On the Optimal Team Incentive Compensation Schemes

Baomin Dong[1]

Department of Economics

Concordia University

Montréal, Québec H3G 1M8

e-mail: don@alcor.concordia.ca

submitted 2001; revised 2001

JEL classification: C61 D82

Keywords and phrases: Moral Hazard, Brownian Motion, Linearity,

Abstract

This paper studies optimal incentive schemes in a team setting and we explicitly derive a closed form solution to the compensation scheme where the underlying output processes are correlated drifted Brownian motions. Some implications of the model are discussed.

[1] I thank David Desjardins and Yongge Tian for some useful discussions on related technical issues.

1 Introduction

Incentive schemes need to be robust in either single agent or multiple agent case. A simple incentive scheme may perform well across a wide range of environments as well as having low writing costs. Intricate schemes are designed for the purpose of inducing agents to exert effort in the principal's best interest in particular environments. However these fine-tuning complex incentive schemes are not very realistic for at least two reasons: first, they could hardly maintain their optimality for even a slightest change in the information structure or technological change; second, intuitively, intricate schemes are hard to be implemented, for instance, agents would take advantage of the complexity of the contracts by arbitrage-taking like behaviours, and this situation can be exaggerated by multiple agents because collusion and deviating from optimal reciprocal help level will be their options. A more rudimentary concern is monotonicity of the incentive contracts. As standard first order approach pointed out, an optimal incentive scheme has not to be monotone if MLRP does not hold. However if free disposal of the output is feasible, a practical scheme then has to be monotone despite any other argument on either information structure or stochastic production technology it may have. If sabotage is feasible to some level, the output would be undermined seriously if the scheme is not monotone. With this concern removed, we move onto the linearity of the optimal contracts.

The robustness of compensation linearity was studied in Holmstrom and Milgrom

(1987) for a single agent case. Mirrlees (1974) claims that a step function scheme can induce optimal e®ort level of the agent as well as linear contract. However the e®ort level chosen by the agent is just one shot. If instead, more generally, the agent choose e®ort level everyday, over the course of a company ¯scal year, if the compensation is a step function of the aggregate output of his account, then his e®ort level chosen will vary day to day. By the end of year, he would certainly do nothing in the o± ce if the output already exceeds the critical level, the same thing if the realized aggregated output is too far below the critical level. To rescue step function in this aspect, one would possibly adjust the scheme into a day to day step function scheme, but this revisionism fails because if the agent labours under the same one-day step contract throughout the year the aggregate compensation itself is a linear function of the aggregate output. Furthermore, step functions do not converge to ¯rst-best if some assumptions on utility functions and technology are not clearly speci¯ed even under one-shot e®ort case, pointed out by Holmstrom and Milgrom (1987).

In the continuous-time principal-agent framework, Holmstrom and Milgrom(1987) and Scattler and Sung(1993, 1997) suggest an optimal sharing rule for the principal given the underlying output process is a Brownian motion. In their models, the optimal sharing rule is a linear function of aggregate output. Kim and Wang(1998) argue that the optimal incentive contract is nonlinear when the agent is risk averse and the principal and agent are set in a double moral hazard situation in which the principal also participates in the production process. Some others constantly cast doubt on the robustness of the linearity. We agree that real world schemes need to be robust. It is not enough for a scheme performs optimally in a limited environment but also a changing environment because constant changes of schemes are simply not feasible. Therefore regularities about the shape of the optimal sharing rule are essential to modeling. Linear contracts are observed across a wide range of technologies, companies and industries, on the other hand, managerial teams, instead of single manager, are on play, in most modern capitalistic and partnership ¯rms.

Several other issues arise together with the optimality of linear incentive contracts in teams. Many have realized that the managers' role is not only exerting e®ort to improve the NPV of the ongoing projects, but also the choice on future projects to be carried out. Even only in the ongoing projects, an entrepreneur's function can be categorized into both mean improving and variance reducing activities, for example, creating strategic alliance with friendly rivals in the industry, spending time with potential clients, strategic actions in ¯nancial markets, everyday on-¯eld management and supervision, or, some other activities like regular maintenance of production facilities, reducing unnecessary risks to the ¯rm by activities such as maintaining a modest relationship with workers' union to avoid strikes, or be careful on every tip of the ongoing projects in order to minimize lawsuits ¯led by unsatis¯ed customers. While some of these activities are not too hard to observe, most are. Sung (1995) realizes this fact and claims that optimal schemes are still linear in ¯nal outcome alone even when the agent is allowed to control variance. Sung further shows that the sensitivity must be low to correct the manager's incentive to choose the right project, o®ering an alternative interpretation for Jensen-Murphy puzzle. Some chunks of our analysis were inspired by this observation, nevertheless, we focus on problems arise in teams, for instance, the optimal level of help between two team members in characterized in the optimal schemes we proposed. The incentive problem concerned in our model is two-fold, ¯rst, an optimal contract should fully exploit agents' comparative advantage on di®erent types of activities, for example, some careful and versatile people may be good at everyday human management and facility maintenance, while some very creative and knowledgable about current business lines may be good at dealing with tough clients or carrying out new projects, it is therefore better to let them work more on the tasks they specialized in; secondly, the relative performance evaluation in our model does exist but is not to ¯lter out noise, instead, to encourage cooperation between the team members by placing correct contractual incentives.

The structure of the rest of the paper follows. Section 2 introduces the framework of the model. Section 3 gives a closed form optimal linear incentive scheme under group performance evaluation. Section 4 solves the optimal incentive scheme under individual performance evaluation. Section 5 discusses team related issues like reciprocal helps and other comparative statics. Section 6 summarizes the essay.

