Matching Information

Hector Chade*and Jan Eeckhout†

February, 2012

Abstract

The role of information aggregation within firms has long been recognized. We analyze the optimal allocation of differentially informed agents to different firms when those firms are in competition. Will the well informed match with those who are well informed or will they mix with the less informed? This is important because it provides a rationale for the observed worker composition of firms. We find that in equilibrium, the allocation consists of maximally diversifying the work force within the firm, while minimizing the informational difference between firms. Configuring diversely informed teams is the most informative. The value of information is submodular in types, and this allocation is a generalization of negative assortative matching (NAM) in one-to-one matching to a multi-agent team setting, *team*-NAM. We analyze the allocation in the presence of endogenous firm size, heterogeneity in productivity, and under correlation of the agents' information.

Keywords. Maximal Diversification. Matching. Information Aggregation. Sorting. Correlation.

*Arizona State University, `hector.chade@asu.edu`
†University College London and UPF, `j.eeckhout@ucl.ac.uk`.

1 Introduction

One of the major roles of competitive markets is the aggregation of decentralized information. The price system aggregates decentralized information upon which a multitude of agents act to allocate resources in such a way that a planner can impossibly collect and process. Yet, within a market economy, firms and organizations exist. Several common rationales for the existence of firms have been proposed – transaction costs, economies of scale, incentives related to moral hazard, adverse selection and monitoring –, each of which contributes to our understanding of the existence of firms. One of the explanations proposed by Marshak and Radner (1972) is the aggregation of decentralized information within the firm itself that does not get aggregated in the price system. The CEO of the company cannot obtain the most detailed information about the activity and logistics of all its employees, nor can she execute those different activities herself. Instead, she relies on the information and judgement of her employees. As Warren Buffet puts it: "We don't get paid for activity, just for being right."[1]

Information aggregation within the firm or a team *in isolation* has extensively been studied following Marschak and Radner (amongst many others, Van Zandt 1998, Garicano 2000, Prat 2002). In this paper, we address the question of the optimal composition of the work force when different teams consisting of multiple agents simultaneously compete for those decision makers in the labor market. We thus analyze teams *in competition* for those decision makers, who differ in their ability to process information and therefore their predictive precision. Our challenge is to investigate how does market competition shape organizations. Will more informed agents, as measured in Blackwell's sense, be paired with other more informed agents to form superstar teams? Or will they be matched with less informed ones and will the information capabilities be diversified? The key insight here is the role that prices, or wages, play in the allocation process. Even if an agent generates superior value of information, her opportunity cost is determined by the assignment to an alternative team, and thus her market price.

We find that in the presence of competition, there is maximal diversification of information *within* teams. At the same time, competitive pressures minimize differences in the informational composition of agents *between* teams. In the baseline model, the composition of an individual firm exactly mirrors that of the economy at large, in the sense that the firm distribution is the same as the economy-wide distribution. The interpretation of this result is that in organizations as well as in collaborations, we typically tend to see a mixture of experienced and non-experienced individuals. This is true for many environments, from the firms as a whole, to the composition of management consulting teams, to members of a research lab or co-authors on a research project. The ability to evaluate uncertain situations differs between junior and senior members. It turns out that in equilibrium and given market prices, it is optimal to diversify the composition of informedness. This diversification is not unlike that of an optimal portfolio of financial assets that maximizes expected return by mirroring the market composition.

[1] 1998 Berkshire Hathaway Annual Meeting.

Our set up and findings also contribute to the literature on matching without search frictions. One can think of the maximal diversification result as a generalization of the well-known Negative Assortative Matching (NAM) in one-to-one matching (Becker, 1973). However, the notion of Negative Assortative Matching does not readily generalize to the case of many-to-many matching in large teams. First, we show that when the assignment is integer, finding the allocation is extremely difficult, in the sense that the problem is akin to the well-known 3-partition problem in discrete optimization, which is NP-complete. Each alternative allocation must be verified against the candidate equilibrium allocation, and the number of such verifications grows exponentially in the number of agents. This is the opposite of the simplicity of finding the equilibrium allocation under one-to-one matching when there is NAM. Nonetheless, we can show that under integer assignment, the optimal team structure is a partition that majorizes the maximal diversification vector. We refer to this notion of negative sorting for multiple agents as *team*-NAM.

Second, we introduce fractional assignment. Often team members spend a small fraction of their time on a series of different tasks thus belonging to different teams, or for researchers to work on different projects in parallel. This is especially true for senior agents. Under fractional assignment, the problem becomes extremely tractable and the notion of *team*-NAM is now much more transparent. It involves minimal differentiation of the informational composition between firms and maximal diversification within firms. Like the discrete assignment version of the model, the solution involves minimizing the difference between the total informativeness of the firms. But in sharp contrast, due to the divisibility the solution is now simple and intuitive because the informativeness can exactly be equated across firms and the optimal solution is the interior solution that satisfies a set of first order conditions.

As in the case of standard NAM, the driving force is the submodularity of the value of information. Moreover, we derive our information aggregation result as in Marshak-Randner in a specific context under normality and independence of the signals and a quadratic loss function. While this setup is restrictive, it allows us to derive a simple and closed form solution for the value of information. That value is submodular in the precision of the agents' signals. Moreover, it is concave in the sum of the agents' precisions. The concavity implies that it is optimal to equalize the total sum of precisions across different firms, hence the minimal differentiation between firms. This can be achieved by mimicking the economy-wide distribution, hence the maximal diversification within the firm.

We derived the baseline results for an exogenously given firm size and for ex ante identical firms. We extend the setup to allow for endogenous firm size in the presence of free entry and find that the size of firms increases in the cost of entry. Marginal product of agents decreases thus decreasing wages, and increasing gross profits, sufficiently to cover the entry cost. When firms are heterogeneous, the composition across firms now differs. More productive firms have an overall more informed set of agents. Thus there is Positive Assortative Matching between the informedness of the workers and the type of the firm. Still, within the firm there is maximal diversification between agents, *team*-NAM. This provides insights into the determinants of the technology and is consistent with observed firm characteristics. Firms differ in productivity and size, yet the distribution of skills within firms, whether they are large or small, highly productive or not, is diverse.

Finally, we investigate the robustness of the information aggregation technology by allowing for correlation in the signals. Agents base their prediction about the unknown state of nature on information gathered. Management consultants for example carefully study the operations and managerial decision making processes. If consultants access common sources, their signals will be correlated, even if senior consultants process information more efficiently and generate a prediction of the unknown state of nature with higher precision than the junior colleague. Consistent with what we find, the results obtained for the case of no correlation are robust to the introduction of some small correlation. There is still *team*-NAM that involves maximal diversification within the firm and minimal difference between the firms.

Quite surprisingly however, we show that for any setup, there always exist values of the correlation coefficient for which there is Positive Assortative Matching (PAM). In particular, for values of the correlation coefficient approaching perfect positive correlation, the value of information is supermodular, and the equilibrium allocation matches likes with likes: high precision types are in teams with other high precision types and the low precision types are together in other teams. This is the polar opposite of *team*-NAM because the allocation minimizes the diversity within the team and maximizes the differences between teams. To our knowledge, identifying the role of correlation both in the information aggregation and in the matching process is novel in the literature.

Literature. Our model builds on Marschak-Radner's (1972) theory of teams. For teams in isolation, there has been extensive research analyzing the role of the optimal decision function for a given information structure. Information acquisition within the firm in particular has received extensive attention.

The role of competition the organizational form is also analyzed in Alonso, Dessein and Matouschek (2011) in a firm that aggregates information. In contrast to our matching approach to analyze the optimal competitive hiring of informed agents, they study the role of competition in the output market and how it shapes the organization of the firm.

In the matching literature of the labor market, partnerships are often formed one-to-one. A notable exceptions are Kelso and Crawford (1982) and for an application to mechanism design, Hatfield and Milgrom (2005). They analyze multi-agent matching problems and the conditions for existence of a core allocation which is guaranteed under gross substitutes. In the literature on search, several models analyze multi-worker firms, but typically without peer effects. Examples include Shi (2001), Shimer (2005), Kaas and Kircher (2011) and Hawkins (2011). And even if directed search is one-to-one ex post, ex ante there are queues making the matching problem effectively a multi-agent matching problem. for example), because several agents apply for one job. But with the exception of Albrecht, Anderson and Vroman (2010) and Guler, Guvenen, and Violante (2011) who consider matching with search frictions where decisions are made by committees or households, akin to teams.

Observe that while our approach is information theoretic, there is no incomplete information. All agents have the same information sets. For recent work on matching in the presence of incomplete information see Hoppe, Moldovanu, and Ozdenoren (2011) and Hoppe, Moldovanu and Sela (2009).

Finally there is a budding literature on peer group effects within multi-agent firms. Board (2009) analyzes peer group effects in a monopolistic firm and in duopoly, Damiano, Hao, and Suen (2010) study a duopolist setting where firms compete for workers who have externalities on their peers.

2 The Model

There is a finite set $\mathcal{I}$ of agents, and N firms consisting of k agents, where with $I = |\mathcal{I}| = kN$. There is a set of types $\mathcal{X}$ and each agent is assigned a type $x : \mathcal{I} \to \mathcal{X}$, where $x(i) \equiv x_i$, $i = 1, 2, ..., kN$ (notice that two agents i and i' in $\mathcal{X}$ can have the same type, i.e., $x_i = x_{i'}$). Without loss of generality, we shall assume that $x_1 \leq x_2 \leq \cdots \leq x_{kN}$. Agents have linear preferences over consumption, and as a result utility is transferable.

There are two stages. In the first stage *team formation*, N teams each of size k are formed based on the agents' precision of their signal. In the second stage of *information aggregation*, and once a team is formed, the agents collect decentralized information and aggregate it at the team level.

