Journal of Economic Theory **91**, 292–301 (2000)
doi:10.1006/jeth.1999.2623, available online at http://www.idealibrary.com on IDEAL®

Asymptotic Efficiency in Dynamic Principal-Agent Problems

Holger M. Müller[1]

Department of Economics, University of Mannheim, A5, 68131 Mannheim, Germany
hmueller@pool.uni-mannheim.de

Received March 31, 1999; final version received October 18, 1999

In a seminal paper, B. R. Holmström and P. R. Milgrom (1987, *Econometrica* **55**, 303–328) examine a principal-agent model in which the agent continuously controls the drift rate of a Brownian motion. Given a stationary environment, they show that the optimal sharing rule is a linear function of aggregated output. This paper considers a variant of the Brownian model in which control revisions take place in discrete time. It is shown that no matter how "close" discrete time is to continuous time, the first-best solution can be approximated arbitrarily closely with a random spot check and a suitably chosen sequence of step functions. *Journal of Economic Literature* Classification Numbers: D82, J33. © 2000 Academic Press

1. INTRODUCTION

In the well known principal-agent model by Holmström and Milgrom [4], an agent continuously controls the drift rate vector of a multi-dimensional Brownian motion. In this model, the optimal sharing rule is a linear function of the end-of-period values of "accounts" that list the number of times that each possible output level has occurred. If the principal cannot observe the time paths of the different accounts, but only that of some coarser linear aggregate (e.g., total output), Holmström and Milgrom moreover show that the optimal sharing rule is a linear function of the end-of-period value of this aggregate. Recently, the Brownian model has been generalized and extended by Schättler and Sung [9] and Sung [10].

[1] This is a revised version of an earlier paper entitled "Randomization in Dynamic Principal-Agent Problems," Working Papers in Economics E-96-5, Hoover Institution, which was written while I was visiting Stanford University. Financial support from Schweizerischer Nationalfonds and Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 504, is gratefully acknowledged. An earlier version of this paper was presented at the Econometric Society European Meeting 1997 in Toulouse. I have benefitted from comments and suggestions by Ralf Becker, Martin Hellwig, Paul Milgrom, Heinz Müller, Klaus Schmidt, Jae Sung, Dimitri Vayanos, and participants in seminars at the Universities of Bonn, Copenhagen, Lund, Mannheim, Munich, Stockholm, and the Stockholm School of Economics.

0022-0531/00 $35.00

This paper considers a version of the Brownian model in which control revisions take place in discrete time. It is shown that, no matter how "close" discrete time is to continuous time, linear sharing rules are not optimal as the principal can approach the first-best solution asymptotically with a random spot check and a suitably chosen sequence of step functions. To see this, suppose the relevant time interval is partitioned into T subintervals of equal length, and the agent can revise his control only at the beginning of each subinterval. Consider the following incentive scheme. Ex post, the principal randomly selects one of the T subintervals and the agent is compensated with a step function based on the output produced in this subinterval. Given a stationary environment, this incentive scheme creates uniform incentives: as the agent does not know in advance which subinterval will be selected, he optimally chooses a constant control. Accordingly, the problem can be treated as a static problem in which the agent controls the mean of a normally distributed random variable and the principal is a priori restricted to step functions. As Mirrlees [5, 6] has shown, this implies that the principal can approximate the first-best solution arbitrarily closely.

The emphasis in this paper is on random spot checks as these—like linear sharing rules—are widely used in practice. Our result that the principal can approach the first-best solution does not hinge on the use of a random sharing rule, however; the first-best solution can be approximated as well if a step function is applied to each of the T outputs separately. What makes this incentive scheme unattractive though is that it becomes prohibitively expensive if auditing is costly and the number of subintervals becomes large.

