The Optimal Assortativity of Teams Inside the Firm*

Ashwin Kambhampati†

Carlos Segura-Rodriguez‡

November 3, 2021

Abstract

How does a profit-maximizing manager form teams and compensate workers when workers have private information about their productivity and exert hidden effort once in a team? We study a team-production model in which positive assortative matching is both efficient and profit-maximizing under pure adverse selection and pure moral hazard. We show that the interaction of adverse selection and moral hazard can lead to non-assortative matching if complementarities are sufficiently weak. When this is the case, the manager may prefer to delegate matching, allowing workers to sort themselves into teams.

Keywords: Assortative Matching, Adverse Selection, Moral Hazard, Delegation, Teams

JEL Classification: C78, D86, L23

*We are grateful to the Editor, David Martimort, and two anonymous referees for constructive feedback. We thank Aislinn Bohren, Ralph Boleslavsky, Kevin He, Brian Hill, Rohit Lamba, Annie Liang, George Lukyanov, Paula Onuchic, Andrew Postlewaite, Steven Matthews, Lones Smith, Juuso Toikka, Rakesh Vohra, and, especially, George Mailath for helpful comments and suggestions. This article incorporates material from the second chapter of the first author's dissertation at the University of Pennsylvania, "Optimal Team Formation Under Asymmetric Information". This chapter is also reproduced as the Penn Institute for Economic Research working paper 20-018.

†United States Naval Academy, kambhamp@usna.edu.

‡Central Bank of Costa Rica, segurarc@bccr.fi.cr.

1 Introduction

Teamwork has increasingly become "a way of life" in many firms (Lazear and Shaw, 2007). For instance, at Google, "the team is the molecular unit where real production happens, where innovative ideas are conceived and tested, and where employees experience most of their work."[1] Yet, forming teams composed of complementary and productive workers is complicated. First, workers often possess private information about their characteristics, such as their ability or willingness to work collaboratively. Second, individual effort is difficult to identify from team output. A profit-maximizing manager must therefore design contracts that simultaneously screen for unobservable characteristics and provide incentives for effort in teams.

What is the optimal assignment of workers to teams? How should a manager remunerate her workers? With few exceptions, the economic theory of teams has answered each question separately. To wit, a large literature in matching, pioneered by Becker (1973), studies the optimal composition of teams, abstracting from incentives. At the same time, a large literature in contract theory, pioneered by Holmström (1982), studies the provision of incentives within a single team, fixing its composition.[2]

By conducting a unified analysis of optimal team composition and incentives, we uncover a novel economic distortion: Even when matching likes with likes — positive assortative matching (PAM) — is productively efficient, creating the right incentives for workers to truthfully reveal their characteristics and exert effort can make implementing PAM prohibitively costly, leading a profit-maximizing manager to match non-assortatively. We identify when and why productive distortions occur, and when a profit-maximizing manager would prefer to allow workers to sort themselves in order

[1] https://rework.withgoogle.com/guides/understanding-team-effectiveness/steps/introduction/

[2] Less related is the work of Marschak and Radner (1972), which investigates the behavior of a fixed team of agents whom share a common prior and objective function, but possess different information when taking actions. Within this framework, Prat (2002) studies the optimal composition of teams.

to save on incentive costs. Together, our results offer a new rationale for non-assortative matching and delegation inside of firms, complementing recent empirical work documenting these patterns (see Adhvaryu, Bassi, Nyshadham, and Tamayo (2021) and Kambhampati, Segura-Rodriguez, and Shao (2021) and the references therein).

We posit a simple model to illustrate the key mechanism. A single risk-neutral manager assigns risk-neutral workers, each protected by limited liability, to teams of two. Each team completes a project, which is either a success or a failure. The success of a project depends on the private type of each worker in the team, high or low, and the costly, hidden effort they exert. In particular, the probability of success is multiplicative in the team's total effort and a strictly increasing, strictly supermodular function of the types of the workers.

The production technology ensures that PAM is efficient (Theorem 1). Moreover, PAM is profit-maximizing under pure adverse selection (see Section 3) and pure moral hazard (Theorem 2). Our main analysis thus concerns the problem of implementing PAM in the mixed model with both adverse selection and moral hazard, and on the optimality of that implementation. Our key finding is that if a team composed of lows is sufficiently less productive than one composed of highs (*types are sufficiently important to production*) and the productivity gain a high yields to a team is almost independent of the the type of his teammate (*complementarities are weak*), then implementing PAM can be prohibitively costly.

The result is a consequence of two different incentive problems. First, in order to incentivize a low to exert a given amount of effort, a manager must pay him a higher bonus than a high for success. Providing effort incentives for lows thereby increases the payoff to highs from misreporting their type. Hence, the downward incentive compatibility constraint binds.

Second, when the manager implements PAM, lows match with lower quality teammates than highs. Hence, if types are sufficiently important to production and complementarities are sufficiently weak, the higher bonus a low receives is not enough to offset the gain he would get from misreporting his type and matching with a more productive teammate. It follows that the upward incentive compatibility constraint also binds (Theorem 3).

To separate highs and lows, a profit-maximizing manager not only pays lows strictly positive base wages, but also offers distortionary bonuses. In fact, when the fraction of highs in the firm is sufficiently small, effort is distorted downwards in all teams. Our main result, Theorem 4, shows that when this is the case, any efficiency gains of implementing PAM are outweighed by its incentive costs, implying that matching distortions are profit maximizing. In fact, randomly assigning workers to teams and paying each a bonus noncontingent upon their reported type yields strictly higher profits.

To conclude our analysis, we consider the implications of our matching distortion result for the optimal internal organization of the firm. We ask the following question: If a manager could commit to *not* asking workers to report their types, so that she would instead have to delegate the problem of sorting to her workers, would she do so? On one hand, such an arrangement entails a *loss of control*: She can no longer tailor wages to reported characteristics. But on the other, the manager can exploit her workers' *local information* about each other's characteristics. We formalize this tradeoff by considering an environment in which there is no reporting stage, but in which workers are endowed with knowledge of one another's types. Using this knowledge, workers then form self-enforcing teams. Theorem 5, our final result, identifies conditions under which a profit-maximizing manager prefers to delegate sorting rather than implement an inefficient matching under centralized assignment.

Literature

We summarize the closest related theoretical literature. Franco, Mitchell, and Vereshchagina (2011) and Kaya and Vereshchagina (2014) consider settings in which a profit-maximizing manager assigns workers to teams subject to moral hazard. Our model enriches these frameworks to include adverse selection. Damiano and Li (2007) and Johnson (2013) find conditions for PAM to be profit-maximizing in environments in which individuals have private information, but match payoffs are specified exogenously.[3] Our model enriches these frameworks to include moral hazard within teams. To distinguish our contribution from these articles, we assume a production specification ensuring that neither moral hazard nor adverse selection alone generates a distortion of PAM.[4] Indeed, all of our results are driven by the *interaction* between the two.

More broadly, building on Becker (1973)'s marriage model, a number of articles have investigated the role of imperfectly transferable utility and costly search in distorting the stability of PAM. For instance, Legros and Newman (2007) find conditions for PAM in general environments with imperfectly transferable utility; Serfes (2005) and Serfes (2007) find conditions for PAM when principals match agents (see also Wright (2004)); Shimer and Smith (2000) and Smith (2006) find conditions for PAM when individuals engage in random search; and Eeckhout and Kircher (2010) find conditions for PAM when individuals engage in directed search. Though we share a similar motivation as these articles, our analysis is quite different due to our focus on profit-maximization under asymmetric information rather than stability under complete information.

Our modeling of the moral hazard in teams problem follows the literature in which limited liability constraints are the source of contracting friction (in addition to Franco, Mitchell, and Vereshchagina (2011) and Kaya and Vereshchagina (2014), see, among many other articles, Sappington (1983), Sappington (1984), Innes (1990), and Che and Yoo (2001)). This ensures that, despite our assumption of risk neutrality, "efficiency wages" must be paid to incentivize effort. Outside of this literature, McAfee and McMillan (1991) consider the interaction of adverse selection and moral hazard within a fixed

[3]The problem of a profit-maximizing platform also bears resemblance to the one faced by the manager in our study. See, for instance, Gomes and Pavan (2016) and the references therein. See also Utgoff (2020), who studies ex post Nash implementation of PAM under incomplete information.

[4]A detailed comparison to Franco, Mitchell, and Vereshchagina (2011) is provided in Section 4.

team in the absence of limited liability constraints. They establish conditions under which incentives are linear in team output, even when individual performance measures are available.

Finally, our modeling of the tradeoff between centralized assignment and delegated matching is inspired by the literature on delegation, i.e. Aghion and Tirole (1997) and Dessein (2002), wherein the benefit of delegation is that workers can utilize superior information to make decisions and the cost is a loss of control. In contrast to these theories, however, the loss of control issue in our model is not related to the misalignment of incentives between workers and the manager. Indeed, under strictly increasing differences, endogenous sorting leads to PAM, the productively-efficient matching.[5] Instead, the loss of control is related to the manager's decision to commit to *not* eliciting reports about workers' types; as wages cannot depend on reports, the manager's ability to extract rents is limited.

2 Model

We describe the environment, timing, and contracts in the case in which matching is centralized, leaving the description of the delegation alternative to Section 6.

Environment

A risk-neutral, profit-maximizing manager (residual claimant) matches a continuum of risk-neutral workers of mass two into teams of two.[6] Each worker i has an outside option of zero and a type θ_i, either high ($h \in (0,1)$) or low ($\ell \in (0,h)$). In what follows, it will be convenient to partition workers into two unit mass continua each with the same proportion of highs and lows. Let the measure of each type in each partition element be

[5]Kambhampati, Segura-Rodriguez, and Shao (2021) studies delegated matching in an informational environment in which endogenous sorting is not productively efficient.

[6]We describe the manager using female pronouns and the workers using male pronouns.

given by $\nu \in \Delta(\Theta)$, where $\Theta := \{h, \ell\}$ and $\nu(h) \in (0,1)$. Finally, within a team, each worker i exerts costly effort $e_i \in [0, \frac{1}{2}]$.

Each team completes a project, which is either a success (valued at 1 to the manager) or a failure (valued at 0 to the manager). Success of a project depends on the types of the workers completing it and the costly effort they exert. In particular, a project succeeds with probability

$$(e_i + e_j) \cdot q(\theta_i, \theta_j),$$

where $j \neq i$ is i's teammate and $q : (0,1)^2 \to (0,1)$ is a symmetric, twice-differentiable, strictly increasing $\left(\frac{\partial q(\theta_i,\theta_j)}{\partial \theta_i} > 0 \text{ and } \frac{\partial q(\theta_i,\theta_j)}{\partial \theta_j} > 0\right)$, and strictly supermodular $\left(\frac{\partial^2 q(\theta_i,\theta_j)}{\partial \theta_i \partial \theta_j} > 0\right)$ function determining the team's quality.[7] Although we assume that effort is perfectly substitutable for simplicity of exposition, all of our results extend to the case of a general constant elasticity of substitution production function provided that workers can coordinate on their preferred Nash equilibrium. We elaborate on this point in Section 7.

The cost of effort is given by a second-degree polynomial function $c : [0, \frac{1}{2}] \to \mathbb{R}_+$ with

$$c(e) := \alpha \cdot e^2 + \beta \cdot e, \quad \alpha + \beta > q(h,h) \text{ and } \beta \in (0, \frac{q(\ell,\ell)}{4}).$$

Notice that c is strictly increasing and strictly convex on $(0,1)$ with $c(0) = 0$, $c'(0) < \frac{q(\ell,\ell)}{4}$, and $c'(\frac{1}{2}) > q(h,h)$.[8] We use only these general properties in much of our analysis and make note of our use of additional properties of this functional form.

Timing and Contracts

The timing of events is as follows:[9]

[7] Notice, these derivatives are well-defined as q is not only defined for $h, \ell \in (0,1)$, but for any other type in the open unit interval.

[8] The last two conditions generalize the usual Inada conditions $\lim_{e \to 0} c'(0) = 0$ and $\lim_{e \to \frac{1}{2}} c'(e) = \infty$.

[9] We need not specify behavior by the manager or the workers in the cases in which some workers reject the contract. The manager could always replicate the performance of such a contract by paying the rejecting workers an unconditional wage of zero and having them participate.

1. The manager proposes a contract.
2. Each worker accepts or rejects the proposed contract.
3. Each worker reports his type to the manager.
4. The manager assigns workers to teams.
5. Each worker learns the type of his assigned teammate.
6. Workers exert effort.
7. The manager observes output and compensates each worker.

A **contract** is a matching and a wage scheme. A **matching** is a coupling $\mu \in \Delta(\Theta \times \Theta)$ with marginal measures $\nu \in \Delta(\Theta)$:

$$\sum_{\theta'} \mu(\theta, \theta') = \nu(\theta) \quad \text{for each } \theta \in \Theta.$$

Let $\mathcal{M}$ denote the set of all matchings. Given $\mu \in \mathcal{M}$, the probability with which a type $\theta \in \Theta$ matches a high is denoted by

$$p_\theta^\mu := \frac{\mu(\theta, h)}{\nu(h)}.$$

We will be interested in distortions of **positive assortative matching (PAM)**, i.e. matchings μ for which $p_h^\mu = 1$.

A **wage scheme** is a tuple

$$\mathbf{w} := (w_h, w_\ell, b_{hh}, b_{h\ell}, b_{\ell h}, b_{\ell\ell}) \in \mathbb{R}_+^6,$$

where $w_{\hat{\theta}}$ is the base wage paid to a worker who announces his type as $\hat{\theta}$ and $b_{\hat{\theta}\hat{\theta}'}$ is the bonus he receives if he matches a worker with reported type $\hat{\theta}'$ and his team succeeds. Notice that base wages and bonuses are non-negative, reflecting limited liability constraints. These constraints imply that all agents have a (weak) incentive to accept any contract offered by the manager. Participation constraints are thus ignored until our analysis of the model without limited liability, discussed in Section 7.

An additional comment is in order. The contracts we consider are without loss of generality among those that elicit information about worker types only in step 3.[10] But, in step 5, due to their close interaction, each worker learns the type of his assigned teammate. Yet, there is no subsequent reporting stage. Contracts are therefore incomplete; mechanisms in which teammates report each other's type are ruled out by assumption.[11] We believe this is a plausible assumption in environments in which the manager is worried about collusion among her workers or if reporting workers fear retaliation by their co-workers for making undesirable reports.[12] Moreover, maintaining this assumption throughout the article allows us to make a meaningful comparison between centralized matching and delegated matching. Indeed, any comparison between the two in which delegated matching strictly outperforms centralized matching *requires* the introduction of contractual incompleteness. Otherwise, by the Revelation Principle, delegated matching would appear as a mechanism in the present environment, so that delegation could never result in strictly higher profits.[13] Notwithstanding this point, in Section 7, we argue that our matching distortion results hold under an alternative timing in which workers do *not* learn about one another's types and contracts are complete.

3 Efficient Allocations

We now describe efficient (surplus-maximizing) effort levels and matchings. We then point out that profit maximization implies efficiency if effort is observable.

[10]See Lemma 1 of Kambhampati and Segura-Rodriguez (2021) for a formal proof in a similar model.

[11]As is well-known, under weak Nash implementation, the manager could obtain full-information profits by punishing workers with zero wages when their reports about the profile of types in their team disagree.

[12]In Section 7, we suggest an approach to endogenize such concerns.

[13]See Poitevin (2000) and Mookherjee (2006) for illuminating discussions.

Efficient Effort

In a team of workers with types (θ_i, θ_j), efficient effort levels solve the following maximization problem:

$$S^{FB}(\theta_i, \theta_j) := \max_{e_i, e_j \in [0, \frac{1}{2}]} \quad (e_i + e_j) \cdot q(\theta_i, \theta_j) - c(e_i) - c(e_j).$$

Because $c'(0) < q(\ell, \ell)$ and $c'(\frac{1}{2}) > q(h, h)$, the first-order conditions are both necessary and sufficient for optimality:

$$[e_i] \quad q(\theta_i, \theta_j) = c'(e_i),$$

i.e. effort must be such that the marginal return to effort equals its marginal cost. Alternatively,

$$e^{FB}(\theta_i, \theta_j) = \phi(q(\theta_i, \theta_j)) > 0,$$

where

$$\phi(e) := (c')^{-1}(e).$$

Notice, ϕ is differentiable and strictly increasing because c'' exists and is strictly positive.

