Journal of Public Economics 249 (2025) 105475

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/pubec

The pond dilemma with heterogeneous relative concerns☆

Paweł Gola

University of Edinburgh, 30 Buccleuch Place, Edinburgh, EH8 9JT, UK

HIGHLIGHTS

- Sorting by skill and status desire creates “proud leader-pragmatic companion” teams.
- Skill-biased technology creates trickle-down effect for pragmatic low-skill workers.
- Heterogeneous status preferences can increase overall wage inequality.
- Firms outsource to avoid harmful social comparisons between worker types.
- Skill-biased technological change explains observed rise in domestic outsourcing.

ARTICLE INFO

JEL classification:
C78
D91
J31

Keywords:
Relative concerns
Multidimensional sorting
Trickle-down effects
Skill-biased technological change
Outsourcing

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 three large classes 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, which provides a compelling explanation for the observed long-term increase in domestic outsourcing.

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 et al., 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

☆ 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.

Email address: pawel.gola@ed.ac.uk.

[1] The literature on routine-biased technological change, started by Autor et al. (2003), posits that low-skill workers in non-routine occupations also gained from technological change. Unless these beneficiaries are intrinsically different from 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 to entry.

https://doi.org/10.1016/j.jpubeco.2025.105475
Received 12 March 2025; Received in revised form 20 July 2025; Accepted 20 August 2025

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 (see, for example, Luttmer, 2005; Card et al., 2012; Perez-Truglia, 2020; Bottan and Perez-Truglia, 2022). Moreover, the strength of these relative concerns—and their close cousin, competitiveness—differs across individuals, affecting their career and location choices (Buser et al., 2014; Bottan and Perez-Truglia, 2022).

It is this heterogeneity in relative concerns which produces heterogeneous trickle-down effects of SBTC. To present the core mechanism cleanly, I focus first on the case without production complementarities, but will later discuss how the interactions between production complementarities and status concerns generate additional insights about outsourcing and inequality. Absent production complementarities, matching is driven by relative concerns only and 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. While a low-skill worker with strong relative concerns would require a large compensation for accepting low status, a low-skill worker who is *pragmatic*, in that they care about their own wage only, requires no such compensation. High-skill workers with strong relative concerns therefore match with these pragmatic low-skill workers and the avoided compensation is split endogenously between them, with the exact split depending on both the production technology and the distribution of preferences. Crucially, this wage premium received by pragmatic low-skill workers is not a 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, pragmatic 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 et al., 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 pragmatic 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. Section 5 explores the economic implications of relative concerns, using three specifications of the model for which analytical solutions can be derived. Section 6 extends the model in a number of directions. Section 7 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-workers' skills, (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 et al. (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.[2] 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 their lower status.[3] For that reason, relative concerns affect wage inequality through sorting only, and the trickle-down effect—critical for my results—is absent.[4]

In addition to that, Cabrales et al. (2008); Cabrales and Calvó-Armengol (2008) consider sorting in the presence of inequity aversion

[2] 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.

[3] 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.

[4] 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) and 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.

rather than relative concerns, and show that inequity aversion leads to more positive and assortative sorting in skill.[5] In contrast to my work, they only allow for additively separable production functions, and so the surprising result that inequity 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 produces using only one of the two dimensions of skill, the other characterisations all leverage bi-linear surplus functions and Gaussian distributions of traits to retain tractability (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 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 distributions is much easier to satisfy.[6]

Theory of the firm. The theory presented in this paper provides a clear answer to a very old question posed by the transaction cost theory of the firm literature (Coase, 1937; Williamson, 1971; Klein et al., 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.[7] 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.[8] 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 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 et al. (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 pragmatism $x_2 \in D_{x_2}$, 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 $D_{\mathbf{x}} \equiv D_{x_1} \times D_{x_2}$, with the support of H denoted by $\mathrm{supp}(H)$. The marginal distribution of x_i will be denoted by H_{x_i}, 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

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

and $x_2 \geq 0$. A low x_2 means that the worker has strong *relative concerns*, whereas a high x_2 makes the worker more *pragmatic*, that is, mostly interested in their own wage rather than the co-workers' wage. In particular, for workers with $x_2^k = 1$, (1) reduces to standard neo-classical

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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 6.3). Section 6.4 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 all workers have neo-classical preferences ($x_2 = 1$), then the model reduces to a standard Sattinger (1979) sorting model. I will refer to this case as the *benchmark* and denote it by the subscript B; the benchmark case is discussed in more detail in Section 5.

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 values 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 on the teammates' pragmatism. It is worth noting that the surplus function is supermodular in the worker's skill x_1^k and their co-worker's pragmatism x_2^j in addition to any production complementarities between the worker's and their co-worker's skills that are inherited from the production function F.

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 workers 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 $\mu^* : D_{\mathbf{x}} \to D_{\mathbf{x}}$ 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).

It is well-established that in two-sided sorting problems with transferable utility the competitive equilibrium coincides with the solution to the planner's problem (Gretsky et al., 1992). McCann and Trokhimtchouk (2010) show that this 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:*

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

and u^* solves its dual problem, that is

$$u^* \in \arg\min_{u} \left(\int_{D_{\mathbf{x}}} u(\mathbf{x}) dH(\mathbf{x}) \quad s.t.\ \forall_{(\mathbf{x}^j, \mathbf{x}^k) \in D_{\mathbf{x}}^2} u(\mathbf{x}^k) + u(\mathbf{x}^j) \geq \Pi(\mathbf{x}^j, \mathbf{x}^k) \right).$$

Theorem 1 states that the set of equilibria coincides with the set of solutions to the social planner's problem. This means that one can characterise the equilibrium by solving the social planner's problem, which is extremely useful because the social planner's problem is typically much easier to tackle.

3.3. Assumptions

As discussed in Section 3.1, the surplus function may feature complementarities between both x_1^k, x_1^j and x_1^k, x_2^j. This means that the social planner has two, typically conflicting goals: They always want to match high-skill workers with pragmatic co-workers, but if the production function is supermodular (submodular) they may also want to match these very same high-skill workers with high-(low-) skill co-workers.

The optimal trade-off between these two conflicting goals is extremely challenging to characterise. However, the characterisation becomes possible if all complementarities between x_1^k and $\mathbf{x}^j$ can be captured using a single-dimensional index of $\mathbf{x}^j$.

Assumption 1 (The Skill-Preference Index). There exist functions $q, s : D_{\mathbf{x}} \to \mathbb{R}$ as well as a strictly supermodular function $t : \mathbb{R} \times D_{x_1} \to \mathbb{R}$ such that

$$0.5F(x_1^j, x_1^k)x_2^j = t(q(\mathbf{x}^j), x_1^k) + s(\mathbf{x}^j). \tag{4}$$

I will denote the rank of worker $\mathbf{x}$ in the distribution of $q(\mathbf{X})$ by $v_1(\mathbf{x})$ and refer to it as the *skill-preference index*, with[9]

$$v_1(\mathbf{x}) \equiv J(q(\mathbf{x})) \equiv \Pr(q(\mathbf{X}) \leq q(\mathbf{x})).$$

Assumption 1 implies that the surplus function can be rewritten as

$$\Pi(\mathbf{x}^k, \mathbf{x}^j) = t(J^{-1}(v_1(\mathbf{x}^k)), x_1^j) + t(J^{-1}(v_1(\mathbf{x}^j)), x_1^k) + s(\mathbf{x}^k) + s(\mathbf{x}^j),$$

where J denotes the cumulative distribution function of $q(\mathbf{X})$. Since t is supermodular, this formulation of surplus makes it very clear that social planner would like to match high-skill workers with co-workers who have high skill-preference index. However, for this positive assortative matching to be feasible we need an additional assumption; and in order to state this assumption, we 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

$$C_{\mathbf{Y}}(\Pr(Y_1 \leq y_1), \Pr(Y_2 \leq y_2)) = \Pr(Y_1 \leq y_1, Y_2 \leq y_2).$$

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 (Exchangeability). The distribution H and the production function F are such that there exists an *exchangeable* copula C_{X_1,V_1} of (X_1, V_1), that is, such a copula that $C_{X_1,V_1}(u,v) = C_{X_1,V_1}(v,u)$ for all $u, v \in [0,1]$.

The exchangeability of C_{X_1,V_1} implies that the copula of X_1, V_1 is the same as the copula of V_1, X_1; this means that a sorting function that assigns high skill workers to co-workers with high skill-preference index (and workers with high skill-preference index to co-workers with high skill) will be feasible.

[9] Here, $\mathbf{x}$ denotes a specific realisation of $\mathbf{X}$, the random vector.

To understand the structure that exchangeability imposes on the joint distribution of x_1, v_1 it is best to restrict attention to continuous random variables and focus on the conditional distributions of x_1 and v_1. Exchangeability means that a worker with skill-preference index equal to v is as likely to rank higher than p in skill as a worker who ranks v in skill is to have a skill-preference index higher than p, and that this is true for any values of v and p. Note that most commonly used copulas are exchangeable: This is true, in particular, for all Archimedean and elliptical copulas.[10]

4. Characterising the equilibrium

In this section, I will characterise the equilibrium sorting, payoff and wage functions.

4.1. Equilibrium sorting

To characterise the equilibrium I will leverage the equivalence between the set of equilibria and the set of solutions to the planner's problem established in Theorem 1, as well as the fact that 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 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

$$\mu^*(x_1,x_2) = [H_{x_1}^{-1}(v_1(x_1,x_2)), z(H_{x_1}^{-1}(v_1(x_1,x_2)), H_{x_1}(x_1)]^T, \tag{5}$$

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

Theorem 2 shows that, in equilibrium, high-skill workers match workers with high skill-preference index. Furthermore, when the distribution of skill is continuous (and hence the inverse of its cumulative distribution function is well-defined), the equilibrium sorting function is unique and can be characterised analytically. To better understand what are the implications of Theorem 2 for sorting in the (x_1,x_2) space, let us first consider the two extreme cases in which the skill-preference index depends only on (a) skill and (b) preference.

If all workers have the same preferences, then the skill-preference index will depend on skill only; in fact, Assumption 1 is then satisfied if and only if the production function is either strictly supermodular or strictly submodular. Under strictly supermodular production, the function t is strictly supermodular in y, x_1^j if and only if v_1 is strictly increasing in x_1^k, so that $v_1(\mathbf{x}^k) = H_{x_1}(x_1^k)$ and we obtain positive assortative matching (PAM) in skills. Similarly, under strictly submodular production v_1 must be strictly decreasing in x_1^k and thus $v_1(\mathbf{x}) = 1 - H_{x_1}(x_1^k)$ 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; indeed, if $x_2 = 1$ for all workers, then the model reduces to the benchmark case.

