INCENTIVES FOR BOUNDEDLY RATIONAL AGENTS

Suren Basov

Department of Economics

The University of Melbourne,

Victoria 3010, Australia

e-mail: s.basov@unimelb.edu.au

Tel. (61 3)-8344-9733

Fax (61 3)-8344-6899

Acknowledgments: I am grateful to Adam Brandenburger, Hsueh-Ling Huynh, Dilip Mookherjee, Robert W. Rosenthal, and Marc Rysman for helpful comments and discussion. None of them is responsible for any mistakes in the paper.

Keywords: bounded rationality, incentives, principal-agent model.

JEL classification numbers: C60, D82, D83

Abstract

This paper develops a theoretical framework for analyzing incentive schemes under bounded rationality. It starts from a standard principal-agent model and then superimposes an assumption of boundedly rational behavior on the part of the agent. Boundedly rational behavior is modeled as an explicit optimization procedure which combines gradient dynamics with a specific form of social learning called imitation of scope. The results predict the underprovision of optimal incentives and deviation from a standard sufficient statistics result from the agency literature. It also allows us to address the question of creating the optimal incentives in a multicultural environment.

1 INTRODUCTION

Agency relationships form an important part of economic life. Among the most common examples are managers acting on behalf of an owner, workers supplying labor to a firm, and customers buying coverage from an insurance company. The common feature of all these examples is that *unobservable* actions undertaken by one party have payoff relevant consequences for another. This creates a *moral hazard problem*. The early papers that incorporated agency relationships into formal economic models were [6], [10-12], [14].

Currently, there exists a big body of literature devoted to optimal contract design under moral hazard. Although some features of real life contracts are captured by this literature, several important deviations are worth mentioning. First, most real life contracts are incomplete. They neither specify actions at each possible contingency nor use all available information. Second, incentives in the real world are too low powered [8]. Third, insufficient noise filtering occurs [3], that is the basic sufficient statistics result ([4], [6], [7], [13]) is violated.

The conventional model is build on the assumptions of full rationality. In this paper I replace this assumption by an assumption of boundedly rational behavior. Introducing bounded rationality into a model of optimal incentives is important for being able properly to address issues of low-powered incentives, insufficient noise filtering, incompleteness of contracts, and robustness of optimal incentive schemes. As Hart and Holmström write: *one will have to go outside the Bayesian framework and introduce bounded rationality in order to capture the true sense in which incentive schemes need to be robust in the real world* [5]. The model developed in this paper allows us to deal with the first two of these problems. Even though the issues of incompleteness and robustness are not addresses here, I believe that the model developed here can be considered as a first step to addressing these issues as well. Some hints along these lines are made in the paper.

Boundedly rational agents have to learn the optimal behavior. Learning is often a social process. It can be considered as a stochastic rule for adjustment of the current behavior that utilizes social information. In expectation an agent adjusts her choice in the direction of utility increase proportionally to the utility gradient. The agents using such a rule can be considered to be procedurally rational; under some conditions the process may converge to the rational outcome. The stochastic component of the rule has the virtue of reducing the probability of getting stuck at a local maximum at a generic environment. Endogenising the random component by making it a function of others' actions can facilitate convergence. The endogenising is achieved by making agents experiment more aggressively when there is less consensus in the population about the optimal course of action. I will call this type of behavior *imitation of scope*.

I find that when a principal knows that a population of agents behaves according to such a process and plans accordingly, the contracts she offers can be very different from those of standard models, and the resulting behavior of both principal and agents can be more realistic than is depicted in standard models.

The model also allows us to discuss the optimal incentive provision in a multicultural environment. Even though the model of bounded rationality developed in this paper may seem rather specific, it turns out to be a special case of a general social learning rule, studied in Basov [2]. Hence, its predictions hold generically.

The rest of the paper is organized as follows: Section 2 reviews a standard principal-agent model, Section 3 considers agents following gradient dynamics (an example of a learning rule). The central part of the paper is Section 4, where a social learning rule is proposed for a population of agents, and the main results about the nature of contract and the resulting agents behavior are presented. Section 5 applies the model of the previous section to address a problem of optimal incentives in a multicultural environment. I revisit the sufficient statistic result in Section 6, and conclude in Section 7.