Efficient Matching

Efficient matching solves the Monge-Kantorovich optimal transportation problem with expected (efficient) team surplus in the objective function:[14]

$$\max_{\mu \in \mathcal{M}} \quad \sum_{\theta_i} \sum_{\theta_j} \mu(\theta_i, \theta_j) \cdot S^{FB}(\theta_i, \theta_j).$$

[14]For background on the Monge-Kantorovich problem, see, for instance, Chapter 2 of Galichon (2016).

We observe that PAM is the unique[15] efficient matching by showing that $S^{FB}(\theta_i, \theta_j)$ is strictly supermodular.

Theorem 1 (Efficient Allocations)**.** *The efficient effort level of each worker in a team with types* (θ_i, θ_j) *is*

$$e^{FB}(\theta_i, \theta_j) = \phi(q(\theta_i, \theta_j))$$

and PAM is the unique efficient matching.

Proof. By the Envelope Theorem,

$$\frac{\partial S^{FB}(\theta_i, \theta_j)}{\partial \theta_i} = 2 \cdot \phi(q(\theta_i, \theta_j)) \cdot \frac{\partial q(\theta_i, \theta_j)}{\partial \theta_i}.$$

Differentiating with respect to θ_j, we see that

$$\frac{\partial^2 S^{FB}(\theta_i, \theta_j)}{\partial \theta_i \partial \theta_j} = 2 \cdot \phi(q(\theta_i, \theta_j)) \cdot \frac{\partial q(\theta_i, \theta_j)}{\partial \theta_i \partial \theta_j} + 2 \cdot \phi'(q(\theta_i, \theta_j)) \cdot \frac{\partial q(\theta_i, \theta_j)}{\partial \theta_j} \frac{\partial q(\theta_i, \theta_j)}{\partial \theta_i}.$$

This expression is strictly positive as q is strictly increasing in both arguments and strictly supermodular, and $\phi'(q(\theta_i, \theta_j)) > 0$. Hence, $S^{FB}(\theta_i, \theta_j)$ is strictly supermodular. By the analysis of Becker (1973), strict supermodularity of $S^{FB}(\theta_i, \theta_j)$ implies that positive assortative matching is the unique efficient matching. □

Profit Maximization and Efficiency

We now point out that a profit-maximizing manager implements efficient allocations if effort is observable.

Full Information (Effort Observable and Type Observable). First, suppose both effort and types are observable, so that contracts can condition on either. Consider a contract

[15] Here and throughout, uniqueness refers only to the value of p_h^{μ}. As there are a continuum of workers, matchings μ and μ' for which $p_h^{\mu} = p_h^{\mu'}$ may differ on negligible sets.

consisting of positive assortative matching (given observed types) and a wage scheme that pays each agent his cost of effort, $c(e^{FB}(\theta_i,\theta_j))$, if he exerts the manager's desired effort level, $e^{FB}(\theta_i,\theta_j)$, and zero otherwise. As this contract maximizes total surplus and yields no rent to the workers, no other contract can do better.

Pure Adverse Selection (Effort Observable and Type Unobservable). Second, suppose effort is observable, but type is unobservable, so that contracts can only condition on effort and reported type. Consider a contract consisting of positive assortative matching (given *reported* types) and a wage scheme that pays each worker i his cost of effort, $c(e^{FB}(\hat{\theta}_i,\hat{\theta}_j))$, if he exerts the manager's desired effort level, $e^{FB}(\hat{\theta}_i,\hat{\theta}_j)$, and zero otherwise, where $\hat{\theta}_i$ is worker i's reported type and $\hat{\theta}_j$ is his teammate's reported type. This scheme is incentive compatible: As the cost of effort does not depend on type, each worker obtains zero utility in any team to which he is assigned, whether or not he reports his type truthfully. As this contract yields the same profits as the full-information optimal contract, we again see that profit maximization implies efficiency.

4 Second-Best Allocations: Pure Moral Hazard

We now consider the case in which effort is unobservable, but type is observable, so that contracts can condition on observed type, but not on effort. We show that the interaction between limited liability and obedience constraints distorts effort levels away from efficiency, as is the case in the many other contacting models. Our contribution in this section is to show that PAM is still the profit-maximizing matching, in spite of these distortions.

Obedience Constraints

Consider the effort decision of worker i in a team with types (θ_i,θ_j). If worker i's bonus in this team is $b_i \geq 0$, then his optimal effort choice, $e(\theta_i,\theta_j,b_i)$, solves the following maximization problem:

$$\max_{e\in[0,\frac{1}{2}]} \quad e\cdot q(\theta_i,\theta_j)\cdot b_i - c(e).$$

His best response is thus given by the function

$$e(\theta_i,\theta_j,b_i) := \begin{cases} 0 & \text{if } c'(0) > q(\theta_i,\theta_j)\cdot b_i \\ \phi(q(\theta_i,\theta_j)\cdot b_i) & \text{if } c'(0) \leq q(\theta_i,\theta_j)\cdot b_i \leq c'(\frac{1}{2}) \\ \frac{1}{2} & \text{if } q(\theta_i,\theta_j)\cdot b_i > c'(\frac{1}{2}). \end{cases} \tag{1}$$

Notice, effort provision is efficient, i.e. $e(\theta_i,\theta_j,b_i) = e^{FB}(\theta_i,\theta_j)$, if and only if a worker receives the entire value of the output his team produces, i.e. $b_i = 1$. As a manager obtains strictly negative profits from any such bonus scheme, this immediately implies that effort will be underprovided if she is profit-maximizing. We show this formally next and discuss the relationship between team quality and second-best optimal effort.

Second-Best Effort

To identify second-best effort levels, we first observe that any optimal wage scheme cannot pay any worker strictly positive base wages. Such schemes can be strictly improved upon by reducing base wages to zero and keeping bonuses constant. Hence, the manager need only select profit-maximizing bonuses in teams with observable types (θ_i,θ_j).

Let the surplus generated in a team with types (θ_i,θ_j) and bonuses (b_i,b_j) be denoted by

$$S(\theta_i,\theta_j,b_i,b_j) := (e(\theta_i,\theta_j,b_i) + e(\theta_j,\theta_i,b_j))\cdot q(\theta_i,\theta_j) - c(e(\theta_i,\theta_j,b_i)) - c(e(\theta_j,\theta_i,b_j)).$$

Let worker i's payoff in this team be denoted by

$$V(\theta_i,\theta_j,b_i,b_j) := (e(\theta_i,\theta_j,b_i) + e(\theta_j,\theta_i,b_j))\cdot q(\theta_i,\theta_j)\cdot b_i - c(e(\theta_j,\theta_i,b_i)).$$

Finally, let the manager's expected profits be given by

$$\Pi(\theta_i, \theta_j, b_i, b_j) := S(\theta_i, \theta_j, b_i, b_j) - V(\theta_i, \theta_j, b_i, b_j) - V(\theta_j, \theta_i, b_j, b_i).$$

As types are observable, optimal bonuses, $(b^{MH}_{\theta_i\theta_j}, b^{MH}_{\theta_j\theta_i})$, solve the maximization problem

$$\max_{(b,b')\in\mathbb{R}^2_+} \quad \Pi(\theta_i, \theta_j, b, b').$$

Manipulating the first-order conditions reveals that interior second-best effort solves the equation

$$e(\theta_i, \theta_j, b^{MH}_{\theta_i\theta_j}) = \phi\left(q(\theta_i, \theta_j)\left[1 + D^{MH}_{\theta_i\theta_j}\right]\right), \tag{2}$$

where

$$D^{MH}_{\theta_i\theta_j} := -\left[\left(\frac{\partial e(\theta_i, \theta_j, b^{MH}_{\theta_i\theta_j})}{\partial b_i}\right)^{-1} \cdot \left(e(\theta_i, \theta_j, b^{MH}_{\theta_i\theta_j}) + e(\theta_j, \theta_i, b^{MH}_{\theta_j\theta_i})\right) + b^{MH}_{\theta_j\theta_i}\right].$$

As $D^{MH}_{\theta_i\theta_j}$ is negative at any solution, second-best optimal effort by any worker in any team (interior or otherwise) is distorted downwards. Using $e(\theta_i, \theta_j, b_{\theta_i\theta_j}) = \phi(q(\theta_i, \theta_j) \cdot b_{\theta_i\theta_j})$, we can also characterize interior second-best optimal bonuses:

$$b^{MH}_{\theta_i\theta_j} = \frac{1}{2} - \frac{\phi(q(\theta_i, \theta_j) \cdot b^{MH}_{\theta_i\theta_j})}{\phi'(q(\theta_i, \theta_j) \cdot b^{MH}_{\theta_i\theta_j})} \cdot \frac{1}{q(\theta_i, \theta_j)}. \tag{3}$$

For general strictly increasing and strictly convex cost functions, the relationship between team quality and bonuses is ambiguous. However, as shown in Theorem 2, the second-degree polynomial function we use ensures that there is a decreasing relationship between the two— higher types receive lower bonus payments. This property pushes second-best optimal matching in the direction of PAM, as we next show in the proof of Theorem 2 and as previously observed by Franco, Mitchell, and Vereshchagina (2011).

Second-Best Matching

Given second-best optimal profit from each team, optimal matching solves the Monge-Kantorovich transportation problem with second-best team profit in the objective function:

$$\max_{\mu \in \mathcal{M}} \quad \sum_{\theta_i} \sum_{\theta_j} \mu(\theta_i, \theta_j) \cdot \Pi(\theta_i, \theta_j, b^{MH}_{\theta_i\theta_j}, b^{MH}_{\theta_j\theta_i}).$$

We observe that PAM is the unique second-best optimal matching by showing that $\Pi(\theta_i, \theta_j, b^{MH}_{\theta_i\theta_j}, b^{MH}_{\theta_j\theta_i})$ is strictly supermodular.

Theorem 2 (Pure Moral Hazard: Optimal Allocations)**.** *If types are observable, but effort is not, then the following properties hold:*

1. *optimal bonuses are strictly decreasing in team quality,*

$$b^{MH}_{\theta_i\theta_j} = \frac{1}{4} + \frac{\beta}{2q(\theta_i, \theta_j)};$$

2. *optimal effort is distorted downwards in all teams,*

$$e(\theta_i, \theta_j, b^{MH}_{\theta_i\theta_j}) = \frac{1}{2}\left[\frac{q(\theta_i, \theta_j)}{4\alpha} - \frac{\beta}{2\alpha}\right] < \left[\frac{q(\theta_i, \theta_j)}{4\alpha} - \frac{\beta}{2\alpha}\right] = e^{FB}(\theta_i, \theta_j); \quad \textit{and,}$$

3. *PAM is the unique optimal matching.*

Proof. The equations for optimal effort levels and bonuses follow from (2), (3), and

$$\phi(e) = \frac{e}{2\alpha} - \frac{\beta}{2\alpha}.$$

We prove that $\Pi(\theta_i, \theta_j, b^{MH}_{\theta_i\theta_j}, b^{MH}_{\theta_j\theta_i})$ is supermodular by exploiting the following properties:

1. $\frac{\partial q(\theta_i, \theta_j) \cdot b^{MH}_{\theta_i\theta_j}}{\partial \theta_i} > 0$ (Higher Types Exert Higher Effort);

2. $\frac{\partial b^{MH}_{\theta_i\theta_j}}{\partial \theta_i} < 0$ (Higher Types Receive Lower Bonuses); and,

3. $c''' \leq 0$ (Marginal Costs are Concave).

To reduce notation, let $q := q(\theta_i, \theta_j)$ and $b := b^{MH}_{\theta_i\theta_j} = b^{MH}_{\theta_j\theta_i}$ in the rest of the proof. By the Envelope Theorem, we have

$$\frac{\partial \Pi(\theta_i, \theta_j, b, b)}{\partial \theta_i} = 2\left[\phi(q \cdot b) \cdot \frac{\partial q}{\partial \theta_i} + \phi'(q \cdot b) \cdot \frac{\partial q}{\partial \theta_i} \cdot b \cdot q\right][1 - 2b].$$

If $\frac{\partial b}{\partial \theta_j} < 0$, then the derivative of the second bracketed expression with respect to θ_j is strictly positive. If $c''' \leq 0$ (so that $\phi''(q \cdot b) \geq 0$) and $\frac{\partial qb}{\partial \theta_i} > 0$, then the derivative of the first bracketed expression with respect to θ_j is strictly positive, too:

$$\frac{\partial}{\partial \theta_j}\left[\phi(q \cdot b) \cdot \frac{\partial q}{\partial \theta_i} + \phi'(q \cdot b) \cdot \frac{\partial q}{\partial \theta_i} \cdot b \cdot q\right] =$$

$$\left(2 \cdot \phi'(q \cdot b) + \phi''(q \cdot b) \cdot qb\right) \frac{\partial q}{\partial \theta_i} \cdot \frac{\partial qb}{\partial \theta_j} + \phi(q \cdot b) \cdot \frac{\partial q}{\partial \theta_i \partial \theta_j} + \phi'(q \cdot b) \cdot \frac{\partial q}{\partial \theta_i \partial \theta_j} \cdot qb > 0.$$

By the product rule and strict positivity of the bracketed expressions, it follows that

$$\frac{\partial^2 \Pi(\theta_i, \theta_j, b, b)}{\partial \theta_i \partial \theta_j} > 0.$$

□

A comment on the literature is in order. Although the third property of the cost function we exploit in our proof that PAM is second-best optimal concerns a primitive of the model—concavity of marginal costs— the first two conditions concern endogenous variables— bonus payments that depend on the cost of effort through (3). Nevertheless, writing the proof in this way facilitates comparison to Franco, Mitchell, and Vereshchagina (2011), who conduct an in-depth analysis of the relationship between pure moral hazard and assortative matching.

Franco, Mitchell, and Vereshchagina (2011) consider a model in which effort is perfectly substitutable, as in our setup, but in which types do not affect the probability with which a team succeeds (instead, types affect the cost of effort). In this setting, they show that if types exerting higher effort receive lower bonuses, then PAM is second-best optimal. The intuition is that expected payments to workers depend on "earned income" attributable to their own effort ($e_i \cdot b_i$) and on "accidental income" attributable to their teammate's effort ($e_j \cdot b_i$). If types exerting higher effort and receiving lower bonuses are matched, then accidental income is minimized. Moreover, when types do not directly affect production, all matchings result in identical expected output. These two observations together imply that PAM is second-best optimal (Proposition 1 of their article).[16]

In our setting, matching structure has non-trivial effects on expected output. In particular, it may be the case that second-best expected output is submodular, even when types exerting higher effort receive lower bonuses. Concave marginal costs ensure this is not the case, making our conditions jointly sufficient for PAM to be second-best optimal.[17] Hence, all distortions of PAM we identify arise solely through the interaction of moral hazard and adverse selection.

5 Second-Best Allocations: The Mixed Model

We now analyze second-best allocations in the mixed model in which both type and effort are unobservable.

[16]Franco, Mitchell, and Vereshchagina (2011) also consider a model in which matching affects expected output, but the assumptions in that model are not satisfied by our production function.

[17]Intuitively, bounding the curvature of the cost function ensures that it is not too costly to make workers in higher quality teams exert more effort.

Incentive Compatibility Constraints

In addition to the obedience constraints of the pure moral hazard model, the manager must now ensure that each worker has an incentive to truthfully report his type. Given a matching μ and wage scheme $\mathbf{w}$, a worker's payoff from reporting his type as $\hat{\theta}$ when his true type is θ (and all other workers truthfully report their type) is

$$U(\hat{\theta}, \theta) := p^{\mu}_{\hat{\theta}} \cdot V(\theta, h, b_{\hat{\theta}h}, b_{h\hat{\theta}}) + (1 - p^{\mu}_{\hat{\theta}}) \cdot V(\theta, \ell, b_{\hat{\theta}\ell}, b_{\ell\hat{\theta}}) + w_{\hat{\theta}}.$$

Hence, the **incentive compatibility constraint** for a worker of type θ, IC_θ, is

$$U(\theta, \theta) \geq U(\hat{\theta}, \theta) \quad \text{where } \hat{\theta} \neq \theta.$$

The **manager's problem** is to choose a contract $(\mu, \mathbf{w})$ to maximize

$$\underbrace{\left[\sum_{\theta_i}\sum_{\theta_j} \mu(\theta_i, \theta_j) \cdot \Pi(\theta_i, \theta_j, b_i, b_j)\right]}_{\text{Sum of Expected Profit From Each Team}} - \underbrace{[2 \cdot \nu(h) \cdot w_h + 2 \cdot \nu(\ell) \cdot w_\ell]}_{\text{Expected Base Wage Payments}}$$

subject to IC_h and IC_ℓ.