The other extreme is the case when the production function features no complementarities. In that case, the supermodularity of t implies that v_1 can depend only on the worker's pragmatism x_2, and $v_1(\mathbf{x}^k) = H_{x_2}(x_2^k)$. As a consequence, high-skill workers with strong relative concerns (high x_1, low x_2) will have very pragmatic low-skill (low x_1, high x_2) co-workers. I will refer to this as the "proud leader—pragmatic companion" sorting pattern. Intuitively, such a sorting pattern is welfare maximising because it allows each member of the team to receive what they care most about: for the "proud leader" that is high within-team status, and for the "pragmatic companion" it is a higher wage than their skill would otherwise warrant.

Finally, let us consider the intermediate case, when relative concerns are heterogenous but the production function is supermodular (submodular). Keeping the worker's skill constant, their skill-preference index will depend positively on how pragmatic they are; similarly, keeping their preference constant, the index will depend positively (negatively) on their skill. Intuitively, 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. In Section 5 we will explore in detail how these two forces offset each other.

4.2. Equilibrium payoffs and wages

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

$$u^*(\mathbf{x}) = u^*(x_1,x_2^*) + 0.5\int_{x_2^*}^{x_2} F(x_1,\mu_1^*(x_1,s))\,\mathrm{d}s. \tag{6}$$

Thus, as long as for every x_1 there exists some x_2^* for which $u^*(x_1,x_2^*)$ can be determined, we will be able to derive $u^*(\mathbf{x})$. The obvious candidates for such (x_1,x_2^*) are workers who same-match in equilibrium, that is, match with a co-worker of the same skill. Co-workers of the same skill must split output equally—otherwise one of them could do strictly better by matching with 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 (6) as long as for every x_1 there exists a worker $(x_1,x_2^*(x_1))$ who same-matches. If skills are continuously distributed, this condition is satisfied if and only if $(H_{x_1}(x_1), H_{x_1}(x_1)) \in \mathrm{supp}(C_{X_1,V_1})$ for all $x_1 \in D_{x_1}$, because with continuous skills a worker same-matches if and only if $v_1(\mathbf{x}) = H_{x_1}(x_1)$.

Theorem 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 $(x_1,x_2^*(x_1)) \in \mathrm{supp}(H)$ and $\mu_1(x_1,x_2^*(x_1)) = x_1$, then the equilibrium payoff function and wage functions u^*, w^* are:*

$$u^*(\mathbf{x}) = \underbrace{0.5F(\mathbf{x_1})x_2}_{\equiv u_S(\mathbf{x})\text{(same-match payoff)}} + \underbrace{0.5\int_{x_2^*(x_1)}^{x_2} x_2 - s\,dF(x_1,\mu_1^*(x_1,s))}_{\text{cross-matching benefit}}, \tag{7}$$

$$w^*(\mathbf{x}) = \underbrace{0.5F(\mathbf{x_1})}_{\equiv w_S(x_1)\text{(same-match wage)}} + \underbrace{0.5\int_{x_2^*(x_1)}^{x_2} 1 - s\,dF(x_1,\mu_1^*(x_1,s))}_{\text{cross-matching payment}}. \tag{8}$$

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

Theorem 3 provides analytical expressions for equilibrium payoffs and wages. Since every worker can always same-match, it is natural to decompose payoffs and wages into the same-match component and the remainder, which I call the cross-matching benefit (for payoffs) or payment (for wages). Considering payoffs first, the same-match component $u_S(\mathbf{x})$ equals half of the output of a same-match multiplied by the agent's pragmatism. Technically, the same-match component increases with the worker's pragmatism—but only because pragmatic workers value the exact same wage more! Since the same-match payoff is available to all workers, the cross-matching benefit must be positive for everyone. However, workers who same-match do not benefit from cross-matching. And indeed, an inspection of (7) reveals that holding skill constant the cross-matching benefit increases with the distance between x_2 and $x_2^*(x_1)$. This means that it is precisely the 'proud leaders' and the 'pragmatic companions' who gain the most from cross-matching.

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

Overall, the increase in the same-match payoff caused by higher pragmatism dominates the possibly lower cross-matching benefit that results, so that holding skill constant more pragmatic workers earn higher payoffs.[11] This however does not imply that pragmatic workers are better off than workers with high relative concerns, because welfare comparisons between workers with different preferences are not meaningful.[12] Indeed, (3) implies that all workers prefer their equilibrium bundle of own and co-worker wages over that of another worker with the same skill but different preferences. In other words, there is no envy between workers of identical skill. This is discussed in more detail in Section 6.6 and Appendix A.

The same-match wage equals half of the output of a same-match and thus does not depend on the worker's pragmatism. In contrast, the cross-matching payment does depend on preferences. Since $v_1(\mathbf{x})$ and $\mu_1^*(\mathbf{x})$ increase in x_2, we have

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

provided that $1 > x_2 > x_2'$. Thus, the more pragmatic of two workers with equal skill will earn the higher wage, provided both of them dislike earning less than their co-worker. The reason for this perhaps slightly counter-intuitive result is that a worker's pragmatism does not affect their co-worker's payoff, which depends only on the co-worker's wage and the output produced by the match. Thus, if a worker $\mathbf{x}^j$ is happy to match with $\mathbf{x}$ and receive wage w^j, they will be equally happy to match with (x_1, x_2') for the same wage. Therefore, workers of the same skill face the same menu of options, and in particular, the same trade-off between within-team status and wages. Naturally, pragmatic workers choose higher wages and lower status from this menu, whereas status-driven workers choose higher status and lower wages.

5. The economic implications of relative concerns

In this section, I evaluate the implications of heterogeneous relative concerns for sorting, the existence of trickle-down effects and wage inequality. I do this by introducing and analytically solving three special cases of the model: the additive case (Section 5.1), the binary-skill case (Section 5.2) and the multiplicative case (Section 5.3).

5.1. Additive production

I start with a special case of the model in which there are no production complementarities.

Assumption 3 (The Additive Case). The production function satisfies $F(x_1^k, x_1^j) = K(x_1^k) + K(x_1^j)$, where K is increasing and differentiable, the distribution of x_1 is absolutely continuous and the copula $C_{\mathbf{X}}$ of (x_1, x_2) is exchangeable, that is, $C_{\mathbf{X}}(u, v) = C_{\mathbf{X}}(v, u)$ for all $(u, v) \in [0, 1]$.

The case of additive production isolates the effect of relative concerns on sorting, because in the absence of production complementarities aggregate output does not depend on the sorting pattern. Both Frank (1984a) and Fershtman et al. (2006) restrict attention to the additive case, presumably for that very reason. The only distributional requirement is that H's copula is exchangeable ($C_{\mathbf{X}}(u, v) = C_{\mathbf{X}}(v, u)$). Since this assumption is about the copula, marginal distributions remain unrestricted.

Before I start the analysis proper, let me quickly discuss what the equilibrium sorting and payoffs are in the benchmark, that is, when $x_2 = 1$ for all workers. As there are no production complementarities, in the additive case any feasible sorting is sustainable in equilibrium. Since every worker can guarantee themselves the same-match payoff, and $F(x_1^k, x_1^j) = u_S(x_1^k) + u_S(x_1^j)$, it follows that all workers earn their same-match payoff and wage in the benchmark, with $w_B(x_1) = u_B(x_1) = K(x_1)$.

5.1.1. Sorting

To see that Assumption 1 is satisfied in the additive case, it suffices to set $q(\mathbf{x}) = x_2$, $t(q, x_1) = 0.5 q x_1$ and $s(\mathbf{x}) = 0.5 x_1 x_2$. Accordingly, $v_1(\mathbf{x}) = H_{x_2}(x_2)$ and thus Assumption 1 is satisfied by the exchangeability of copula $C_{\mathbf{X}}$. This implies that

$$\mu^*(\mathbf{x}) = [H_{x_1}^{-1}(H_{x_2}(x_2)), H_{x_2}^{-1}(H_{x_1}(x_2))]. \tag{9}$$

Sorting in the additive case exhibits the very same 'proud leader-pragmatic companion' pattern as discussed in Section 4.1: A worker with skill of rank $H_{x_1}(x_1^k) = p$ matches a co-worker with pragmatism of the exact same rank, $H_{x_2}(x_2^j) = p$. As a consequence, $C_{X_1, \mu_1^*(\mathbf{X})}$, the copula of $X_1, \mu_1^*(\mathbf{X})$ is equal to $C_{\mathbf{X}}$, the copula of $\mathbf{X}$. Given that $C_{X_1, \mu_1^*(\mathbf{X})}$ captures the direction and strength of matching in skills (Anderson and Smith, 2024), in the additive case sorting in skills is solely determined by the interdependence between skill and pragmatism—for example, sorting is perfectly positive (negative) assortative in skills if and only if x_2 increases (decreases) deterministically in x_1. Of course, if production complementarities were present, then they would matter for the direction and strength of sorting in skill as well—however, as we will see in Sections 5.2.1 and 5.3.1, even then the distribution of relative concerns continues to play a (possibly dominant) role.

5.1.2. The trickle-down effect of technological change

In this section, I derive the impact of skill-biased technological change (SBTC) on payoffs and wages.

Definition 3. A technological change is a change in the production function from $F(\cdot, \cdot; \theta_1)$ to $F(\cdot, \cdot; \theta_2)$. A technological change is *skill-biased* if, for all $x_1' \geq x_1$ and all y,

$$F(x_1', y; \theta_2) - F(x_1, y; \theta_2) \geq F(x_1', y; \theta_1) - F(x_1, y; \theta_1). \tag{10}$$

My definition of skill-biased technological change is very general and requires only that the difference in output produced by workers of higher skill increases compared to workers of lower skill. Of course, in the additive case this definition of SBTC reduces to the requirement that $K(x_1'; \theta_2) - K(x_1; \theta_2) > K(x_1'; \theta_2) - K(x_1; \theta_1)$.

Let me start by deriving analytical expressions for equilibrium payoffs and wages. In the additive case, only workers for whom $H_{x_1}(x_1) = H_{x_2}(x_2)$ same-match; thus the premise of Theorem 3 is satisfied as long as $(x_1, H_{x_2}^{-1}(H_{x_1}(x_1)) \in \mathrm{supp}(H)$ for every $x_1 \in D_{x_1}$. If that is the case, then (7) and (8) simplify to

$$u^*(\mathbf{x}) = 0.5 F(\mathbf{x_1}) x_2 + \int_{x_1}^{H_{x_1}^{-1}(H_{x_2}(x_2))} \left(x_2 - H_{x_2}^{-1}(H_{x_1}(y)) \right) K'(y) \, \mathrm{d}y, \tag{11}$$

$$w^*(\mathbf{x}) = 0.5 F(\mathbf{x_1}) + \int_{x_1}^{H_{x_1}^{-1}(H_{x_2}(x_2))} \left(1 - H_{x_2}^{-1}(H_{x_1}(y)) \right) K'(y) \, \mathrm{d}y. \tag{12}$$

Thus, in addition to the worker's own type $\mathbf{x}$, payoffs and wages depend on their co-worker's skill ($H_{x_1}^{-1}(H_{x_2}(x_2))$), the additional output ($K'(y)$) produced by workers with skill between the worker's own skill and their co-worker's skill and on the pragmatism of the co-workers of those workers ($H_{x_2}^{-1}(H_{x_1}(y))$). The following result follows by inspection of (11) and (12).