2. THE MODEL AND MAIN RESULT

The model is a one-dimensional version of the Brownian model by Holmström and Milgrom [4] in which control revisions take place in discrete time. In the time interval $[0, 1]$ the agent controls a publicly observable output process X with boundary condition $X_0 = 0$ and stochastic differential equation

$$dX_t = f(u_t)\, dt + \sigma\, dB_t, \tag{1}$$

where $f(u_t)$ is the instantaneous mean, $u_t = u_t(t, X)$ is the agent's control at date t, σ is the diffusion rate, and B is a standard Brownian motion. The control u is not observable and can only be revised at times $t = 0, 1/T, ..., (T-1)/T$, where T denotes the total number of revisions. The length of a subinterval with constant control is then $\Delta t \equiv 1/T$. Moreover, u is an $\mathcal{F}_t$-predictable process which takes values in some open bounded control set $U \subseteq \mathscr{R}_+$.[2] Denote the class of all such control processes by $\mathscr{U}$. The "production function" $f(\cdot)$ is bounded with derivatives $f'(\cdot) > 0$ and $f''(\cdot) < 0$, and the diffusion rate σ lies in a bounded subset of $\mathscr{R}_{++}$. Additionally, the agent incurs effort cost $c(u_t)$, where $c(\cdot)$ is bounded with derivatives $c'(\cdot) > 0$ and $c''(\cdot) \geqslant 0$. For convenience, we assume that the principal is risk neutral. The agent has negative exponential von Neumann–Morgenstern utility with coefficient of risk aversion r.

In principle, the sharing rule $S(\cdot)$ may depend on the entire history of X. The principal maximizes expected profits subject to two constraints: (i) the agent weakly prefers the random allocation $(S(X), u)$ to the certain income W_A, and (ii) given the sharing rule $S(X)$, the agent's optimal control is u. Formally,

$$\max_{S(X),\, u} E[X_1 - S(X)] \tag{2}$$

s.t.

$$dX_t = f(u_t)\, dt + \sigma\, dB_t, \tag{3}$$

$$E\left[-\exp\left\{-r\left(S(X) - \sum_{t=0}^{(T-1)/T} c(u_t)\, \Delta t\right)\right\}\right] \geqslant -\exp\{-rW_A\}, \tag{4}$$

and

$$u \in \arg\max_{\hat{u} \in \mathscr{U}} E\left[-\exp\left\{-r\left(S(X) - \sum_{t=0}^{(T-1)/T} c(\hat{u}_t)\, \Delta t\right)\right\}\right]. \tag{5}$$

In the absence of incentive constraints, the principal's first-best problem is characterized by (2)–(4). The following lemma follows from Müller [8].

LEMMA 1. *In the discrete-time version of the Brownian model, the first-best sharing rule is*

$$S_{FB} = W_A + c(u_{FB}), \tag{6}$$

where the first-best control $u_{tFB} = u_{FB}$ *is unique and constant over time.*

Since the cumulative output in subintervals $[t, t + \Delta t]$ is normally distributed, the Brownian model with discrete-time control revisions is conceptionally equivalent to a discrete-time model in which the agent repeatedly

[2] Any left-continuous, measurable process that is adapted to the filtration $\{\mathscr{F}_t\}$ is $\mathscr{F}_t$-predictable. See [1] for details.

chooses the mean $f(\mu_t)\,\Delta t$ of a normally distributed random variable ΔX_t. Among other things, this implies that each subinterval has a "Mirrlees structure," i.e., powerful MLRP tests can be performed in the tails of the distribution that almost perfectly reveal the agent's effort choice. In a static context, Mirrlees [5, 6] has shown that in this environment the first-best outcome can be approximated arbitrarily closely with a suitably chosen sequence of step functions. In a dynamic context, this is not necessarily true. As an illustration, consider the case where $T=2$ and suppose a step function is applied to the aggregated end-of-period output X_1. From the perspective of the second (and final) revision date $t=\frac{1}{2}$, the end-of-period output X_1 is normally distributed with mean $f(\mu_{1/2})\,\Delta t + X_{1/2}$ and variance $\sigma^2\,\Delta t$. Consequently, the agent's optimal second-period control $\mu_{1/2}$ will depend on the underlying state $X_{1/2}$. In particular, if $X_{1/2}$ is high, the agent will choose a low control, and if $X_{1/2}$ is low, he will choose a high control. In the face of this fact, Hart and Holmström [2, p. 93] conclude that "step functions will induce a path of effort that will be both erratic and, on average, low.... This suggests that the optimality of step functions is highly sensitive to the assumption that the agent chooses his labor input only once."