We start by describing the second stage, that of information aggregation. The team faces an uncertain environment, denoted by the state s. This represents for example uncertainty about the firm's demand, about the aggregate economic environment in which the firm operates, or about the outcome of investment in technology. We assume that the prior belief about the state s in each team is given by a density $h(\cdot)$ that is normal with mean μ and precision τ, i.e., $s \sim N(\mu, \tau^{-1})$.[2] The firm hires agents who bring their expertise in estimating the unknown state to the team. An agent spends time in the firm gathering information and analyzing the economic and business environment. Her type x indicates how precise she can estimate the true but unobservable state s. Formally, an agent of type x_i observes a signal σ_i drawn from $f(\cdot|s, x_i)$ that is normal with mean s and precision x_i, i.e., $\sigma_i \sim N(s, x_i^{-1})$. In other words, an agent's type is the informativeness of her signal: better types are endowed with more informative signals (in Blackwell's sense). Different partners draw signals that are conditionally independent.

The objective of the team is to use all the team members' individual information and take a joint action to maximizes the team's profit. First, the team observes the realization of the signals of all k agents in the team. Then the team chooses a single action a contingent on the signal realization in order to maximize the expected value of $\pi - (a - s)^2$, where $\pi > 0$ is a constant representing a maximal level of profits that is reduced by the error in estimating the unknown state. A team whose members have types given by $(x_1, ..., x_k)$ (suitably renumbered so that x_i is the type of the i-th partner, etc.),

[2]Recall that precision is the inverse of the variance.

chooses a measurable function $a : \mathbb{R}^k \to \mathbb{R}$ to maximize

$$V(x_1, ..., x_k) = \max_{a(\cdot)} \quad \pi - \mathbb{E}_{s,x_1,...,x_k} \left[a(\sigma_1, ..., \sigma_k) - s\right]^2 \tag{1}$$

where $\mathbb{E}_{s,x_1,...,x_k} \left[a(\sigma_1, ..., \sigma_k) - s\right]^2 = \int \cdots \int \left[a(\sigma_1, ..., \sigma_k) - s\right]^2 \left(\prod_{i=1}^k f(\sigma_i|s, x_i)\right) h(s) \prod_{i=1}^k d\sigma_i ds$, and $V(x_1, ..., x_k)$ is the maximum expected payoff of the team.

Conditional on knowing the team composition $(x_1, ..., x_k)$, we know the optimal expected value generated by its team members and denoted by $V(x_1, ..., x_k)$. Before any information is generated, there is competition by the all the teams for the different types of agents. In what follows, we will analyze the planner's problem as it coincides with the competitive allocation. The optimal matching problem consists in finding a partition $\mathcal{P}$ of $\mathcal{X}$ into the N teams, each of size k agents. The partition maximizes the sum of the expected payoffs of the teams. Formally, let $P(\mathcal{X})$ be the set of all those partitions. The problem is

$$\max_{\mathcal{P} \in P(\mathcal{X})} \sum_{n \in \mathcal{P}} V(x_1^n, ..., x_k^n). \tag{2}$$

Since utility is fully transferable, suitable transfers can be found that will sustain the optimal matching it as a stable team structure (core allocation) or as the outcome of a competitive equilibrium.

3 Solution and Results

3.1 Solving the Team Problem

Despite the complex appearance of the information aggregation problem (1), it has a remarkably simple closed form solution. This is due to the combination of the assumptions on our setup: normally distributed signals and a quadratic loss function. Recall Bayesian updating with normal prior and observations implies that the posterior distribution is also normal with mean given by a weighted average of the prior mean and the observation, and precision equal to the sum of the precisions of the prior and the signal. This extrapolates straightforwardly to the case with multiple observations of conditionally independent signals. That is, after observing $(\sigma_1, ..., \sigma_k)$, the posterior density function $h(\cdot|\sigma_1, ..., \sigma_k; x_1, ..., x_k)$ is normal with mean and precision:

$$s(\sigma_1, ..., \sigma_k; x_1, ..., x_k) \sim \mathcal{N}\left(\frac{\mu\tau + \sum_{i=1}^k \sigma_i x_i}{\tau + \sum_{i=1}^k x_i}, \tau + \sum_{i=1}^k x_i\right).$$

Upon observing the realization of the signals, the team chooses $a \in \mathbb{R}$ to solve the problem $\max_{a \in \mathbb{R}} \pi - \int (a - s)^2 h(s|\sigma_1, ..., \sigma_k; x_1, ..., x_k) ds$, and from the first-order conditions it follows immediately that the optimal action is

$$a^{\star}(\sigma_1, ..., \sigma_k) = \int sh(s|\sigma_1, ..., \sigma_k; x_1, ..., x_k)ds = \mathbb{E}\left[s|\sigma_1, ..., \sigma_k; x_1, ..., x_k\right].$$

Solving for the expected value of information $V(x_1, ..., x_k)$ of a team as in equation (1), given optimal actions, involves solving for $k+1$ integrals (one for each of the k signals over the conditional distributions and one for the prior s over the prior distribution). For general distributions and loss functions, there is no closed form solution to this integral. Building on Marshak and Radner (1972), the normality of the distributions of the prior and the signals in conjunction with the quadratic loss function imply that the optimal action $a^{\star}$ is linear in the signal precision and the prior. This implies we can find a simple expression for the value of a team that consists of agents $x_1, ..., x_k$.

Claim 1 *The value of a match between k agents with types $x_1, ..., x_k$ is given by:*

$$V(x_1, ..., x_k) = \pi - \left(\frac{1}{\tau_s + \sum_{i=1}^{k} x_i}\right)$$

Proof. In Appendix. ■

It is easy to show that $V(x_1, ..., x_k) = \pi - \left(\tau_s + \sum_{i=1}^{k} x_i\right)^{-1}$ has the following properties. First, it depends only on the sum of the precisions of the team members, i.e., the *team's precision.* Second, and intuitively, it is strictly increasing in team precision, i.e., more information is better. Third, it is strictly *concave* in the precision of the team. Fourth, an immediate implication of the first and third properties is that the function $V(\cdot)$ is strictly *submodular* in $x_1, ..., x_k$.[3]

Since for a given firm n, $V(\cdot)$ depends on $x_1, ..., x_k$ only through the sum denoted by $X_n = \sum_{i=1}^{k} x_i$, we define $v(X_n) \equiv V(x_1, ..., x_k)$ and use it henceforth.

3.2 Solving the Matching Problem: Integer Assignment

It is instructive to start with the special case of $k = 2$. Since $v(\sum x_i^S)$ strictly submodular, it follows from Becker (1973) that the optimal sorting pattern exhibits negative assortative matching (NAM). That is, an agent with the most informative signal matches with one with the least informative one, and so on until all pairs are formed. More precisely, the optimal matching must satisfy the following condition: whenever we have types $x_i > x_j \geq x_\ell > x_m$, then total payoff is highest if we match x_i with x_m and x_j with x_ℓ, i.e., $v(x_i + x_m) + v(x_j + x_\ell) > \max\{v(x_i + x_j) + v(x_m + x_\ell), v(x_i + x_m) + v(x_j + x_\ell)\}$. This example also illustrates the economic intuition underlying NAM in the $k = 2$ case and diversification in the general one. By strict concavity, the increase in total payoff when x_i is added to x_m is bigger than that when it is added to x_j or x_ℓ. Thus, x_m can outbid both x_j and x_ℓ when competing for x_i, and the stable team structure must pair x_i with x_m.

[3]Alternatively, the partial derivative of V with respect to x_i and x_j, $V_{ij} = \frac{-2}{(\tau + \sum_{i=1}^{k} x_i^S)^3} < 0$.

In the $k = 2$ case there is a straightforward procedure (algorithm) to construct an optimal partition of Υ into teams of size two: given $x_1 \leq x_2 \leq \cdots \leq x_{kN}$, match x_1 with x_{kN}, x_2 with x_{kN-1}, and so forth. To see that the resulting matching is optimal, notice that any four types ordered as above will be negatively sorted; by strict submodularity, the objective function of (2) is maximized.

Things are more complex when $k > 2$. It is clear that strict submodularity of $v(\sum x_i^S)$ rules out positive assortative matching (PAM), i.e., the k agents with the highest precision match together, the next k form another team, etc.. For in this case one could take any two consecutive teams, swap the agent with the highest type in one of them with the one with the lowest type in the other and increase total payoff. Unlike the $k = 2$ case, however, submodularity does not provide an obvious sorting pattern or an algorithm. But it is clear that the optimal sorting of agents into teams entails *diversification* of types.

Indeed, diversification follows from the strict concavity of the objective function of (2). To see this, let us go back to the $k = 2$ case and consider $x_i > x_j \geq x_\ell > x_m$. Let $\bar{x} = (x_i + x_j + x_\ell + x_m)/2$, $x_{pq} = x_p + x_q$, $p, q = i, j, \ell, m$, $p \neq q$. Notice that all the partitions of these types have the same mean team precision $\bar{x}$. To illustrate, consider the PAM partition $\{\{x_i, x_j\}, \{x_\ell, x_m\}\}$. Then the team precisions are x_{ij} and $x_{\ell m}$, respectively, and thus its mean is $(x_{ij} + x_{\ell m})/2 = \bar{x}$. Similarly for $\{\{x_i, x_\ell\}, \{x_j, x_m\}\}$ and the NAM partition $\{\{x_i, x_m\}, \{x_j, x_\ell\}\}$. But in terms of their variance, the NAM partition that mixes high and low types has the smallest one and the PAM partition the largest one. By strict concavity, NAM maximizes the total payoff. Notice that one can order the four partitions from starting from PAM according to their variances.