2 A SIMPLE PRINCIPAL-AGENT MODEL

In this section I will consider a simple conventional principal-agent model. Let the gross profit of the principal be given by

$$\Pi = z + \varepsilon, \tag{1}$$

where z is effort undertaken by the agent, and ε is random noise with zero mean and variance σ^2. Only Π is observable by the principal and verifiable by both parties. The utility of the agent is given by:

$$U = E(w) - \frac{\phi}{2} Var(w) - \frac{z^2}{2}, \tag{2}$$

where w is the agent's payment (wage) conditioned on z through Π. The principal wants to maximize expected profits net of the wage, subject to the incentive compatibility constraint:

$$z \in \arg\max(E(w) - \frac{\phi}{2}Var(w) - \frac{z^2}{2}) \tag{3}$$

and the individual rationality (participation) constraint:

$$E(w) - \frac{\phi}{2}Var(w) - \frac{z^2}{2} \geq 0. \tag{4}$$

I will concentrate attention on affine payment schemes:

$$w = \alpha\Pi + \beta. \tag{5}$$

It is straightforward to show that the optimal affine contract has:

$$\alpha = \frac{1}{1+\phi\sigma^2}, \qquad \beta = \frac{\phi\sigma^2 - 1}{2(1+\phi\sigma^2)^2}. \tag{6}$$

To see this, note that α is chosen to maximize a total surplus W defined as

$$W = E(U + \Pi - w), \tag{7}$$

subject to (3), and β is chosen to insure that (4) holds. Since in this case the objective function of the agent is strictly concave, the incentive constraint can be replaced by the first order condition $z = \alpha$. Plugging this into (7), solving the maximization program, and using (4) to obtain β, yields (6).

The net profit of the principal and the utility of the agent under the optimal affine compensation scheme are given by:

$$E(\Pi - w) = \frac{1}{2(1+\phi\sigma^2)}, \qquad U = 0. \tag{8}$$

One can see that the slope α of the optimal compensation scheme and the profit of the principal are decreasing in σ, while the utility of the agent is determined by the reservation utility, which is normalized at zero here. Hence, noise damps incentives and dissipates social surplus.

3 OPTIMAL INCENTIVES UNDER GRADIENT DYNAMICS

In this section I assume that a single agent is boundedly rational and adjusts her effort in the direction of increasing payoff in proportion to the value of the derivative of her utility. Let the general structure of the model be the same as in Section 2. As there, restrict attention to affine compensation schemes. The main difference between the model in this section and that of Section 2 is that the agent, instead of responding optimally to the compensation scheme, adjusts her choices according to the differential equation:

$$\frac{dz}{dt} = \alpha(t) - z, \tag{9}$$

conditionally on the decision to participate. The function on the right hand side of (9) is the derivative of the agent's utility with respect to z. This is the reason I call equation (9) the gradient dynamics. It is worth mentioning that units of utility have meaning in this framework since they determine the speed of adaptation. This contrasts with the rational paradigm where utility units are arbitrary.

The agent participates if her instantaneous utility at time t is nonnegative, that is if:

$$\alpha z - \frac{z^2}{2} - \frac{1}{2}\phi\alpha^2\sigma^2 + \beta \geq 0. \tag{10}$$

(To obtain the left-hand side of (10), plug the compensation scheme (5) into the agent's utility function (2) and use the assumed expectation and variance.)

The principal seeks to maximize the discounted expected present value of net profits, subject to (9) and (10). Solving (10) with equality for β, one gets the following optimal control problem for the principal:

$$Max \int_{0}^{\infty} e^{-\rho t}(z - \frac{z^2}{2} - \frac{\phi[\alpha(t)]^2\sigma^2}{2})dt \tag{11}$$

subject to (9). The integrand is the discounted total surplus.