While in the above example step functions indeed fail to implement the (constant) first-best control, this is not because effort is chosen repeatedly, but because the agent's compensation is based exclusively on aggregated end-of-period output X_1. With a more elaborate incentive scheme based on intermediate output values, however, it is not only possible to implement any desired constant control, but also to approximate the first-best solution arbitrarily closely. To see this, consider the following incentive scheme.

DEFINITION: RANDOM SPOT CHECK. At time 1, the principal randomly selects one of the T output increments $\Delta X_t \equiv X_{t+\Delta t} - X_t$. Subsequently, the agent is compensated with a step function based on ΔX_t.

The randomizing device in question is assumed to be symmetric. Thus, to perform a random spot check the principal need only observe the value of one randomly selected output increment ΔX_t. In particular, the principal need not observe the values of the other $T-1$ increments or the time order in which these increments were generated.[3] In other words, it is not necessary that the principal knows the entire history of the output process X.

[3] In many practical examples, obtaining such information is either too costly or not possible. Consider, for instance, a quality control department where random spot checks are used to test the quality of transistors. Typically, the tester faces a pile of transistors from which he randomly selects a few candidates. Both the quality of the remaining transistors and the time order in which the transistors were generated remain unknown, however.

Given a random spot check, the agent faces the same control problem in each subinterval. Accordingly the agent's overall control problem can be expressed as a simple multivariate optimization problem. For instance, if there are only two subintervals, the agent's overall control problem is

$$\begin{aligned} \max_{u_0,\, u_{1/2}} & -\tfrac{1}{2} F(\overline{\Delta X} \mid u_0) \exp\{-r(\underline{s} - c(u_0)\, \Delta t - c(u_{1/2})\, \Delta t)\} \\ & -\tfrac{1}{2}(1 - F(\overline{\Delta X} \mid u_0)) \exp\{-r(\bar{s} - c(u_0)\, \Delta t - c(u_{1/2})\, \Delta t)\} \\ & -\tfrac{1}{2} F(\overline{\Delta X} \mid u_{1/2}) \exp\{-r(\underline{s} - c(u_0)\, \Delta t - c(u_{1/2})\, \Delta t)\} \\ & -\tfrac{1}{2}(1 - F(\overline{\Delta X} \mid u_{1/2})) \exp\{-r(\bar{s} - c(u_0)\, \Delta t - c(u_{1/2})\, \Delta t)\}, \end{aligned} \tag{7}$$

where $\overline{\Delta X}$ denotes the cutoff, $\underline{s}$ is the payment if $\Delta X_t \leqslant \overline{\Delta X}$, $\bar{s}$ is the payment if $\Delta X_t > \overline{\Delta X}$, and $F(\overline{\Delta X} \mid u_t)$ is the probability that $\Delta X_t \leqslant \overline{\Delta X}$ conditional upon the fact that u_t is chosen. In (7), the first two rows represent the agent's expected utility if the selected increment is $\Delta X_t = \Delta X_0$, weighted with the probability $\frac{1}{2}$ that the first subinterval is selected. Analogously, the last two rows represent the agent's expected utility if $\Delta X_t = \Delta X_{1/2}$, weighted with the probability $\frac{1}{2}$ that the second subinterval is selected.

Like linear sharing rules, random spot checks provide the agent with constant incentives over time. As the following lemma shows, the principal can implement any desired constant control provided that the payments $\underline{s}$ and $\bar{s}$ are chosen appropriately and the cutoff $\overline{\Delta X}$ is sufficiently small.

