Team production, endogenous learning about abilities and career concerns*

Evangelia Chalioti

Yale University†

Abstract

This paper studies career concerns in teams where the support a worker receives depends on fellow team members' efforts *and* abilities. In this setting, by exerting effort and providing support, a worker can influence her own and her teammates' project outputs in order to bias the learning process in her favor. To manipulate the market's assessment, we argue that in equilibrium, a worker has incentives to help or even sabotage her colleagues in order to signal that she is of higher ability. In a multiperiod stationary framework, we show that the stationary level of work effort is above and help effort is below their efficient levels.

Keywords: career concerns, team incentives, incentives to help, incentives to sabotage

Jel codes: D83, J24, M54

*I am indebted to Lambros Pechlivanos for the very useful discussions on this topic. I am also grateful to Dan Bernhardt, Marco Celentani, George Deltas, Johannes Hörner, Patrick Rey, and Spyros Vassilakis for their comments and suggestions. I give special thanks to Costas Arkolakis, Roberto Burguet, Mehmet Ekmekci, Simona Fabrizi, Eiichiro Kazumori, Ting Liu, Chrisovalantou Milliou, Michael Powell, Larry Samuelson, Konstantinos Serfes, Spyridon Skouras, Juuso Välimäki, Anastasios Xepapadeas, and the participants at various conferences and meetings. Financial support from the State Scholarships Foundation in Greece is also acknowledged.

†Contact: Yale University, Department of Economics, 28 Hillhouse Ave., Room 208, New Haven, CT, 06511, USA. Email: evangelia.chalioti@yale.edu, Tel.: +1-217-607-3535

1 Introduction

Modern corporations launch innovative employment practices in the workplace, including teamwork, job rotation, and problem-solving groups, to raise productivity and profits.[1] However, providing team incentives creates challenges. Workers who may be subject to explicit incentives that arise from compensation contracts, may also have implicit incentives arising from career concerns: concerns about the effect of reputation on external labor markets and thus on future remuneration.[2,3] This model analyzes reputation incentives of team workers when their individual performance is observable and depends on the quality of fellow members. This is likely to happen in research collaborations or even on sports teams. For instance, in engineering, researchers work in labs and get valuable feedback from their colleagues while running their own projects. Due to the collaborative nature of this activity, the individual outcome also depends on the contributions and thus the qualities of fellow team members. Two professors sharing or exchanging course materials face the same challenges in terms of the effect of their interactions on their individual performance and their reputations.

This paper studies career concerns in teams where the support a worker receives depends on fellow team members' efforts *and* abilities. We address the questions of how the learning process regarding a worker's ability is shaped by teamwork interactions and how career concerns arise in this setting. A worker's effort and ability are inputs in her teammates' production functions. Thus, by exerting effort and providing support, a worker can influence her own *and* her teammates' performance in order to manipulate the market's assessment of her own ability. We argue that in equilibrium, a worker has incentives either to help or even to sabotage her colleagues, in order to bias the learning process in her favor. The existing literature on career concerns in teams, based on Auriol, Friebel & Pechlivanos (2002), assumes that a teammate's support depends exclusively on her teammates' effort (not abilities). The learning process of a worker's ability is therefore independent of the quality of fellow team members, and her career concerns depend exclusively on her own performance.

We employ Holmström's (1982, 1999) career concerns framework, in which neither the workers nor the market know workers' innate abilities, and both learn from past performances. We consider a simple setting with two agents who work and interact for two periods. Agents consider work and help

[1] The 5th European Working Conditions Survey (2012) reports that the pace of respondents' work depends on direct control of their boss (43% in all workplaces), production or performance targets (47%), and work done by colleagues (45%). The 2011 Workplace Employment Relations Study about British workplaces finds that the incidence of methods for knowledge transmission and teamwork interactions are considerable; i.e., meetings involving all staff (in 80% of workplaces), team briefings (60%), and problem solving groups (14%).

[2] Explicit incentives to perform a job or a task are provided through explicit contractual commitments by a principal. However, implicit incentives arise when principals competing in a labor market have some *ex post discretion* how to respond to an agent's performance. This agent has implicit incentives to change her current effort in order to influence the learning process regarding her ability and thus increase her future payments.

[3] For example, Hann, Roberts, Slaughter & Fielding (2004) argue that star programmers are an order of magnitude more productive than their peers. Thus, they contribute to the development of open sourse software because there is much to signal. Kogut & Metiu (2000) state that many programmers reportedly believe that being a member of the LINUX community "commands a $10,000 premium on annual wages". The Apache project makes a point of recognizing all contributors on its website, http://httpd.apache.org/contributors/#colm.

as two separate tasks and have task-specific cost functions. This is a crucial assumption because it captures the benefits of influencing the teammate's project output. Such incentives will not arise with a total-effort-cost function as in Holmström & Milgrom (1991), where there are negative externalities between the tasks.[4] A worker's "project" output is observable and linear in her own innate ability and "work" effort, her teammate's support, and a transitory shock. The support a teammate provides also depends on her own "help" effort *and* ability; i.e., the teammate's ability matters for an agent's performance. Agents' abilities and the transitory shocks are independently and normally distributed. Additionally, we consider different degrees of *initiated* and *received* teamwork interactions; i.e., the fraction of a teammate's support that is appropriated by an agent may differ from the fraction of an agent's help that increases a teammate's production.

The dependence of future rewards on past performance plays a key role in agents' labor supply. The market draws inferences about the levels of agents' abilities via current project outputs. Since labor is a substitute for ability, an agent can influence the learning process in her favor by distorting both her efforts upwards.[5] Because both teammates' abilities are inputs in the production function, an agent's project output as a signal of her own ability is noisier, thus complicating inferences. However, her colleague's output also conveys information.

By exerting work and help effort, an agent can influence both performance measures and manipulate market perceptions. If initiated interactions are strong enough relative to received interactions, resulting in an agent's support having a great impact on her colleague's production, the market attributes high performance by a teammate to the agent's ability and revises its assessment about ability upwards. In this case, we argue that an agent has the incentive to work *and* help her colleague in order to build up her reputation. The opposite occurs if received interactions are strong enough relative to initiated interactions. High performance by a teammate is attributed to the teammate's ability. This causes the market to put a negative weight on that performance when forecasting an agent's ability. In this case, an agent's help will increase the teammate's performance further, which biases the learning process against her. Thus, an agent now has incentives to sabotage her colleague. She can induce an upward revision of her own ability only by destroying some part of her teammate's production.

This analysis shows that what matters for career concerns is how many components of the production and learning process an agent can affect in order to shape the market's assessment. An agent cashes in a reputational bonus that increases with effort exertion and support provision or sabotage. Holmström (1982) studies career concerns when there are no interactions, while Auriol et al. (2002)

[4] With total-effort-cost functions, as an agent increases the effort directed towards affecting a teammate's project output, the marginal cost of effort to improve her own performance will also increase. Thus, providing support crowds out effort allocated to an agent's own task.

[5] Empirical studies find evidence of the existence of career concerns for professionals (Gibbons & Murphy (1992)), and for economists (Coupe, Smeets & Warzynski (2006)); i.e., past performance and the probability of promotion are positively related. Also, the sensitivity of promotion to performance declines with experience, indicating the presence of a learning process. Borland (1992) provides a survey.

assume that the support an agent receives depends exclusively on her colleague's effort. In their model, by looking at a teammate's performance, the market cannot draw any additional information about an agent's ability. The process of inference about each teammate's ability is independent, and the quality of fellow team members has no effect on an agent's reputation incentives.

This paper also investigates Fama's (1980) conjecture that career concerns induce agents to behave efficiently. Holmström (1999) formalizes this idea by considering a "stationary" single-agent model where ability is not fixed but fluctuates over time, thereby preventing the market from fully learning its level. He states that without discounting, Fama's result is correct: agents exert the efficient level of work effort. We argue that in a multi-agent model where there are teamwork interactions and the quality of fellow team members matters for an agent's decisions, this result does not necessarily hold. In particular, if initiated interactions occur, despite the absence of discounting, the stationary work effort is higher and help effort lower than their efficient levels. Because we add noise to the learning process, both performance measures become more vague. An agent can more effectively shape the market's assessments by increasing her own project output, thus distorting the work effort upwards, and the help effort downwards. The balance between the reputation incentives in a stationary model indicates that an agent is oriented towards focusing on tasks that increase her own project output, dragging her attention away from helping or sabotaging her teammate. In a stationary equilibrium, career concerns induce an agent to over-provide work effort.

An agent's stationary effort levels are efficient only in two cases, provided there is no discounting. On the one hand, this happens as long as an agent's ability is not an input in her teammate's production function as in the settings of Holmström (1999) and Auriol et al. (2002), although received interactions may occur. A teammate's output as a performance measure should not convey any information about an agent's ability and hence has no effect on reputation incentives. In this case, the supplied work effort is efficient regardless of the intensity of received interactions or of how noisy the signal of an agent's performance is about her own ability. Exerting zero help effort is also efficient. On the other hand, efficient effort levels are obtained in a stationary setting as long as both initiated and received interactions are perfect, implying that an agent's work and help efforts need to be equally productive.

This paper contributes to the existing literature on career concerns in teams where either the ratchet effect or sabotage incentives arise. Lazear (1989) considers sabotage incentives in tournaments. In Auriol et al. (2002), explicit contracts are also provided, and the source of sabotage incentives is a lack of commitment by the principal. In a two-agent model, Meyer & Vickers (1997) use Holmström's (1999) production function where an agent's effort and ability matters only for her own outcome. Thus, an agent cannot influence another's production. However, the learning process depends on whether agents' abilities are correlated. They argue that on account of a positive externality, each agent free-rides on the effort of the other to enhance reputation. Due to free-riding, reputation incentives are weakened, and the ratchet effect arises. Agents have a decreasing willingness to work. In our setting, the teammates' innate characteristics are independent. However, due to teamwork interactions, an agent has incentives to take action in order to affect her teammates' performance. Even in the absence of explicit contracts, an agent exerts effort either to help or sabotage her teammate by destroying some part of her production. Incentives to sabotage arise when the market puts a negative weight on a teammate's performance when predicting an agent's ability.

The literature on moral hazard problems remains narrow in its focus on whether market forces alone suffice to remove them. Fama (1980) states that explicit contracts are unnecessary to solve principal-agent conflicts. The market already provides efficient implicit contracts, inducing the "right" level of labor supply. Holmström (1999) shows that risk-aversion and discounting place limitations on the market's ability to engender adequate incentives. However, if these limitations are lifted in a stationary model, agents exert efficient effort levels. Bar-Isaac & Hörner (2014) consider an agent who has different abilities - specialized and generalized abilities - to perform two tasks. They compare the value of specializing with acting as a generalist in an infinite-horizon model and find that without discounting, the stationary level of effort is also efficient. Bonatti & Hörner (2014) consider a dynamic framework with exponential learning. In our model, we show that when teammates' abilities affect their reputation incentives, the stationary levels of efforts on both tasks are inefficient. While the stationary work effort is higher, the help effort is lower than its efficient level.

This paper is also tied to the literature on team incentives when the degree of visibility of an agent's characteristics is an issue. In team production models, the market only observes the team output and uses this (single) measure to infer the level of workers' abilities. Ortega (2003) examines the effect of power allocation within the firm on workers' career concerns. Since power confers visibility, as one agent becomes more visible, the visibility of her colleague must decline. He argues that uneven allocation of authority is optimal. Jeon (1996) shows the optimality of equal sharing of team output among workers as well as the advantage of mixing young and old workers in a team. Bar-Isaac (2007) analyzes workers' incentives to work for their own reputations when young but for their firms' reputation when old. Arya & Mittendorf (2011) examine the desirability of aggregate performance measures in models with reputation incentives. They assume that an agent can impact multiple dimensions of a firm's operation, and that the output of each operation depends on her own effort and ability. Effort can influence all signals to varying degrees. There are no teamwork interactions, and the process of inference of an agent's ability depends only on her own efforts. They argue that an aggregate signal of the outputs of these operations can improve efficiency. In a single agent model, Dewatripont, Jewitt & Tirole (2000) consider multitasking and claim that increasing the number of tasks reduces the total effort because performance becomes noisier. Dewatripont, Jewitt & Tirole (1999) use a production function in which an agent's effort and ability are multiplicative and argue that market expectations about focus on a task matter, not the observability of tasks as in an additive case. They also examine incentives under a "fuzzy mission", in which case the market is ignorant about the allocation of an agent's effort across tasks. In a different setting, Effinger &

Polborn (2001) assume that an agent is most valuable if she is the only smart agent. If this value is sufficiently large, the other expert opposes her predecessor's report. 'Antiherding' may result.