Corollary 1. *Under Assumption 3 and the premise of Theorem 3, skill-biased technological change increases $u^*(\mathbf{x}) - u_S(x_1)$ (and $w^*(\mathbf{x}) - w_S(x_1)$) for all $\mathbf{x}$ such that $H_{x_2}(x_2) > H_{x_1}(x_1)$ (and $x_2 < 1$).*

Corollary 1 establishes that the benefits of skill-biased technological change trickle-down to all workers who rank higher in pragmatism

[11] This follows from the fact that $\frac{\partial}{\partial x_2} u^*(\mathbf{x}) = 0.5 F(x_1, \mu_1^*(\mathbf{x})) > 0$.

[12] We could, for example, divide everyone's utility by some sufficiently steep increasing function of x_2. The resulting model would not have transferable utility, but would result in the same sorting and wages—and the payoffs in such a model would be decreasing in x_2.

than they do in skill—that is, SBTC increases their payoffs and wages above and beyond any increase in their own productivity (as captured by changes to their same-matching payoffs and wages). Equivalently, any change in technology that makes 'proud leaders' more productive benefits their 'pragmatic companions' as well.

To better understand these results, consider a change in K that increases $K'(x_1)$ for all x_1 above some cut-off $\hat{x}$, but leaves K unchanged otherwise. In the presence of relative concerns, such a change in K raises the payoffs of very pragmatic low-skill workers, as their co-workers have skill above the cut-off. These 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 case, 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 cut-off, and thus do not see any increase in payoff. Even worse, if the cut-off $\hat{x}$ is very low, then workers with skill just above the cut-off will have low skill and yet, if their pragmatism is lower still, their wages and payoffs will increase less than in the benchmark. Thus, the welfare impact of SBTC is starkly different for low-skill workers with weak relative concerns than for those with strong ones.

The trickle-down effect triggered by SBTC can be very substantial indeed. For comparison, assume $K(\underline{x})$ remains unchanged and denote by ΔK the *per capita* increase in output due to technological change. Notice that, because the sorting pattern does not matter for output in the additive case, we have

$$\Delta K \equiv \int_{D_{x_1}} \Delta K(y) \mathrm{d}H_{x_1}(y),$$

where $\Delta K(x) = K(x;\theta_2) - K(x;\theta_1)$. Integrating ΔK by parts and rearranging yields

$$w^*(\underline{x}_1, \bar{x}_2; \theta_2) - w^*(\underline{x}_1, \bar{x}_2; \theta_1) - \Delta K = \int_0^1 \frac{\partial}{\partial x_1} \Delta K(H_{x_1}^{-1}(s))(s - 0.5(H_{x_2}^{-1}(s) + 1)) \mathrm{d}H_{x_1}^{-1}(s).$$

In the case of skill-biased technological change, the expression inside the integral has to be negative for any $s \leq 0.5$. For $s > 0.5$, however, the expression is positive if $H_{x_2}^{-1}$ takes low enough values, that is, when workers have sufficiently strong relative concerns. This implies, in particular, that if $\hat{x}$ is sufficiently high, so that $\frac{\partial}{\partial x_1}\Delta K(H_{x_1}^{-1}(s)) = 0$ for low s, then the wage increase of the most pragmatic low-skill worker may be greater than if they simply received the *per capita* increase in output. Part (ii) of the following proposition formalises this insight; part (i) states the analogous result for changes in payoffs.

Proposition 1. *Suppose that Assumption 3 and the premise of Theorem 3 are satisfied and technological change is such that there exists some $\hat{x} \in (\underline{x}_1, \bar{x}_1)$ such that $\Delta K(x) = \frac{\partial}{\partial x_1}\Delta K(x) = 0$ for $x \leq \hat{x}$ and $\frac{\partial}{\partial x_1}\Delta K(x) > 0$ for all $x > \hat{x}$. (i) If $H_{x_2}(\bar{x}_2(2x_2 - 1)) > x_2$ for all $x_2 \in (H_{x_1}(\hat{x}), 1)$, then $u^*(\underline{x}_1, \bar{x}_2; \theta_2) \geq u^*(\underline{x}_1, \bar{x}_2; \theta_1) + \bar{x}_2\Delta K$. (ii) If $H_{x_2}(2x_2 - 1) > x_2$ for all $x_2 \in (H_{x_1}(\hat{x}), 1)$, then $w^*(\underline{x}_1, \bar{x}_2; \theta_2) - w^*(\underline{x}_1, \bar{x}_2; \theta_1) \geq \Delta K$.*

Proposition 1(i) shows that, as long as the population as a whole cares about relative concerns sufficiently strongly, SBTC that affects exclusively workers of very high-skill benefits the most pragmatic low-skill worker more than a policy that redistributes the gains from SBTC equally among all workers. Intuitively, if most workers have much stronger relative concerns 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 is very pragmatic, she does not mind the increase in the gap between her own and her co-workers' wages all that much.

Proposition 1(ii) shows an analogous result for the wage of the most pragmatic, lowest-skilled worker, with a slightly modified condition regarding the distribution of relative concerns. Note that the distributional conditions imposed by Proposition 1 are not empty. For instance, they are both satisfied if $H_{x_2}(x_2) = x_2^a$, $a < 0.5$ and $H_{x_1}(\hat{x})$ is sufficiently close to 1.

5.1.3. Wage inequality

Intuitively, as long as all workers dislike earning less than their co-workers (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 the 'proud leaders', whose same-match wages are high, will earn less than their same-match wages, whereas the 'pragmatic companions', whose same-match wages are low, will earn more. The following result formalises this intuition.

Proposition 2. *Under Assumption 3 and the premise of Theorem 3, if $\bar{x}_2 \leq 1$, then $Var(w_S) = Var(w_B) \geq Var(w^*)$.*

Proposition 2 shows that in the absence of production complementarities, equilibrium wages in the model with relative concerns are less unequal than same-match wages. Since under Assumption 3 all workers earn the same-match wage in the benchmark, this implies that relative concerns decrease wage inequality.

5.2. Binary-skills

The workers in this model care both about their relative position within the team and about their own wages. If high- and low-skill workers are either complements or substitutes, the sorting pattern determines how high is the average wage in the economy. Therefore, the equilibrium sorting must be shaped by the interplay between these production complementarities and relative concern heterogeneity. I first analyse this interplay in a setting with just two levels of skill (high and low) before moving to continuous skills in Section 5.3.

Assumption 4 (The Binary-Skill Case). Skills are binary, with $x_1 \in \{l, h\}$, where $h > l$, and $\Pr(X_1 = h) = 0.5$.[13]

Restricting skills to take only two values is the simplest way to introduce production complementarities into the model. Specifically, the production function is supermodular (high-skill workers are complements with other high-skill workers) if

$$a_F \equiv \frac{F(h,h) - F(l,h)}{F(h,l) - F(l,l)} > 1;$$

conversely, the production function is submodular (i.e., high-skill workers are complements with low-skill workers) if $a_F < 1$. If $a_F = 1$ then no production complementarities are present, exactly as in the additive case of Section 5.1.

As a benchmark, consider what happens in the binary-skill case if $x_2 = 1$ for all workers. If F is strictly supermodular ($a_F > 1$), then all workers same match, so that $\mu_1^*(x_1) = x_1$. Naturally, each worker receives then the same-match payoff/wage, with $u_S(x_1) = w_S(x_1) = 0.5F(x_1, x_1)$. If F is strictly submodular ($a_F < 1$), then there is NAM in skills and all workers cross-match. Any division of the additional output that is produced by cross-matching can be sustained in equilibrium, which implies that the function

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

is a valid equilibrium payoff and wage function for any $\alpha_l, \alpha_h \in [0,1]$ such that $\alpha_h + \alpha_l = 1$. The parameters α_l, α_h can be equivalently thought of

[13] If $\Pr(X_1 = h) \neq 0.5$ 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 $\Pr(X_1 = h) = 0.5$. Thus, the assumption that $\Pr(X_1 = h) = 0.5$ is without loss for the characterisation of equilibrium.

as a way of denoting the different possible equilibria or as the bargaining power of low- and high-skill workers respectively. In either case, any comparative statics results involving cases where payoffs and wages are not unique will assume that α_l, α_h remain unchanged.

5.2.1. Sorting with production complementarities

Let me now characterise equilibrium sorting in the binary-skill case. Setting $q(\mathbf{x}^k) = 0.5\left(F(x_1^k, h) - F(x_1^k, l)\right)x_2^k$, $t(q, x_1) = q(x_1 - l)/(h - l)$ and $s(\mathbf{x}^k) = 0.5F(x_1^k, l)x_2^k$ ensures that Assumption 1 is satisfied; I show in Appendix A that Assumption 2 is satisfied as well.

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

$$v_1(\mathbf{x}) = \sum_{j\in\{l,h\}} 0.5\left[G_j\left(x_2(F(x_1,h) - F(x_1,l))/(F(j,h) - F(j,l))\right)\right]. \qquad (13)$$

By Theorem 2 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

$$a_F = \frac{G_l^{-1}(1-\bar{y})}{G_h^{-1}(\bar{y})} \qquad (14)$$

otherwise. A rearrangement of (13) yields then that high-skill workers with $x_2 \le G_h^{-1}(\bar{y})$ match low-skill workers with $x_2 \ge G_l^{-1}(1-\bar{y})$ and all remaining workers same-match.

As we can see, the 'proud leader-pragmatic companion' sorting pattern continues to hold even with production complementarities, in the sense that high-skill workers with very strong relative concerns match very pragmatic low-skill workers. However, how "proud" the leader and how "pragmatic" the companion need to be for such a match to happen depends very much on the strength of complementarities between high- and low-skill workers. If high-skill workers become stronger complements to low-skill workers (a_F decreases), then the difference in the pragmatism between low- and high-skill workers needed to warrant a match also decreases, leading to an increase in the number of cross-matching teams.