The same is true in the general case. Define a partition $\mathcal{P}$ to be *more mixed than* a partition $\mathcal{P}'$ ($\mathcal{P} \succeq \mathcal{P}'$) if it can be obtained from P by an elementary swap of two types from two different teams as in the example above. Starting from the PAM partition, one can *completely* order the set $P(\Upsilon)$ using $\succeq$. Intuitively, a more mixed partition reduces the dispersion in the team precisions. Since the set of partitions is completely ordered, it has a maximal element. By strict concavity, the maximal element is the optimal partition. (All this needs to be proved.)

Proposition 1 *The optimal team structure is the partition that exhibits the maximal diversification of types. Equivalently, it is the one that minimizes the dispersion of the team precisions.*

Sketch of Proof. The proof would follow from these steps:

1. Define the majorization partial order $\succeq_m$ on $P(\Upsilon)$, i.e., $P \succeq_m P'$ if $(x_{S_1}, ..., x_{S_N})$, $x_{S_i} = \sum_{\ell \in S_i} x_i$, is majorized by $(x_{S'_1}, ..., x_{S'_N})$.

2. The vector corresponding to PAM majorizes all vectors ordered this way, while the vector with maximal diversification is majorized by every other vector.

3. The objective function is Schur-concave in $(x_{S_1}, ..., x_{S_N})$, x_{S_i} for any partition (it is the sum of concave functions of x_{S_i}).

4. By definition, the maximal diversification vector maximizes this function among all vectors that are ordered by majorization.

5. It remains to prove that any other vector in $P(\Upsilon)$ also majorizes the maximal diversification vector. $\square$

Although it is intuitively clear that the optimal assignment of types to teams exhibits maximal diversification, finding an efficient algorithm (which finds the solution in polynomial time) seems futile except in the $k = 2$ case. For some perspective, consider the $k = 3$ case. Proposition 1 suggest that finding an optimum is tantamount to finding a partition where team precisions are equalized as much as possible. Viewed this way, the problem seems awfully similar to the well-known 3-partition problem, which has been shown to be NP-complete.

3.3 Fractional Assignment

Many assignment problems involve fractional dedication time. Management consultants at McKinsey for example or lawyers in a partnership dedicate time to different projects that run in parallel. And researchers collaborate on different research projects with different co-authors at the same time. Applying this insight to our setting, we endow each agent with a time budget normalized to one, and she can dedicate a fraction of her time to different firms. The constraint is of course that the total time dedicated across all teams does not exceed one.

We follow the long tradition in the matching literature[4] and let the precision be *proportional* to the time dedication. In any team where the agent participates, his contribution is the observation of a signal with precision equal to the fraction of time dedicated times his full time precision. This interpretation is consistent with the view that the precision of information only obtains after time is dedicated, and that dedicated time is a perfect substitute: double the time dedication, double the precision of the signal.

Slightly changing notation, let $x(I) = \{x_1, ..., x_J\}$ be the set of *distinct* types of agents, and denote by m_j the number of agents of type x_j, and by X the sum of all types $\sum_{j=1}^{J} m_j x_j$. Clearly, $\sum_{j=1}^{J} m_j = kN$.[5] Denote by $\mu_{jn} \geq 0$ the fractional assignment of type-j agents to team n. Feasibility requires that $\sum_{j=1}^{J} \mu_{jn} = k$ for every n (i.e., the sum of the allocations of types to team n must equal the fixed team-size k), and $\sum_{n=1}^{N} \mu_{jn} = m_j$ for every j (i.e., the fraction of type-j agents allocated to all the teams must equal the total number of them m_j).

[4]The Monge-Kantorovic (and also Beckmann-Koopmans 1957, and Becker 1973) assignment problem considers fractional assignment, without the restriction that the allocation be integer.

[5]If k is larger relative to the number of types J, there will be several agents of a given type matched in the same partnership and only the marginal type will have a fractional assignment. In the limit as $k/J \to \infty$, the discrete assignment converges to the fractional assignment.

The fractional assignment problem is:

$$\max_{\{\mu_{jn}\}_{j,n}} \sum_{n=1}^{N} v\left(\sum_{j=1}^{J} \mu_{jn} x_j\right)$$

$$\text{s.t.} \quad \sum_{n=1}^{N} \mu_{jn} = m_j, \quad \sum_{j=1}^{J} \mu_{jn} = k \ , \quad \mu_{jn} \geq 0 \tag{3}$$

We will show that this problem has a unique solution. To this end, let $X_n \equiv \sum_{j=1}^{J} \mu_{jn} x_j$, and notice that if we multiply both sides of the first constraint (3) by x_j and sum with respect to j, we obtain $\sum_{n=1}^{N} X_n = X$.

Consider the following 'relaxed problem':

$$\max_{\{X_n\}_{n=1}^{N}} \sum_{n=1}^{N} v\left(X_n\right)$$

$$\text{s.t.} \quad \sum_{n=1}^{N} X_n = X$$

It is immediate from the necessary and sufficient first-order conditions that $v'(X_{n'}) = v'(X_{n''})$ for all $n' \neq n''$. Thus, $X_n = X/n$ for all n. Intuitively, concavity yields equalization of marginal benefits of team precision, which obviously implies the equalization of team precisions across all teams.

If the unique solution to the relaxed problem solves the original fractional assignment problem, then it is a solution to that problem. The task is to find values for the μ_{jn}s such that the equal precision rule X/N for all n satisfies all the feasibility constraints. Let $\mu_{jn} = m_j/N$ for all j, n. Then $X_n = \sum_{j=1}^{J} \mu_{jn} x_j = \sum_{j=1}^{J} (m_j/N) x_j = X/N$ for all n. Moreover, (6) is clearly satisfied. Regarding (6), notice that $\sum_n \mu_{jn} = \sum_n m_j/N = N m_j/N = m_j$. Finally, $\sum_j \mu_{jn} = \sum_j m_j/N = kN/N = k$. Thus, $\mu_{jn} = m_j/N$ for all j, n *solves* the fractional assignment problem. Since the solution distributes the total number of agents of a given type equally among all teams, we call this assignment *perfect diversification*.

We can show the following result:

Proposition 2 *Minimal differentiation between firms, i.e., identical firm precision X_n across firms is the unique equilibrium. Moreover, perfect diversification within firms $\mu_{jn} = \frac{m_j}{N}$ for all j, n, that implements X_n for all firms always exists.*

Notice that this is a natural generalization of NAM to the case of fractionally matching teams. We now establish the second part of the Proposition, whether perfect diversification is the unique solution to the fractional assignment problem. Suppose that, besides $\mu_{jn} = m_j/N$ for all j, n, there is another solution to the problem, say $\hat{\mu}_{jn}$. Define $\delta_{jn} \equiv \mu_{jn} - \hat{\mu}_{jn}$ as the difference between the two solutions. Since any solution must involve equalization of precision across teams, $\sum_j \mu_{jn} x_j = \sum_j \hat{\mu}_{jn} x_j$ for all n, or

$$\sum_{j=1}^{J}\delta_{jn}x_j = 0, \quad \sum_{n=1}^{N}\delta_{jn} = 0, \quad \sum_{j=1}^{J}\delta_{jn} = 0 \tag{4}$$

Notice that ((4) constitutes a homogeneous system of $2N + J - 2$ linear equations in δ_{jn}: there are $2N+J$ equations, of which 2 are dependent since feasibility requires that the total of the team capacities sums up to the number of agents, and the total of all X_n equals the total sum of individual x_j over all agents. Obviously, $\delta_{jn} = 0$ is a solution (the trivial one). If the number of equations is greater than or equal to the number of unknowns $J \times N$, then this is the unique solution, i.e., if $J \leq 2$. If not, then by a well-known result for homogeneous systems of linear equations there are nontrivial solutions as well.

The main message of this Proposition is that total precision is equalized across different teams, and that different *perfect diversification* can always be implemented with a proportional allocation, i.e., one where the distribution of types within each firm is identical across all firms and therefore identical to the economy wide distribution. Yet, given equal precision across firms, there can be different allocations that implement constant total precision. The multiplicity derives from the perfect substitutability of signals in value of information (it only depends on the total sum of precisions) in conjunction with the linear time constraints. The multiplicity will therefore most likely disappear whenever any of these properties does not hold. Below, we investigate correlated signals and find that precisions are not perfect substitutes in the value of information.

4 Extending the Basic Setup

In this section we consider two relevant extensions of the basic setup. First, we endogenize the number of teams in the presence of free entry of firms. Second, we allow for heterogeneity of firms. Firms differ in productivity and this obviously affects the value of information.

4.1 Endogenous Size of Teams

Consider a technology where the size of the firm is not fixed ex ante. There is no exogenous capacity constraint k on each firm. Rather, at stage zero and by incurring an entry cost F, identical entrepreneurs simultaneously enter the market. They hire workers of type x_j at a competitive wage rate w_j. As before, let there be J types with precision x_j, and there are m_j agents of each type. X still denotes the total precision in the economy. Denote by N the endogenous number of firms.