2 The Model

We start with a two-dimensional Brownian motion where the two agents control the drift rate of the Brownian motions and the Wiener processes are correlated. Both agents and principal have negative exponential utilities with different coefficients of absolute risk aversion. Time is normalized to one.

Formally, the processes are governed by a simultaneous differential equation system of the form,

$$dX_1 = a_1(t)dt + \sigma_1 dB_1(t);$$

$$dX_2 = a_2(t)dt + \sigma_2(\rho dB_1(t) + \sqrt{1 - \rho^2}dB_2(t));$$

where $dB_1; dB_2(t)$ are independent Brownian motions. Production technology is additive, i.e.,

$$dY = dX_1 + dX_2$$

$$= [a_1 + a_2]dt + dB_4(t) \times (\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho)^{\frac{1}{2}};$$

where $dB_4(t) = [(\sigma_1 + \sigma_2\rho)dB_1(t) + \sqrt{1 - \rho^2}dB_2(t)]/(\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho)^{\frac{1}{2}}$:

The principal's problem is to find a sharing rule $S_i(Y)$ and controls $a_{ii}; a_{ij}$ to maximize his utility.

$$\max_{S(Y);a_{ii};a_{ij}} E[-\exp\{-R(Y(1) - \sum_{i=1}^{2} S_i(Y)]; \tag{2.1}$$

subject to

$$
\begin{aligned}
dX_1 &= a_1(t)dt + \sigma_1 dB_1(t); &(2.2)\\
dX_2 &= a_2(t)dt + \sigma_2(\rho dB_1(t) + \sqrt{1-\rho^2}dB_2(t));
\end{aligned}
$$

$$E[-\exp\{-r(S_i(Y) - \int_0^1 c(a_i)dt] \geq -\exp\{-rW_{CE}\}; \quad \text{and} \tag{2.3}$$

$$a_i \in \arg\max_{a_i \in A} E[-\exp\{-r(S_i(Y) - \int_0^1 c(a_i)dt] \tag{2.4}$$

where W_{CE} is the agents' certainty equivalent at time 0 in monetary terms; $c(\cdot)$ is the disutility function for the agents; a_i is the agent i's effort;:and $R; r$ are the coefficients of absolute risk aversion to the principal and agents, respectively.

3 Optimal Group Performance Evaluation

For the utility maximization problem $(2{:}1)$; we have the value function as the following,

$$V\left(t; \sum x_i\right) = \max_{S;\, a} E\left[\left\{\exp(-R(-dS(\sum x_i)))\right\} V(t+dt; \sum x_i\right] : \tag{3.1}$$

Note that Ito's Lemma can be written

$$
\begin{aligned}
dG &= \left(\frac{\partial G}{\partial x}\mu + \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}\sigma^2\right)dt + \frac{\partial G}{\partial x}\sigma dZ\\
&= \frac{\partial G(x)}{\partial x}(\mu dt + \sigma dZ) + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}\sigma^2 dt\\
&= \frac{\partial G(x)}{\partial x}dx + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}\sigma^2 dt;
\end{aligned}
$$

where $G(x)$ is a function of the drifted Brownian motion $dx = \mu dt + \sigma dZ$. Therefore for a vector drifted Brownian motion and a scalar function of the vector below,

$$
\begin{aligned}
dx(t)G &= a(t)dt + \Sigma(t)dB(t) \qquad \text{and} \\
y(t) &= \psi(t; x(t));
\end{aligned}
$$

Ito's Lemma is

$$
\begin{aligned}
dy(t) &= d\psi(t; x(t)) = \left\{ \psi_t(t; x(t)) + \psi_x(t; x(t))a(t) + \frac{1}{2}\mathrm{tr}\,\{\psi_{xx}(t; x(t))\Sigma(t)\Sigma'(t)\} \right\} dt \\
&\quad + \psi_x(t; x(t))\Sigma(t)dB(t).
\end{aligned}
$$

Therefore we have the first best compensation scheme's total differentiate as

$$
\begin{aligned}
dS(t; x(t)) &= \left\{ \frac{\partial S}{\partial t} + \left(\frac{\partial S}{\partial x_1}; \frac{\partial S}{\partial x_2} \right) \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \frac{1}{2}\mathrm{tr} \left\{ \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} [\sigma_1; \sigma_2] \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix} \right\} \right\} dt \\
&\quad + \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} \left(\sigma_1 \quad \sigma_2 \right) d \begin{bmatrix} B_1(t) \\ \rho B_1(t) + \sqrt{1 - \rho^2} B_2(t)) \end{bmatrix} \\
&= \alpha dt + \beta dx
\end{aligned}
$$

where $S_i \equiv \frac{\partial S}{\partial x_1}$; $S_{ij} \equiv \frac{\partial S}{\partial x_i \partial x_j}$; $\alpha = \frac{1}{2}\mathrm{tr}\left\{ \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} [\sigma_1; \sigma_2] \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix} \right\}$ and $\beta = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}$.

