Motivating innovation with a structured incentives scheme under continuous states

Chengli Zheng [1], Yan Chen [2]

Abstract –*The problem of incentive is an important component of the separation of ownership and control. A large literature focuses on the problem of how to use pay-for-performance schemes to both inspirit agents to exert effort and to deter agent-based resource tunneling. Manso (2011) proposes the use of structured incentive schemes with two periods to motivate innovation under discrete states. In combining these two perspectives, this paper propose a version with continuous states and point that agent can simultaneously innovate while exerting effort to obtain greater output per unit time. By offered the suitable incentive contract, the agent will take the action plan of exploration though he may get a failure. At the mean time, he will exert his all effort to get more production which determines his reward. This will explain many things in managerial compensation, such as a combination of stock options with long vesting periods, option re-pricing, golden parachutes and managerial entrenchment.*

Keywords*: Motivating innovation, Structured incentive scheme, Exploitation and exploration*

Authors' information

Chengli Zheng, School of Economics, Huanzhong Normal University, China
Applied Mathematics & Statistics, Stony Brook University, USA.
Email1: zhengchengli168@gmail.com

Yan Chen, School of Economics, Huanzhong Normal University, China.
Email2: stchenyan88@163.com

Motivating innovation with a structured incentives scheme under continuous states

Chengli Zheng [1], Yan Chen [2]

Abstract –*The problem of incentive is an important component of the separation of ownership and control. A large literature focuses on the problem of how to use pay-for-performance schemes to both inspirit agents to exert effort and to deter agent-based resource tunneling. Manso (2011) proposes the use of structured incentive schemes with two periods to motivate innovation under discrete states. In combining these two perspectives, this paper propose a version with continuous states and point that agent can simultaneously innovate while exerting effort to obtain greater output per unit time. By offered the suitable incentive contract, the agent will take the action plan of exploration though he may get a failure. At the mean time, he will exert his all effort to get more production which determines his reward. This will explain many things in managerial compensation, such as a combination of stock options with long vesting periods, option re-pricing, golden parachutes and managerial entrenchment.*

***Keywords**: Motivating innovation; Structured incentive scheme; Exploitation and exploration; continuous state*

I. Introduction

Since Berle and Means' (1932) point the drawbacks with the separation of ownership and control, the incentive problem becomes the interest of this field, see Harris and Raviv (1978) and Holmstrom (1979) and the sequel. Most of them focus on the problems of how to inspirit the agent to exert effort or deter the agent from tunneling resources away from the corporation by applying principal-agent models. Manso(2011) presents a different view. He studies how to build a certain structure of incentives to motivate the agent to be more innovative with a two-period model. He shows that incentive schemes that motivate innovation should be structured differently from standard pay-for-performance schemes used to induce effort or avoid tunneling. Innovation involves the exploration of new untested approaches that are likely to fail. Therefore, standard pay-for-performance schemes that punish failures with low rewards and termination may in fact have adverse effects on innovation. In contrast, the optimal incentive scheme that motivates innovation exhibits substantial tolerance (or even reward) for early failure and reward for long-term success. Under this incentive scheme, compensation depends not only on total performance, but also on the path of performance; an agent who performs well initially but poorly later earns less than an agent who performs poorly initially but well later or even an agent who performs poorly repeatedly.

Based on the framework of Manso(2011), this paper study the incentives for innovation with non-fixed reward for the agent. Our model absorbs the advantages of the two directions above mentioned: incentive schemes for motivating innovation and standard pay-for -performance schemes. We give the standard of success, and the reward of the agent depends on the amount of the excess output over the baseline. The fixed wage and non-fixed wage (wage rate) are designed. These structured incentives can motivate the agent to select a more innovative work method and stimulate the agent to exert effort to get a better output at the mean time. The reward of the agent comprise of two parts: one fixed part which is independent at any situations, and another non-fixed part which depends on the output. The fixed part mainly be used to tolerate the failure of the exploration, and the on-fixed part is used to stimulate the agent to engage the innovation action and to exert his all effort to get the best reward.

Similar to Manso(2011), we use a two-period innovation process to deal with the incentives problem. To model the innovation process, we use a class of Bayesian decision models known as bandit problems. We focus on the central concern that arises in bandit problems: the tension between the exploration of new untested actions and the exploitation of well-known actions. The related literature see Holmstrom (1989), Aghion and Tirole (1994), Arrow (1969), March (1991), Moscarini and Smith (2001), Hellmann and Thiele (2009), Tian and Wang (2010), Ederer and Manso (2010) and other literature cited in Manso(2011).However, there are differences here, too. The model of Manso(2011) just consider two states: success and failure, and the optimal contracts depend only on the probability of success or failure, not on the amount of outputs. Our model is treated under the continuous states, and the optimal contracts depend on the distribution of the production---not only on the probability of success or failure, but also on the amount of outputs.