Lemma 2. *Let $F(\overline{\Delta X} \mid \bar{u})$ denote the value of the distribution function of a normally distributed random variable ΔX with mean $f(\bar{u})\, \Delta t$ and variance $\sigma^2\, \Delta t$ at $\Delta X = \overline{\Delta X}$. Given an arbitrary constant control $\bar{u}$, there exists a random spot check with payments*

$$\underline{s} = W_A + c(\bar{u}) - \frac{1}{r} \ln[1 - rc'(\bar{u})(1 - F(\overline{\Delta X} \mid \bar{u}))/F_u(\overline{\Delta X} \mid \bar{u})] \tag{8}$$

and

$$\bar{s} = W_A + c(\bar{u}) - \frac{1}{r} \ln[1 + rc'(\bar{u})\, F(\overline{\Delta X} \mid \bar{u})/F_u(\overline{\Delta X} \mid \bar{u})] \tag{9}$$

such that (i) *the agent chooses $u_t = \bar{u}$ in each subinterval, and* (ii) *the agent's participation constraint* (4) *holds with equality.*

Proof. Note that both payments are well defined for sufficiently small $\overline{\Delta X}$ since $F_u(\overline{\Delta X} \mid \bar{u}) < 0$ and $\lim_{\overline{\Delta X} \to -\infty} F(\overline{\Delta X} \mid \bar{u})/F_u(\overline{\Delta X} \mid \bar{u}) = 0$. Given a random spot check, the agent's control problem is

$$\max_{u_0, \ldots, u_{(T-1)/T}} -\frac{1}{T} \sum_{t=0}^{(T-1)/T} F(\overline{\Delta X} \mid u_t) \exp\left\{-r\left(\underline{s} - \sum_{t=0}^{(T-1)/T} c(u_t)\, \Delta t\right)\right\}$$

$$-\frac{1}{T} \sum_{t=0}^{(T-1)/T} (1 - F(\overline{\Delta X} \mid u_t)) \exp\left\{-r\left(\bar{s} - \sum_{t=0}^{(T-1)/T} c(u_t)\, \Delta t\right)\right\}. \quad (10)$$

As is shown in the working paper version [7], the agent's objective function is strictly concave in u_t if $\overline{\Delta X}$ is sufficiently small. Differentiating (10) with respect to u_t and rearranging yields the set of first-order conditions

$$\frac{F_u(\overline{\Delta X} \mid u_t) + rc'(u_t)\, \Delta t \sum_{t=0}^{(T-1)/T} F(\overline{\Delta X} \mid u_t)}{F_u(\overline{\Delta X} \mid u_t) - rc'(u_t)\, \Delta t \sum_{t=0}^{(T-1)/T} (1 - F(\overline{\Delta X} \mid u_t))} = \exp\{-r(\bar{s} - \underline{s})\}. \quad (11)$$

By symmetry, $u_t = u$ for all t, i.e., the agent chooses the same control in each subinterval. Inserting (8)–(9) in (11) shows that the agent's first-order condition is satisfied at $u = \bar{u}$. The agent's participation constraint is then

$$-\exp\{-r(\underline{s} - c(\bar{u}))\}\, F(\overline{\Delta X} \mid \bar{u}) - \exp\{-r(\bar{s} - c(\bar{u}))\}(1 - F(\overline{\Delta X} \mid \bar{u}))$$

$$\geqslant -\exp\{-rW_A\}, \quad (12)$$

which holds with equality if $\underline{s}$ and $\bar{s}$ are replaced by (8) and (9), respectively. ∎

We are now in the position to show that the principal can approximate the first-best solution arbitrarily closely. Consider a random spot check with payments $\underline{s}$ and $\bar{s}$ given by (8)–(9) where $\bar{u} = u_{FB}$ and where $\overline{\Delta X}$ is sufficiently small. By Lemma 2, this incentive scheme implements the (constant) first-best control, i.e., the agent chooses u_{FB} in each subinterval and his participation constraint is satisfied. The principal's expected utility is then

$$E[X_1 \mid u_{FB}] - \underline{s}F(\overline{\Delta X} \mid u_{FB}) - \bar{s}(1 - F(\overline{\Delta X} \mid u_{FB})). \quad (13)$$

In particular, since $\bar{s} > \underline{s}$, the expected payment to the agent is bounded from above by $\bar{s}$, which implies that (13) is bounded from below by

$$E[X_1 \mid u_{FB}] - W_A - c(u_{FB}) + \frac{1}{r} \ln\left[1 + rc'(u_{FB}) \frac{F(\overline{\Delta X} \mid u_{FB})}{F_u(\overline{\Delta X} \mid u_{FB})}\right]. \quad (14)$$