In our two-agent model, individual project outputs are observable (separate signals) and subject to market shocks that are independent of each other. Teammates' abilities are also uncorrelated. The degree of visibility of agents' abilities changes with teamwork interactions that also make individual production noisier. This happens because a teammate's ability affects an agent's project output. However, a teammate's performance also conveys information about an agent's ability, and it is likely that the signals will be jointly more informative. In our model, it is not the amount of available information about teammates' abilities per se that drives optimal reputation incentives, but how agents' project outputs are related. An agent's attempts to shape the market's assessment may induce her to help or sabotage her teammate, even in the absence of explicit motivation.[6,7]

The paper is organized as follows. Section 2 presents the model. It discusses the process of learning about abilities and the effect of teamwork interactions on the amount of available information. Section 3 derives teammates' reputation incentives. The optimal incentives to help or sabotage are analyzed. We also discuss the reputation incentives when market shocks are correlated. In section 4, we consider a multiperiod model and focus on the stationary level of labor supply. Section 5 concludes.

2 The model

In this section, we assume that there are two effort-averse agents 1 and 2, indexed by i and j where $i \neq j$. Agents are also rational and forward-looking. Employment lasts for two periods indexed by $t = \{1, 2\}$, and at each period, each agent carries out her own project.

2.1 Production technology

Agents are engaged in a stochastic production process. At each period t, agent i's "project" output, z_t^i, depends on her own innate ability, θ^i, her "work" effort, e_t^i, and a transitory shock, ε_t^i. In addition, z_t^i depends on the teammate's support, $\theta^j + a_t^j$, weighted by a parameter h_j, where $0 \leq h_j \leq 1$:

$$z_t^i = \theta^i + e_t^i + h_j \left(\theta^j + a_t^j\right) + \varepsilon_t^i. \tag{1}$$

The teammate's innate ability, θ^j, and her "help" effort, a_t^j, increase agent i's project output in an additive way. Thus, each agent exerts *work* effort to accomplish her own project as well as *help* effort

[6] Heterogeneous teams in terms of seniority or learning by doing are beyond the scope of this analysis.

[7] Milgrom & Oster (1987) study the role of a worker's visibility in the job market: the abilities of visible workers are known to all parties while those of invisible ones are concealed by an employer from other potential employers. Mukherjee (2008) examines a firm's decision to disclose information about its workers' productivity.

to improve her colleague's performance. The help effort can even be negative, indicating that an agent may choose to sabotage her colleague instead of helping her.[8]

When agent i enters the labor market, her ability is not known with certainty. However, all parties share the common prior that abilities are independently and identically distributed, where θ^i is drawn from a normal distribution with mean m_1^i and variance σ_i^2. Prendergast & Topel (1996) consider θ^i as the fit between the agent and her job that is contingent on some systemic variation, (symmetrically) unknown to all parties at each stage.[9] The parameter h_j measures the degree of *received* teamwork interactions - the fraction of agent j's support that is appropriated by agent i - and h_i indicates the degree of *initiated* interactions - the fraction of agent i's support that contributes to agent j's production. These parameters may differ. They are also exogenous and lie in $[0, 1]$, implying that teamwork interactions are value-creating. Their intensity also depends on the characteristics of the technology used by each agent or, for instance, the degree of tacit knowledge required in production. The fact that h_i and h_j are less than one reflects the imperfect nature of teamwork interactions: providing help to a fellow member of the team is (somewhat) less productive than putting effort into one's own task. The random terms ε_t^i, ε_t^j are also independently and normally distributed, across agents and periods, with zero mean and variance σ_ε^2.

2.2 Learning process

In multi-agent career concerns models with uncorrelated shocks, the market updates from an agent's past performance in order to infer the level of her ability. In our model, teamwork interactions occur and support depends on the ability of the fellow member. Since the unknown θ^j enters agent i's production function, agent i's project output, z_t^i, as a signal of her ability, θ^i, becomes noisier. Teamwork interactions weaken the link between an agent's performance and her ability, implying that this relationship becomes less autonomous and accountable. However, θ^i is also an input in the teammate's production function. Therefore, z_1^j also conveys information about agent i's ability. The market has two performance measures from which to draw inference about an agent's ability.

In Holmström's (1999) model where $z_t^i = \theta^i + e_t^i + \varepsilon_t^i$, there are no interactions, while in a two agent version of Auriol et al. (2002) model, agent i's production function is $z_t^i = \theta^i + e_t^i + h_j a_t^j + \varepsilon_t^i$. Thus, the support an agent receives depends exclusively on her colleague's effort. There is no link between θ^j and z_t^i, $corr\left(\theta^i, z_1^j\right) = 0$. The processes of inference of θ^i and θ^j are completely independent.

Following DeGroot (1970), Lemma 1 specifies the mean and variance of the conditional distribution of abilities after the realizations of z_1^i and z_1^j. All parties (the market and the two teammates) observe the outputs of both projects realized at the end of the first period. With $\widehat{e}_1^i$ and $\widehat{a}_1^i$, we denote the market conjectures about agent i's first period efforts.

[8] The price of the outputs is normalized to one, and the scale of production is identical in all t periods.

[9] Laffont & Tirole (1988), among others, analyze the optimal incentives when an agent has private information about her own ability before she goes to the market.

Lemma 1 (Conditional distribution of abilities) *Given the realization of the first-period project outputs, z_1^i and z_1^j, the mean and variance of the conditional distribution of θ^i in period 2 are*

$$\begin{aligned} m_2^i &\equiv E\left\{\theta^i \mid z_1^i, z_1^j\right\} = \mu_1^i m_1^i + \rho_1^{ii}\left(z_1^i - \widehat{e}_1^i - h_j \widehat{a}_1^j - h_j m_1^j\right) + \rho_1^{ij}\left(z_1^j - \widehat{e}_1^j - m_1^j - h_i \widehat{a}_1^i\right), \\ \sigma_{i,2}^2 &\equiv var\left\{\theta^i \mid z_1^i, z_1^j\right\} = \sigma_i^2\left(1 - \rho_1^{ii} - h_i \rho_1^{ij}\right), \end{aligned}$$

where $\mu_1^i \equiv 1 - \rho_1^{ii} - h_i \rho_1^{ij}$. The conditional correlation coefficients of z_1^i and z_1^j are, respectively,

$$\begin{aligned} \rho_1^{ii} &\equiv corr\left(\theta^i, z_1^i \mid z_1^j\right) = \frac{\sigma_i^2}{\lambda_1}\left[\sigma_\varepsilon^2 + \left(1 - h_i h_j\right)\sigma_j^2\right], \\ \rho_1^{ij} &\equiv corr\left(\theta^i, z_1^j \mid z_1^i\right) = \frac{\sigma_i^2}{\lambda_1}\left[h_i \sigma_\varepsilon^2 - \left(1 - h_i h_j\right) h_j \sigma_j^2\right], \end{aligned}$$

where $\lambda_1 \equiv \sigma_\varepsilon^4 + \left(1 - h_i h_j\right)^2 \sigma_i^2 \sigma_j^2 + \sigma_\varepsilon^2\left[\left(1 + h_i^2\right)\sigma_i^2 + \left(1 + h_j^2\right)\sigma_j^2\right]$ for all h_i and h_j.

Proof. In appendix $(A.1)$. ■

Provided that all parties have rational expectations, the equilibrium conjectures must be correct: $\widehat{e}_1^i = e_1^{i*}$ and $\widehat{a}_1^i = a_1^{i*}$. There is no off-equilibrium realization of observables because of the presence of noise. Each agent is compelled to exert the equilibrium effort levels that are expected of her, since working less will bias the learning process against her. Remark 1 highlights the informativeness of the signals about an agent's ability.[10]

Remark 1 (Informativeness of signals) *(a) Given z_1^j, the conditional correlation between agent i's ability, θ^i, and her own project output, z_1^i, is always positive: $\rho_1^{ii} > 0$ for all h_i and h_j.*

(b) Given z_1^i, the conditional correlation between agent i's ability, θ^i, and her teammate's project output, z_1^j, is positive as long as initiated interactions are substantial:

$$\rho_1^{ij} > 0 \text{ if and only if } h_i > \frac{h_j \sigma_j^2}{\sigma_\varepsilon^2 + h_j^2 \sigma_j^2}.$$

The coefficient ρ_1^{ii} represents the correlation of agent i's ability and her own project output, given a teammate's performance; i.e., the linear dependence between θ^i and z_1^i, given z_1^j.[11] This correlation coefficient is always positive, $\rho_1^{ii} > 0$, because $\frac{cov\left(\theta^i, z_1^i\right)}{cov\left(\theta^i, z_1^j\right)} > \frac{cov\left(z_1^i, z_1^j\right)}{var\left(z_1^j\right)} \Leftrightarrow \frac{1}{h_i} > \frac{h_j \sigma_j^2}{\sigma_\varepsilon^2 + \sigma_j^2}$ for all h_i and h_j.

[10] If the estimate of θ^i is based only on z_1^i, we have $E\left\{\theta^i \mid z_1^i\right\} = (1-\xi) m_1^i + \xi\left(z_1^i - \widehat{e}_1^i - h_j \widehat{a}_1^j - h_j m_1^j\right)$ and $Var\left\{\theta^i \mid z_1^i\right\} = \sigma_i^2(1-\xi)$, where $\xi \equiv \sigma_i^2\left[\sigma_i^2 + h_j^2 \sigma_j^2 + \sigma_\varepsilon^2\right]^{-1}$. ξ exceeds ρ_1^{ii}, $\xi \geq \rho_1^{ii}$, implying that the market puts a lower weight on z_1^i to perceive the level of θ^i if another signal is also available. However, the two signals are jointly more informative, allowing for a better estimate: $\rho_1^{ii} + h_i \rho_1^{ij} \geq \xi$ for all h_i and h_j.

[11] The correlation coefficients of the unconditional distribution of θ^i are $corr\left(\theta^i, z_t^i\right) = \sigma_i\left[\sigma_i^2 + h_j^2 \sigma_j^2 + \sigma_\varepsilon^2\right]^{-\frac{1}{2}}$ and $corr\left(\theta^i, z_t^j\right) = h_i corr\left(\theta^i, z_t^i\right)$. Both are positive.

Thus, given the realization of her colleague's project output, an agent's high "own" performance signals high "own" ability and vice versa. If there are no teamwork interactions, as in Holmström (1999), or if support depends only on teammate's effort, as in Auriol et al. (2002), the variance of agent i's ability after the observation of z_1^i, $var\left(\theta^i \mid z_1^i\right)$, is independent of σ_j^2 and equal to $\frac{\sigma_\varepsilon^2 \sigma_i^2}{\sigma_\varepsilon^2+\sigma_i^2}$. The correlation of θ^i and z_1^j is also zero.

In our model, z_1^j conveys information about θ^i, but the sign of the (conditional) correlation coefficient ρ_1^{ij} is less straightforward. The sign of ρ_1^{ij} depends on the *relative* intensity of the degrees of teamwork interactions (rather than on their absolute values) as well as on the variance of θ^j and ε_1^i; these two inputs are apart from agent i's characteristics and beyond her control. A positive ρ_1^{ij} requires $\frac{cov\left(\theta^i, z_1^j\right)}{cov\left(\theta^i, z_1^i\right)} > \frac{cov\left(z_1^i, z_1^j\right)}{var\left(z_1^i\right)} \Leftrightarrow h_i > \frac{h_j \sigma_j^2}{\sigma_\varepsilon^2+h_j^2 \sigma_j^2}$. Initiated interactions must be strong enough so that agent j's performance is sensitive to θ^i, while received interactions, h_j, must be weak (z_1^i must not be sensitive to θ^j). If this is the case, both signals are more likely to reflect the level of θ^i. Thus, given z_1^i, higher z_1^j is "good news" for agent i's ability. The market perceives that a high z_1^j is due to agent i's high ability and updates its assessment upwards. In the polar case where $h_j = 0$, ρ_1^{ij} is positive for all h_i.

The opposite occurs if received interactions, h_j, are large enough while initiated interactions, h_i, are small. In this case, as z_1^j increases, $E\left[\theta^j \mid z_1^i, z_1^j\right]$ will increase for a given fixed z_1^i. Hence, a larger proportion of this z_1^i will also be attributed to θ^j rather than θ^i, so that $E\left[\theta^i \mid z_1^i, z_1^j\right]$ will decrease. In particular, if h_i is small and the variance of θ^j is large enough, it is more likely that both performance measures indicate the level of θ^j. Thus, if both agents perform well, the market attributes these outcomes to high θ^j, causing the estimate of θ^i to be updated downwards. Agent j is now perceived as the high-quality member of the team. In the polar case where $h_i = 0$, θ^i does not contribute to agent j's project output at all. However, the market still uses this performance measure to draw valuable information about θ^j (and indirectly about θ^i). Under these conditions, given z_1^i, the market always puts a negative weight on z_1^j to assess θ^i: if $h_i = 0$, $\rho_1^{ij} < 0$ for all h_j.