The rest of the paper is arranged as follows: section Π gives the bandit problem for tension between exploration and exploitation; section Ш presents the principal-agent problem about the tension; section IV gives the solutions of the principal-agent problem, namely the optimal incentive contracts for exploration and exploitation, respectively; and the last section concludes.

II. The Bandit Problem for tension between Exploration and Exploitation

Here, We review the two-armed bandit problem with one known arm as Manso(2011) and Zheng&Chen(2012). It illustrates the tension between exploration and exploitation. The original models are under discrete states. We extend it to be one model with continuous states.

We assume that the agent lives for only two periods. In each period $t \in T = \{1,2\}$, the agent takes an action $i \in I$, producing output R_{ti}, which is a random variable with cumulative distribution function $F_{R_{ti}}(x) = P[R_{ti} \le x]$. The principal gives the baseline B_t of output for each period $t \in T$ to evaluate the performance of the agent. If $R_{ti} > B_t$, the agent is judged as "success"; if $R_{ti} \le B_t$ the agent is judged as "failure". The cumulative distribution function $F_{R_{ti}}(x)$ may be unknown for some actions. To obtain information about $F_{R_{ti}}(x)$ for these actions, the agent needs to engage in experimentations in the first period. We let $h(R_{ti})$ denote the return function on output R_{ti}. And we let $E[h(R_{ti})]$ denote the unconditional expectation of $h(R_{ti})$, let $E[h(R_{ti}) | R_{t-1j} > B_{t-1}]$ denote the conditional expectation of $h(R_{ti})$ given a success on action j in last period, and $E[h(R_{ti}) | R_{t-1j} \le B_{t-1}]$ denote the conditional expectation of $h(R_{ti})$ given a failure on action j in last period. When the agent takes action $i \in I$ in period $t \in T$, he only learns about the information for the distribution of R_{t+1i} for the next period, so that

$$E[h(R_{ti})] = E[h(R_{ti}) | R_{t-1j}] \text{ for } i \ne j$$

This means that if the agent wants to know the information for the distribution of R_{t+1i} for the next period, he must engage in experimentation of action i with unknown distribution in this period.

Because there is no new information for unconditional expectation of $h(R_{ti})$, namely, it is independent of time, so we denote $E[h(R_{ti})] = E[h(R_i)]$ in this situation.

Our main interest focus on the tension between two actions: action 1 is exploration and action 2 is exploitation. We assume that in each period $t \in T$ the agent chooses between these two actions. Action 1 is the conventional work method, has a known distribution of R_{t1} in any period $t \in T$, namely $R_{t1} = R_1$, such that

$$E[h(R_{t1})] = E[h(R_{t1}) | R_{t-11}] = E[h(R_1)]$$

Action 2 is the new work method, has an unknown distribution of R_{t2} such that[1]

$$E[h(R_{t2}) | R_{t-12} \le B_{t-1}] < E[h(R_{t2})] < E[h(R_{t2}) | R_{t-12} > B_{t-1}]$$

This means that if the agent observes a success with the new work method, then he updates his beliefs that there is more possibility the new work method will succeed. Or, if the agent observes a failure with the new work method, then he updates his beliefs that there is more possibility the new work method will fail.

We assume that Action 2 has exploratory nature. This means that when the agent experiments with the new work method, he is initially not as likely to succeed as when he conforms to the conventional work method. However, if the agent observes a success with the new work method, then he updates his beliefs about the probability of success with the new work method, so that the new work method becomes perceived as better than the conventional work method. This is captured as follows:

$$E[h(R_2)] < E[h(R_1)] < E[h(R_{t2}) | R_{t-12} > B_{t-1}]$$

In fact, the agent may shirk, he do not choose any of the two work method above mentioned. This action 0 is allowed in the model. Shirking has zero private cost, but has a lower expected return than either of the two work methods. Here, we assume that action 0(shirking) has a return R_0 with known distribution in any period $t \in T$. Without lose generality, we assume that there exist stochastic dominances relationship as follows:

[1] Here we assume that $h(R_{t2})$ is increasing function on R_{t2}.

$$\left(R_{t2}\middle|R_{t-12} > B_{t-1}\right) \overset{FSD}{\succ} R_1 \overset{FSD}{\succ} R_2 \overset{FSD}{\succ} \left(R_{t2}\middle|R_{t-12} \le B_{t-1}\right) \overset{FSD}{\succ} R_0$$

Where $X \overset{FSD}{\succ} Y$ means that X stochastically dominates Y in first order, namely $F_X(\eta) \le F_Y(\eta)$ for all $\eta \in \mathrm{R}$.

