Economic Review 61*(2), 112–123.