To obtain better insight, we also examine how the variances of θ^i and θ^j affect how much weight the market puts on outputs in estimating teammates' abilities. In particular, we have: (i) $\frac{\partial \rho_1^{ii}}{\partial \sigma_i^2} > 0$ for all h_i and h_j; (ii) $\frac{\partial \rho_1^{ij}}{\partial \sigma_i^2} > 0$ if and only if $\rho_1^{ij} > 0$. As long as initiated interactions are strong enough relative to the degree of received interactions so that a higher z_1^i or z_1^j is attributed to a higher θ^i, an increase in the variance of agent i's ability, σ_i^2, will trigger the market to rely more on both signals. The market will be willing and able to learn more about θ^i. On the other hand, we have: (i) $\frac{\partial \rho_1^{ii}}{\partial \sigma_j^2} < 0$ if and only if $\rho_1^{ji} \equiv corr\left(\theta^j, z_1^i \mid z_1^j\right) > 0$ (see Lemma 1); (ii) $\frac{\partial \rho_1^{ij}}{\partial \sigma_j^2} < 0$ for all h_i and h_j. Thus, for strong received interactions (large h_j), a teammate's ability is key for an agent's performance and $\rho_1^{ji} > 0$. In this case, as σ_j^2 increases, z_1^i is more likely to reflect the level of θ^j, while as a signal of θ^i, it becomes more vague, $\frac{\partial \rho_1^{ii}}{\partial \sigma_j^2} < 0$. The opposite occurs when received interactions are weak (small h_j). To interpret this case, let us assume $h_j = 0$, implying that agent i's output is now independent of θ^j and $\rho_1^{ji} < 0$ for all h_i. The negative sign of ρ_1^{ji} indicates that given z_1^j, a higher z_1^i is "bad news" for agent j. The market attributes a higher z_1^i to a higher θ^i. An increase in σ_j^2 now works in favor of agent i and induces the market to rely more on an agent's project output, z_1^i, to perceive the level of her ability, θ^i. We have $\frac{\partial \rho_1^{ii}}{\partial \sigma_j^2} > 0$.

The variance of agents' abilities and the degrees of teamwork interactions also affect the "total" amount of available information in the market. Learning about abilities is captured by a decrease in the variance of the posterior estimate of the θs, and thus by an increase in

$$\rho_1^{ii} + h_i \rho_1^{ij} = \frac{\sigma_i^2}{\lambda_1} \left[\left(1 + h_i^2\right) \sigma_\varepsilon^2 + \left(1 - h_i h_j\right)^2 \sigma_j^2 \right],$$

where λ_1 is given in Lemma 1. The market can obtain a better estimate of agent i's ability as σ_i^2 increases and σ_j^2, σ_ε^2 decrease. Remark 2 shows that h_i also increases learning as long as ρ_1^{ij} is positive.[12]

Remark 2 (Information extraction & teamwork interactions) *Given z_1^i and z_1^j, the conditional variance of θ^i: (i) decreases with initiated interactions h_i, $\frac{\partial \left(\rho_1^{ii} + h_i \rho_1^{ij}\right)}{\partial h_i} > 0$, if and only if such interactions are strong enough so that $\rho_1^{ij} > 0$; (ii) increases with received interactions h_j, $\frac{\partial \left(\rho_1^{ii} + h_i \rho_1^{ij}\right)}{\partial h_j} < 0$, for all h_i.*

[Figures 1 are about here.]

As h_j increases, the joint signal $\left(z_1^i, z_1^j\right)$ about θ^i becomes more vague. The market finds it harder to disentangle the contribution of agent i's ability to both teammates' project outputs and the information conveyed by z_1^i and z_1^j about θ^i is less pronounced. The market relies less on the performance measures to assess agent i's ability as the impact of θ^j on z_1^i increases. Similarly, as long as $\rho_1^{ij} < 0$, a small increase in h_i prevents the market from learning, since it makes the joint signal $\left(z_1^i, z_1^j\right)$ about θ^i to reveal less information. Nevertheless, if h_i exceeds a threshold such that ρ_1^{ij} becomes positive, the conditional variance of θ^i decreases. In this regime, ρ_1^{ii} decreases with h_i. However, as agent i's help matters more for agent j's performance, z_1^j will become more informative. The effect of h_i on z_1^j exceeds that of z_1^i, making both signals jointly "speak" more about ability. Higher h_i helps the market to learn, resulting in better estimates of θ^i.

2.3 Agents' preferences and objectives

In carrying out her own task and providing support to her teammate, agent i incurs task specific disutility. The cost functions of work effort and help effort are $\psi\left(e_t^i\right)$ and $\psi\left(a_t^i\right)$, respectively. The

[12] We have $\frac{\partial \left(\rho_1^{ii} + h_i \rho_1^{ij}\right)}{\partial h_j} = -\frac{2\sigma_i^2 \sigma_j^2 \sigma_\varepsilon^2}{\lambda_1^2} \left(h_i + h_j\right) \left[\sigma_\varepsilon^2 + \left(1 - h_i h_j\right) \sigma_j^2\right] < 0$ for all h_i, σ_i^2, σ_j^2 and σ_ε^2. The derivative with respect to h_i gives $\frac{\partial \left(\rho_1^{ii} + h_i \rho_1^{ij}\right)}{\partial h_i} = \frac{2\sigma_i^2 \sigma_\varepsilon^2}{\lambda_1^2} \left[\sigma_\varepsilon^2 + \left(1 + h_j^2\right) \sigma_j^2\right] \left[h_i \sigma_\varepsilon^2 - \left(1 - h_i h_j\right) h_j \sigma_j^2\right]$. Note that $sign\left\{h_i \sigma_\varepsilon^2 - \left(1 - h_i h_j\right) h_j \sigma_j^2\right\} = sign\left\{\rho_1^{ij}\right\}$.

function $\psi(.)$ is twice continuously differentiable and convex, implying that there are diminishing returns to scale in the production process. We also assume that $\psi'(0) = 0$, $\lim_{e_t^i \to \infty} \psi'(e_t^i) = \infty$ and $\lim_{a_t^i \to \infty} \psi'(a_t^i) = \infty$. Note that the agents incur a cost for exerting negative help effort which is also clearly higher than that associated with zero effort.

Task-specific cost functions are used in multi-agent models, such as Auriol et al. (2002) and Itoh (1992). However, they are in stark contrast to other multitask models based on Holmström & Milgrom (1991) which assume $\psi(e_t^i + a_t^i)$. In the latter models, the cross-partial derivatives with respect to two efforts are positive. That is, tasks are (perfect) substitutes in an agent's cost function. These total-effort-cost functions introduce negative externalities between a given agent's tasks. As an agent increases the effort devoted to one task, the marginal cost of effort to the other task will grow. Thus, providing support to a teammate would be costly to an agent, and it crowds out effort directed to her own task, thereby decreasing her own project output. Agents care about the sum of effort exerted. The allocation of effort between the tasks depends on the relative benefits an agent derives from these two tasks. In fact, the agent must equate the marginal return to effort in both tasks. These models focus on the allocation of an agent's "attention" between the tasks.

In our model with task-specific cost functions, disaggregated information, and separation of tasks - work effort and help effort are inputs in different production functions - benefits of providing help or sabotage emerge. Allocating a given total effort to both tasks entails lower disutility. The cross-partial of the cost function is zero, hence the cost of exerting effort to perform a given task is independent of the other task. An agent can focus on eliciting effort to affect her teammate's project output without simultaneously having to consider technologically founded externalities. For those engineers who work in the same lab but in different projects, or the professors who share class material, putting effort into a task does not require diverting effort away from the other task.

Agent i is risk neutral and receives the reward w_t^i. She derives utility

$$U^i = \sum_{t=1}^{2} \left[w_t^i - \psi\left(e_t^i\right) - \psi\left(a_t^i\right) \right]. \tag{2}$$

This function is additively separable across periods, implying that agents behave as if they have access to perfect capital markets. They also do not discount the future.

Agent i's reward is determined in equilibrium and depends on the available information conveyed by both agents' past performance measures. A competitive market will set

$$w_t^i = (1 + h_i) E\left\{\theta^i \mid z_{t-1}^i, z_{t-1}^j\right\} + \widehat{e}_t^i + h_i \widehat{a}_t^i \equiv \widetilde{\theta}_t^i. \tag{3}$$

Each agent receives a fixed payment equal to the reputational bonus she can claim for her contribution to both teammates' project outputs.[13,14] This bonus is the total rent an agent can get by exerting effort *and* providing support. Given the available information, her payment increases with an upward revision of the market's estimate of her own ability.

3 Reputation incentives

We now solve the two-period game and derive the teammates' optimal efforts. The conventional wisdom in career concerns models is that an agent works harder at the beginning of her career in order to improve her own performance and thus manipulate market assessment about her ability. In our multi-agent model where an agent's ability inserts into a fellow member's production function, we show that additional reputation incentives arise. To influence the learning process, under certain conditions, an agent has incentives either to help or to sabotage her colleague. Then, we perform this analysis when the output shocks are correlated.

3.1 Work and help effort

In period 2, agent i receives $w_2^i = (1+h_i)\, E\left\{\theta^i \mid z_1^i, z_1^j\right\} + \widehat{e}_2^i + h_i \widehat{a}_2^i$. However, this reward does not depend on her current actions. There are no career concerns and thus she exerts zero effort: $e_2^{i*} = 0$ and $a_2^{i*} = 0$. In period 1, agent i maximizes her current and future utility:

$$E\left\{w_1^i\right\} - \psi\left(e_1^i\right) - \psi\left(a_1^i\right) + E\left\{w_2^i \mid z_1^i, z_1^j\right\} - \psi\left(e_2^{i*}\right) - \psi\left(a_2^{i*}\right).$$

The reward w_1^i is independent of e_1^i and a_1^i because $z_0^i = \emptyset$ and $E\left\{\theta^i\right\} = m_1^i$. Given also that $\psi\left(e_2^{i*}\right)$ and $\psi\left(a_2^{i*}\right)$ are zero, agent i's problem reduces to maximizing

$$-\psi\left(e_1^i\right) - \psi\left(a_1^i\right) + (1+h_i)\, E\left\{\theta^i \mid z_1^i, z_1^j\right\}.$$

Career concerns arise because the levels of current project outputs, z_1^i and z_1^j, affect the reputational bonus (wage) in the second period. As long as ability is unknown, there are returns to supplying labor, since past performance will influence the markets' perception of θ^i. Labor is a substitute for ability. Thus, by increasing labor supply, an agent can potentially bias the process of inference in her favor. Proposition 1 presents the optimal efforts.[15]

---

[13] Recall that $t = \{1, 2\}$. If employment lasts for T periods, where $T > 2$, the market's perceptions of abilities will depend on all past performances. The reputational bonus will be $w_t^i = (1+h_i)\, E\left\{\theta^i \mid z_1^i, z_1^j, ..., z_{t-1}^i, z_{t-1}^j\right\} + \widehat{e}_t^i + h_i \widehat{a}_t^i$.

[14] The principal maximizes the sum of outputs minus the payments to the agents. However, the competition among them will drive their profits down to zero, and each agent will receive her reputational bonus.

[15] One can consider the normalization $z_t^i = (1-h_i)\left(\theta^i + e_t^i\right) + h_j\left(\theta^j + a_t^j\right) + \varepsilon_t^i$, where h_i and h_j lie in $\left[0, \frac{1}{2}\right]$. The reputational bonus now is $E\left\{\theta^i \mid z_{t-1}^i, z_{t-1}^j\right\} + (1-h_i)\, \widehat{e}_t^i + h_i \widehat{a}_t^i$. This normalization serves to guarantee that agents tend to put effort into both tasks exactly in order to manipulate the market's perception rather than because

Proposition 1 (Career concerns) *In equilibrium, agent i has reputation (implicit) incentives to work, thereby increasing her own project output, as well as incentives to help or sabotage her teammate's production:*

$$\psi'\left(e_1^{i*}\right) = (1+h_i)\,\rho_1^{ii} \text{ and } \psi'\left(a_1^{i*}\right) = \underbrace{(1+h_i)\,h_i\rho_1^{ij}}_{\text{help or sabotage}},$$

where ρ_1^{ii} and ρ_1^{ij} are given in Lemma 1.

The optimal efforts are contingent on the measures the market uses to draw inferences about ability. In line with the literature, career concerns depend on the weight the market puts on outputs when estimating ability. However, we argue that what also matters for career concerns is how many components of the production process and the learning process an agent can influence in order to manipulate the market's perception in her favor, and how many "pieces" of future remuneration depend on an agent's current actions. By exerting work effort in the current period and providing support, an agent affects both teammates' performance measures, z_1^i and z_1^j, in order to induce an upward revision of the market's estimate of her own ability. Thus, an agent has two tools available to shape the market's assessment. In Auriol et al. (2002) where the support an agent receives depends only on her teammate's effort (not on her ability), and the market shocks are not correlated, market assessment of agent i's ability only depends on her own performance. Thus, providing support has no effect on an agent's future remuneration. Agent i's utility-maximizing help effort is zero. Her work effort is independent of the degrees of teamwork interactions and equal to $\frac{\sigma_i^2}{\sigma_i^2+\sigma_\varepsilon^2}$.