Crucially, however, relative concerns/pragmatism continue to matter for sorting. Specifically, when high-skill workers become more pragmatic (G_h^{-1} increases), the number of cross-matching pairs decreases; conversely, if it is the low-skill workers who become more pragmatic, the number of cross-matching teams increases. Importantly, the impact of relative concerns on sorting can dominate that of production complementarities: We can fix the production function and yet produce any degree of sorting in skill in equilibrium, simply by altering the distribution of preferences conditional on skill (G_h and G_l). 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 all high-skill workers are more pragmatic than the most pragmatic low-skill worker, that is, $G_h^{-1}(0)/G_l^{-1}(1) \ge 1$. In that case, a high-skill worker faces a trade-off between maximising their own wage by matching a low-skill worker, and minimising the within-firm wage differential by same-matching. If high-skill workers are much more pragmatic than low-skill workers ($G_h^{-1}(0)/G_l^{-1}(1) \ge 1/a_F$), 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.

5.2.2. Trickle down effects of NAM-biased technological change

In general, by Theorem 3, technological change affects payoffs and wages both directly, through changes to $F(x_1', y) - F(x_1, y)$, and indirectly, through its impact on the skill component of the matching function μ_1^* and the preference of the same-matching workers $x_2^*(x_1)$. In the additive case, the indirect channel was absent—because sorting did not depend on the production function; in the binary-skill case, however, this is no longer the case. Consequently, the (possible) trickle-down effects of technological change will depend not only on whether technological change is skill-biased, but also on whether it is biased toward PAM or NAM in skills.

Definition 4. A technological change is *NAM-biased* (*PAM-biased*) if, for $x' > x, y' > y$,

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

Any change in technology that decreases (increases) a_F will be called NAM- (PAM-) biased, because a decrease (increase) in a_F raises the number of cross-matching (same-matching) teams. Let me now derive analytical expressions for the cross-matching benefit and payment in the binary-skill case (using Theorem 3).[14] Evaluating the integrals in (7) and (8) yields

$$u^*(\mathbf{x}) - u_S(\mathbf{x}) = \frac{F(\mu_1^*(\mathbf{x}), x_1) - F(\mathbf{x_1})}{2}\left(x_2 - z(x_1, 0.5)\right) \qquad (15)$$

$$w^*(\mathbf{x}) - w_S(x_1) = \frac{F(\mu_1^*(\mathbf{x}), x_1) - F(\mathbf{x_1})}{2}\left(1 - z(x_1, 0.5)\right). \qquad (16)$$

In the binary-skill case, skill-biased technological change is certain to generate a trickle-down effect for 'pragmatic companions' only when it is also NAM-biased. More formally, any technological change that is both skill- and NAM-biased increases both the cross-matching benefit and the cross-matching payment (provided $z(l, 0.5) < 1$) for all low-skill workers who end up with high-skill co-workers after the change—specifically, low-skill workers with $x_2 > G_l^{-1}(1 - \bar{y}(\theta_2))$. To understand this mechanism, consider the marginal low-skill worker, $(l, z(l, 0.5))$, who is indifferent between same- and cross-matching. When this marginal worker cross-matches, they earn less than their high-skill co-worker and therefore (if $z(l, 0.5) < 1$) must earn more than the same-match wage to remain indifferent. The required compensation depends positively on the high-skill worker's productivity advantage $(F(h, l) - F(l, l))$ and negatively on the marginal low-skill worker's pragmatism ($z(l, 0.5) = G_l^{-1}(1 - \bar{y})$). As it is always the most pragmatic low-skill workers that cross-match, NAM-biased technological change, by increasing the number of cross-matching teams, must decrease the pragmatism of the marginal low-skill worker and thus increase the compensation they receive. Additionally, some low-skill workers who previously same-matched will transition to cross-matching, raising their cross-matching payment and benefit from zero to positive values.

Conversely, PAM-biased technological change dampens the trickle-down effect of skill-biased technological change by increasing the marginal low-skill worker's pragmatism and forcing some low-skill workers into same-matching arrangements. When the bias toward PAM is sufficiently strong, all low-skill workers may gain less from SBTC than they would have earned under same-matching. This outcome occurs, for instance, when technological change increases only $F(\mathbf{h})$ to such an extent that a_F rises above $G_l^{-1}(1)/G_h^{-1}(0)$. Such a change would induce perfect PAM, making all previously cross-matching low-skill workers worse off while leaving their same-matching payoff unchanged.

5.2.3. Wage inequality

If the production function is supermodular, then relative concerns' heterogeneity decreases wage inequality in comparison to the benchmark—which was also the case in the additive case. If the production function is submodular, however, then the presence of relative concerns can actually increase wage inequality.

[14] The premise of Theorem 3 is clearly satisfied as long as $\bar{y} \neq 1$, because high-skill workers with $x_2 > G_h^{-1}(\bar{y})$ and low-skill workers with $x_2 < G_l^{-1}(1 - \bar{y})$ same-match.

Proposition 3. *Consider an economy (F, H) which satisfies Assumption 4. (i) If $\bar{x}_2 \leq 1$, then $Var(w_S) \geq Var(w^*)$. If, in addition, $Var(w_B) \geq Var(w_S)$, which is always satisfied for $a_F \geq 1$, then $Var(w_B) \geq Var(w^*)$. (ii) If, instead, $Var(w_B) < Var(w_S)$, then there exists a distribution of traits $\tilde{H}$, such that (a) economy $(F, \tilde{H})$ satisfies Assumption 4, (b) the marginal distribution of skill is the same under H and $\tilde{H}$ ($H_{x_1} = \tilde{H}_{x_1}$), and (c) $Var(w_B) < Var(w^*)$.*

Proposition 3(i) shows that, much as in the additive case, in the binary-case the presence of heterogeneity decreases wage inequality in comparison to the case where all workers earn same-match wages. The intuition also remains the same: In any match between a 'proud leader' and a 'pragmatic companion' the leader earns less than the same-match wage, whereas the companion earns more. The immediate consequence is that if the production function is supermodular, so that all workers earn a same-match wage in the benchmark, then wage inequality is lower in a model with relative concerns heterogeneity than in the benchmark.

Proposition 3(ii) shows that, in contrast, if the benchmark wage distribution is less unequal than the same-match one, then there must exist a preference distribution that increases wage inequality above benchmark levels. The reason is that, as discussed in Section 5.2.1, if high-skill workers are sufficiently more pragmatic than low-skill workers, then same-matching obtains in equilibrium. Finally, notice that if the production function is submodular, then $\mathrm{Var}(w_B) = \mathrm{Var}(w_S)(1 + (1 - 2\alpha_l)(1 - a_F)/(1 + a_F))^2$, which is strictly less than $\mathrm{Var}(w_S)$ if and only if $\alpha_l > 0.5$. Therefore, Proposition 3(ii) implies that if $a_F < 1$, $\alpha_l > 0.5$ and high-skill workers are sufficiently more pragmatic than low-skill workers, then heterogenous relative concerns increase wage inequality in comparison to the benchmark.

5.3. Multiplicative production

The last specification allows for both continuously distributed skills (unlike the binary-skill case) and production complementarities (unlike the additive case), but at the cost of making specific functional form assumptions.

Assumption 5. The production function is $F(x_1^k, x_1^j) = A + ((x_1^k x_1^j)^c - 1)/c$, 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 $i = \sqrt{-1}$.

Reassuringly, the functional form assumed in this specification is very natural. $F(x_1^k, x_1^j) = A + ((x_1^k x_1^j)^c - 1)/c$ is the standard multiplicative production function commonly used in assignment models (e.g., Costrell and Loury, 2004; Tervio, 2008) but it is extended to accommodate not only super- ($c > 0$) but also submodularity ($c < 0$). The family 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 contains the log-normal, the log-t-student, and the (log) scale-mixtures. Log-elliptical distributions are widely used in asset pricing (e.g., Kim, 1998), portfolio choice (e.g., Owen and Rabinovitch, 1983), and machine learning (e.g., Louizos et al., 2017), as they combine the appealing properties of log-normal distributions with the ability to model heavy tails.

In the benchmark, $c > 0$ implies PAM in skills and workers earning their same-matching payoffs and wages, with $u_S(x_1) = w_S(x_1) = 0.5(A + (x_1^{2c} - 1)/c)$. Conversely, $c < 0$ implies NAM in skills, with $H_{x_1}(\mu_1^*(x_1)) = 1 - H_{x_1}(x_1)$. Even though the premise of Theorem 3 is not satisfied here, one can apply the Envelope Theorem to (3) to derive $\frac{\partial}{\partial x_1}u(\mathbf{x})$ (instead of $\frac{\partial}{\partial x_2}u(\mathbf{x})$) and then use the fact that the worker of median skill same-matches under NAM to show that

$$u_B(x_1^k) = w_B(x_1^k) = 0.5(Ac + \exp(2c\delta_1) - 1)/c + \exp(2c\delta_1)(\ln x_1 - \delta_1).$$

In my discussion of the multiplicative case, I will focus on those economic implications that differ from the binary-skill case: Namely, the impact of technology on sorting and the trickle-down effect. The impact of relative concerns on wage inequality is qualitatively the same as in the binary-skill case—Proposition 3 holds analogously for the multiplicative case (see Appendix A for a proof). What is worth noting, however, is that it remains possible that $\mathrm{Var}(w_B) < \mathrm{Var}(w_S)$ if production is submodular.[15] This demonstrates that the presence of relative concerns can increase wage inequality compared to the benchmark even when skills are continuous, and thus this possibility is a general phenomenon rather than an artifact of the non-uniqueness of wages in the binary-skill case.

5.3.1. Sorting when skills are continuous

The multiplicative case satisfies Assumption 1, with, for example, $q(\mathbf{x}) = x_1^c x_2$, $t(q, x_1) = 0.5qx_1^c/c$ and $s(\mathbf{x}) = 0.5(A - 1/c)x_2$. It follows that $v_1(\mathbf{x}^k)$ is equal to worker's $\mathbf{x}^k$ rank in the distribution of $\bar{v}_1 \equiv c\ln x_1 + \ln x_2$. 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 $(\ln x_1, \bar{v}_1) \sim EC_2(\mathbf{A}\Delta, \mathbf{A}\Omega\mathbf{A}^T; \phi)$, where $A \equiv \left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right)$. The copula of (x_1, v_1) must thus be an elliptical copula; and because all elliptical copulas are exchangeable, Assumption 2 is satisfied as well.

