Team Incentives with Imperfect Mutual Inference

Alice Peng-Ju Su*

Department of Economics, National Taipei University
August, 2018

Abstract

I study the optimal team incentive when the agents can coordinate private actions through repeated interaction with imperfect public monitoring. The agents are able to imperfectly infer each other's private actions via the stochastically correlated measurements. Correlation of measurement noise, besides its risk sharing role in the conventional multiple-agent moral hazard problem, is crucial to the accuracy of each agent's inference. The principal's choice of performance pay to provide incentive via inducing competition or coordination among the agents thus exhibits the trade-off between risk sharing and mutual inference between the agents. I characterize the optimal form of performance pay with respect to the correlation of measurement noise. Whether the conventional theoretical prediction holds depends on how the agents form the mutual inference, based on an exogenous standard or formed endogenously.

Keywords: Team incentive, Relative performance evaluation, Joint performance evaluation, Imperfect monitoring, Collusion

JEL Classification: D86, D82, J33

*E-mail: apjs@gm.ntpu.edu.tw. Address: 151 University Rd., Sanxia Dist., Department of Economics, National Taipei University, New Taipei City, 23741 Taiwan. Suggestions by the co-editor, the associate editor and two anonymous referees have greatly improved this paper. I am grateful to Fahad Khalil, Jacques Lawarrée, Charles Hill, Wei Li, Dennis O'Dea, Luis Rayo, Lan Shi, and Quan Wen, as well as participants at the 44th EARIE, for their helpful comments and suggestions. Financial support from the Ministry of Science and Technology (grant number MOST 105-2410-H-305-012-) is gratefully acknowledged.

1 Introduction

In the studies of team incentives with agents who may coordinate private actions through side contracting or repeated interaction[1], each agent observes an exogenous signal of each other's action. Such observation is independent of the measurement noise of their private actions. I study the optimal form of contract given an alternative on how the agents observe each other's private action. They infer it through the publicly observable and contractible stochastic measurement of each agent's private action. Each agent, knowing his own private action, knows to what extent the measurement is attributed to noise. If the measurement noise of each agent's action is correlated to that of another's, by observing another agent's measurement of action he is able to update his belief on that agent's action. The principal does not have this information as she cannot distinguish the private action from the measurement noise of any agent. Her decision on the optimal contract incorporates whether to induce the agents to make this mutual inference as an implicit incentive, or to deter them from making the mutual inference to game the system.

Consider for example, two salespersons of the same real estate company. If the wage scheme is based on relative performance of the two in additional to personal performance, the salespersons may reach an informal agreement between each other to shirk. When the monthly performance is announced, if the salesperson who followed the agreement to shirk observes an exceptionally good performance of the other, he updates his belief that it is very likely that his colleague has worked hard instead. He would thus have a second thought on whether to reach any informal agreement with this colleague in the rest of his career. Alternatively, if the real estate company pay the salespersons based on joint performance in addition to personal performance, the salespersons may reach an informal agreement between each other to work hard. When the monthly performance is announced, if the salesperson who followed the agreement to work hard observes an exceptionally poor performance of the other, he updates his belief that it is very likely that his colleague has shirked instead. He would thus have a second thought on whether to reach any informal agreement with this colleague in the rest of his career.

In this example, the coalition between the salespersons are enforced with a mutual

[1]Please refer to Holmström (1982), Mookherjee (1984), and Ramakrishnan and Thakor (1991) for pioneer research on team incentive with non-cooperative agents, and Holmström and Milgrom (1990), Itoh (1992, 1993), Che and Yoo (2001), Kim and Vikander (2015), and Rayo (2007) for that with agents coordinating actions with either a side contract or repeated interaction with perfect monitoring. Fleckinger and Roux (2012) surveyed this literature.

inference that is only available among the salespersons themselves, because each of them knows whether he had really shirked or worked, which is a hidden action for the real estate company. Accuracy of such mutual inference is related to whether the salespersons face correlated market shocks. Suppose that the salespersons are paid contingent on joint performance and reach an agreement to work hard. If the salespersons work in the same housing market so that their performances are highly correlated, when one of them performs exceptionally poorly, they are more convinced that this poorly-performed salesperson has deviated from their informal agreement. If on the other hand the salespersons are allocated to different housing markets with low stochastic correlation, each of them would also attribute the poor performance of the colleague to the cold market where he works.

The optimal form of team incentive to either induce the agents to coordinate actions or to deter the agents from collusion of actions relates to how the agents form the mutual inference to enforce their coalition,[2] as well as to factors that determines the accuracy of mutual inference. Specifically, I borrow the model from Holmström and Milgrom (1990) and incorporate infinitely repeated interaction among the agents when the principal's equilibrium response is to offer a stationary contract in every period as assumed in Che and Yoo (2001). The key feature of this paper that distinguishes itself from the related literature is that the repeated interaction is enforced by the imperfect inference the agents make on each other's private action. The agents' hidden actions are imperfectly measured by correlated individual performance measurements.

Higher correlation of measurement noise in this context potentially has the following effects. First, it allows for more efficient risk sharing under a contract with relative performance evaluation (RPE). More correlated are the agents, the principal is able to design the contract with RPE to better filter out the common measurement noise so that incentive is provided with a lower risk premium. This is the risk sharing effect highlighted by Holmström and Milgrom (1990). Second, more correlated the agents are in the measurement noise, each agent better distinguishes the private action from the measurement noise of another agent. Higher correlation of measurement noise improves the accuracy of the agents' mutual inference and better enforces the agents' coordination or collusion. Improved mutual inference lowers the cost to induce coordination under a contract with joint performance evaluation (JPE), while it increases the cost to deter collusion and provide incentive under a contract with RPE. This is

[2] For clarity, throughout this paper I use the term coordination when the principal induces the agent to coordinate actions, while the term collusion is used when the principal deters such activity. The term coalition is referred to the team of agents who coordinate or collude.

the monitoring effect emphasized by Rayo (2007).

The principal's choice of contract and her decision on whether to induce the agents to coordinate or to deter them from collusion exhibit a trade-off between these two effects. Risk sharing under a contract with RPE is distorted to deter collusion among the agents. Inducing the agents to coordinate actions via mutual inference relies on a contract with JPE, which is at the expense of higher risk premium. The optimal contract as a result of this trade-off is related to how the agents form the mutual inference, which shapes the enforcement of the agents' coalition through the accuracy of inference without directly affecting the risk sharing role of the contract.

If the agents form their mutual inference based on an exogenous standard of statistical significance, correlation of measurement noise plays an essential role in the update of mutual inference. The optimal contract may not be monotonic in the correlation parameter. For relatively small correlations, a higher correlation has a sharper effect on the improvement of mutual inference than that on the improvement of risk sharing. The optimal contract has collusion-proof RPE for lower correlations and coordination-inducing JPE for higher correlations. This contradicts to Holmström and Milgrom (1990)'s claim that RPE is optimal for more correlated measurement noises, yet it is consistent to Rayo (2007) who suggested that a more accurate signal is in favor of JPE that induces implicit incentive. For relatively large correlations of measurement noise, a higher correlation has a sharper effect on the improvement of risk sharing than that on the improvement of mutual inference. The optimal contract has coordination-inducing JPE for lower correlations and collusion-proof RPE for higher correlations, even though a higher correlation implies better accuracy of inference. This is consistent to Holmström and Milgrom (1990), yet it contradicts to Rayo (2007).

If the agents form their mutual inference based on an endogenously chosen standard of significance, the optimal choice is where the accuracy of mutual inference is maximized. The endogenous inference amplifies the enforcement of coalition. A coordination-inducing JPE is less costly for the principal while the collusion-proof RPE is more costly. In addition, the optimal standard of significance replaces the correlation of measurement noise as the essential role to update inference. At any level of correlation, the standard of significance is chosen such that the mutual inference is most accurate. The improvement of mutual inference is less prominent than that with exogenous inference whereas the risk sharing effect is not directly affected by how inference is formed. Even if a higher correlation improves the accuracy of inference, such improvement is outweighed by the improvement of risk sharing. The optimal contract is less likely to exhibit non-monotonicity in the correlation of measurement noise. It has coordination-inducing JPE for lower correlations to utilize the agents' endogenous inference and collusion-proof RPE for higher correlations to benefit from more efficient risk sharing.

When it comes to incentive provision for teams, monitoring and risk sharing among the agents are often treated in the literature as two independent aspects that affect the optimal contract. In this paper, they are both related to the correlation of measurement noise. The optimal contract in theory depends on the relative magnitude of the two effects instead of one of the effects with the other fixed. The relative magnitude of the two effects also relate to whether the inference is formed endogenously or exogenously. Identification of how the inference is formed and how essential stochastic correlation is in improving the inference is thus important when designing the optimal contract.

1.1 Related Literature

The paper builds on a vast literature devoted to the study of optimal team incentive, with the option that agents are able to coordinate private actions through an explicit side contract (Holmström and Milgrom (1990), Itoh (1992, 1993)), relational side payment (Rayo (2007)), or enforced by repeated interaction (Che and Yoo (2001) and Kim and Vikander (2015)). It also relates to Deb, Li, and Mukherjee (2016), who considered collusion-free team incentives when each agent receives and reports private inference (peer evaluation) to the principal. They, however, assume that private actions are either contractible among the agents or observable through an exogenous signal. Contractibility or observation among the agents in these papers are independent of how the measurement of each agent's private action correlates to that of another. I take one step further to argue that each agent's observation on the others' private actions is an inference drawn from stochastically correlated measurements of actions. This attributes the accuracy of observation to the correlation of measurement noise, and yields new conclusions to the literature.

Repeated game with imperfect public monitoring has been developed in the industrial organizational literature (Green and Porter (1984) and Abreu, Pearce, and Stacchetti (1986)) and made general by Abreu, Pearce, and Stacchetti (1990) and Fudenberg, Levine, and Maskin (1994). I apply the concept to contract theory as an enforcement of the agents' coordination, with imperfect monitoring inferred from correlated measurements of private actions, to study the optimal provision of team incentives. Cai and Obara (2009) also apply it to study horizontal integration of the firms, in which the imperfect monitoring of the customers relate to the size of integration. The fact that the optimal organizational form relates to the accuracy of imperfect monitoring is similar in spirit to this paper, yet with a different formation of information structure on imperfect monitoring.

On optimal form of contracts, Fleckinger (2012) has similar findings that RPE is favored with lower (equilibrium) correlation of noise, a contradiction to Holmström and Milgrom (1990). However, he assumes that correlation of measurement noise depends on the level of effort, so that equilibrium correlation not only provides the principal with information on stochastic environment (informative correlation) but also signals all agents' private actions (technological correlation). I show in this paper that even if correlation of measurement noise is productive independent of the agents' actions (purely informative) as in Holmström and Milgrom's model, the principal may still benefit from inducing the agents to coordinate with a higher correlation. The essential condition is that the agents's exogenous mutual inference on each other's private action is prominently updated from the stochastically correlated measurement of actions. Meyer and Vickers (1997) discuss incentive gain or loss of RPE with respect to different levels of cross-sectional and inter-temporal correlations with the presence of career concern. Their focus is on the interaction between ratchet effect and insurance, whether they reinforce or oppose each other. Collusion and coordination between the agents are absent in both papers.

Ishiguro (2004) and Kim (2011) characterize the optimal collusion-proof contract in a tournament, assuming that the risk neutral agents with limited liability can write side contract contingent on verifiable outcome before exerting private actions. There is no loss of generality to consider only collusion-proof contracts in their models as the agents do not possess superior information to that of the principal. They conclude that shutting down one agent non-anonymously is essential to deter side contracting in a tournament with limited liability, where performance pay contingent on actual realization of outcome is impossible.