The present-value Hamiltonian for this problem has the form:

$$H = z - \frac{z^2}{2} - \frac{\phi\alpha^2\sigma^2}{2} + \mu(\alpha - z). \tag{12}$$

The evolution of the costate variable μ is governed by:

$$\frac{d\mu}{dt} = (1+\rho)\mu - 1 + z, \tag{13}$$

together with the transversality condition:

$$\lim_{t\to\infty} \mu(t)e^{-\rho t} = 0. \tag{14}$$

The time discount rate ρ will be assumed to be small, in fact I will fix it at zero after using the transversality condition. This is to make the comparison of the steady state with the outcome of the static model meaningful. (In the opposite case with a large discount rate the solution is trivial: take the level of effort as given and pay the constant wage that ensures that the participation constraint is satisfied. Indeed, since the principal does not care about the future and since current effort is given, the only problem is one of optimal risk sharing, which implies the above outcome since the principal is risk neutral and the agent risk averse. Admittedly, this case is not very interesting.)

The maximum principle states that the present-value Hamiltonian should be maximized with respect to α, which implies

$$\alpha(t) = \frac{\mu(t)}{\phi\sigma^2}. \tag{15}$$

Combining (15) with (13)-(14) gives the complete system characterizing the optimal incentive schedule:

$$\frac{d\alpha}{dt} = (1+\rho)\alpha + \frac{z}{\phi\sigma^2} - \frac{1}{\phi\sigma^2}, \tag{16}$$

$$\frac{dz}{dt} = \alpha - z, \tag{17}$$

$$z(0) = z_0, \ \lim_{t\to\infty} \alpha(t)e^{-\rho t} = 0, \tag{18}$$

where z_0 is the initial effort exerted by the agent.

Define γ by the expression

$$\gamma = \frac{\sqrt{(\rho+2)^2 + \frac{4}{\phi\sigma^2}} - \rho}{2}. \tag{19}$$

Then the only solution to the system (16)-(18) is:

$$\alpha(t) = \frac{1}{1+\phi\sigma^2(1+\rho)} + (z_0 - \frac{1}{1+\phi\sigma^2(1+\rho)})(1-\gamma)e^{-\gamma t} \tag{20}$$

$$z(t) = \frac{1}{1+\phi\sigma^2(1+\rho)} + (z_0 - \frac{1}{1+\phi\sigma^2(1+\rho)})e^{-\gamma t}. \tag{21}$$

Two things are worth mentioning here. First, the slope of the compensation scheme $\alpha(t)$ converges to a stationary value, which coincides with the slope of the optimal compensation scheme for a rational agent in the static model if $\rho = 0$. Second, on the dynamic path $z(t) > \alpha(t)$ provided that z_0 is greater than the steady-state level; that is, the agent exerts more effort than would be myopically optimal, though the difference shrinks in time.

It is straightforward to show that the present value of expected profits of the principal as a function of the initial conditions has the form:

$$\Pi(z_0) = Az_0^2 + Bz_0 + C, \tag{22}$$

with $A < 0$ and $C < 0$. (One has simply to plug (20)-(21) into (11) and carry out the integration.) This implies that profits would be negative if initial effort is too high. In that case the principal would prefer to stay out of business. This makes intuitive sense since the principal has to compensate the agent for wasted effort due to the participation constraint.

Note that the model of this section provides an additional rationale for simple linear schemes. Under suitable assumptions such schemes provide the agent with a strictly concave objective function. Agents who search myopically would eventually learn their globally maximizing effort level independently of the initial effort level. Since this level of effort maximizes the total surplus and the agent's utility is fixed by the participation constraint, the principal wants the agent to learn the global maximizer. With nonlinear compensation schemes, the agent's objective function need not be concave and hence may have local maxima which are not the global maxima. In that case, what the agent learns might depend on the initial effort level.

4 OPTIMAL INCENTIVES WHEN AGENTS ALSO IMITATE

In this section I consider the principal's problem of designing an optimal compensation scheme when a population of agents is engaged in a social adaptation process. Assume that there is a continuum of identical agents working for the same principal. Assume that the principal can pay a wage based only on the output produced by an agent but not on relative performance, and is bound to pay different agents the same wage for the same performance. Each agent chooses effort x from the interval $[c, d]$ $(\infty > d > c > 0)$ at each point in time.[1] As in Section 2, the instantaneous payoff to the agent from choosing x at t is given by the function:

$$U(x,t) = \alpha(t)x - \frac{x^2}{2} + const. \tag{23}$$

The constant term in (23) does not depend on x but may depend on α.