Each firm chooses the employment configuration μ_{jn}:

$$\max_{\mu_{jn}} v\left(X_n\right) - \sum_j \mu_{jn}w_j - F$$

$$\text{s.t.} \quad v\left(X_n\right) - \sum_j \mu_{jn}w_j - F = 0 \quad \text{and} \quad X_n = \sum_j \mu_{jn}x_j$$

where the first constraint is the zero profit condition from free entry. The first order condition implies, just as before, that X_n equalizes across all firms, given common wages w_j. Since in equilibrium there are N firms, this implies $X_n = \frac{X}{N}$. We can then write the zero profit condition as:

$$\pi - \frac{1}{\tau + X_n} - w_1\mu_{11} - w_2\mu_{12} - F = 0$$

From the first order condition, we obtain that $w_j = v'(X_n)\, x_i$ where $v' = \frac{1}{(\tau+X_n)^2}$, and $\mu_{jn} = \frac{m_j}{N}$.[6] We can then rewrite the zero profit condition along the equilibrium allocation as:

$$\pi - F \;=\; \frac{N}{N\tau + X} + \frac{NX}{(N\tau + X)^2}.$$

where we have substituted $\sum_j x_j m_j = X$, and $X_n = \frac{X}{N}$. This equation in one variable N completely determines the equilibrium. Let $N^\star$ be the solution where (5) holds with equality. In reality with indivisible firms, the number of firms will be the smallest positive integer. We treat the continuous variable $N^\star$ as a proxy for the discrete number of firms. The following result holds:

Proposition 3 *Provided $\pi - F < \frac{1}{\tau}$, there is a unique solution $N^\star$ that determines the number of firms and the firm size $k^\star = \frac{M}{N^\star}$. The firm size $k^\star$ is: 1. increasing in the entry cost F: $\frac{\partial k^\star}{\partial F} > 0$; 2. decreasing in the prior precision τ: $\frac{\partial k^\star}{\partial \tau} < 0$; 3. decreasing in the aggregate signal precision X: $\frac{\partial k^\star}{\partial X} < 0$.*

Proof. In Appendix. ■

With a positive entry cost, firms will compete away any surplus obtained from entry in the market. If the cost of entry tends to zero, there will be an infinite number of firms each hiring an infinitesimal measure of workers. A higher entry cost implies more profits need to be made to pay back the entry cost. Fewer firms enter which pushes up the firm size and in turn lowers the marginal product. A lower marginal product implies lower wages and thus higher profits.

In addition to the positive relation between firm size $k^\star$ and the fixed cost F, there is also a negative relation between firm size and both prior precision τ and total worker precision in the economy X. Prior precision increases the marginal product of additional workers at a given firm size. As a result, firms make higher profits, generating new entry, and thus lower firm size. The same logic applies for an increase in X.

This logic obviously applies to more general cost functions. In particular, we can introduce a cost that captures Hayek's view that a cost of communication to centralize information will impose bounds on the size of teams. That is, if within the team the cost of interaction between any two agents is c,

[6] Observe that even if there are multiple allocations, the total wage bill $\sum_j \mu_{jn} w_j$ is constant since there is uniqueness in payoffs. By Proposition 2, X_n is constant across different allocations, and therefore also the total wage bill.

then the total cost in a team of size k is $C(k) = \frac{1}{2}ck(k-1)$.[7] For expositional simplicity but without loss, we assume that $C(k) = \frac{1}{2}ck^2$, with $c'(k) = ck$ and let $J = 2$. Observe that exogenous size is equivalent to a weakly convex cost equal to zero if the size is less than k and equal to ∞ if the size is larger than k.

First let there be N firms, exogenously given. Now the problem of each firm n is:

$$\max_{\mu_{jn}} \quad v\left(X_n\right) - \sum_j \mu_{jn} w_j - C\left(\sum_j \mu_{jn}\right)$$

subject to the constraint that $\sum_n \sum_j \mu_{jn} = kN$. The first order conditions are:

$$v'\left(X_n\right) x_j - w_j - C'(k_n) = 0, \quad \forall j, n.$$

It is easily verified that a symmetric equilibrium with the Hayek cost gives (or any convex cost in general) $X_n = \frac{X}{N}$, provided profits are non-negative, i.e.

$$c < \left(\pi - \frac{1}{\tau + X/N} - \frac{X}{N}\frac{1}{(\tau + X/N)^2}\right)\left(\frac{N}{I}\right)^2,$$

where $I = kN$ denotes the cardinality of the set of agents. With free entry, the number of firms (again assuming divisibility of N), is determined by the condition:

$$\pi - \frac{1}{\tau + X/N} - \frac{X}{N}\frac{1}{(\tau + X/N)^2} - c\left(\frac{I}{N}\right)^2 = 0.$$

4.2 Firm Heterogeneity

How will information match if the firms are heterogeneous? Consider for example that firms are tackling problems of varying difficulty. Competing management consulting firms consult for clients of different size and market value. Then the value information at each of those firms will differ, and as a result so will the demand for informed agents.

Suppose there is a distribution of firm types y, where y indicates the firm's productivity, say due to its capital, or technology. Without loss of generality, let $y_1 \leq y_2 \leq \cdots \leq y_N$. Workers are heterogeneous in x and the information aggregation problem is exactly as before[8], which gives a value of information $v(X)$. We show below that this leads to an allocation with that resembles perfect diversification within firms. Let there be complementarity in y, v, where the total output can be represented by $yv(X_n)$ where v is determined as before.

[7] Notice that we treat k here as the total amount of units of time, which does not necessarily mean number of different people.

[8] For the purpose here, we fix the firm size to an exogenous k, though we could also endogenize k as above.

The planner solves:

$$\max_{X_1, X_2, \cdots, X_N} \sum_{n=1}^{N} y_n v\left(X_n\right)$$
$$\text{s.t.} \sum_n X_n = X$$

The first order conditions for this problem are:

$$y_n v'\left(X_n\right) = y_{n'} v'\left(X_{n'}\right), \quad \forall n \neq n'.$$

It is immediately apparent here that there is no simple solution that equates $X_n = X_{n'}$ as before, and that the solution in general will depend on the sequence of y_n. The next result establishes that we can solve for all the equilibrium conditions sequentially, yielding a simple and intuitive expression for the value of a match given y_n.

Claim 2 *Given N firms of type $y_1, ..., y_N$, the equilibrium allocation X_n at firm n is given by*

$$X_n = \frac{\tau\left[N\left(\frac{y_n}{y_1}\right)^{1/2} - \sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2}\right] + \left(\frac{y_n}{y_1}\right)^{1/2} X}{\sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2}}$$

Proof. In Appendix. ■

Observe that a special case where $y_n = y$ for all n implies $X_n = \frac{X}{N}$.[9] The above allows us to characterize the equilibrium allocation in the Proposition 4 below. Interestingly, higher y have higher X in equilibrium. More productive firms have a higher value of information and therefore demand on average higher types. Now there is no longer equalization of information across firms. As such, there is positive assortative matching between the firm type and the total informational precision of the agents. Interestingly, at the same time there is still diversification of information within the firm, a generalized form of negative assortative matching between agents of the same type.

This insight may be instructive in understanding differences across firms. There are clearly firms that are more productive than other ones. And they are typically also larger, though this is absent here in the present model. More importantly though is that we observe a great diversity of skills within the firm even if the demand for skills across firms differs. There is little evidence in empirical work of a systematic relation between the skill distribution within firms and the characteristics of the firm, such as productivity, size,... This is consistent with the form of generalized negative assortative matching

[9] While we consider y as exogenous, we could also imagine a technology that needs groups of people with complementary inputs (say sales and marketing), yet each team needs to solve its own information aggregation problem. In that case, the output can be written as $\pi - v\left(X^A\right) v\left(X^B\right)$, where A and B are the distinct populations (salesmen and marketeers).

and diversification of skills within the competitive firm, as derived in this model. Firms may differ substantially in productivity, yet they all diversify the work force.

Proposition 4 also establishes that a change in the prior precision τ affects the equilibrium precision X_n the low productivity firms negatively, and that of the high productivity firms positively. Not surprisingly, the dispersion in firm precision increases in the prior precision. In contrast, an increase in the aggregate precision increases all firms' precision as well as its dispersion.

Proposition 4 *Higher firm types y have higher aggregate precision: $X_1 \leq X_2 \leq \cdots \leq X_N$. Moreover:*

1. *When there is an increase in prior precision τ: 1. there exists $n^\star$ such that for all $n \leq n^\star$, X_n decreases, and X_n increases otherwise; 2. there is an increase in the dispersion in X_n.*

2. *An increase in the aggregate signal precision X: 1. increases all X_n; 2. decreases dispersion: $X_n - X_{n-1}$ increases in X*

3. *The firm level precision $X_n, \forall n$ is unique. The following allocation is always an equilibrium:*

$$\mu_{jn} = \frac{m_j}{N} + a_j \left(X_n - \frac{X}{N} \right)$$

where $a_j = \frac{(-1)^j}{\sum_{j=1}^J (-1)^j x_j}$ if J is even, and $a_j = \frac{d_j}{\sum_{j=1}^J d_j x_j}$ if J is odd, where d_j=-1 if j is odd and $d_j = \frac{J+1}{J-1}$ if j is even.

Proof. In Appendix. [**THESE RESULTS NEED DOUBLE CHECKING!!**] ■

An Example: $J = 2, k = 2$. Consider a simple example with $N = 2, x_1 = 1, x_2 = 2, m_1 = 2, m_2 = 1, k = 3/2, y_2 = 4, y_1 = 1$. Then $X = 4$

$$X_1 = \frac{2X + \tau}{3} = \frac{8 + \tau}{3}$$
$$X_2 = \frac{X - \tau}{3} = \frac{4 - \tau}{3}$$

so it is obviously the case that $X_1 + X_2 = X$. In the case of $J = 2$, the unique allocation is

$$\mu_{jn} = \frac{m_j}{2} + \frac{(-1)^j}{x_2 - x_1} \left(X_n - \frac{X}{2} \right)$$

and we can write the allocation matrix as:

$$\mu = \begin{pmatrix} \mu_{11} & \mu_{21} \\ \mu_{12} & \mu_{22} \end{pmatrix} = \begin{pmatrix} 1 + \frac{X+2\tau}{6} & \frac{1}{2} - \frac{X+2\tau}{6} \\ 1 - \frac{X+2\tau}{6} & \frac{1}{2} + \frac{X+2\tau}{6} \end{pmatrix}.$$

There is first-order stochastic dominance in the distributions of types across firms. The cumulative distribution of x's in the high y firm is everywhere below that in the low y firm.

5 Correlated Types

So far we considered matching of agents who obtain information about the state of the word that is independent of each others'. In reality, signals might well be correlated.[10] For example, while the ability to process information determines the precision of a signal, different agents may need to consult sources on which to base their decision. Management consultants scan different departments and auditors investigate the balance sheets of different division. Positive correlation can be the result of the fact that collaborating agents tend to access the same sources in order to form an opinion. Alternatively, negative correlation can result from the fact that agents necessarily access different sources: say there are two sources (the finance division and the logistics division), and only one of them will give an insight about s, but it is not ex ante known which one. If one agent investigates finance and the other logistics, the signals may be negatively correlated.