The value function (3.1) is hence rewritten as

$$
\begin{aligned}
& V(t; x_i) && (3.2)\\
=\; & \max_{S;\,a} E\left[\left(1 + R(\alpha dt + \beta dx) + \frac{1}{2}R^2(\alpha dt + \beta dx)^2\right)\right.\\
& \left.\left(V(\cdot) + \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial x}dx + \frac{1}{2}\mathrm{tr}\left(\frac{\partial^2 V}{\partial x^2}\sigma\sigma'\right)dt\right)\right].
\end{aligned}
$$

In a second best world where individual output is not observable and hence compensation could only be based on aggregate output, that is

$$
\begin{aligned}
dY(t) &= (a_1 + a_2)dt + \sigma_1 dB_1 + \rho\sigma_2 dB_1 + \sqrt{1-\rho^2}\,\sigma_2 dB_2\\
&= C dt + D\begin{bmatrix} dB_1 \\ dB_2 \end{bmatrix}
\end{aligned}
$$

where $C = a_1 + a_2;\ D = \left[\sigma_1 + \rho\sigma_2;\ \sqrt{1-\rho^2}\,\sigma_2\right]$.

The value depends on a scalar Y instead of the vector x, i.e.,

$$
\begin{aligned}
& V\left(t; \sum x_i\right) && (3.3)\\
=\; & \max_{S;\,a} E\left[\left(1 + R(\alpha_1 dt + \beta_1 dY) + \frac{1}{2}R^2(\alpha_1 dt + \beta_1 dY)^2\right)\right.\\
& \left.\left(V(t; Y) + \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial Y}dY + \frac{1}{2}\left(\frac{\partial^2 V}{\partial Y^2}DD'\right)dt\right)\right].
\end{aligned}
$$

Rewriting (3.3) yields

$$
\begin{aligned}
0 \equiv\; & \frac{\partial V}{\partial t}dt + \frac{1}{2}\left(\frac{\partial^2 V}{\partial Y^2}DD'\right)dt\\
& + \max_{\alpha_1;\,\beta_1;\,a}\left[\frac{\partial V}{\partial Y}(a_1 + a_2 + R\beta_1 DD')dt + RV(\cdot)\left(\beta_1(a_1 + a_2) + \alpha_1 + \frac{R}{2}\beta_1^2 DD'\right)dt\right.
\end{aligned}
$$

with transversality condition $V(1;Y) \equiv -\exp\{-RY(1)\}$; and agents' participation constraints:

$$E[-\exp\{-r(\frac{1}{2}dS(Y) - c(a_i)dt]\} \ge -1;$$

for $i = 1;2$:

The agents' participation constraints imply,

$$0 \equiv \frac{\alpha_1}{2}dt + \frac{\beta_1}{2}(a_1 + a_2)dt - c(a_i)dt - \frac{r}{8}\beta_1^2 DD'dt$$

which can be rewritten as

$$\alpha_1^* \equiv 2c(a_i^*) - 2\beta_1 a_i^* + \frac{1}{4}r\beta_1^2 DD': \tag{3.4}$$

Once β_1 is obtained, optimal α is calculated from the transformation of the agents' participation constraint $(3{:}4)$:

With the agents' participation constraint, we now rewrite the dynamic programming problem as

$$\begin{aligned} 0 \equiv\ & \frac{\partial V}{\partial t}dt + \frac{1}{2}\left(\frac{\partial^2 V}{\partial Y^2}DD'\right)dt \\ & + \max_{\beta_1;a}\left[\frac{\partial V}{\partial Y}(a_1 + a_2 + R\beta_1 DD')dt + RV(\cdot)\left(2c(a_i^*) + \frac{2R + r}{4}\beta_1^2 DD'\right)dt\right]: \end{aligned}$$

The value function can be treated functionally as the principal's certainty equivalent or any linear transformation of the certainty equivalent. Again by Taylor's expansion, we have

$$0 \equiv \frac{\partial \Psi}{\partial t}dt + \frac{1}{2}\left(\frac{\partial^2 \Psi}{\partial Y^2}DD^{\top}\right)dt - \frac{R}{2}\left(\frac{\partial \Psi}{\partial Y}\right)^2 DD^{\top}dt \tag{3.5}$$
$$+\max_{\beta_1, a}\left[\frac{\partial \Psi}{\partial Y}(a_1 + a_2 + R\beta_1 DD^{\top})dt - 2c(a_i^*)dt - \frac{2R + r}{4}\beta_1^2 DD^{\top}dt\right],$$

where $\Psi \equiv \Psi(t, Y)$, $V(t, Y) = -\exp\{-R\Psi(t, Y)\}$, with terminal condition $\Psi(1, Y) = Y(1)$.

Differentiating (3.5) yields the first order conditions for the principal's utility maximization problem, and these conditions are

$$\frac{\partial \Psi}{\partial Y}RDD^{\top}dt - \frac{2R + r}{2}\beta_1^2 DD^{\top}dt = 0 \Rightarrow \beta_1^* \equiv \frac{R}{R + \frac{r}{2}}\frac{\partial \Psi}{\partial Y}, \tag{3.6}$$

$$\frac{\partial \Psi}{\partial Y} = 2c'(a_i^*). \tag{3.7}$$

Substituting the first order conditions above back to the maximand (3.5) yields

$$0 \equiv \frac{\partial \Psi}{\partial t}dt + \frac{1}{2}\left(\frac{\partial^2 \Psi}{\partial Y^2}DD^{\top}\right)dt - \frac{R}{2}\left(\frac{\partial \Psi}{\partial Y^2}\right)^2 DD^{\top}dt \tag{3.8}$$
$$+\max_{\beta_1, a}\left[\begin{array}{l} 2c'(a_i^*)\left(a_1 + a_2 + 2c'(a_i^*)\frac{R^2}{R + \frac{r}{2}}DD^{\top}\right)dt \\ - 2c(a_i^*)dt - 2\left(c'(a_i^*)\right)^2 \frac{R^2}{R + \frac{r}{2}}DD^{\top}dt \end{array}\right]$$
$$= \frac{\partial \Psi}{\partial t} + \frac{1}{2}\frac{\partial^2 \Psi}{\partial Y^2}DD^{\top} - \frac{R}{2}\left(\frac{\partial \Psi}{\partial Y^2}\right)^2 DD^{\top}$$
$$+2c'(a_i^*)\left[c'(a_i^*)\frac{R^2}{R + \frac{r}{2}}DD^{\top} + a_1 + a_2\right] - 2c(a_i^*).$$

