Optimal Multi-Agent Performance Measures for Team Contracts

Hyeng Keun Koo   Gyoocheol Shim   Jaeyoung Sung*

April 4, 2007

Abstract

We present a continuous-time contracting model under moral hazard with many agents. The principal contracts many agents as a team, and they jointly produce correlated outcomes. We show the optimal contract for each agent is linear in outcomes of all other agents as well as his/her own. The structure of the optimal contract strikingly reveals that the optimal aggregate performance measure in general can be orthogonally decomposed into two statistics: one is a sufficient statistic, and the other a non-sufficient statistic. As a consequence, the optimal aggregate performance measure in general is not a sufficient statistic, unless the principal is risk neutral. We further discuss agents' optimal effort choices using a "quadratic-cost" example, which also strikingly suggests that team contracts sometimes provide lower-powered effort incentives than individually separate contracts do.

JEL Classification: C61, C73, D86, G30, L14, L25
Key Words: principal-agent problem, many agents, moral hazard, team, performance measure, contracts, continuous-time model, martingale method.

*Koo: School of Business Administration, Ajou University, Suwon, 443-749, Korea; hkoo@ajou.ac.kr. Shim: The Bank of Korea, Seoul, 100794, Korea; gshim@bok.or.kr. Sung: Department of Finance (M/C 168), University of Illinois at Chicago, 601 South Morgan Street, Chicago, Illinois 60607-7124; jsung@uic.edu.

1 Introduction

We present a continuous-time principal-agent model under moral hazard with many agents. Since Holmstrom and Milgrom [1987], continuous-time approaches have been powerfully applied to various principal-agent problems, particularly because continuous-time approaches offer tractable ways to resolve well-known serious technical difficulties existing in the discrete-time literature, and produce relatively simple forms of optimal contracts.[1] However, thus far, no continuous-time multi-agency models have been rigorously attempted, even though there has already been a significant amount of the discrete-time multi-agency literature.[2]

Our model is a continuous-time counterpart of Holmstrom's [1982] discrete-time model, and can be viewed as an extension of Holmstrom and Milgrom [1987] with a team of many agents. We use Schättler and Sung's [1993] martingale method to model both the principal and agents' problems in which all agents jointly choose probability distributions of given outcome processes. All agents exhibiting constant absolute risk aversion (CARA) are assumed to be Nash game players: namely, each agent chooses his/her effort level as if all other agents' effort choices have already been made. We show that as is the case with Holmstrom and Milgrom's standard principal-agent model, optimal team contracts for many agents in this paper are also linear in all outcomes. For the linearity result, the Nash-game assumption appears to be important, because under the assumption, our model preserves Holmstrom and Milgrom's stationary decision-making environment for the principal, which is critical to produce linear contracts as optimal contracts.

We utilize the linearity result to examine multi-agency/team effects on the structure of optimal performance measures and on agents' effort decisions. First, we reconfirm the well-known Holmstrom-Mookherjee's sufficient statistics results when the principal is risk neutral.[3] However, we argue that the same results may not hold, when the principal is risk averse.

In particular, optimal team contracts in this paper strikingly reveal that the optimal performance measure for each agent consists of two orthogonal metrics: one is a sufficient statistic for effort incentive and the other a non-sufficient statistic for risk sharing. It is noteworthy that the latter is orthogonal to the sufficient statistic, containing zero information content about the agent's marginal effort level. As a special case, if the principal is risk neutral, the performance measure for each agent turns out to be a sufficient statistic, because the risk-neutral principal is not concerned with risk-sharing. However, if

[1] See Schättler and Sung [1993] for a detailed discussion of the resolution of technical difficulties that have long existed in the discrete-time literature. More discussion on methodologies for continuous-time principal agent problems can be found in Sung [1995, 1997], Cadenillas, Cvitanic and Zapatero [2003], Williams [2003], and Cvitanic, Wan and Zhang [2005].

[2] See Holmstrom [1982] and Mookherjee [1984] for discrete-time multi-agency models under moral hazard. The relative performance evaluation literature is also related to the multi-agency models. See Lazear and Rosen [1981], Nalebuff and Stiglitz [1983], Gibbons and Murphy [1990], Dye [1992], Core and Guay [2003], and Ou-Yang [2005].

[3] Holmstrom [1982] and Mookherjee [1984] derive the celebrated "sufficient statistics result", which states that the optimal salary function of each agent in a team under moral hazard should depend only on sufficient statistics for his/her effort, not on the effort outcomes of other agents. However, as Holmstrom noted, the result may depend on the assumption that the principal is risk neutral.

the principal is risk averse, the performance measure for each agent becomes a combination of sufficient and non-sufficient statistics, and thus it fails to be a sufficient statistic.

In order to investigate team effects on contracting, we compare our team contracts with individually separate contracts. In this paper, each of the team contracts is allowed to depend on all agents' outcomes in the team, whereas each of individually separate contracts is allowed to depend only on the individual' own outcome, but not on other agents'. Team contracting in general improves both incentives and risk-sharing. Our quadratic-cost example suggests that risk-sharing can be improved because team contracting enables the principal to distribute among agents the risks that are unrelated to effort incentives.[4] Incentives can also be improved because correlations among individual outcomes can be utilized to reduce the volatility of the effort performance measure for each agent.

However, it is also striking that team contracts do not always provide higher effort incentives than individually separate contracts do. This result is somewhat counterintuitive, because it is tempting to conclude the contrary based on a reasoning as follows: team contracts typically can be used to reduce volatilities of agents' performance risks, and thus effort incentives under team contracts should be always higher than or equal to those under individually separate contracts. We argue that this reasoning is not always true.

To see the possibility of lower effort incentives under team contracts than those under individually separate contracts, consider a case where individual-effort outcomes of all agents with CARA preferences in a team are independent of each other and there are no joint-effort outcomes. Note that, under individually separate contracts, risk-averse principal's risk-sharing with each agent can only be arranged by increasing the contract sensitivity to each agent's own outcome. However, under team contracts, the performance measure for each agent can consist of his own outcome for effort incentives and other agents' outcomes for risk-sharing. Since the other agents' outcomes are independent of each agent's effort, his effort incentive is unaffected by the principal's risk-sharing motivation. As a result, individually separate contracts sometimes stipulate higher-powered incentives than team contracts do.

This paper is organized as follows. In the next section, we describe the general model of the paper. General forms of optimal contracts are derived in Section 3, and then special cases are discussed in Section 4. Finally, in Section 5 we provide a brief summary of results of the paper.

2 A Multi-Agency Problem in Continuous Time

We investigate optimal contracts in a continuous-time principal-agent problem where there are one principal (investors) and multiple agents (managers) under moral hazard. The time horizon of interest is the unit interval $[0,1]$. The principal is endowed with a set of N production tasks. At time zero,

[4] The distribution of unrelated risks across agent in a team can be eventually related to a diversification effect, which is studied in detail in our companion paper, Koo, Shim and Sung [2006].

the principal and a team of N agents sign on N contracts $\{S^a, a \in A\}$ for the N tasks: agent a, $a \in A := \{1, ..., N\}$, is in charge of task a during the time period $[0,1]$, and will be compensated at time 1 according to the compensation scheme specified by contract S^a. The principal and agents' preferences are characterized by exponential utility functions, and the principal's and agent a's coefficients of absolute risk aversion are, respectively, $R \geq 0$ and $R^a > 0$ for $a \in A$.

All agents carry out their tasks under uncertainty. There are $M(> N)$ sources of uncertainty represented by an M-vector of standard independent Wiener processes $W_t = (W_t^1, \cdots, W_t^M)^\top$ on the probability space $(\Omega, \mathcal{F}_t, P)$, where $^\top$ denotes the transpose of a matrix or a vector, and $\mathcal{F}_t$ is the augmentation of the filtration generated by W_t. Outcomes of agents' effort over time are described by an N-vector of outcome processes $\{Y_t\}$ which evolves as follows.

$$dY_t = \sigma dW_t,$$

where $Y_t = (Y_t^1, \cdots, Y_t^N)^\top$ is an N-dimensional outcome vector at the time t, and $\sigma(\in \mathcal{R}^{N \times M})$ is a bounded $N \times M$ volatility matrix. We let σ^a, $a \in A$, denote the a-th row of matrix σ, and Q an $N \times N$ matrix of instantaneous variances and covariances of Y_t, i.e., $Q = \sigma\sigma^\top$.