In our model, additional reputation incentives arise. Agent i exerts effort to increase her future remuneration by $M_1^{ii} \equiv (1+h_i)\,\rho_1^{ii}$ through her work and by $M_1^{ij} \equiv (1+h_i)\,h_i\rho_1^{ij}$ through help or sabotage. In particular, if initiated interactions are strong enough (large h_i) relative to the degree of received interactions h_j so that $\rho_1^{ij} > 0$, agent i anticipates that good teammate performance (high z_1^j) will engender an upward revision of the market's estimate of her own ability, θ^i. Therefore, she has additional incentives to help her colleague, $M_1^{ij} > 0$. However, for a small h_i so that $\rho_1^{ij} < 0$, such reputation incentives are reversed, $M_1^{ij} < 0$. If initiated interactions are weak, a higher z_1^j is attributed to θ^j, and the market updates its assessment of θ^i downwards. Thus, by helping a teammate to further increase her project output, agent i will induce market inferences to be revised against her. Instead, bad performance by her teammate will be a good signal of her own ability. A decrease in z_1^j will increase agent i's reputation so that she now has incentives to sabotage her colleague. We can interpret negative effort as hiding, stealing, or even destroying some part of a teammate's project output. In the polar case where "one-way" teamwork interactions occur - $h_i > 0$ while $h_j = 0$ - agent i always has incentives to help.

---

the size of the "pie" increases by helping a teammate. Qualitatively, *all* our results also hold in this setting. The optimal work effort will satisfy $\psi'\left(e_1^{i*}\right) = (1-h_i)\,corr\left(\theta^i, z_1^i \mid z_1^j\right)$, and the optimal help effort is given by $\psi'\left(a_1^{i*}\right) = h_i corr\left(\theta^i, z_1^j \mid z_1^i\right)$, which can be either positive or negative.

This analysis boils down to the following: agent i has stronger reputation incentives as more pieces of information during the learning process depend on current actions, and as the impact of the estimate of θ^i on future remuneration increases. An agent always has incentives to exert work effort in order to increase her own project output. As long as a teammate's performance is sensitive to the agent's own ability so that $\rho_1^{ij} > 0$, we argue that this agent has additional incentives to help her colleague in order to build up her own reputation. In contrast, if the impact of agent i's support to her teammate is insignificant so that the market puts a negative weight on her teammate's output to estimate her own ability, $\rho_1^{ij} < 0$, an increase in z_1^j will bias the learning process against her. Thus, incentives to sabotage her teammate arise.

[Figures 2 are about here.]

We can also compare the teammates' effort decisions, given the differences in the variance of their abilities.[16] In particular, suppose that received and initiated interactions are identical, $h_i = h_j$, and the coefficients ρ_1^{ij} and ρ_1^{ji} are positive. The agent with the higher variance of ability, say $\sigma_i^2 > \sigma_j^2$, exerts more work effort, $\psi'\left(e_1^{i*}\right) > \psi'\left(e_1^{j*}\right)$, and help effort, $\psi'\left(a_1^{i*}\right) > \psi'\left(a_1^{j*}\right)$. Due to higher σ_i^2, the market is able to draw additional information about θ^i, and agent i's attempts to manipulate market perception are more effective. More generally, as long as the interactions initiated by the agent with the higher variance, $\sigma_i^2 > \sigma_j^2$, are large enough relative to the intensity of received interactions, this agent exerts more work and help effort than her colleague. The market anticipates that this agent's efforts are key determinants of both project outputs and relies more on both signals that likely reflect the level of her ability.[17]

3.2 Correlated output shocks

We analyze the reputation incentives when the transitory shocks, ε_t^i and ε_t^j, are correlated. Suppose that $\phi \equiv \frac{cov\left(\varepsilon_t^i, \varepsilon_t^j\right)}{\sigma_\varepsilon^2}$ denotes the correlation coefficient, where $|\phi| < 1$. The type of correlation (positive or negative) may depend on whether the team members use similar or different technologies in production.[18] Now, there are two "forms" of correlation between the team members' project outputs: one due to teamwork interactions and the other due to the correlation of the random terms.

[16] We have $\psi'\left(e_1^{i*}\right) > \psi'\left(e_1^{j*}\right)$ if and only if $\sigma_i^2 > \frac{(1+h_j)\sigma_j^2\sigma_\varepsilon^2}{(1+h_i)\sigma_\varepsilon^2+(h_i-h_j)(1-h_ih_j)\sigma_j^2}$. Additionally, $\psi'\left(a_1^{i*}\right) > \psi'\left(a_1^{j*}\right)$ if and only if $\sigma_i^2 > \frac{h_j^2}{h_i}\frac{(1+h_j)\sigma_j^2\sigma_\varepsilon^2}{h_i(1+h_i)\sigma_\varepsilon^2+h_j(h_j-h_i)(1-h_ih_j)\sigma_j^2}$ for any h_i, h_j and σ_ε^2.

[17] One can also consider the degrees of teamwork interactions to be decision variables; i.e., an agent decides how much of a teammate's support she will appropriate. Agent i's "appropriation" effort, (say) b_t^i, and θ^j are multiplicative, $z_t^i = \theta^i + e_t^i + b_t^i\left(\theta^j + a_t^j\right) + \varepsilon_t^i$. There are now multiple equilibria. The optimal efforts satisfy $\psi'\left(e_1^{i*}\right) = \left(1 + b_1^{j*}\right)\rho_1^{ii}$, $\psi'\left(a_t^{i*}\right) = \left(1 + b_1^{j*}\right) b_1^{j*}\rho_1^{ij}$ and $\psi'\left(b_1^{i*}\right) = \left(1 + b_1^{j*}\right)\rho_1^{ii}\left(m_j + e_1^{j*}\right)$ for any i and j. Dewatripont et al. (1999) assume that agent i's (work) effort is multiplied with her 'own' ability, and thus career concerns depend only on the mean of θ^i, and not on θ^j as in our setting.

[18] For instance, one can consider a team that produces hard disks, but the team members use different technologies; i.e., magnetic and holographic. A market shock may hit the project outputs based on these two technologies differently.

Given the realized performances z_1^i and z_1^j, the correlation coefficients of the (conditional) distribution of θ^i are

$$\widetilde{\rho}_1^{ii} = \frac{\sigma_i^2}{\widetilde{\lambda}_1}\left[\left(1-h_i\phi\right)\sigma_\varepsilon^2 + \left(1-h_ih_j\right)\sigma_j^2\right] \text{ and } \widetilde{\rho}_1^{ij} = \frac{\sigma_i^2}{\widetilde{\lambda}_1}\left[\left(h_i-\phi\right)\sigma_\varepsilon^2 - \left(1-h_ih_j\right)h_j\sigma_j^2\right],$$

where $\widetilde{\lambda}_1 = \sigma_\varepsilon^4\left(1-\phi^2\right) + \left(1-h_ih_j\right)^2\sigma_i^2\sigma_j^2 + \sigma_\varepsilon^2\left[\left(1+h_i^2\right)\sigma_i^2 + \left(1+h_j^2\right)\sigma_j^2 - 2\phi\left(h_i\sigma_i^2 + h_j\sigma_j^2\right)\right]$.

The correlation coefficient $\widetilde{\rho}_1^{ii}$ is always positive, $\widetilde{\rho}_1^{ii} > 0$, but the effect of an increase in ϕ on $\widetilde{\rho}_1^{ii}$ and thus on the intensity of an agent's (utility-maximizing) work effort, e_1^{i*}, is not straightforward. For example, let $\sigma_\varepsilon^2 = \sigma_i^2 = \sigma_j^2 = 1$ and $h_j = 0$ in order to isolate the effects of h_i and ϕ on agent i' reputation incentives. If $\phi = 0.9$ while $h_i = 0.1$, we have $\frac{\partial\widetilde{\rho}_1^{ii}}{\partial\phi} > 0$: agent i's contribution in z_1^j is negligible, but the observation of this additional signal effectively reduces the variance of the "noise" of her own project, ε_1^i, allowing the market to put a higher weight on z_1^i in estimating θ^i. Thus, an increase in the correlation between the output shocks leads an agent to exert higher work effort in order to build up her reputation.[19]

The relationship between e_1^{i*} and ϕ becomes negative, $\frac{\partial\widetilde{\rho}_1^{ii}}{\partial\phi} < 0$, when $\phi = 0.1$ while $h_i = 0.9$. Assuming that agent i does not receive any help while her support is critical to her teammate's performance, high project outputs are mainly attributed to her own ability. The market perceives both signals indicating the level of θ^i and thus, given z_1^j, z_1^i is a good estimate of its level. However, as ϕ increases, and the market accumulates more information about the market conditions, a lower weight is put on z_1^i in estimating θ^i. In particular, as the 'prior' variance of the noise terms decreases, and market factors affect teammates' project outputs the same way (recall $h_j = 0$), the market anticipates that both teammates' good performance is influenced by market factors, revising the estimate of θ^i downwards. Thus, higher correlation between the shocks will decrease agent i's optimal work effort. However, if a teammate's support in agent i's project output is significant (say $h_j = 1$), the derivative $\frac{\partial\widetilde{\rho}_1^{ii}}{\partial\phi}$ becomes positive, because now additional information about the market environment will be nothing else but useful. If cross-agent teamwork interactions are intensive, the market finds it harder to perceive the levels of the θs. Thus, as ϕ increases, the market can better identify whether outputs signal the level of teammates' abilities or are influenced by marketwide factors.

This analysis highlights that, given the available information, a larger ϕ will discourage agent i to exert work effort if this increase leads to a worse market estimate of θ^i. More precisely, an increase in a small $\phi > 0$ will decrease agent i's optimal work effort when initiated interactions, h_i, are strong enough while received interactions, h_j, are weak: $\frac{\partial\psi'\left(e_1^{i*}\right)}{\partial\phi} < 0$ if and only if

$$\phi < \frac{\sigma_\varepsilon^2 + \left(1-h_ih_j\right)\sigma_j^2 - \left(\sigma_\varepsilon^2 + h_i^2\sigma_i^2 + \sigma_j^2\right)^{\frac{1}{2}}\left[\left(1-h_i^2\right)\sigma_\varepsilon^2 + \left(1-h_i^2h_j^2\right)\sigma_j^2\right]^{\frac{1}{2}}}{h_i\sigma_\varepsilon^2}.$$

[19]Under the assumption that $\sigma_\varepsilon^2 = \sigma_i^2 = \sigma_j^2 = 1$ and $h_j = 0$, for $\phi < 0$, we have $\frac{\partial\widetilde{\rho}_1^{ii}}{\partial\phi} < 0$ for any h_i.

Meyer & Vickers (1997) also examine the relationship between reputation incentives and the correlation of the output shocks. They consider a two-agent setting in which each agent's output depends only on her own effort and ability, as in Holmström (1999). They find that when agents' output shocks are correlated (while their abilities are independent), a larger correlation ϕ, when $\phi > 0$, leads an agent to exert higher effort, $\widetilde{e}_1^{i*}$, in order to increase her reputation. There is a negative externality and some rivalry between agents. The observation of another agent's outcome exactly reduces the variance of the "noise" and allows the market to rely more on an agent's performance to infer the level of her ability.[20] This effect is also present in our setting where teamwork interactions occur. However, we argue that this relationship can turn out to be negative when h_i is large enough and h_j small. In this case, an increase in ϕ induces the market to decrease the weight it puts on z_1^i to perceive the level of agent i's ability.

The sign of $\widetilde{\rho}_1^{ij}$ is also not clear cut. It depends on the *relative* intensity of the two forms of correlation between the project outputs. For $\widetilde{\rho}_1^{ij}$ to be positive, initiated interactions, h_i, must be sufficiently large in order for θ^i to be a key determinant of z_1^j. For instance, if the market shocks vary substantially (high σ_ε^2) and are negatively correlated, $\phi < 0$, $\widetilde{\rho}_1^{ij}$ is more "likely" to be positive. A high realization of z_1^i should be associated with a low z_1^j. However, if agent j's project output is also high, this is attributed to high θ^i, especially for relatively intensive initiated interactions h_i. In turn, agent i cashes in an increase in her reputational bonus due to a higher z_1^j and thus, she has incentives to help her fellow member.