Binding Constraints and Effort Distortions

Fixing a matching exhibiting PAM, we now identify when each incentive constraint binds and characterize second-best optimal wage schemes. Given this wage scheme, we identify the sign of effort distortions using the first-order conditions with respect to bonuses:

$$e(\theta_i, \theta_j, b_{\theta_i\theta_j}) = \phi\left(q(\theta_i, \theta_j)\left[1 + D^{MH}_{\theta_i\theta_j} + D^{AS}_{\theta_i\theta_j}\right]\right),$$

where the distortion due to adverse selection, $D^{AS}_{\theta_i\theta_j}$, depends on the quality function q and the optimized Lagrange multipliers associated with IC_h and IC_ℓ.

In the analysis that follows, we fix $0 < q(\ell,\ell) < q(h,h) < 1$ and consider varying $q(h,\ell) = q(\ell,h) \in \left(q(\ell,\ell), \frac{1}{2}(q(h,h)+q(\ell,\ell))\right)$ so that strict increasingness and supermodularity of the quality function q are preserved. When $q(h,\ell)$ is close to $q(\ell,\ell)$, complementarities are *strong*; the gain a high gives to another high is much larger than the gain a high gives to a low. On the other hand, when $q(h,\ell)$ is close to $\frac{1}{2}(q(h,h)+q(\ell,\ell))$, complementarities are *weak*; the gain a high gives to another high is almost the same as the gain a high gives to a low. We link complementarities in the firm to properties of the optimal wage scheme implementing PAM in the following Theorem.

Theorem 3 (Mixed Model: Optimal Wage Scheme Implementing PAM)**.**

1. *If $q(h,\ell)$ is sufficiently close to $q(\ell,\ell)$, then at the unique optimal wage scheme implementing PAM, the following properties hold:*
   (a) *Neither IC_ℓ nor IC_h bind.*
   (b) $D^{AS}_{\ell\ell} = D^{AS}_{hh} = 0$.
   (c) $w_\ell = w_h = 0$.
2. *There exists an $M > 0$ such that if $\frac{q(\ell,\ell)}{q(h,h)} \leq M$, $q(h,\ell)$ is sufficiently close to $\frac{1}{2}\left(q(h,h)+q(\ell,\ell)\right)$, and $\nu(h)$ is sufficiently small, then at the unique optimal wage scheme implementing PAM, both types of workers exert strictly positive effort and the following properties hold:*
   (a) *Both IC_ℓ and IC_h bind.*
   (b) $D^{AS}_{\ell\ell} < 0$ *and* $D^{AS}_{hh} < 0$.
   (c) $w_\ell > w_h = 0$.

Proof. See Appendix A. □

The first property of the Theorem indicates that the pure moral hazard optimal wage scheme is incentive compatible when complementarities are sufficiently strong. The intuition is as follows. Under PAM and the pure moral hazard optimal wage scheme, lows receive strictly higher bonuses than highs upon truthfully reporting their type. As complementarities become arbitrarily strong, the productivity of a low matched with a high becomes arbitrarily close to the productivity of a low matched with a low. Hence, lows strictly prefer a higher bonus payment over the benefit of matching with a teammate that yields him only a small increase in productivity. It follows that IC_ℓ does not bind. Reversing this logic, as complementarities become strong, the productivity of a high matched with a low becomes arbitrarily close to the productivity of a low matched with a low. This implies that the utility a high receives upon masquerading as a low becomes arbitrarily close to the utility of a truth-telling low. Under PAM and the pure moral hazard optimal wage scheme, however, a high receives a strictly higher expected utility than a low. It follows that IC_h does not bind.

The second property of the Theorem identifies conditions under which the pure moral hazard optimal wage scheme is *not* incentive compatible. In its proof, we show that if $\nu(h)$ is sufficiently small, then the manager finds it optimal to induce strictly positive effort by both types of workers, i.e. there is an interior solution to the manager's problem.[18] Moreover, at any interior solution, if highs are sufficiently important to production, i.e. the quality of two highs is sufficiently higher than the quality of two lows, and complementarities are weak, so that a low matched with a high can realize as much of the gains a high gives to a team as another high, then lows do not have an incentive to truthfully report their type. To satisfy this incentive constraint, the manager must reduce bonus payments to highs relative to their pure moral hazard levels. But, doing so leads to a violation of the incentive compatibility constraint for highs. As a consequence, *both* incentive compatibility constraints— IC_h and IC_ℓ — bind at any optimal wage scheme

[18] Absent this condition, the manager may find it optimal *not* to incentivize any effort by lows, i.e. set $b_{\ell\ell} = 0$.

and effort is distorted downwards for both types. The manager resolves this cyclicality by giving lows "lower-powered" incentives, increasing their base wage above zero and decreasing their bonus payment.

Second-Best Matching

When types are sufficiently important to production and complementarities are weak, separating highs and lows in the mixed model is costly because lows have a strong incentive to masquerade as a high in order to match with a more productive teammate. But providing optimal incentives for truth-telling by lows tightens the incentive compatibility constraint for highs, causing both incentive compatibility constraints to bind when it is optimal for lows to exert a non-trivial quantity of effort. We prove now that the manager can increase her profits by distorting the matching she implements in order to relax her incentive constraints, provided that the parameters of the model are in the regions identified in the second part of Theorem 3. We remark that the condition that $\nu(h)$ is sufficiently small now plays two roles: (i) it ensures that incentivizing effort by lows is profit-maximizing and (ii) it ensures that the efficiency gains of positive assortative matching are sufficiently small.

Theorem 4 (Mixed Model: Optimal Matching)**.**

1. *If $q(h,\ell)$ is sufficiently close to $q(\ell,\ell)$, then PAM is the unique optimal matching.*
2. *There exists an $M > 0$ such that if $\frac{q(\ell,\ell)}{q(h,h)} \leq M$, $q(h,\ell)$ is sufficiently close to $\frac{1}{2}\left(q(h,h) + q(\ell,\ell)\right)$, and $\nu(h)$ is sufficiently small, then PAM is strictly suboptimal.*

Proof.

1. Immediate from the first property of Theorem 3 (optimal pure moral hazard bonuses implement PAM) and the third property of Theorem 2 (PAM is the unique profit-maximizing under optimal pure moral hazard bonuses).

2. Suppose the manager implements a random matching, i.e. a matching μ for which $p_h^\mu = \nu(h) = p_\ell^\mu$, with base wages $w_h = w_\ell = 0$ and bonuses $b_{\theta\theta'} = b_{\ell\ell}^{MH}$ for all θ and θ'. Such a wage scheme is incentive compatible; a worker's report has no influence on his assigned teammate's type or on his (or his teammate's) compensation. As $\nu(h)$ approaches zero, observe that the manager's expected profits under this contract approach those under pure moral hazard. This follows because (i) expected output under random matching approaches expected output under PAM (given pure moral hazard bonuses) and (ii) expected wage payments approach their values under pure moral hazard.

   On the other hand, the proof of Theorem 3 establishes that distortions due to adverse selection are non-vanishing when the manager (optimally) implements PAM. In particular, the optimal value of $b_{\ell\ell}$ is bounded above by a number strictly less than $b_{\ell\ell}^{MH}$ as $\nu(h)$ becomes small (see the proof that $D_{\ell\ell}^{AS} < 0$). Hence, the manager's expected profits under this contract approach a value strictly less than under pure moral hazard.

   As we have identified a feasible contract that outperforms the best contract involving PAM, we have proven the result.

□

Although we have used limiting arguments to establish our matching distortion result, Figure 1 presents numerical examples in which distortions are quantitatively large even in the absence of extreme parameters.[19] This possibility is consistent with the analytical results of our companion article Kambhampati and Segura-Rodriguez (2021). In that article, we provide exact conditions under which random matching and negative assortative matching are uniquely optimal in a more tractable binary-effort setting in which the manager desires to implement high effort by all workers in all teams. We

[19] The Mathematica code is in Online Appendix A, available on our personal websites.

also provide sufficient conditions, similar to those of Theorem 4, under which positive assortative matching is strictly outperformed by negative assortative matching when the manager's effort implementation is allowed to vary.

Another comment is worth mentioning. Our Theorem shows that PAM is suboptimal when complementarities are sufficiently weak. In such cases, the productive value of forming teams is reduced. However, there are other reasons, such as saving on monitoring costs, to form teams in the first place (an issue beyond the scope of our analysis).

6 Delegation

As we have seen, it may be costly to implement efficient matching in a centralized workplace in which a manager asks each worker to report his type, and uses these reports to assign workers to teams. But what if, instead, the manager did not ask for reports, and simply allowed workers to sort themselves?

A tradeoff arises. On one hand, such an arrangement entails a *loss of control* for the manager: She can no longer tailor wages to reports. But on the other hand, the manager can exploit *local information*: Workers may possess superior information about one another's characteristics and use this information to sort efficiently.

We formalize this tradeoff by considering an environment in which there is no reporting stage, but in which, during the process of finding a teammate, workers commonly learn the true type profile. Using this knowledge, workers then form self-enforcing teams.

Timing and Contracts

The environment is the same as in Section 2. The timing, in contrast, is as follows:

1. The manager proposes a delegation contract.

2. Each worker accepts or rejects the proposed contract.

3. Workers commonly learn the type of each worker in the firm.

4. Workers form teams.

5. Workers exert effort.

6. The manager observes output and compensates each worker.

A **delegation contract** is simply a bonus $b \in \mathbb{R}_+$ paid to each worker when his team successfully completes its project. Though not formally part of a contract, we may think of the manager as choosing a matching, μ, in addition to a bonus, b. In contrast to the centralization environment, however, to implement a matching μ, μ must be self-enforcing given b in the sense that it is in the **core**. That is, there does not exist a non-zero measure set of workers who can match with one another and obtain a strictly higher payoff in some Nash equilibrium. We call any matching satisfying this property **stable**. The **manager's delegation problem** is to choose a matching and a bonus, (μ, b), to maximize profits, subject to the constraint that μ is stable with respect to b.

Optimal Delegation Contract

We now show that any non-trivial bonus leads to efficient sorting — the local information benefit of delegation.

Lemma 1. *For any delegation contract $b \in \mathbb{R}_+$ yielding the manager strictly positive profits, only PAM is stable with respect to b.*

Proof. By (1), any bonus $b \leq \frac{\beta}{q(h,h)}$ results in zero effort by any worker in any team. Hence, the manager obtains zero profit. In what follows, suppose that $b > \frac{\beta}{q(h,h)}$. We show that only PAM is stable with respect to any such b to prove the Lemma.

First, observe that a high's unique Nash equilibrium payoff is strictly increasing in his teammate's type $\theta_j \in \{\ell, h\}$. To see why, notice that both $e(h, \theta_j, b)$ and $e(\theta_j, h, b)$ are strictly increasing in θ_j and worker i's payoff is strictly increasing in worker j's effort,

$e(\theta_j, h, b)$. Revealed (strict) preference for higher effort by worker i given an increase in θ_j thus implies that worker i must receive a strictly higher payoff when θ_j increases from ℓ to h.

Second, observe that for any matching that is not PAM, there is a positive measure of teams composed of a high and a low. This matching cannot be stable with respect to b because, then, there exists a positive measure of highs, each matched to a low, who can form deviating teams and obtain strictly higher payoffs (by the observation in the preceding paragraph).

Finally, observe that PAM is stable with respect to b; no high matched with a high prefers to match a low and only a negligible set of highs may be matched with a low under PAM. □

Utilizing the previous Lemma, we identify the optimal delegation contract.

Lemma 2. *The optimal delegation contract is*

$$b_D := \frac{1}{4} + \frac{\beta}{2\gamma} \in (b_{hh}^{MH}, b_{\ell\ell}^{MH}),$$

where[20]

$$\gamma := \frac{\nu(\ell)q^2(\ell,\ell) + \nu(h)q^2(h,h)}{\nu(\ell)q(\ell,\ell) + \nu(h)q(h,h)} \in (q(\ell,\ell), q(h,h)).$$

Hence, effort in teams composed of two lows (highs) is strictly lower (higher) than under pure moral hazard.

Proof. From Lemma 1, for any $b \in [\frac{\beta}{q(h,h)}, \frac{1}{2}]$, the manager's expected profit is

$$\nu(h)\left[2 \cdot e(h,h,b) \cdot q(h,h) \cdot (1-2b)\right] + \nu(\ell)\left[2 \cdot e(\ell,\ell,b) \cdot q(\ell,\ell) \cdot (1-2b)\right].$$

[20]The mediant inequality establishes that for any strictly positive real numbers a, b, c, and d for which $a/c < b/d$,

$$\frac{a}{c} < \frac{a+b}{c+d} < \frac{b}{d}.$$

From (1), the above expression is a continuous function of b. Hence, it must have a maximizer in the compact set $[\frac{\beta}{q(h,h)}, \frac{1}{2}]$. On the interval $[\frac{\beta}{q(\ell,\ell)}, \frac{1}{2}]$, the necessary first-order condition for an optimal interior solution is

$$2 \cdot \nu(h) \cdot q(h,h) \cdot \left[\phi'(q(h,h) \cdot b) \cdot q(h,h) \cdot (1-2b) - 2 \cdot \phi(q(h,h) \cdot b)\right] +$$

$$2 \cdot \nu(\ell) \cdot q(\ell,\ell) \cdot \left[\phi'(q(\ell,\ell) \cdot b) \cdot q(\ell,\ell) \cdot (1-2b) - 2 \cdot \phi(q(\ell,\ell) \cdot b)\right] = 0.$$

Algebraic manipulation shows that its unique solution is b_D (as defined in the statement of the Lemma), which lies in the interior of the interval $[\frac{\beta}{q(h,h)}, \frac{\alpha+\beta}{q(h,h)}]$ and yields the manager a strictly positive expected payoff.

As any bonus $b \in [0, \frac{\beta}{q(h,h)}] \cup [\frac{\alpha+\beta}{q(h,h)}, \infty)$ yields the manager at most zero expected payoff, no such bonus can be optimal. It remains to consider bonuses $b \in [\frac{\beta}{q(h,h)}, \frac{\beta}{q(\ell,\ell)}]$ at which lows exert zero effort. But, $\frac{\beta}{q(\ell,\ell)} < b_{hh}^{MH}$ for any $\beta \in (0, \frac{q(\ell,\ell)}{4})$. Hence, the manager can increase her profits from highs (and lows) by increasing her bonus b strictly above $\frac{\beta}{q(\ell,\ell)}$. We have thus established that b_D is the unique optimal bonus. □

The optimal delegation contract is a weighted average of the optimal pure moral hazard bonuses paid to high quality and low quality teams, b_{hh}^{MH} and $b_{\ell\ell}^{MH}$. Intuitively, as β becomes larger, so that the gap between b_{hh}^{MH} and $b_{\ell\ell}^{MH}$ grows large, the inability of the manager to condition her bonuses on reported types becomes more costly — the loss of control cost of delegation.

Optimality of Delegation

Delegation is optimal if and only if the local information benefit outweighs this loss of control. We find conditions on the primitives of the model under which each force dominates.

Theorem 5 (Mixed Model: Delegation)**.**

1. *(Loss of Control Outweighs Local Information Benefit)*

   If $q(h, \ell)$ is sufficiently close to $q(\ell, \ell)$, then centralized assignment yields strictly higher profits than delegation.

2. *(Local Information Benefit Outweighs Loss of Control)*

   There exists an $M > 0$ such that if $\frac{q(\ell,\ell)}{q(h,h)} \leq M$, $q(h, \ell)$ is sufficiently close to $\frac{1}{2}(q(h, h) + q(\ell, \ell))$, and β is sufficiently close to zero, then delegation yields strictly higher profits than centralized assignment.

Proof.

1. For any $q(h, \ell)$, Lemma 2 shows that the manager's inability to tailor wages to types results in strict over provision of effort in teams composed of two highs and strict under provision of effort in teams composed of two lows (relative to the case of pure moral hazard). Hence, the manager's expected profits are strictly lower than under pure moral hazard. On the other hand, from the first property of Theorem 3, if $q(h, \ell)$ is sufficiently close to $q(\ell, \ell)$, expected profits under centralized assignment in the mixed model equal expected profits under pure moral hazard. The result follows.