So, if $h(\bullet)$ is non-decreasing function, we have

$$E[h(R_0)] < E[h(R_{t2})|R_{t-12} \le B_{t-1}] < E[h(R_2)] < E[h(R_1)] < E[h(R_{t2})|R_{t-12} > B_{t-1}] \qquad (1)$$

In fact, the model is a three-armed bandit problem, namely $I = \{0,1,2\}$, but we only consider the tension between exploration and exploitation. The agent is risk-neutral and has a discount factor normalized to one. The agent thus chooses an action plan $\langle i_k^j \rangle$ to maximize his total expected payoff. Where $i \in I$ is the first-period action, $j \in I$ is the second-period action in the case of success in the first period, and $k \in I$ is the second-period action in the case of failure in the first period.

Two action plans need to be considered. Action plan $\langle 1_1^1 \rangle$, which Manso(2011) call exploitation, is just the repetition of the conventional work method. Action plan $\langle 2_1^2 \rangle$, which Manso call exploration, is to initially try the new work method, stick to the new work method in the case of success in the first period, and revert to the conventional work method in the case of failure in the first period. Apparently, the total payoff of action plan $\langle 2_1^2 \rangle$ from exploration is higher than that of action plan $\langle 1_1^1 \rangle$ from exploitation if and only if

$$E[R_2] > E[R_1] - E\{1_{R_{12} > B_1}(E[R_{22}|R_{12} > B_1] - E[R_1])\}$$

If the agent tries the new work method, he obtains information about R_{t2}. This information is useful for the agent's decision in the second period, since the agent can switch to the conventional work method if he learns that the new work method is not worth pursuing. The agent may thus be willing to try the new work method even though the initial expected return $E[h(R_2)]$ with the new work method is lower than expected return $E[h(R_1)]$ with the conventional work method.

III. The Principal-agent Problem

In this section, we introduce incentive problems to the three-armed bandit problem with two known arms as reviewed in the previous section.

The principal hires an agent to perform the task described in the previous section. In each period, the agent incurs private costs $c_i \ge 0$ if he takes action $i = 1,2$, but can avoid these private costs by taking action $i = 0$, shirking ($c_0 = 0$).

We assume that the principal does not observe the actions taken by the agent. As such, before the agent starts working, the principal offers the agent a contract $\langle \vec{\lambda}, \vec{w} \rangle = \{\langle \lambda_1, w_1 \rangle, \langle \lambda_2, w_2 \rangle, \langle \lambda_3, w_3 \rangle\}$ that specifies the agent's wages contingent on future performance. The agent has limited liability, meaning that his wages can not be negative. Here, w_s ($s = 1,2,3$) are fixed wages, which are the minimum wages in any situations. And λ_s is the wage rate for extra return at situation of success. This means that if it is a failure, the agent will get a fixed wage w_s, if it is a success, he will get a fixed wage w_s plus flexible wage $\lambda_s(R_s - B_s)1_{R_s > B_s}$. Specifically, $\langle \lambda_1, w_1 \rangle$ is the wage rate and fixed wage in the first period, respectively. $\langle \lambda_2, w_2 \rangle$ is the wage rate and fixed wage in the second period conditional on the situation of success in the first period, respectively. And $\langle \lambda_3, w_3 \rangle$ is the wage rate and fixed wage in the second period conditional on the situation of failure in the first period, respectively.

Different from that of Manso(2011), the contract $\langle \vec{\lambda}, \vec{w} \rangle$ in our model is not a fixed wage. While fixed wage in the situation of failure, but $\vec{\lambda}$ is fixed wage rate in the situation of success. When the agent succeed in one period t, according to the baseline of success B_t given by the principal in advance, he will get a payoff w_s plus $\lambda_s(R_s - B_s)1_{R_s > B_s}$, $s = 1,2,3$, which is dependent of the output. The more output it produces, the more wage reward he gets. So, the contract $\langle \vec{\lambda}, \vec{w} \rangle$ of our Principal-agent model has two functions: one is to motivate the agent to be more innovative and the other is to inspirit the agent to exert effort.

And it is different from that of Zheng&Chen(2012), where the w_s is not minimum wage, which may lead to the situation that the wage in success will be lower that in failure. Here, we revise this fault.

In addition to these differences, another feature is that the models here are built with continuous states. To illustrate the process with reward structure, see the figure 1 as follows.