Remark 3 (Different forms of correlation of performance measures & help effort) *An agent will have reputation incentives to help a teammate when initiated interactions are substantially larger than the correlation of the output shocks:* $\psi'\left(a_1^{i*}\right) > 0$ *if and only if* $h_i - \frac{\sigma_j^2}{\sigma_\varepsilon^2} h_j \left(1 - h_i h_j\right) > \phi$.

[Figures 3 are about here.]

In a setting where the random shocks are positively correlated, $\phi > 0$, but ϕ exceeds h_i, $\phi > h_i$, then $\widetilde{\rho}_1^{ij}$ is negative. This happens because the contribution of θ^i in z_1^j is relatively small and teammates' high project outputs are mainly attributed to market factors. The market believes that the teammates act in a favorable environment and updates its assessment of θ^i downwards. Therefore, there is some rivalry between the agents and, as a result, incentives to sabotage arise.

4 Multiperiod models

We now focus on career concerns when employment extends to many periods, and the output shocks are uncorrelated. We also use a stationary model, as in Holmström (1999), to examine whether

[20] Meyer & Vickers (1997) argue that this relationship between reputation incentives and the correlation of the output shocks is the counterpart of the insurance effect in a static principal-agent model where "comparative performance information" compensation schemes are provided. The observation of another agent's output increases the precision with which an agent's effort is estimated, leading the principal to provide additional motivation.

the equilibrium efforts are efficient under the assumption that the quality of a fellow team member matters for an agent's reputation.

4.1 The T-period case

At each period t, the market's assessment of abilities now depends on the history of agent i's and j's project outputs $z_1^i, z_1^j, ..., z_{t-1}^i, z_{t-1}^j$. The optimal efforts satisfy the equations $\psi'(e_t^{i*}) = (1+h_i)\sum_{\tau=t}^{\infty}\rho_\tau^{ii}$ and $\psi'(a_t^{i*}) = (1+h_i)h_i\sum_{\tau=t}^{\infty}\rho_\tau^{ij}$. The signals are

$$\rho_\tau^{ii} = \frac{\sigma_i^2}{\lambda_\tau}\left[\sigma_\varepsilon^2 + (\tau-1)(1-h_ih_j)\sigma_j^2\right] \text{ and } \rho_\tau^{ij} = \frac{\sigma_i^2}{\lambda_\tau}\left[h_i\sigma_\varepsilon^2 - (\tau-1)(1-h_ih_j)h_j\sigma_j^2\right],$$

where $\lambda_\tau \equiv \sigma_\varepsilon^4 + (\tau-1)^2(1-h_ih_j)^2\sigma_i^2\sigma_j^2 + (\tau-1)\sigma_\varepsilon^2\left[(1+h_i^2)\sigma_i^2 + \left(1+h_j^2\right)\sigma_j^2\right]$.

In line with the literature, the signal ρ_τ^{ii} is always positive but decreasing in τ. The more uncertainty there is about an agent's ability, the bigger the returns to her work effort. Thus, early in the process when there is less available information, the market puts more weight on the most recent output observation when updating its assessment of θ^i. Eventually, θ^i will be revealed almost completely, and new output observations will have little impact on market perceptions. For small h_i, the presence of teamwork interactions impedes the learning process about θ^i. However, agent i's attempts to influence output are only temporarily effective (only early in career). In this multi-period setting, the signal ρ_τ^{ij} deserves special attention.

Remark 4 (Signals over the periods) *For strong initiated interactions (large h_i), ρ_τ^{ij} is positive only in the early stages of agent i's career: $\rho_\tau^{ij} > 0$ if and only if $1 + \frac{h_i\sigma_\varepsilon^2}{(1-h_ih_j)h_j\sigma_j^2} > \tau$.*

The signal ρ_τ^{ij} may even switch signs, from positive to negative, as τ increases. Note that although the variance of a performance measure depends on the variance of its transitory shock - i.e., $var(z_t^i) = \sigma_i^2 + h_j^2\sigma_j^2 + \sigma_\varepsilon^2$ - the covariance of project outputs realized in the same or different periods depends only on the variance of teammates' abilities: $cov\left(z_t^i, z_{t+1}^i\right) = \sigma_i^2 + h_j^2\sigma_j^2$ and $cov\left(z_t^i, z_{t+1}^j\right) = h_i\sigma_i^2 + h_j\sigma_j^2$. Thus, over the periods, the noise in the performance measures driven by the output shocks becomes (relatively) less significant in the process of estimating abilities. To put it differently, the signals incorporate information about the covariances of all project outputs that have been realized in the past. Under the assumption of independently distributed random terms, such covariances depend solely on σ_i^2 and σ_j^2 (not σ_ε^2), implying that over the periods, the noise introduced by teammates' abilities matters more in the learning process and for reputation incentives. Thus, even when teamwork interactions are such that the market puts a positive weight on agent j's project outputs to infer the level of θ^i early in the process, as performance observations accumulate, ρ_τ^{ij} diminishes. At later stages of an agent's career, as θ^j becomes key in predicting θ^i, this signal can turn out to be negative. As the market learns more about θ^j by observing z_t^j, agent i's reputation incentives reverse and, in fact, she has incentives to sabotage. Even if early in the process an agent has incentives to help her colleague, sabotage incentives can arise for those agents who are about to retire.

4.2 The stationary case

We now investigate the relationship between the intensity of reputation incentives over time and the efficient level of efforts in a stationary setting where teammates' abilities remain unknown to the parties. In this setting, we can examine whether agents' desires to shape market perceptions in order to increase future remuneration can induce them to exert the "right" level of efforts. We also need to assume that the agents discount the future by some factor ζ. For smaller ζ, agents put a lower weight on the future and thus value the "delayed" payments less. Provided that career concerns arise exactly because of agents' attempts to increase their reputation and to seek higher future monetary payments, such incentives will be stronger in a setting without discounting. However, the presence of discounting in this analysis will allow for additional insights on whether the market forces alone can remove the moral hazard problems and provide adequate incentives for workers to perform.

In line with the literature based on Holmström (1999), we assume that ability fluctuates over an agents' working life, according to the process

$$\theta^i_{t+1} = \theta^i_t + \eta^i_t,$$

where η^i_t is independently and normally distributed with zero mean and variance σ^2_η. Thus, at period $t+1$, agent i's project output is $z^i_{t+1} = \theta^i_t + \eta^i_t + e^i_{t+1} + h_j\left(\theta^j_t + \eta^j_t + a^j_{t+1}\right) + \varepsilon^i_{t+1}$. The shocks η^i_t and η^j_t add uncertainty that prevents agents' abilities from becoming fully known. Lemma 2 derives the variance of θ^i_{t+1} in this stationary setting, where $\widehat{\sigma}^2_{i,t} \equiv \sigma^2_{i,t}\left(1-\rho^{ii}_t - h_i\rho^{ij}_t\right)$ is the variance of θ^i_{t+1} before observing the realizations of z^i_{t+1} and z^j_{t+1}.

Lemma 2 (Variance along the path to the stationary state) *After observing z^i_{t+1} and z^j_{t+1}, the variance along the path to the stationary state of agent i's ability at $t+1$ is*

$$\overline{\sigma}^2_{i,t+1} = \left(\widehat{\sigma}^2_{i,t} + \sigma^2_\eta\right)\left(1-\widehat{\rho}^{ii}_t - h_i\widehat{\rho}^{ij}_t\right),$$

where $\widehat{\rho}^{ii}_t \equiv \dfrac{\widehat{\sigma}^2_{i,t}+\sigma^2_\eta}{\widehat{\lambda}^i_t}\left[\sigma^2_\varepsilon + (1-h_ih_j)\left(\widehat{\sigma}^2_{j,t}+\sigma^2_\eta\right)\right]$, $\widehat{\rho}^{ij}_t \equiv \dfrac{\widehat{\sigma}^2_{i,t}+\sigma^2_\eta}{\widehat{\lambda}^i_t}\left[h_i\sigma^2_\varepsilon - (1-h_ih_j)\,h_j\left(\widehat{\sigma}^2_{j,t}+\sigma^2_\eta\right)\right]$,

and $\widehat{\lambda}^i_t \equiv \sigma^4_\varepsilon + (1-h_ih_j)^2\left(\widehat{\sigma}^2_{i,t}+\sigma^2_\eta\right)\left(\widehat{\sigma}^2_{j,t}+\sigma^2_\eta\right) + \sigma^2_\varepsilon\left[\left(1+h_i^2\right)\left(\widehat{\sigma}^2_{i,t}+\sigma^2_\eta\right) + \left(1+h_j^2\right)\left(\widehat{\sigma}^2_{j,t}+\sigma^2_\eta\right)\right]$.

The learning process becomes

$$m^i_{t+1} = \widehat{\mu}^i_t m^i_t + \widehat{\rho}^{ii}_t\left(z^i_t - \widehat{e}^i_t - h_j\widehat{a}^j_t - h_j m^j_t\right) + \widehat{\rho}^{ij}_t\left(z^j_t - \widehat{e}^j_t - m^j_t - h_i\widehat{a}^i_t\right),$$

where $\widehat{\mu}_t^i = 1 - \widehat{\rho}_t^{ii} - h_i \widehat{\rho}_t^{ij}$. The shocks η_t^i and η_t^j prevent the market from learning, and thus the variance of abilities declines deterministically with t but does not go to zero. The optimal efforts satisfy

$$\psi'\left(e_t^{i*}\right) = (1+h_i)\,\widehat{\rho}_t^{ii} \sum_{s=t+1}^{\infty} \zeta^{s-t} \Pi_{\kappa=t+1}^{s} \mu_\kappa^i \equiv T_{e_t^i}, \tag{4}$$

$$\psi'\left(a_t^{i*}\right) = (1+h_i)\,h_i \widehat{\rho}_t^{ij} \sum_{s=t+1}^{\infty} \zeta^{s-t} \Pi_{\kappa=t+1}^{s} \mu_\kappa^i \equiv T_{a_t^i}. \tag{5}$$

In period 1, we have $\psi'\left(e_1^{i*}\right)$ to be given by the sum of the terms $\zeta\,(1+h_i)\,\widehat{\rho}_1^{ii}$, $\zeta^2\,(1+h_i)\,\widehat{\rho}_1^{ii}\widehat{\mu}_2^i$, $\zeta^3\,(1+h_i)\,\widehat{\rho}_1^{ii}\widehat{\mu}_2^i\widehat{\mu}_3^i$, etc. In the stationary case where $\widehat{\rho}_{t+1}^{ii} = \widehat{\rho}_t^{ii} = \widehat{\rho}^{ii*}$, $\widehat{\rho}_{t+1}^{ij} = \widehat{\rho}_t^{ij} = \widehat{\rho}^{ij*}$ and $\widehat{\mu}^{i*} = 1 - \widehat{\rho}^{ii*} - h_i \widehat{\rho}^{ij*}$, equation (4) becomes

$$\psi'\left(e_1^{i*}\right) = \zeta\,(1+h_i)\,\widehat{\rho}^{ii*}\left[1 + \zeta\widehat{\mu}^{i*} + \zeta^2\left(\widehat{\mu}^{i*}\right)^2 + \zeta^3\left(\widehat{\mu}^{i*}\right)^3 + \ldots\right],$$

where the sum in the brackets is equal to $\frac{1}{1-\zeta\mu^{i*}}$. Thus, we obtain the stationary work and help effort levels:

$$\psi'\left(e^{i*}\right) = \frac{\zeta\,(1+h_i)\,\widehat{\rho}^{ii*}}{1-\zeta\left(1-\widehat{\rho}^{ii*} - h_i\widehat{\rho}^{ij*}\right)} \text{ and } \psi'\left(a^{i*}\right) = \frac{\zeta\,(1+h_i)\,h_i\widehat{\rho}^{ij*}}{1-\zeta\left(1-\widehat{\rho}^{ii*} - h_i\widehat{\rho}^{ij*}\right)}. \tag{6}$$

Holmström (1999) formalizes Fama's (1980) major conclusion that the market induces agents to exert the efficient effort levels. In a single-agent model, he shows that this happens if there is no discounting. In our model where the team members interact, and an agent's individual performance depends on the quality of her team, even if there is no discounting, for $h_i > 0$, we argue that Fama's conclusion generically fails. The stationary levels of efforts are above or below their efficient levels. Let us first discuss two cases where teammates' stationary effort levels are also efficient in our model.

Under full information (first-best), agent i's remuneration is a fixed payment equal to the sum of the disutilities of work and help efforts, $\psi\left(e_t^i\right) + \psi\left(a_t^i\right)$, and the efficient effort levels $e_t^{i,fb}$ and $a_t^{i,fb}$ satisfy the conditions $\psi'\left(e_t^{i,fb}\right) = 1$ and $\psi'\left(a_t^{i,fb}\right) = h_i$, respectively. The first-best reward at each period t is the reward that is optimal in a one-shot game.