2. The proof of Theorem 3 establishes that if $M \leq \frac{1}{4}$ and $q(h, \ell)$ is sufficiently close to $\frac{1}{2}(q(h, h) + q(\ell, \ell))$, then optimal bonuses under pure moral hazard are not incentive compatible for any $\beta \in [0, \frac{q(\ell,\ell)}{4}]$ and any prior over types ν. Hence, profits are bounded away from those under pure moral hazard. On the other hand, as β approaches zero, the manager's expected profits under delegation approach those under pure moral hazard. The result follows.

$\square$

An immediate implication of Theorem 4 part 2 and Theorem 5 part 2 is that delegation emerges as a profitable alternative to distorting PAM when β and $\nu(h)$ are both small. Rather than commit to a distorted matching and elicit reports, the manager would rather allow workers to sort themselves because workers, utilizing common knowledge of each other's types, will sort efficiently. In other words, if a manager contemplates distorting the efficient matching in order to reduce expected wage payments, there may be another option: write a "pay-for-performance" contract that does not depend on non-verifiable reports, and simply allow workers to sort themselves.

7 Extensions

We discuss the robustness of the results to various modifications of the model.

Alternative Production Functions

We first justify the claim made in Section 2 that all of our results extend under a general constant elasticity of substitution (CES) production function. The production function we have considered is a special case of the constant-returns-to-scale CES production function

$$(e_i^\rho + e_j^\rho)^{1/\rho} \cdot q(\theta_i, \theta_j),$$

with $\rho = 1$. We now consider the case in which ρ can take on any value in $(-\infty, 1] \backslash \{0\}$. We show that if workers can coordinate on a Pareto dominant Nash equilibrium in the complete information game played within each team, then our results hold as stated. This implies that distorting PAM is optimal for arbitrary complementarities in effort (in the limit as $\rho \to -\infty$, the CES production function converges to one exhibiting perfect complements).

We make three preliminary observations. First, observe that if two teammates receive a bonus payment of $b > \frac{2 \cdot \alpha \cdot \beta}{q(\theta_i, \theta_j)}$, then there exists a Nash equilibrium in which workers exert the same effort as they do in the case of perfect substitutes,

$$e_i(\theta_i, \theta_j, b) = e_j(\theta_i, \theta_j, b) = \frac{q(\theta_i, \theta_j)}{2\alpha} \cdot b - \beta,$$

yielding each an equilibrium payoff of

$$V(\theta_i, \theta_j, b, b).$$

Moreover, the manager obtains the same expected profit in this equilibrium for any complementarity parameter ρ (symmetry of effort cancels out the effect of complementarity on expected output). Second, for any $b > \frac{2 \cdot \alpha \cdot \beta}{q(\theta_i, \theta_j)}$ and any $e_i > 0$, observe that the marginal return to effort is strictly increasing in e_j and that worker i's payoff is strictly increasing in the effort of his teammate j. Hence, the unique Pareto efficient Nash equilibrium must be the largest symmetric strategy profile. Finally, observe that no symmetric strategy profile in which $e_i = e_j > e_i(\theta_i, \theta_j, b)$ can be a Nash equilibrium; for any such profile and any ρ, the marginal cost of effort for each worker strictly exceeds the marginal return to effort.

In light of the preceding paragraph, we see that if workers can select their preferred Nash equilibrium and each receives the same bonus payment, then both workers and the manager obtain the same payoff for all feasible ρ. Because under full information, pure adverse selection, pure moral hazard, and delegation, any two workers in the same team receive the same bonus, it follows that all our results in these settings hold. In addition, when the manager implements PAM in the mixed model, she only forms teams composed of two highs and two lows, with workers of each type receiving the same bonus. Hence, Theorem 3 holds. Finally, we have proven Theorem 4 by showing that the manager can do strictly better than any implementation of PAM by implementing a random matching and offering all workers the same bonus. As profits under such a contract remain the same for any value of ρ, Theorem 4 holds as well.

Two qualifying remarks are in order. First, although our PAM distortion result holds for production functions exhibiting any degree of complementarities in effort, it does not characterize the *degree* to which matchings are distorted as complementarities in effort increase. One may expect that as complementarities increase, PAM becomes more attractive; the productivity loss of breaking up a team composed of two highs increases. Motivated by this observation, in a companion article Kambhampati and Segura-Rodriguez (2021), we explore a production environment featuring binary, perfectly complementary effort. In that model, we solve for optimal wage schemes and matchings in closed form as complementarities between *types* vary. However, it remains an open question to derive the comparative statics of second-best optimal matching as complementarities in *effort* vary.

Second, we have assumed a production specification in which an aggregate function of types multiplies total effort. If, instead, a worker's own type affects his individual return to effort in addition to a common return, as in our model, then matching a low with a high could dampen effort incentives for the high. This force might push second-best optimal matching in the direction of PAM.[21] We leave a complete analysis of this extension to future research.

Alternative Timing of Events

We do not investigate mechanisms that provide information to the manager after teams have been formed. Although relaxing this assumption is theoretically interesting, we believe this restriction is plausible in environments in which peer evaluation is ineffective at generating reliable reports. As previously mentioned, two explanations for why this may be the case are that peer reports may be subject to collusion and reporting parties may fear retaliation by their co-workers (see Che and Yoo (2001), who make a similar non-contractability assumption, for a discussion of the former point and Chassang

[21] We thank an anonymous referee for this observation.

and Zehnder (2019) for recent work related to the latter point). An interesting, though challenging, direction for future research would be to study whether the distortion we identify holds in a dynamic contracting environment in which information about one's teammate arrives over time and the manager demands contracts to be robust to collusion and/or that individual reports cannot be identified by the manager's chosen matching.

We remark, however, that our matching distortion result, Theorem 4, holds when allowing for complete contracts under the following alternative learning environment and timing:[22]

1. The manager proposes a contract.
2. Each worker accepts or rejects the proposed contract.
3. Each worker reports his type to the manager.
4. The manager assigns workers to teams.
5. Workers (not knowing their teammate's type) exert effort.
6. The manager observes output and compensates each worker.

The key subtlety that must be accounted for under this alternative timing is that each worker's effort can only depend on his beliefs about the quality of his team given the matching, his report, and knowledge of his own type. But, the matching and wage scheme proposed in the proof of Theorem 4 that dominates any implementation of PAM continues to be incentive compatible and again approaches pure moral hazard profits as $\nu(h)$ grows small (as $\nu(h)$ grows small, each worker's belief about his teammate's type approaches the point mass on ℓ). At the same time, profits under PAM remain bounded away from those under pure moral hazard.[23] Hence, Theorem 4 holds as stated.

[22]Notice, Maskin-type mechanisms are no longer effective in this environment because workers do not learn one another's type.

[23]We prove this claim in Online Appendix B, available on our personal websites. The steps of the proof mirror those in Appendix A (only slight modifications of the first-order conditions are required).

Unlimited Liability

Finally, we point out that limited liability constraints are, in fact, critical for our results. In Appendix B, we study the manager's problem when she can offer negative base wages. We obtain the following results:

1. Under observable effort, the manager can obtain first-best profits using the same contracts described in Section 3.

2. Under pure moral hazard, the manager can "sell each team" to its workers and obtain first-best profits.

3. In the mixed model with both moral hazard and adverse selection, distorting PAM is never profit maximizing if $\nu(h)$ is sufficiently small, even though first-best profits are unattainable.

Matching distortions in the mixed model are no longer optimal because the manager can use negative base wages to separate types; highs are willing to accept a lower base wages in exchange for higher quality teammates. This results in lower effort distortions for both high and low quality workers when the manager implements PAM; in fact, as highs disappear from the population, effort levels approach their efficient values. This starkly contrasts with the case in which limited liability constraints are present.

In summary, it is the interaction of all of the constraints in our model— obedience constraints, incentive compatibility constraints, and limited liability constraints — that gives rise to inefficient matching.

8 Conclusion

Our analysis identifies a new channel by which asymmetric information distorts efficient matching inside of firms. When types are sufficiently important to production and complementarities are weak, less productive workers have an incentive to misreport their type in order to match with more productive teammates if the manager implements positive assortative matching. But increasing the wages of less productive workers incentivizes more productive workers to misreport their type. When incentivizing effort by less productive workers is sufficiently valuable, both incentive compatibility constraints bind at any optimal wage scheme. In order to relax these constraints, a profit-maximizing manager may decrease the assortativity of the matching she implements. We investigate the implications of this result for the optimal management of teams inside the firm and find conditions under which delegating the sorting problem to workers themselves outperforms centralized assignment. Together, our results offer new explanations for non-assortative and delegated matching inside of firms.

Appendix A

We prove Theorem 3 using the following steps. First, we prove property 1 by establishing that the optimal wage scheme under pure moral hazard is incentive compatible when complementarities are strong. This scheme, of course, satisfies 1 (a)-(c). Second, fixing an arbitrary matching with assortativity p_h^μ, we establish that there exists a solution to the manager's optimal wage problem for any parameters of the model. Third, fixing a matching with assortativity $p_h^\mu = 1$ (PAM), we show that if the solution to the manager's problem involves interior effort by both types of workers, then it is unique and satisfies the properties 2 (a)-(c). Fourth, fixing a matching with assortativity $p_h^\mu = 1$ (PAM), we show that non-interior effort cannot be optimal if $\nu(h)$ is small. Together with the third step, this proves property 2.

Incentive Compatibility Under Strong Complementarities

Suppose bonuses are given by their values under pure moral hazard,

$$b_{\theta_i\theta_j} = \frac{1}{4} + \frac{\beta}{2q(\theta_i, \theta_j)}.$$

For any $\beta > 0$, lows receive strictly higher bonus payments than highs, i.e. $b_{\ell\ell} > b_{hh}$. Hence, if $q(h, \ell)$ is sufficiently close to $q(\ell, \ell)$, so that the productivity gain a low receives when matching with a high instead of a low is made arbitrarily small, IC_ℓ must be satisfied.

We show that if $q(h, \ell)$ is sufficiently close to $q(\ell, \ell)$, then IC_h is strictly satisfied as well. First, observe that that

$$\underbrace{2 \cdot e(h, h, b_{hh}) \cdot q(h, h) \cdot b_{hh} - c(e(h, h, b_{hh}))}_{=U(h,h)} \geq \underbrace{2 \cdot e(h, \ell, b_{\ell\ell}) \cdot q(h, \ell) \cdot b_{\ell\ell} - c(e(h, \ell, b_{\ell\ell}))}_{=U(h,\ell)}$$

$$\iff \left(\frac{q(h,h)}{8\alpha}-\frac{\beta}{4\alpha}\right)\frac{q(h,h)}{2}-\alpha\left(\frac{q(h,h)}{8\alpha}-\frac{\beta}{4\alpha}\right)^2 \geq$$

$$\left(\frac{q(h,\ell)}{8\alpha}+\frac{\beta}{2\alpha}\left(\frac{1}{2}\frac{q(h,\ell)}{q(\ell,\ell)}-1\right)\right)\left(\frac{q(h,\ell)}{2}+\beta\left(\frac{q(h,\ell)}{q(\ell,\ell)}-1\right)\right)$$

$$-\alpha\left(\frac{q(h,\ell)}{8\alpha}+\frac{\beta}{2\alpha}\left(\frac{1}{2}\frac{q(h,\ell)}{q(\ell,\ell)}-1\right)\right)^2.$$

Second, observe that as $q(h,\ell)$ approaches $q(\ell,\ell)$, the right-hand becomes arbitrarily close to

$$\left(\frac{q(\ell,\ell)}{8\alpha}-\frac{\beta}{4\alpha}\right)\frac{q(\ell,\ell)}{2}-\alpha\left(\frac{q(\ell,\ell)}{8\alpha}-\frac{\beta}{4\alpha}\right)^2.$$

Third, observe that this expression is strictly smaller than

$$\left(\frac{q(h,h)}{8\alpha}-\frac{\beta}{4\alpha}\right)\frac{q(h,h)}{2}-\alpha\left(\frac{q(h,h)}{8\alpha}-\frac{\beta}{4\alpha}\right)^2$$

because

$$\frac{\partial}{\partial q}\left[\left(\frac{q}{8\alpha}-\frac{\beta}{4\alpha}\right)\frac{q}{2}-\alpha\left(\frac{q}{8\alpha}-\frac{\beta}{4\alpha}\right)^2\right]=$$

$$\frac{3}{4}\frac{q}{8\alpha}-\frac{\beta}{8\alpha},$$

which is strictly positive whenever $0<\beta<\frac{q}{2}$ and $0<\alpha$. Hence, for $q(h,\ell)$ sufficiently close to $q(\ell,\ell)$, IC_h holds strictly.

Existence of a Global Solution

First, observe that it is without loss of generality to constrain bonuses $b_{\theta\theta'}$ to lie in the compact set $\{0\}\cup[\frac{\beta}{q(\theta,\theta')},\frac{\alpha+\beta}{q(\theta,\theta')}]$, where $\frac{\beta}{q(\theta,\theta')}$ is the bonus at which zero effort is optimal for a worker of type θ matched with one of type θ' and $\frac{\alpha+\beta}{q(\theta,\theta')}$ is the bonus at which $\frac{1}{2}$ effort is optimal. Any value between 0 and $\frac{\beta}{q(\theta,\theta')}$ results in zero effort by such a worker. Hence, if such a bonus were to be optimal, it would also be optimal to set $b_{\theta\theta'}=0$. This modification not only keeps effort fixed at zero, but also minimizes the payoff a worker of type θ' obtains when masquerading as a worker of type θ, thereby relaxing $IC_{\theta'}$. Similarly, for any bonus $b_{\theta\theta'} > \frac{\alpha+\beta}{q(\theta,\theta')}$, the manager can reduce $b_{\theta\theta'}$ and increase w_θ. This modification not only keeps effort fixed at $\frac{1}{2}$, but also minimizes the payoff a worker of type θ' obtains when masquerading as a worker of type θ, thereby relaxing $IC_{\theta'}$.

Second, observe that it is without loss of generality to bound w_h and w_ℓ above by a large value $W > 0$ at which the manager could obtain at most strictly negative profits by setting either w_h and w_ℓ larger than or equal to W. As the manager can always obtain zero expected profits by setting all wages and bonuses equal to zero, such base wages could never be optimal.

Third, observe that the intersection of the boundary constraints on transfers and the incentive compatibility constraints is compact and the manager's expected payoff is continuous in her choice variables. The Weierstrass Theorem thus implies the existence of a solution to the manager's maximization problem.

Candidate Maximizer with Interior Effort

We first show that the constant rank constraint qualification for the necessity of the KKT optimality conditions holds at any maximizer with interior effort by workers of both types. By construction, the upper bound constraints on base wages never bind and, by hypothesis, we have assumed that no boundary constraints on bonuses bind. Let the choice variables be ordered by $(w_h, w_\ell, b_{hh}, b_{\ell\ell})$ and the remaining constraints be ordered by (LL_h, LL_ℓ, IC_h, IC_ℓ), where LL_θ is the non-negativity constraint on w_θ. Then, the

Jacobian with respect to the constraints is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & -1 & p_h^\mu q(h,h)\left(2e(h,h,b_{hh}) + \frac{\partial e(h,h,b_{hh})}{\partial b_i} b_{hh}\right) & -(1-p_\ell^\mu) q(h,\ell)\left(2e(h,\ell,b_{\ell\ell}) + \frac{\partial e(h,\ell,b_{\ell\ell})}{\partial b_i} b_{\ell\ell}\right) \\ -1 & 1 & -p_h^\mu q(h,\ell)\left(2e(h,\ell,b_{hh}) + \frac{\partial e(h,\ell,b_{hh})}{\partial b_i} b_{hh}\right) & (1-p_\ell^\mu) q(\ell,\ell)\left(2e(\ell,\ell,b_{\ell\ell}) + \frac{\partial e(\ell,\ell,b_{\ell\ell})}{\partial b_i} b_{\ell\ell}\right) \end{bmatrix}.$$

The last two rows are linearly independent whenever either $b_{hh} > \frac{\beta}{q(h,h)}$ or $b_{\ell\ell} > \frac{\beta}{q(\ell,\ell)}$. Hence, the Jacobian has full rank at any interior solution.