As ΔX is normally distributed with mean $f(u_{FB})\,\Delta t$ and variance $\sigma^2\,\Delta t$, we have

$$\lim_{\overline{\Delta X} \to -\infty} \frac{F(\overline{\Delta X} \mid u_{FB})}{F_u(\overline{\Delta X} \mid u_{FB})} = 0, \tag{15}$$

which implies that the rightmost term in (14) converges to zero as $\overline{\Delta X} \to -\infty$. In other words for any $\varepsilon > 0$, the principal can find a cutoff $\overline{\Delta X}(\varepsilon)$ such that his expected utility is equal to the first-best utility $E[X_1 \mid u_{FB}] - W_A - c(u_{FB})$ minus ε. As in [5, 6], this is due to the fact that the likelihood ratio $f_u(\overline{\Delta X} \mid u_{FB})/f(\overline{\Delta X} \mid u_{FB})$ tends to $-\infty$ as $\overline{\Delta X} \to -\infty$, implying that low values of ΔX are an extremely accurate signal that the agent has shirked. We thus have the following result.

THEOREM 1. *In the discrete-time version of the Brownian model, the principal can approach the first-best outcome asymptotically.*

3. EXTENSIONS AND DISCUSSION

3.1. *Private Information about Timing of Control Revisions*

While Section 2 assumes that the timing of the agent's control revisions is common knowledge, our results continue to hold if the agent is privately informed about the number of control revisions as long as T is finite (that is, as long as the agent cannot revise his effort continuously) and the principal knows the partition of $[0, 1]$ associated with each possible value of T. As an illustration, consider the case where the principal only knows that $T \in \{1, 2, 3\}$ and that all subintervals have equal length. By using a random spot check with "test interval" length $\frac{1}{6}$ (i.e., the randomly selected increment ΔX_t is chosen from one of the subintervals $[0, 1/6]$, ..., $[5/6, 1]$), the principal can ensure that during any of the six possible test intervals the agent's control is constant. As a consequence, the agent faces the same control problem in each subinterval, which once again implies that the principal can implement any desired constant control. More generally, if it is common knowledge that $T \in \{\underline{T}, ..., \overline{T}\}$, where $\overline{T} < \infty$, any constant control can be implemented by choosing a test interval of length $1/(\overline{T}(\overline{T}-1))$. The rest is analogous to Section 2.

If the partition of $[0, 1]$ associated with a particular value of T is not common knowledge, the principal can no longer approximate the first-best outcome. For instance, if it is common knowledge that the agent can revise his control T times but the dates at which the revisions take place are private information, step functions fail to implement the constant first-best control—very much like in the example at the beginning of Section 2 where step functions were based on aggregated end-of-period output X_1.

3.2. *Implementation via Alternative Incentive Schemes*

In many applications of interest, performing a random spot check after the entire output has been produced is impossible. For instance, if production process is organized as a *kanban* system, work-in-progress is not accumulated but transported directly to the production unit where it is used next. Nevertheless, the principal can implement any desired constant control by randomly selecting a subinterval $[t, t+\Delta t]$ at time 0 and measuring the corresponding increment ΔX_t immediately after it has been produced, i.e., at time $t+\Delta t$. As long as the agent remains in the dark about the timing of the principal's inspection, he faces the same control problem in each subinterval and optimally selects a constant control. Consequently, both the random draw at time 0 and the inspection at time $t+\Delta t$ must be secret but verifiable ex post. In practice, such "secret inspections" are possible if, e.g., the output is forwarded to a different work unit where it can be checked by the principal without the agent's noticing it.

So far, attention has been restricted to random spot checks as these —like linear sharing rules—are widely used in practice. However, this does not imply that random spot checks are the *only* incentive schemes allowing the principal to approximate the first-best solution in the discrete-time version of the Brownian model. For instance, consider an incentive scheme where *each* of the T subintervals is checked and where the agent receives $\underline{s}$ for each subinterval where $\Delta X_t \leqslant \overline{\Delta X}$ and $\bar{s}$ for each subinterval where $\Delta X_t > \overline{\Delta X}$. As the agent's control problem is the same in each subinterval, he optimally chooses a constant control. Moreover, by letting $\overline{\Delta X} \to -\infty$, the principal can approach the first-best solution asymptotically by straightforward analogy with the analysis in Section 2. What makes this scheme unattractive though is that for arbitrarily small inspection costs $\varepsilon > 0$ per subinterval, the costs of providing constant incentives explodes as the number of subintervals goes to infinity.