Agent i's stationary work and help efforts are efficient as long as there is no discounting, $\zeta = 1$, and an agent's ability *does not* affect her teammate's project output, $h_i = 0$, as in Holmström (1999) and Auriol et al. (2002), although received interactions may occur, $h_j > 0$ (agent j's stationary efforts will be inefficient). Therefore, $\widehat{\rho}^{ij*}$ as a signal should play no role in agent i's reputation decisions. In turn, the stationary work effort is efficient and equal to one: $\psi'\left(e^{i*}\right) = \psi'\left(e_t^{i,fb}\right) = 1$. Since an agent's help effort does not affect the process of inference about her own ability, any incentive to influence a teammate's performance disappears. Exerting zero help effort in this stationary model is also efficient: $\psi'\left(a^{i*}\right) = \psi'\left(a_t^{i,fb}\right) = 0$.

It is rather striking that in our model where teammates' abilities matter for reputation concerns, the stationary effort levels can also be efficient as long as the initiated *and* received interactions are *perfect*, $h_i = h_j = 1$. Now, providing help to a colleague is as productive as putting effort into one's own task. An agent's help and work efforts weight equally to both performance measures. Under full information, the efficient effort levels are given by $\psi'\left(e_t^{i,fb}\right) = \psi'\left(a_t^{j,fb}\right) = 1$. The stationary efforts also satisfy $\psi'(e^{i*}) = \psi'(a^{i*}) = \psi'(e^{j*}) = \psi'(a^{j*}) = 1$. They become efficient as soon as we add any amount of noise into the learning process about agents' ability levels. There is a balance between the incentives to work and help in order to build up reputation. Proposition 2 establishes that in our model, under any other condition, Fama's conclusion does not hold.

Proposition 2 (Stationary labor supply) *In the stationary model where $\sigma_\eta^2 > 0$ and $\sigma_\varepsilon^2 > 0$, for all $h_j \in (0,1)$, if initiated teamwork interactions occur, $h_i > 0$, agent i exerts*

(a) higher work effort than its efficient level: $\psi'(e^{i*}) > \psi'\left(e_t^{i,fb}\right)$;

(b) lower help effort than its efficient level: $\psi'(a^{i*}) < \psi'\left(a_t^{i,fb}\right)$.

In the stationary case where $h_j < 1$, for any h_i, an agent's work effort is higher, and help effort is lower than its efficient level. Agent i has incentives to distort her work effort upwards in order to signal that she is of higher ability and induce the market to revise its beliefs in her favor. The stationary reputation incentives indicate that an agent is oriented to exert more work effort in order to improve her own performance, rather than to focus on helping or sabotaging her colleague. The optimal a^{i*} is distorted downwards. Thus, agents will overprovide work effort and underprovide help effort. For small initiated interactions such that $\widehat{\rho}^{ij*} < 0$, the stationary level of help effort can even be negative, while its efficient level is always positive.

To complete this analysis, we also need to examine the convergence to the stationary state. We need to explore reputation incentives before a stationary state is reached. In Holmström (1999), learning follows the process $m_{t+1}^i = v_t^i m_t^i + (1 - v_t^i)\, z_t^i$, where $z_t^i = \theta^i + e_t^i + \varepsilon_t^i$ and $v_t^i > 0$, implying that the convergence of an agent's effort to the stationary state is directly related to the dynamics of v_t^i. In our model, $\widehat{\mu}_t^i$ incorporates both signals $\widehat{\rho}_t^{ii}$ and $\widehat{\rho}_t^{ii}$. Thus, the sequence of $\widehat{\mu}_t^i$ will depend on the amount of information extracted by both signals, $\widehat{\rho}_t^{ii} + h_i \widehat{\rho}_t^{ii}$, in each period. However, the convergence of an agent's work effort will depend on the dynamics of $\widehat{\rho}_t^{ii}$, while the convergence of help effort will depend on the dynamics of $\widehat{\rho}_t^{ij}$.

We show in Appendix $(A.2)$ that the sequence of agent i's optimal work effort $\{e_t^{i*}\}$ converges to the stationary state level e^{i*}. In equation (4), when $t = 1$, each term is increasing in $\widehat{\rho}_1^{ii}$, hence so is $T_{e_1^i}$. If $\left\{\widehat{\rho}_t^{ii}\right\}$ is an increasing sequence, then so is $\left\{T_{e_t^i}\right\}$, and the convergence of $\{e_t^{i*}\}$ is from below. Similarly, if $\widehat{\rho}_t^{ii}$ is a decreasing sequence, the convergence of $\{e_t^{i*}\}$ is from above. The same dynamics govern the convergence of $\{a_t^{i*}\}$ to its stationary level.

5 Conclusion

We examine career concerns in teams in a setting with interactions between fellow team members where the help an agent receives depends on both her colleague's effort and innate ability. Teamwork interactions affect the learning process and are at the heart of this analysis. By exerting effort and providing support, an agent can affect both her own and her teammate's project outputs. Thus, she can use both performance measures to induce the market to revise its assessment of her own ability upwards.

We show that career concerns depend on how many signals the agent can influence in order to manipulate the market's inference. In particular, we argue that if initiated interactions are substantial enough to make an agent's support a key determinant of her teammate's production, the agent has incentives to work and help her colleague. By providing support, an agent can signal that she is a high-productivity agent. In contrast, if initiated interactions are weak and received interactions are intensive enough for the market to update its beliefs about an agent's ability downwards upon the colleague's strong performance, sabotage incentives arise. This happens because an agent's higher help effort increases a teammate's performance, which biases the process of inference against her. Thus, the agent has incentives to sabotage her teammate in order to signal that she is of higher ability and increase her reputation. In a stationary model where we add uncertainty into the performance measures in order for abilities to remain unknown, initiated interactions induce the agents to supply work effort above its efficient level and help effort below its efficient level. The optimal implicit incentives are distorted as long as teamwork interactions are imperfect, and there is no discounting.

This model can be used to analyze reputation incentives of team workers when their individual performance is observable and depends on the quality of fellow members. This is likely to happen in research collaborations, sports, or finance reams. There are also extensions and directions for future work that are of special interest, using the present model as a reference point. For instance, one can consider differences in the mean of the distribution of teammates' abilities and address the question of whether a researcher has incentives to be teamed with senior or junior colleagues. We can also assume that a worker contributes to multiple projects and is teamed with different workers in each of them. We can then examine if she has incentives to work in projects where teammates' abilities are more visible or in projects where teammates are of lower productivity. The size of the team with heterogeneous workers is another key determinant of career concerns. For instance, biotechnology requires large teams and may lack the ability to break up large projects into small independent modules, as is possible in the software industry.

Market conditions may also alter team workers' incentives to signal their abilities. For instance, the existence of a dominant competitor tends to align the goals of the team members, and thus sabotage incentives may be weakened or even reversed. Competition may necessitate cooperation within a heterogeneous group and reputation incentives may encourage support provision in order to ensure the success of the projects. If explicit contracts are provided, by allowing for side payments between the agents as well as for different allocations of the bargaining power, one may boost this analysis further.

References

Arya, A. & Mittendorf, B. (2011), 'The benefits of aggregate performance metrics in the presence of career concerns', *Management Science* **57**(8), 1424–1437.

Auriol, E., Friebel, G. & Pechlivanos, L. (2002), 'Career concerns in teams', *Journal of Labor Economics* **20**(2), 289–307.

Bar-Isaac, H. (2007), 'Something to prove: reputation in teams', *RAND Journal of Economics* **38**(2), 495–511.

Bar-Isaac, H. & Hörner, J. (2014), 'Specialized careers', *Journal of Economics & Management Strategy* **23**(3), 601–627.

Bonatti, A. & Hörner, J. (2014), 'Career concerns with exponential learning', *working paper, MIT and Yale University* .

Borland, J. (1992), 'Career concerns: Incentives and endogenous learning in labour markets', *Journal of Economic Surveys* **6**(3), 251–270.

Coupe, T., Smeets, V. & Warzynski, F. (2006), 'Incentives, sorting and productivity along the career: Evidence from a sample of top economists', *Journal of Law, Economics and Organization* **22**(1), 137–167.

DeGroot, M. H. (1970), *Optimal Statistical Decisions*, McGraw-Hill, New York.

Dewatripont, M., Jewitt, I. & Tirole, J. (1999), 'The economics of career concerns, Part II: Application to missions and accountability of government agencies', *Review of Economic Studies* **66**(1), 199–217.

Dewatripont, M., Jewitt, I. & Tirole, J. (2000), 'Multitask agency problems: Focus and task clustering', *European Economic Review* **44**(4), 869–877.

Effinger, M. R. & Polborn, M. K. (2001), 'Herding and anti-herding: A model of reputational differentiation', *European Economic Review* **45**(3), 385–403.

Fama, E. F. (1980), 'Agency problems and the theory of the firm', *Journal of Political Economy* **88**(2), 288–307.

Gibbons, R. & Murphy, K. J. (1992), 'Optimal incentive contracts in the presence of career concerns: Theory and evidence', *Journal of Political Economy* **100**(3), 468–505.

Hann, I.-H., Roberts, J., Slaughter, S. & Fielding, R. (2004), 'An empirical analysis of economic returns to open source participation', *Faculty Development Grant und Carnegie Bosch Institute, Carnegie Mellon University* .

Holmström, B. (1982), *Managerial incentives - A dynamic perspective*, Essays in Economics and Management in Honor of Lars Wahlbeck, Helsinki, Finland: Swedish School of Economics.

Holmström, B. (1999), 'Managerial incentive problems: A dynamic perspective', *Review of Economic Studies* **66**, 169–182.

Holmström, B. & Milgrom, P. (1991), 'Multitask principal agent analyses: Incentive contracts, asset ownership and job design', *Journal of Law, Economics, and Organization* **7**, 24–52.

Itoh, H. (1992), 'Cooperation in hierarchical organizations: An incentive perspective', *Journal of Law, Economics, and Organization* **8**, 321–345.

Jeon, S. (1996), 'Moral hazard and reputational concerns in teams: Implications for organizational choice', *International Journal of Industrial Organization* **14**, 297–315.

Kogut, B. & Metiu, A. (2000), 'The emergence of e-innovation: Insights from open source software development', *Philadelphia: Wharton School, University of Pennsylvania* .

Laffont, J. J. & Tirole, J. (1988), 'The dynamics of incentive contracts', *Econometrica* **56**(5), 1153–1175.

Lazear, E. (1989), 'Pay equality and industrial politics', *Journal of Political Economy* **97**, 561–580.

Meyer, M. A. & Vickers, J. (1997), 'Performance comparisons and dynamic incentives', *Journal of Political Economy* **105**(3), 547–581.

Milgrom, P. & Oster, S. (1987), 'Job discrimination, market forces, and the invisibility hypothesis', *Quarterly Journal of Economics* pp. 453–476.

Mukherjee, A. (2008), 'Sustaining implicit contracts when agents have career concerns: the role of information disclosure', *RAND Journal of Economics* **39**(2), 469–490.

Ortega, J. (2003), 'Power in the firm and managerial career concerns', *Journal of Economics and Management Strategy* **12**(1), 1–29.

Prendergast, C. & Topel, R. H. (1996), 'Favoritism in organizations', *Journal of Political Economy* **104**(5), 958–978.