Not only is the correlated case relevant *per se*, it also serves as a robustness check for the analysis of independent signals. In the absence of correlation, we know that the matching pattern induces maximal diversification of types within the company. We therefore investigate how the matching pattern is affected by correlation.

In line with the set up developed earlier, we consider teams of two agents, where signals are now drawn from a bivariate normal distribution and where $\rho \in [-1, 1]$ is the correlation between signals σ_1 and σ_2:

$$(\sigma_1, \sigma_2) \sim f(\sigma_1, \sigma_2) \sim \mathcal{N}\left(\begin{bmatrix} s \\ s \end{bmatrix}, \begin{bmatrix} \frac{1}{x_1} & \frac{\rho}{\sqrt{x_1 x_2}} \\ \frac{\rho}{\sqrt{x_1 x_2}} & \frac{1}{x_2} \end{bmatrix}\right)$$

and

$$s \sim \mathcal{N}\left(\mu, \frac{1}{\tau}\right) \sim h(s).$$

As before, the objective function of the team is to choose a joint action a to maximize $\pi - (a - s)^2$, where s is unknown but the two correlated signals σ_1, σ_2 are used to form a belief about s. The following result established the value of information V if agents with precision x_1 and x_2 are matched.[11]

Claim 3 *Let $\rho \in [-1, 1]$. The value of information of matching two agents with precision x_1 and x_2 is:*

$$V(x_1, x_2) = \frac{-1}{\tau + \frac{x_1 + x_2 - 2\rho(x_1 x_2)^{1/2}}{1-\rho^2}}$$

Proof. In Appendix. ■

[10] Some simple yet natural degree of correlation is of course already built in. The draws of each of the agents has a *common* mean equal to the state of nature s. Both agents learn about the same event, and therefore the distribution from which they draw is centered around the same mean. Here, we aim to introduce correlation in the informativeness of the signals, which is captured by the (co)variance.

[11] In the Appendix we derive the value of information with correlated signals for teams of size k.

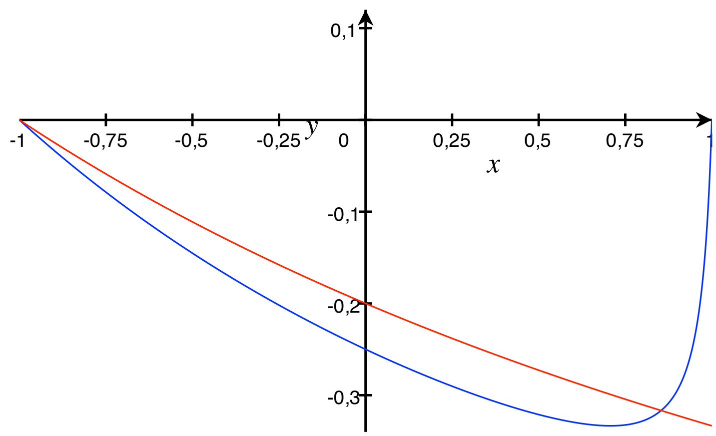

Figure 1: The Value of information $V(x_1, x_2)$ (vertical axis) as a function of ρ (horizontal axis): different precision ($x_1 = 1, x_2 = 2$, in blue); equal precision ($x_1 = x_2 = 1$, in red); parameter values: $\pi = 0, \tau = 1$.

Observe that the value of information is a function of the Blackwell informativeness index **[CITATIONS???]**, given by:

$$\mathcal{B}(x_1, x_2) = \frac{x_1 + x_2 - 2\rho(x_1 x_2)^{1/2}}{1 - \rho^2} \quad \Rightarrow \quad V(x_1, x_2) = \pi - \frac{1}{\tau + \mathcal{B}(x_1, x_2)}.$$

Intuitively, positive correlation is bad for information. To see this, consider the special case where $x_1 = x_2 = x$. Then we can write the Blackwell informativeness index as $\mathcal{B}(x, x) = \frac{2x}{1+\rho}$, and therefore the value of information is:

$$V(x, x) = \pi - \frac{1}{\tau + \frac{2x}{1+\rho}}.$$

It is immediately verified that $\mathcal{B}(x, x)$ and therefore $V(x, x)$ is everywhere decreasing in the correlation ρ. Compared to independence, positively correlated signals *subtract* from $\mathcal{B}$ (that is, $\frac{2x}{1+\rho} < 2x$) and therefore the value of information since correlated signals duplicate some of the information. In fact, two perfectly (positively) correlated signals ($\rho = 1$) are just as good as having only one, and the Blackwell informativeness index is equal to x, just what it is when one signal only is drawn. The opposite is true with negative correlation. Now signals *add* informativeness relative to independence (that is, $\frac{2x}{1+\rho} > 2x$ when ρ is negative). The best way to hedge risk is to take a position in a negatively correlated asset. When there is perfect negative correlation ($\rho = -1$), $\mathcal{B} \to \infty$: the two signals are perfectly informative and allow for the exact determination of the state, i.e., the value of information is maximized $V(x, x) = \pi$. This is illustrated in Figure 5 (in red).

When signal precision differs amongst agents, the notion that correlation subtracts from informativeness continues to be true. This can immediately be seen from inspection of $\mathcal{B}$, where the term $x_1 + x_2 - 2\rho(x_1 x_2)^{1/2} < x_1 + x_2$ under positive correlation. However, this is only part of the impact of correlation. There is also a confounding impact due to the updating with different precision. Correlation affects the prior precision τ and Blackwell informativeness differentially through the term $(1-\rho)^2$. As a result, there is a trade off between the prior precision and the team member's precision. As the correlation is high enough, the effect of the prior starts to dominate. When signals are close to perfectly (positively) correlated, the difference in precision becomes increasingly informative. If both agents observe different signals, the updating rule will lead to a very big movement in the posterior, which is highly informative. Observe that this effect of the informativeness never kicks in when the precision of both agents is the same. Under perfect correlation, their updating rule is colinear, rendering the second signal redundant. As a result, there is a discontinuity in $V(x_1, x_2)$ at $\rho = 1$ as x_1 approaches x_2.

More importantly, correlation affects the sorting pattern. We first state the following general result

Proposition 5 *There always exist values of ρ for which there is Negative Assortative Matching and there always exist different value of ρ for which there is Positive Assortative Matching.*

In the light of the zero correlation result, this result is quite surprising. In particular the fact that there always exists values of ρ for which there is Positive Assortative Matching. To get an idea for the underlying motive for this result, we first prove it by means of two Lemma's that characterize the equilibrium allocation. In order to get an idea of the matching pattern, we derive the sign of the cross-partial $V_{12}(x_1, x_2)$. The cross-partial of V is given by:

$$V_{12} = \frac{\mathcal{B}_{12}(\tau + \mathcal{B}) - 2\mathcal{B}_1\mathcal{B}_2}{(\tau + \mathcal{B})^3},$$

where

$$\begin{aligned}
\mathcal{B}_1 &= \frac{1}{1-\rho^2}\left[1 - \rho x_1^{-1/2} x_2^{1/2}\right] \\
\mathcal{B}_2 &= \frac{1}{1-\rho^2}\left[1 - \rho x_1^{1/2} x_2^{-1/2}\right] \\
\mathcal{B}_{12} &= -\frac{1}{2}\frac{\rho}{1-\rho^2} x_1^{-1/2} x_2^{-1/2}
\end{aligned}$$

After a change of variables $x_1^{1/2} = y_1$ and $x_2^{1/2} = y_2$ and canceling $(1-\rho^2)^2$ in numerator and denominator, we can write V_{12} as:

$$V_{12} = (1-\rho^2)\frac{-\frac{\rho}{2}\frac{1}{y_1 y_2}\left(\tau(1-\rho^2) + y_1^2 + y_2^2 - 2\rho y_1 y_2\right) - 2\left[1 - \rho\frac{y_2}{y_1}\right]\left[1 - \rho\frac{y_1}{y_2}\right]}{\left(\tau(1-\rho^2) + y_1^2 + y_2^2 - 2\rho y_1 y_2\right)^3}.$$

Now we can establish the first Lemma that leads to the proof of Proposition 5:

Lemma 1 *Consider any general matching problem with a bounded prior, $\tau < \infty$. Then:*

1. *In an open neighborhood of $\rho = 0$, there is always negative assortative matching;*

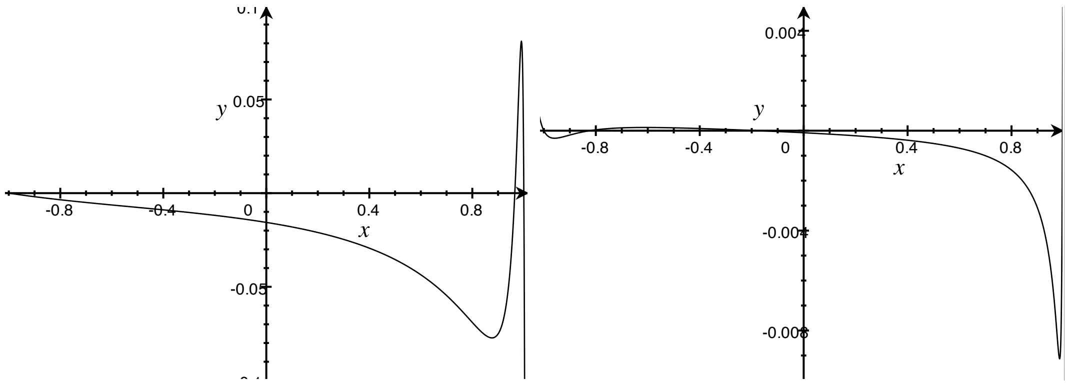

Figure 2: $V_{12}(x_1, x_2)$ (vertical axis) as a function of ρ (horizontal axis): A. $\tau = 1$; B. $\tau = 50$; parameter values: $x_1 = 1, x_2 = 2, \pi = 0, \tau = 1$.