2 Model

A principal contracts with two agents $i = 1, 2$ to execute a project at each period $t = 1, 2, \ldots$. Each agent exerts private actions $a_{it} \in \{a_L, a_H\}$, $a_L < a_H$, at a cost $c(a_k) \equiv c_k$, $k = L, H$ and $0 \leq c_L < c_H$. For future reference, denote the additional cost due to a higher level of private action as $\eta \equiv \frac{c_H - c_L}{a_H - a_L}$. The actions are productive independent in the sense that agent i's private action generates a publicly observable and contractible output x_{it} which is an imperfect measurement of his private action. Let $x_{it} = a_{it} + \varepsilon_{it}$, where ε_{it} is the stochastic nature agent i faced at period t that is time-independently distributed.[3] Output (x_{1t}, x_{2t}) jointly follows normal distribution with unit variance and covariance $0 \leq \sigma \leq 1$ across agents,

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}, \begin{pmatrix} 1 & \sigma \\ \sigma & 1 \end{pmatrix} \right).$$

The principal is risk neutral with time separable utility $u_{pt} = \sum_i (vx_{it} - w_{it}(x_{1t}, x_{2t}))$, where $v > 0$ measures the marginal revenue of output and $w_{it}(x_{1t}, x_{2t})$ is the transfer paid to agent i contingent on realization of outputs at period t. My goal is to study the optimal form of contract to induce the agents to or to deter them from coordinating private actions, and how such form of contract depends on the agents' mutual inference. I thus assume throughout that implementing $a_i = a_H$ for $i = 1, 2$ is optimal, i.e. the marginal revenue of output is sufficiently large relative to the cost of actions. In addition, the agents' coordination is enforced in a repeated relationship with their ability to make inference on each other's action through correlated imperfect measurement, as will be specified in the next subsection. To emphasize on such enforcement of coordination, in the main sections I restrict attention to the situation where the principal offers the stationary contract $w_{it}(x_{1t}, x_{2t}) = w_i(x_{1t}, x_{2t})$.[4]

Each agent is risk averse and has a time-separable exponential utility $u_{it} = 1 - \exp(-r(w_i(x_{1t}, x_{2t}) - c(a_{it})))$, where $r > 0$ denotes the agent's constant absolute risk aversion. Assumption of exponential utility allows us to take advantage of its convenience without losing too much insight. Specifically, in each agent's optimization problem given an accepted contract, maximization of the expected utility with an exponential form is equivalent to maximization of its certainty equivalence, denoted as

[3] An alternative way to model is to assume that the agents exert private actions to produce a non-contractible noisy output, and that each agent's private action is measured separately and imperfectly by a contractible indicator, which is stochastically correlated among the agents. Modeling in this way, conclusions would be qualitatively the same even with synergy in production.

[4] A stationary contract can be justified as the existence of a sequentially optimal short-term contract. Fudenberg, Holmström, and Milgrom (1990) constructed sufficient conditions for such existence. In brief, the principal and the agents are symmetrically informed of future history-independent contingent outcomes and technologies, and that preferences are time separable, in which case a long-term contract has no extra value beyond a sequence of its short-term counterpart. Given the nature of the agents' repeated game in this paper as will be introduced shortly, the future outcomes are not history-independent and the team of agents is better informed of these future outcomes than is the principal. I thus briefly discuss qualitative robustness when relaxing the restriction to allowing a menu of stationary contracts later as well.

$\bar{u}_{it} = E(w_i(x_{1t}, x_{2t})) - c(a_{it}) - \frac{r}{2}Var(w_i(x_{1t}, x_{2t}))$, which allows us to directly model the risk sharing concern, captured by $R(w_i(x_{1t}, x_{2t})) \equiv \frac{r}{2}Var(w_i(x_{1t}, x_{2t}))$. In addition, with exponential utilities and time-separable production, Holmström and Milgrom (1987) have shown the optimality of linear contract, $w_i(x_{1t}, x_{2t}) = \alpha_i x_{it} + \beta_i x_{jt} + \gamma_i$, in which β_i as the transfer rate contingent on the other agent's realization of output determines the contractual form I intend to study in this paper. In addition, the agents are different only in the realization of measurement noise, I thus focus on the non-discriminative contract where $(\alpha_i, \beta_i, \gamma_i) = (\alpha, \beta, \gamma)$ for $i = 1, 2$ to implement (a_H, a_H).

The optimal contract may take the form of relative performance evaluation, joint performance evaluation, or independent piece rate, defined respectively as the following. The contract has a relative performance evaluation (RPE) if a better realization of one agent's output lowers the other's payoff. In the current context, this corresponds to $\beta < 0$. It has a joint performance evaluation (JPE) if a better realization of one agent's output increases the other's payoff. In the current context, this corresponds to $\beta > 0$. It consists of an independent piece rate (IPR) if each agent's payoff is independent of the other's realization of output, i.e. if $\beta = 0$.

The principal-agents relationship is summarized as the following repeated contracting relationship. The principal proposes a stationary contract $w_i(x_{1t}, x_{2t})$ to agent $i = 1, 2$, each of whom either accepts or rejects the contract. If it is rejected, each player earns reservation payoff zero. If it is accepted, the agents decide in each period t whether to coordinate their private actions, followed by exerting action a_{it} simultaneously and independently. Output (x_{1t}, x_{2t}) is realized and publicly observable at the end of each period, and transfer payment $w_i(x_{1t}, x_{2t})$ is made accordingly at the end of each period. All players have perfect recall.

2.1 Imperfect Inference via Correlated Noise

After signing the contract and before production, the agents can coordinate with each other and agree on a non-verifiable action profile (a_1^0, a_2^0), when the contract implements static Nash equilibrium $(a_1^{'}, a_2^{'})$. The actions are not directly observable to each other, but given a production function with separable action and measurement noise, each agent is aware of to what extent his own realization of output is attributed to measurement noise, which is correlated to that of the other agent. Hence, each agent has partial information on the other's output realization that is attributed to the measurement noise instead of to the productive action. This information allows them to deduce imperfectly the private action exerted by each other. Such deduction is not available to the principal, who does not observe any agent's action, and thus cannot distinguish realization of output due to productive action or measurement noise of any agent.

Specifically, I model this ability to make inference through correlated measurement noise as what follows. Agent $j \neq i$, knowing his own private action, forms a belief on agent i's production, as a conditional density of x_{it} on x_{jt},

$$x_{it}(a_{it})|x_{jt}(a_j^0) \sim \mathcal{N}\left(a_{it} + \sigma(x_{jt} - a_j^0),\, 1 - \sigma^2\right).$$

When the realization of agent i's output falls far from the mean, at an extreme tail of the above conditional density, agent j believes that agent i has deviated from their agreement (a_1^0, a_2^0). This is precisely defined in the following assumption.

Assumption 1. *Agent j believes that agent i has unilaterally deviated downward from their non-verifiable agreement (a_{1t}^0, a_{2t}^0) if the realization of x_{it} conditional on x_{jt} is $s > 0$ standard deviations lower than the conditional mean, i.e. if $x_{it}(a_{it})|x_{jt}(a_j^0) < a_i^0 + \sigma(x_{jt} - a_j^0) - s\sqrt{1-\sigma^2}$, and has unilaterally deviated upward from their agreement if the realization of x_{it} conditional on x_{jt} is $s > 0$ standard deviations higher than the conditional mean, i.e. if $x_{it}(a_{it})|x_{jt}(a_j^0) > a_i^0 + \sigma(x_{jt} - a_j^0) + s\sqrt{1-\sigma^2}$.*

Parameter s can be interpreted as a commonly adopted standard for significance in statistical inferences with the null hypothesis being non-deviation. In Economics, for example, a 1% level of significance is commonly adopted in research, which approximately corresponds to $s = 2.33$. In manufacturing for another example, a well-adopted six-sigma rule in quality control and management translates to $s = 6$. With this interpretation, s is assumed to be an exogenous threshold of believed deviation. No individual can easily alters the commonly adopted standard. Parameter s can also be regarded as an intrinsic trust between the agents. If the agents place more trust on each other, each believes that the other has cheated when the realization of output is farther from the conditional mean, represented by a larger s. This latter interpretation invites the concern of an endogenously chosen threshold of believed deviation within the team of agents. We will first study the case with an exogenous threshold of believed deviation, as a comparison with the follow-up discussion on the optimal team incentive when this threshold is chosen endogenously by the team of agents.

For future reference, let q_i^d be the probability of agent j detecting agent i's deviation correctly, and q_i^n be the probability of false detection in the case of no deviation. That is,

$$q_i^d = \begin{cases} Pr\left(x_{it}(a_i^{'})|x_{jt}(a_j^0) < a_i^0 + \sigma(x_{jt} - a_j^0) - s\sqrt{1-\sigma^2}\right) & for\, a_i^{'} < a_i^0 \\ Pr\left(x_{it}(a_i^{'})|x_{jt}(a_j^0) > a_i^0 + \sigma(x_{jt} - a_j^0) + s\sqrt{1-\sigma^2}\right) & for\, a_i^{'} > a_i^0 \end{cases},$$

$$q_i^n = \begin{cases} Pr\left(x_{it}(a_i^0)|x_{jt}(a_j^0) < a_i^0 + \sigma(x_{jt} - a_j^0) - s\sqrt{1-\sigma^2}\right) & for\, a_i^{'} < a_i^0 \\ Pr\left(x_{it}(a_i^0)|x_{jt}(a_j^0) > a_i^0 + \sigma(x_{jt} - a_j^0) + s\sqrt{1-\sigma^2}\right) & for\, a_i^{'} > a_i^0 \end{cases},$$

where $a_i^{'}$ is the optimal action to deviate given the contract. The difference between the two, $\Delta \equiv q_i^d - q_i^n$, measures the accuracy of the agents' mutual inference. To see how this inference is incorporated into the agents' repeated interaction, suppose that the principal proposes a contract $\mathbb{C}$ that implements a static game between the agents in the form of a prisoner's dilemma: the non-cooperative Nash equilibrium is $(a_1^{'}, a_2^{'})$, yet the aggregate utility maximizing action profile is (a_1^0, a_2^0).

Assumption 2. *The agents play the grim trigger strategy with a bang-bang property. They continue to exert the coordinated actions (a_1^0, a_2^0) if a unilateral deviation is not detected; otherwise, the coalition breaks and the agents play the non-cooperative Nash equilibrium strategy $(a_1^{'}, a_2^{'})$ infinitely.*

In the repeated game between the agents described by Assumption 1 and 2, a contract $\mathbb{C}$ enforces (a_1^0, a_2^0) as a perfect public equilibrium (PPE)[5] if

$$\begin{aligned}(1-\delta)\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) + \delta\left((1-q_i^n)\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) + q_i^n\overline{u}_i(a_i^{'}, a_j^{'}|\mathbb{C})\right) \\ \geq (1-\delta)\overline{u}_i(a_i^{'}, a_j^0|\mathbb{C}) + \delta\left((1-q_i^d)\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) + q_i^d\overline{u}_i(a_i^{'}, a_j^{'}|\mathbb{C})\right),\end{aligned}$$

where $\delta \in (0,1)$ is the common discount factor. Note that with imperfect monitoring and a symmetric setting as in this model, punishment for both agents must occur on the equilibrium path to enforce (a_1^0, a_2^0), making the most punishing grim trigger strategy inefficient.[6] Nevertheless, it is the form of contract that admits (a_1^0, a_2^0) instead

[5]On general repeated games with imperfect public monitoring, please refer to Abreu, Pearce, and Stacchetti (1990) and Fudenberg, Levine, and Maskin (1994). Threshold grim trigger strategies were traditionally adopted in the study of repeated Cournot games to enforce collusion, e.g. Porter (1983) and Green and Porter (1984) as pioneer works and Aoyagi and Fréchette (2009) for a recent theoretical and experimental work.