Each agent starts at some exogenously specified effort level $x_0 \in \Omega$ and adjusts it at times $k\Delta t$, where k ranges over the natural numbers and Δt

[1] The symbol z will now be used for mean effort level in the agent population.

is some fixed time interval length. To describe the adaptation rule it is necessary to specify the information available to the agent at the moment of adaptation and the rules of information processing.

I will assume that each agent knows the gradient (i.e. the derivative) of the utility function at the point of her current choice, the population mean z and population variance v of choices. The utility function is assumed to be continuously differentiable and globally concave. The agent's choices follow a stochastic differential equation:

$$x_{t+\Delta t} = (1 - \gamma(\Delta t))(x_t + \frac{\partial U(x_t)}{\partial x}\Delta t) + \gamma(\Delta t)(y_t - x_t). \qquad (24)$$

The first term represents gradient dynamics, while the second term represents imitation. In the above formula the imitation weight $\gamma(\Delta t)$ is assumed to be a nondegenerate random variable with compact support, and such that $E(\gamma(\Delta t)) = 0$. This implies a particular form of imitation: *imitation of scope*. Since $\gamma(\Delta t)$ assumes both positive and negative values with positive probability, the agent does not imitate directly the choice of the other agent. Instead, she opens a search window, the width of which is determined by the degree of disagreement between her current behavior and the observed choice of another agent, that is by $(y_t - x_t)$. Intuitively, since the observation the agent makes is the choice of another agent who is also boundedly rational, there is no good reason to imitate the choice directly. On the other hand, the spread of the choices in the population indicates that society as a whole does not know the optimal choice, and hence that there may be returns to experimentation. The second term in (24) embodies a simple version of this intuition: $|y_t - x_t|$ increases probabilistically in the population's spread.

It will be convenient to assume $\gamma(\Delta t) = \sqrt{\Delta t} u$, where u is a random variable, independent of x, with a compact support and such that $E(u) = 0$, and $Var(u) = 1$. To capture the intuition that it is more likely to move in the observed direction than in the opposite one, one may allow $E(\gamma(\Delta t)) = b\Delta t$ with $b > 0$. However, for a strictly concave utility function this term will be subsumed by the first one, and hence will not make any qualitative difference, while making the analysis more complicated.

Let $f(x, t)$ denote the density of the choices in the population of agents at time t. (If we normalize the mass of the population to be one, an equivalent interpretation of the function $f(x, t)$ would be the probability density of the choice of an individual at time t.) Its evolution is described by the following theorem:

Theorem 1 *Let the adaptation rule be given by (4.26). Then the continuous-*

time evolution of the density $f(x,t)$ is well-defined and is governed by:

$$\frac{\partial f}{\partial t}+\frac{\partial}{\partial x}((\alpha(t)-x)f) = \frac{1}{2}\frac{\partial^2}{\partial x^2}((x-z(t))^2+v(t)f) \tag{25}$$

$$z(t) = \int_c^d xf(x,t)dx \tag{26}$$

$$v(t) = \int_c^d (x-z(t))^2 f(x,t)dx. \tag{27}$$

Proof. Let $P(x,U,\Delta t)$ be the probability $x_{t+\Delta t}\in U$ if $x_t\in U\subset R$, where U is a Borel set. Since $U(\cdot)$ is continuously differentiable, $x_t, y_t\in[c,d]$, and u has a compact support,

$$\lim_{\Delta t\to 0}(x_{t+\Delta t}-x_t)=0$$

uniformly in (x_t,y_t,u_t), where u_t is the realization of u at time t. But then $\forall\varepsilon>0\ \exists\delta>0$ such that $0<\Delta t<\delta$ implies $|x_{t+\Delta t}-x_t|<\varepsilon$. Let U_ε^c be the complement of the $\varepsilon-$neighborhood of point x. Then $P(x,U_\varepsilon^c,\Delta t)=0$ for $\Delta t\in(0,\delta)$. This implies that