To solve the partial differential equation (3.8), we guess a solution of the form $\Psi(t, Y) = \lambda(t)Y^3 + \phi(t)Y^2 + \kappa(t)Y + \psi(t)$ with the terminal constraints $\lambda(1) =$

$0; \phi(1) = 0; \kappa(1) = 1;$ and $\psi(1) = 0.$ Therefore we have

$$\frac{\partial \Psi}{\partial Y} = 3\lambda(t)Y^2 + 2\phi(t)Y + \kappa(t);$$

$$\frac{\partial \Psi}{\partial t} = \dot{\lambda}(t)Y^3 + \dot{\phi}(t)Y^2 + \dot{\kappa}(t)Y + \dot{\psi}(t);$$

$$\frac{\partial^2 \Psi}{\partial Y^2} = 6\lambda(t)Y + 2\phi(t);$$

$$\left(\frac{\partial \Psi}{\partial Y}\right)^2 = 9\lambda^2(t)Y^4 + 12\lambda(t)\phi(t)Y^3 + (4\phi^2(t) + 6\lambda(t)\kappa(t))Y^2 + 2\phi(t)\kappa(t)Y + \kappa^2(t)$$

and the utility maximization is then

$$\begin{aligned}
0 \equiv\ & 9\lambda^2(t)Y^4\left(-\frac{R}{2}\right)DD^{\top} + \left(\dot{\lambda}(t) - \frac{R}{2}DD^{\top}12\lambda(t)\phi(t)\right)Y^3 \\
& + \left(\dot{\phi}(t) - \frac{R}{2}DD^{\top}\left[4\phi^2(t) + 6\lambda(t)\kappa(t)\right]\right)Y^2 \\
& + \left(\dot{\kappa}(t) - \phi(t)\kappa(t)RDD^{\top} + \frac{R^2}{2}6\lambda\right)Y + \dot{\psi}(t) + DD^{\top}\phi(t) \\
& - \frac{R}{2}DD^{\top}\kappa^2(t) + 2c'_-(\cdot)\left[c'_-(\cdot)\frac{R^2}{R + \frac{r}{2}}DD^{\top} + a_1 + a_2\right] - 2c_-(\cdot)
\end{aligned}$$

Collecting the terms yields a system of ordinary differential equations:

$$0 = -\frac{9}{2}\lambda^2(t)RDD^{\top};$$

$$0 = \dot{\lambda}(t) - \frac{R}{2}DD^{\top}12\lambda(t)\phi(t);$$

$$0 = \dot{\phi}(t) - \frac{R}{2}DD^{\top}(4\phi^2(t) + 6\lambda(t)\kappa(t));$$

$$0 = \dot{\kappa}(t) - \phi(t)\kappa(t)RDD' + 3\lambda DD'$$

$$0 = \dot{\psi}(t) - \frac{R}{2}\kappa(t)DD' + 2c_i'(\cdot)\left[c_i'(\cdot)\frac{R^2}{R+\frac{r}{2}}DD' + a_1 + a_2\right] - 2c_i(\cdot).$$

The above ordinary differential equation system, together with the boundary conditions, yield a set of solutions: $\lambda(t) = 0$; $\phi(t) = 0$; $\kappa(t) = 1$; and $\dot{\psi}(t) = \frac{R}{2}DD'\kappa^2(t) - 2c_i'(\cdot)\left[c_i'(\cdot)\frac{R^2}{R+\frac{r}{2}}DD' + a_1 + a_2\right] + 2c_i(\cdot)$. Therefore the

$$\Psi(t;Y) = Y + \left\{\frac{R}{2}DD'\kappa^2(t) - 2c_i'(\cdot)\left[c_i'(\cdot)\frac{R^2}{R+\frac{r}{2}}DD' + a_1 + a_2\right] + 2c_i(\cdot)\right\}(t-1). \tag{3.9}$$

Differentiating (3.9) with respect to Y gives

$$\frac{\partial\Psi(t;Y)}{\partial Y} = 1.$$

Recall the first order condition (3.7), we can have

$$c_i' = \frac{1}{2}. \tag{3.10}$$

By (3.10) and the first order condition (3.6), we then have

$$\beta_1 = \frac{R}{2R+r}. \tag{3.11}$$

To get α_1, simply substitute (3.11) into (3.4), and we have

$$S^* = S(0;Y) + 2\int_0^1 c(a_i)dt + \int_0^1 \frac{R}{2R+r}(dY - (a_1+a_2)dt) + r\int_0^1 \left(\frac{R}{2R+r}\right)^2 \frac{DD'}{4}dt.$$

We present the linear sharing rule result below.

Proposition 1 (Efficient sharing rule) The solution to the principal's problem (1.1)-(1.4) is

$$
\begin{aligned}
S^*(Y) = w_{CE} + 2\int_0^1 c(a_i)dt &+ \int_0^1 \frac{R}{2R+r}\Big(dY - \sum_i a_i(t)dt\Big) \\
&+ r\int_0^1 \left(\frac{R}{2R+r}\right)^2 \frac{DD'}{4}dt
\end{aligned}
\tag{3.12}
$$

where the effort level a_i is uniquely and implicitly defined by $c'(a_i) = 1$.