2. *In an open left-neighborhood of $\rho = 1$, there is always positive assortative matching;*

3. *In an open right-neighborhood of $\rho = -1$, there is always negative assortative matching.*

Proof. In Appendix. ■

The proof of Lemma 1 result immediately demonstrates the result in Proposition 5.

The cross-partial goes positive when ρ becomes large because positive correlation introduces a complementarity in the value. We know that V_{12} goes to zero as ρ approaches 1, yet in the left-neighborhood of $\rho = 1$, the cross-partial is positive. In the Blackwell precision $\mathcal{B}$, the term $-2\rho(x_1 x_2)^{1/2}$ is maximized for ρ close to one. This term introduces a complementarity (the cross-partial is negative and $\mathcal{B}$ enters in V as the inverse), and as the proof establishes, it always dominates the substitutability due to the concavity. Yet the magnitude of the complementarity depends on the difference between x_1 and x_2. As they become closer, the complementarity becomes smaller. When they are equal, the complementarity is zero, hence the cross-partial is everywhere negative for identical x.

The next Lemma establishes that PAM can also occur for negative values of ρ (bounded away from -1 and 0), provided the prior τ is sufficiently high.

Lemma 2 *There exist values of τ sufficiently high, such that for some $\rho < 0$ there is Positive Assortative Matching.*

Proof. In Appendix. ■

Positive Assortative Matching can arise under negative correlation and for sufficiently high values of the prior precision. An insight into the intuition for this can be found in investigating the "riskiness" of the value. $-V''/V'$ is decreasing in τ, therefore for large τ the planner is approximately "risk neutral". Of course, there is a close relationship between V'' and V_{12} that depends on ρ. When $\rho = 0$, $V'' = V_{12}$

and no matter how large τ, V_{12} is negative. Instead, when $\rho \neq 0$, V_{12} can become positive, while V'' remains negative, provided $-V"/V'$ gets close enough to zero from below. This occurs when $\mathcal{B}$ is large enough relative to the contribution by the precision of the signals. This occurs when τ is sufficiently large for negative, intermediate values of ρ. There is therefore a fundamental trade off between prior precision and and team precision.

Appendix

Proof of Claim 1

Proof. Inserting the expression for $a^\star$ in the objective function of (1), and denoting the unconditional joint density of the signals by $f(\vec{\sigma}|\vec{x}^S) \equiv \int \left(\prod_{i=1}^k f(\sigma_i|s, x_i^S)\right) h(s)ds$, we obtain

$$\begin{aligned}
V(\vec{x}^S) &= \pi - \int \cdots \int \left(\mathbb{E}[\tilde{s}|\vec{\sigma}, \vec{x}^S] - s\right)^2 \left(\prod_{i=1}^k f(\sigma_i|s, x_i^S)\right) h(s) \prod_{i=1}^k d\sigma_i ds \\
&= \pi - \int \cdots \int \left(\int \left(s - \mathbb{E}[\tilde{s}|\vec{\sigma}, \vec{x}^S]\right)^2 h(s|\vec{\sigma}, \vec{x}^S) ds\right) f(\vec{\sigma}|\vec{x}^S) \prod_{i=1}^k d\sigma_i \\
&= \pi - \int \cdots \int \left(\frac{1}{\tau_s + \sum_{i=1}^k x_i^S}\right) f(\vec{\sigma}|\vec{x}^S) \prod_{i=1}^k d\sigma_i \\
&= \pi - \left(\frac{1}{\tau_s + \sum_{i=1}^k x_i^S}\right), \qquad (5)
\end{aligned}$$

where the second equality follows from $h(s|\vec{\sigma}, \vec{x}^S) f(\vec{\sigma}|\vec{x}^S) = h(s) \prod_{i=1}^k f(\sigma_i|s, x_i^S)$, the third by replacing the expression for the variance of the posterior density, and the last equality from the fact that the variance is independent of the realizations of the signals. ■

Proof of Proposition 3

Proof. We can write the equilibrium condition as:

$$\pi - F = \frac{(N^\star)^2\tau + 2N^\star X}{(N^\star\tau + X)^2}.$$

The LHS is a positive constant; denote the RHS by $f(N^\star)$, a function of $N^\star$. At zero, $f(0) = 0$, while at infinity, $\lim_{N^\star \to \infty} f(N^\star) = \frac{1}{\tau}$. First, observe that there is a lower bound on τ for the existence of a zero profit equilibrium: if the prior precision is very high, the value of the team will be positive no matter the precision of the team members and no matter how many teams enter the market. Second, $f(N^\star)$ is increasing in $N^\star$:

$$\begin{aligned}
f'(N^\star) &= \frac{(2N^\star\tau + 2X)(N^\star\tau + X) - ((N^\star)^2\tau + 2XN^\star)2\tau}{(N^\star\tau + X)^3} \\
&= \frac{2X^2}{(N^\star\tau + X)^3} > 0.
\end{aligned}$$

As a result, provided $\pi - F < \frac{1}{\tau}$, there is a unique solution N and therefore $k = \frac{M}{(}N^\star)$.

Writing (5) as $f(N^\star) - \pi + F = 0$ where $f'(N^\star) > 0$. Applying the implicit function theorem in each case:

1.

$$\frac{\partial N^\star}{\partial F} = -\frac{1}{f'(N^\star)} < 0.$$

Therefore:

$$\frac{\partial k^\star}{\partial F} = \frac{\partial \frac{M}{N^\star}}{F} = -\frac{M}{(N^\star)^2}\frac{\partial N^\star}{\partial F} > 0.$$

2.

$$\begin{aligned}\frac{\partial N^\star}{\partial \tau} &= -\frac{\frac{(N^\star)^2(N^\star\tau+X)-((N^\star)^2\tau+2XN^\star)2N^\star}{(N^\star\tau+X)^3}}{f'(N^\star)} \\ &= \frac{(N^\star)^2(N^\star\tau+3X)}{2X^2} > 0.\end{aligned}$$

Therefore:

$$\frac{\partial k^\star}{\partial \tau} = \frac{\partial \frac{M}{N^\star}}{\tau} = -\frac{M}{(N^\star)^2}\frac{\partial N^\star}{\partial \tau} < 0.$$

3.

$$\begin{aligned}\frac{\partial N^\star}{\partial X} &= -\frac{\frac{2N^\star(N^\star\tau+X)-((N^\star)^2\tau+2XN^\star)2}{(N^\star\tau+X)^3}}{f'(N^\star)} \\ &= \frac{N^\star}{X} > 0.\end{aligned}$$

Therefore:

$$\frac{\partial k^\star}{\partial X} = \frac{\partial \frac{M}{N^\star}}{X} = -\frac{M}{(N^\star)^2}\frac{\partial N^\star}{\partial X} < 0.$$

■

Proof of Claim 2

Proof. Given the value of information, $V'(X_n) = \frac{1}{\tau+X_n}$, we can write the first order conditions as:

$$\frac{\tau+X_{n-1}}{\tau+X_n} = \left(\frac{y_{n-1}}{y_n}\right)^{1/2} = \frac{1}{b_n}, \forall n = 2,\ldots,N,$$

where we use the shorthand notation b_n for the ratio. Then we can write the following sequence:

$$\begin{aligned}X_2 &= b_2X_1 + \tau(b_2-1) \\ X_3 &= b_3\left(b_2X_1+\tau(b_2-1)\right)+\tau(b_3-1) \\ X_4 &= b_4b_3b_2X_1+\tau\left(b_4\left(b_3(b_2-1)+b_3-1\right)+b_4-1\right) \\ &= b_4b_3b_2X_1+(b_4b_3b_2-1)\tau \\ \ldots \\ X_n &= \prod_{l=2}^{n} b_lX_1+\left(\prod_{l=2}^{n} b_l-1\right)\tau\end{aligned}$$

Then we can write:

$$
\begin{aligned}
X_1 &= X - \sum_{n=2}^{N} X_n \\
&= X - \sum_{n=2}^{N} \left(\prod_{l=2}^{n} b_l X_1 + \left(\prod_{l=2}^{n} b_l - 1 \right) \tau \right)
\end{aligned}
$$

or

$$
X_1 = \frac{X - \tau \sum_{n=2}^{N} \left(\prod_{l=2}^{n} b_l - 1 \right)}{1 + \sum_{n=2}^{N} \prod_{l=2}^{n} b_l}
$$

Observe that

$$
\prod_{l=2}^{n} b_l = \left(\frac{y_2}{y_1} \right)^{1/2} \left(\frac{y_3}{y_2} \right)^{1/2} \cdots \left(\frac{y_n}{y_{n-1}} \right)^{1/2} = \left(\frac{y_n}{y_1} \right)^{1/2}
$$

and therefore we can write

$$
X_n = \left(\frac{y_n}{y_1} \right)^{1/2} X_1 + \left(\left(\frac{y_n}{y_1} \right)^{1/2} - 1 \right) \tau.
$$

Then,

$$
\begin{aligned}
X_n &= \left(\left(\frac{y_n}{y_1} \right)^{1/2} - 1 \right) \tau + \left(\frac{y_n}{y_1} \right)^{1/2} \frac{X - \tau \sum_{n=2}^{N} \left(\left(\frac{y_n}{y_1} \right)^{1/2} - 1 \right)}{1 + \sum_{n=2}^{N} \left(\frac{y_n}{y_1} \right)^{1/2}} \\
&= \left(\left(\frac{y_n}{y_1} \right)^{1/2} - 1 \right) \tau + \left(\frac{y_n}{y_1} \right)^{1/2} \frac{X - \tau \sum_{n=1}^{N} \left(\left(\frac{y_n}{y_1} \right)^{1/2} - N \right)}{\sum_{n=1}^{N} \left(\frac{y_n}{y_1} \right)^{1/2}} \\
&= \frac{-\tau \sum_{n=1}^{N} \left(\frac{y_n}{y_1} \right)^{1/2} + \left(\frac{y_n}{y_1} \right)^{1/2} X + \left(\frac{y_n}{y_1} \right)^{1/2} \tau N}{\sum_{n=1}^{N} \left(\frac{y_n}{y_1} \right)^{1/2}} \\
&= \frac{\tau \left[N \left(\frac{y_n}{y_1} \right)^{1/2} - \sum_{n=1}^{N} \left(\frac{y_n}{y_1} \right)^{1/2} \right] + \left(\frac{y_n}{y_1} \right)^{1/2} X}{\sum_{n=1}^{N} \left(\frac{y_n}{y_1} \right)^{1/2}}
\end{aligned}
$$

■

Proof of Proposition 4

Proof. Since $y_1 < y_2 < \cdots < y_N$, it immediately follows that $(N\tau + X) \left(\frac{y_n}{y_1} \right)^{1/2}$ is increasing in n and therefore, $X_1 \leq X_2 \leq \cdots \leq X_N$.