Denote the square root of the ratio of variances of $\bar{v}_1$ and $\ln x_1$ by r: it follows then from Theorem 1 and some linear algebra that:

$$\begin{aligned} \mu_1^*(\mathbf{x}) &= \left(x_1^c x_2\right)^{1/r} e^{\delta_1\left(1-\frac{c}{r}\right)-\frac{\delta_2}{r}}, \\ \mu_2^*(\mathbf{x}) &= e^{\left(1+\frac{c}{r}\right)[\delta_1(c-r)+\delta_2]}\left(\left(x_1^{\frac{r^2}{c}-c}\right)/x_2\right)^{c/r}. \end{aligned} \tag{17}$$

As usual, if $c > 0$ ($c < 0$), and thus if the production function 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 pragmatic co-workers. This can be seen very clearly by examining the correlation between the worker's skill x_1 and their co-worker's skill $\mu_1^*(\mathbf{x})$, denoted by ρ and equal to

$$\rho = \frac{c\sqrt{\omega_{11}} + \rho_{12}\sqrt{\omega_{22}}}{\sqrt{c^2\omega_{11} + 2c\rho_{12}\sqrt{\omega_{11}\omega_{22}} + \omega_{22}}} = \begin{cases} \frac{c\sqrt{\omega_{11}}+\sqrt{\omega_{22}}}{|c\sqrt{\omega_{11}}+\sqrt{\omega_{22}}|} & \text{if } \rho_{12} = 1 \\ \frac{c\sqrt{\omega_{11}}-\sqrt{\omega_{22}}}{|c\sqrt{\omega_{11}}-\sqrt{\omega_{22}}|} & \text{if } \rho_{12} = -1, \end{cases} \tag{18}$$

where ω_{ij} is the i-row and j column element of the covariance matrix Ω and ρ_{12} is the correlation between x_1 and x_2. Similarly to the binary-skill case, the strength and the direction of sorting in skill depend both on production complementarities (through c), the correlation between skills and preferences ρ_{12} and on the variances of skill (ω_{11}) and relative concerns (ω_{22}). Specifically, as $\frac{\partial}{\partial c}\rho = (1 - \rho^2)/r$, a decrease in c (which is a NAM-biased change in technology) makes sorting more negative and assortative in skills. However, if $\rho_{12} = -1$ ($\rho_{12} = 1$) and $\sqrt{\omega_{22}} > |c|\sqrt{\omega_{11}}$ then equilibrium sorting exhibits perfect NAM (PAM) in skill. Therefore, it remains possible to fix both c and ω_{11} (i.e., the production side of the economy), and yet produce any degree of sorting in skills by just varying the correlation between skill and preference and the variance of x_2.

In broad strokes, therefore, production complementarities and relative concerns affect sorting similarly in the continuous skill case as in the binary skill case. There, is however, one noteworthy difference, which is that when skills are log-elliptically distributed, the measure of same-matching workers is 0 as long as Ω is of full rank. This is because, instead of same-matching, a very highly skilled and pragmatic (but not very pragmatic!) worker can now match a very pragmatic and highly (but not very highly!) skilled co-worker. As a consequence, the impact of NAM-biased technological change on the matches of individual workers can be quite different from that in the binary-skill case, even if its

[15] To see this, assume that $c = -1$ and skills are log-normally distributed. Then $\mathrm{Var}(w_B) - \mathrm{Var}(w_S) = 0.25\exp(4\delta_1)\left(4\omega_{11} - \exp(8\omega_{11}) + \exp(4\omega_{11})\right)$. Clearly, at $\omega_{11} = 0$ both the expression itself and its derivative with respect to ω_{11} are equal to zero. In addition, this expression is concave in ω_{11}. It follows, therefore, that $\mathrm{Var}(w_B) < \mathrm{Var}(w_S)$.

impact on the sorting pattern as a whole is similar. This can be seen by taking the derivative of μ_1^* with respect to c:

$$\frac{\partial}{\partial c}\ln(\mu_1^*(\mathbf{x})) = \left[(1-\rho)(\ln x_1 - \delta_1) + \rho\left(\ln x_1 - \ln\mu_1^*(\mathbf{x})\right)\right]/r. \quad (19)$$

Evidently, if there is PAM in skill in equilibrium ($\rho > 0$), then NAM-biased technological change (a decrease in c) improves the skill of the co-worker for any worker with below median skill $\ln x_1 < \delta$ who is matched to a more skilled co-worker $\mu_1^*(\mathbf{x}) > \ln x_1$. If, however, there is NAM in equilibrium, then a change in technology that makes sorting even more negative and assortative will decrease the co-worker's skill for below median-skill workers whose co-workers are sufficiently skilled (i.e., if μ_1^* is much greater than x_1).

As this result has important consequences for the trickle-down effect, let me build understanding with an example. Consider the worker $(\exp(\delta_1), H_{x_2}^{-1}(0.99))$—a very pragmatic worker of median skill. If c is very high, then production complementarities are the dominant force determining sorting, and $(\exp(\delta_1), H_{x_2}^{-1}(0.99))$ will match a co-worker of very similar (if a bit higher) skill. As c decreases, relative concerns start playing an outsized role in matching, and thus the high pragmatism of $(\exp(\delta_1), H_{x_2}^{-1}(0.99))$ will net them co-workers with higher than median skill. This process will reach its peak for $c \approx 0$, in which case $(\exp(\delta_1), H_{x_2}^{-1}(0.99))$ will be matched to a co-worker of skill $\approx H_{x_1}^{-1}(0.99)$. After that, as c keeps decreasing, production complementarities start taking over again: and because our worker has median skill, same-matching remains their output maximising match, so that the skill of their co-worker will start to decrease again.

5.3.2. The trickle-down effect revisited

Technology continues to affect the cross-matching benefits and payments through the direct and indirect channels that were discussed in Section 5.2.2. To see this, first notice that $x_2^*(x_1) = x_1^{r-c}\exp(\delta_2 - (r-c)\delta_1)$. As long as x_1 and x_2 are not perfectly correlated there will exist a worker of type $(x_1, x_1^{r-c}\exp(\delta_2 - (r-c)\delta_1))$ for every x_1, in which case the premise of Theorem 3 is satisfied and (7) and (8) can be written as

$$u^*(\mathbf{x}) = 0.5F(\mathbf{x_1})x_2 + 0.5\int_{x_1}^{\mu_1^*(x_1,x_2)}\left(x_2 - z(x_1; H_{x_1}(s))\right)\frac{\partial}{\partial x_1^j}F(x_1,s)\,\mathrm{d}s, \quad (20)$$

$$w^*(\mathbf{x}) = 0.5F(\mathbf{x_1}) + 0.5\int_{x_1}^{\mu_1^*(x_1,x_2)}\left(1 - z(x_1; H_{x_1}(s))\right)\frac{\partial}{\partial x_1^j}F(x_1,s)\,\mathrm{d}s. \quad (21)$$

Recall that $z(x_1;\cdot)$ is the inverse of v_1 with respect to x_2. Naturally, skill-biased technological change increases $\frac{\partial}{\partial x_1^j}F(x_1,s)$, just as in the additive and binary-skill cases. As I discussed in the preceding section, NAM-biased technological change need not increase μ_1^* (and thus decreases $z(x_1; H_{x_1}(s))$) if the equilibrium features NAM in skills.

Proposition 4. *Consider some* $\mathbf{x}$ *such that* $\ln x_1 \le \min\{\delta_1, \mu_1^*(\mathbf{x})\}$. *Under Assumption 5, a small decrease in* c *increases* $u^*(\mathbf{x}) - u_S(x_1)$ *(and* $w^*(\mathbf{x}) - w_S(x_1)$*) as long as* $\ln x_1 - \delta_1 \le (\ln\mu_1^*(\mathbf{x}) - \ln x_1)\rho/(1-\rho)$, $x_1\mu_1^*(\mathbf{x}) < 1$ *(and* $x_2 < 1$*).*

Proposition 4 states that the benefits of a fall in c trickle-down to low-skill workers with more skilled co-workers if (i) their skill is sufficiently low compared to the median skill and the difference between the co-worker's and their own skill ($\ln x_1 \le \delta_1 + (\ln\mu_1^*(\mathbf{x}) - \ln x_1)\rho/(1-\rho)$) and (ii) the product of the two teammates' skills is less than 1. The first condition is needed to guarantee that the decrease in c improves the co-worker's skill, which then ensures that the indirect channel produces the trickle-down effect. It follows from the discussion in Section 5.3.1 that this condition is implied by $\ln x_1 < \min\{\delta_1, \mu_1^*(\mathbf{x})\}$ if the correlation between workers' and their co-workers' skills is positive. If, instead, the correlation is negative, then the indirect effect may affect the cross-matching benefit negatively for pragmatic workers with skills slightly below median, making the overall effect ambiguous. The second condition is needed to ensure that the decrease in c actually constitutes a skill-biased technological change for this team.

6. Extensions

In this section, I discuss a number of extensions and robustness checks, the chief of which are introducing outsourcing as a mechanism to mitigate harmful social comparisons (Section 6.1) and expanding the binary-skill case to incorporate inequity aversion rather than relative concerns (Section 6.2).

6.1. 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. Since 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 $c \ge 0$: 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 5. *Suppose that Assumption 4 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 Section 5.2.1.*

(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*

$$\frac{G_l^{-1}(1-y^o)}{G_h^{-1}(y^o)} = \min\left\{\frac{G_l^{-1}(1)}{G_h^{-1}(0)}, \max\left\{\frac{G_l^{-1}(0)}{G_h^{-1}(1)}, b_F\right\}\right\}. \quad (22)$$

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* $x_2 > G_h^{-1}(y^o)$*, while the non-outsourced low-skill workers match with high-skill workers of* $x_2 \le G_h^{-1}(y^o)$.

Proposition 5 characterises equilibrium sorting and outsourcing decisions. The crucial piece of intuition is that teams formed by 'proud leaders' and 'pragmatic companions' are not interested in outsourcing, because social comparisons increase their joint welfare: The status-sensitive high-skill worker gains more utility from their high relative position than the pragmatic low-skill worker loses. In contrast, in teams formed by 'pragmatic leaders' and 'proud companions' social comparisons decrease welfare; accordingly, these teams—and these teams only—may benefit from outsourcing.

Consider first the extreme case in which outsourcing is costless. If high- and low-skill workers are substitutes in production ($s_F \le 0$), then there is no reason for 'pragmatic leaders' to ever match 'proud companions; thus, no teams outsource and the sorting pattern is exactly the same as in Section 5.2.1. If, instead, high- and low-skill workers are complements in production ($s_F > 0$), then 'pragmatic leaders' want to match 'proud companions' in order to produce more output—and outsourcing allows them do to so without suffering detrimental social comparisons!

In this case of costless outsourcing, the *marginal team*—that is a team which is indifferent between forming one or two firms—will consist of two workers with different skills but identical preferences. Of course, high-skill workers forming (non-)outsourcing teams will have weaker (stronger) relative concerns than the high-skill worker in the marginal team; and *vice versa* for the low-skill workers.