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}P(x,U_\varepsilon^c,\Delta t)=0.$$

Hence

$$P(x, U_\varepsilon^c, \Delta t) = o(\Delta t).$$

Define

$$\begin{aligned} v(x,t) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} E(x_{t+\Delta t} - x_t) \\ \Gamma(x,t) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} Var(x_{t+\Delta t} - x_t). \end{aligned}$$

Then the population density of choices is governed by an equation [9]

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial x}(fv) = \frac{1}{2}\frac{\partial^2}{\partial x^2}(\Gamma f) = 0.$$

Direct calculation proves the Theorem. ■

To make a specific prediction, the system (25)-(27) should be supplemented by initial and boundary conditions. The initial condition is arbitrary, but I will impose the following boundary condition:

$$(\alpha - x)f - \frac{1}{2}\frac{\partial}{\partial x}((x-z)^2 + v)f) = 0 \text{ for } x = c, d, \ \forall t \geq 0. \tag{28}$$

The boundary condition (28) guarantees conservation of the full probability.

An important feature of this model is the existence of special kinds of solutions: *wave packets* and *quasistationary states.* Intuitively, a wave packet is a solution to (25)-(27) in which the mean moves according to the gradient dynamics and the variance shrinks so slowly that in a first-order approximation it can be considered to be constant. As the mean approaches its steady state value under the gradient dynamics, a wave packet converges to a quasistationary state, that is to a distribution with a very slowly changing mean and variance.

To demonstrate the existence of wave packets and quasistationary states, I need to assume that the choice space is large in the sense that

$$c << z_0, \ \frac{1}{1+\phi(\sigma^2+v_0)} << d, \ c << a, \ b << d, \ \sqrt{v_0} << d-c \qquad (29)$$

Here $<<$ means "much less," z_0 and v_0 are the mean and variance of the population's effort distribution at time zero, respectively; and the numbers a and b ($a < b$) are bounds on the support of z_0 which guarantee that the expected profit of the principal at time zero is positive. The inequalities in (29) say that the initial distribution and quasistationary state are concentrated far from the boundary points, and that the principal will never force a large probability mass close to the boundary.[2]

Under (29) I can derive differential equations that govern the evolution of $z(t)$ and $v(t)$. To do this, differentiate equations (26) and (27) with respect to time, integrate by parts, and use the boundary condition (28). This yields:

$$\frac{dz}{dt} = \alpha(t) - z - \frac{1}{2}g(z,d,v)f(d,t) - g(z,c,v)f(c,t) \tag{30}$$

$$\frac{dv}{dt} = -(d-z)g(z,d,v)f(d,t) - (z-c)g(z,c,v)f(c,t) \tag{31}$$

where $g(z,\zeta,v) = (\zeta - z)^2 + v$. Under the conditions imposed on the initial density function, the boundary terms will be small. Indeed, they are small at time zero by assumption and remain small because the variance shrinks in time (due to (31)) and the principal has no incentive to push probability mass close to the boundary. Hence, the system (30)-(31) can be rewritten approximately in the form:

[2] The argument for this last point is the same as in the next-to-last paragraph of Section 3.

$$\frac{dz}{dt} = \alpha(t) - z \tag{32}$$

$$\frac{dv}{dt} = 0. \tag{33}$$

System (32)-(33) implies that the mean follows the gradient dynamics while the variance remains constant (I call the solution to this system a wave packet).

Next, to formulate the principal's problem, I will specify the participation constraint as follows. I will assume that each agent observes the variance of the current output in the population, uses it to evaluate wage variance and hence, her expected utility, and participates as long as it is greater than zero. Since an agent is assumed to be boundedly rational, she is assumed to be incapable of evaluating the expected discounted present value of the incentive scheme. However, she realizes that her instantaneous gratification is a poor indicator of her future utility stream and hence, chooses to rely on the population information when evaluating the value of the incentive scheme. This pins down $\beta(t)$. Given this, either all agents will decide to participate or all will drop out. Hence, the principal's problem is