[6]Fudenberg, Levine, and Maskin (1994) derived conditions under which a grim trigger strategy can yield efficient payoff. In brief, not only a unilateral deviation of a given player is detected stochastically as in this paper (individual full rank), but also the player who has deviated can be detected stochastically (pairwise full rank).

of the punishment strategy in the agents' coalition that is of interest. I thereby adopt the simplest.

Exogenous Inference. When the team of agents adopt an exogenous standard of significance in making inferences on each other's action, denote the accuracy of inference at an exogenous threshold s as Δ_s. A higher correlation between the measurement noise of actions allows the agents to better infer each other's private action. At the extreme, if the agents' stochastic outputs are perfectly correlated, they have perfect mutual inference among each other. At the other extreme, if the agents' stochastic outputs are independently distributed, they know as much about each other's action as the principal does.

Lemma 1. *The probability of correct detection of deviation q_i^d is increasing in σ and the probability of false detection q_i^n is independent of σ. Hence, the exogenous accuracy of mutual inference Δ_s is increasing in σ.*

Proof. Appendix A.1. □

More stochastic correlated are the agents, each of them correctly detects a unilateral deviation of action with a higher probability, while the probability of false detection is independent of the correlation of measurement noise. It is pure error and only depends on the standard of significance s. Hence, correlation of measurement noise improves the accuracy of the agents' mutual inference.

Endogenous Inference. When the agents' belief on each other's unilateral deviation is endogenously derived from an intrinsic trust between each other, knowing that their actions are measured with a highly correlated noise, the agents adjust their intrinsic trust accordingly. When the threshold of believed deviation s is an endogenous consensus within the coalition, the optimal s^* when the agents coordinate/collude will be such that the expected payoff under non-deviation is maximized subject to enforcement, i.e.

$$s^* \in \arg\max_{s} \sum_{i=1,2} \left((1-\delta)\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) + \delta \left((1-q_i^n)\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) + q_i^n \overline{u}_i(a_i^{'}, a_j^{'}|\mathbb{C}) \right) \right)$$

subject to

$$(1-\delta)\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) + \delta\left((1-q_i^n)\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) + q_i^n\overline{u}_i(a_i^{'}, a_j^{'}|\mathbb{C})\right)$$
$$\geq (1-\delta)\overline{u}_i(a_i^{'}, a_j^0|\mathbb{C}) + \delta\left((1-q_i^d)\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) + q_i^d\overline{u}_i(a_i^{'}, a_j^{'}|\mathbb{C})\right),$$

in which the agents agree to coordinate/collude to exert (a_i^0, a_j^0) while the non-cooperative Nash equilibrium is $(a_i^{'}, a_j^{'})$. Given enforcement, the contract must satisfies team incentive compatibility, $\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) > \overline{u}_i(a_i^{'}, a_j^{'}|\mathbb{C})$, so the problem is reduced to

$$s^* \in \arg\min_s q_i^n$$

subject to

$$\Delta \geq \frac{1-\delta}{\delta}\left(\frac{\overline{u}_i(a_i^{'}, a_j^0|\mathbb{C}) - \overline{u}_i(a_i^0, a_j^0|\mathbb{C})}{\overline{u}_i(a_i^0, a_j^0|\mathbb{C}) - \overline{u}_i(a_i^{'}, a_j^{'}|\mathbb{C})}\right) \qquad (EF).$$

The optimal threshold of believed unilateral deviation is such that the probability of false detection is minimized, subject to that the difference in correct and false detection is sufficiently large to ensure the enforcement of coalition, with the lower bound depending on the contract. This is as if the principal can not only induce coordination, but also implement the level of trust the agents place on each other, which results in binding enforcement constraint at the optimal s^* for the agents. Denote the endogenous accuracy of inference resulted from the optimal threshold of believed unilateral deviation as Δ^*.

2.2 The Contracting Problems

Collusion-proof Contract. A contract as a triplet $\mathbb{C} = (\alpha, \beta, \gamma)$ implementing (a_H, a_H) is incentive compatible and collusion-proof if (a_H, a_H) is the Nash equilibrium among the non-cooperative agents, and neither (a_L, a_L), (a_H, a_L), nor (a_L, a_H) can be enforced as a perfect public equilibrium of the agents' repeated game described in Assumptions 1 and 2. It is individually rational if each agent earns at least his reservation payoff in each period. A collusion-proof contract (α, β, γ) satisfies the following individual rationality, incentive compatibility, and collusion-proofness constraints,

$$\overline{u}_i(a_H, a_H|\mathbb{C}) \geq 0 \qquad (IR_n),$$

$$\overline{u}_i(a_H, a_j|\mathbb{C}) \geq \overline{u}_i(a_L, a_j|\mathbb{C}) \quad a_j \in \{a_H, a_L\} \qquad (IC_n),$$

$$(1-\delta)\overline{u}_i(a_L, a_L|\mathbb{C}) + \delta\left((1-q_i^n)\overline{u}_i(a_L, a_L|\mathbb{C}) + q_i^n\overline{u}_i(a_H, a_H|\mathbb{C})\right)$$
$$\leq (1-\delta)\overline{u}_i(a_i = a_H, a_j = a_L|\mathbb{C}) + \delta\left((1-q_i^d)\overline{u}_i(a_L, a_L|\mathbb{C}) + q_i^d\overline{u}_i(a_H, a_H|\mathbb{C})\right), \quad (1)$$

and

$$(1-\delta)\overline{u}_i(a_i = a_L, a_j = a_H|\mathbb{C}) + \delta\left((1-q_i^n)\overline{u}_i(a_i = a_L, a_j = a_H|\mathbb{C}) + q_i^n\overline{u}_i(a_H, a_H|\mathbb{C})\right)$$
$$\leq (1-\delta)\overline{u}_i(a_H, a_H|\mathbb{C}) + \delta\left((1-q_i^d)\overline{u}_i(a_i = a_L, a_j = a_H|\mathbb{C}) + q_i^d\overline{u}_i(a_H, a_H|\mathbb{C})\right). \quad (2)$$

Constraint (2) holds when (IC_n) is satisfied, so only (1) is the relevant collusion-proofness constraints. Constraint (1) given contract $\mathbb{C}$ can be written as

$$\alpha + h(\delta, \Delta)\beta \geq \eta \qquad (CP),$$

where $h(\delta, \Delta) \equiv \frac{\delta\Delta}{1-\delta+\delta\Delta}$ measures the quality of the coalition if it were ever formed, which is increasing in the discount factor and the accuracy of the agents' mutual inference. With exogenous inference, $\Delta = \Delta_s$. With endogenous inference $\Delta = \Delta^*$ given s^*, and the (EF) constraint is identical to the (CP) constraint.

The fixed payment in the contract, γ, is adjusted to satisfy constraint (IR_n) without affecting incentive compatibility nor collusion-proofness. The principal's problem to find the optimal collusion-proof contract is to minimize expected transfer to the agents, which is reduced to minimizing the risk premium given binding (IR_n), subject to incentive compatibility and collusion-proofness constraints, i.e.

$$\mathscr{P}_{CP}: \qquad \min_{\alpha,\beta} R(\alpha, \beta) \equiv \frac{r}{2}(\alpha^2 + \beta^2 + 2\sigma\alpha\beta)$$

subject to (IC_n) and (CP). Note that any $\beta > 0$ raises the risk premium without providing additional incentive. JPE ($\beta > 0$) is never optimal to deter collusion.

When $\delta \to 0$, collusion-proofness is redundant given incentive compatibility. We obtain the non-cooperative benchmark in which collusion is impossible. The non-cooperative benchmark contract is in the form of RPE due to its better risk sharing across the agents, unless the measurement noise is stochastically independent.[7] It is straightforward from (CP) that the benchmark RPE is not collusion-proof as long as the discount factor is positive.

[7] As raised by Ishiguro (2002), under the non-cooperative benchmark, the implemented action profile (a_H, a_H) is not the unique Nash equilibrium with binding (IC_n). Nevertheless, optimality of RPE would still hold, because the principal can increase α_n by an infinitesimal amount to implement (a_H, a_H) as the unique Nash equilibrium within the team, which does not change the fact that the contract exhibits RPE.

Coordination-inducing Contract. A contract as a triplet $\mathbb{C} = (\alpha, \beta, \gamma)$ implementing (a_H, a_H) induces coordination between the agents and is incentive compatible for the coalition if (a_H, a_H) is enforced as a perfect public equilibrium of the agents' repeated game described in Assumptions 1 and 2 while (a_L, a_L) is the non-cooperative Nash equilibrium. A coordination-inducing contract (α, β, γ) satisfies the following team incentive compatibility and coordination constraints,

$$\overline{u}_1(a_H, a_H|\mathbb{C}) + \overline{u}_2(a_H, a_H|\mathbb{C}) \geq \overline{u}_1(a_1, a_2|\mathbb{C}) + \overline{u}_2(a_1, a_2|\mathbb{C})$$
$$\forall (a_1, a_2) \in \{(a_H, a_L), (a_L, a_H), (a_L, a_L)\} \qquad (IC_c)$$

and

$$\begin{aligned}&(1-\delta)\overline{u}_i(a_H, a_H|\mathbb{C}) + \delta\left((1-q_i^n)\overline{u}_i(a_H, a_H|\mathbb{C}) + q_i^n\overline{u}_i(a_L, a_L|\mathbb{C})\right)\\ &\quad\geq (1-\delta)\overline{u}_i(a_i = a_L, a_j = a_H|\mathbb{C}) + \delta\left((1-q_i^d)\overline{u}_i(a_H, a_H|\mathbb{C}) + q_i^d\overline{u}_i(a_L, a_L|\mathbb{C})\right). \quad (3)\end{aligned}$$

Given the model specification, (3) can be written as

$$\alpha + h(\delta, \Delta)\beta \geq \eta \qquad (CI),$$

where $h(\delta, \Delta) \equiv \frac{\delta\Delta}{1-\delta+\delta\Delta}$ measures the quality of the coalition, which is increasing in the discount factor and the accuracy of the agents' mutual inference. With exogenous inference, $\Delta = \Delta_s$. With endogenous inference $\Delta = \Delta^*$ given s^*, and the (EF) constraint is identical to the (CI) constraint.