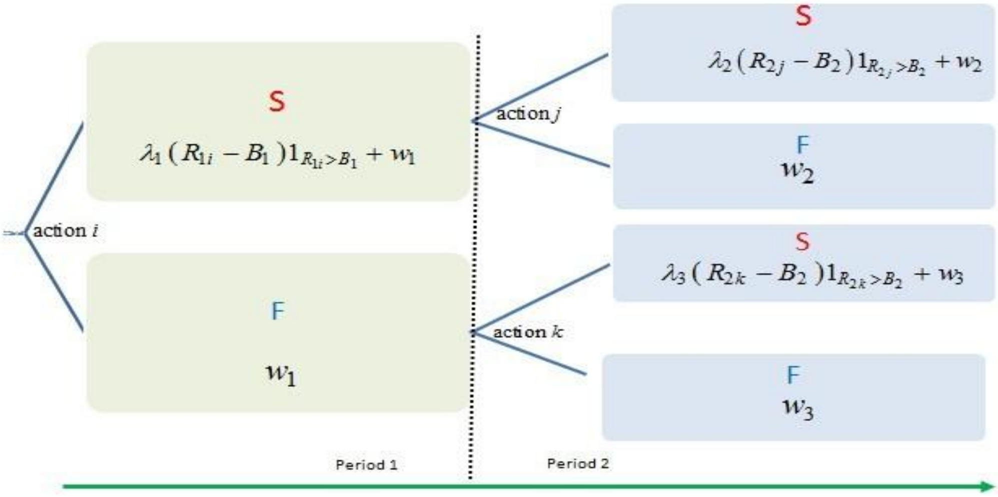

Figure 1 structured reward of action plan $\langle i_k^j \rangle$

S—success, F--failure

We assume that both the principal and the agent are risk-neutral and have a discount factor of one, just for simplicity. When the principal offers the agent a contract $\langle \vec{\lambda}, \vec{w} \rangle$ and the agent takes action plan $\langle i_k^j \rangle$, the total expected payments from the principal to the agent are given by

$$\begin{aligned} W\left(\vec{\lambda},\vec{w},\langle i_k^j\rangle\right) &= E[\lambda_1(R_{1i}-B_1)1_{R_{1i}>B_1}+w_1] \\ &+E\left\{1_{R_{1i}>B_1}E[\lambda_2(R_{2j}-B_2)1_{R_{2j}>B_2}+w_2|R_{1i}>B_1]\right\} \\ &+E\left\{1_{R_{1i}\le B_1}E[\lambda_3(R_{2k}-B_2)1_{R_{2k}>B_2}+w_3|R_{1i}\le B_1]\right\} \end{aligned} \tag{2}$$

Apparently, the model of Manso (2011) and Zheng & Chen (2013) are special discrete cases of our model.

Because $E\left[(R_{ti}-B_t)1_{R_{ti}>B_t}\right]$ can be viewed as a call option whose underlying asset is output R_{ti} and strike price is B_t, we denote $op_{ti}=E\left[(R_{ti}-B_t)1_{R_{ti}>B_t}\right]$.

Similarly, we denote

$\overline{op_{2j}^{1i}}=E[(R_{2j}-B_2)1_{R_{2j}>B_2}|R_{1i}>B_1]$, and $\underline{op_{2k}^{1i}}=E[(R_{2k}-B_2)1_{R_{2k}>B_2}|R_{1i}\le B_1]$.

So the equation (2) can be rewritten as

$$\begin{aligned} W\left(\vec{\lambda},\vec{w},\langle i_k^j\rangle\right) &= \lambda_1 op_{1i}+w_1 \\ &+E\left\{1_{R_{1i}>B_1}(\lambda_2\overline{op_{2j}^{1i}}+w_2)\right\} \\ &+E\left\{1_{R_{1i}\le B_1}(\lambda_3\underline{op_{2k}^{1i}}+w_3)\right\} \end{aligned} \tag{3}$$

It means that the total expected payments comprise of a series of options.

According to the assumptions in the previous section, we have

$$op_{t0}<\underline{op_{22}^{12}}<op_{t2}<op_{t1}<\overline{op_{22}^{12}} \tag{4}$$

When the agent takes action plan $\langle i_k^j \rangle$, the total expected costs incurred by the agent are given by

$$C\left(\langle i_k^j\rangle\right)=c_i+E\left[1_{R_{1i}>B_1}\right]c_j+E\left[1_{R_{1i}\le B_1}\right]c_k \tag{5}$$

Here we consider a non-cooperative game (Stackelberg game). It needs to be pointed that the model assumes a common knowledge framework in which all information is known to both agents. This assumption is because of the nature of Stackelberg game. However, the problem here is a little different from the standard solution. We only want to know what kind of wage structure can lead the agent to take the objective action plan $\langle i_k^j \rangle$, such as the innovative action plan $\langle 2_1^2 \rangle$ or conventional action plan $\langle 1_1^1 \rangle$.

We say that contract $\langle \vec{\lambda}, \vec{w} \rangle$ is an optimal contract that implements action plan $\langle i_k^j \rangle$ if it minimizes the total expected payments from the principal to the agent,

$$W\left(\vec{\lambda},\vec{w},\langle i_k^j\rangle\right) \tag{6}$$

subjected to the incentive compatibility constraints,

$$W\left(\vec{\lambda},\vec{w},\langle i_k^j\rangle\right)-C\left(\langle i_k^j\rangle\right)\geq W\left(\vec{\lambda},\vec{w},\langle l_n^m\rangle\right)-C\left(\langle l_n^m\rangle\right)\quad(IC_{\langle l_n^m\rangle})$$

This is a linear program with six unknowns and 27 constraints because $l,m,n\in I$. When more than one contract solves this program, we restrict attention to the contract that pays the agent earlier as Manso(2011).