$$Max \int_0^\infty e^{-\rho t}\left(z - \frac{z^2 + v(t)}{2} - \frac{\phi[\alpha(t)]^2(\sigma^2 + v(t))}{2}\right)dt \tag{34}$$

$$s.t. \frac{dz}{dt} = \alpha(t) - z. \tag{35}$$

The variance $v(t)$ will be taken to be constant according to the approximation (33); hence I will omit the argument t in the function $v(t)$ below. The derivation of (34) is similar to the derivation of (11). The only difference is that effort is now stochastic. This randomness is reflected in the expected cost of effort ($E(x^2) = z^2 + v$) and the variance of output, which now has two components, exogenous σ^2 and endogenous v. The solution to (34)-(35) is given by:

$$\alpha(t) = \frac{1}{1 + \phi(\sigma^2 + v)(1 + \rho)} + \left(z_0 - \frac{(1 - \delta)}{1 + \phi(\sigma^2 + v)(1 + \rho)}\right)e^{-\delta t} \tag{36}$$

$$z(t) = \frac{1}{1 + \phi(\sigma^2 + v)(1 + \rho)} + \left(z_0 - \frac{1}{1 + \phi(\sigma^2 + v)(1 + \rho)}\right)e^{-\delta t}, \tag{37}$$

where

$$\delta = \frac{\sqrt{(\rho+2)^2 + \frac{4}{\phi(\sigma^2+v)}} - \rho}{2} > 0. \tag{38}$$

From (36)-(38) one can see that the steady-state incentive $\alpha(\infty)$ is lower than under the pure gradient dynamics, and convergence takes longer. The principal's profits are given by

$$\Pi = \frac{1}{2(1+\phi(\sigma^2+v))}. \tag{39}$$

From (39) it is apparent that the steady state profits under bounded rationality are smaller than under full rationality, while agents get the same utility on average. Hence, the cost of boundedly rational behavior is completely born by the principal.

5 OPTIMAL INCENTIVES IN A MULTI-CULTURAL ENVIRONMENT

In this section I study optimal incentives in a multicultural environment. Assume there are two distinguishable groups of agents such that $f_1(x,0) \neq f_2(x,0)$. I will refer to the difference in the initial effort distributions as a difference in culture. The principal has a choice whether to pool the groups, so the members of both groups face the same wage schedule and can observe each others choices, or to keep them separate so they cannot observe each other and have different wage schedules. I will refer to the last choice as a differential treatment.

As long as there is no direct skill transmission from one group to another, the conventional model will predict that this choice makes no difference. The model of Section 3 agrees with the conventional one on this issue. However, using the model of Section 4, one can argue that, in the absence of positive productive externalities, the differential treatment is always optimal. To see why, denote the population fraction of group one by $\alpha \in (0, 1)$ and observe that under the pooling regime the agent population is characterized by the density

$$f(x, t) = \alpha f_1(x, t) + (1 - \alpha) f_2(x, t).$$

The population variance of the effort is given by

$$v = \alpha v_1 + (1 - \alpha) v_2 + \alpha(1 - \alpha)(z_1 - z_2)^2,$$

where z_i and v_i are the population's i mean and variance of effort respectively.

Hence, under the pooling regime the principal's profits satisfy

$$\Pi(v) = \frac{1}{2(1+\phi(\sigma^2+v))} < \frac{1}{2(1+\phi(\sigma^2+\alpha v_1+(1-\alpha)v_2))}.$$

But now, noting that $\Pi(\cdot)$ is convex and using Jensen's inequality,

$$\Pi(v) < \alpha\Pi(v_1) + (1-\alpha)\Pi(v_2).$$

Since the right hand side represents the profits under differential treatment the claim is proven.

The above result is very intuitive. It states, that given the cultural differences between different groups in population, it is better to capitalize on them than try to smooth them. Formula (38) implies that under the differential treatment the more uniform group will face stronger incentives. Assuming both groups have the same reservation utility, this implies that the more diverse group will recieve a higher fixed payment. These predictions distinguish the explanation of the low powered incentives based on the model of bounded rationality with individualistic preferences from the explanation based on the model of social preferences with perfect rationality. If the latter is the case, some preliminary results obtained by the author indicate that the more uniform group should face weaker incentives.