Carrying their tasks, agent a, $a \in A$, exerts effort $\mu_t^a \in U$, $t \in [0,1]$, incurring personal (monetary) cost at a rate of $c^a(\mu_t^a)$, where U is a bounded open interval. We assume no discounting and thus agent a's total cost of effort during the contract period is $\int_0^1 c^a(\mu_t^a)dt$. We also assume all cost functions are strictly increasing and convex in effort, i.e., $c_\mu^a > 0$, and $c_{\mu\mu}^a > 0$ for $a \in A$, where subscripts μ and $\mu\mu$ denote the first and second order partial derivatives with respect to μ_t^a, respectively. Let $\mu_t = (\mu_t^1, ..., \mu_t^N)^\top$, and $u = \{\mu_t\}$. Then μ_t is an N-vector of all agents' effort at time t, and u is an N-vector process of all agents' effort.

By exerting effort u, all agents jointly change the probability measure of $\{W_t\}$ and thus that of $\{Y_t\}$ from P to P_u, where

$$\frac{dP_u}{dP} = e^{\int_0^1 (\phi(\mu_t))^\top dW_t - \frac{1}{2}\int_0^1 \|\phi(\mu_t)\|^2 dt},$$

and $\phi : U^N \to \mathcal{R}^M$ is a bounded function such that $\sigma\phi(\mu_t)$ is increasing and concave in μ_t. We assume that P_u cannot be observed/verified. This assumption leads to a nontrivial agency problem involving conflicts of interest between the principal and agents over effort choices.

Under P_u,

$$B_t^u = W_t - \int_0^t \phi(\mu_t)ds$$

is an M-vector of standard Wiener processes, and

$$dY_t = \sigma\left(dB_t^u + \phi(\mu_t)dt\right).$$

Let

$$f(\mu_t) \equiv (f^1(\mu_t), ..., f^N(\mu_t))^\top := \sigma\phi(\mu_t).$$

Then dynamics of Y can be equivalently written as

$$dY_t = f(\mu_t)dt + \sigma dB_t^u.$$

These equivalent dynamics allow us to alternatively interpret agents' effort as follows: each agent's effort affects the drift of the outcome process $\{Y_t\}$. However, with this alternative interpretation, it is important to assume that the principal cannot observe and verify B_t^μ separately from the outcome process $\{Y_t\}$, because otherwise our agency problems can be trivially resolved.

In particular, we make the following informational assumptions. Both the principal and agents observe the whole outcome processes $\{Y_t\}$, and thus the diffusion rate matrix σ is common knowledge.[5] Each agent can observe his/her own effort levels, but cannot directly observe the other agents' effort levels. Moreover, the principal can directly observe/verify none of agents' effort levels.[6]

Given the above informational constraint and given a contract S^a, agent $a(\in A)$ chooses an effort-level process $\{\mu_t^a\}$ to maximize his/her expected utility of compensation net of the cost of effort. We assume that agent a's reservation utility at time zero is $-e^{-r\mathcal{W}_0^a}$, $a \in A$, and that each agent chooses his/her effort levels over time as if all other agents effort decisions had already been made and his/her effort decisions would not affect other agents' current and future effort decisions. That is, all agents behave like Nash game players. On the other hand, taking into account all agents' reservation utility levels and behaviors in their effort choices, the principal would like to design a set of team contracts $\{S^a, a \in A\}$ to maximize her expected utility of the total final output after compensations to all agents. We also assume that admissible contracts for agent a, $a \in A$, have the following structure:

$$S^a(Y) = S_1^a(Y) + \int_0^1 \alpha^a(t,Y)dt + \int_0^1 (\beta^a)^\top(t,Y)dY_t,$$

where S_1 is a bounded $\mathcal{F}_1$-measurable random variable; and α^a and β^a are bounded $\mathcal{F}_t$-predictable processes. Note that we allow all individual components of S^a to be non-Markovian, i.e., $S_1^a(Y)$, $\alpha^a(t,Y)$ and $(\beta^a)^\top(t,Y)$ can be non-Markovian.

3 Optimal Contracts

Let E_u denote the expectation operator under probability measure P_u, and $E_{(\hat{u}^a;u^{-a})}$ under $P_{(\hat{u}^a;u^{-a})}$, where $(\hat{u}^a;u^{-a}) \equiv \{(\{\mu_t^1\},...,\{\mu_t^{a-1}\},\{\hat{\mu}_t^a\},\{\mu^{a+1}\},...,\{\mu^N\})\}$. Then the principal's problem is now stated as:

[5] See Sung (1995) for a case of one principal and one agent where the principal can only observe the final outcome Y_1 and the agent is allowed to control the diffusion rate.

[6] If the agents' effort levels are observable to the principal, then the first best is obtainable and it can be shown that the optimal salary functions of agents are linear functions of total output by extending the argument for the case where there is a single agent as in Sung (1991) and Müller (1998)

Problem 3.1 *Choose u (or probability measure P_u) and compensation schemes $\{S^a, a \in A\}$ to*

$$\max \quad E_u\left[-\exp\left\{-R\left\{Y_1^T - \sum_{a\in A} S^a(Y)\right\}\right\}\right]$$

$$s.t. \quad (1) \quad dY_t = \sigma dW_t,$$

(2) $\forall a \in A$, *given $\{S^a, a \in A\}$ and u^{-a}, $u^a (= \{\mu_t^a\})$ is a solution to the following problem:*

$$\max_{\hat{u}^a} E_{(\hat{u}^a;u^{-a})}\left[-\exp\left\{-R^a\left\{S^a(Y) - \int_0^1 c^a(\hat{\mu}_t^a)dt\right\}\right\}\right],$$

$$(3) \quad \forall a \in A, \quad E_u\left[-\exp\left\{-R^a\left\{S^a(Y) - \int_0^1 c^a(\mu_t^a)dt\right\}\right\}\right] \geq -\exp\left\{-R^a\mathcal{W}_0^a\right\}.$$

Note that constraints (2) and (3) are agents' incentive compatibility and participation conditions. In constraint (2), agent a chooses u^a, as if all other agents' choices u^{-a} were given. That is, we try to investigate Nash equilibrium solutions.

3.1 Agents' Problems

Let us start with agents' incentive compatibility conditions. Given u^{-a} and S^a with (α^a, β^a), agent a's problem is as follows:

$$\begin{aligned}
\max_{\hat{u}^a} \quad & E_{(\hat{u}^a;u^{-a})}\left[-\exp\left\{-R^a\left(S^a(Y) - \int_0^1 c^a(\hat{\mu}_t^a)dt\right)\right\}\right] \\
&= E_{(\hat{u}^a;u^{-a})}\left[-\exp\left\{-R^a\left(S_1^a(Y) + \int_0^1 \left(\alpha^a + (\beta^a)^\top f(\hat{\mu}_t^a;\mu_t^{-a}) - c^a(\hat{\mu}_t^a)\right)dt\right.\right.\right. \\
&\qquad \left.\left.\left. + \int_0^1 (\beta^a)^\top \sigma dB_t^{(\hat{u}^a;u^{-a})}\right)\right\}\right].
\end{aligned}$$

Recall that in the above stochastic control problem, α^a, β^a and $S_1^a(Y)$ may not be Markovian. We examine this potentially non-Markovian stochastic control problem using the same martingale method as in Schättler and Sung [1993].

Lemma 3.1 *(Schättler and Sung [1993, Theorem 3.1]) Given u^{-a} and S^a with (α^a, β^a), let μ_t^{a*}, $t \in [0,1]$, be optimal for agent a's problem. Then, there exist $\mathcal{F}_t$-adapted processes $\mathcal{V}_t$ and $\nabla\mathcal{V}_t^a \equiv (\nabla\mathcal{V}_t^{a1}, ..., \nabla\mathcal{V}_t^{aN})^\top$ such that optimal μ_t^a maximizes*

$$\begin{aligned}
& H_t^a(\hat{\mu}_t^a; \mu^{-a}, \sigma, \alpha^a, \beta^a, \mathcal{V}_t^a, \nabla\mathcal{V}_t^a) \\
&:= (\nabla\mathcal{V}_t^a)^\top \left\{f(\hat{\mu}^a;\mu^{-a}) - R^a\sigma\sigma^\top\beta^a\right\} \\
&\quad + \mathcal{V}_t^a\left\{R^a\left\{c^a(\mu_t^a) - \alpha^a - (\beta^a)^\top f(\hat{\mu}^a;\mu^{-a})\right\} + \frac{1}{2}(R^a)^2\|\sigma^\top\beta^a\|^2\right\}. \qquad (1)
\end{aligned}$$

Furthermore, $\mathcal{V}_t$ is the value process, i.e.,

$$\begin{aligned}
\mathcal{V}_t^a \;=\; & \max_{\hat{\mu}_s^a, s\in[t,1]} E_{(\hat{u}^a;u^{-a})}\left[-\exp\left\{-R^a\left(S_1^a(Y)\right.\right.\right. \\
& \left.\left.\left. + \int_t^1 \left(\alpha^a + (\beta^a)^\top f(\hat{\mu}_t^a;\mu_t^{-a}) - c^a(\hat{\mu}_s^a)\right)dt + \int_t^1 (\beta^a)^\top\sigma dB_s^{(\hat{u},u^{-a})}\right)\right\}\middle|\, \mathcal{F}_t\right],
\end{aligned}$$

with the following dynamics:

$$\mathcal{V}_t^a = \mathcal{V}_0^a - \int_0^t H_s^a ds + \int_0^t (\nabla \mathcal{V}_s^a)^\top dY_s, \tag{2}$$

where $H_t^a := \max_{\hat{\mu}_t^a} H_t^a(\hat{\mu}_t^a; \mu^{-a}, \sigma, \alpha^a, \beta^a, \mathcal{V}_t^a, \nabla \mathcal{V}_t^a)$.