Let λ_θ be the Lagrange multiplier on IC_θ and γ_θ be the Lagrange multiplier on the non-negativity constraint of w_θ. The first-order condition with respect to w_h is

$$-2 \cdot \nu(h) + (\lambda_h - \lambda_\ell) + \gamma_h = 0.$$

The first-order condition with respect to w_ℓ is

$$-2 \cdot \nu(\ell) - (\lambda_h - \lambda_\ell) + \gamma_\ell = 0.$$

Summing, we see that

$$\gamma_h + \gamma_\ell = 2.$$

Hence, either $\gamma_h > 0$ or $\gamma_\ell > 0$ (or both). By complementary slackness, it must be that either $w_h = 0$ or $w_\ell = 0$ (or both). The first-order conditions with respect to bonuses result in the following characterization of second-best effort:

$$e(\theta_i, \theta_j, b_{\theta_i\theta_j}) = \phi\left(q(\theta_i, \theta_j)\left[1 + D^{MH}_{\theta_i\theta_j} + D^{AS}_{\theta_i\theta_j}\right]\right),$$

where

$$D^{MH}_{\theta\theta} := -\left[2 \cdot \left(\frac{\partial e(\theta,\theta,b^{MH}_{\theta\theta})}{\partial b_i}\right)^{-1} \cdot e(\theta,\theta,b^{MH}_{\theta\theta}) + b^{MH}_{\theta\theta}\right] \quad \text{and}$$

$$D^{AS}_{\theta\theta} := \frac{\lambda_\theta}{\nu(\theta)}\left[\frac{e(\theta,\theta,b_{\theta\theta})}{\frac{\partial e(\theta,\theta,b_{\theta\theta})}{\partial b_i}} + \frac{b_{\theta\theta}}{2}\right] - \frac{\lambda_{\theta'}}{\nu(\theta)}\left[\frac{q(\theta,\theta')}{q(\theta,\theta)}\left(\frac{e(\theta,\theta',b_{\theta\theta})}{\frac{\partial e(\theta,\theta,b_{\theta\theta})}{\partial b_i}} + \frac{\frac{\partial e(\theta,\theta',b_{\theta\theta})}{\partial b_i}}{\frac{\partial e(\theta,\theta,b_{\theta\theta})}{\partial b_i}}\frac{b_{\theta\theta}}{2}\right)\right],$$

with $\theta' \neq \theta$. We use the calculations

$$e(\theta,\theta',b) = \frac{q(\theta,\theta')\cdot b}{2\alpha} - \frac{\beta}{2\alpha} \text{ and}$$
$$\frac{\partial e(\theta,\theta',b)}{\partial b_i} = \frac{q(\theta,\theta')}{2\alpha}$$

to simplify these expressions. In particular, moral hazard distortions are given by

$$D^{MH}_{hh} = -3 \cdot b_{hh} + \frac{2\beta}{q(h,h)} \quad \text{and}$$
$$D^{MH}_{\ell\ell} = -3 \cdot b_{\ell\ell} + \frac{2\beta}{q(\ell,\ell)}.$$

Adverse selection distortions are given by

$$D^{AS}_{hh} = \frac{\lambda_h}{\nu(h)}\left[\frac{3}{2}\cdot b_{hh} - \frac{\beta}{q(h,h)}\right] - \frac{\lambda_\ell}{\nu(h)}\left[\frac{3}{2}\frac{q^2(h,\ell)}{q^2(h,h)}b_{hh} - \beta\frac{q(h,\ell)}{q^2(h,h)}\right] \quad \text{and}$$
$$D^{AS}_{\ell\ell} = \frac{\lambda_\ell}{\nu(\ell)}\left[\frac{3}{2}\cdot b_{\ell\ell} - \frac{\beta}{q(\ell,\ell)}\right] - \frac{\lambda_h}{\nu(\ell)}\left[\frac{3}{2}\frac{q^2(h,\ell)}{q^2(\ell,\ell)}b_{\ell\ell} - \beta\frac{q(h,\ell)}{q^2(\ell,\ell)}\right].$$

Manipulating the first-order conditions with respect to b_{hh} and $b_{\ell\ell}$, we see that

$$b_{hh} = \frac{\nu(h)q^2(h,h) + 2\beta\nu(h)q(h,h) + \lambda_\ell\beta q(h,\ell) - \lambda_h\beta q(h,h)}{4\nu(h)q^2(h,h) + \lambda_\ell\frac{3}{2}q^2(h,\ell) - \lambda_h\frac{3}{2}q^2(h,h)}$$

and

$$b_{\ell\ell} = \frac{\nu(\ell)q^2(\ell,\ell) + 2\beta\nu(\ell)q(\ell,\ell) - \lambda_\ell\beta q(\ell,\ell) + \lambda_h\beta q(h,\ell)}{4\nu(\ell)q^2(\ell,\ell) - \lambda_\ell\frac{3}{2}q^2(\ell,\ell) + \lambda_h\frac{3}{2}q^2(h,\ell)}$$

at any interior solution to the manager's maximization problem.

Proof that both IC_h and IC_ℓ bind.

Suppose $\frac{q(\ell,\ell)}{q(h,h)} \leq M \leq \frac{1}{4}$. We argue that it must be the case that $\lambda_h > 0$ and $\lambda_\ell > 0$ when $q(h,\ell)$ is sufficiently close to $\frac{1}{2}(q(h,h)+q(\ell,\ell))$ by deriving contradictions in the three other cases: (i) $\lambda_\ell = \lambda_h = 0$; (ii) $\lambda_\ell = 0$ and $\lambda_h > 0$; and (iii) $\lambda_\ell > 0$ and $\lambda_h = 0$.

(i) If $\lambda_\ell = \lambda_h = 0$, then $\gamma_\ell > 0$ and $\gamma_h > 0$. Hence, $w_\ell = w_h = 0$ and the first-order conditions imply

$$b_{\theta_i\theta_j} = \frac{1}{4} + \frac{\beta}{2q(\theta_i,\theta_j)} \in (\frac{1}{4}, \frac{1}{2})$$

at any point satisfying the KKT conditions. But, IC_ℓ is satisfied if and only if

$$\frac{3}{4}q(\ell,\ell)^2 b_{\ell\ell}^2 - \beta q(\ell,\ell)b_{\ell\ell} - \frac{3}{4}q(h,\ell)^2 b_{hh}^2 + \beta q(h,\ell)b_{hh} \geq 0,$$

which is satisfied only if

$$b_{hh} \leq \frac{q(\ell,\ell)}{q(h,\ell)} b_{\ell\ell} := \bar{b}.$$

As $q(h,\ell)$ approaches $\frac{1}{2}(q(h,h)+q(\ell,\ell))$, $\bar{b}$ becomes arbitrarily close to

$$2b_{\ell\ell}\frac{q(\ell,\ell)}{q(h,h)+q(\ell,\ell)}.$$

As $b_{\ell\ell} \leq \frac{1}{2}$ for any parameters at any point satisfying the KKT conditions, and $\frac{q(\ell,\ell)}{q(h,h)+q(\ell,\ell)} < \frac{q(\ell,\ell)}{q(h,h)} \leq \frac{1}{4}$, $\bar{b}$ must eventually lie strictly below $\frac{1}{4}$. Hence, b_{hh} must lie strictly below $\frac{1}{4}$ in order to satisfy IC_ℓ. But, this contradicts our initial expression for b_{hh}, which is always strictly greater than $\frac{1}{4}$.

(ii) If $\lambda_h > 0$ and $\lambda_\ell = 0$, then $\gamma_\ell > 0$ and $w_\ell = 0$. If $w_h = 0$, then, as argued in (i), it must be that

$$b_{hh} \leq \bar{b}$$

in order to satisfy IC_ℓ, where $\bar{b}$ approaches a value strictly smaller than $\frac{1}{4}$ as $q(h,\ell)$

approaches $\frac{1}{2}(q(h,h)+q(\ell,\ell))$. However, by the necessary first-order conditions, it must be that

$$b_{hh} > \frac{1-\frac{\beta}{q(h,h)}\frac{\lambda_h}{\nu(h)}}{4-\frac{3}{2}\frac{\lambda_h}{\nu(h)}} > \frac{1}{4}.$$

To see why, notice that the inequality holds if and only if

$$4\frac{\beta}{q(h,h)}\frac{\lambda_h}{\nu(h)} < \frac{3}{2}\frac{\lambda_h}{\nu(h)}.$$

This always holds because

$$4\frac{\beta}{q(h,h)}\frac{\lambda_h}{\nu(h)} < 2\frac{q(\ell,\ell)}{q(h,h)}\frac{\lambda_h}{\nu(h)} < \frac{1}{2}\frac{\lambda_h}{\nu(h)},$$

where the first inequality follows from $\beta < \frac{q(\ell,\ell)}{2}$ and the second from $\frac{q(\ell,\ell)}{q(h,h)+q(\ell,\ell)} < \frac{q(\ell,\ell)}{q(h,h)} \leq \frac{1}{4}$. Hence, we have arrived at our desired contradiction for the case in which $w_h = 0$.

If $w_h > 0$, on the other hand, then $\gamma_h = 0$ and $\lambda_h = 2\nu(h)$. In this case,

$$b_{hh} = 1 \text{ and}$$

$$b_{\ell\ell} = \frac{\nu(\ell)q^2(\ell,\ell)+2\nu(\ell)\beta q(\ell,\ell)+2\nu(h)\beta q(h,\ell)}{4\nu(\ell)q^2(\ell,\ell)+3\nu(h)q^2(h,\ell)} < 1.$$

As $w_\ell < w_h$ and $q(\ell,\ell) < q(h,\ell)$, however, IC_ℓ cannot be satisfied.

(iii) If $\lambda_\ell > 0$ (so that IC_ℓ holds with equality) and $\lambda_h = 0$, then $\gamma_h > 0$ and $w_h = 0$. If $w_\ell = 0$, then, as argued in (i), it must be that

$$b_{hh} \leq \bar{b}$$

in order to satisfy IC_ℓ, where $\bar{b}$ approaches a value strictly smaller than $\frac{1}{4}$ as $q(h,\ell)$ approaches $\frac{1}{2}(q(h,h)+q(\ell,\ell))$. However, by the necessary first-order conditions, it must be that

$$b_{hh} = \frac{1 + \frac{2\beta}{q(h,h)} + \lambda_\ell \left(\frac{\beta}{\nu(h)} \frac{q(h,\ell)}{q^2(h,h)} \right)}{4 + \lambda_\ell \left(\frac{3q^2(h,\ell)}{2\nu(h)q^2(h,h)} \right)},$$

a decreasing function of λ_ℓ. As the expression is larger than $\frac{1}{4}$ when $\lambda_\ell = 0$, we have arrived at our desired contradiction.

If, on the other hand, $w_\ell > 0$, then $\gamma_\ell = 0$ and $\lambda_\ell = 2\nu(\ell)$. In this case,

$$b_{hh} = \frac{\nu(h)q^2(h,h) + 2\nu(h)\beta q(h,h) + 2\nu(\ell)\beta q(h,\ell)}{4\nu(h)q^2(h,h) + 3\nu(\ell)q^2(h,\ell)} < \frac{1}{2} \text{ and}$$

$$b_{\ell\ell} = 1.$$

In addition, IC_h requires that

$$3q^2(h,h)b_{hh}^2 - 4\beta b_{hh}q(h,h) - 3q^2(h,\ell) + 4\beta q(h,\ell) - 4w_\ell \geq 0.$$

Whenever $q(h,\ell) > \frac{q(h,h)}{2}$, which always holds for $q(h,\ell)$ close to $\frac{1}{2}(q(h,h) + q(\ell,\ell))$, the inequality cannot be satisfied for any $b_{hh} < \frac{1}{2}$.

Proof that $w_\ell > 0$ and $w_h = 0$.

From the first-order conditions with respect to w_h and w_ℓ, we know that either $w_\ell = 0$, $w_h = 0$, or both. We show by contradiction that it cannot be that $w_\ell = 0$ and $w_h \geq 0$ by demonstrating that IC_ℓ cannot be satisfied when the conditions of the Theorem are met.

We use the following preliminaries: By complementary slackness, $w_\ell = 0$ and $w_h > 0$ implies that $\gamma_h = 0$. Further, the first-order condition with respect to w_h implies that

$\lambda_h - \lambda_\ell = 2\nu(h)$. The first-order conditions with respect to b_{hh} and $b_{\ell\ell}$ yield

$$b_{hh} = \frac{\nu(h)q^2(h,h) - \lambda_\ell \beta(q(h,h) - q(h,\ell))}{\nu(h)q^2(h,h) + \lambda_\ell \frac{3}{2}(q^2(h,\ell) - q^2(h,h))} \quad \text{and}$$
$$b_{\ell\ell} = \frac{\nu(\ell)q^2(\ell,\ell) + 2\beta\nu(\ell)q(\ell,\ell) - \lambda_\ell \beta q(\ell,\ell) + \lambda_h \beta q(\ell,\ell)}{4\nu(\ell)q^2(\ell,\ell) - \lambda_\ell \frac{3}{2}q(\ell,\ell)^2 + \lambda_h \frac{3}{2}q(h,\ell)^2}.$$

As $\beta \in [0, q(\ell,\ell)/2]$ and $\lambda_h > \lambda_\ell$, we have $\frac{3}{2}\lambda_h q^2(h,\ell) - \frac{3}{2}\lambda_\ell q^2(\ell,\ell) > \lambda_h \beta q(h,\ell) - \lambda_\ell \beta q(\ell,\ell)$. Hence, $b_{\ell\ell} < 1$ for any q.

Using the definition of b_{hh}, we have that

$$U(\hat{\theta} = h, \theta = \ell) = \frac{3}{4}q^2(h,\ell)b_{hh}^2 - \beta b_{hh} q(h,\ell) + w_h > 0 \iff$$

$$\left(\frac{3}{4}q(h,\ell) - \beta\right)\nu(h)q^2(h,h) + \frac{3}{4}\lambda_\ell \beta q(h,h)(q(h,h) - q(h,\ell))$$
$$+\frac{3}{4}\lambda_\ell \beta(q^2(h,h) - q^2(h,\ell)) + w_h > 0,$$

which always holds because $\beta < \frac{q(\ell,\ell)}{2}$ and $q(\ell,\ell) < q(h,\ell) < q(h,h)$. Therefore, if a low masquerades as a high, he receives a utility that is positive and bounded away from 0. On the other hand, if he reports his type truthfully, he receives a utility equal to

$$U(\ell,\ell) = \frac{3}{4}q^2(\ell,\ell)b_{\ell\ell}^2 - \beta b_{\ell\ell} q(\ell,\ell),$$

which converges to zero as $q(\ell,\ell)$ becomes small because $b_{\ell\ell}$ is bounded above by 1. Hence, there exists an $M > 0$ such that if the conditions of the Theorem are satisfied, then it cannot be both that $w_\ell = 0$ and $w_h \geq 0$ without violating either IC_ℓ or the KKT necessary conditions for optimality. This completes the proof.

Proof that the interior maximizer is unique.

Plugging in $\lambda_\ell = 2\nu(\ell) + \lambda_h$, we have

$$b_{hh} = \frac{\nu(h)q^2(h,h) + 2\beta\nu(h)q(h,h) + 2\beta\nu(\ell)q(h,\ell) - \lambda_h\beta(q(h,h) - q(h,\ell))}{4\nu(h)q^2(h,h) + 3\nu(\ell)q^2(h,\ell) - \lambda_h\frac{3}{2}(q^2(h,h) - q^2(h,\ell))} \quad \text{and}$$
$$b_{\ell\ell} = \frac{\nu(\ell)q^2(\ell,\ell) + \lambda_h\beta(q(h,\ell) - q(\ell,\ell))}{\nu(\ell)q^2(\ell,\ell) + \lambda_h\frac{3}{2}(q^2(h,\ell) - q^2(\ell,\ell))}.$$

Algebraic manipulation yields $\frac{\partial b_{\ell\ell}}{\partial\lambda_h} < 0$ and $\frac{\partial b_{hh}}{\partial\lambda_h} > 0$. Summing the binding constraints IC_h and IC_ℓ, we see that incentive compatible bonuses, b_{hh} and $b_{\ell\ell}$, must satisfy

$$\frac{3}{4}(q(h,h)^2 - q(h,\ell)^2)b_{hh}^2 - \beta(q(h,h) - q(h,\ell))b_{hh} =$$

$$\frac{3}{4}(q(h,\ell)^2 - q(\ell,\ell)^2)b_{\ell\ell}^2 - \beta(q(h,\ell) - q(\ell,\ell))b_{\ell\ell}.$$

From this condition, we conclude that if $q(h,\ell)$ is close to $\frac{q(h,h)+q(\ell,\ell)}{2}$, then (i) $b_{hh} < b_{\ell\ell}$, (ii) the left-hand side is strictly increasing in b_{hh}, and (iii) the right-hand side is strictly increasing in $b_{\ell\ell}$. Hence, as $\frac{\partial b_{\ell\ell}}{\partial\lambda_h} < 0$, the left-hand side is strictly increasing in λ_h and, because $\frac{\partial b_{hh}}{\partial\lambda_h} > 0$, the right-hand side is strictly decreasing in λ_h. Therefore, there exists a unique λ_h at which the equality holds. As the optimal wage scheme is uniquely determined by λ_h, the interior KKT solution is unique.