6 SUFFICIENT STATISTICS REVISITED

One of the main results of conventional contract theory is that when several signals of effort are observed, optimal contracts should be based on a sufficient statistic for effort. Since in the first best the principal should compensate the agent only for effort, it seems quite intuitive that second-best compensation is based on the best possible estimate of effort available.[3] However, this is not the case under bounded rationality. Intuitively, the reason is that under bounded rationality, effort can be only partly attributed to the incentive scheme. The other part comes from the social adaptation process.

Formally, consider a model similar to the model of the previous section but allow for two measures of output, Π_1 and Π_2. For example, Π_1 may be profit and Π_2 sales. Let them be determined as follows:

[3] This intuition is a little misleading since in the equilibrium the principal knows the effort. However, to create correct incentives, she should pretend that she does not know and behave as a statistician who tries to estimate effort from available data.

$$\Pi_1 = x + \varepsilon_1 \quad (40)$$

$$\Pi_2 = x + \varepsilon_2. \quad (41)$$

Here x is effort exerted by the agent, ε_1 and ε_2 are independent normal random variables with zero means and variances σ_1^2 and σ_2^2, respectively. Suppose the principal is interested in Π_1 only.[4] The optimal contract under perfect rationality should be based on a sufficient statistic for x, namely on

$$\frac{\Pi_1}{\sigma_1^2} + \frac{\Pi_2}{\sigma_2^2}. \quad (42)$$

This is rather intuitive: in the optimal contract, an observation that conveys more information should be given higher weight. This is reflected by the fact that the ratio of the coefficients before Π_1 and Π_2 is σ_2^2/σ_1^2. In the case of bounded rationality, going through the same calculations as in the previous section, one can conclude that this ratio changes to $(\sigma_2^2+v)/(\sigma_1^2+v)$[5]. That

[4]Everything below will be true if the principal were interested in any convex combination of Π_1 and Π_2.

[5]To obtain this result, one has to assume that agents observe variances of both measures of output and treat them as independent when calculating the variance of the wage. In fact, the behavior of the agents creates a correlation between x_1 and x_2 (though they are independent conditional on effort) but the agents fail to understand this. If we assumed

is, the optimal contract will have the form

$$w(t) = \alpha(t)(\frac{\Pi_1}{\sigma_1^2 + v} + \frac{\Pi_2}{\sigma_2^2 + v}) + \beta(t). \tag{43}$$

The optimal slope α will be given by

$$\alpha(t) = \frac{1}{1 + \phi\zeta(1+\rho)} + (z_0 - \frac{(1-\eta)}{1+\phi\zeta(1+\rho)})e^{-\eta t} \tag{44}$$

where

$$\zeta = \frac{(\sigma_1^2 + \sigma_2^2 + v)}{(\sigma_1^2 + v)(\sigma_2^2 + v)}, \; \eta = \frac{\sqrt{(\rho+2)^2 + \frac{4}{\phi\zeta}} - \rho}{2}. \tag{45}$$

The intercept β is given by the participation constraint. Its exact value is not interesting here. (To obtain all these results, one needs to go through calculations similar to those of the previous two sections.)

It is worth mentioning that the ratio $(\sigma_2^2 + v)/(\sigma_1^2 + v)$ is less sensitive to changes in the variances of exogenous noise than the ratio σ_2^2/σ_1^2. If social noise dominates technological uncertainty, that is $v >> \max(\sigma_1^2, \sigma_2^2)$, then (4.49) can be rewritten approximately as

---

instead that agents observe wage variance directly, the results would change quantitatively but not qualitatively.

$$w(t) = \frac{2\alpha(t)}{v}(\frac{\Pi_1 + \Pi_2}{2}) + \beta(t). \quad (46)$$

$$w(t) = \frac{2\alpha(t)}{v}(\frac{\Pi_1 + \Pi_2}{2}) + \beta(t). \quad (47)$$

In this case, the agents' payments depend on the arithmetic mean of the signals. The intuitive reason for this is quite straightforward: the principal wants to pay only for the part of effort that responds to the incentive scheme. To do so, she must filter out both technological uncertainly and social noise. Since social noise is common to both signals, taking the arithmetic mean is the best way to filter it. When social noise is much greater than technological uncertainty, the issue of filtering out social noise dominates the issue of filtering out technological uncertainty.