When outsourcing is very costly relative to the strength of production complementarities ($c_F > s_F$) then, naturally, no outsourcing takes place, because any hypothetical 'pragmatic leader—proud companion' teams will prefer to same-match instead. If, however, the cost is positive but low in comparison to production complementarities ($c_F \in (0, s_F)$), then only teams in which the high-skill worker is sufficiently more pragmatic than the low-skill worker will outsource. To see why, consider the marginal team again. As outsourcing comes at a cost, social comparisons must also be costly within that team, and hence the high-skill worker must be strictly more pragmatic than their low-skill companion.

6.1.1. The impact of SBTC on outsourcing and sorting

Let us 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 (22) in Proposition 5 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.

6.1.2. Discussion

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 for 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.

A compelling feature of this extension 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.

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' skills is large and positive (see Fig. 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 assortative matching in skills, as long as low-skill workers have stronger relative concerns than high-skill workers (see Sections 5.2.1 and 5.3.1). 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 et al. (2021) as the outcome of within-team problem solving.

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 5 that this would have the same qualitative impact on outsourcing and sorting as SBTC.

Finally, 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.

6.2. 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:

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

Here, $\gamma \geq 0$, $\beta \in [0, 1)$—to ensure that, keeping the team's output constant, utility increases in own wage—and both γ and β differ across individuals. In general, a sorting model in which workers have this utility function is different from mine. However, if skills are binary, then the inequity aversion model is isomorphic to my model if $x_2 = 1/(1 - \beta)$ for high-skill workers and $x_2 = 1/(1 + \gamma)$ 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 workers 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.

Economic implications of inequity aversion. In the class of inequity aversion equivalent preferences skills are perfectly negatively correlated with relative concerns. Therefore, for the reasons discussed in Section 5.2.1, inequity aversion equivalent preferences push sorting to be more positive and assortative in skills. Thus, the following result follows directly from the discussions above, in Sections 5.2.1 and 5.2.3.

Corollary 2. *Suppose that all workers have Fehr and Schmidt (1999) utility and Assumption 4 is satisfied. Denote the lowest value of β (γ) among high- (low-)skill workers by $\underline{\beta}_h$ ($\underline{\gamma}_l$). If $a_F \in (\frac{1-\underline{\beta}_h}{1+\underline{\gamma}_l}, 1)$ and $\alpha_l > 0.5$, then $Var(w_B) < Var(w^*)$.*

In other words, if low- and high-skill workers are complements to each other, and same-match wages are more unequal than benchmark wages ($\alpha_l > 0.5$), then *sufficiently strong inequity aversion among workers increases wage inequality in the economy*. This 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.

Finally, because inequity aversion equivalent preferences satisfy the premise of Proposition 5, the impact of SBTC on outsourcing and sorting under inequity aversion is the same as under relative concerns.[16]

6.3. 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:

$$U(w^k, \bar{w}^{k,j}, \bar{w}; x_2^k, x_3^k) \equiv x_2^k w^k + (1 - x_2^k)(w^k - \bar{w}^{k,j}) + x_3^k(\bar{w}^{k,j} - \bar{w}), \quad (23)$$

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.

Defining a new random variable $\tilde{x}_2 \equiv x_2 + x_3$ lets us rewrite (23) as

$$U(w^k, \bar{w}^{k,j}, \bar{w}; \tilde{x}_2^k, x_3^k) = \tilde{x}_2^k w^k + (1 - \tilde{x}_2^k)(w^k - \bar{w}^{k,j}) - x_3^k \bar{w}.$$

Since any pair of workers is of measure zero, the sorting decision of any individual worker has no impact on $x_3^k \bar{w}$, and thus this term can be dropped. Therefore, 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.[17]

6.4. 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).[18] 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.

The simple KUJ specification used here is needed to fully characterise sorting in settings with continuously distributed skills. As discussed in Sections 2 and 3.3, even with perfectly transferable utility, characterising 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.

6.5. 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- and 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. Since 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, if skills are continuous, then Assumption 1 is difficult to satisfy for production functions constructed from asymmetric ϕ functions. 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 6.1 (see Section 6.1.2 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 continuous cases serves as a reassurance that the symmetry of ϕ is not driving the main results.

6.6. 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.[19] I show this formally in Appendix A, but the intuition is simple: As discussed in Section 4.2, workers' strength of relative concerns does not affect their co-workers' 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 who offer them the same wage, and thus workers have no incentive to lie about the strength of their relative concerns.

7. 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.

Since 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 implications for wage inequality depend critically on the production technology. Under supermodular production, wage inequality

[16] Provided that the cost of outsourcing is neither too high nor too low, so that both outsourcing and non-outsourcing teams co-exist.

[17] 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.

[18] Apart from the articles discussed in the main body, KUJ utility functions have been adopted, for example, in Clark and Oswald (1998); Ghiglino and Goyal (2010); Barnett et al. (2019) and Langtry and Ghinglino (2025).

[19] 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 et al., 2014; Liu, 2020, 2024), this literature remains silent on one-sided problems.

is lower compared to a model without social comparisons. Under submodular production, however, the presence of heterogeneous relative concerns may increase wage inequality.

Finally, 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.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author used Claude.ai to improve language and readability. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Paweł Gola reports that financial support was provided by UK Research and Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Omitted proofs and derivations

Proof of Theorem 2. Define $\mathbf{v}(\mathbf{x}^k, \mathbf{x}^j) \equiv (v_1(\mathbf{x}^k), x_1^j)$, 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_1^j is a complement to $v_1(\mathbf{x}^k)$ and x_1^k 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. First, define

$$\pi(\mathbf{v}(\mathbf{x}^k, \mathbf{x}^j), \mathbf{v}(\mathbf{x}^j, \mathbf{x}^k)) \equiv t(J^{-1}(v_1(\mathbf{x}^k)), x_1^j) + t(J^{-1}(v_1(\mathbf{x}^j)), x_1^k)$$

$$V_{P_\pi}(\mu) \equiv \int_{D_{\mathbf{x}}} \pi(\mathbf{v}(\mu(\mathbf{x}), \mathbf{x}), \mathbf{v}(\mathbf{x}, \mu(\mathbf{x})))\mathrm{d}H(\mathbf{x}).$$

Assumption 1 implies that

$$\Pi(\mathbf{x}^k, \mathbf{x}^j) = \pi(\mathbf{v}(\mathbf{x}^k, \mathbf{x}^j), \mathbf{v}(\mathbf{x}^j, \mathbf{x}^k)) + s(\mathbf{x}^k) + s(\mathbf{x}^j).$$

Since 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

$$\max_{\mu \in S(H)} V_{P\mathbf{x}}(\mu) = 2E_H(s(\mathbf{x})) + \max_{\mu \in S(H)} V_{P\pi}(\mu).$$

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 $V_{P_\pi}(\mu_g)$ are the bivariate distributions of $\mathbf{v}(\mathbf{x}^k, \mathbf{x}^j)$ and $\mathbf{v}(\mathbf{x}^j, \mathbf{x}^k)$ it induces. By Assumption 1, the function t is strictly supermodular and J^{-1} is strictly increasing, so that $\pi(\mathbf{v}(\mathbf{x}^k, \mathbf{x}^j), \mathbf{v}(\mathbf{x}^j, \mathbf{x}^k))$ is supermodular in $\mathbf{v}(\mathbf{x}^k, \mathbf{x}^j)$. Therefore, by standard results (see, for example, Theorem 4.3 in Galichon, 2016), $V_{P_\pi}(\mu_g)$ cannot reach a value higher than that achieved for sortings that satisfy (b), that is, sortings which ensure that v_1^j increases deterministically in x_1^k. Since $v_1(\mu^*(\mathbf{x}))$ strictly increases in x_1, $\mu_1^*(\mathbf{x})$ must also increase in $v_1(\mathbf{x})$; this also 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 $(v_1(\mathbf{x}), x_1)$, which by Assumption 2 is the same as the copula of $(x_1, v_1(\mathbf{x}))$. Thus μ^* is feasible by Sklar's Theorem and (a). It follows that $\mu^* \in \arg\max_{\mu \in S(H)} V_{P\mathbf{x}}(\mu)$. Finally, with strictly increasing H_{x_1} only sorting μ^* satisfies (a) and (b).

Proof of Theorem 3. (7) follows from substituting $u^*(x_1, x_2^*) = 0.5F(x_1, x_1)x_2^*(x_1)$ into (6), integration by parts and some rearranging. (8) follows then from substituting (1) into (7) and rearranging.

Assumption 4 implies Assumption 2. I will show that Assumption 4 $\Rightarrow$ Assumption 2 by constructing copula C_h of x_1, v_1 that satisfies Assumption 2. First, define the functions

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

The candidate copula is then

$$C_h(x_1, v_1) \equiv C^1(x_1, v_1) + (C^2(x_1, v_1) + C^2(v_1, x_1)) + C^3(x_1, v_1). \tag{24}$$

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 (24) and the fact that $\Pr(X_1 \le l, v_1 \le v_1) = 0.5G(v)$.

Proof of Proposition 1. Integrating ΔK by parts and rearranging yields

$$u^*(\underline{x}_1, \bar{x}_2; \theta_2) - u^*(\underline{x}_1, \bar{x}_2; \theta_1) - x_2\Delta K$$
$$= \int_{H_{x_1}(\hat{x})}^1 \frac{\partial}{\partial x_1}\Delta K(H_{x_1}^{-1}(s))(\bar{x}_2 s - 0.5(H_{x_2}^{-1}(s) + \bar{x}_2))\mathrm{d}H_{x_1}^{-1}(s)$$
$$w^*(\underline{x}_1, \bar{x}_2; \theta_2) - w^*(\underline{x}_1, \bar{x}_2; \theta_1) - \Delta K$$
$$= \int_{H_{x_1}(\hat{x})}^1 \frac{\partial}{\partial x_1}\Delta K(H_{x_1}^{-1}(s))(s - 0.5(H_{x_2}^{-1}(s) + 1))\mathrm{d}H_{x_1}^{-1}(s).$$

The results follow then by a simple rearrangement of the conditions $\bar{x}_2 s - 0.5(H_{x_2}^{-1}(s) + \bar{x}_2)$ and $s - 0.5(H_{x_2}^{-1}(s) + 1)$, respectively.

Proof of Proposition 2. Notice that, by (12), (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 (that is, x_1 increases deterministically in x_2) then each worker same-matches and receives the same-match wage. Thus, $\mathrm{Var}(w^*) < \mathrm{Var}(w_S)$ 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 additively separable and (under the assumption that $\underline{x}_2 \le 1$) increases in both variables, and thus $w(\mathbf{x})^2$ is supermodular.