We utilize Lemma 3.1 to examine necessary conditions for optimal contracts. First, let us define

$$v_t^a \equiv (v_t^{a1}, ..., v_t^{aN})^\top = \beta^a - \frac{1}{R^a \mathcal{V}^a} \nabla \mathcal{V}_t^a,$$

$$v_t = (v_t^1, ..., v_t^N)^\top,$$

and

$$f_{\mu_t^a}(\mu_t) = \left(f_{\mu_t^a}^1(\mu_t), ..., f_{\mu_t^a}^N(\mu_t) \right)^\top,$$

where $f_{\mu_t^a}^b$ denotes the partial derivative of f^b with respect to μ_t^a, i.e., $f_{\mu_t^a}$ is an N-vector of marginal products of agent a's effort. As will be seen in the following proposition, v_t^a turns out to be an N-vector of sensitivities of the optimal contract for agent a to outcome vector Y. For brevity we let $\nu := (\nu^1, ..., \nu^N)$ and $\nu^a := \{v_t^a\}$.

Proposition 3.1 *Given an admissible contract* $S^a(Y)$ *and* $\{\mu_t^{-a}\}$*, let, for all* t*,*

$$\mu_t^a \in \arg\max_{\hat{\mu}} H_t^a(\hat{\mu}; \mu_t^{-a}, \sigma, \alpha^a, \beta^a, \mathcal{V}_t^a, \nabla \mathcal{V}_t^a).$$

Then agent a*'s salary function* $S^a(Y)$ *can also be expressed in the following form:*

$$S^a(Y; u^a, \nu^a; u^{-a}) = \mathcal{W}_0^a + \int_0^1 \left(c^a(\mu_t^a) - (v_t^a)^\top f(\mu_t) + \frac{R^a}{2} \left\| (v_t^a)^\top \sigma \right\|^2 \right) dt + \int_0^1 (v_t^a)^\top dY_t, \tag{3}$$

where μ_t^a *satisfies the following equation.*

$$(v_t^a)^\top f_{\mu_t^a}(\mu_t) - c_{\mu^a}^a(\mu_t^a) = 0. \tag{4}$$

Proof: Note that condition (4) immediately follows from agent a's first order condition (FOC) to maximize agent a's Hamiltonian (1) in Lemma 3.1. On the other hand, define the certainty equivalent wealth process $\mathcal{W}_t^a$ as follows:

$$\mathcal{W}_t^a = -\frac{1}{R^a} \log(-\mathcal{V}^a(t, Y_t)).$$

Then, by Itô's rule,

$$\mathcal{W}_t^a = \mathcal{W}_0^a - \int_0^t \frac{1}{R^a \mathcal{V}^a} d\mathcal{V}_s^a + \frac{1}{2} \int_0^t \frac{1}{R^a (\mathcal{V}^a)^2} d\langle \mathcal{V}^a, \mathcal{V}^a \rangle_s.$$

By substituting (2) into the above, we have

$$\begin{aligned}
\mathcal{W}_t^a - \mathcal{W}_0^a &= \int_0^t \frac{1}{R^a\mathcal{V}^a}\left(H_t^a dt - \nabla\mathcal{V}_t^a dY_t\right) + \frac{1}{2}\int_0^t \frac{1}{R^a(\mathcal{V}^a)^2}\|\nabla\mathcal{V}_t^a\sigma\|^2 dt \\
&= \int_0^t \left(\frac{1}{R^a\mathcal{V}^a}\nabla\mathcal{V}_t^a\left\{f(\mu_t) - R^a\sigma\sigma^\top\beta^a\right\} + c^a(\mu_t^a) - \alpha^a - (\beta^a)^\top f(\mu_t) + \frac{1}{2}R^a\|\sigma^\top\beta^a\|^2\right. \\
&\qquad \left. + \frac{1}{2}\frac{1}{R^a(\mathcal{V}^a)^2}\|\nabla\mathcal{V}_t^a\sigma\|^2\right) dt - \int_0^t \frac{1}{R^a\mathcal{V}^a}\nabla\mathcal{V}_t^a dY_t \\
&= \int_0^t \left(c^a(\mu_t^a) - \alpha^a + \left\{\frac{1}{R^a\mathcal{V}^a}\nabla\mathcal{V}_t^a - (\beta^a)^\top\right\} f(\mu_t) - \frac{1}{R^a\mathcal{V}^a}\nabla\mathcal{V}_t^a R^a\sigma\sigma^\top\beta^a + \frac{1}{2}R^a\|\sigma^\top\beta^a\|^2\right. \\
&\qquad \left. + \frac{1}{2}\frac{1}{R^a(\mathcal{V}^a)^2}\|\nabla\mathcal{V}_t^a\sigma\|^2\right) dt - \int_0^t (\beta^a)^\top dY_t - \int_0^t \left\{\frac{1}{R^a\mathcal{V}^a}\nabla\mathcal{V}_t^a - (\beta^a)^\top\right\} dY_t.
\end{aligned}$$

Since $\mathcal{W}_1^a$ is S_1^a by construction, the above equation implies (3). □

Remark 1: The salary function in (3) together with (4) is a multi-agent generalization of Schättler and Sung [1993, Theorem 4.1].

Remark 2: If $N = M = 1$ and $f(\mu_t) = \mu_t$, then Eq. (4) implies that $v_t^{aa} = c_\mu(\mu_t)$, $v_t^{-a} = 0_{N-1}$, and the representation of the salary function S^a in (3) reduces to that of Holmstrom and Milgrom [1987].

Since (3) subject to (4) is a necessary condition for optimal contracts, we henceforth only focus on this class of contracts. Eq.(4) simply describes agent a's behavior over his/her effort choice: that is, the agent chooses the effort in such a way that the marginal cost of effort is equal to the marginal expected compensation.

However, an issue can arise, because the class of salary functions described by (3) subject to (4) can still be too large, and not all contracts in the class can be implementable. That is, given an arbitrary contract $S^a(Y;\bar{\mu}^a,\bar{\nu}^a;\bar{\mu}^{-a})$ of the form (3), agent a may not optimally choose effort controls $\bar{\mu}^a$. In this case, $S^a(Y;\bar{\mu}^a,\bar{\nu}^a;\bar{\mu}^{-a})$ may not be implementable, and the principal can safely ignore such non-implementable contracts, because she can always find another contract of the same form (3) with implementable controls without sacrificing her expected utility. See Sung [1997, Proposition 1]. The following proposition helps eliminate such non-implementable contracts from class (3), and we only consider the class (3) of contracts with implementable controls.

Proposition 3.2 *If, for all $\beta \in \mathcal{R}^N$ and $\mu^{-a} \in \mathcal{R}^{N-1}$, $\beta f(\mu) - c^a(\mu^a)$ is concave in μ^a, then salary function $S^a(Y;u^a,\nu^a;u^{-a})$ that is given in the form of (3) satisfying (4) is implementable, i.e., given the salary function, agent a optimally chooses $u^a(=\{\mu_t^a\})$.*

Remark: The concavity assumption is always satisfied if $f(\mu)$ is linear in μ, and c is convex.