Proof that $D_{\ell\ell}^{AS} < 0$ and $D_{hh}^{AS} < 0$.

To show that $D_{\ell\ell}^{AS} < 0$ and $D_{hh}^{AS} < 0$, we must show that $b_{\ell\ell} < b_{\ell\ell}^{MH} = \frac{1}{4} + \frac{\beta}{2q(\ell,\ell)}$ and $b_{hh} < b_{hh}^{MH} = \frac{1}{4} + \frac{\beta}{2q(h,h)}$ in the parameter regions identified in the statement of the Theorem.

Let $b_{hh} = \frac{q(h,\ell)+q(\ell,\ell)}{q(h,h)+q(h,\ell)}b_{\ell\ell}$ and denote by λ_h^1 the (unique) multiplier at which the first-order conditions would yield this bonus. As $\nu(h)$ approaches zero and $q(h,\ell)$ approaches

$\frac{1}{2}(q(h,h)+q(\ell,\ell))$, λ_h^1 approaches

$$\frac{\eta(1+\eta)(\frac{3}{4}(1+\eta)(1+3\eta)-\delta\eta(3+\eta))}{\frac{(1-\eta)}{2}(\frac{3}{2}\delta(1+\eta)(1+3\eta)+\frac{3}{4}\eta(3+\eta)(1+3\eta)-\delta\eta^2(3+\eta))},$$

where $\eta = \frac{q(\ell,\ell)}{q(h,h)}$ and $\delta = \frac{\beta}{q(\ell,\ell)}$. In addition, the right-hand side of the sum of the incentive compatibility constraints is strictly larger than the left-hand side. Hence, the optimal value of λ_h must be larger than λ_h^1.

Now, let λ_h^2 be the value of λ_h at which $b_{\ell\ell} = b_{\ell\ell}^{MH}$. As $\nu(h)$ approaches zero and $q(h,\ell)$ approaches $\frac{1}{2}(q(h,h)+q(\ell,\ell))$, λ_h^2 approaches

$$\frac{\left(\frac{3}{2}-\delta\right)\eta^2}{\frac{1-\eta}{8}(\delta\eta+3\delta+\frac{3}{2}(1+3\eta))}.$$

Comparing the two expressions, we see that if $\frac{q(\ell,\ell)}{q(h,h)} \leq M$ for M sufficiently small, so that η is sufficiently close to zero, then $\lambda_h^2 < \lambda_h^1$. It follows that the optimal value of λ_h, λ_h^*, satisfies $\lambda_h^* > \lambda_h^1 \geq \lambda_h^2$. Since $\frac{\partial b_{\ell\ell}}{\partial \lambda_h} < 0$, we conclude that $b_{\ell\ell} < b_{\ell\ell}^{MH}$.

Finally, denote by λ_h^3 the value that makes $b_{hh} = b_{hh}^{MH}$. As $\nu(h)$ approaches zero and $q(h,\ell)$ approaches $\frac{1}{2}(q(h,h)+q(\ell,\ell))$, λ_h^3 approaches

$$\frac{2\delta\eta(1+\eta)-(\frac{1}{2}+\delta\eta)\frac{3}{4}(1+\eta)^2}{\frac{1-\eta}{8}(8\delta\eta-3(3+\eta)(\frac{1}{2}+\delta\eta))}.$$

Plugging this value into the first-order condition expression for $b_{\ell\ell}$ and b_{hh}, we obtain that $b_{\ell\ell}$ is approximately equal to

$$\frac{4(4\eta(3+\eta)+3\delta(1+\eta)^2)}{16(3+\eta)\eta^2+9(1+3\eta)(1+\eta)^2}$$

and b_{hh} is approximately equal to

$$\frac{4(4\delta(3+\eta)-3\delta(1+\eta))}{3(3+\eta)(1+\eta)}.$$

Therefore, if $\frac{q(\ell,\ell)}{q(h,h)} \leq M$ for M sufficiently small, so that η is sufficiently close to zero, then $b_{\ell\ell}$ is close to $\frac{\delta}{3}$ and b_{hh} is close to δ, a contradiction since, by the KKT conditions, it is necessary that $b_{\ell\ell} > b_{hh}$. Since $\frac{\partial b_{\ell\ell}}{\partial \lambda_h} < 0$ and $\frac{\partial b_{hh}}{\partial \lambda_h} > 0$, it has to be that $\lambda_h^* < \lambda_h^3$, which, in turn, implies that $b_{hh} < b_{hh}^{MH}$.

Candidate Maximizers with Non-Interior Effort

We first argue that we need only consider two candidate maximizers: $b_{hh} = 0$ and $b_{\ell\ell} > \frac{\beta}{q(\ell,\ell)}$, or $b_{\ell\ell} = 0$ and $b_{hh} > \frac{\beta}{q(h,h)}$. If $b_{hh} \in (0, \frac{\beta}{q(h,h)})$, then the manager could reduce b_{hh} and weakly increase her profits. If $b_{\ell\ell} \in (0, \frac{\beta}{q(\ell,\ell)})$, then the manager could reduce $b_{\ell\ell}$ and increase w_ℓ to keep low workers' utility constant and relax IC_h, thereby keeping her profits constant. Finally, if both $b_{\ell\ell} < \frac{\beta}{q(\ell,\ell)}$ and $b_{hh} < \frac{\beta}{q(h,h)}$, then the manager obtains at most zero expected profit. As the manager can always obtain strictly positive profits for any matching she implements, say, by setting $b_{hh} = \frac{\beta}{q(h,\ell)}$, $b_{h\ell} = \frac{\beta}{q(\ell,\ell)}$, and $b_{\ell h} = b_{\ell\ell} = w_h = w_\ell = 0$, such bonuses can never be optimal.

We now argue that $b_{hh} = 0$ and $b_{\ell\ell} > \frac{\beta}{q(\ell,\ell)}$ can never be optimal. Observe first that a high masquerading as a low always obtains a higher utility as a low under her wage scheme. As lows obtain strictly positive expected utility from honesty and obedience when $b_{\ell\ell} > \frac{\beta}{q(\ell,\ell)}$, this implies that $w_h > 0$ in order to satisfy IC_h. In addition, any wage scheme in which $w_h > 0$ and $w_\ell > 0$ cannot possibly be optimal; the manager can reduce both wages by a constant and strictly increase her profits. But, if $w_\ell = 0$, then any $w_h > 0$ satisfying IC_h cannot satisfy IC_ℓ. We have arrived at a contradiction.

We proceed to analyze the case in which $b_{\ell\ell} = 0$ and $b_{hh} > \frac{\beta}{q(h,h)}$. First, observe that any wage scheme with $b_{hh} > \frac{\beta}{q(h,\ell)}$ and $b_{\ell\ell} = 0$ cannot be optimal if $\nu(h)$ is sufficiently small. If such a scheme were used, then it must be that w_ℓ is bounded below by a strictly positive value for any $\nu(h)$. Hence, as $\nu(h)$ approaches zero, the profits the manager obtains from highs approaches zero, while the profits the manager obtains from lows remains strictly negative. Second, observe that $b_{hh} < \frac{\beta}{q(h,\ell)}$ and $b_{\ell\ell} = 0$ can never be optimal. If the manager did choose such wage scheme, then she could set $w_h = w_\ell = 0$ and increase b_{hh} without violating any incentive compatibility constraints. This strictly increases the manager's profit provided that $\frac{\beta}{q(h,\ell)} < b_{hh}^{MH}$, which holds if and only if

$$\beta < \frac{q(h,h)q(h,\ell)}{4q(h,h) - 2q(h,\ell)}.$$

If $\frac{q(\ell,\ell)}{q(h,h)}$ is small and $q(h,\ell)$ is sufficiently close to $\frac{1}{2}\left(q(h,h) + q(\ell,\ell)\right)$, however, the upper bound on β in the displayed equation is larger than $q(\ell,\ell)/2$, ensuring that such an increase is profitable.

From the preceding paragraph, there is only one remaining case to consider: $b_{\ell\ell} = 0$ and $b_{hh} = \frac{\beta}{q(h,\ell)}$. We argue now that the manager can strictly increase her profits by increasing both $b_{\ell\ell}$ and b_{hh}. In particular, suppose that $b_{\ell\ell} \geq \frac{\beta}{q(\ell,\ell)}$ and $b_{hh} = \frac{q(\ell,\ell)}{q(h,\ell)} b_{\ell\ell}$ so that both incentive compatibility constraints are satisfied. We show that strictly increasing $b_{\ell\ell}$ above $\frac{\beta}{q(\ell,\ell)}$ strictly increases the manager's expected profits,

$$\nu(\ell)\hat{e}(\ell,\ell,b_{\ell\ell})(1 - 2b_{\ell\ell}) + \nu(h)\hat{e}(h,h,\frac{q(\ell,\ell)}{q(h,\ell)}b_{\ell\ell})q(h,h)\left(1 - 2\frac{q(\ell,\ell)}{q(h,\ell)}b_{\ell\ell}\right),$$

where $\hat{e}(\theta,\theta',b) := \phi(q(\theta,\theta') \cdot b)$. Differentiating with respect to $b_{\ell\ell}$ and evaluating the resulting expression at $b_{\ell\ell} = \frac{\beta}{q(\ell,\ell)}$ yields a number larger than

$$\nu(h)\left[\frac{q^2(h,h)q(\ell,\ell)}{q(h,\ell)}\left(1 - \frac{q(\ell,\ell)}{q(h,\ell)}\right) + 2\beta\frac{q(h,h)q(\ell,\ell)}{q(h,\ell)}\left(1 - \frac{q(\ell,\ell)q(h,h)}{2q(h,\ell)}\right)\right],$$

which is strictly positive.

Appendix B

Without limited liability constraints, a **wage scheme** is a tuple

$$\mathbf{w} := (w_h, w_\ell, b_{hh}, b_{h\ell}, b_{\ell h}, b_{\ell\ell}) \in \mathbb{R}^2 \times \mathbb{R}^4_+,$$

where $w_{\hat{\theta}}$ is the base wage paid to a worker who announces his type as $\hat{\theta}$ and $b_{\hat{\theta}\hat{\theta}'}$ is the bonus he receives if he matches a worker with reported type $\hat{\theta}'$ and his team succeeds. We make two observations. First, notice that base wages may now be strictly negative. Second, it is without loss of generality to restrict bonuses to be non-negative. If a worker's bonus were to be negative, then he would optimally exert zero effort and receive a strictly positive payment if his team's project fails. But then, we could increase this bonus to zero and decrease the worker's base wage without changing the principal's expected wage payments or any worker's incentives.

Profit Maximization and Efficiency

We first point out that a profit-maximizing manager implements efficient allocations if either effort or type is observable.

Full Information (Effort and Type Observable). First, suppose both effort and types are observable, so that contracts can condition on either. Consider a contract consisting of positive assortative matching (given observed types) and a wage scheme that pays each agent his cost of effort, $c(e^{FB}(\theta_i, \theta_j))$, if he exerts the manager's desired effort level, $e^{FB}(\theta_i, \theta_j)$, and zero otherwise. No other contract can do strictly better in terms of profits without violating workers' participation constraints. Hence, profit maximization implies efficiency.

Pure Adverse Selection (Effort Observable and Type Unobservable). Second, suppose effort is observable, but type is unobservable, so that contracts can only condition on effort and reported type. Consider a contract consisting of positive assortative matching (given *reported* types) and a wage scheme that pays each worker i his cost of effort, $c(e^{FB}(\hat{\theta}_i, \hat{\theta}_j))$, if he exerts the manager's desired effort level, $e^{FB}(\hat{\theta}_i, \hat{\theta}_j)$, and zero otherwise, where $\hat{\theta}_i$ is worker i's reported type and $\hat{\theta}_j$ is his teammate's reported type. This scheme implements PAM and efficient effort at minimal cost without violating incentive compatibility or any participation constraints; because the cost of effort does not depend on type, each worker obtains zero utility in any team to which he is assigned, whether or not he reports his type truthfully. Again, profit maximization implies efficiency.

Pure Moral Hazard (Effort Unobservable and Type Observable). Finally, suppose effort is unobservable, but type is observable, so that contracts can condition on observed type, but not effort. Consider a contract consisting of positive assortative matching (given observed types) and a wage scheme in which the manager sells each worker her contribution to team output: each worker in a team with observed types (θ_i, θ_j) receives a negative base wage

$$-\left[e^{FB}(\theta_i, \theta_j) \cdot q(\theta_i, \theta_j) - c(e^{FB}(\theta_i, \theta_j))\right]$$

and a bonus of 1. This scheme implements PAM and efficient effort at minimal cost without violating any participation constraints; each worker receives an expected payoff of zero in his assigned team. Once again, profit maximization implies efficiency.

Binding Constraints

In the absence of limited liability constraints, the manager must ensure that agents have an incentive to participate in the firm. The **interim individual rationality constraint** for a worker of type θ, IR_θ, ensures that each receives an expected utility larger than his outside option of zero:

$$U(\theta, \theta) \geq 0.$$

In light of these additional constraints, the **manager's problem** is to choose a contract $(\mu, \mathbf{w})$ to maximize

$$\overbrace{\sum_{\theta_i} \sum_{\theta_j} \mu(\theta_i, \theta_j) \cdot S(\theta_i, \theta_j, b_{\theta_i \theta_j}, b_{\theta_j \theta_i})}^{\text{Expected Surplus}} - \overbrace{2 \sum_{\theta} \nu(\theta) \cdot U(\theta, \theta)}^{\text{Expected Rent}}$$

subject to IC_h, IC_ℓ, IR_h, and IR_ℓ.

We now identify which constraints bind and when.

Lemma 3 (Binding Constraints)**.**

1. *Any wage scheme that satisfies IC_h and IR_ℓ also satisfies IR_h.*
2. *In any second-best optimal wage scheme, IC_h and IR_ℓ always bind.*
3. *There exists a $\bar{p} < 1$ such that IC_ℓ binds if and only if $p_h^\mu \leq \bar{p}$.*

To show that IR_h is implied by IC_h and IR_ℓ, it suffices to observe that, because highs are more productive than lows, any high masquerading as a low obtains a higher utility than a truth-telling low. Hence, if highs have an incentive to truthfully report their type, i.e. IC_h is satisfied, and lows obtain at least their reservation utility when truthfully reporting their type, i.e. IR_ℓ is satisfied, then highs must also obtain at least their reservation utility when truthfully reporting their type, i.e. IR_h is satisfied. Upon removing IR_h from the manager's optimization problem, we see that both IC_h and IR_ℓ must bind because, if either does not, the manager can reduce the base wages w_h and w_ℓ in a way that satisfies all incentive constraints and strictly reduces rent payments. Notice, this argument does *not* work in the presence of limited liability constraints.