Proof of Proposition 3. (i) If $\bar{y} = 0$ then $w_S = w^*$ and the result is immediate. If $\bar{y} = 1$ and $\bar{x}_2 \le 1$ then all low- (high) skill workers must earn more (less) than $w_S(l)$ ($w_S(h)$) and the result follows as well. For $\bar{y} \in (0,1)$, denote the wage received by a worker of skill x_1 from cross-matching by $w^c(x_1)$. Of course, worker $(x_1, z(x_1, 0.5))$ is indifferent between same- and cross-matching, so that

$w^c(x_1) - 0.5(1 - z(x_1, 0.5))F(h,l) = z(x_1, 0.5)w_S(x_1)$. After some rearranging, one can show that this implies that $(z(h,0.5) + z(l,0.5))\Delta w^c = 2\, z(h,0.5)z(l,0.5)\Delta w_S$, and $F(h,l) - (F(\mathbf{h}) + F(\mathbf{l}))/2 = 0.5\Delta w_S(z(h,0.5) - z(l,0.5))/(z(h,0.5) + z(l,0.5))$, where $\Delta w_S \equiv w_S(h) - w_S(l) = 0.5(F(h) - F(l))$ and $\Delta w^c \equiv w^c(h) - w^c(l)$. Clearly then, $\mathrm{Var}(w_S) = 0.25(\Delta w_S)^2$ and

$$\frac{\mathrm{Var}(w^*)}{\mathrm{Var}(w_S)} = 1 - \bar{y} + \frac{4\bar{y}(z(h,0.5)z(l,0.5))^2 + (1-\bar{y})\bar{y}(z(h,0.5) - z(l,0.5))^2}{(z(h,0.5) + z(l,0.5))^2},$$

Elementary algebra reveals that if $z(h,0.5)z(l,0.5) \le 1$ then $\mathrm{Var}(w^*) - \mathrm{Var}(w_S) \le 0$. Thus, of course, if $\mathrm{Var}(w_B) \ge \mathrm{Var}(w_S)$—as is the case if $a_F \ge 1$—then $\mathrm{Var}(w_B) \ge \mathrm{Var}(w^*)$, as required.

(ii) Consider $G_h^{-1}(y) = 1/(2-y)$ and $G_l^{-1}(y) = a_F/((1-y)a_F + 2)$; clearly then $G_l^{-1}(1)/G_h^{-1}(0) = a_F$ so that by (14) all workers same-match in equilibrium and $w^*(\mathbf{x}) = w_S^*(x_1)$; then the result follows from the fact that $\mathrm{Var}(w_S) > \mathrm{Var}(w_B)$.

Proof that Proposition 3 holds under Assumption 5. (i) It suffices to show that $\mathrm{Var}(w^*) \le \mathrm{Var}(w_S)$, the second part follows immediately.

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

$$\mathrm{BWI}(w,\mu) = \mathrm{Var}(w(\mathbf{x}) + w(\mu(\mathbf{x})),$$

$$\mathrm{WWI}(w,\mu) = 0.25\mathrm{E}\big(w(\mathbf{x}) - w(\mu(\mathbf{x}),\mu_2(\mathbf{x}))\big)^2;$$

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

The first part is easy. It follows directly from (8) that worker $\mathbf{x}$ earns more (less) than their benchmark wage if $\mu_1^*(\mathbf{x}) \ge (\le) x_1$ and $z(x_1, H_{x_1}(s)) \le 1$ 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 $\max\{x_2, \mu_2^*(x_1,x_2)\} \le 1$ and $\mu_1^*(\mathbf{x}) \ge (\le)\mathbf{x}$ then $w(\mu^*(\mathbf{x})) \le w_B(\mu^*(\mathbf{x}))$ and $w(\mathbf{x}) \ge w_B(\mathbf{x})$. The assumption that $\bar{x}_2 \le 1$ ensures that $\max\{x_2, \mu_2^*(x_1,x_2)\} \le 1$, and it thus follows that $\mathrm{WWI}(w_B,\mu^*) \ge \mathrm{WWI}(w^*,\mu^*)$.[20]

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

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

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 $\mathrm{BWI}(w^*,\mu^*) \le \mathrm{BWI}(w_S,\mu^*)$.[21] 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

$$\frac{\partial^2}{\partial\alpha^2}p(z,r;\alpha) = 2(z^2e^{2\alpha z} + r^2e^{2\alpha r} - (r+z)^2e^{\alpha(r+z)}),$$

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

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

Next, note that

$$\begin{aligned}&\frac{\partial}{\partial z}\frac{\partial^2}{\partial\alpha^2}P(z,r;\alpha) \ge 0 \Rightarrow \frac{\partial^2}{\partial z^2}\frac{\partial^2}{\partial\alpha^2}P(z,r;\alpha) > 0,\\
&K(z,r;\alpha) \equiv k(z,r;\alpha) + k(-z,r;\alpha) + k(z,-r;\alpha) + k(-z,-r;\alpha)\\
&\qquad \ge \frac{\partial^2}{\partial\alpha^2}P(z,r;\alpha),\end{aligned}$$

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 $z \ge \max\{0,r\}$. 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 $\frac{\partial^2}{\partial\alpha^2}P(z,r;\alpha) \ge 0$ 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 $\frac{\partial^2}{\partial\alpha^2}P(z,r^*;\alpha), \frac{\partial}{\partial z}\frac{\partial^2}{\partial\alpha^2}P(z,r^*;\alpha) \ge 0$. This implies that $\frac{\partial}{\partial z}\frac{\partial^2}{\partial\alpha^2}P(z',r^*;\alpha) > 0$; contradiction!

(ii) The result follows if, for any $c < 0$ and ω_{11}, there exists a preference structure ρ_{12} and ω_{22} such that $w^*(\mathbf{x}) = w_S(x_1)$. This follows immediately from (18) by setting $\rho_{12} = 1$ and $\omega_{22} > -c\omega_{11}$.

Proof of Proposition 4. Starting with the direct effect, notice that $\frac{\partial}{\partial x_1^j}F(x_1,s) = ((x_1,s)^c)/x_1$. Thus, a decrease in c raises $\frac{\partial}{\partial x_1^j}F(x_1,s)$ as long as $x_1 s < 1$. Jointly, $x_1\mu_1^*(\mathbf{x}) < 1$ and $\ln x_1 < \ln \mu_1^*(\mathbf{x})$ imply that $x_1 s < 1$ for any $s \in (x_1, \mu_1^*(\mathbf{x}))$, and thus the direct effect increases $u^*(\mathbf{x}) - u_S(x_1)$. Moreover, as $z(x_1, H_{x_1}(s))$ is the inverse of μ_1^* with respect to x_2, and μ_1^* increases in x_2, z must increase in s. Thus, if $x_2 < 1$, then $z(x_1, H_{x_1}(s)) < 1$ for $s \in (x_1, \mu_1^*(\mathbf{x}))$, because $z(x_1, H_{x_1}(\mu_1^*(\mathbf{x}))) = x_2$. It follows that the direct effect increases $w^*(\mathbf{x}) - w_S(x_1)$.

To prove that the indirect effect increases $u^*(\mathbf{x}) - u_S(x_1)$ (and, if $x_1 < 1$, also $w^*(\mathbf{x}) - w_S(x_1)$) it suffices to show that a decrease in c lowers $z(x_1, H_{x_1}(s))$ for all $s \in [x_1, \mu_1^*(\mathbf{x})]$, as this implies that it raises $\mu_1^*(\mathbf{x})$. Since

$$\frac{\partial}{\partial c}z(x_1, H_{x_1}(s)) = -\frac{\partial}{\partial c}\mu_1^*(x_1, z(x_1, H_{x_1}(s)))/\frac{\partial}{\partial x_2}\mu_1^*(x_1, z(x_1, H_{x_1}(s))),$$

this will be true as long as $\ln x_1 - \delta_1 \le (\ln \mu_1^*(x_1, z(x_1, H_{x_1}(s))) - \ln x_1)\rho/(1-\rho)$. Importantly, note that because z increases in s and μ_1^* increases in x_2, $\mu_1^*(x_1, z(x_1, H_{x_1}(s)))$ increases in s. If $\rho < 0$ this immediately implies that

$$(\ln \mu_1^*(x_1, z(x_1, H_{x_1}(s))) - \ln x_1)\rho/(1-\rho) \ge (\ln \mu_1^*(\mathbf{x}) - \ln x_1)\rho/(1-\rho) \ge \ln x_1 - \delta_1$$

for any $s \in (x_1, \mu_1^*(\mathbf{x})$, as required. If, instead, $\rho > 0$, then the result follows because $z(x_1, H_{x_1}(x_1)) = x_2^*(x_1)$ and thus $\ln x_1 \le \ln \mu_1^*(\mathbf{x})$ implies that

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

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

$\ln x_1 \leq \ln \mu_1^*(x_1, z(x_1, H_{x_1}(s)))$ for all $s \in (x_1, \mu_1^*(\mathbf{x}))$, which together with $\ln x_1 \leq \delta_1$ ensures that $\ln x_1 - \delta_1 \leq (\ln \mu_1^*(x_1, z(x_1, H_{x_1}(s))) - \ln x_1)\rho/(1-\rho)$ is satisfied.

Proof of Proposition 5. (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

$$w^o_{x_1} = 0.5F(\mathbf{x}_1) + \alpha_{x_1}(s_F - c), \tag{25}$$

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 cut-off level of x_2 such all high (low) skill workers with x_2 greater (lower) that the cut-off 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

$$\begin{aligned} w^n_h &= G_h^{-1}(y^o) w^o_h + 0.5(1 - G_h^{-1}(y^o))F(h,l), \\ w^n_l &= G_l^{-1}(1-y^o) w^o_l + 0.5(1 - G_l^{-1}(1-y^o))F(h,l). \end{aligned} \tag{26}$$

(22) follows then from adding up (26) for the two skill types, using $w^n_h + w^n_l = F(h,l)$, $w^o_h + w^o_l = F(h,l) - c$, (25) and some algebra.[22]

The argument from Section 6.6. Truth-telling about one's 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 worker's true type. Finally, on top of the requirements specified in Definition 1, we impose the *truth-telling condition*:

$$\begin{aligned} u^*(x_1, x_2) &= \max_{\hat{x}_2} w^*(x_1, \hat{x}_2) - 0.5(1 - x_2)F(x_1, \mu_1^*(x_1, \hat{x}_2)) \\ &= \max_{\hat{x}_2} 0.5(1 + x_2)F(x_1, \mu_1^*(x_1, \hat{x}_2)) - w^*(\mu(x_1, \hat{x}_2)), \end{aligned} \tag{27}$$

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 (27). Therefore, the requirement of truth-telling does not change the equilibrium conditions.