Comparative Statics. From the implicit function theorem,

$$\frac{\partial X_n}{\partial \tau} = \frac{\left[N\left(\frac{y_n}{y_1}\right)^{1/2} - \sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2}\right]}{\sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2}}$$

There is a critical type $n^\star$ such that $N\left(\frac{y_n^\star}{y_1}\right)^{1/2} - \sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2} \leq 0$ and with opposite inequality when $n = n^\star + 1$. As a result, $\frac{\partial X_n}{\partial \tau} < 0$ if $n \leq n^\star$ and $\frac{\partial X_n}{\partial \tau} > 0$ otherwise.

Moreover, since y_n is increasing in n and $\sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2}$ is a constant, it follows that

$$\frac{\partial X_n}{\partial \tau} - \frac{\partial X_{n-1}}{\partial \tau} = N\frac{\left(\frac{y_n}{y_1}\right)^{1/2} - \left(\frac{y_{n-1}}{y_1}\right)^{1/2}}{\sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2}} > 0.$$

Again, applying the implicit function theorem:

$$\frac{\partial X_n}{\partial X} = \frac{\left(\frac{y_n}{y_1}\right)^{1/2}}{\sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2}} > 0,$$

i.e., all X_n increase in X. Also dispersion in X_n increases:

$$\frac{\partial X_n}{\partial X} - \frac{\partial X_{n-1}}{\partial X} = \frac{\left(\frac{y_n}{y_1}\right)^{1/2} - \left(\frac{y_{n-1}}{y_1}\right)^{1/2}}{\sum_{n=1}^{N}\left(\frac{y_n}{y_1}\right)^{1/2}} > 0.$$

The Allocation. Since X_n is defined as:

$$\begin{aligned}
X_n &= \sum_{j=1}^{J}\mu_{jn}x_j \\
&= \sum_{j=1}^{J}\left[\frac{m_j x_j}{N} + a_j x_j\left(X_n - \frac{X}{N}\right)\right] \\
&= \frac{X}{N} + \left(X_n - \frac{X}{N}\right)\sum_{j=1}^{J}a_j x_j \\
&= X_n
\end{aligned}$$

since both

$$\sum_{j=1}^{J}a_j x_j = \sum_{j=1}^{J}\frac{(-1)^j x_j}{\sum_{j=1}^{J}(-1)^j x_j} = 1$$

if J is even, and

$$\sum_{j=1}^{J} a_j x_j = \frac{d_j x_j}{\sum_{j=1}^{J} d_j x_j} = 1$$

if J is odd.

The allocation must also satisfy feasibility:

$$\begin{aligned}
\sum_{n=1}^{N} \mu_{jn} &= m_j \;\; \forall j \\
\sum_{j=1}^{J} \mu_{jn} &= k \quad \forall n \\
\mu_{jn} &\geq 0 \quad \forall j, n
\end{aligned}$$

The first condition implies

$$\begin{aligned}
\sum_{n=1}^{N} \left[\frac{m_j}{N} + a_j \left(X_n - \frac{X}{N} \right) \right] &= m_j \\
m_j + a_j \sum_{n=1}^{N} \left(X_n - \frac{X}{N} \right) &= m_j \\
m_j + a_j \sum_{n=1}^{N} (X - X) &= m_j
\end{aligned}$$

The second condition becomes:

$$\begin{aligned}
\sum_{j=1}^{J} \mu_{jn} &= k \\
\sum_{j=1}^{J} \left[\frac{m_j}{N} + a_j \left(X_n - \frac{X}{N} \right) \right] &= k \\
\frac{M}{N} + \left(X_n - \frac{X}{N} \right) \sum_{j=1}^{J} a_j &= k
\end{aligned}$$

where both

$$\sum_{j=1}^{J} a_j = \sum_{j=1}^{J} \frac{(-1)^j}{\sum_{j=1}^{J} (-1)^j x_j} = 0$$

if J is even, and

$$\sum_{j=1}^{J} a_j = \frac{d_j}{\sum_{j=1}^{J} d_j x_j} = 0$$

if J is odd, where $d_j = -1$ if j is odd and $d_j = \frac{J+1}{J-1}$ if j is even. Therefore the condition becomes $\frac{M}{N} = k$ which is satisfied by construction since all firms are identical. Observe that the condition that $\mu_{jn} \geq 0$ is always satisfied. **[CHECK]** ■

Proof of Claim 3

Proof. As before, from the first order condition, the optimal action $a^\star = \mathbb{E}s$. The objective therefore is to find the distribution of s given the signals: $h(s|\sigma_1, \sigma_2)$. We will proceed in two steps, where we can treat the signal as being generated sequentially.[12]

1. Calculate $h(s|\sigma_i)$. Since $h(s)$ is normal and $f(\sigma_i|s) \sim \mathcal{N}\left(s, \frac{1}{x_i}\right)$, we get

$$h(s|\sigma_1) \sim \mathcal{N}\left(\frac{\mu\tau + x_1\sigma_1}{\tau + x_1}, \frac{1}{\tau + x_1}\right) \sim \mathcal{N}\left(\mu', \frac{1}{\tau'}\right) \tag{6}$$

2. We calculate $f(\sigma_2|\sigma_1, s)$. To that end, we incorporate the conditioning on the correlated signal σ_1. Conditional on observing σ_1, the mean and variance of the distribution of σ_2 depend on the realization of σ_1. The variance decreases by $1-\rho^2$. The mean now depends on the sum of the unconditional expected value of σ_2, which is s, and on the expected value of s given σ_1 **(WE HAVE TO ADD WHERE THE MEAN COMES FROM):**

$$\begin{aligned} f(\sigma_2|\sigma_1, s) &\sim \mathcal{N}\left(s + \left(\frac{x_1}{x_2}\right)^{1/2} \rho(\sigma_1 - s), \frac{1-\rho^2}{x_2}\right) \\ &\sim \mathcal{N}\left(\left(1 - \left(\frac{x_1}{x_2}\right)^{1/2} \rho\right) s + \left(\frac{x_1}{x_2}\right)^{1/2} \rho\sigma_1, \frac{1-\rho^2}{x_2}\right) \end{aligned}$$

   It is apparent from the mean of this normal distribution that $\mathbb{E}[\sigma_2|\sigma_1, s] \neq s$. If we write

$$\sigma_2 = as + b + \varepsilon$$

   where $a = 1 - \left(\frac{x_1}{x_2}\right)^{1/2} \rho$, $b = \left(\frac{x_1}{x_2}\right)^{1/2} \rho\sigma_1$, $\varepsilon \sim \mathcal{N}\left(0, \frac{1-\rho^2}{x_2}\right)$, then the transformed signal $\hat{\sigma}_2$ defined as

$$\hat{\sigma}_2 = \frac{\sigma_2 - b}{a} = s + \frac{\varepsilon}{a}$$

[12] To see this, observe that from updating, and given that $h(s|\sigma_1) = \frac{f(\sigma_1|s)h(s)}{\int f(\sigma_1|t)h(t)dt}$ and $f(\sigma_2|\sigma_1, s) = \frac{f(\sigma_1,\sigma_2|s)}{f(\sigma_1|s)}$, we can write:

$$h(s|\sigma_1, \sigma_2) = \frac{f(\sigma_2|\sigma_1, s)h(s|\sigma_1)}{\int f(\sigma_2|\sigma_1, t)h(t|\sigma_1)dt} = \frac{\frac{f(\sigma_1,\sigma_2|s)}{f(\sigma_1|s)} \frac{f(\sigma_1|s)h(s)}{\int f(\sigma_1|t)h(t)dt}}{\int \frac{f(\sigma_1,\sigma_2|t)}{f(\sigma_1|t)} \frac{f(\sigma_1|t)h(t)}{\int f(\sigma_1|t)h(t)} dt} = \frac{f(\sigma_1|\sigma_2, s)h(s|\sigma_2)}{\int f(\sigma_1|\sigma_2, t)h(t|\sigma_2)dt}$$

is normal with mean $\mathbb{E}[\hat{\sigma}_2] = s$ and variance $\text{Var}[\hat{\sigma}_2] = \frac{\text{Var}(\varepsilon)}{a^2}$. Therefore

$$\begin{aligned}
\hat{\sigma}_2 &\sim \mathcal{N}\left(s, \frac{\frac{1-\rho^2}{x_2}}{\left(1-\left(\frac{x_1}{x_2}\right)^{1/2}\rho\right)^2}\right) && (7)\\
&\sim \mathcal{N}\left(s, \frac{1-\rho^2}{\left(x_2^{1/2}-x_1^{1/2}\rho\right)^2}\right) && (8)\\
&\sim \mathcal{N}\left(s, \frac{1}{\hat{\tau}}\right) && (9)
\end{aligned}$$