4 Optimal Individual Performance Evaluation

In this section, we study the optimal performance evaluation scheme in a two-person team where individual performance for each agent is observable. To make our analysis more vivid, we work on a more general setting where agents' control variables include the diffusion term related ones, and, the output processes are interrelated.

Let $X_1(t)$ and $X_2(t)$ be governed by the following stochastic differential equation system:

$$
\begin{aligned}
dX_1 = {} & (f(a_{11}(t))X_1(t) + f(a_{12}(t))X_2(t) + \mu_1)dt \\
& + \sigma_1\sqrt{h(b_{11}(t))X_1(t) + h(b_{12}(t))X_2(t) + \varpi}\, dB_1(t);
\end{aligned}
\tag{4.1}
$$

$$
\begin{aligned}
dX_2 = {} & (g(a_{21}(t))X_1(t) + g(a_{22}(t))X_2(t) + \mu_2)dt \\
& + \sigma_2\sqrt{l(b_{21}(t))X_1(t) + l(b_{22}(t))X_2(t) + \varpi}\left(\rho dB_1(t) + \sqrt{1-\rho^2}dB_2(t)\right);
\end{aligned}
\tag{4.2}
$$

where $dB_1; dB_2(t)$ are once again independent Brownian motions. To save some notations, we define

$$Z_1 \equiv h(b_{11}(t))X_1(t) + h(b_{12}(t))X_2(t) + \varpi;$$

$$Z_2 \equiv l(b_{21}(t))X_1(t) + l(b_{22}(t))X_2(t) + \varpi;$$

and

$$dB_3(t) \equiv \rho dB_1(t) + \sqrt{1 - \rho^2}\, dB_2(t):$$

We then immediately have

$$dX_1 dX_1 = \sigma_1^2 Z_1 dt; \qquad dX_1 dX_2 = \rho\sigma_1\sigma_2\sqrt{Z_1 Z_2}\, dt; \qquad dX_2 dX_2 = \sigma_2^2 Z_2 dt:$$

We know that under the optimal sharing rule, the agents will implement the effort level that the principal anticipates and they are constants over time. Therefore, (4:1) and (4:2) can be rewritten as

$$dX_1 = (f(a_{11})X_1(t) + f(a_{12})X_2(t) + \mu_1)dt + \sigma_1\sqrt{Z_1(t)}\, dB_1(t); \tag{4.3}$$

$$dX_2 = (g(a_{21})X_1(t) + g(a_{22})X_2(t) + \mu_2)dt + \sigma_2\sqrt{Z_2(t)}\, dB_3(t); \tag{4.4}$$

The value at time $t \le \infty$ of a contract executed from time 0, to the agent, is

$$E\left[-\exp\left\{ -r\left(\int_t^\infty S_i(X_i(s); X_j(s))ds - \int_t^\infty c_i(a_{ij}; b_{ij}; a_{ji}; b_{ji})ds \right) \right\} \Big| F(t) \right]$$

Since the coefficients in (4.3) and (4.4) are not time dependent but only the expected maturity $\tau \equiv 1 - t$, and the process pair (X_1, X_2) is of Markov, there exists a function $H(X_1, X_2, 1 - t)$ such that.

$$H(X_1, X_2, 1 - t) = E\left[-\exp\left\{ -r\left(\int_t^1 S_i(X_i(s), X_j(s))ds - \int_t^1 c_i(a_{ij}, b_{ij}, a_{ji,} b_{ji})ds \right) \right\} \Big| F(t) \right],$$

where $\{F(t)\}$ is a filtration in which $\{X_1\}$ and $\{X_2\}$ are adapted.

By the tower property, we have that

$$\left\{ -\exp\left(-r \int_0^t S_i(X_i(s), X_j(s))ds - \int_0^t c_i(a_{ij}, b_{ij}, a_{ji,} b_{ji})ds \right) \right\} H(X_1, X_2, 1 - t)$$

is a martingale. Likewise,

$$\left\{ -\exp\left(-r \int_0^t S_i(X_i(s), X_j(s))ds \right) \right\} M(X_1, X_2, 1 - t)$$

is also a martingale where

$$M(X_1, X_2, 1 - t) \equiv -\exp\left\{ -r\left(\int_t^1 S_i(X_i(s), X_j(s))ds \right) \right\}.$$

Therefore, we have

$$
\begin{aligned}
& d\left(-\exp\left\{-r\int_0^t S_i(X_i(s),X_j(s))ds\right\}M(X_1,X_2,1-t)\right)\\
= \;& r\exp\left\{-r\int_0^t S_i(X_i(s),X_j(s))ds\right\}[X_1Mdt - M_{X_1}dX_1 - M_{X_2}dX_2 + M_\tau dt\\
& -\frac{1}{2}rM_{X_1X_1}(dX_1)^2 - \frac{1}{2}rM_{X_1X_2}dX_1dX_2 - \frac{1}{2}rM_{X_2X_2}(dX_2)^2]\\
= \;& r\exp\left\{-r\int_0^t S_i(X_i(s),X_j(s))ds\right\}[(X_1M - (f(a_{11}(t))X_1(t) + f(a_{12}(t))X_2(t)\\
& +\mu_1)M_{X_1} - (g(a_{21}(t))X_1(t) + g(a_{22}(t))X_2(t) + \mu_2)M_{X_2} + M_\tau\\
& -\frac{1}{2}r\sigma_1^2(h(b_{11}(t))X_1(t) + h(b_{12}(t))X_2(t) + \varpi)M_{X_1X_1} - r\rho\sigma_1\sigma_2\sqrt{Z_1Z_2}M_{X_1X_2}\\
& -\frac{1}{2}r\sigma_2^2(l(b_{21}(t))X_1(t) + l(b_{22}(t))X_2(t) + \varpi)M_{X_2X_2})dt + \sigma_1\sqrt{Z_1}M_{X_1}dB_1\\
& +\sigma_2\sqrt{Z_2}M_{X_2}dB_3]
\end{aligned}
$$