A APPENDIX

A.1 Proof of Lemma 1: conditional distribution of abilities

The variance-covariance matrix of the multivariate normal distribution of θ^i, z_t^i and z_t^j is

$$\begin{pmatrix} \sigma_i^2 & \sigma_i^2 & h_i\sigma_i^2 \\ \sigma_i^2 & \sigma_i^2 + h_j^2\sigma_j^2 + \sigma_\varepsilon^2 & h_i\sigma_i^2 + h_j\sigma_j^2 \\ h_i\sigma_i^2 & h_i\sigma_i^2 + h_j\sigma_j^2 & \sigma_j^2 + h_i^2\sigma_i^2 + \sigma_\varepsilon^2 \end{pmatrix}.$$

Following DeGroot (1970), after the observation of z_1^i and z_1^j, the conditional mean of θ^i has the form

$$\begin{aligned} E\left\{\theta^i \mid z_1^i, z_1^j\right\} &= m_1^i + \sigma_{i1}'\Sigma_{i1}^{-1}\begin{pmatrix} z_1^i - \widehat{e}_1^i - m_1^i - h_j\left(\widehat{a}_1^j + m_1^j\right) \\ z_1^j - \widehat{e}_1^j - m_1^j - h_i\left(\widehat{a}_1^i + m_1^i\right) \end{pmatrix}, \\ \text{where } \sigma_{i1}' &= \begin{pmatrix} cov\left(\theta^i, z_1^i\right) & cov\left(\theta^i, z_1^j\right) \end{pmatrix} \text{ and } \Sigma_{i1} = \begin{pmatrix} var\left(z_1^i\right) & cov\left(z_1^i, z_1^j\right) \\ cov\left(z_1^i, z_1^j\right) & var\left(z_1^j\right) \end{pmatrix}. \end{aligned}$$

Thus, $\Sigma_{i1}^{-1} = \frac{1}{\lambda_1}\begin{pmatrix} var\left(z_1^j\right) & -cov\left(z_1^i, z_1^j\right) \\ -cov\left(z_1^i, z_1^j\right) & var\left(z_1^i\right) \end{pmatrix}$, where $\lambda_1 = var\left(z_1^i\right)var\left(z_1^j\right) - \left[cov\left(z_1^i, z_1^j\right)\right]^2$. The correlation coefficient matrix is

$$\sigma_{i1}'\Sigma_{i1}^{-1} = \begin{pmatrix} \rho_1^{ii} & \rho_1^{ij} \end{pmatrix} = \frac{1}{\lambda_1}\begin{pmatrix} \sigma_i^2 & h_i\sigma_i^2 \end{pmatrix}\begin{pmatrix} \sigma_j^2 + h_i^2\sigma_i^2 + \sigma_\varepsilon^2 & -\left(h_i\sigma_i^2 + h_j\sigma_j^2\right) \\ -\left(h_i\sigma_i^2 + h_j\sigma_j^2\right) & \sigma_i^2 + h_j^2\sigma_j^2 + \sigma_\varepsilon^2 \end{pmatrix},$$

where $\lambda_1 = \sigma_\varepsilon^4 + (1 - h_ih_j)^2\sigma_i^2\sigma_j^2 + \sigma_\varepsilon^2\left[\left(1 + h_i^2\right)\sigma_i^2 + \left(1 + h_j^2\right)\sigma_j^2\right]$. The conditional variance of θ^i has as

$$var\left\{\theta^i \mid z_1^i, z_1^j\right\} = var\left(\theta^i\right) - \sigma_{i1}'\Sigma_{i1}^{-1}\sigma_{i1} \text{ where } \sigma_{i1}'\Sigma_{i1}^{-1}\sigma_{i1} = \sigma_i^2\left(\rho_1^{ii} + h_i\rho_1^{ij}\right).$$

A.2 Convergence to the stationary effort levels

We first elaborate the dynamics of the learning process. The sequence of $\widehat{\mu}$s reveals how fast updating about agent i's ability occurs. In particular, the recursive relationship between the $\widehat{\mu}$s is $\widehat{\mu}_{t+1}^i = \frac{1}{k_i - \widehat{\mu}_t^i}$, where $k_i \geq 2$.[21] We solve this equation and get

$$\widehat{\mu}_t^i = \frac{\sqrt{4 + k_i^2}\widehat{\mu}_0^i\beta + \left(2 - k_i\mu_0^i\right)\gamma}{\sqrt{4 + k_i^2}\beta + \left(2\mu_0^i + k_i\right)\gamma},$$

[21] We have $k_i = 2 + \delta_i$, where $\delta_i \equiv \frac{(h_i+h_j)^2\sigma_\varepsilon^4\sigma_j^4 v_i + \lambda_1^i(1-h_ih_j)^2\sigma_\eta^4 + \sigma_\eta^2\sigma_\varepsilon^2\left[\left(1+h_i^2\right)\left[\sigma_\varepsilon^4 + 2(1-h_ih_j)^2\sigma_i^2\sigma_j^2\right] + \left(1-h_j^2-2h_ih_j+h_i^2+h_i^4\right)\sigma_i^2\sigma_\varepsilon^2 + \sigma_j^2\sigma_\varepsilon^2 l\right]}{\sigma_\varepsilon^4\left[\left(1+h_j^2\right)\sigma_\eta^2+\sigma_\varepsilon^2\right]\left[\left(1+h_i^2\right)\sigma_i^2+\sigma_\varepsilon^2\right] + \sigma_j^2\sigma_\varepsilon^2\left[\sigma_\varepsilon^2\left(1+h_j^2\right)\left[\left(1+h_j^2\right)\sigma_\eta^2+2\sigma_\varepsilon^2\right] + \sigma_i^2\left[\left(1+h_j^2\right)(1-h_ih_j)^2\sigma_\eta^2 + l\sigma_\varepsilon^2\right]\right]}$, $v_i \equiv \frac{\sigma_i^2}{\lambda_1^i}\left[\left(1 + h_j^2\right)\sigma_\varepsilon^2 + (1 - h_ih_j)^2\sigma_i^2\right]$ and $l \equiv 2(1 - h_ih_j) + \left(1 + 2h_i^2\right)h_j^2 + \left(1 + 2h_j^2\right)h_i^2$.

where $\beta \equiv \left(k_i + \sqrt{4+k_i^2}\right)^t + \left(k_i - \sqrt{4+k_i^2}\right)^t$, and $\gamma \equiv \left(k_i + \sqrt{4+k_i^2}\right)^t - \left(k_i - \sqrt{4+k_i^2}\right)^t$.[22] The difference between two subsequent $\widehat{\mu}$s gives

$$\widehat{\mu}_{t+1}^i - \widehat{\mu}_t^i = \frac{4\left(4+k_i^2\right)\left(k_i + \sqrt{4+k_i^2}\right)^t \left(k_i - \sqrt{4+k_i^2}\right)^t \left[1 - \widehat{\mu}_0^i\left(k_i + \widehat{\mu}_0^i\right)\right]}{\left[\beta\sqrt{4+k_i^2} + \gamma\left(k_i + 2\widehat{\mu}_0^i\right)\right]\left[\left(k_i\gamma + \beta\sqrt{4+k_i^2}\right)\left(k_i + \widehat{\mu}_0^i\right) + 2\gamma\right]}.$$

We can consider a numerical example when $t = 2$. If the above difference is positive, we get $\widehat{\mu}_0^i = -\frac{53}{16}$ and $k_i = 3$, while when it is negative, $\widehat{\mu}_0^i = -\frac{27}{8}$ and $k_i = 3$. Thus, for $k_i = 3$, as $t \to \infty$, this difference is equal to $\frac{1}{2}\left(-3+\sqrt{13}\right)$. Stationarity requires $\widehat{\mu}_{t+1}^i = \widehat{\mu}_t^i = \widehat{\mu}^{i*}$, implying that $\widehat{\mu}^{i*} = \frac{1}{2}\left(-k_i + \sqrt{4+k_i^2}\right)$, which is also equal to $\frac{1}{2}\left(-3+\sqrt{13}\right)$ when $k_i = 3$.[23] We now get

$$\begin{aligned}
\widehat{\mu}_t^i - \widehat{\mu}^{i*} &= \frac{\left(k_i - \sqrt{4+k_i^2}\right)^t \left[\left(k_i + 2\widehat{\mu}_0^i\right)\sqrt{4+k_i^2} - 4 - k_i^2\right]}{\beta\sqrt{4+k_i^2} + \gamma\left(k_i + 2\widehat{\mu}_0^i\right)}, \\
\frac{\widehat{\mu}_t^i - \widehat{\mu}^{i*}}{\left(k_i - \sqrt{4+k_i^2}\right)^t} &= \frac{\left(k_i + 2\widehat{\mu}_0^i\right)\sqrt{4+k_i^2} - 4 - k_i^2}{\sqrt{4+k_i^2} - \left(k_i + 2\widehat{\mu}_0^i\right) + \left[\sqrt{4+k_i^2} + \left(k_i + 2\widehat{\mu}_0^i\right)\right]\left(\frac{k_i+\sqrt{4+k_i^2}}{k_i-\sqrt{4+k_i^2}}\right)^t}.
\end{aligned}$$

The numerator is a positive constant. The denominator depends on the fraction $\frac{k_i+\sqrt{4+k_i^2}}{k_i-\sqrt{4+k_i^2}}$, which is greater than one, implying that $\left(\frac{k_i+\sqrt{4+k_i^2}}{k_i-\sqrt{4+k_i^2}}\right)^t$ diverges. Thus, the sequence will converge to the stationary state.

The dynamics of agent i's work effort are revealed in equation (4), implying that

$$T_{e_1^i} = \zeta\left(1+h_i\right)\widehat{\rho}_1^{ii} + \zeta^2\left(1+h_i\right)\widehat{\rho}_1^{ii}\widehat{\mu}_2^i + \zeta^3\left(1+h_i\right)\widehat{\rho}_1^{ii}\widehat{\mu}_2^i\widehat{\mu}_3^i + \ldots \tag{7}$$

We can show that each term in equation (7) is increasing in $\widehat{\rho}_1^{ii}$, hence so is $T_{e_1^i}$. This positive relationship holds by induction. Let $\xi_s\left(\widehat{\rho}_1^{ii}\right) \equiv \widehat{\rho}_1^{ii}\widehat{\mu}_2^i\widehat{\mu}_3^i \ldots \widehat{\mu}_s^i$, which is increasing in $\widehat{\rho}_1^{ii}$, and thus, $\xi_s\left(\widehat{\rho}_2^{ii}\right) = \widehat{\rho}_2^{ii}\widehat{\mu}_3^i\widehat{\mu}_4^i \ldots \widehat{\mu}_{s+1}^i$. We get

$$\xi_{s+1}\left(\widehat{\rho}_1^{ii}\right) = \widehat{\rho}_1^{ii}\widehat{\mu}_2^i\widehat{\mu}_3^i \ldots \widehat{\mu}_{s+1}^i = \frac{\widehat{\rho}_1^{ii}}{\widehat{\rho}_2^{ii}}\widehat{\mu}_2^i \xi_s\left(\widehat{\rho}_2^{ii}\right).$$

The correlation coefficients $\widehat{\rho}_1^{ii}$ and $\widehat{\rho}_2^{ii}$ are both positive. By the inductive hypothesis, $\xi_s(.)$ is in-

[22] If $h_i = h_j = 0$, as in Holmström (1999), δ_i is equal to $\frac{\sigma_\eta^2}{\sigma_\varepsilon^2}$. If σ_ε^2 is sufficiently larger than σ_η^2 so that δ_i is close to zero, the stationary level μ^{i*} is close to 1, implying that learning occurs slowly. In contrast, if $\sigma_\varepsilon^2 = \sigma_\eta^2$, then $\delta_i = 1$ and updating about m_t^i occurs quickly. In our model, if $h_i = 0$ and $h_j = 1$, teamwork interactions slow down learning since $\delta_i < 1$, while they speed up learning if $h_i = 1$ and $h_j = 0$ since $\delta_i > 1$. δ_i exceeds $\frac{\sigma_\eta^2}{\sigma_\varepsilon^2}$ and thus $\widehat{\mu}^{i*}$ is lower than the stationary level of this measure in Holmström's model. The updating of m_t occurs faster, diminishing the negative effects of discounting.

[23] Note that the solution of the equation $\widehat{\mu}^{i*} = \frac{1}{k_i - \widehat{\mu}^{i*}}$ gives two fixed points, $\widehat{\mu}^{i*} = \frac{1}{2}\left(-k_i \pm \sqrt{4+k_i^2}\right)$. However, if the initial condition is other than $\frac{1}{2}\left(-k_i - \sqrt{4+k_i^2}\right)$, the sequence converges to $\frac{1}{2}\left(-k_i + \sqrt{4+k_i^2}\right)$.

creasing. Note that $\widehat{\mu}_2^i = 1 - \widehat{\rho}_2^{ii} - h_i^{ii}\widehat{\rho}_2^{ij} = \frac{\sigma_\varepsilon^2}{\widehat{\lambda}_2^i}\left[\sigma_\varepsilon^2 + \left(1+h_j^2\right)\left(\widehat{\sigma}_{j,2}^2 + \sigma_\eta^2\right)\right]$. Thus, we have $\frac{1}{\widehat{\rho}_2^{ii}}\widehat{\mu}_2^i = \frac{\sigma_\varepsilon^2\left[\sigma_\varepsilon^2 + \left(1+h_j^2\right)\left(\widehat{\sigma}_{j,2}^2+\sigma_\eta^2\right)\right]}{\left(\widehat{\sigma}_{i,2}^2+\sigma_\eta^2\right)\left[\sigma_\varepsilon^2+(1-h_ih_j)\left(\widehat{\sigma}_{j,2}^2+\sigma_\eta^2\right)\right]}$, which is decreasing in μ_1^{ii} and thus increasing in $\widehat{\rho}_1^{ii}$. In turn, the positive coefficient $\frac{\widehat{\rho}_1^{ii}}{\widehat{\rho}_2^{ii}}\widehat{\mu}_2^i$ is also increasing in $\widehat{\rho}_1^{ii}$. It follows that $\xi_{s+1}\left(\widehat{\rho}_1^{ii}\right)$ and thus $\psi'\left(e_1^{i*}\right)$ are also increasing as functions of $\widehat{\rho}_1^{ii}$. It boils down to the following: if $\left\{\widehat{\rho}_t^{ii}\right\}$ is an increasing (decreasing) sequence, $\left\{T_{e_t^i}\right\}$ is also an increasing (decreasing) sequence. Thus, the sequence of agent i's optimal work effort $\{e_t^{i*}\}$ converges to the stationary state level e^{i*}. If $\left\{\widehat{\rho}_t^{ii}\right\}$ is an increasing sequence, the convergence of $\{e_t^{i*}\}$ is from below. Similarly, if $\left\{\widehat{\rho}_t^{ii}\right\}$ is a decreasing sequence, the convergence of $\{e_t^{i*}\}$ is from above.