Then using the result from the two steps, (6) and (8), we obtain:

$$h(s|\sigma_1, \sigma_2) \sim \mathcal{N}\left(\frac{\tau'\mu' + \hat{\tau}\hat{\sigma}_2}{\tau' + \hat{\tau}}, \frac{1}{\tau' + \hat{\tau}}\right)$$

where

$$\begin{aligned}
\tau' + \hat{\tau} &= \tau + x_1 + \frac{\left(x_2^{1/2} - x_1^{1/2}\rho\right)^2}{1-\rho^2}\\
&= \tau + \frac{x_1 + x_2 - 2\rho(x_1x_2)^{1/2}}{1-\rho^2}
\end{aligned}$$

Since as before, we can write the value of information as a function of the precision, it follows that:

$$V(x_1, x_2) = \pi - \frac{1}{\tau + \frac{x_1+x_2-2\rho(x_1x_2)^{1/2}}{1-\rho^2}}$$

■

Derivation of the Value of Information with Correlated Signals for Size k Teams

$$V(x_1, ..., x_k) = \pi - \frac{1}{\tau + \mathcal{B}(x_1, ..., x_k)}$$

where

$$\mathcal{B}(x_1, ..., x_k) = \frac{(1 + (k-2)\rho)\sum_{i=1}^k x_i - 2\rho\left((x_1x_2)^{1/2} + ... + (x_1x_k)^{1/2} + (x_2x_3)^{1/2} + ... + (x_{k-1}x_k)^{1/2}\right)}{(1-\rho)(1+(k-1)\rho)}$$

Confirm the special cases:

1. $\rho = 0$

$$V(x_1, ..., x_k) = \pi - \frac{1}{\tau + \sum_{i=1}^k x_i}$$

2. $x_1 = ... = x_k = x$

$$V(x, ..., x) = \pi - \frac{1}{\tau + \frac{kx}{1+(k-1)\rho}}$$

Proof of Lemma 1

Proof. We consider each of the three values of ρ:

1. $\rho = 0$
   Evaluating V_{12} at $\rho = 0$, we obtain:

$$V_{12} = \frac{-2}{(\tau + x_1 + x_2)^3} < 0.$$

   When τ is bounded, then V_{12} is bounded away from zero. V_{12} is continuous in ρ and as a result, in the neighborhood (left and right) of $\rho = 0$ V_{12} is negative and there is Negative Assortative Matching.

2. $\rho = 1$

$$\begin{aligned} V_{12} &= (1-\rho^2)\frac{-\frac{1}{2}\frac{1}{y_1 y_2}(y_1 - y_2)^2 - \frac{2}{y_1 y_2}[y_1 - y_2][y_2 - y_1]}{(y_1 - y_2)^6} \\ &= (1-\rho^2)\frac{-\frac{1}{2}\frac{1}{y_1 y_2}(y_1 - y_2)^2 + \frac{2}{y_1 y_2}[y_1 - y_2]^2}{(y_1 - y_2)^6} \\ &= (1-\rho^2)\frac{\frac{3}{2}\frac{1}{y_1 y_2}}{(y_1 - y_2)^4} \end{aligned}$$

   This is zero at $\rho = 1$ for all $x_1 \neq x_2$.

   Now we show that at $\rho = 1$, the derivative of V_{12} wrt ρ is negative (we obviously take the derivative of $V_{12}(\rho)$ for all ρ and then evaluate at $\rho = 1$):

$$\begin{aligned} \frac{dV_{12}}{d\rho} &= -2\rho\frac{\frac{3}{2}\frac{1}{y_1 y_2}}{(y_1 - y_2)^4} \\ \left.\frac{dV_{12}}{d\rho}\right|_{\rho=1} &= -\frac{-3\frac{1}{y_1 y_2}}{(y_1 - y_2)^4} < 0 \end{aligned}$$

   Therefore, in the left-neighborhood of $\rho = 1, V_{12} > 0$, and there is PAM.

3. $\rho = -1$
   As in the case of $\rho = 1$, we can immediately establish that $V_{12} = 0$ for $\rho = -1$:

$$\begin{aligned} V_{12} &= (1-\rho^2)\frac{\frac{1}{2}\frac{1}{y_1 y_2}(y_1 + y_2)^2 - \frac{2}{y_1 y_2}[y_1 + y_2][y_2 + y_1]}{(y_1 + y_2)^6} \\ &= (1-\rho^2)\frac{-\frac{3}{2}\frac{1}{y_1 y_2}}{(y_1 + y_2)^4}. \end{aligned}$$

   Again, we calculate the derivative with respect to ρ:

$$\begin{aligned} \frac{dV_{12}}{d\rho} &= -2\rho\frac{-\frac{3}{2}\frac{1}{y_1 y_2}}{(y_1 - y_2)^4} \\ \left.\frac{dV_{12}}{d\rho}\right|_{\rho=-1} &= \frac{-3\frac{1}{y_1 y_2}}{(y_1 - y_2)^4} < 0 \end{aligned}$$

and therefore in the right-neighborhood of $\rho = -1, V_{12} < 0$ and there is NAM.

■

5.1 Proof of Lemma 2

Proof.

$$V_{12} = (1-\rho^2)\frac{-\frac{\rho}{2}\frac{1}{y_1 y_2}\left(\tau(1-\rho^2) + y_1^2 + y_2^2 - 2\rho y_1 y_2\right) - 2\left[1-\rho\frac{y_2}{y_1}\right]\left[1-\rho\frac{y_1}{y_2}\right]}{\left(\tau(1-\rho^2) + y_1^2 + y_2^2 - 2\rho y_1 y_2\right)^3}.$$

The denominator is always positive. Let $\phi = -\rho$, then the numerator can be written as:

$$\begin{aligned}
&= \frac{\phi}{2}\frac{1}{y_1 y_2}\left(\tau(1-\phi^2) + y_1^2 + y_2^2 + 2\phi y_1 y_2\right) - \frac{2}{y_1 y_2}\left[y_1 + \phi y_2\right]\left[y_2 + \phi y_1\right] \\
&= \frac{1}{2y_1 y_2}\left[\phi\left(\tau(1-\phi^2) + y_1^2 + y_2^2 + 2\phi y_1 y_2\right) - 4\left((1+\phi^2)y_1 y_2 + \phi y_1^2 + \phi y_2^2\right)\right] \\
&= \frac{1}{2y_1 y_2}\left[\tau\phi(1-\phi^2) - 3\phi y_1^2 - 3\phi y_2^2 - 2(2+\phi^2)y_1 y_2\right]
\end{aligned}$$

The numerator is positive provided

$$\tau > \frac{3\phi y_1^2 + 3\phi y_2^2 + 2(2+\phi^2)y_1 y_2}{\phi(1-\phi^2)}$$

Except when $\phi = -\rho$ approaches 0 or -1, there exists a large and finite τ for any bounded precisions x_1, x_2. ■

References

[1] Albrecht, James, Axel Anderson, and Susan Vroman, "Search by Committee," *Journal of Economic Theory* **145**, 2010, 13861407.

[2] Alonso, Ricarod, Wouter Dessein, and Niko Matouschek, "Organize to Compete," mimeo, 2011.

[3] Board, Simon, "Monopolistic Group Design with Peer Effects," *Theoretical Economics* **4(1)**, 2009, 89-125.

[4] Burguet, Roberto, and Sákovics, "Imperfect Competition in Auction Designs," *International Economic Review* **40(1)**,1999, 231-247.

[5] Cole, Harold L., George J. Mailath and Andrew Postlewaite, "Social Norms, Savings Behavior, and Growth, Journal of Political Economy 100, 10921125.

[6] Damiano, Ettore, Hao Li and Wing Suen, "Competing for Talents," mimeo, 2010.

[7] Garicano, Luis, "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 2000, **108 (5)**, 874-904.

[8] Guler, Bulent, Fatih Guvenen, and Gianluca Violante, "Joint Search Theory: New Opportunities and New Frictions," forthcoming *Journal of Monetary Economics*, 2011.

[9] Hatfield, John W. and Paul R. Milgrom, "Matching with Contracts," *The American Economic Review* **95(4)**, 2005, 913-935.

[10] Hoppe, Heidrun, Benny Moldovanu, and Emre Ozdenoren "Coarse Matching with Incomplete Information," *Economic Theory*, **47(1)**, 2011, 75-104.

[11] Hoppe, Heidrun, Benny Moldovanu, and Emre Ozdenoren "The Theory of Assortative Matching Based on Costly Signals," *Review of Economic Studies* **76(1)**, 2009, 253-281.

[12] Kaas, Leo and Philipp Kircher, "Efficient Firm Dynamics in a Frictional Labor Market", mimeo, 2011.

[13] Kelso, Alexander S., and Vincent P. Crawford, "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica* **50(6)**, 1982, 1483-1504.

[14] Legros, Patrick, and Patrick Newman, "Beauty is a Beast, Frog is a Prince - Assortative Matching with Nontransferabilities," *Econometrica*, July 2007.

[15] Marshak, Jakob, and Roy Radner, "Economic Theory of Teams," Yale University Press, New Haven, Connecticut, 1972.

[16] Moldovanu, Benny, Aner Sela, and Xianwen Shi (2007), "Contests for status," *Journal of Political Economy*, **115**, 2007, 338363.

[17] Prat, Andrea, "Should a team be homogeneous?", *European Economic Review* **46**, 2002, 1187-1207.

[18] Shi, Shouyong, "Frictional Assignment. I. Efficiency," *Journal of Economic Theory* **98**, 2001, 232-260.

[19] Shimer, Robert, "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy*, **113(5)**, 2005, 996-1025.