and

$$
\begin{aligned}
0 = \;& X_1M - (f(a_{11}(t))X_1(t) + f(a_{12}(t))X_2(t) + \mu_1)M_{X_1} - (g(a_{21}(t))X_1(t) + M_\tau\\
& +g(a_{22}(t))X_2(t) + \mu_2)M_{X_2} - \frac{1}{2}r\sigma_1^2(h(b_{11}(t))X_1(t) + h(b_{12}(t))X_2(t) + \varpi)M_{X_1X_1}\\
& -\sqrt{h(b_{11}(t))X_1(t) + h(b_{12}(t))X_2(t) + \varpi}\sqrt{l(b_{21}(t))X_1(t) + l(b_{22}(t))X_2(t) + \varpi}\\
& \times r\rho\sigma_1\sigma_2M_{X_1X_2} - \frac{1}{2}r\sigma_2^2(l(b_{21}(t))X_1(t) + l(b_{22}(t))X_2(t) + \varpi)M_{X_2X_2}.
\end{aligned} \tag{4.5}
$$

Our premise for the solution is of the form

$$
M(X_1,X_2,1-t) = -\exp\{-r(X_1^2C_1(\tau) + X_2^2C_2(\tau) + X_1C_3(\tau) + X_2C_4(\tau) + A(\tau))\}, \tag{4.6}
$$

The initial conditions must be incentive compatible, that is,

$$C_1(0) = C_2(0) = C_3(0) = C_4(0) = A(0) = 0; \tag{4.7}$$

as $\tau = 0$ corresponds to $t = 1$:

In order to find $C_1(\tau); C_2(\tau); C_3(\tau); C_4(\tau)$; and $A(\tau)$ for $\tau > 0$; first we find the derivatives for $M(\cdot)$ as follows

$$M_\tau(\cdot) = -r(X_1^2 C_1'(\tau) + X_2^2 C_2'(\tau) + X_1 C_3'(\tau) + X_2 C_4'(\tau) + A'(\tau))M(\cdot); \tag{4.8}$$

$$M_{X_1}(\cdot) = -r(2X_1 C_1(\tau) + C_3(\tau))M(\cdot); \tag{4.9}$$

$$M_{X_2}(\cdot) = -r(2X_2 C_2(\tau) + C_4(\tau))M(\cdot); \tag{4.10}$$

$$M_{X_1X_1}(\cdot) = 2r^2 C_1(\tau)(2X_1 C_1(\tau) + C_3(\tau))M(\cdot); \tag{4.11}$$

$$M_{X_2X_2}(\cdot) = 2r^2 C_2(\tau)(2X_2 C_2(\tau) + C_4(\tau))M(\cdot); \tag{4.12}$$

$$M_{X_1X_2}(\cdot) = r^2(2X_2 C_2(\tau) + C_4(\tau))(2X_1 C_1(\tau) + C_3(\tau))M(\cdot): \tag{4.13}$$

Therefore (4:5) can be rewritten into

$$
\begin{aligned}
0 \;=\;& M(\cdot)\{X_1 - r(X_1^2 C_1'(\tau) + X_2^2 C_2'(\tau) + X_1 C_3'(\tau) + X_2 C_4'(\tau) + A'(\tau)) \qquad (4.14)\\
&+ r[f(a_{11}(t))X_1(t) + f(a_{12}(t))X_2(t) + \mu_1](2X_1C_1(\tau) + C_3(\tau))\\
&+ r[(g(a_{21}(t))X_1(t) + g(a_{22}(t))X_2(t) + \mu_2)(2X_2C_2(\tau) + C_4(\tau))\\
&- \frac{1}{2} r^3 \sigma_1^2 (h(b_{11}(t))X_1(t) + h(b_{12}(t))X_2(t) + \varpi) 2C_1(\tau)(2X_1C_1(\tau) + C_3(\tau))\\
&- r^3 \rho \sigma_1 \sigma_2 \sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\,\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\\
&\times (2X_2C_2(\tau) + C_4(\tau))(2X_1C_1(\tau) + C_3(\tau))\\
&- \frac{1}{2} r^3 \sigma_2^2 [l(b_{21}(t))X_1(t) + l(b_{22}(t))X_2(t) + \varpi) 2C_2(\tau)(2X_2C_2(\tau) + C_4(\tau)]\}.
\end{aligned}
$$