In addition, the agents on the equilibrium path enter the punishment phase in their repeated relationship with a slightly positive probability q_i^n. It then matters how the contract shapes the punishment phase even on the equilibrium path. Suppose that the agents are protected by the law to have the option to opt out from the contract at the end of each period t, had them expected the contract to yield negative payoffs. If the agents were induced to coordinate by a contract satisfying binding (CI) and (IR_n), the agents at the punishment phase of their repeated game would have incentive to leave the contractual relationship and seek outside options from which each of them earns exactly the reservation utility. That is, the punishment phase to enforce the coalition given a contract satisfying (IR_n) is not credible. The relevant participation constraint to induce coordination is thus the interim individual rationality at the punishment phase[8], i.e.

$$\overline{u}_i(a_L, a_L|\mathbb{C}) \geq 0 \qquad (IR_c).$$

The fixed payment in the contract, γ, is adjusted to satisfy constraint (IR_c) without affecting team incentive compatibility nor the coordination constraint. The principal's problem to find the optimal coordination-inducing contract is to minimize expected transfer to the agents, which is reduced to minimizing the sum of risk premium and expected loss of transfer due to a false detection of unilateral deviation given binding (IR_c), subject to team incentive compatibility and coordination constraints, i.e.

$$\mathscr{P}_{CI}: \qquad \min_{\alpha,\beta}(\alpha+\beta)(a_H-a_L)+R(\alpha,\beta)\equiv(\alpha+\beta)(a_H-a_L)+\frac{r}{2}(\alpha^2+\beta^2+2\sigma\alpha\beta)$$

subject to (IC_c) and (CI). Note that any $\beta < 0$ tightens team incentive compatibility as well as enforcement of coalition. RPE ($\beta < 0$) is never optimal to induce coordination.

When $\delta \to 1$, the coordination constraint is redundant given team incentive compatibility. We obtain the cooperative benchmark in which the agents can write a verifiable side contract on actions. The cooperative benchmark contract has the form of JPE, to motivate the team of agents as one single agent exerting multiple productive independent actions, which is independent of the correlation of measurement noise. It is straightforward from (CI) that the benchmark JPE is vulnerable to free-riding as long as each agent's future payoff is discounted.

As a brief discussion, the optimality for the principal offering a stationary contract relies on the symmetric information of history-independent future outcomes. This holds if collusion is deterred, but fails if coordination is induced, because of the imperfect nature of the agents' mutual inference. The agents know whether they are in a continuation or in a punishment phase in their repeated relationship, whereas the principal does not. The outcome of imperfect monitoring is public among the agents, but private to the coalition of agents. The principal thus has incentive to screen the continuation phase from the punishment phase, by offering a menu of contract $\{\mathbb{C}_{co}, \mathbb{C}_p\} = \{(\alpha_{co}, \beta_{co}, \gamma_{co}), (\alpha_p, \beta_p, \gamma_p)\}$ such that the coalition of agents in the continuation phase voluntarily act according to $\mathbb{C}_{co}$, and the agents in the punishment

[8]If the agents are not protected by the legal right to opt out from the contract at an interim stage, what is relevant to the coordination-inducing contract is the ex-ante participation constraint, $(1-\delta)\overline{u}_i(a_H, a_H|\mathbb{C}) + \delta\left((1-q_i^n)\overline{u}_i(a_H, a_H|\mathbb{C}) + q_i^n\overline{u}_i(a_L, a_L|\mathbb{C})\right) \geq 0$. This allows negative payoff in the punishment phase and reduces the cost to induce the agents to coordinate actions, without affecting the principal's optimal contract to deter collusion.

phase voluntarily follow $\mathbb{C}_p$.

$$\overline{u}_i(a_H, a_H|\mathbb{C}_{co}) \geq \overline{u}_i(a_L, a_L|\mathbb{C}_p) \qquad (TT_{co})$$

$$\overline{u}_i(a_H, a_H|\mathbb{C}_p) \geq \overline{u}_i(a_L, a_L|\mathbb{C}_{co}) \qquad (TT_p)$$

For any $\mathbb{C}_p$ such that (TT_p) slacks, the principal is better off lowering γ_p such that (TT_p) is binding, for it relaxes (TT_{co}) and better enforce the coalition. Constraint (TT_p) is thus binding, given which constraint (CI) is unaffected with the presence of screening. If the phases of the repeated game are publicly observable, the coordination-inducing contract solving $\mathscr{P}_{CI}$ is chosen to induce coordination in the continuation phase, and the non-cooperative benchmark contract is offered to minimize risk premium in the punishment phase. Given binding (TT_p) and (CI), (TT_{co}) satisfies if $(1-h(\delta,\Delta))\beta_{CI} \geq -\sigma\eta$. Otherwise, the contract has $\beta_{co} \geq \beta_{CI} > 0$ and $0 > \beta_p \geq -\sigma\eta$ instead. The solution of $\mathscr{P}_{CI}$ is thus qualitatively robust to screening the punishment phase.

3 Optimal Team Incentives

3.1 Mutual Inference versus Risk Sharing

To see the trade-off behind the principal's decision on the optimal contract, observe the optimality condition with respect to the linear contracting parameters (α, β). Deterring collusion, (α, β) is optimally decided from $\mathscr{P}_{CP}$ such that

$$\underbrace{\frac{R_\beta(\alpha,\beta)}{R_\alpha(\alpha,\beta)}}_{MRS\,in\,Premium} = \underbrace{h(\delta,\Delta)}_{Quality\,of\,Coalition}. \tag{4}$$

Inducing coordination, (α, β) is optimally decided decided from $\mathscr{P}_{CI}$ such that

$$\underbrace{\frac{(a_H - a_L) + R_\beta(\alpha,\beta)}{(a_H - a_L) + R_\alpha(\alpha,\beta)}}_{MRS\,in\,Modified\,Premium} = \underbrace{h(\delta,\Delta)}_{Quality\,of\,Coalition}. \tag{5}$$

Subscripts denote partial derivative to avoid cluster, and $\Delta = \Delta_s$ if inference is formed exogenously while $\Delta = \Delta^*$ with endogenous inference.

The left-hand-side of (4) is straightforward and well-explored in the literature, i.e. the marginal rate of substitution in minimizing risk premium to the agents when collusion is deterred. The left-hand-side of (5) is the marginal rate of substitution in minimizing a modified risk premium to the agents when coordination is induced. There are two sources of risk premium when coordination is induced, one attributed to the stochastic realization of output on which the transfer is contingent (captured by the second term) and the other attributed to the credible punishment phase from a false detection of unilateral deviation (captured by the first term). These mark that one consideration of the choice of contract is to provide incentive through improvement of risk sharing.

At the right-hand side of both equations is the quality of the agents' coalition independent of contract offered, $h(\delta, \Delta) \equiv \frac{\delta\Delta}{1-\delta+\delta\Delta}$. Its interpretation is based on the observation that it increases in the interaction of the agents' patience and their accuracy of mutual inference ($\delta\Delta$), which in turn measures the discounted difference in likelihood to enter a punishment phase in the agents' repeated game. Such quality of coalition is increasing in the correlation of measurement noise through the accuracy of inference Δ. Higher $h(\delta, \Delta)$, stronger is the expected punishment for unilateral deviation, and the coordinated actions are enforced more firmly. This marks another consideration of the choice of contract, to provide incentive through inducing the agents' coordination or through deterring their collusion enforced with their mutual inference.

At the incentive compatible IPR with $\beta = 0$ and $\alpha = \eta$, if $h(\delta, \Delta) < \frac{R_\beta(\eta,0)}{R_\alpha(\eta,0)} = \sigma$, the quality of the agents' coalition is sufficiently weak. Offering a contract with $\beta < 0$ to deter collusion is valuable to the principal as its improvement in risk sharing is more prominent than its increased cost on deterring collusion. Collusion-proof RPE is optimal if $h(\delta, \Delta) < \sigma$. If $h(\delta, \Delta) > \frac{(a_H-a_L)+R_\beta(\eta,0)}{(a_H-a_L)+R_\alpha(\eta,0)} = \frac{(a_H-a_L)+r\eta\sigma}{(a_H-a_L)+r\eta} \equiv \varphi(\sigma)$, the quality of the agents' coalition is sufficiently firm. Offering a contract with $\beta > 0$ to induce coordination is valuable to the principal as it induces implicit incentive governed by the agents' mutual inference, which is sufficiently accurate, at a relatively small cost from inefficient risk sharing. Coordination-inducing JPE is optimal if $h(\delta, \Delta) > \varphi(\sigma)$. Otherwise, if $\sigma < h(\delta, \Delta) < \varphi(\sigma)$, the coalition is not only insufficiently weak for the cost of deterring collusion under $\beta < 0$ to be relatively small comparing to the benefit from risk sharing, but also insufficiently firm for the benefit of implicit incentive governed by the agents' mutual inference to be relatively large to outweigh the cost from inefficient risk sharing under $\beta > 0$. IPR is thus optimal if $\sigma < h(\delta, \Delta) < \varphi(\sigma)$.

Note that the principal's contracting problem differs when the agents form exogenous or endogenous inference only on the coordination (or the collusion-proofness) constraint through the quality of coalition $h(\delta, \Delta)$. How mutual inference is made shapes the quality of coalition differently without directly affecting the risk sharing role of the contract.

3.2 Team Incentives with Exogenous Inference

The quality of the agents' coalition relates to the agents' patience as well as the correlation of measurement noise. The optimal contract with exogenous inference is characterized in terms of the former in the following proposition.

Proposition 1. *It is optimal to deter collusion with a RPE if* $\delta < \delta_{CP} \equiv \frac{\sigma}{(1-\sigma)\Delta_s+\sigma}$*, to induce coordination with a JPE if* $\delta > \delta_{CI} \equiv \frac{r\sigma\eta+(a_H-a_L)}{r((1-\sigma)\Delta_s+\sigma)\eta+(a_H-a_L)}$*, and to provide incentive by IPR if* $\delta_{CP} < \delta < \delta_{CI}$*. For* $\delta < \delta_{CP}$*, the optimal collusion-proof contract* $\mathbb{C}_{CP} = (\alpha_{CP}, \beta_{CP}, \gamma_{CP})$ *has* $\beta_{CP} = \frac{(h(\delta,\Delta_s)-\sigma)\eta}{1+h(\delta,\Delta_s)^2-2\sigma h(\delta,\Delta_s)} < 0$*. For* $\delta > \delta_{CI}$*, the optimal coordination-inducing contract* $\mathbb{C}_{CI} = (\alpha_{CI}, \beta_{CI}, \gamma_{CI})$ *has* $\beta_{CI} = \frac{r(h(\delta,\Delta_s)-\sigma)\eta-(1-h(\delta,\Delta_s))(a_H-a_L)}{r(1+h(\delta,\Delta_s)^2-2\sigma h(\delta,\Delta_s))} > 0$*.*

Proof. Appendix A.2. □

Characterization of contract by the discount factor is a natural first thought with repeated interaction available. What's more to our interest is, given a fixed discount factor, how the trade-off behind the principal's decision depends on the correlation of measurement noise. Comparing with the conventional literature, in which the agents' monitoring ability is assumed to be independent of how they are stochastically correlated ($h(\delta, \Delta_s)$ does not depend on σ), it is optimal to offer coordination-inducing JPE for sufficiently small correlations, collusion-proof RPE for sufficiently large correlations, and IPR otherwise. However, with the agents' monitoring inferred from correlated stochastic output, the accuracy of inference depends on the correlation of measurement noise ($h(\delta, \Delta_s)$ depends on σ through Δ_s). Define the correlation elasticity of exogenous inference as $\frac{\partial \Delta_s}{\partial \sigma}\frac{\sigma}{\Delta_s}$. The optimal form of contract may not be monotonic in the correlation parameter, as suggested in the following proposition.