The principal's expected profit from implementing action plan $\langle i_k^j\rangle$ is given by

$$\Pi\left(\langle i_k^j\rangle\right)=Y\left(\langle i_k^j\rangle\right)-W\left(\vec{\lambda}\left(\langle i_k^j\rangle\right),\vec{w}\left(\langle i_k^j\rangle\right),\langle i_k^j\rangle\right)\quad(7)$$

Where

$$Y\left(\langle i_k^j\rangle\right)=E[R_{1i}]+E\{1_{R_{1i}>B_1}E[R_{2j}\,|\,R_{1i}>B_1]\}+E\{1_{R_{1i}>B_1}E[R_{2k}\,|\,R_{1i}\leq B_1]\}\quad(8)$$

is the principal's total expected revenue when the agent uses action plan $\langle i_k^j\rangle$ and $\langle\vec{\lambda}\left(\langle i_k^j\rangle\right),\vec{w}\left(\langle i_k^j\rangle\right)\rangle$ is the optimal contract that implements action plan $\langle i_k^j\rangle$ .The principal thus chooses the action plan $\langle i_k^j\rangle$ that maximizes $\Pi\left(\langle i_k^j\rangle\right)$.

The assumptions in the principal-agent problem studied here are standard except that there is learning about the technology being employed. This gives rise to the tension between exploration and exploitation, since there is nothing to be learned about the conventional technology, but a lot to be learned about the new technology.

IV. Incentives for Exploration and Exploitation

Here we presents the optimal contracts that implement exploration and exploitation.

IV.1. Incentives for Exploitation

Recall from Section II that exploitation represented by action plan $\langle 1_1^1\rangle$.

$$\begin{aligned}W\left(\vec{\lambda},\vec{w},\langle 1_1^1\rangle\right)&=\lambda_1 op_{11}+w_1\\&+E\left\{1_{R_{11}>B_1}(\lambda_2\overline{op_{21}^{11}}+w_2)\right\}\\&+E\left\{1_{R_{11}\leq B_1}(\lambda_3\underline{op_{21}^{11}}+w_3)\right\}\end{aligned}\quad(9)$$

Given the goal of action plan $\langle 1_1^1\rangle$, the principal must offer the optimal contracts that the agent implement the exploitation. The optimal contracts $\langle\vec{\lambda},\vec{w}\rangle$ must maximizes $\Pi\left(\langle 1_1^1\rangle\right)$, namely

$$\text{minimizes}\quad W\left(\vec{\lambda},\vec{w},\langle 1_1^1\rangle\right)$$

subject to the incentive compatibility constraints,

$$W\left(\vec{\lambda},\vec{w},\langle 1_1^1\rangle\right)-C\left(\langle 1_1^1\rangle\right)\geq W\left(\vec{\lambda},\vec{w},\langle l_n^m\rangle\right)-C\left(\langle l_n^m\rangle\right)\qquad(IC_{\langle l_n^m\rangle})$$

Now we derive the optimal contract that implements exploitation. The following definitions will be useful in stating Proposition 1:

$$\beta_0=\frac{1}{1+E\left[1_{R_2>B_1}\right]}\left(\frac{E[1_{R_2>B_1}(\overline{op_{22}^{12}}-op_{20})]}{op_{21}-op_{20}}+\frac{op_{12}-op_{10}}{op_{11}-op_{10}}\right)$$

Because the distribution of return R_2 in the first period is unknown, so we use expectation to $E[1_{R_2>B_1}]$ to denote it. And we denote $p_0=E[1_{R_0>B_1}]$, $p_1=E[1_{R_1>B_1}]$ directly.

PROPOSITION 1: The optimal contract $\langle\vec{\lambda},\vec{w}\rangle_1^*$ that implements exploitation is such that

$$w_1=w_2=w_3=0,\ \lambda_2=\lambda_3=\frac{c_1}{op_{21}-op_{20}}$$

$$\lambda_1=\frac{c_1}{op_{11}-op_{10}}+\frac{\left(1+E\left[1_{R_2>B_1}\right]\right)c_1}{op_{11}-op_{12}}\left(\beta_0-\frac{c_2}{c_1}\right)^+$$

where $(x)^+=\max(x,0)$.

The formal proofs of all the propositions are omitted limited to the length. However, the main intuition behind Proposition 1 is as follows. To implement exploitation, the principal must prevent the agent from both shirking and exploring. If c_2 is high relative to c_1, only shirking constraints are binding, and thus the optimal contract that implements exploitation is similar to the optimal contract used to induce the agent to exert effort in a standard word-shirk principal-agent model. If c_2 is low relative to c_1, the exploration constraint is binding. To prevent exploration, the principal must pay the agent an extra premium in the case of success in the first period. This extra premium is decreasing in c_2 / c_1, since as c_2 / c_1 increases the agent becomes less inclined to explore.