Without loss of generality, let f_{11}; f_{12}; g_{11}; g_{12}; h_{11}; h_{12}; l_{11}; l_{12} denote $f(a_{11}(t))$; $f(a_{12}(t))$; $g(a_{21}(t))$; $g(a_{22}(t))$; $h(b_{11}(t))$; $h(b_{12}(t))$; $l(b_{21}(t)$; $l(b_{22}(t))$; respectively. Then we can further rewrite (4.14) into

$$
\begin{aligned}
0 = {} & X_1 M(\cdot)\{1 + r f_{11}C_3 - rC_3' + 2r\mu_1 C_1 + r g_{11}C_4 + 2r\mu_2 C_2 - r^3\sigma_1^2 h_{11}C_1C_3 \\
& - 2r^3\sigma_1^2\varpi C_1^2 - 2r^3\rho\sigma_1\sigma_2\sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\,C_1C_4 \\
& - r^3\sigma_2^2 l_{11}C_2C_4\} + rX_2M(\cdot)\{-C_4' + f_{12}C_3 + g_{22}C_4 + 2\mu_2C_2 - r^2\sigma_1^2h_{12}C_1C_3 \\
& - 2r^2\rho\sigma_1\sigma_2\sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\,C_2C_3 - 2r^2\sigma_2^2C_2^2\varpi \\
& - r^2\sigma_2^2C_4C_2l_{22}\} + rX_1^2M(\cdot)\{-C_1' + 2f_{11}C_1 - 2r^2\sigma_1^2h_{11}C_1^2\} \\
& + rX_2^2M(\cdot)\{-C_2' + 2g_{22}C_2 - 2r^2\sigma_2^2l_{22}C_2^2\} + rX_1X_2M(\cdot)\{2f_{12}C_1 + g_{11}C_2 \\
& - 2r^2\sigma_1^2h_{22}C_1^2 - 4r^2\rho\sigma_1\sigma_2\sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\,C_1C_2 \\
& - 2r^2\sigma_2^2l_{21}C_2^2\} + rM(\cdot)\{-A' + \mu_1C_3 + \mu_2C_4 - r^2\sigma_1^2C_1C_3\varpi \\
& - r^2\rho\sigma_1\sigma_2\sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\,C_3C_4 \\
& - r^2\sigma_2^2C_2C_4\varpi\}.
\end{aligned}
\tag{4.15}
$$

We then from (4.15) have five equations:

$$C_1' = 2f_{11}C_1 - 2r^2\sigma_1^2h_{11}C_1^2; \tag{4.16}$$

$$C_2' = 2g_{22}C_2 - 2r^2\sigma_2^2l_{22}C_2^2; \tag{4.17}$$

$$
\begin{aligned}
C_3' = {} & r^{-1} + f_{11}C_3 + 2\mu_1C_1 + g_{11}C_4 + 2\mu_2C_2 - r^2\sigma_1^2h_{11}C_1C_3 \\
& - 2r^2\sigma_1^2\varpi C_1^2 - 2r^2\rho\sigma_1\sigma_2\sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\,C_1C_4 \\
& - r^2\sigma_2^2l_{11}C_2C_4;
\end{aligned}
\tag{4.18}
$$

$$
\begin{aligned}
C_4' = \; & f_{12}C_3 + g_{22}C_4 + 2\mu_2 C_2 - r^2\sigma_1^2 h_{12}C_1C_3 - r^2\sigma_2^2 C_4C_2 l_{22} \\
& - 2r^2\rho\sigma_1\sigma_2 \sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\, C_2C_3 - 2r^2\sigma_2^2 C_2^2 \varpi;
\end{aligned} \tag{4.19}
$$

$$
\begin{aligned}
& 2r^2\rho\sigma_1\sigma_2 \sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\, C_1C_2 \\
& = f_{12}C_1 + g_{11}C_2 - r^2\sigma_1^2 h_{22}C_1^2 - r^2\sigma_2^2 l_{21}C_2^2.
\end{aligned} \tag{4.20}
$$

$$
\begin{aligned}
A' = \; & \mu_1 C_3 + \mu_2 C_4 - r^2\sigma_1^2 C_1C_3\varpi - r^2\sigma_2^2 C_2C_4\varpi \\
& - r^2\rho\sigma_1\sigma_2 \sqrt{h_{11}X_1 + h_{12}X_2 + \varpi}\sqrt{l_{21}X_1 + l_{22}X_2 + \varpi}\, C_3C_4;
\end{aligned} \tag{4.21}
$$

Combined with initial conditions (4.7); we solve (4.16) to (4.20) simultaneously, and then integrate (4.21) to obtain $A(\tau)$.

We propose a solution that is compatible with the differential equation system (4.16) to (4.21) in the following proposition. This solution is analogous to its' single agent counterpart.

Proposition 2 Suppose both drift and diffusion terms are control variables to the agents; agents' disutility of effort $c(\cdot)$ is convex in drift enhancing efforts and diffusion reducing efforts; and let $(f_{11}^*; f_{12}^*; h_{11}^*; h_{12}^*)$ and $(g_{21}^*; g_{22}^*; l_{21}^*; l_{22}^*)$ be the control vectors of agent 1 and 2, respectively, that maximize the principal's expected utility and also incentive compatible. Then a linear optimal compensation scheme to agent 1 is given by

$$S^* = S(0, X_1, X_2) + c_1(f^*_{11}, f^*_{12}, h^*_{11}, h^*_{12}) + \frac{c^*_{f_{11}}}{f'(a^*_{11})}[X_1 - f^*_{11} - g^*_{21} - \mu_1] + \frac{c^*_{f_{12}}}{f'(a^*_{12})}[X_2 - f^*_{12} - g^*_{22} - \mu_2] + \frac{r}{2}\left[\left(\frac{c^*_{f_{11}}}{f'(a_{11})}\right)^2 \sigma_1^2 Z_1^* + \left(\frac{c^*_{f_{12}}}{f'(a^*_{12})}\right)^2 \sigma_2^2 Z_2^*\right]. \quad (4.22)$$

The scheme to agent 2 is similar.