The dynamics of agent i's help effort follow by studying equation (5). We have

$$T_{a_1^i} = \zeta\left(1+h_i\right)\widehat{\rho}_1^{ij} + \zeta^2\left(1+h_i\right)\widehat{\rho}_1^{ij}\widehat{\mu}_2^i + \zeta^3\left(1+h_i\right)\widehat{\rho}_1^{ij}\widehat{\mu}_2^i\widehat{\mu}_3^i + ... \tag{8}$$

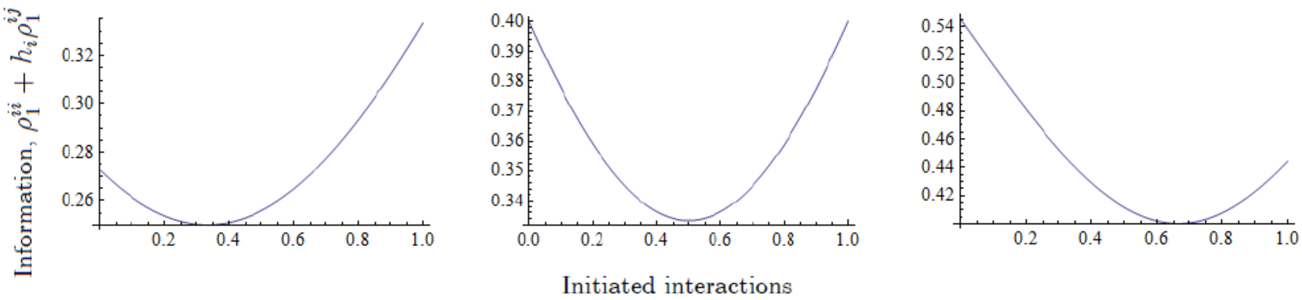
The sign of each term in equation (8) is the same as the sign of $\widehat{\rho}_1^{ij}$. First, we will examine the convergence to the stationary level of help effort when this is positive. If each term in (8) is increasing in (the positive) $\widehat{\rho}_1^{ij}$, the same is true for $T_{a_1^i}$. As above, we prove this statement by induction. Suppose $\xi_s\left(\widehat{\rho}_1^{ij}\right) = \widehat{\rho}_1^{ij}\widehat{\mu}_2^i\widehat{\mu}_3^i...\widehat{\mu}_s^i$, which implies $\xi_s\left(\widehat{\rho}_2^{ij}\right) = \widehat{\rho}_2^{ij}\widehat{\mu}_3^i\widehat{\mu}_4^i...\widehat{\mu}_{s+1}^i$. These equations give

$$\xi_{s+1}\left(\widehat{\rho}_1^{ij}\right) = \widehat{\rho}_1^{ij}\widehat{\mu}_2^i\widehat{\mu}_3^i...\widehat{\mu}_{s+1}^i = \frac{\widehat{\rho}_1^{ij}}{\widehat{\rho}_2^{ij}}\widehat{\mu}_2^i\xi_s\left(\widehat{\rho}_2^{ij}\right).$$

We have $\frac{1}{\widehat{\rho}_2^{ij}}\widehat{\mu}_2^i = \frac{\sigma_\varepsilon^2\left[\sigma_\varepsilon^2+\left(1+h_j^2\right)\left(\widehat{\sigma}_{j,2}^2+\sigma_\eta^2\right)\right]}{\left(\widehat{\sigma}_{i,2}^2+\sigma_\eta^2\right)\left[h_i\sigma_\varepsilon^2-(1-h_ih_j)h_j\left(\widehat{\sigma}_{j,2}^2+\sigma_\eta^2\right)\right]}$, which is increasing in $\widehat{\rho}_1^{ij}$. It follows that the coefficient $\frac{\widehat{\rho}_1^{ij}}{\widehat{\rho}_2^{ii}}\widehat{\mu}_2^i$, the product $\xi_{s+1}\left(\widehat{\rho}_1^{ij}\right)$, and thus the optimal help effort a_1^{i*} are also increasing in $\widehat{\rho}_1^{ij}$. Therefore, if $\left\{\widehat{\rho}_t^{ij}\right\}$ is a positive and increasing sequence, $\left\{T_{a_t^i}\right\}$ is also an increasing sequence. The convergence of agent i's optimal help effort $\{a_t^{i*}\}$ to the stationary state level a^{i*} will be from below. If $\left\{\widehat{\rho}_t^{ij}\right\}$ is a decreasing sequence, the convergence of $\{a_t^{i*}\}$ is from above.

We can perform the same analysis to examine the convergence to the stationary level of help effort when this is negative. Now, $\xi_{s+1}\left(\widehat{\rho}_1^{ij}\right)$ and $\xi_s\left(\widehat{\rho}_2^{ij}\right)$ are also negative. However, the coefficient $\frac{\widehat{\rho}_1^{ij}}{\widehat{\rho}_2^{ij}}\mu_2^i$ is positive and increasing in $\widehat{\rho}_1^{ij}$. This coefficient reinforces the dynamics of $\xi_s(.)$. Thus, the sequence of the help effort also converges to the stationary state.

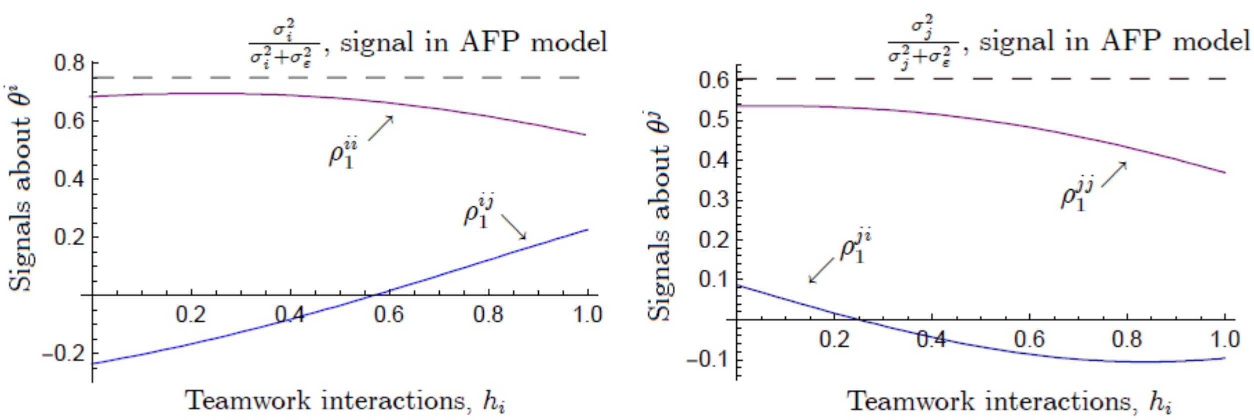
Figures 1. Effect of h_i on learning, captured by $\rho_1^{ii} + h_i\rho_1^{ij}$.



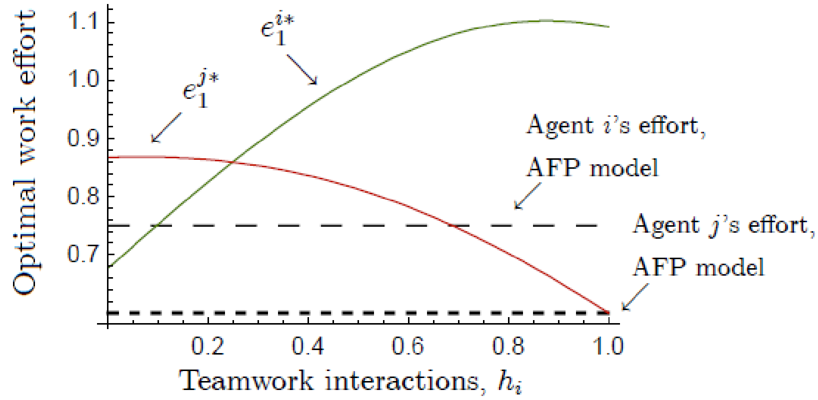
$\sigma_i^2 = \sigma_j^2 = 2$, $\sigma_\varepsilon^2 = 4$ $\sigma_i^2 = \sigma_j^2 = 4$, $\sigma_\varepsilon^2 = 4$ $\sigma_i^2 = \sigma_j^2 = 4$, $\sigma_\varepsilon^2 = 2$

Figures 1 show the changes in $\rho_1^{ii} + h_i\rho_1^{ij}$ as h_i increases under different assumptions about the level of σ_i^2, σ_j^2 and σ_ε^2. In all three figures, it is also assumed that $h_j = 1$.

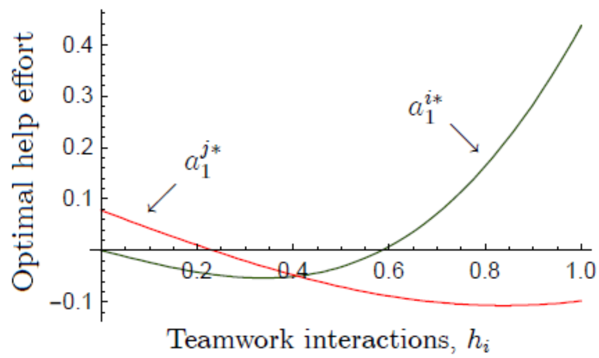
Figures 2. Effect of h_i on the signals about abilities and teammates' optimal efforts.



2.1 2.2



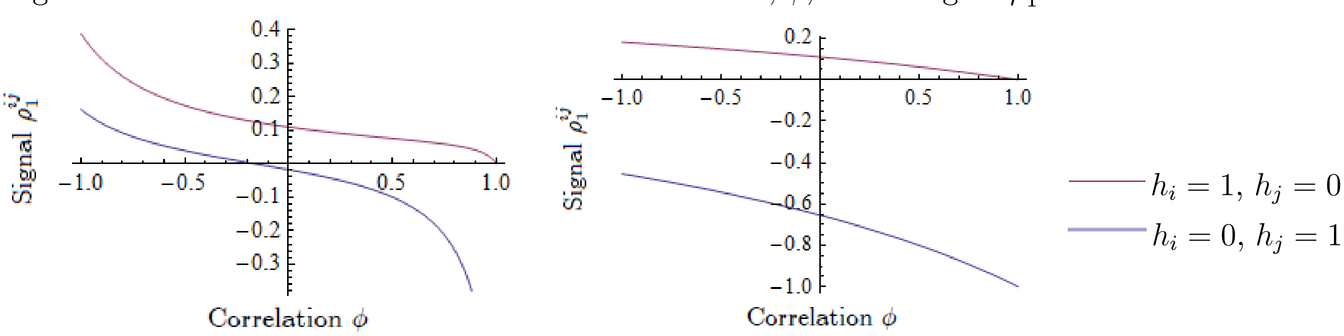
2.3



2.4

Figures 2.1 and 2.2 show the effect of teamwork interactions h_i on the signals about θ^i and θ^j, respectively, under the assumptions that $\sigma_i^2 = 6$, $\sigma_j^2 = 3$, $\sigma_\varepsilon^2 = 2$ and $h_j = 0.6$. In Auriol et al. (2002) and Holmström (1999), $corr\left(\theta^i \mid z_1^i\right) = \frac{\sigma_i^2}{\sigma_\varepsilon^2+\sigma_i^2}$ and $corr\left(\theta^j \mid z_1^j\right) = \frac{\sigma_j^2}{\sigma_\varepsilon^2+\sigma_j^2}$, while $corr\left(\theta^i \mid z_1^j\right) = corr\left(\theta^j \mid z_1^i\right) = 0$. Figures 2.3 and 2.4 show how the optimal work efforts and help efforts change with h_i. We assume that $\psi\left(e_t^i\right) = \frac{1}{2}\left(e_t^i\right)^2$and $\psi\left(a_t^i\right) = \frac{1}{2}\left(a_t^i\right)^2$.

Figures 3. Effect of the correlation of the random terms, ϕ, on the signal ρ_1^{ij}.



$\sigma_i^2 = \sigma_j^2 = 1$, $\sigma_\varepsilon^2 = 6$ $\qquad$ $\sigma_i^2 = \sigma_j^2 = 6$, $\sigma_\varepsilon^2 = 1$

Figures 3 show the changes in the sign of ρ_1^{ij} as ϕ increases, under certain assumptions about σ_i^2, σ_j^2, σ_ε^2, h_i and h_j.