Proposition 2. *There exists some exogenous parameters such that the correlation elasticity of exogenous inference is sufficiently high that* $h(\delta, \Delta_s) > \varphi(\sigma)$ *for* $\sigma \in (\widehat{\sigma}, \overline{\sigma})$*,* $0 < \widehat{\sigma} < \overline{\sigma} < 1$*, and* $h(\delta, \Delta_s) < \varphi(\sigma)$ *otherwise. That is, for* $\sigma \in (\widehat{\sigma}, \overline{\sigma})$*, coordination-inducing JPE is optimal.*

Proof. Appendix A.3. □

The proposition essentially suggests that there exists some parametric values of the difference in the level of private actions ($a_H - a_L$), those of the commonly-adopted standard for statistical significance (s), and those of the discount factor (δ), given which the optimal contract is not monotonic in the correlation of measurement noise. Specifically, these parameters are such that the correlation elasticity of inference accuracy defined as $\frac{\partial \Delta_s}{\partial \sigma}\frac{\sigma}{\Delta_s}$ is sufficiently high for some intermediate levels of correlation. This intuitively implies that stochastic correlation plays a crucial role in the improvement of mutual inference, which results in a prominent improvement on the quality of coalition ($h(\delta, \Delta_s)$). It is then optimal for the principal to induce the sufficiently patient agents to coordinate with a JPE for intermediate levels of correlation.

For a counter example, if a deviation is too obvious even with a low correlation ($a_H - a_L$ being too large), the magnitude of deviation itself is a key source of inference accuracy. Correlation of measurement noise is essential in improvement of mutual inference only when such deviation is not too obvious from the outset. For another counter example, if the agents adopt a weak standard for the detection of deviation (s being too small), even with a low correlation the agents are convinced of the others' deviation with a high probability. It is then the conservative standard for detection of deviation that plays a crucial role in enforcing the agents' coalition. Correlation of measurement noise is essential in improvement of the quality of coalition only when the agents adopt a relatively strict standard for statistical significance.

Figure 1 illustrates one of the examples that Proposition 4 holds. The upper panel illustrates the parametric space of the discount factor ($\delta(\sigma)$) and correlation (σ) such that coordination-inducing JPE, IPR, or collusion-proof RPE is optimal. The lower panel illustrates the benefit from mutual inference (or symmetrically, the cost to deter collusion) captured by the quality of coalition ($h(\delta, \Delta_s)$), the cost from risk sharing ($\varphi(\sigma)$) under JPE, as well as the benefit from risk sharing (σ) under RPE. By relating the two panels we are able to characterize how the relative magnitude of the effects on improvement of quality of coalition and improvement of risk sharing determines the optimal form of contract.

Fix the agents' patience at the level indicated by the dotted line in the upper panel. For sufficiently small correlations of measurement noise ($\sigma < \underline{\sigma}$), neither the effect of risk sharing nor that of mutual inference is significant, $\sigma < h(\delta, \Delta_s) < \varphi(\sigma)$ as illustrated in the lower panel. IPR is optimal. For relatively small correlations, an increase in the correlation improves risk sharing under RPE more prominently than its improvement on quality of coalition via accuracy of mutual inference to an extent that $\sigma > h(\delta, \Delta_s)$. The benefit from risk sharing exceeds the cost from deterring collusion, so collusion-proof RPE is optimal. For relatively large correlations, improvement on quality of coalition via accuracy of mutual inference with respect to a higher correlation

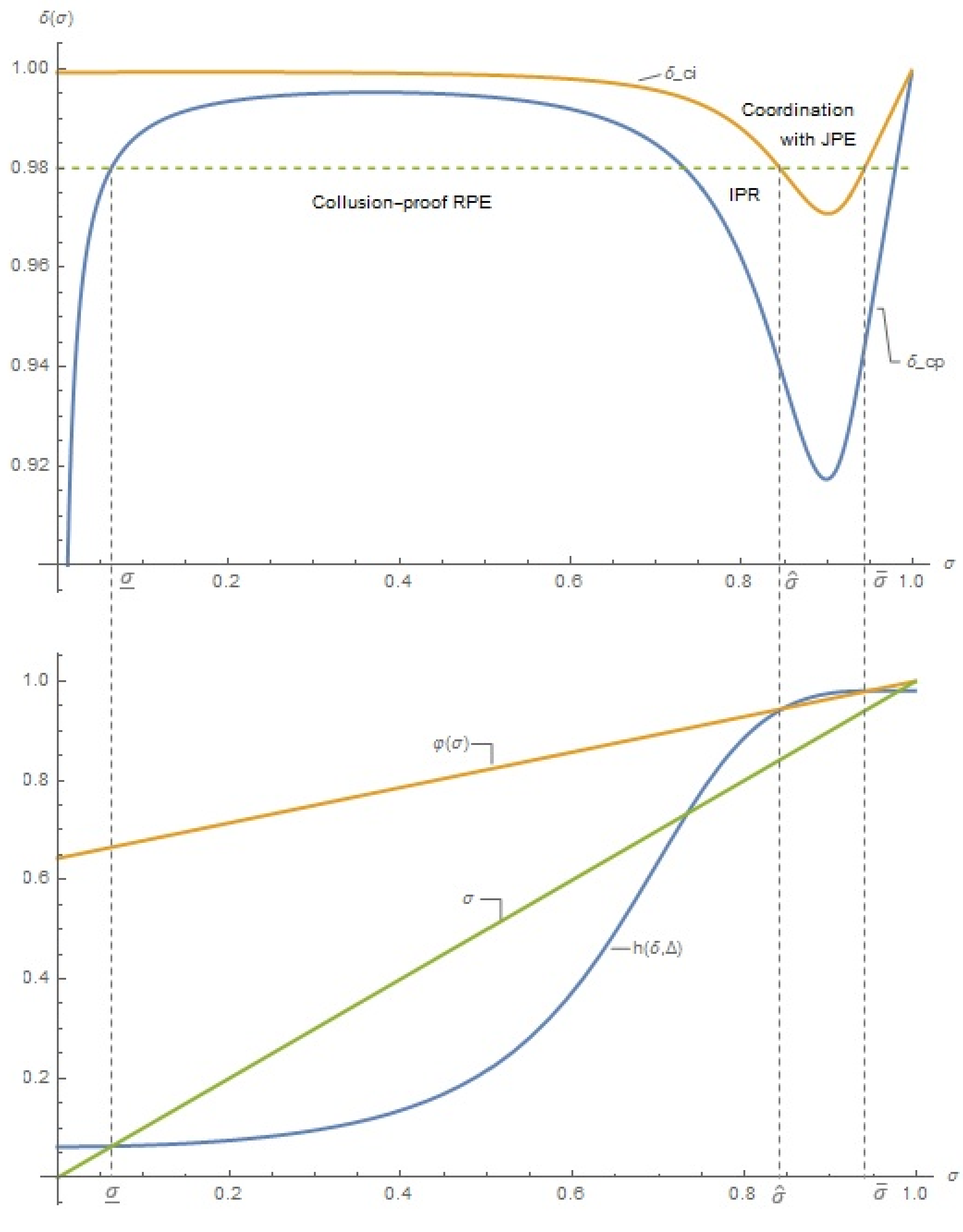

Figure 1: Optimal Contract
Figure 1 is drawn with $s = 6$, $a_H - a_L = 3$, $c_H - c_L = 5$.

is relatively significant to the risk sharing effect under JPE to an extent that $h(\delta, \Delta_s) > \varphi(\sigma)$. The benefit from implicit incentive through mutual inference exceeds the cost from inefficient risk sharing, so coordination-inducing JPE is optimal. For sufficiently large correlations, the agents' mutual inference is so accurate that it has little room for improvement, and the risk sharing effect of a higher correlation exceeds the effect from improvement of inference accuracy. Collusion-proof RPE is then optimal.

The above result stems from the two roles of a higher correlation of measurement noise, which are in favor of different contractual forms respectively. A higher correlation of measurement noise contributes to a more efficient risk sharing along with RPE. RPE, however, introduces incentive for the agents to collude to exert a lower level of effort, and with high correlation of measurement noise such collusion can be firmly enforced with a relatively accurate mutual inference. A higher correlation of measurement noise also contributes to the improved accuracy of mutual inference, which better enforces the agents' coordination along with JPE. JPE, nevertheless, is subject to a higher cost for the principal in terms of less efficient risk sharing. If stochastic correlation does not play an essential role in the improvement of mutual inference, its dominant effect is on improving risk sharing and lowering risk premium under RPE. There would be monotonicity in the optimal form of contract. With the essential role of the stochastic correlation in improvement of both quality of coalition and risk sharing, how the optimal form of contract differs in the correlation of measurement noise depends on how powerful each improvement effect is.

For intermediate correlations $\sigma \in (\underline{\sigma}, \overline{\sigma})$, $\overline{\sigma} < 1$, a higher correlation results in a sharper improvement on quality of coalition via accuracy of mutual inference than its improvement on risk sharing, captured by an overall steeper $h(\delta, \Delta_s)$ than both $\varphi(\sigma)$ and σ in the lower panel. Collusion-proof RPE is optimal for sufficiently small correlations within this range, IPR is optimal for intermediate levels of correlation, and the principal finds it optimal to induce coordination among the agents with a JPE for sufficiently large correlations to take advantage of their mutual inference at the expense of efficient risk sharing. In comparison with the literature, this is consistent to Rayo (2007)'s claim that a signal of higher accuracy is in favor of JPE that induces implicit incentive. It meanwhile contradicts to the conventional proposition introduced by Holmström and Milgrom (1990) that a higher correlation of measurement noise is in favor of RPE for its risk efficiency. This is due to the sharper effect on improvement of mutual inference that outweighs the risk sharing concern emphasized by Holmström and Milgrom (1990).

For sufficiently high correlations $\sigma \in (\widehat{\sigma}, 1)$, $0 < \widehat{\sigma} < \overline{\sigma}$, a higher correlation results in a sharper improvement on risk sharing than its improvement on quality of coalition, captured by an overall flatter $h(\delta, \Delta_s)$ than both $\varphi(\sigma)$ and σ. Coordination-inducing JPE is optimal for sufficiently small correlations within this range, IPR is optimal for intermediate levels of correlation, and the principal finds it optimal to deter collusion with a RPE for sufficiently large correlations as the cost to induce coordination from inefficient risk sharing is relatively higher than the benefit from mutual inference. In comparison with the literature, this is consistent to the proposition of Holmström and Milgrom (1990) that a higher correlation of measurement noise is in favor of RPE. It meanwhile contradicts to the proposition of Rayo (2007) as more accurate signal due to a higher correlation does not favor coordination-inducing JPE, because a sharper improvement on risk sharing dominates.

From the perspective of informativeness, using the agents mutual inference implicitly or using a contract with RPE explicitly are both imperfectly informative of the agents' hidden actions. Only for intermediate levels of correlation $(\widehat{\sigma}, \overline{\sigma})$ would the agents' mutual inference be relatively more informative than the contracted outputs. For sufficiently low correlations, mutual inference is insufficiently accurate to outweigh the cost of inefficient risk sharing for a contract with JPE. For sufficiently high correlations, both mutual inference and the contracted outputs are highly informative, yet using the latter in a contract with RPE is relatively risk efficient.

3.3 Team Incentives with Endogenous Inference

When the agents make endogenous inference, shown in the following lemma that they choose the optimal threshold of believed deviation which maximizes the accuracy of mutual inference.

Lemma 2. *The optimal threshold of believed unilateral deviation is* $s^* \in \arg\max_s \Delta$, *at which* (EF) *is binding. Both* s^* *and* Δ^* *are increasing in* σ.

Proof. Appendix A.4. □

More correlated is the measurement noise, the agents are able to make a more accurate inference given any fixed threshold s. This allows them to trust each other more in terms of relaxing the standard of a believed deviation. This further improves the accuracy of mutual inference not only through increasing the probability of correct detection of a unilateral deviation but also through reducing the probability of false detection.