Appendix 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:

$$U(w^k, \bar{w}^{k,j}; x_2), \tag{28}$$

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

[22] 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.

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* $\psi : D^2_{\mathbf{x}} \times \mathbb{R} \to \mathbb{R}$, 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 defined, we can now state the assumption that ensures that x_2 captures the inverse of the strength of relative concerns.

Assumption 6. 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

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

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 et al. (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 6 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

$$u(\mathbf{x}^k) = \max_{\mathbf{x}^j} \psi(\mathbf{x}^j, \mathbf{x}^k, u(\mathbf{x}^j)). \tag{29}$$

The binary skills case. In the binary skills case equilibrium sorting can be fully characterised even for general KUJ utility.[23]

Proposition 6. *Suppose that Assumption 4 is satisfied and workers' utility function is given by* (28)*. 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*

[23] If skills are continuously distributed, then one can show that Assumption 6 implies that $\mu_1^*(\mathbf{x})$ is increasing in x_2 by adapting the standard arguments from Sattinger (1979) (proof available on request).

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 5, 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 cut-off value of x_2, such that all high-skill workers with x_2 above the cut-off will same-match, and all those with x_2 below the cut-off will cross-match. Clearly, $b_h(x_2)$ is strictly 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 6.

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 cut-off 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

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

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. □

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 cut-off $\tilde{y}$: In the baseline model, this was a function of a simple sufficient statistic of F, a_F; in the general case, the cut-off still depends on all possible production plans ($F(\mathbf{l}, F(h, l), F(\mathbf{h}))$ but the relationship is much more complicated.

Data availability

No data were used for the research described in the article.

References

Abraham, K.G., Taylor, S.K., 1996. Firms' use of outside contractors: theory and evidence. J. Labor Econ. 14 (3), 394–424.
Aghion, P., Bergeaud, A., Blundell, R.W., Griffith, R., 2019. The innovation premium to soft skills in low-skilled occupations. Available At SSRN 3489777.
Anderson, A., Smith, L., 2024 Mar. The comparative statics of sorting. Am. Econ. Rev. 114 (3), 709–751.
Autor, D.H., Levy, F., Murnane, R.J., 2003 Nov. The skill content of recent technological change: an empirical exploration*. Q. J. Econ. 118 (4), 1279–1333.
Barnett, R.C., Bhattacharya, J., Bunzel, H., 2019. The fight-or-flight response to the joneses and inequality. J. Econ. Dyn. Control 101, 187–210.
Becker, G.S., 1973 Jul-Aug.. A theory of marriage: Part I. J. Polit. Econ. 81 (4), 813–846.
Bergeaud, A., Malgouyres, C., Mazet-Sonilhac, C., Signorelli, S., 2024. Technological change and domestic outsourcing. J. Labor Econ. https://doi.org/10.1086/730166
Boerma, J., Tsyvinski, A., Zimin, A.P., 2021 Sep. Sorting with Team Formation. Working Paper 29290, National Bureau of Economic Research.
Bojilov, R., Galichon, A., 2016. Matching in closed-form: equilibrium, identification, and comparative statics. Econ. Theory 61 (4), 587–609.
Bottan, N.L., Perez-Truglia, R., 2022 Sep. Choosing your pond: location choices and relative income. Rev. Econ. Stat. 104 (5), 1010–1027.
Bound, J., Johnson, G., 1992. Changes in the structure of wages in the 1980's: an evaluation of alternative explanations. Am. Econ. Rev. 82 (3), 371–392.
Buser, T., Niederle, M., Oosterbeek, H., 2014 May. Gender, competitiveness, and career choices*. Q. J. Econ. 129 (3), 1409–1447.
Cabrales, A., Calvó-Armengol, A., 2008. Interdependent preferences and segregating equilibria. J. Econ. Theory 139 (1), 99–113.
Cabrales, A., Calvó-Armengol, A., Pavoni, N., 2008. Social preferences, skill segregation, and wage dynamics. Rev. Econ. Stud. 75 (1), 65–98.
Card, D., Mas, A., Moretti, E., Saez, E., 2012 Oct. Inequality at work: the effect of peer salaries on job satisfaction. Am. Econ. Rev. 102 (6), 2981–3003.
Chade, H., Eeckhout, J., Smith, L., 2017. Sorting through search and matching models in economics. J. Econ. Lit. 55 (2), 493–544.
Clark, A.E., Oswald, A.J., 1998. Comparison-concave utility and following behaviour in social and economic settings. J. Public Econ. 70 (1), 133–155.
Coase, R.H., 1937. The nature of the firm. Economica 4 (16), 386–405.
Costrell, R.M., Loury, G.C., 2004 Dec. Distribution of ability and earnings in a hierarchical job assignment model. J. Polit. Econ. 112 (6), 1322–1363.
Fang, K.W., 1990. Symmetric Multivariate and Related Distributions, 1st ed. Chapman and Hall/CRC.
Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 114 (3), 817–868.
Fershtman, C., Hvide, H.K., Weiss, Y., 2006. Cultural diversity, status concerns and the organization of work. In: The Economics of Immigration and Social Diversity. Emerald group Publishing Limited.
Fershtman, C., Weiss, Y., 1993 Jul. Social status, culture and economic performance. Econ. J. 103 (419), 946–959.
Frank, R., 1984a Sep. Are workers paid their marginal products? Am. Econ. Rev. 74 (4), 549–571.
Frank, R., 1984b. Interdependent preferences and the competitive wage structure. RAND J. Econ. 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. J. Money Credit Bank. 26 (1), 1–8.
Galichon, A., 2016. Optimal Transport Methods in Economics. Princeton University Press.
Ghiglino, C., Goyal, S., 2010. Keeping up with the neighbors: social interaction in a market economy. J. Eur. Econ. Assoc. 8 (1), 90–119.
Gola, P., 2021. Supply and demand in a two-sector matching model. J. Polit. Econ. 129 (3), 940–978.
Gola, P., 2024. On the importance of social status for occupational sorting. Econ. J. 134 (661), 2009–2040.
Goldschmidt, D., Schmieder, J.F., 2017. The rise of domestic outsourcing and the evolution of the German wage structure. Q. J. Econ. 132 (3), 1165–1217.
Gretsky, N.E., Ostroy, J.M., Zame, W.R., 1992 Jan. The nonatomic assignment model. Econ. Theory 2 (1), 103–127.
Grossman, S.J., Hart, O.D., 1986. The costs and benefits of ownership: a theory of vertical and lateral integration. J. Polit. Econ. 94 (4), 691–719.
Hart, O., Moore, J., 1990. Property rights and the nature of the firm. J. Polit. Econ. 98 (6), 1119–1158.
Holmström, B., 1999. Managerial incentive problems: a dynamic perspective. Rev. Econ. Stud. 66 (1), 169–182.
Holmstrom, B., Milgrom, P., 1994. The firm as an incentive system. Am. Econ. Rev. 84 (4), 972–991.
Hopkins, E., 2024 Mar. Cardinal sins? conspicuous consumption, cardinal status and inequality. J. Eur. Econ. Assoc. 22 (5), 2374–2413.
Hopkins, E., Kornienko, T., 2004 Sep. Running to keep in the same place: consumer choice as a game of status. Am. Econ. Rev. 94 (4), 1085–1107.
Hopkins, E., Kornienko, T., 2009 Nov. Status, affluence, and inequality: rank-based comparisons in games of status. Games Econ. Behav. 67 (2), 552–568.
Juhn, C., Murphy, K.M., Pierce, B., 1993. Wage inequality and the rise in returns to skill. J. Polit. Econ. 101 (3), 410–442.
Katz, L., Murphy, K.M., 1992. Changes in relative wages, 1963–1987: supply and demand factors. Q. J. Econ. 107 (1), 35–78.
Kim, C., 1998. Stochastic dominance, pareto optimality, and equilibrium asset pricing. Rev. Econ. Stud. 65 (2), 341–356.
Klein, B., Crawford, R.G., Alchian, A.A., 1978. Vertical integration, appropriable rents, and the competitive contracting process. J. Law Econ. 21 (2), 297–326.
Kőszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. Q. J. Econ. 121 (4), 1133–1165.
Kremer, M., Maskin, E., 1996. Wage inequality and segregation by skill.
Langtry, A., 2023. Keeping up with "the joneses": reference-dependent choice with social comparisons. Am. Econ. J. Microecon. 15 (3), 474–500.
Langtry, A., Ghinglino, C., 2025. Status substitution and conspicuous consumption.

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

Lindenlaub, I., 2017 Jan. Sorting multidimensional types: theory and application. Rev. Econ. Stud. 84 (2), 718–789.

Liu, Q., 2020 Aug. Stability and Bayesian consistency in two-sided markets. Am. Econ. Rev. 110 (8), 2625–2666.

Liu, Q., 2024. Chapter 5 - matching with incomplete information. In: Che, Y.-K., Chiappori, P.-A., Salanié, B. (Eds.), Handbook of the Economics of Matching, vol. 1. North-Holland, pp. 237–274.

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

Louizos, C., Ullrich, K., Welling, M., 2017. Bayesian compression for deep learning. In: Guyon, I., Von Luxburg, U., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., (Eds.), Advances in Neural Information Processing Systems, vol. 30. Curran Associates, Inc..

Luttmer, E.F.P., 2005. Neighbors as negatives: relative earnings and well-being. Q. J. Econ. 120 (3), 963–1002.

Mani, A., Mullin, C.H., 2004 Oct. Choosing the right pond: social approval and occupational choice. J. Labor Econ. 22 (4), 835–862.

Manning, A., 2004. We can work it out: the impact of technological change on the demand for low-skill workers. Scot. J. Polit. Econ. 51 (5), 581–608.

Mazzolari, F., Ragusa, G., 2013. Spillovers from high-skill consumption to low-skill labor markets. Rev. Econ. Stat. 95 (1), 74–86.

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

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

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

Perez-Truglia, R., 2020. The effects of income transparency on well-being: evidence from a natural experiment. Am. Econ. Rev. 110 (4), 1019–1054.

Sattinger, M., 1979 Mar. Differential rents and the distribution of earnings. Oxf. Econ. Pap. 31 (1), 60–71.

Sklar, M., 1959. Fonctions de répartition à n dimensions et leurs marges. In: Annales de l'ISUP, vol. 8. pp. 229–231.

Tervio, M., 2008 Jun. The difference that CEOs make: an assignment model approach. Am. Econ. Rev. 98 (3), 642–668.

Tinbergen, J., 1956. On the theory of income distribution. Weltwirtschaftl. Arch. 77, 155–175.

Williamson, O.E., 1971. The vertical integration of production: market failure considerations. Am. Econ. Rev. 61 (2), 112–123.