Similarly, the baseline B_t will affect the result. If $B_1 \geq B_2$, then $\lambda_1 \geq \lambda_2 = \lambda_3$. This can be interpret as that when the baseline of standard for success decreases, the difficulty for success in second period decreases, the exploration constraint may be binding. To prevent exploration, the principal must pay the agent an extra premium in the case of success in the first period. However, if $B_1 < B_2$, the difficulty for success in second period increases, the exploitation constraint may be binding, the principal may not need to pay the agent an extra premium in the case of success in the first period. It means that the following $\lambda_1 < \lambda_2 = \lambda_3$ may be hold in this time.

To encourage the agent to take the conventional method, there are no any fixed minimum wages. This means that no any failure is tolerated in the whole process.

IV.2. Incentives for Exploration

Proposition 2 derives the optimal contract that implements exploration. Recall from Section II that exploration is given by action plan $\langle 2_1^2 \rangle$.

$$W\left(\vec{\lambda}, \vec{w}, \langle 2_1^2 \rangle\right) = \lambda_2 op_{12} + w_1 + E\left\{1_{R_{12} > B_1}(\lambda_2 \overline{op_{22}^{12}} + w_2)\right\} + E\left\{1_{R_{12} \leq B_1}(\lambda_3 \underline{op_{21}^{12}} + w_3)\right\} \quad (10)$$

Given the goal of action plan $\langle 2_1^2 \rangle$, the principal must offer the optimal contracts that implement the exploration. The optimal contracts $\langle \vec{\lambda}, \vec{w} \rangle$ must maximizes $\prod\left(\langle 2_1^2 \rangle\right)$, namely

$$\text{minimizes } W\left(\vec{\lambda}, \vec{w}, \langle 2_1^2 \rangle\right)$$

subject to the incentive compatibility constraints,

$$W\left(\vec{\lambda}, \vec{w}, \langle 2_1^2 \rangle\right) - C\left(\langle 2_1^2 \rangle\right) \geq W\left(\vec{\lambda}, \vec{w}, \langle l_n^m \rangle\right) - C\left(\langle l_n^m \rangle\right) \qquad (IC_{\langle l_n^m \rangle})$$

The form of the optimal contract that implements exploration will depend on whether exploration is moderate or radical.

DEFINITION 1: Exploration is radical if

$$\frac{E\left[1_{R_2 \leq B_1}\right]}{E\left[1_{R_1 \leq B_1}\right]} \geq \frac{E\left[1_{R_2 > B_1} \overline{op_{22}^{12}}\right]}{E\left[1_{R_1 > B_1}\right] op_{21}}$$

and moderate otherwise.

Exploration is radical if the likelihood ratio between exploration and exploitation of a failure in the first period is greater than the reward ratio between exploration and exploitation of two consecutive successes. We call this exploration radical because it has a high expected probability of failure in the first period relative to the probability of failure of the conventional action.

The following definitions will also be useful in stating Proposition 2:

$$\beta_1 = \frac{E[1_{R_2 > B_1}(\overline{op_{22}^{12}} - op_{20})]}{(1 + E[1_{R_2 > B_1}])(op_{21} - op_{20})}$$

$$\beta_2 = \beta_1 + \frac{1}{1 + E\left[1_{R_2 > B_1}\right]} \frac{E\left[1_{R_2 > B_1} \overline{op_{22}^{12}}\right] - p_0 op_{21}}{(p_1 - p_0) op_{21}}$$

PROPOSITION 2: The optimal contract $\langle \vec{\lambda}, \vec{w} \rangle_2^*$ that implements exploration is such that

$$\lambda_1 = 0, \; \lambda_3 = \frac{c_1}{op_{21} - op_{20}} \text{ and } w_2 = w_3 = 0$$

If exploration is moderate, then $w_1 = 0$ and

$$\lambda_2=\frac{c_1}{op_{21}-op_{20}}-\frac{(1+E[1_{R_2>B_1}])c_1}{E\left[1_{R_2>B_1}\overline{op_{22}^{12}}\right]-p_0op_{20}}\left(\beta_1-\frac{c_2}{c_1}\right)^+$$
$$+\frac{(1+E[1_{R_2>B_1}])c_1}{E\left[1_{R_2>B_1}\overline{op_{22}^{12}}\right]-p_0op_{21}}\left(\frac{c_2}{c_1}-\beta_1\right)^+ +$$
$$\frac{(1+E[1_{R_2>B_1}])(p_1-p_0)op_{21}c_1}{\left(E\left[1_{R_2>B_1}\overline{op_{22}^{12}}\right]-p_1op_{21}\right)\left(E\left[1_{R_2>B_1}\overline{op_{22}^{12}}\right]-p_0op_{21}\right)}\left(\frac{c_2}{c_1}-\beta_2\right)^+$$