3.3. *General Discussion*

This paper has shown that, if the agent acts in discrete time rather than in continuous time, linear sharing rules are no longer optimal in the Brownian model by Holmström and Milgrom [4] as the first-best solution can then be approximated arbitrarily closely with a random spot check and a suitably chosen sequence of step functions. Intuitively, this is because in the discrete-time version of the Brownian model, the principal's compensation options are no longer smaller than the agent's action options, implying that the principal has many ways to implement any desired action.[4]

In their paper, Holmström and Milgrom already indicate that the Brownian model may be viewed as the limit of different processes with very different discrete-time behavior.[5] The issue is therefore not so much whether the linearity result is robust with respect to the choice of discrete-time approximation, but to what extent the differences in results obtained from looking at different discrete-time approximations help us to understand the relation between the economic and technical aspects of the Brownian model. In the present (i.e., discrete-time) version, each output ΔX_t is associated with a unique control u_t, which may be viewed as, e.g., the choice of input quality used in the production of the output. If, for whatever reasons, the input quality cannot be changed during the production of ΔX_t (e.g., because of technological constraints), the discrete-time model may be descriptively more relevant. On the other hand, if the input quality can be adjusted while ΔX_t is being produced, the continuous-time version of the Brownian model may be better suited in representing the situation.

REFERENCES

1. R. J. Elliot, "Stochastic Calculus and Applications," Springer-Verlag, New York, 1982.
2. O. D. Hart and B. R. Holmström, The theory of contracts, *in* "Advances in Economic Theory, Fifth World Congress" (T. Bewley, Ed.), Cambridge Univ. Press, Cambridge, UK, 1987.
3. M. F. Hellwig and K. M. Schmidt, "Discrete-Time Approximations of the Holmström–Milgrom Brownian Motion Model of Intertemporal Incentive Problems," Discussion Paper 98-06, SFB504, University of Mannheim, 1998.
4. B. R. Holmström and P. R. Milgrom, Aggregation and linearity in the provision of intertemporal incentives, *Econometrica* **55** (1987), 303–328.
5. J. A. Mirrlees, Notes on welfare economics, information and uncertainty, *in* "Essays on Economic Behavior under Uncertainty" (M. Balch, D. McFadden, and S.-Y. Wu, Eds.), North-Holland, Amsterdam, 1974.
6. J. A. Mirrlees, The theory of moral hazard and unobservable behavior, Part I, *Rev. Econ. Stud.* **66** (1999), 3–21.

[4] The linearity result can be restored by allowing the agent to act *after* privately observing the realization of the stochastic variable. In such a hidden information model, the agent's action can be thought of as a contingent strategy mapping observed signals into effort levels. The agent's action set is then of higher dimension than the principal's compensation options, which means that Holmström and Milgrom's uniqueness result (Theorem 3) again applies. For a discussion, see Holmström and Milgrom [4, p. 308].

[5] In the paper by Holmström and Milgrom, the Brownian model is derived as the limit of a sequence of multi-period discrete-time models in which the agent controls the probability vector of a multinomial distribution. For a discussion of the relation between the discrete-time and the continuous-time model, see [3].

7. H. M. Müller, "Randomization in Dynamic Principal-Agent Problems," Discussion Paper 97-37, SFB504, University of Mannheim, 1997.
8. H. M. Müller, The first-best sharing rule in the continuous-time principal-agent problem with exponential utility, *J. Econ. Theory* **79** (1998), 276–280.
9. H. Schättler and J. Sung, The first-order approach to the continuous-time principal-agent problem with exponential utility, *J. Econ. Theory* **61** (1993), 331–371.
10. J. Sung, Linearity with project selection and controllable diffusion rate in continuous-time principal-agent problems, *Rand. J. Econ.* **26** (1995), 720–743.