We then use the binding constraints IC_h and IR_ℓ to considerably simplify the manager's problem. The binding constraint IR_ℓ implies $U(\ell, \ell) = 0$, allowing us to eliminate rent payments to lows from the manager's objective function. The binding constraint IC_h allows us to express rent payments to highs as their expected gains relative to lows upon masquerading as a low. In particular, let

$$G(\theta_j, b_i, b_j) := V(h, \theta_j, b_i, b_j) - V(\ell, \theta_j, b_i, b_j)$$

denote the gain a high receives relative to a low when matched with a teammate of type θ_j and bonuses in the team are (b_i, b_j). Then,

$$U(h,h) = \underbrace{U(\ell,\ell)}_{=0} + p_\ell^\mu \cdot G(h, b_{h\ell}, b_{\ell h}) + (1 - p_\ell^\mu) \cdot G(\ell, b_{\ell\ell}, b_{\ell\ell}).$$

Using these observations, we show that the manager's problem is simply to choose bonuses $(b_{hh}, b_{h\ell}, b_{\ell h}, b_{\ell\ell}) \in \mathbb{R}_+^4$ to maximize

$$\sum_{\theta_i} \sum_{\theta_j} \mu(\theta_i, \theta_j) \cdot S(\theta_i, \theta_j, b_{\theta_i\theta_j}, b_{\theta_i\theta_j})$$

$$-2 \cdot \nu(h) \cdot \left[p_\ell^\mu \cdot G(h, b_{\ell h}, b_{h\ell}) + (1 - p_\ell^\mu) \cdot G(\ell, b_{\ell\ell}, b_{\ell\ell})\right]$$

subject to

$$p_\ell^\mu \cdot G(h, b_{\ell h}, b_{h\ell}) + (1 - p_\ell^\mu) \cdot G(\ell, b_{\ell\ell}, b_{\ell\ell}) \leq$$
$$p_h^\mu \cdot G(h, b_{hh}, b_{hh}) + (1 - p_h^\mu) \cdot G(\ell, b_{h\ell}, b_{\ell h}),$$

with base wages w_h and w_ℓ determined by IC_h and IR_ℓ. In this formulation, the single incentive constraint IC_ℓ is expressed as an upper bound on expected rent payments to highs. This bound depends on the gains highs receive relative to lows when *truthfully* reporting their type. To show that this constraint does not bind if matching is sufficiently assortative, i.e. if p_h^μ is above some threshold $\bar{p} < 1$, we observe that bonuses that solve a relaxed problem without IC_ℓ must have $b_{hh} > b_{\ell\ell}$. Hence, $G(h, b_{hh}, b_{hh}) > G(\ell, b_{\ell\ell}, b_{\ell\ell})$ so that IC_ℓ is satisfied whenever p_h^μ is large enough.

Second-Best Effort

We now state our characterization of interior second-best optimal effort.

Lemma 4 (Second-Best Optimal Effort without Limited Liability)**.** *Interior second-best optimal effort levels are characterized by the equations*

$$\begin{aligned}
e(h,h,b_{hh}) &= \phi\left(q(h,h) + D^{\lambda_\ell}_{hh}\right), \\
e(h,\ell,b_{h\ell}) &= \phi\left(q(h,\ell) - \frac{\nu(h)}{\nu(\ell)}\left[\frac{\partial e(h,\ell,b_{h\ell})}{\partial b_i}\right]^{-1}\left[\frac{\partial G(h,b_{\ell h},b_{h\ell})}{\partial b_j}\right] + D^{\lambda_\ell}_{h\ell}\right), \\
e(\ell,h,b_{\ell h}) &= \phi\left(q(h,\ell) - \frac{\nu(h)}{\nu(\ell)}\left[\frac{\partial e(\ell,h,b_{\ell h})}{\partial b_i}\right]^{-1}\left[\frac{\partial G(h,b_{\ell h},b_{h\ell})}{\partial b_i}\right] + D^{\lambda_\ell}_{\ell h}\right), \textit{ and} \\
e(\ell,\ell,b_{\ell\ell}) &= \phi\left(q(\ell,\ell) - \frac{\nu(h)}{\nu(\ell)}\left[\frac{\partial e(\ell,\ell,b_{\ell\ell})}{\partial b_i}\right]^{-1}\left[\frac{\partial G(\ell,b_{\ell\ell},b_{\ell\ell})}{\partial b_i} + \frac{\partial G(\ell,b_{\ell\ell},b_{\ell\ell})}{\partial b_j}\right] + D^{\lambda_\ell}_{\ell\ell}\right),
\end{aligned}$$

where $D^{\lambda_\ell}_{hh}$, $D^{\lambda_\ell}_{\ell h}$, $D^{\lambda_\ell}_{\ell\ell}$, and $D^{\lambda_\ell}_{h\ell}$ are non-zero only if IC_ℓ binds and λ_ℓ, its optimized KKT multiplier, is strictly positive:

$$\begin{aligned}
D^{\lambda_\ell}_{hh} &:= \frac{\lambda_\ell}{\nu(h)}\left[\frac{\partial e(h,h,b_{hh})}{\partial b_{hh}}\right]^{-1}\left[\frac{\partial G(h,b_{hh},b_{hh})}{\partial b_i} + \frac{\partial G(h,b_{hh},b_{hh})}{\partial b_j}\right], \\
D^{\lambda_\ell}_{h\ell} &:= \frac{\lambda_\ell}{2}\left[\frac{\partial e(h,\ell,b_{h\ell})}{\partial b_i}\right]^{-1}\left[\frac{1}{\nu(h)}\cdot\frac{\partial G(\ell,b_{h\ell},b_{\ell h})}{\partial b_i} - \frac{1}{\nu(\ell)}\cdot\frac{\partial G(h,b_{\ell h},b_{h\ell})}{\partial b_j}\right], \\
D^{\lambda_\ell}_{\ell h} &:= \frac{\lambda_\ell}{2}\left[\frac{\partial e(\ell,h,b_{\ell h})}{\partial b_i}\right]^{-1}\left[\frac{1}{\nu(h)}\cdot\frac{\partial G(\ell,b_{h\ell},b_{\ell h})}{\partial b_j} - \frac{1}{\nu(\ell)}\cdot\frac{\partial G(h,b_{\ell h},b_{h\ell})}{\partial b_i}\right], \textit{ and} \\
D^{\lambda_\ell}_{\ell\ell} &:= -\frac{\lambda_\ell}{2\cdot\nu(\ell)}\left[\frac{\partial e(\ell,\ell,b_{\ell\ell})}{\partial b_i}\right]^{-1}\left[\frac{\partial G(\ell,b_{\ell\ell},b_{\ell\ell})}{\partial b_i} + \frac{\partial G(\ell,b_{\ell\ell},b_{\ell\ell})}{\partial b_j}\right].
\end{aligned}$$

Interior effort distortions depend on two forces: (i) the incentive of a high to masquerade as a low (IC_h) and (ii) the incentive of a low to masquerade as a high (IC_ℓ). When the manager implements a sufficiently assortative matching, we know from Lemma 3 that IC_ℓ does *not* bind, so that only (i) is relevant. Inspecting the equations in Lemma 4, we see that, in this case, teams composed of two highs exert efficient effort, i.e. there is "no distortion at the top". On the other hand, effort is generally distorted in mixed teams composed of a high and a low and in teams composed of two lows. The distortions in these teams allow the manager to reduce her expected rent payment; lower levels of implemented effort entail lower bonuses, decreasing the utility highs receive in these teams when they masquerade as lows.

Second-Best Matching

We now turn to the problem of profit-maximizing matching. For any matching μ, the optimal wage scheme $\mathbf{w}(\mu)$ dictates that expected rent payments are given by

$$2 \cdot \nu(h) \cdot \left(p_\ell^\mu \cdot G(h, b_{\ell h}^*, b_{h\ell}^*) + (1 - p_\ell^\mu) \cdot G(\ell, b_{\ell\ell}^*, b_{\ell\ell}^*)\right),$$

where starred bonuses are optimal. We call the difference

$$G(\ell, b_{\ell\ell}^*, b_{\ell\ell}^*) - G(h, b_{\ell h}^*, b_{h\ell}^*)$$

the **degree of disassortative incentives**; as this difference increases, positive assortative matching, i.e. setting $p_h^\mu = 1$ so that $p_\ell^\mu = \frac{\nu(h)}{\nu(\ell)}(1 - p_h^\mu) = 0$, becomes less attractive from the standpoint of minimizing expected rent payments. Taking the derivative of profits with respect to p_h^μ reveals that profits increase in matching assortativity if and only if the degree of supermodularity of second-best surplus outweighs the degree of disassortative incentives:

$$\begin{aligned}&\overbrace{S(h,h,b_{hh}^*,b_{hh}^*) + S(\ell,\ell,b_{\ell\ell}^*,b_{\ell\ell}^*) - 2\cdot S(h,\ell,b_{h\ell}^*,b_{\ell h}^*)}^{\text{Degree of Supermodularity}} \geq \\ &2 \cdot \frac{\nu(h)}{\nu(\ell)} \cdot \underbrace{(G(\ell,b_{\ell\ell}^*,b_{\ell\ell}^*) - G(h,b_{h\ell}^*,b_{\ell h}^*))}_{\text{Degree of Disassortative Incentives}}.\end{aligned}$$

We show that this equation is always satisfied when $\nu(h)$ is sufficiently small to obtain the following Theorem.

Theorem 6 (Optimal Matching Without Limited Liability Constraints)**.** *If $\nu(h)$ is sufficiently small, then PAM is the unique optimal matching.*

Proofs for Appendix B

Proof of Lemma 3

1. IC_h and IR_ℓ imply IR_h by the following inequalities:

$$\begin{aligned} U(h,h) &\geq p_\ell^\mu \cdot V(h,h,b_{\ell h},b_{h\ell}) + (1-p_\ell^\mu) \cdot V(h,\ell,b_{\ell\ell},b_{\ell\ell}) + w_\ell \\ &\geq p_\ell^\mu \cdot V(\ell,h,b_{\ell h},b_{h\ell}) + (1-p_\ell^\mu) \cdot V(\ell,\ell,b_{\ell\ell},b_{\ell\ell}) + w_\ell \\ &\geq 0, \end{aligned}$$

   where the first inequality is IC_h, the second holds because $\frac{\partial V(\theta_i,\theta_j,b_i,b_j)}{\partial \theta_i} \geq 0$, and the third is IR_ℓ.

2. Upon removing IR_h, we now show that, in any solution to the second-best problem, both IC_h and IR_ℓ bind. Denote λ_h as the KKT multiplier corresponding to IC_h, λ_ℓ as the multiplier corresponding to IC_ℓ, and γ_ℓ as the multiplier corresponding to IR_ℓ. The first-order conditions with respect to w_h and w_ℓ are

$$-2 \cdot \nu(h) + \lambda_h - \lambda_\ell = 0$$

   and

$$-2 \cdot \nu(\ell) - (\lambda_h - \lambda_\ell) + \gamma_\ell = 0.$$

   Combining, we obtain

$$\gamma_\ell = 2 > 0$$

and,

$$\lambda_h = 2 \cdot \nu(h) + \lambda_\ell > 0.$$

By complementary slackness, IC_h and IR_ℓ must therefore bind at any critical point.

Simplifying the problem. Using the observation that IC_h and IR_ℓ always bind, we may considerably simplify the manager's problem. Let

$$G(\theta_j, b_i, b_j) := V(h, \theta_j, b_i, b_j) - V(\ell, \theta_j, b_i, b_j)$$

denote the gain a high receives relative to a low when matched with a teammate of type θ_j and bonuses in the team are (b_i, b_j). As IR_ℓ binds, we know that any solution to the second-best problem has $U(\ell, \ell) = 0$. Using the binding constraint IC_h, we thus have,

$$U(h,h) = \underbrace{U(\ell,\ell)}_{=0} + p_\ell^\mu \cdot G(h, b_{h\ell}, b_{\ell h}) + (1 - p_\ell^\mu) \cdot G(\ell, b_{\ell\ell}, b_{\ell\ell}).$$

Re-arranging, this equation implies that

$$\begin{aligned} w_h = \; & p_\ell^\mu \cdot G(h, b_{\ell h}, b_{h\ell}) + (1 - p_\ell^\mu) \cdot G(\ell, b_{\ell\ell}, b_{\ell\ell}) \\ & - \left[p_h^\mu \cdot V(h, h, b_{hh}, b_{hh}) + (1 - p_h^\mu) \cdot V(h, \ell, b_{h\ell}, b_{\ell h}) \right]. \end{aligned}$$

Finally, we may simplify IC_ℓ by substituting in for $U(\ell, \ell) = 0$ and w_h:

$$\begin{gathered} p_\ell^\mu \cdot G(h, b_{\ell h}, b_{h\ell}) + (1 - p_\ell^\mu) \cdot G(\ell, b_{\ell\ell}, b_{\ell\ell}) \leq \\ p_h^\mu \cdot G(h, b_{hh}, b_{hh}) + (1 - p_h^\mu) \cdot G(\ell, b_{h\ell}, b_{\ell h}), \end{gathered}$$

We are thus left with the problem of choosing bonuses $(b_{hh}, b_{h\ell}, b_{\ell h}, b_{\ell\ell}) \in \mathbb{R}_+^4$ to maximize

$$\sum_{\theta_i}\sum_{\theta_j}\mu(\theta_i,\theta_j)\cdot S(\theta_i,\theta_j,b_{\theta_i\theta_j},b_{\theta_i\theta_j})$$

$$-2\cdot\nu(h)\cdot\left(p_\ell^\mu\cdot G(h,b_{\ell h},b_{h\ell})+(1-p_\ell^\mu)\cdot G(\ell,b_{\ell\ell},b_{\ell\ell})\right)$$

subject to IC_ℓ.

3. We first consider a relaxed problem without IC_ℓ and non-negativity constraints on b_{hh}. In this problem, the first-order condition with respect to b_{hh} is

$$2\cdot\mu(h,h)\cdot\frac{\partial e(h,h,b_{hh})}{\partial b_i}\left[e(h,h,b_{hh})\cdot q(h,h)-c'(e(h,h,b_{hh}))\right]=0.$$

Hence,

$$e(h,h,b_{hh})=\phi\left(q(h,h)\right)\iff b_{hh}=1.$$

Denote by $\gamma_{\ell\ell}$ the KKT multiplier corresponding to the non-negativity constraint on $b_{\ell\ell}$. The first-order condition with respect to $b_{\ell\ell}$ is

$$2\cdot\mu(\ell,\ell)\cdot\frac{e(\ell,\ell,b_{\ell\ell})}{\partial b_i}\left[e(\ell,\ell,b_{\ell\ell})\cdot q(\ell,\ell)-c'(e(\ell,\ell,b_{\ell\ell}))\right]$$
$$-2\cdot\nu(h)\cdot(1-p_\ell^\mu)\left[\frac{\partial G(\ell,b_{\ell\ell},b_{\ell\ell})}{\partial b_i}+\frac{\partial G(\ell,b_{\ell\ell},b_{\ell\ell})}{\partial b_j}\right]+\gamma_{\ell\ell}=0.$$

As $\mu(\ell,\ell)=\nu(\ell)\cdot(1-p_\ell^\mu)$, we may re-arrange the equation to obtain,

$$e(\ell,\ell,b_{\ell\ell})=\phi\left(q(\ell,\ell)-\frac{\nu(h)}{\nu(\ell)}\cdot\left[\frac{e(\ell,\ell,b_{\ell\ell})}{\partial b_i}\right]^{-1}\left[\frac{\partial G(\ell,b_{\ell\ell},b_{\ell\ell})}{\partial b_i}+\frac{\partial G(\ell,b_{\ell\ell},b_{\ell\ell})}{\partial b_j}\right]+\frac{\gamma_{\ell\ell}}{\alpha}\right),$$

where $\alpha>0$. If $b_{\ell\ell}>0$, then $\gamma_{\ell\ell}=0$ and the first-order condition implies that $b_{\ell\ell}<1$. Hence, in any solution, $b_{\ell\ell}<1$.

We now show that the solution to the relaxed problem satisfies IC_ℓ whenever p_h^μ is close to one. As $b_{\ell\ell} < b_{hh} = 1$,

$$\begin{aligned} G(h, b_{hh}, b_{hh}) &= e(h, h, b_{hh}) \cdot b_{hh} \cdot (q(h, h) - q(l, h)) \\ &> e(h, \ell, b_{\ell\ell}) \cdot b_{\ell\ell} \cdot (q(h, \ell) - q(\ell, \ell)) \\ &= G(\ell, b_{\ell\ell}, b_{\ell\ell}), \end{aligned}$$

where the inequality follows because $\frac{\partial e(\theta_i, \theta_j, b)}{\partial \theta_j} > 0$, $\frac{\partial e(\theta_i, \theta_j, b_i)}{\partial b_i} > 0$, and $q(h, h) - q(h, \ell) > q(h, \ell) - q(\ell, \ell) > 0$ by strict supermodularity of q. Now, if p_h^μ is close to one, then

$$p_\ell^\mu = \frac{\nu(h)}{\nu(\ell)}(1 - p_h^\mu)$$

is close to zero. Hence, IC_ℓ holds strictly above some threshold $\bar{p} < 1$.