Proof: Let S^a be a salary function given to agent a with parameters $\{\mu_t^a\}$, and $\{v_t^a\}$ in the form of (3) satisfying (4). Then, agent a's problem is to choose $\{\hat{\mu}_t^a\}$ to maximize

$$
\begin{aligned}
&E_{(\hat{u}^a;u^{-a})}\left[-\exp\left\{-R^a\left(S^a-\int_0^1 c^a(\hat{\mu}_t^a)dt\right)\right\}\right]\\
&=E_{(\hat{u}^a;u^{-a})}\left[-\exp\left\{-R^a\left(\mathcal{W}_0^a+\int_0^1\left(c^a(\mu_t^a)-(v_t^a)^\top f(\mu_t)+\frac{R^a}{2}\left\|(v_t^a)^\top\sigma\right\|^2-c^a(\hat{\mu}_t^a)\right)dt\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.+\int_0^1(v_t^a)^\top dY_t\right)\right\}\right]\\
&=E_{(\hat{u}^a;u^{-a})}\left[-\exp\left\{-R^a\left(\mathcal{W}_0^a+\int_0^1\left(c^a(\mu_t^a)-(v_t^a)^\top f(\mu_t)-c^a(\hat{\mu}_t^a)+(v_t^a)^\top f(\hat{\mu}_t^a;\mu_t^{-a})\right)dt\right)\right\}\right.\\
&\qquad\qquad\left.\times\exp\left\{-\frac{(R^a)^2}{2}\int_0^1\|(v_t^a)^\top\sigma\|^2dt-R^a\int_0^1(v_t^a)^\top\sigma dB_t^u\right\}\right].
\end{aligned}
$$

Recall that S^a with parameters u^a, u^{-a} and v^a in the form of (3) is constructed to satisfy $-c^a_{\mu^a}(\mu_t^a)+(v_t^a)^\top f_{\mu^a}(\mu)=0$. Thus, the concavity assumption implies $c^a(\mu_t^a)-(v_t^a)^\top f(\mu_t)-c^a(\hat{\mu}_t^a)+(v_t^a)^\top f(\hat{\mu}_t^a;\mu_t^{-a})\leq 0$ with equality at $\hat{\mu}_t^a=\mu_t^a$. That is, given S^a and u^{-a}, the agent's expected utility for all admissible $\hat{u}^a(=\{\hat{\mu}_t^a\})$ is less than or equal to $-e^{-R^a\mathcal{W}_0^a}$, and the equality, i.e., the maximum expected utility, is achieved at $\hat{\mu}_t^a=\mu_t^a$. Therefore, given the salary function S^a, the agent optimally chooses $\{\mu_t^a\}$, and S^a is implementable. □

3.2 The Principal's Problem

In fact, Eq.'s (3) and (4) in Proposition 3.1 presents necessary conditions for agents' incentive and participation constraints, and conditions for Proposition 3.2 are sufficient for Eq.'s (3) and (4) to satisfy agents' incentive compatibility conditions. Thus, we simplify and solve the principal's problem by substituting Eq.(3) for $S^a(Y)$ in her expected utility. Then the solution becomes optimal if it satisfies the implementability condition in Proposition 3.2.

Let us define constant N-vectors $\mu:=(\mu^1,...,\mu^N)$ and $v^a=(v^{a1},...,v^{aN})^\top$, for μ^a, $v^{ab}\in\mathcal{R}$, and a constant $N\times N$-matrix $v=(v^1,...,v^N)$. Also define the principal's Hamiltonian as follows.

$$
\begin{aligned}
&H^P(\mu,v;\sigma,R,\{R^a,a\in A\})\\
&=\sum_{a\in A}f^a(\mu)-\sum_{a\in A}\left(c^a(\mu^a)+\frac{R^a}{2}\left\|(v^a)^\top\sigma\right\|^2\right)-\frac{R}{2}\left\|\left(1^\top-\sum_{a\in A}(v^a)^\top\right)\sigma\right\|^2. \qquad (5)
\end{aligned}
$$

Theorem 3.1 *For all $a\in A$, $\mu^{-a}\in\mathcal{R}^{N-1}$, and $\beta\in\mathcal{R}^N$, assume $\beta f(\mu)-c^a(\mu^a)$ is concave in μ^a. Suppose that there exist $\mu^*=(\mu^{1*},...,\mu^{N*})^\top$ and $v^*=(v^{1*},...,v^{N*})$ maximizing*

$$
H^P(\mu,v;\sigma,R,\{R^a,a\in A\}),\quad s.t.\quad c^a_{\mu^a}(\mu^a)-(v^a)^\top f_{\mu^a}(\mu)=0,\quad\forall a\in A. \qquad (6)
$$

Then μ^* and v^* are the principal's optimal controls for (μ_t,v_t) for all t, the principal's expected utility is given by $-\exp\left\{-R\left(-\sum_{a\in A}\mathcal{W}_0^a+H^P(\mu^*,v^*;\sigma,R,\{R^a,a\in A\})\right)\right\}$, and optimal contracts for all *agents are linear in the final outcome vector* Y_1 *as follows: for all* $a \in A$,

$$S^a(Y_1) = \mathcal{W}_0^a + c^a(\mu^{a*}) - (v^{a*})^\top f(\mu^*) + \frac{R^a}{2}\left\|(v^{a*})^\top \sigma\right\|^2 + (v^{a*})^\top Y_1, \tag{7}$$

where the pair (μ^{a*}, v^{a*}) *satisfies* $(v^{a*})^\top f_{\mu^{a*}}(\mu^*) - c^a_{\mu^a}(\mu^{a*}) = 0$.

Proof: The principal's problem is to choose $u(=\{\mu_t\})$ and $v(=\{v_t\})$ to maximize the following quantity subject to $c^a_{\mu^a}(\mu^a_t) - (v^a_t)^\top f_{\mu^a_t}(\mu) = 0$:

$$\begin{aligned}
&E_u\left[-\exp\left\{-R\left(Y^T - \sum_{a\in A} S^a\right)\right\}\right] \\
&= E_u\Bigg[-\exp\Bigg\{-R\Bigg(-\sum_{a\in A}\mathcal{W}_0^a - \int_0^1 \sum_{a\in A}\left(c^a(\mu^a_t) - (v^a_t)^\top f(\mu) + \frac{R^a}{2}\left\|(v^a_t)^\top\sigma\right\|^2\right)dt \\
&\qquad + \int_0^1 (1^\top - \sum_{a\in A}(v^a_t)^\top)dY_t\Bigg)\Bigg\}\Bigg] \\
&= \tilde{E}_u\left[-\exp\left\{-R\left(-\sum_{a\in A}\mathcal{W}_0^a + \int_0^1 H^P(\mu_t, v_t; \sigma, R, \{R^a, a\in A\})dt\right)\right\}\right],
\end{aligned}$$

where expectation $\tilde{E}$ in the last equality is taken under $\tilde{P}^u$ which is defined such that

$$\frac{d\tilde{P}^u}{dP^u} = \exp\left\{-\frac{R^2}{2}\int_0^1 \left\|\left(1^\top - \sum_{a\in A}(v^a_t)^\top\right)\sigma\right\|^2 dt - R\int_0^1\left(1^\top - \sum_{a\in A}(v^a_t)^\top\right)\sigma dB^u_t\right\}.$$

However,

$$\begin{aligned}
&\tilde{E}_u\left[-\exp\left\{-R\left(-\sum_{a\in A}\mathcal{W}_0^a + \int_0^1 H^P(\mu_t, v_t; \sigma, R, \{R^a, a\in A\})dt\right)\right\}\right] \\
&\le -\exp\left\{-R\left(-\sum_{a\in A}\mathcal{W}_0^a + H^P(\mu^*, v^*; \sigma, R, \{R^a, a\in A\})\right)\right\},
\end{aligned}\tag{8}$$

because H^P is independent of time and state, (μ^*, v^*) maximizes $H^P(\mu, v; \sigma, R, \{R^a, a \in A\})$, and $E[d\tilde{P}^u/dP^u] = 1$. However, since the inequality (8) becomes equality when the constant pair (μ_t, v_t) is chosen to be (μ^*, v^*) for all t, the constant control pair is the principal's optimal control pair for (μ_t, v_t) for all t, i.e., $(\mu^*, v^*) = (\mu^*_t, v^*_t)$. Thus, the principal's optimal utility is as stated in the proposition. Furthermore, since the optimal control pair (μ^*_t, v^*_t) is constant over time, (3) implies that optimal contracts are linear in Y_1. □