If exploration is radical, then

$$w_1=\frac{c_1(1+E[1_{R_2>B_1}])op_{21}}{E\left[1_{R_2>B_1}(\overline{op_{22}^{12}}-op_{21})\right]}\left(\frac{c_2}{c_1}-\beta_2\right)^+$$

And

$$\lambda_2=\frac{c_1}{op_{21}-op_{20}}-\frac{(1+E[1_{R_2>B_1}])c_1}{E\left[1_{R_2>B_1}\overline{op_{22}^{12}}\right]-p_0op_{20}}\left(\beta_1-\frac{c_2}{c_1}\right)^+$$
$$+\frac{(1+E[1_{R_2>B_1}])c_1}{E\left[1_{R_2>B_1}\overline{op_{22}^{12}}\right]-p_0op_{21}}\left(\frac{c_2}{c_1}-\beta_1\right)^+ +\left(\frac{c_2}{c_1}-\beta_2\right)^+ *$$
$$\frac{(1+E[1_{R_2>B_1}])(E[1_{R_2>B_1}]-p_0)op_{21}c_1}{E\left[1_{R_2>B_1}(\overline{op_{22}^{12}}-op_{21})\right]\left(E\left[1_{R_2>B_1}\overline{op_{22}^{12}}\right]-p_0op_{21}\right)}$$

To implement exploration, the principal must prevent the agent from shirking or exploiting. The principal does not make payments to the agent after a failure in the second period, since this only gives incentives for the agent to shirk. Moreover, the principal does not make payments to the agent after a success in the first period for two reasons. First, rewarding first-period success gives the agent incentives to employ the conventional work method in the first period, since the initial expected probability $E[p_2]$ of success with the new work method is lower than the probability p_1 of success with the conventional work method. Second, in the case of success in the first period, additional information about the first-period action is provided by the second-period performance, since the expected probability of success with the new work method in the second period depends on the action taken by the agent in the first period. Delaying compensation to obtain this additional information is thus optimal.

Anyway, the principal expect the agent choose conventional work method in the second period after a failure in the first period. To prevent the agent from shirking in this situation, the principal pays the agent $\lambda_3=\frac{c_1}{op_{21}-op_{20}}$.

Then, at last, to encourage exploration the principal must reward the agent second-period success after a success in the first period. The wage rate λ_2 depends on the difficulty of implement exploration relative to exploitation. With the increase of c_2/c_1, the difficulty of implement exploration relative to exploitation increases, and wage rate λ_2 must increase, too.

If $c_2/c_1<\beta_1$, then exploitation is too costly for the agent, but exploration is not costly for the agent. At this situation, the principal pays the agent $\lambda_2<\lambda_3$. If $c_2/c_1\geq\beta_1$, then exploitation is not too costly for the agent, but exploration is costly for the agent. At this situation, the principal must pays the agent $\lambda_2\geq\lambda_3$. When $c_2/c_1\geq\beta_2$, the rage rate λ_2 must increase further. At this case, if $\frac{E[1_{R_2\leq B_1}]}{E[1_{R_1\leq B_1}]}\geq\frac{E\left[1_{R_2>B_1}\overline{op_{22}^{12}}\right]}{E[1_{R_1>B_1}]op_{21}}$, namely Exploration is radical, it has a high expected probability of failure in the first period relative to the probability of failure of the conventional action, expected reward for exploration of two consecutive successes can not compensate the risk of failure. So, the principal must pay the agent a higher λ_2, and reward the agent for failure in the first period at the same time.

Similarly, the baseline B_t will affect the result. If $B_2\geq B_1$, then λ_3 and λ_2 increase. This can be interpret as that when the baseline of standard for success increases, the difficulty for success in second period increases, the exploitation constraint may be binding. To prevent exploitation, the principal must pay the agent an extra premium in the case of success in the second period.

To illustrate the differences of the optimal contracts between these two action plan, see figure 2 as follows.