It is interesting to examine some empirical evidence in the light of this result. Bertrand and Mullainathan [3] found that the actual payment does not filter out all technological uncertainty in a variety of contexts. For example, compensation of managers in the petroleum industry responded to changes in profits in the same way, regardless of whether this change came from better marketing or from changes in the price of oil.

7 CONCLUSION

This paper develops a theoretical framework for analyzing incentive schemes when agents behave in a boundedly rational manner.

Underprovision of optimal incentives, deviations from the sufficient statistics result in the direction of more equal weights on the performance measures, and gift-exchange behavior are all ubiquitous in agency relationships in the real world. The first two phenomena are directly predicted by the model studied here.

Furthermore, under the behavioral assumption that replaces rationality in the model of Section 4, the agents' choices follow some continuous stochastic process. The stochastic component of the adaptation rule is determined by a social adaptation rule. The important property of this adaptation rule is that it allows for endogenous variance in the distribution of choices in the steady state. This in turn leads to a dissipation of the social surplus and the principal's profits, and it allows us to study the optimal incentives in a multi-cultural environment.

Even though gift-exchange behavior is not a direct consequence of the model, the model helps to shed light on its prevalence. The fact that under social adaptation, incentive contracts become less attractive, confirms the intuition expressed, for example, in [1] that gift-exchange is a result of social interaction. I study it in more detail in Basov [2].

Even though a lot of questions (for example, the robustness of optimal incentive schemes) remain unsolved in this paper, it can be viewed as a first step in incorporating bounded rationality into incentive problems.

REFERENCES

1. G. A. Akerlof, Labor Contracts as Partial Gift Exchange, *Quarterly Journal of Economics*, 97, (1982), 543-569.

2. S. Basov, "Bounded Rationality, Reciprocity, and Their Economic Consequences," Ph.D. Thesis, Boston University, Boston, May 2001.

3. M. Bertrand, M. and S. Mullainathan "Are CEOs Rewarded for Luck? A Test of Performance Filtering," NBER Working Paper 7604, March 2000.

4. M. Dewartripont, I. Jewitt, and J. Tirole, The Economics of Career Concerns, Part I, *Review of Economic Studies*, 66, (1999), 183-198.

5. O. Hart and B. Holmström, The Theory of Contracts, *in* "Advances in Economic Theory, Fifth World Congress," (T. Bewley, Ed.), New York: Cambridge University Press, 1987.

6. B.Holmström, Moral Hazard and Observability, *Bell Journal of Economics*, 10, (1979), 74-91

7. B. Holmström, Moral Hazard in Teams, *Bell Journal of Economics*, 13, (1982), 324-340.

8. M. C. Jensen and K. Murphy, Performance Pay and Top-Management Incentives, *Journal of Political Economy* 98 (1990), 225-264.

9. D. Kanan, "An Introduction to Stochastic Processes," Elsevier North Holland, Inc., 1979.

10. J. Mirrlees, Notes on Welfare Economics, Information and Uncertainty, *in* "Essays in Economic Behavior Under Uncertainty," (M. Balch, D. McFadden, and S. Wu, Eds.), pp.243-258, Amsterdam: North-Holland, 1974.

11. J. Mirrlees, The Optimal Structure of Authority and Incentives Within an Organization, *Bell Journal of Economics*, 7, (1976), 105-131.

12. S. Ross, The Economic Theory of Agency: The Principal's Problem, *American Economic Review*, 63, (1973), 134-139.

13. S. Shavell, Risk Sharing and Incentives in the Principal and Agent Relationship, *Bell Journal of Economics*, 10, (1979), 55-73.

14. M. Spence and R. Zeckhauser, Insurance, Information and Individual Action, *American Economic Review: Papers and Proceedings*, 61, (1971) 380-387