Next, in order to examine conditions for the existence of (μ^*, v^*), we directly try to solve the principal's constrained maximization (6). For simplification, we first eliminate the constraints (agents' FOCs) by directly substituting them into the principal's objective function. Note that the constraint for agent a's FOC implies

$$v^{aa} = \tilde{c}^a(\mu) - (v^a_{-a})^\top \tilde{f}^a(\mu), \tag{9}$$

where $v^a_{-a} = (v^{a1}, ..., v^{a,a-1}, v^{a,a+1}, ..., v^{aN})^\top$,

$$\tilde{c}^a(\mu) := \frac{c^a_{\mu^a}(\mu^a)}{f^a_{\mu^a}(\mu)}, \quad \text{and} \quad \tilde{f}^a(\mu) := \left(\frac{f^1_{\mu^a}(\mu)}{f^a_{\mu^a}(\mu)}, \cdots, \frac{f^{a-1}_{\mu^a}(\mu)}{f^a_{\mu^a}(\mu)}, \frac{f^{a+1}_{\mu^a}(\mu)}{f^a_{\mu^a}(\mu)} \cdots, \frac{f^N_{\mu^a}(\mu)}{f^a_{\mu^a}(\mu)} \right)^\top.$$

The substitution of v^{aa} in (9) into the principal's objective function in (5) reduces the principal's choice variables to μ and $\{v^a_{-a}; a \in A\}$. With the substitution, the performance-based part of the salary function can be rewritten as

$$(v^a)^\top Y_1 = \tilde{c}^a(\mu) Y^a_1 + (v^a_{-a})^\top \left(Y^{-a}_1 - \tilde{f}^a(\mu) Y^a_1 \right).$$

Note that the $Y^{-a}_1 - \tilde{f}^a(\mu)Y^a_1$ is a vector of the portions of other agents' outcomes that are independent of marginal changes in μ^a.

For brevity, let C_{ab} be an $(N-1) \times (N-1)$ matrix of covariances between $Y^{-a} - \tilde{f}^a Y^a$ and $(Y^{-b} - \tilde{f}^b Y^b)$, i.e.,

$$C_{ab} := Cov\left((Y^{-a} - \tilde{f}^a Y^a), (Y^{-b} - \tilde{f}^b Y^b) \right),$$

and q_{ab} be an $(N-1)$-vector of covariances between Y^a and $Y^{-b} - \tilde{f}^b Y^b$, i.e.,[7]

$$q_{ab} := Cov\left(Y^a, (Y^{-b} - \tilde{f}^b Y^b) \right).$$

Also let $D(\mu)$ be an $(N-1)N \times (N-1)N$-matrix such that

$$D(\mu) = \begin{bmatrix} C_{11} & \frac{R}{R^1+R} C_{12} & \frac{R}{R^1+R} C_{13} & \cdots & \frac{R}{R^1+R} C_{1N} \\ \frac{R}{R^2+R} C_{21} & C_{22} & \frac{R}{R^2+R} C_{23} & \cdots & \frac{R}{R^2+R} C_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{R}{R^N+R} C_{N1} & \frac{R}{R^N+R} C_{N2} & \frac{R}{R^N+R} C_{N3} & \cdots & C_{NN} \end{bmatrix}.$$

Now, we are ready to state a condition for the existence of v^{a*}_{-a}, $a \in A$, the optimal contract sensitivities for agent a to other agents' effort outcomes.

Theorem 3.2 *Assume that the matrix $D(\mu)$ is invertible for all $\mu \in U^N$. Then $\{v^{a*}_{-a}, a \in A\}$ is the unique solution to the following system of linear equations: for all $a \in A$,*

$$C_{aa} v^a_{-a} + \frac{R}{R^a + R} \sum_{b \in A, b \neq a} C_{ab} v^b_{-b} = \frac{R}{R^a + R} \sum_{b \in A} (1 - \tilde{c}^b) q_{ba} - \frac{R^a \tilde{c}^a}{R^a + R} q_{aa}. \tag{10}$$

[7] Both C_{ab} and q_{ab} can be related to the variance-covariance matrix $Q = \sigma\sigma^\top$ as follows.

$$C_{ab} = Q_{-ab} - \tilde{f}^a p^\top_{ab} - p_{ba} (\tilde{f}^b)^\top + \sigma^a (\sigma^b)^\top \tilde{f}^a (\tilde{f}^b)^\top,$$

and

$$q_{ab} = p_{ab} - \sigma^a (\sigma^b)^\top \tilde{f}^b,$$

where Q_{-ab} be a submatrix of Q obtained by deleting the a-th row and the b-th column from Q, and p_{ab} is a vector of instantaneous covariances between Y^a and Y^{-b}. Note that both C_{ab} and q_{ab} are functions of μ_t.

Proof: First, we substitute agents' FOCs (9) into the Euclidean norms appearing in the principal's Hamiltonian (5).

$$\begin{aligned}
\left\|(v^a)^\top\sigma\right\|^2 &= Var((v^a)^\top Y) = Var\left\{\tilde{c}^a Y^a + (v^a_{-a})^\top\left(Y^{-a} - \tilde{f}^a Y^a\right)\right\} \\
&= (\tilde{c}^a)^2\sigma^a(\sigma^a)^\top + 2\tilde{c}^a(v^a_{-a})^\top q_{aa} + (v^a_{-a})^\top C_{aa}u^a_{-a}. \\
\left\|\left(1_N^\top - \sum_{a\in A}(v^a)^\top\right)\sigma\right\|^2 &= Var\left(\left(1_N^\top - \sum_{a\in A}(v^a)^\top\right)Y\right) \\
&= Var\left(\sum_{a\in A}\left\{(1-\tilde{c}^a(\mu))Y^a - (v^a_{-a})^\top(Y^{-a} - \tilde{f}^a(\mu)Y^a)\right\}\right) \\
&= \sum_{a\in A}\sum_{b\in A}\left\{(1-\tilde{c}^a)(1-\tilde{c}^b)\sigma^a(\sigma^b)^\top + (v^a_{-a})^\top C_{ab}v^b_{-b} - 2(1-\tilde{c}^b)(v^a_{-a})^\top q_{ba}\right\}.
\end{aligned}$$

Thus, the principal's problem is to choose μ and $\{v^a_{-a}; a \in A\}$ to maximize

$$\begin{aligned}
&\sum_{a\in A} f^a(\mu) - \sum_{a\in A} c^a(\mu^a) - \sum_{a\in A}\frac{R^a}{2}\left\{(\tilde{c}^a)^2\sigma^a(\sigma^a)^\top + 2\tilde{c}^a(v^a_{-a})^\top q_{aa} + (v^a_{-a})^\top C_{aa}v^a_{-a}\right\} \\
&\quad -\frac{R}{2}\sum_{a\in A}\sum_{b\in A}\left\{(1-\tilde{c}^a)(1-\tilde{c}^b)\sigma^a(\sigma^b)^\top + (v^a_{-a})^\top C_{ab}v^b_{-b} - 2(1-\tilde{c}^b)(v^a_{-a})^\top q_{ba}\right\}. \qquad (11)
\end{aligned}$$

Then, the FOC with respect to v^a_{-a} for the above principal's problem is

$$0_{N-1} \quad = \quad R^a\left(\tilde{c}^a q_{aa} + C_{aa}v^a_{-a}\right) + R\sum_{b\in A}\left\{C_{ab}v^b_{-b} - (1-\tilde{c}^b)q_{ba}\right\}.$$

This FOC can be rearranged as in (10) for all $a \in A$. Since D is invertible for all $\mu \in U^N$ by assumption, a unique solution exists, and the solution should be equal to $\{v^{a*}_{-a}, a \in A\}$. □

Theorem 3.2 tells us that given μ^*, v^{a*}_{-a}, the optimal sensitivities of each managerial contract to other managerial effort outcomes, are determined by solving the linear system of equations in Eq.(10). Then optimal μ^* can be found by maximizing the principal Hamiltonian (5) over μ after eliminating v^a_{-a} and v^{aa} from the Hamiltonian by substitution of solutions to Eq.'s (10) and (9) given μ. In the next section we use special cases to explore properties of v^{a*} and μ^*.

4 Special Cases

Arguably, one of the most important issues with a multi-agent principal-agent model is about the structure of the optimal performance measure for each agent. We examine the structure of optimal performance measures using two special cases in the following two subsections. In the first case, we assume the principal is risk neutral, and in the second case, the principal is allowed to be either risk-neutral or risk-averse, and agents are identical in terms of their risk aversion and effort production functions. Then, we further present a subcase of the second case with two identical agents. With the subcase, we examine team effects on each agent's optimal effort choices.

4.1 A Risk-neutral Principal

In this subsection we assume $R = 0$, i.e., the principal is risk-neutral.