Proof of Lemma 4 As we are characterizing interior solutions, we ignore non-negativity constraints on bonuses. Denote by λ_ℓ the KKT multiplier on IC_ℓ. We identify the equations in the Theorem using the first-order conditions.

The first-order condition with respect to b_{hh} is

$$\begin{aligned} 2 \cdot \mu(h, h) \cdot \frac{\partial e(h, h, b_{hh})}{\partial b_i} &\left[e(h, h, b_{hh}) \cdot q(h, h) - c'(e(h, h, b_{hh}))\right] \\ &+ \lambda_\ell \cdot p_h^\mu \left[\frac{\partial G(h, b_{hh}, b_{hh})}{\partial b_i} + \frac{\partial G(h, b_{hh}, b_{hh})}{\partial b_j}\right] = 0 \end{aligned}$$

As $\mu(h, h) = \nu(h) \cdot p_h^\mu$, we obtain the following expression upon re-arranging terms:

$$e(h, h, b_{hh}) = \phi\left(q(h, h) + D_{hh}^{\lambda_\ell}\right),$$

where

$$D_{hh}^{\lambda_\ell} := \frac{\lambda_\ell}{\nu(h)} \left[\frac{\partial e(h, h, b_{hh})}{\partial b_{hh}}\right]^{-1} \left[\frac{\partial G(h, b_{hh}, b_{hh})}{\partial b_i} + \frac{\partial G(h, b_{hh}, b_{hh})}{\partial b_j}\right]$$

The first-order condition with respect to $b_{\ell\ell}$ is

$$2 \cdot \mu(\ell,\ell) \cdot \frac{\partial e(\ell,\ell,b_{\ell\ell})}{\partial b_i} \left[e(\ell,\ell,b_{\ell\ell}) \cdot q(\ell,\ell) - c'(e(\ell,\ell,b_{\ell\ell})) \right] \\ - (2 \cdot \nu(h) + \lambda_\ell) \cdot (1 - p^\mu_\ell) \left[\frac{\partial G(\ell, b_{\ell\ell}, b_{\ell\ell})}{\partial b_i} + \frac{\partial G(\ell, b_{\ell\ell}, b_{\ell\ell})}{\partial b_j} \right] = 0.$$

As $\mu(\ell,\ell) = \nu(\ell) \cdot (1 - p^\mu_\ell)$, we obtain the following expression upon re-arranging terms:

$$e(\ell,\ell,b_{\ell\ell}) = \phi \left(q(\ell,\ell) - \frac{\nu(h)}{\nu(\ell)} \left[\frac{\partial e(\ell,\ell,b_{\ell\ell})}{\partial b_i} \right]^{-1} \left[\frac{\partial G(\ell, b_{\ell\ell}, b_{\ell\ell})}{\partial b_i} + \frac{\partial G(\ell, b_{\ell\ell}, b_{\ell\ell})}{\partial b_j} \right] + D^{\lambda_\ell}_{\ell\ell} \right),$$

where

$$D^{\lambda_\ell}_{\ell\ell} := -\frac{\lambda_\ell}{2 \cdot \nu(\ell)} \left[\frac{\partial e(\ell,\ell,b_{\ell\ell})}{\partial b_i} \right]^{-1} \left[\frac{\partial G(\ell, b_{\ell\ell}, b_{\ell\ell})}{\partial b_i} + \frac{\partial G(\ell, b_{\ell\ell}, b_{\ell\ell})}{\partial b_j} \right]$$

The first-order condition with respect to $b_{h\ell}$ is

$$\mu(h,\ell) \frac{\partial e(h,\ell,b_{h\ell})}{\partial b_i} \left[q(h,\ell) - c'(e(h,\ell,b_{h\ell})) \right] \\ - (2 \cdot \nu(h) + \lambda_\ell) \cdot p^\mu_\ell \left[\frac{\partial G(h, b_{\ell h}, b_{h\ell})}{\partial b_j} \right] + \lambda_\ell \cdot (1 - p^\mu_h) \left[\frac{\partial G(\ell, b_{h\ell}, b_{\ell h})}{\partial b_i} \right] = 0.$$

As $\mu(h,\ell) = 2 \cdot \nu(\ell) \cdot p^\mu_\ell = 2 \cdot \nu(h) \cdot (1 - p^\mu_h)$, we obtain the following expression upon re-arranging terms:

$$e(h,\ell,b_{h\ell}) = \phi \left(q(h,\ell) - \frac{\nu(h)}{\nu(\ell)} \left[\frac{\partial e(h,\ell,b_{h\ell})}{\partial b_i} \right]^{-1} \left[\frac{\partial G(h, b_{\ell h}, b_{h\ell})}{\partial b_j} \right] + D^{\lambda_\ell}_{h\ell} \right),$$

where

$$D^{\lambda_\ell}_{h\ell} := \frac{\lambda_\ell}{2} \left[\frac{\partial e(h,\ell,b_{h\ell})}{\partial b_i} \right]^{-1} \left[\frac{1}{\nu(h)} \cdot \frac{\partial G(\ell, b_{h\ell}, b_{\ell h})}{\partial b_i} - \frac{1}{\nu(\ell)} \cdot \frac{\partial G(h, b_{\ell h}, b_{h\ell})}{\partial b_j} \right].$$

Finally, a similar manipulation of the first-order condition with respect to $b_{\ell h}$ yields

$$e(\ell, h, b_{\ell h}) = \phi\left(q(h,\ell) - \frac{\nu(h)}{\nu(\ell)}\left[\frac{\partial e(\ell, h, b_{\ell h})}{\partial b_i}\right]^{-1}\left[\frac{\partial G(h, b_{\ell h}, b_{h\ell})}{\partial b_i}\right] + D^{\lambda_\ell}_{h\ell}\right),$$

where

$$D^{\lambda_\ell}_{\ell h} := \frac{\lambda_\ell}{2}\left[\frac{\partial e(\ell, h, b_{\ell h})}{\partial b_i}\right]^{-1}\left[\frac{1}{\nu(h)} \cdot \frac{\partial G(\ell, b_{h\ell}, b_{\ell h})}{\partial b_j} - \frac{1}{\nu(\ell)} \cdot \frac{\partial G(h, b_{\ell h}, b_{h\ell})}{\partial b_i}\right].$$

Proof of Theorem 6 We now identify closed-form expressions for bonuses using the following calculations:

$$\begin{aligned}
\phi(e) &= \frac{e}{2\alpha} - \frac{\beta}{2\alpha}, \\
e(\theta_i, \theta_j, b_i) &= \frac{q(\theta_i, \theta_j) \cdot b_i}{2\alpha} - \frac{\beta}{2\alpha}, \\
\frac{\partial e(\theta_i, \theta_j, b_i)}{\partial b_i} &= \frac{q(\theta_i, \theta_j)}{2\alpha}, \\
\frac{\partial V(\theta_i, \theta_j, b_i, b_j)}{\partial b_i} &= \frac{q^2(\theta_i, \theta_j)(b_i + b_j)}{2\alpha} - q(\theta_i, \theta_j)\frac{2\beta}{2\alpha}, \\
\frac{\partial G(\theta_j, b_i, b_j)}{\partial b_i} &= \frac{(q^2(h, \theta_j) - q^2(\ell, \theta_j))(b_i + b_j)}{2\alpha} - \left(q(h, \theta_j) - q(\ell, \theta_j)\right)\frac{2\beta}{2\alpha}, \\
\frac{\partial V(\theta_i, \theta_j, b_i, b_j)}{\partial b_j} &= \frac{q^2(\theta_i, \theta_j) \cdot b_i}{2\alpha}, \text{ and} \\
\frac{\partial G(\theta_j, b_i, b_j)}{\partial b_j} &= \frac{(q^2(h, \theta_j) - q^2(\ell, \theta_j)) \cdot b_i}{2\alpha}.
\end{aligned}$$

Using the formulas in the statement of Theorem 4, we obtain the following solutions for b_{hh} and $b_{\ell\ell}$ when IC_ℓ does not bind:

$$\begin{aligned}
b^*_{hh} &= 1 \quad \text{and} \\
b^*_{\ell\ell} &= \frac{\nu(\ell)q^2(\ell,\ell) + 2\beta\nu(h)\left(q(h,\ell) - q(\ell,\ell)\right)}{\nu(\ell)q^2(\ell,\ell) + 3\nu(h)\left(q^2(h,\ell) - q^2(\ell,\ell)\right)}.
\end{aligned}$$

As $\nu(h)$ approaches zero, $b^*_{\ell\ell}$ approaches one so that effort in teams composed of two highs and two lows approach their efficient values. Hence, the degree of supermodularity

$$S(h,h,b^*_{hh},b^*_{hh}) + S(\ell,\ell,b^*_{\ell\ell},b^*_{\ell\ell}) - 2 \cdot S(h,\ell,b^*_{h\ell},b^*_{\ell h})$$

approaches a strictly positive value for any values of $b^*_{h\ell}$ and $b^*_{\ell h}$ (interior or otherwise). On the other hand, the degree of disassortative incentives

$$2 \cdot \frac{\nu(h)}{\nu(\ell)}(G(\ell,b^*_{\ell\ell},b^*_{\ell\ell}) - G(h,b^*_{h\ell},b^*_{\ell h}))$$

approaches zero because $b^*_{h\ell}$ and $b^*_{\ell h}$ are bounded above uniformly over $\nu(h)$. As IC_ℓ does not bind when the manager implements PAM (part 3 of Lemma 3), and because interior values of b_{hh} and $b_{\ell\ell}$ are optimal, the result follows.

References

Adhvaryu, A., Bassi, V., Nyshadham, A., and Tamayo, J. "No Line Left Behind: Assortative Matching Inside the Firm." *SSRN Electronic Journal*, (2021).

Aghion, P. and Tirole, J. "Formal and Real Authority in Organizations." *Journal of Political Economy*, Vol. 105 (1997), pp. 1–29.

Becker, G. "A Theory of Marriage: Part I." *Journal of Political Economy*, Vol. 81 (1973), pp. 813–846.

Chassang, S. and Zehnder, C. "Secure Survey Design in Organizations: Theory and Experiments." Working Paper 25918, National Bureau of Economic Research, 2019.

Che, Y.K. and Yoo, S.W. "Optimal Incentives for Teams." *American Economic Review*, Vol. 91 (2001), pp. 525–541.

Damiano, E. and Li, H. "Price Discrimination and Efficient Matching." *Economic Theory*, Vol. 30 (2007), pp. 243–263.

Dessein, W. "Authority and Communication in Organizations." *Review of Economic Studies*, Vol. 69 (2002), pp. 811–838.

Eeckhout, J. and Kircher, P. "Sorting and Decentralized Price Competition." *Econometrica*, Vol. 78 (2010), pp. 539–574.

Franco, A., Mitchell, M., and Vereshchagina, G. "Incentives and the Structure of Teams." *Journal of Economic Theory*, Vol. 146 (2011), pp. 2307 – 2332.

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

Gomes, R. and Pavan, A. "Many-to-Many Matching and Price Discrimination." *Theoretical Economics*, Vol. 11 (2016), pp. 1005–1052.

Holmström, B. "Moral Hazard in Teams." *Bell Journal of Economics*, Vol. 13 (1982), pp. 324–340.

Innes, R.D. "Limited Liability and Incentive Contracting With Ex-ante Action Choices." *Journal of Economic Theory*, Vol. 52 (1990), pp. 45 – 67.

Johnson, T. "Matching Through Position Auctions." *Journal of Economic Theory*, Vol. 148 (2013), pp. 1700 – 1713.

Kambhampati, A. and Segura-Rodriguez, C. "Optimal Team Formation Under Asymmetric Information." PIER Working Paper Archive 20-018, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania, 2021.

Kambhampati, A., Segura-Rodriguez, C., and Shao, P. "Matching to Produce Information: A Model of Self-Organized Research Teams." *SSRN Electronic Journal*, (2021).

Kaya, A. and Vereshchagina, G. "Partnerships Versus Corporations: Moral Hazard, Sorting, and Ownership Structure." *American Economic Review*, Vol. 104 (2014), pp. 291–307.

Lazear, E. and Shaw, K. "Personnel Economics: The Economist's View of Human Resources." *Journal of Economic Perspectives*, Vol. 21 (2007), pp. 91–114.

Legros, P. and Newman, A. "Beauty is a Beast, Frog is a Prince: Assortative Matching With Nontransferabilities." *Econometrica*, Vol. 75 (2007), pp. 1073–1102.

Marschak, J. and Radner, R. *Economic Theory of Teams.* New Haven, Conn.: Yale University Press., 1972.

McAfee, P. and McMillan, J. "Optimal Contracts for Teams." *International Economic Review*, Vol. 32 (1991), pp. 561–577.

Mookherjee, D. "Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective." *Journal of Economic Literature*, Vol. 44 (2006), pp. 367–390.

Poitevin, M. "Can the Theory of Incentives Explain Decentralization?" *Canadian Journal of Economics / Revue canadienne d'Economique*, Vol. 33 (2000), pp. 878–906.

Prat, A. "Should a Team Be Homogeneous?" *European Economic Review*, Vol. 46 (2002), pp. 1187–1207.

Sappington, D. "Limited Liability Contracts Between Principal and Agent." *Journal of Economic Theory*, Vol. 29 (1983), pp. 1 – 21.

Sappington, D. "Incentive Contracting With Asymmetric and Imperfect Precontractual Knowledge." *Journal of Economic Theory*, Vol. 34 (1984), pp. 52 – 70.

Serfes, K. "Risk Sharing vs. Incentives: Contract Design Under Two-Sided Heterogeneity." *Economics Letters*, Vol. 88 (2005), pp. 343 – 349.

Serfes, K. "Endogenous Matching in a Market With Heterogeneous Principals and Agents." *International Journal of Game Theory*, Vol. 36 (2007), pp. 587–619.

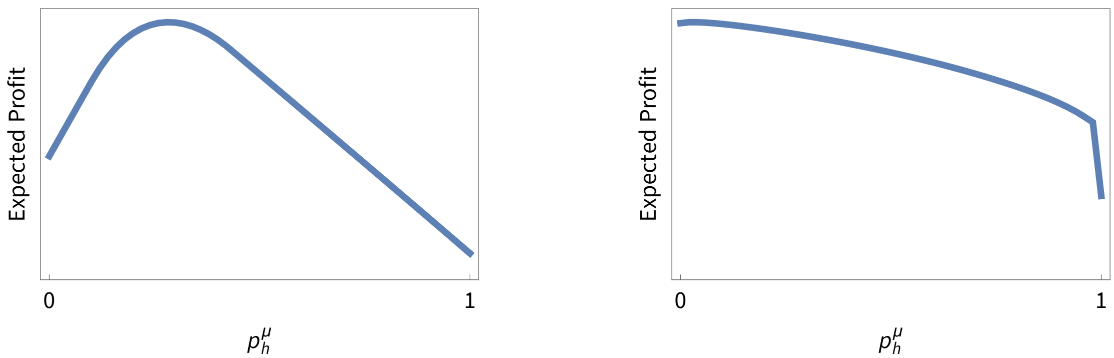
Shimer, R. and Smith, L. "Assortative Matching and Search." *Econometrica*, Vol. 68 (2000), pp. 343–369.

Smith, L. "The Marriage Model With Search Frictions." *Journal of Political Economy*, Vol. 114 (2006), pp. 1124–1144.

Utgoff, N. "Implementation of Assortative Matching Under Incomplete Information." *Journal of Economic Theory*, Vol. 188 (2020).

Wright, D. "The Risk and Incentives Trade-off in the Presence of Heterogeneous Managers." *Journal of Economics*, Vol. 83 (2004), pp. 209–223.

Figures



(a) $\nu(h) = \frac{1}{10}$.

(b) $\nu(h) = \frac{1}{100}$.

Figure 1: *Numerical optimization with parameters* $q(\ell, \ell) = \frac{1}{10}$, $q(h, h) = \frac{9}{10}$, $q(h, \ell) = \frac{49}{100}\left(q(h, h) + q(\ell, \ell)\right)$, $\alpha = q(h, h)$, *and* $\beta = \frac{1}{100} q(\ell, \ell)$. *See our personal website*