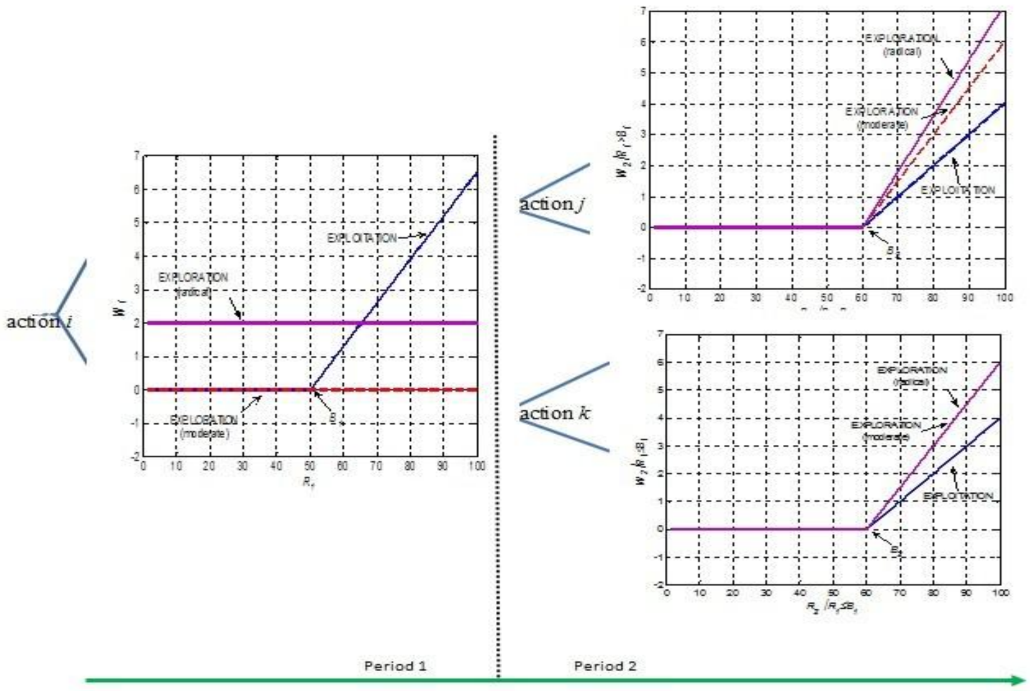

Figure 2. structured reward of action plan $\langle i_k^j \rangle$

The blue line for action plan $\langle 1_1^1 \rangle$ (exploitation), the red dashed line for action plan $\langle 2_1^2 \rangle$ (moderate exploration)

the purple line for action plan $\langle 2_1^2 \rangle$ (radical exploration)

V. Conclusion and limitations

Based on the framework of Manso(2011), this paper study the incentives for innovation with non-fixed reward for the agent. We give the standard of success, and the reward of the agent depends on the amount of the excess output over the baseline. The fixed wage and wage rate for success are designed. These structured incentives can motivate the agent select a more innovative work method and stimulate the agent to exert effort to get a better output.

The optimal contract that implements both exploitation and exploration comprise of a series of options, which are structured. To stimulate exploration, the principal must offer a proper fixed reward to tolerate the possibility of failure; at the same time, the non-fixed reward must not be offered. The optimal contract depends on the baseline of success and the private cost of the agent, especially for the cost ratio of exploration and exploitation.

There are some limitations for the paper. 1). we only consider the first-order stochastic dominances relationship between the returns. They are may be second-order or higher-order. So, more real distributions need to discuss on the problem.

2)In the paper, the information is assumed symmetry . In fact, the information may be asymmetry, which will impact the results severely.

3) the interest rate and time preference are not considered. The span of periods may have important impact on the solutions.

4)Some of the predictions of the model remain untested though, and additional empirical work seems wanted.

Acknowledgements

This work is supported by National Natural Science Foundation of China (NSFC), Grant No.71171095 and Chinese Scholarship Council (grant no.2013).

References

[1] Aghion, Philippe, and Jean Tirole, 1994, The management of innovation, *Quarterly Journal of Economics* 109, 1185–1209.
[2] Arrow, Kenneth, 1969, Classificatory notes on the production and diffusion of knowledge, *American Economic Review* 59, 29–35.
[3] Berle, Adolph, and Gardiner Means, 1932, The Modern Corporation and Private Property (Macmillan, New York, NY).
[4] Ederer, Florian, and Gustavo Manso, 2010, Is pay-for-performance detrimental to innovation? Working paper, MIT.

[5] Harris, Milton, and Artur Raviv, 1978, Some results on incentive contracts with applications to education, insurance, and law enforcement, *American Economic Review* 68, 20–30.
[6] Hellmann, Thomas, and Veikko Thiele, 2008, Incentives and innovation inside firms: A multi-tasking approach, Working paper, University of British Columbia.
[7] Holmstrom, Bengt, 1979, Moral hazard and observability, *Bell Journal of Economics* 10, 74–91.
[8] Holmstrom, Bengt, 1989, Agency costs and innovation, Journal of Economic Behavior and Organization 12, 305–327.
[9] Manso, Gustavo, 2011, Motivating Innovation, *Journal of Finance vol.* 66, pp. 1823-1869.
[10] March, James, 1991, Exploration and exploitation in organizational learning, *Organization Science* 2, 71–87.
[11] Moscarini, Giuseppe, and Lones Smith, 2001, The optimal level of experimentation, *Econometrica* 69, 1629–1644.
[12] Tian, Xuan, and Tracy Wang, 2010, Tolerance for failure and corporate innovation, Working paper, Indiana University.
[13] Zheng Chengli and Yan Chen, 2012, Motivating innovation with a structured incentives scheme – Extending the Manso model, working paper, Huazhong Normal University.