Proposition 4.1 *Suppose that $R = 0$ and the matrix $D(\mu)$ is invertible for all $\mu \in U^N$. Then the optimal salary function is linear in Y_1 as in Eq.(7) with*

$$v_{-a}^{a*} = -\tilde{c}^a(\mu^*)C_{aa}(\mu^*)^{-1}q_{aa}(\mu^*).$$

Proof: The proposition follows directly from Eq.(10). □

When the principal is risk-neutral, Proposition 4.1 suggests that the performance based part of the optimal salary function (7) for agent a takes the following form:

$$\tilde{c}^a(\mu^*)\Big[Y_1^a - C_{aa}(\mu^*)^{-1}q_{aa}(\mu^*)(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)\Big].$$

One may interpret the quantity inside the square bracket as the optimal measure of agent a's effort performance. The structure of the measure is intuitive. Recall that C_{aa} is the variance-covariance matrix of $(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)$ and q_{aa} is a covariance vector between Y_1^a and $(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)$. Thus, $C_{aa}(\mu^*)^{-1}q_{aa}(\mu^*)(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)$ is a projection of Y_1^a onto the space generated by $Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a$, or a multiple regression of Y^a on $Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a$. The quantity $Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a$ is a vector of outcomes of other agents' effort after adjusting for agent a's marginal contribution to the outcomes, and becomes independent of agent a's marginal effort, although it might still be related to the noise of Y^a, i.e., $\sigma^a B_1^{\mu^a}$. Thus, having the projection subtracted from Y^a, the resulting performance measure for agent a becomes free from both marginal contributions by other agents to Y^a, and systematic noises that are present in outcomes across all agents, while keeping the information content about μ^a. In other words, the measure is *a sufficient statistic* for μ^a.

To see that the measure is indeed a sufficient statistic, recall the sufficient statistics result in Holmstrom [1982], and the definition of such a statistic in mathematical statistics. First, in mathematical statistics, a function $T^a(Y)$ is said to be sufficient for μ^a if there exist functions $h^a(\cdot), p^a(\cdot)$ such that

$$g(Y, \mu) = h^a(Y, \mu^{-a})p^a(T^a(Y), \mu) \tag{12}$$

for all Y and μ values in the support of the probability distribution function g of Y. Holmstrom also defines $T(Y)$ to be *sufficient at* μ^a if for all $a \in A$ and T^a

$$\frac{g_{\mu^a}(Y_1, \mu)}{g(Y_1, \mu)} = \frac{g_{\mu^a}(Y_2, \mu)}{g(Y_2, \mu)} \tag{13}$$

for almost all $Y_1, Y_2 \in \{Y | T^a(Y) = T^a\}$. If $T(Y)$ is sufficient for all $\mu \in U$, then it is said to be *globally sufficient*. The global sufficiency is equivalent to sufficiency defined by Eq.(12).

Proposition 4.2 *Suppose that $R = 0$. Then, $S(Y)$ is sufficient for μ^a at $\mu^a = \mu^{a*}$, $a \in A$. Furthermore, if f is linear, then $S(Y)$ is globally sufficient for all μ^a.*

Proof: At $\mu = \mu^*$, the performance measure $Y_1^a - C_{aa}(\mu^*)^{-1}q_{aa}(\mu^*)(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)$ is independent of $Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a$ for each $a \in A$. Thus $S^a(Y)$ is independent of $Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a$, because $S^a(Y)$ is linear in $Y_1^a - C_{aa}(\mu^*)^{-1}q_{aa}(\mu^*)(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)$. Therefore, we know that

$$g(Y) = p^a(S(Y))h^a(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)$$

for some $p^a(\cdot)$ and $h_a(\cdot)$. On the other hand, note that since, at $\mu = \mu^*$,

$$\frac{\partial}{\partial \mu^a}\{Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a\} = f_{\mu^a}^{-a}(\mu^*) - \tilde{f}^a(\mu^*)f_{\mu^a}^a(\mu^*) = 0,$$

we have

$$h_{\mu^a}^a(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a) = 0.$$

Now consider Y and Y' such that $S^a(Y) = S^a(Y')$. Then clearly $p^a(S(Y), \mu^*) = p^a(S(Y'), \mu^*)$, and $p_S^a(S(Y), \mu^*) = p_S^a(S(Y'), \mu^*)$. Moreover, $S_{\mu^a}(Y) = \tilde{c}(\mu^*) = S_{\mu^a}(Y')$. Therefore,

$$\frac{p_S^a(S(Y, \mu^*)S_{\mu^a}(Y)}{p(S(Y), \mu^*)} = \frac{p_S^a(S(Y', \mu^*)S_{\mu^a}(Y')}{p(S(Y'), \mu^*)},$$

and the condition (13) is satisfied. This proves the first assertion.

To prove the second assertion, suppose now that f is linear. Then it can be easily shown that $Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a$ is a function of μ^{-a} only, and is completely independent of μ^a. Therefore, for any μ^a, $g(Y)$ can be written as

$$g(Y) = p^a(S(Y), \mu^*)h^a(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a),$$

where $h^a(.)$ is a function of μ^{-a} and independent of μ^a. Since $S(Y)$ is linear in μ^a, using the same argument as in the proof of the first assertion, one can show that condition (13) is also satisfied even when agent a choose μ^a that is different from μ^{a*}. □

Proposition 4.2 reconfirms Holmstrom's *sufficient statistics* result. However, this result depends on the assumption of the risk-neutrality of the principal. In the next section, we examine the structure of optimal performance measure in more generality.

4.2 Identical Agents

In this section, the principal can be either risk-averse or risk-neutral, i.e., $R \geq 0$, and all agents are identical in the following sense. For all $a \in A$, $R^a = \bar{R}$, $c^a(\cdot) = c(\cdot)$, and $f^a(\mu) = f(\mu^a, \mu^{-a})$ where f is a symmetric function of μ^{-a}. Moreover

$$Q = \begin{bmatrix} \sigma_c^2+\sigma_i^2 & \sigma_c^2 & \sigma_c^2 & \cdots & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2+\sigma_i^2 & \sigma_c^2 & \cdots & \sigma_c^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_c^2 & \sigma_c^2 & \sigma_c^2 & \cdots & \sigma_c^2+\sigma_i^2 \end{bmatrix},$$

where σ_c is the volatility of a common shock affecting outputs of all agents identically and σ_i is that of an idiosyncratic shock affecting only one agent. As a result, one may write $Y_1^a = f(\mu^a;\mu^{-a})+\sigma_c B_1^c+\sigma_i B_1^a$, where B^c and B^a are independent standard Brownian motions/normal random variables. Then, we have the following proposition.

Proposition 4.3 *Suppose that agents are identical and that the matrix $D(\mu)$ is invertible for all $\mu \in U^N$. Then the optimal salary function for agent a is linear in Y_1 as in Eq.(7) with*

$$\begin{aligned} v_{-a}^{a*} &= \Big[I + \frac{R}{R+\overline{R}}\sum_{b\neq a} C_{aa}(\mu^*)^{-1}C_{ab}(\mu^*)\Big]^{-1} \\ &\quad \times\Big[\frac{R(1-\tilde{c}(\mu^*))}{R+\overline{R}}C_{aa}(\mu^*)^{-1}\sum_{b\in A} q_{ab}(\mu^*) - \frac{\overline{R}\tilde{c}^a}{R+\overline{R}}C_{aa}(\mu^*)^{-1}q_{aa}(\mu^*)\Big]. \end{aligned}$$

Proof: Since all agents are identical, $v^{a*} = v^{b*}$ for $a, b \in A$, and the proposition follows from equation Eq.(10). □

When all agents are identical, Proposition 4.3 indicates that the performance-based part of the optimal salary function of agent a is given in the following form:

$$\tilde{c}(\mu^*)Y_1^a + (\gamma^1+\gamma^2)^\top\Big(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a\Big), \tag{14}$$

where

$$\begin{aligned} \gamma^1 &= -\frac{\overline{R}\tilde{c}(\mu^*)}{R+\overline{R}}\Big[I + \frac{R}{R+\overline{R}}\sum_{b\neq a} C_{aa}(\mu^*)^{-1}C_{ab}(\mu^*)\Big]^{-1} C_{aa}(\mu^*)^{-1}q_{aa}(\mu^*), \\ \gamma^2 &= \frac{R(1-\tilde{c}(\mu^*))}{R+\overline{R}}\Big[I + \frac{R}{R+\overline{R}}\sum_{b\neq a} C_{aa}(\mu^*)^{-1}C_{ab}(\mu^*)\Big]^{-1} C_{aa}(\mu^*)^{-1}\sum_{b\in A} q_{ab}(\mu^*). \end{aligned}$$

Note that $C_{aa}(\mu^*)^{-1}q_{aa}(\mu^*)(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)$ is a projection of Y_1^a onto the space generated by $Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a$, and that $C_{aa}(\mu^*)^{-1}\sum_{b\in A} q_{ba}(\mu^*)(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a)$ is a projection of total output Y^T onto the space generated by $Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a$ and therefore can be regarded as a transfer of risk by the principal to the agent.