The quality of the agents' coalition relates to the agents' patience as well as the correlation of measurement noise. The optimal contract with endogenous inference is characterized in terms of the former in the following proposition.

Proposition 3. *It is optimal to deter collusion with a RPE if* $\delta < \widehat{\delta}_{CP} \equiv \frac{\sigma}{(1-\sigma)\Delta^*+\sigma}$*, to induce coordination with a JPE if* $\delta > \widehat{\delta}_{CI} \equiv \frac{r\sigma\eta+(a_H-a_L)}{r((1-\sigma)\Delta^*+\sigma)\eta+(a_H-a_L)}$*, and to provide incentive by IPR if* $\widehat{\delta}_{CP} < \delta < \widehat{\delta}_{CI}$*. With an endogenous inference, coordination-inducing JPE is optimal for a larger set of exogenous parameters and collusion-proof RPE is optimal for a smaller set of parameters as* $\widehat{\delta}_{CI} \leq \delta_{CI}$ *and* $\widehat{\delta}_{CP} \leq \delta_{CP}$*. For* $\delta < \widehat{\delta}_{CP}$*, the optimal contract has* $0 > \widehat{\beta}_{CP} > \beta_{CP}$*, and for* $\delta > \widehat{\delta}_{CI}$*, the optimal contract has* $\widehat{\beta}_{CI} > \beta_{CI} > 0$*.*

Proof. Appendix A.5. □

With endogenous inference, the agents adopt the threshold of believed unilateral deviation that maximizes the accuracy of mutual inference, which enforces their coalition at a lower cost from the agents' perspective. On the other hand, the endogenous threshold of believed unilateral deviation does not directly affect the risk sharing role of the correlation of measurement noise. From the principal's perspective, an endogenous inference lowers the contracting cost of a coordination-inducing JPE and increases that of a collusion-proof RPE. It is thus optimal to induce coordination for a larger set of exogenous environment and to deter collusion for a smaller set of exogenous environment.

When it is optimal to induce coordination, the principal proposes a contract with a steeper JPE than that with exogenous inference. The steeper JPE provides more powerful incentive for the agents to utilize the relatively accurate inference. When it is optimal to deter collusion, the principal proposes a contract with a flatter RPE than that with exogenous inference. The flatter RPE deters the agents from utilizing the more accurate endogenous inference at the expense of less efficient risk sharing.

Define the correlation elasticity of endogenous inference as $\frac{\partial \Delta^*}{\partial \sigma}\frac{\sigma}{\Delta^*}$. The endogenous inference is less elastic than its exogenous counterpart, except for extremely high or low difference in the level of actions ($a_H - a_L$ extremely large or small) given which how inference is made is trivial. When the optimal contract with exogenous inference is non-monotonic in correlation, the optimal contract with endogenous inference can be monotonic in the correlation parameter, as suggested in the following proposition.

Proposition 4. *The endogenous inference is less elastic than the exogenous counterpart. The optimal contract is more likely to exhibit monotonicity in the correlation of measurement noise, when endogenous inference is adopted.*

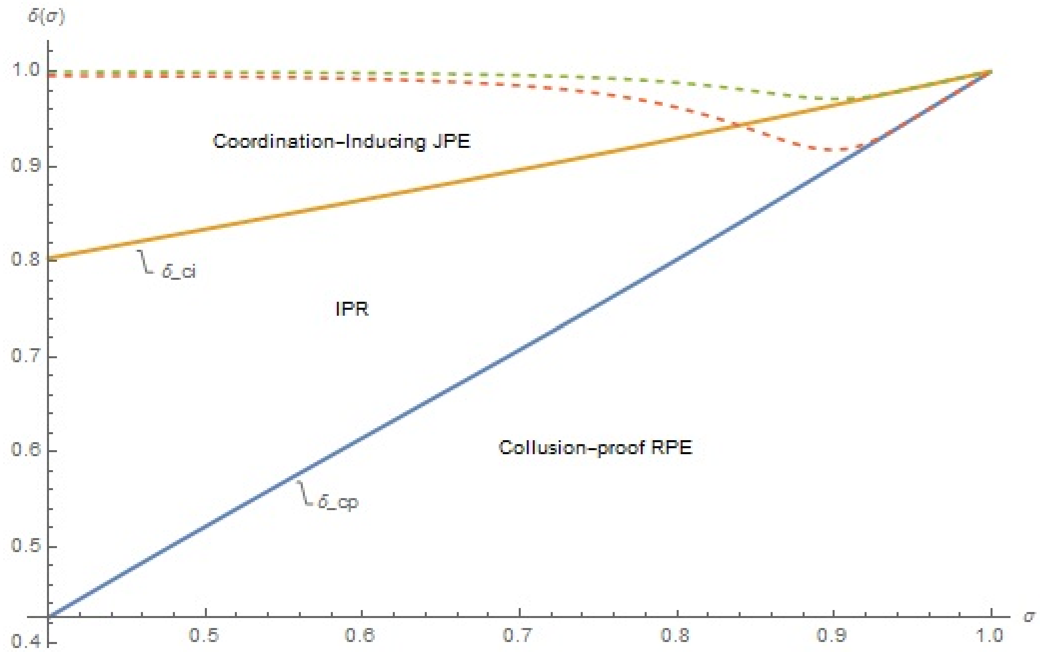

Figure 2: Endogenous Belief of Deviation
Figure 2 is drawn with $a_H - a_L = 3$, $c_H - c_L = 5$.

Proof. Appendix A.6. □

Recall that the non-monotonicity result in Proposition 2 relies on the correlation of measurement noise playing an essential role to improve the accuracy of inference. With endogenous inference, this role is replaced by the optimal choice of intrinsic trust (the optimal threshold of believed unilateral deviation). A tighter (more relaxed) threshold of believed deviation is chosen when the agents are less (more) stochastically correlated, so that the endogenous inference is most accurate at every level of correlation. Improvement of the accuracy of inference with respect to the correlation of measurement noise is thus relatively mild than that of an exogenous inference.

Given the same parameters as in Figure 1, Proposition 3 and 4 are illustrated in Figure 2. An endogenously inference yields a larger region in the σ-δ space for coordination-inducing JPE to be optimal and a smaller region for collusion-proof RPE to be optimal. In addition, the optimal contract is monotonic in the correlation of measurement noise at sufficiently high discount factor. Coordination-inducing JPE is optimal for sufficiently low correlations, and collusion-proof RPE is optimal for sufficiently correlated agents. Intuitively, the agents endogenously set a stricter standard of a believed deviation when they are less correlated. This greatly improves their quality of coalition with a small correlation of measurement noise that outweighs the cost from inefficient risk sharing under a contract with JPE. With a sufficiently high correlation of measurement noise, this improvement in the quality of coalition is relatively mild while the cost efficiency of a contract with RPE in terms of risk sharing increases more prominently.

With endogenous inference, the form of optimal team incentive with respect to the correlation of measurement noise qualitatively aligns with the prediction of Holmström and Milgrom (1990). It contradicts to the prediction of Rayo (2007) that coordination-inducing JPE is optimal for sufficiently accurate signals of actions between the agents. A higher correlation improves the accuracy of the endogenous inference, yet this improvement is outweighed by the stronger effect of improved risk sharing. Collusion-proof RPE is optimal even when the inference becomes more accurate.

4 Conclusion

Explicitly modeling how the agents make inference on each other through the correlated stochastic measurements of private actions, I argue that the agents' monitoring accuracy is increasing in the correlation of measurement noise. This is in contrast to the independency assumed in the conventional studies of optimal team incentives subject to the agents' coordination/collusion. More correlated the measurement noise is, incentive provision through filtering out common noise is less costly, so is the enforcement of the agents' coalition. Thus, the following two propositions in the literature do not co-exist: i) collusion-proof RPE is optimal for sufficiently large correlations of measurement noise due to its risk sharing advantage, emphasized by Holmström and Milgrom (1990), and ii) coordination-inducing JPE is optimal for sufficiently accurate monitoring among the agents, emphasized by Rayo (2007). For some parameters, I showed consistency to Holmström and Milgrom (1990) for relatively large correlations of measurement noise and consistency to Rayo (2007) for relatively small correlations of measurement noise, when the agents make inference based on an exogenous standard of significance. With endogenous belief of unilateral deviation among the agents, the optimal contract regains monotonicity in the correlation of measurement noise consistent to the prediction of Holmström and Milgrom (1990) and in contradiction to that of Rayo (2007).

This paper attempted to incorporate repeated interaction with imperfect public monitoring into the study of optimal provision of team incentives, and it is expecting a generalization in the future. Specifically, a general condition on the functional forms and the density of measurement noise for which the proposition of this paper holds is of interest. In addition, with a continuum of actions, the problem becomes more challenging as the accuracy of mutual inference is related to the magnitude of unilateral deviation as well as to the correlation of measurement noise. It is unclear whether the accuracy of mutual inference is positively related to the correlation of measurement noise with a continuum of hidden actions.

Appendix

A Proof of Propositions

A.1 Proof of Lemma 1

Suppose that the agents coordinate to exert (a_i^0, a_j^0) when the non-cooperative Nash equilibrium is $(a_i^{'}, a_j^{'})$. Agent j detects agent i's deviation correctly with probability

$$q_i^d = \begin{cases} Pr\left(x_{it}(a_i^{'})|x_{jt}(a_j^0) < a_{it}^0 + \sigma(x_{jt} - a_{jt}^0) - s\sqrt{1-\sigma^2}\right) = \Phi\left(\frac{a_i^0 - a_i^{'}}{\sqrt{1-\sigma^2}} - s\right) & for\, a_i^0 > a_i^{'} \\ Pr\left(x_{it}(a_i^{'})|x_{jt}(a_j^0) > a_{it}^0 + \sigma(x_{jt} - a_{jt}^0) + s\sqrt{1-\sigma^2}\right) = 1 - \Phi\left(\frac{a_i^0 - a_i^{'}}{\sqrt{1-\sigma^2}} + s\right) & for\, a_i^0 < a_i^{'} \end{cases},$$

where $\Phi(\cdot)$ is the standard normal CDF. It is increasing in correlation as $\frac{\partial q_i^d}{\partial \sigma} = \phi(\cdot)\frac{\sigma|a_i^0 - a_i^{'}|}{\sqrt{(1-\sigma^2)^3}} > 0$, where $\phi(\cdot)$ is the standard normal PDF. Agent j mistakenly detects agent i of deviation with probability

$$q_i^n = \begin{cases} Pr\left(x_{it}(a_i^0)|x_{jt}(a_j^0) < a_{it}^0 + \sigma(x_{jt} - a_{jt}^0) - s\sqrt{1-\sigma^2}\right) = \Phi(-s) & for\, a_i^0 > a_i^{'} \\ Pr\left(x_{it}(a_i^0)|x_{jt}(a_j^0) > a_{it}^0 + \sigma(x_{jt} - a_{jt}^0) + s\sqrt{1-\sigma^2}\right) = 1 - \Phi(s) & for\, a_i^0 < a_i^{'} \end{cases},$$

which is straightforward that $\frac{\partial q_i^n}{\partial \sigma} = 0$.