This optimal linear sharing rule can be interpreted as follows: the first two terms in (4.22) remunerate the agent a certainty opportunity cost of working in a team for the effort exerted in the unit time period; the third term provides the agent with appropriate incentives for value increasing activities in his own part of the work however it is net of the expectation, note the mathematical expectation for this term is 0; the fourth term provides proper incentives for the agent to exert effort in helping his team mate in project NPV mean improving activities, also net of the expectation; the last term is a risk premium paid to the agent for participating two risky projects with zero means. Note if the agent is risk neutral, in other words, $r = 0$, then the last term drops.

5 Extension and Discussion (to be extended)

In this section we will present a number of results in the forms of lemmas and corollaries of Proposition 2 and sequels of martingale analysis in the previous section. The content is categorized into the following:

1. Comparison between some other non-linear schemes and linear scheme. This is complicate by having diffusion as an extra control variable. We will show that although some nonlinear schemes can slightly outperform the linear scheme, the unnecessary risk added by nonlinear contracts will reduce their credential substantially therefore making the linear or nearly linear scheme thrive in practice.

2. Comparison between competition oriented and attribution oriented linear schemes. Previous research emphasizes the competition amongst the agents within an agency where yardstick or tournament like schemes are proclaimed to be efficient. Yet yardstick competition and tournaments have their virtues when the information is coarse and existing literature has very well exploited it. However our suspicion aroused when we do not find such extreme schemes that could severely penalize an agent given peer performance fluctuates. In a fast paced techno-metabolic and changing business world nowadays, sharp outliers of business performance either individual personal or firm level could be very possibly happening, however not all of the team mates are laid-off, nor are the companies going bankruptcy. In fact, small sized teams emphasizing cooperation, synchronization and synergy are favoured by many industries. Heterogeneity among the agents is recognized easily by the principal but instead of aligning the heterogeneities, firms prefer to best utilize the idiosyncracy of the employ-

ees by assigning them di®erent jobs to bring their comparative advantages into play. Therefore, in our modeling, even if the sensitivity of the scheme is similar to that in other analyses, our interpretation di®ers. Furthermore, since cooperation and synergy are encouraged here, our policy suggestion on job allocation would be di®erent.

3. Comparative statics. The slope of the sensitivity of the optimal incentive scheme is the major concern. We calculate its' derivatives on own and cross control variables, its' derivatives on agents' idiosyncratic parameters or functions. Some illustrative and intuitive conjectures are tested. The optimal linear incentive scheme proposed can instruct the agents to best help their counterparts reciprocally.

4. Finally, grouping is once again a typical issue we will visit. The fundamental question here is, given the production technology is additive, should a team constituted by homogeneous or heterogeneous agents?

6 Concluding Remarks

This section will be based on the results in section 3, 4 and 5.

Our analysis shows that either individual output is observable or not, linear incentive schemes can always outperform other schemes or approximate optimal schemes.

If individual performance is observable, we discuss some interesting topics including best relative performance comparison schemes, own and cross e®ort e®ect on incentive sensitivity in optimal linear schemes, teaming etc. We ¯nd that the main theme in a team where the team members and the owner are all risk averse, cooperation is fundamental to the team's prosperity.

References

1. Chen, Jingliang et al, Modern Applied Analysis (in Chinese), Tsinghua Press, Beijing, (1998),

2. Diamond, Peter, Managerial Incentives: On the Near Linearity of Optimal Compensation, J.P.E. 106(5): 931-957, (1987),

3. Durrett, Rick, Stochastic Calculus: A Practical Introduction, CRC, (1996),

4. Hart, Oliver and B. Holmstrom, The Theory of Contract, in T. Bewley, (eds), Advances in Economic Theory: Fifth World Congress, Cambridge University Press (1987),

5. Hellwig, Martin and K. Schmidt, Discrete-time Approximation of the Holmstrom-Milgrom Brownian-motion Model of Intertemporal Incentive Provision, Working Paper, University of Mannheim, (1998),

6. Holmstrom, Bengt and P. Milgrom, Aggregation and Linearity in the Provision of Intertemporal Incentives, Econometrica 55, 303-328, (1987),

7. Itoh, Hideshi, Incentives to Help in Multi-Agent Situation, Econometrica, 59: 611-636, (1991)

8. Karatzas, Ioannis and S. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, New York, (1991),

9. Kim, Son Ku and S. Wang, Linear Contracts and the Double Moral Hazard, J.E.T., vol 82:342-378 (1998),

10. Meyer, Margaret, The Dynamic of Learning with Team Production: Implications for Task Assignment, Q.J.E., 1157-1184 (1994),

11. Mirrlees, Jim, Notes on Welfare Economics, Information and Uncertainty, in Essays on Economic Behavior Under Uncertainty, ed. by M. Balch, D. McFadden and S. Wu. North-Holland, Amsterdam: 243-258 (1974),

12. ____ and J. Vickers, Performance Comparisons and Dynamic Incentive, J.P.E. , 105: 547-581 (1997),

13. Scattler, Heinz and J. Sung, The First-Order Approach to the Continuous-Time Principal-Agent Problem with Exponential Utility, J.E.T., vol. 61: 331-371, (1993),

14. ____ and J. Sung, On Optimal Sharing Rules in Discrete- and Continuous-time Principal-agent Problems with Exponential Utility, J. of Econ. Dynamics and Control, vol. 21: 551-574, (1997),

15. Sung, Jaeyoung, Linearity with Project Selection and Controllable Di®usion Rate in Continuous-time Principal-agent Problems, Rand J. of Econ., vol. 26 (4): 720-743, (1995),