The performance-based part (14) can also be rearranged as follows:

$$\begin{aligned} &\tilde{c}(\mu^*)Y_1^a + (\gamma^1+\gamma^2)^\top\Big(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a\Big) \\ &= \tilde{c}(\mu^*)\Big\{Y_1^a - \left(C_{aa}^{-1}q_{aa}\right)^\top\left(Y^{-a} - \tilde{f}^a Y^a\right)\Big\} \\ &\qquad + \left\{\tilde{c}(\mu^*)\left(C_{aa}^{-1}q_{aa}\right) + \gamma^1 + \gamma^2\right\}^\top\Big(Y_1^{-a} - \tilde{f}^a(\mu^*)Y_1^a\Big). \end{aligned} \tag{15}$$

This rearrangement reveals a striking structure of optimal performance measures. Note that the first component of the RHS of Eq.(15) is a sufficient statistic for μ^a, and the second is a non-sufficient statistic that is orthogonal to the sufficient statistic. In particular, note that the second component vanishes, when the principal is risk-neutral, i.e., $R = 0$, and therefore, the performance measure turns out to be a sufficient statistic. However, if the principal is risk averse, i.e., $R \neq 0$, then the second component becomes nontrivial. That is, if $R \neq 0$, then the resulting aggregate performance measure is not a sufficient statistic. To recapitulate, *the optimal aggregate performance measure for each agent can be orthogonally decomposed into two statistics; one is a sufficient statistic and the other a non-sufficient statistic.*

The structure in Eq.(15) suggests that in constructing the optimal performance measure for agent a, the principal uses effort performance outcomes of other agents for two purposes: one to construct a sufficient statistic for the agent's effort incentives, and the other to construct a non-sufficient statistic for sharing risks with the agent without directly affecting the agent's optimal incentives.

Thus far, we have discussed how optimal contracts and optimal performance measures are constructed, however without explicit solutions. Finally, in next subsection, we provide an explicit solution to the principal's problem with two identical agents, and examine optimal effort decisions under team contracts.

4.3 Two Identical Agents

Let us assume that there are two identical agents with $c^a(\cdot) = c(\cdot)$, $R^a = \bar{R}$ for $a \in A = \{1, 2\}$, and that expected outcomes of agents' effort, f^a's, have identical structures as follows;

$$f^1(\mu) = \mu^1 + \kappa\mu^2, \quad f^2(\mu) = \mu^2 + \kappa\mu^1,$$

where $0 \leq \kappa < 1$ is a constant. We also assume that the variance-covariance matrix Q of the outcomes is given by

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = \begin{bmatrix} \sigma_c^2 + \sigma_i^2 & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2 + \sigma_i^2 \end{bmatrix}.$$

The above set-up with two identical agents implies that each agent a mutually affects the other agent's expected effort outcome by $\kappa\mu^a$, and that outcomes of both agents' effort are correlated with each other through the common uncertainty with a covariance of σ_c^2.

Let $c_x(\cdot)$ and $c_{xx}(\cdot)$ denote the first and second derivatives of $c(\cdot)$, respectively, with respect to the argument. Then, we have

$$\tilde{c}^1(\mu) = c_x(\mu^1), \quad \tilde{c}^2(\mu) = c_x(\mu^2),$$

$$C_{11} = (1 + \kappa^2)Q_{11} - 2\kappa Q_{12} = C_{22},$$

$$C_{12} = (1 + \kappa^2)Q_{12} - 2\kappa Q_{11} = C_{21},$$

$$q_{11} = Q_{12} - \kappa Q_{11} = q_{22},$$

and

$$q_{12} = Q_{11} - \kappa Q_{12} = q_{21}.$$

By (10), the principal's FOCs with respect to v^{12} and v^{21} are as follows:

$$\begin{aligned}
(\bar{R}+R)C_{11}v^{12} + RC_{12}v^{21} &= R\big(1-c_x(\mu^1)\big)q_{11} + R\big(1-c_x(\mu^2)\big)q_{21} - \bar{R}c_x(\mu^1)q_{11}, \\
(\bar{R}+R)C_{22}v^{21} + RC_{12}v^{12} &= R\big(1-c_x(\mu^2)\big)q_{22} + R\big(1-c_x(\mu^1)\big)q_{12} - \bar{R}c_x(\mu^2)q_{22}.
\end{aligned}$$

Substituting these FOCs back into the principal's Hamiltonian (11), and maximizing it with respect to μ^1 and μ^2, we have

$$\mu^1 = \mu^2 = \mu^*,$$

$$v^{12} = v^{21} = \frac{R\big(1-c_x(\mu^*)\big)(q_{11}+q_{12}) - \bar{R}c_x(\mu^*)q_{11}}{(\bar{R}+R)C_{11} + RC_{12}},$$

where μ^* satisfies the following FOC:

$$\begin{aligned}
0 &= 1+\kappa - c_x(\mu^*) - (\bar{R}+R)c_x(\mu^*)c_{xx}(\mu^*)Q_{11} + Rc_{xx}(\mu^*)Q_{11} \\
&\quad +\big(1-c_x(\mu^*)\big)Rc_{xx}(\mu^*)Q_{12} \\
&\quad -\left\{\bar{R}q_{11} + R(q_{11}+q_{12})\right\} c_{xx}(\mu^*)\frac{R\big(1-c_x(\mu^*)\big)(q_{11}+q_{12}) - \bar{R}c_x(\mu^*)q_{11}}{(\bar{R}+R)C_{11} + RC_{12}}.
\end{aligned} \tag{16}$$

The first five terms of the RHS of the above equation are identical to those which would also appear in the FOC for a risk-averse principal's problem with a single agent. Alternatively, the first *six* terms of the RHS would have been the FOC for a risk-averse principal's problem, had she contracted each agent based on the agent's own outcome alone. We call this alternative case *the case of individually separate contracts* under which the contract for agent a is given in the form of *constant* $+\, c_x(\mu_I)Y^a$, where μ_I satisfies

$$\begin{aligned}
0 &= 1+\kappa - c_x(\mu_I) - (\bar{R}+R)c_x(\mu_I)c_{xx}(\mu_I)Q_{11} + Rc_{xx}(\mu_I)Q_{11} \\
&\quad +\big(1-c_x(\mu_I)\big)Rc_{xx}(\mu_I)Q_{12}.
\end{aligned} \tag{17}$$

Then, comparing team contracts in this paper with individually separate contracts, one may say the last term of the RHS of (16) captures "team" contracting effects. We shall compare explicit solutions resulting from the two different contracting regimes.

For explicit solutions, assume quadratic costs such that $c(x) = \frac{K}{2}x^2$. Then, by using (16), we have

$$K\mu^* = \frac{NUM}{DEN}, \tag{18}$$

where

$$
\begin{aligned}
NUM &= 1+\kappa+RK(Q_{11}+Q_{12})-\frac{RK(q_{11}+q_{12})\{R(q_{11}+q_{12})+\bar{R}q_{11}\}}{R(C_{11}+C_{12})+\bar{R}C_{11}}, \\
DEN &= 1+RK(Q_{11}+Q_{12})+\bar{R}KQ_{11}-\frac{K\{R(q_{11}+q_{12})+\bar{R}q_{11}\}^2}{R(C_{11}+C_{12})+\bar{R}C_{11}}.
\end{aligned}
$$

It is easy to show that $R(C_{11}+C_{12})+\bar{R}C_{11}>0$. By some calculation one can also show that both NUM and DEN are strictly positive. On the other hand, by using (17), we compute optimal effort levels under individually separate contracts, denoted by μ_I, as follows:

$$
K\mu_I=\frac{1+\kappa+RK(Q_{11}+Q_{12})}{1+RK(Q_{11}+Q_{12})+\bar{R}KQ_{11}}. \tag{19}
$$

In order to examine team effects on optimal levels of effort, we compare (18) and (19). The comparison is summarized in the following proposition.