□

A.2 Proof of Proposition 1

The optimal contract to implement (a_H, a_H) and to deter collusion solves the expected transfer minimization problem $\min_{\alpha,\beta,\gamma} \mathbb{E}(\alpha x_i + \beta x_j + \gamma)$ subject to (IR_n), (IC_n), and (CP), $i = 1, 2$ and $i \neq j$. Given binding (IR_n), $(\alpha + \beta)a_H + \gamma = R(\alpha, \beta)$, this contracting problem is reduced to

$$\mathscr{P}_{CP}: \qquad \min_{\alpha,\beta} R(\alpha, \beta) \equiv \frac{r}{2}(\alpha^2 + \beta^2 + 2\sigma\alpha\beta)$$

subject to (IC_n) and (CP). $\Delta = \Delta_s$ with exogenous inference and $\Delta = \Delta^*$ with endogenous inference.

Constraint (IC_n) can be reduced to $\alpha \geq \eta$, where $\eta \equiv \frac{c_H - c_L}{a_H - a_L}$. If $\beta \geq 0$, $\alpha + h(\delta,\Delta)\beta \geq \alpha \geq \eta$ as $0 \leq h(\delta,\Delta) \equiv \frac{\delta\Delta}{1-\delta+\delta\Delta} \leq 1$. (IC_n) is binding while (CP) slacks except at $\beta = 0$. With binding (IC_n), ignoring (CP) and $\beta \geq 0$, the problem is that of the non-cooperative benchmark, in which the optimal $\beta_n = \arg\min_\beta \eta^2 + \beta^2 + 2\sigma\eta\beta = -\sigma\eta < 0$. Among all $\beta \geq 0$, $\beta_{CP} = 0$ is thus optimal to deter collusion. If $\beta < 0$, $\alpha > \alpha + h(\delta,\Delta)\beta \geq \eta$. (CP) is the only binding constraint, i.e. $\alpha + h(\delta,\Delta)\beta = \eta$. Solving the first order condition, $(h(\delta,\Delta)\beta - \eta)\, h(\delta,\Delta) + \beta + \sigma\,(\eta - 2h(\delta,\Delta)\beta) = 0$, the optimal contract with exogenous inference has $\beta_{CP} = \frac{(h(\delta,\Delta_s)-\sigma)\eta}{1+h(\delta,\Delta_s)^2-2\sigma h(\delta,\Delta_s)} < 0$ if $h(\delta,\Delta_s) < \sigma$, and $\beta_{CP} = 0$ if $h(\delta,\Delta_s) \geq \sigma$. Rearranging $h(\delta,\Delta_s) < \sigma$ yields $\delta < \delta_{CP} \equiv \frac{\sigma}{(1-\sigma)\Delta_s+\sigma}$.

The optimal contract to implement (a_H, a_H) and to induce coordination solves the expected transfer minimization problem $\min_{\alpha,\beta,\gamma} \sum_{i=1,2} (\mathbb{E}(\alpha x_i + \beta x_j + \gamma))$ subject to (IR_c), (IC_c), and (CI). Given binding (IR_c), $(\alpha+\beta)a_L + \gamma = R(\alpha,\beta)$, this contracting problem is reduced to

$$\mathscr{P}_{CI}: \qquad \min_{\alpha,\beta} (\alpha+\beta)(a_H - a_L) + R(\alpha,\beta) \equiv (\alpha+\beta)(a_H - a_L) + \frac{r}{2}(\alpha^2 + \beta^2 + 2\sigma\alpha\beta)$$

subject to (IC_c) and (CI). $\Delta = \Delta_s$ with exogenous inference and $\Delta = \Delta^*$ with endogenous inference.

Constraint (IC_c) can be reduced to $\alpha + \beta \geq \eta$. If $\beta \leq 0$, $\alpha + h(\delta,\Delta)\beta \geq \alpha + \beta \geq \eta$ as $0 \leq h(\delta,\Delta) \equiv \frac{\delta\Delta}{1-\delta+\delta\Delta} \leq 1$. $(IC_c)$is binding while (CI) slacks except at $\beta = 0$. With binding (IC_c), ignoring (CI) and $\beta \leq 0$, the problem is reduced to that of the cooperative benchmark in terms of the choice of (α,β), in which the optimal $\beta_c = \arg\min_\beta \eta^2 - 2(1-\sigma)(\eta - \beta)\beta = \frac{\eta}{2} > 0$. Among all $\beta \leq 0$, $\beta_{CI} = 0$ is thus optimal to induce coordination. If $\beta > 0$, $\alpha + \beta > \alpha + h(\delta,\Delta)\beta \geq \eta$. (CI) is the only binding constraint, $\alpha + h(\delta,\Delta)\beta = \eta$. Solving the first order condition, $(a_H - a_L)(1 - h(\delta,\Delta)) + r\,((h(\delta,\Delta)\beta - \eta)\, h(\delta,\Delta) + \beta + \sigma\,(\eta - 2h(\delta,\Delta)\beta)) = 0$, the optimal contract with exogenous inference has $\beta_{CI} = \frac{r(h(\delta,\Delta_s)-\sigma)\eta-(1-h(\delta,\Delta_s))(a_H-a_L)}{r(1+h(\delta,\Delta_s)^2-2\sigma h(\delta,\Delta_s))} > 0$ if $h(\delta,\Delta_s) - \sigma > \frac{(1-h(\delta,\Delta_s))(a_H-a_L)}{r\eta}$, and $\beta_{CI} = 0$ if otherwise. Rearranging $h(\delta,\Delta_s) - \sigma > \frac{(1-h(\delta,\Delta_s))(a_H-a_L)}{r\eta}$ yields $\delta > \delta_{CI} \equiv \frac{r\sigma\eta+(a_H-a_L)}{r((1-\sigma)\Delta_s+\sigma)\eta+(a_H-a_L)}$.

As $\frac{\sigma}{(1-\sigma)\Delta_s+\sigma} < 1$, $\delta_{CI} > \delta_{CP}$. For $\delta < \delta_{CP}$, the optimal collusion-proof RPE is preferred to an IPR and the optimal coordination-inducing contract has IPR. Collusion-proof RPE is thus preferred to coordination-inducing JPE for $\delta < \delta_{CP}$. For $\delta >$

δ_{CI}, the optimal coordination inducing JPE is preferred to an IPR and the optimal collusion-proof contract has IPR. Coordination-inducing JPE is thus preferred to collusion-proof RPE for $\delta > \delta_{CI}$. For $\delta_{CI} \geq \delta \geq \delta_{CP}$, coordination is induced with IPR and collusion is deterred with IPR; IPR is optimal.

□

A.3 Proof of Proposition 2

It is easy to verify that $\delta \gtreqless \delta_{CI}$ is merely a rearrangement of $h(\delta, \Delta_s) \gtreqless \varphi(\sigma)$ and $\delta \lesseqgtr \delta_{CP}$ is a rearrangement of $h(\delta, \Delta_s) \lesseqgtr \sigma$. Proposition 2 holds if there exists parameters s and $a_H - a_L$ such that δ_{CI} and δ_{CP} are diminishing in σ for $\sigma_0 < \sigma < \sigma^*$, and are increasing in σ for $\sigma_1 > \sigma > \sigma^*$, where $(\sigma_0, \sigma_1) \subseteq [0, 1]$. Differentiate $\delta_{CI} \equiv \frac{r\sigma\eta + (a_H - a_L)}{r((1-\sigma)\Delta_s + \sigma)\eta + (a_H - a_L)}$ with respect to σ finds $\frac{\partial \delta_{CI}}{\partial \sigma} \gtreqless 0$ if $\frac{\partial \Delta_s}{\partial \sigma} \frac{1-\sigma}{\Delta_s} \lesseqgtr \frac{r\eta + (a_H - a_L)}{\sigma r\eta + (a_H - a_L)} \equiv \frac{1}{\varphi(\sigma)}$, or $\frac{\partial \Delta_s}{\partial \sigma} \frac{\sigma}{\Delta_s} \lesseqgtr \frac{1}{\varphi(\sigma)} \frac{\sigma}{1-\sigma}$ for $\sigma \in (0, 1)$. Differentiate $\delta_{CP} \equiv \frac{\sigma}{((1-\sigma)\Delta_s + \sigma)}$ with respect to σ finds $\frac{\partial \delta_{CP}}{\partial \sigma} \gtreqless 0$ if $\frac{\partial \Delta_s}{\partial \sigma} \frac{1-\sigma}{\Delta_s} \lesseqgtr \frac{1}{\sigma}$, or $\frac{\partial \Delta_s}{\partial \sigma} \frac{\sigma}{\Delta_s} \lesseqgtr \frac{1}{1-\sigma}$ for $\sigma \in (0, 1)$. For all $\sigma \in (0, 1)$ such that δ_{CI} is increasing in σ, δ_{CP} is increasing in σ as well, and for all $\sigma \in (0, 1)$ such that δ_{CP} is diminishing in σ, δ_{CI} is diminishing in σ as well, because $\frac{\sigma}{\varphi(\sigma)} < 1$. To see the non-monotonicity result, it is sufficient to focus on the existence of parameters such that $\frac{\partial \delta_{CI}}{\partial \sigma} > 0$ for sufficiently high correlations and $\frac{\partial \delta_{CP}}{\partial \sigma} < 0$ for some intermediate correlations.

Since $\frac{1}{\varphi(\sigma)} \geq 1$ with equality at $\sigma = 1$, and $\lim_{\sigma \to 1} \Delta_s = 1$ if s is finite, $\frac{\partial \Delta_s}{\partial \sigma} \frac{\sigma}{\Delta_s} < \frac{1}{\varphi(\sigma)} \frac{\sigma}{1-\sigma}$ and $\frac{\partial \delta_{CI}}{\partial \sigma} > 0$ for sufficiently correlated measurement noise. If $a_H - a_L$ and s are such that there is sufficiently high percentage improvement in the accuracy of inference with respect to higher correlation, captured by $\frac{\partial \Delta_s}{\partial \sigma} \frac{\sigma}{\Delta_s} > \frac{1}{1-\sigma}$, $\frac{\partial \delta_{CP}}{\partial \sigma} < 0$ for some correlation. Given these parameters, $\frac{\partial \delta_{CP}}{\partial \sigma} < 0$ and $\frac{\partial \delta_{CI}}{\partial \sigma} < 0$ for $\sigma \in (\sigma_0, \sigma^*)$, with σ^* satisfying $\frac{\partial \Delta_s}{\partial \sigma} \frac{\sigma}{\Delta_s} = \frac{1}{1-\sigma}$. There then exists $\delta \in (\delta(\sigma^*), \min\{\delta(\sigma_0), \delta(\sigma_1)\})$ at which $h(\delta, \Delta_s) > \varphi(\sigma)$ for $\sigma \in (\widehat{\sigma}, \overline{\sigma})$ and $h(\delta, \Delta_s) < \varphi(\sigma)$ otherwise, where $(\widehat{\sigma}, \overline{\sigma}) \subseteq (\sigma_0, \sigma_1)$. Examples of such parameters are easily found, Figure 1 as one of which.

□

A.4 Proof of Lemma 2

Given a contract that deters collusion with accuracy of inference $\Delta < \max_s \Delta$, the agents have incentive to adjust s to raise the accuracy of inference to $\widetilde{\Delta} > \Delta$ such that (EF) is satisfied at $\widetilde{\Delta}$, which allows them to benefit from collusion. A contract is collusion-proof only if (EF) is weakly violated at $\max_s \Delta$. To induce coordination, the principal is better off to implement a higher Δ that satisfies (EF). If the contract $\mathbb{C}$ is such that (EF) holds with equality at $\Delta < \max_s \Delta$, the principal has incentive to lower β, anticipating that the agents will have incentive to adjust s to raise the accuracy of inference to $\widetilde{\Delta} > \Delta$ such that (EF) is satisfied at $\widetilde{\Delta}$. Coalition would be enforced to exert (a_H, a_H) at a lower transfer. If it is optimal for her to induce coordination, it is optimal for her to implement the level of trust such that Δ is maximized. Thus, the optimal coordination-inducing contract is where (EF) is binding at $\max_s \Delta$.