Proposition 4.4 *$\mu^*<,(=,>)\mu_I$ if and only if $DET>(=,<)0$, where*

$$
\begin{aligned}
DET &= \left\{R(1-\kappa)(Q_{11}+Q_{12})+\bar{R}(Q_{12}-\kappa Q_{11})\right\} \\
&\quad\times\left[R(Q_{11}+Q_{12})\left\{\bar{R}K(Q_{11}-Q_{12})-\kappa(1-\kappa)\right\}-(1+\kappa)\bar{R}\left\{Q_{12}-\kappa Q_{11}\right\}\right].
\end{aligned}
$$

Proof: The proof is immediate from the following algebraic relation: for real numbers x,y,z,w, if $x>0$ and $x-z>0$, then the sign of $\frac{y}{x}-\frac{y-w}{x-z}$ is equal to that of $xw-yz$. Using this, the sign of $K\mu_I-K\mu^*$ is equal to that of $\frac{K}{R(C_{11}+C_{12})+\bar{R}C_{11}}\times DET$. □

As a corollary to Proposition 4.4, we consider two sub-special cases: (i) risk-neutral principal, i.e., $R=0$, and (ii) no common uncertainty and joint productions, i.e., $\sigma_c=\kappa=0$. The first case enables us to isolate team contract effects on effort incentives, and the second case to examine those on risk-sharing, which in turn can also, albeit indirectly, affect optimal effort levels.

Corollary 4.1

i. Suppose that $R=0$ and $q_{11}=Q_{12}-\kappa Q_{11}\neq 0$. Then $DET=-(1+\kappa)\bar{R}^2q_{11}^2<0$, and thus $\mu^*>\mu_I$. If $R=q_{11}=0$, then $\mu^*=\mu_I$.

ii. Suppose $\sigma_c=\kappa=0$. In this case, if $R>(=)0$, then $DET=R^2\bar{R}KQ_{11}^3>(=)0$, and thus $\mu^*<(=)\mu_I$.

Recall that when $R=0$, the principal is only concerned with agents' effort incentives, and when $q_{11}\neq 0$, the other manager's outcome after adjusting from agent 1's contribution is still correlated with the agent's outcome. Thus, the principal can utilize this correlation to reduce the volatility of the agent 1's effort performance measure, and as a result, team contracts can be higher-powered in effort incentives than individually separate contracts. However, if $q_{11} = 0$ and $R = 0$, then team contracts cause no effects on effort incentives.

When $\sigma_c = \kappa = 0$, the other agent's outcome is not useful in reducing the volatility of agent 1's effort performance measure. However, the other agent's outcome can still be utilized for risk-sharing purposes, reducing risk-burden on the principal. Intuitively, when $R > 0$, under individually separate contracts, the risk-averse principal can share outcome risks using each agent's own outcome only, and thus when additional risk-sharing is desired, she has to increase effort incentives excessively more than she would, were she risk-neutral. However, under team contracts, the principal can utilize the agent's own outcome mostly for incentives and other agents' outcomes mostly for risk-sharing. As a result, each agent's own outcome is less excessively used for risk-sharing purposes than it is under individually separate contracting arrangements. Thus, when $\sigma_c = \kappa = 0$, effort incentives under team contracts are lower than those under individually separate contracts.

The two benchmark cases described in Corollary 4.1 suggest that team contracts can improve both incentives and risk sharing. Risk-sharing can be improved because team contracts enable the principal to distribute risks unrelated to effort incentives across agents.[8] Incentives can be also improved because correlations among individual and joint outcomes can be utilized to reduce the volatility of the effort performance measure for each agent. However, improvements in both incentives and risk-sharing under team contracts can result in either higher- or lower-powered effort incentives than those under individually-separate contracting can. For, improvement in risk-sharing drives down the contract sensitivity to effort-performance measure, whereas improvement in incentives due to reduction in the volatility drives up the sensitivity.

5 Conclusion

In this paper we have presented a multi-agent principal-agent moral-hazard problem in continuous time. We have derived optimal salary functions for all agents as a team and shown that they are jointly linear in all effort outcomes. Based on this linearity result, we have examined the general structure of optimal performance measures, and have shown that the optimal aggregate performance measure for each agent can be decomposed into two orthogonal statistics: one being a sufficient statistic mainly for incentives and the other a non-sufficient statistic for risk-sharing.

In particular, we have shown that Holmstrom's sufficient statistics result is valid if the principal is risk neutral. However, when the principal is risk averse, the agent's salary function depends not only on a sufficient statistic but on a non-sufficient statistic that is independent of the agent's marginal effort. The inclusion of the non-sufficient statistic in each contract is from the principal's motivation to share

[8]This distribution effect is consistent with the diversification effect extensively discussed in our companion paper, Koo, Shim and Sung [2006].

total-output risks with agents. As a result, the aggregate performance measure, in general, cannot be a sufficient statistic.

We have examined team effects on contracting by comparing our team contracts with individually separate contracts. Team contracting helps improve both incentives and risk-sharing, by enabling the principal to reduce volatilities of effort performance measures, and to distribute total-output risks among many agents. However, team contracts do not always provide higher effort incentives than individually separate contracts do. The reason is that as compared with individually-separate contracting, improvements in both incentives and risk-sharing under team contracting affect optimal effort incentives in the opposite directions: the improvement in incentives increases contract sensitivities to effort performance measures, but the improvement in risk-sharing decreases the sensitivities. We have shown that if the principal is risk averse and each outcome contains no joint outputs and is independent of all other outcomes, then each contract sensitivity to the effort performance measure under team contracting is lower than that under individually separate contracting.

References

1. Cadenillas A., J. Cvitanic, and F. Zapatero, "Leverage decision and manager compensation with choice of effort and volatility," *Journal of Financial Economics*, 73, 2003, 71-92.

2. Core, J. E. and W. Guay, "When efficient contracts require risk-averse executives to hold equity: Implications for option valuation, for relative performance evaluation, and for the corporate governance debate ," Working Paper, 2003, The Wharton School, University of Pennsylvania.

3. J. Cvitanic, X. Wan and J. Zhang, "Continuous-Time Principal-Agent Problems with Hidden Action: The Weak Formulation," Working Paper, 2005, California Technology of Institute.

4. Dye, R. A., "Relative Performance Evaluation and Project Selection," *Journal of Accounting Research* 30, Spring 1992, 27-52.

5. Gibbons, R. and K. Murphy, "Relative performance evaluation for chief executive officers," *Industrial and Labor Relations Review* 43, 1990, 30-51.

6. Groves, T. "Incentives in Teams," *Econometrica* 41, 617–631.

7. Holmstrom B. "Moral Hazard in Teams," *Bell Journal of Economics*, 1982, 324-340.

8. Holmstrom B. and P. Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55, 1986, 303-328.

9. Koo, H., G. Shim and J. Sung, "Hierarchical Contract, Firm Size, and Pay Sensitivity," Working Paper, 2006.

10. Lazear, E. and S. Rosen, "Rank-order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 1981, 841-864.

11. Melumad, N.D., D. Mookherjee, and S. Reichelstein, "Hierarchical Decentralization of Incentive Contracts," *Rand Journal of Economics*, 26, 1995, 654-672.

12. McAfee R.P., and J. McMillan, "Optimal Contracts for Teams," *International Economic Review*, 32, 1991, 561-577.

13. Mookherjee, D., "Optimal Incentive Schemes with Many Agents," *Review of Economic Studies*, 51, 1984, 433–446.

14. Müller, H., "The First-Best Sharing Rule in the Continuous-Time Principal-Agent Problem with Exponential Utility," *Journal of Economic Theory*, 79, 1998, 276-280.

15. Nalebuff, B. and J. Stiglitz, "Prizes and Incentives: toward a general theory of compensation and competition," *Bell Journal of Economics*, 1983, 21-43.

16. Ou-Yang, H., "An Equilibrium Model of Asset Pricing and Moral Hazard," *Review of Financial Studies*, 18, 2005, 1253-1303.

17. Schättler, H., and J. Sung, "The First-Order Approach to Continuous-Time Principal-Agent Problem with Exponential Utility," *Journal of Economic Theory*, 61, 1993, 331-371.

18. Sung, J. *Managerial Compensation and Corporate Management under Moral Hazard*, unpublished Ph.D. thesis, Washington University in St. Louis, 1991.

19. Sung, J., "Linearity with Project Selection and Controllable Diffusion Rate in Continuous-Time Principal-Agent Problems," *Rand Journal of Economics*, 26, 1995, 720-743.

20. Sung, J., "Corporate Insurance and Managerial Incentives," *Journal of Economic Theory*, Vol.74, 1997, 297-332.

21. Williams, N., "On Dynamic Principal-Agent Problems in Continuous Time," Working paper, 2003, Princeton University and NBER.