Given standard normal distribution of the measurement noise,

$$\Delta = \begin{cases} \Phi\left(\frac{a_i^0 - a_i^{'}}{\sqrt{1-\sigma^2}} - s\right) - \Phi(-s) & for\, a_i^0 > a_i^{'} \\ -\Phi\left(\frac{a_i^0 - a_i^{'}}{\sqrt{1-\sigma^2}} + s\right) + \Phi(s) & for\, a_i^0 < a_i^{'} \end{cases},$$

where the agents collude/coordinate to exert (a_i^0, a_j^0) when the non-cooperative Nash equilibrium is $(a_i^{'}, a_j^{'})$. A contract that deters collusion implements $a_i^{'} > a_i^0$ while a contract that induce coordination implements $a_i^0 > a_i^{'}$. In the former case, $\frac{\partial \Delta}{\partial s} \gtreqless 0$ if $\phi(s) \gtreqless \phi\left(s - \frac{a_H - a_L}{\sqrt{1-\sigma^2}}\right)$, which holds if $s \lesseqgtr \frac{a_H - a_L}{2\sqrt{1-\sigma^2}}$. In the latter case, $\frac{\partial \Delta}{\partial s} \gtreqless 0$ if $\phi(-s) \gtreqless \phi\left(\frac{a_H - a_L}{\sqrt{1-\sigma^2}} - s\right)$, which holds if $s \lesseqgtr \frac{a_H - a_L}{2\sqrt{1-\sigma^2}}$. The optimal threshold of believed unilateral deviation satisfies $s^*\sqrt{1-\sigma^2} = \frac{a_H - a_L}{2}$, at which the maximum accuracy of mutual inference is $\Delta^* = \Phi\left(\frac{a_H - a_L}{2\sqrt{1-\sigma^2}}\right) - \Phi\left(-\frac{a_H - a_L}{2\sqrt{1-\sigma^2}}\right)$. It is then straightforward that the optimal threshold of believed unilateral deviation s^* and the maximum accuracy of mutual inference Δ^* are both increasing in the correlation of measurement noise.

□

A.5 Proof of Proposition 3

The enforcement constraint is rearranged as $\Delta^* \geq \frac{1-\delta}{\delta}\left(\frac{\eta - \alpha}{\alpha + \beta - \eta}\right)$ when collusion is deterred or when coordination is induced to implement (a_H, a_H) in equilibrium. Rearranging (CP) and (CI) given Δ^* results in the same constraint. The optimal team incentive is then straightforward from Lemma 2 and the proof of Proposition 1 in A.2, with $\widehat{\beta}_{CP} = \frac{(h(\delta,\Delta^*)-\sigma)\eta}{1+h(\delta,\Delta^*)^2 - 2\sigma h(\delta,\Delta^*)} < 0$ when $\delta < \widehat{\delta}_{CP} \equiv \frac{\sigma}{(1-\sigma)\Delta^* + \sigma}$ and $\widehat{\beta}_{CI} = \frac{r(h(\delta,\Delta^*)-\sigma)\eta - (1-h(\delta,\Delta^*))(a_H - a_L)}{r(1+h(\delta,\Delta^*)^2 - 2\sigma h(\delta,\Delta^*))}$ when $\delta > \widehat{\delta}_{CI} \equiv \frac{r\sigma\eta + (a_H - a_L)}{r((1-\sigma)\Delta^* + \sigma)\eta + (a_H - a_L)}$, IPR is optimal if otherwise. As $\Delta^* = \max_s \Delta \geq \Delta_s$ for any σ by definition, $h(\delta, \Delta^*) > h(\delta, \Delta_s)$, so $\widehat{\beta}_{CP} > \beta_{CP}$ and $\widehat{\beta}_{CI} > \beta_{CI}$. In addition, $\widehat{\delta}_{CI} \leq \delta_{CI}$ and $\widehat{\delta}_{CP} \leq \delta_{CP}$ as $\Delta^* \geq \Delta_s$.

□

A.6 Proof of Proposition 4

By the proof of Proposition 2 in A.3, $\widehat{\delta}_{CP}$ and $\widehat{\delta}_{CI}$ are increasing (decreasing) in σ if $\frac{\partial \Delta^*}{\partial \sigma}\frac{\sigma}{\Delta^*} < (>) \frac{1}{1-\sigma}$ and $\frac{\partial \Delta^*}{\partial \sigma}\frac{\sigma}{\Delta^*} < (>) \frac{1}{\varphi(\sigma)}\frac{\sigma}{1-\sigma}$. Since $\frac{\sigma}{\varphi(\sigma)} < 1$, for all σsuch that $\widehat{\delta}_{CI}$ is increasing in σ, $\widehat{\delta}_{CP}$ is also increasing in σ. As $\lim_{\sigma \to 1} \Delta^* = 1$ and $\frac{1}{\varphi(\sigma)} \geq 1$ with equality at $\sigma = 1$, $\frac{\partial \Delta^*}{\partial \sigma}\frac{1-\sigma}{\Delta^*} < \frac{1}{\varphi(\sigma)}$ and $\widehat{\delta}_{CI}$ is increasing for sufficiently high correlations. As $\lim_{\sigma \to 0}\left(\frac{\partial \Delta^*}{\partial \sigma}\frac{1-\sigma}{\Delta^*}\right) = 0 < \frac{1}{\varphi(0)}$, $\widehat{\delta}_{CI}$ is increasing for sufficiently low correlations as well.

For intermediate level of correlations, due to the special form of normal CDF, I proceed by fixing the exogenous threshold s and σ at the levels such that δ_{CP} is decreasing in σ and compare $\Delta^*_\sigma \equiv \frac{\partial \Delta^*}{\partial \sigma} = \phi\left(\frac{a_H - a_L}{2\sqrt{1-\sigma^2}}\right)\frac{\sigma(a_H - a_L)}{2\sqrt{(1-\sigma^2)^3}}$ with $\Delta^s_\sigma \equiv \frac{\partial \Delta_s}{\partial \sigma} = \phi\left(\frac{a_H - a_L}{\sqrt{1-\sigma^2}} - s\right)\frac{\sigma(a_H - a_L)}{\sqrt{(1-\sigma^2)^3}}$. The ratio $\frac{\Delta^s_\sigma}{\Delta^*_\sigma} \gtreqless 1$ if $\frac{\phi(2x-s)}{\phi(x)} \gtreqless \frac{1}{2}$ where $x \equiv \frac{a_H - a_L}{2\sqrt{1-\sigma^2}}$. Define y such that $\phi(y) \equiv \frac{1}{2} \cdot \phi\left(\frac{a_H - a_L}{2\sqrt{1-\sigma^2}}\right)$. For $(a_H - a_L)$ such that $2x - s \in (-y, y)$, $\frac{\Delta^s_\sigma}{\Delta^*_\sigma} > 1$, so that whenever δ_{CI} and δ_{CP} are increasing, $\widehat{\delta}_{CI}$ and $\widehat{\delta}_{CP}$ are increasing as well, and when δ_{CI} and δ_{CP} are decreasing, it is not necessary that $\widehat{\delta}_{CI}$ and $\widehat{\delta}_{CP}$ are decreasing as well. This condition rules out extreme cases in which $a_H - a_L$ is so large or so small that how inference is made becomes trivial.

$\square$

References

[1] Abreu, Dilip, David Pearce, and Ennio Stacchetti (1986), "Optimal Cartel Equilibrium with Imperfect Monitoring," *Journal of Economic Theory*, 39, pp. 251-269

[2] Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58(5), pp. 1041-1063

[3] Aoyagi, Masaki and Guillaume Fréchette (2009), "Collusion as Public Monitoring Becomes Noisy: Experimental Evidence," *Journal of Economic Theory*, 144, pp. 1135-1165

[4] Cai, Hongbin and Ichiro Obara (2009), "Firm Reputation and Horizontal Integration," *RAND Journal of Economics*, 40(2), 340–363

[5] Che, Yeon-Koo and Seung-Weon Yoo (2001), "Optimal Incentive for Teams," *American Economic Review*, 91(3), pp.525-541

[6] Deb, Joyee, Jin Li, and Arijit Mukherjee (2016), "Relational Contracts with Subjective Peer Evaluations," *RAND Journal of Economics*, 47(1), pp. 3-28

[7] Fleckinger, Pierre (2012), "Correlation and Relative Performance Evaluation," *Journal of Economic Theory*, 147(1), pp.93-117

[8] Fleckinger, Pierre and Nicolas Roux (2012), "Collective versus Relative Incentives: the Agency Perspective," *Working Paper*

[9] Fudenberg, Drew, David Levine, and Eric Maskin (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62(5), pp.997-1039

[10] Fudenberg, Drew, Bengt Holmström, and Paul Milgrom (1990), "Short-Term Contracts and Long-Term Agency Relationships," *Journal of Economic Theory*, 51(1), pp. 1-31

[11] Green, Edward, and Robert Porter (1984), "Non-cooperative Collusion Under Imperfect Price Information," *Econometrica*, 52(1), pp. 87-100

[12] Holmström, Bengt (1982), "Moral Hazard in Teams," *The Bell Journal of Economics*, 13(2), pp. 324-340

[13] Holmström, Bengt and Paul Milgrom (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55(2), pp. 303-328

[14] Holmström, Bengt and Paul Milgrom (1990), "Regulating Trade among Agents," *Journal of Institutional and Theoretical Economics*, 146, pp. 85-105

[15] Ishiguro, Shingo (2002), "Optimal Incentives for Teams: Comment," *American Economic Review*, 92, pp. 1711

[16] Ishiguro, Shingo (2004), "Collusion and Discrimination in Organizations," *Journal of Economic Theory*, 116, pp. 357-369

[17] Itoh, Hideshi (1992), "Cooperation in Hierarchical Organizations: An Incentive Perspective," *Journal of Law, Economics, & Organization*, 8(2), pp. 321-345

[18] Itoh, Hideshi (1993), "Coalitions, Incentives, and Risk Sharing," *Journal of Economic Theory*, 60, pp. 410-427

[19] Kim, Jungwook (2011), "Collusion and Discrimination in Organizations: Comment," *Journal of Economic Research*, 16, pp. 309-317

[20] Kim, Jin-Hyuk and Nick Vikander (2015), "Team-Based Incentives in Problem-Solving Organizations," *Journal of Law, Economics, & Organization*, 31(2), pp. 262-286

[21] Meyer, Margaret and John Vickers (1997), "Performance Comparisons and Dynamic Incentives," *Journal of Political Economy*, 105(3), pp.547-581

[22] Mookherjee, Dilip (1984), "Optimal Incentive Schemes with Many Agents," *Review of Economic Studies*, 51(3), pp.433-446

[23] Porter, Robert H., (1983), "Optimal Cartel Trigger Price Strategies," *Journal of Economic Theory*, 29(2), pp. 313-338

[24] Ramakrishnan, Ram and Anjan Thakor (1991), "Cooperation versus Competition in Agency," *Journal of Law, Economics, and Organization*, 7(2), pp.248-283

[25] Rayo, Luis (2007), "Relational Incentives and Moral Hazard in Teams," *Review of Economic Studies*, 74, pp. 937-963