SIAM J. CONTROL AND OPTIMIZATION
Vol. 22, No 2, March 1984

0363-0129/84/2202-0003 $01.25/0

AFFINE INCENTIVE SCHEMES FOR STOCHASTIC SYSTEMS WITH DYNAMIC INFORMATION*

TAMER BAŞAR†

Abstract. In this paper we study the derivation of optimal incentive schemes in two-agent stochastic decision problems with a hierarchical decision structure, in a general Hilbert space setting. The agent at the top of the hierarchy is assumed to have access to the value of other agent's decision variable as well as to some common and private information, and the second agent's loss function is taken to be strictly convex. In this set-up, it is shown that there exists, under some fairly mild structural restrictions, an optimal incentive policy for the first agent, which is affine in the dynamic information and generally nonlinear in the static (common and private) information. Certain special cases are also discussed and a numerical example is solved.

Key words. stochastic systems, decision problems with multiple decision makers, incentive schemes, hierarchical information patterns, stochastic nonzero-sum games, Stackelberg solution

1. Introduction. Consider the general class of two-agent stochastic dynamic decision problems with a hierarchical decision structure, wherein one of the agents (called the leader) has access to both the decision value and observation of the other agent (called the follower), and the objective is verification of existence and derivation of optimal strategies for the leader under which the follower's optimal response (based on the minimization of his expected cost function) leads to a desired "optimal" performance for the leader. Such problems are known as Stackelberg problems [1]–[5] or incentive design problems [8], [20]–[23] and have recently attracted considerable attention in the literature, because of the nonstandard nature of the optimization problem faced by the leader, when he has access to dynamic information [6]–[18]; for a survey and unification of some of the available results in the literature on deterministic and stochastic dynamic Stackelberg problems we refer to [8], [12] and [19], and also to [26] for a general discussion.

A recent reference [15] has shown that in deterministic dynamic incentive problems with perfect or partial dynamic information, and when the follower's cost function is strictly convex (but not necessarily quadratic), there exists an optimal incentive strategy for the leader which is affine in the dynamic information. The object of this paper is to provide a nontrivial extension of this result to stochastic decision problems in which there is available some common information on the unknown state of Nature to both agents as well as some private information to the leader; the leader has also access to the value of the follower's decision variable. The problem is formulated in general Hilbert spaces with the follower's loss function taken to be strictly convex in both agents' decision variables. In this general framework, we establish existence of an optimal incentive strategy for the leader, which is affine in the dynamic information, and in general nonlinear in the static (common and private) information; we also obtain an analytic expression for the optimal solution and consider some special cases of the general problem.

* Received by the editors May 28, 1982, and in revised form January 20, 1983. This work was supported in part by the Joint Services Electronics Program under contract N00014-79-C-0424 and in part by the Electric Energy Systems Division, Department of Energy under contract DE-AC01-81RA-50658 with Dynamic Systems, Urbana, Illinois 61801. It was presented at the American Automatic Control Conference, Arlington, Virginia, June 14–16, 1982.

† Department of Electrical Engineering and Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801.

Section 2 provides a precise problem formulation for the case when only common information is available to the leader, whose solution is obtained in § 3 (cf. Proposition 1). Section 4 extends the formulation and results of §§ 2 and 3 to the more general case when the leader has also access to some private information, and a characterization of the complete affine solution is provided in Proposition 2. Section 5 contains a numerical example that serves to illustrate some salient aspects of the solution, and the paper ends with the concluding remarks of § 6.

2. Problem formulation. Let $(\Omega, \mathscr{F}, \mathscr{P})$ be an underlying probability space, on which are defined two random variables x and z taking values in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Let U and V be given Hilbert spaces, denoting the decision spaces of DM1 (the leader) and DM2 (the follower) respectively, and let Γ_1 and Γ_2 be the corresponding strategy spaces defined as

$$\Gamma_2 = \{\text{measurable } \gamma_2 : \mathbb{R}^m \to V, \text{ such that } \underset{z}{E}\{\langle \gamma_2(z), \gamma_2(z)\rangle\} < \infty\}, \tag{1}$$

$$\Gamma_1 = \{\text{measurable } \gamma_1 : V \times \mathbb{R}^m \to U, \text{ such that } \underset{z}{E}\{\langle \gamma_1[\gamma_2(z), z], \gamma_1[\gamma_2(z), z]\rangle_u\} < \infty \ \forall \gamma_2 \in \Gamma_2\}. \tag{2}$$

Furthermore let $\Gamma_1^s \subset \Gamma_1$ denote the set of all static policies for the leader, i.e.

$$\Gamma_1^s = \{\text{measurable } \gamma_1 : \mathbb{R}^m \to U, \text{ such that } \underset{z}{E}\{\langle \gamma_1(z), \gamma_1(z)\rangle_u\} < \infty\}. \tag{3}$$

Here, $\langle \cdot, \cdot \rangle_u$ (respectively, $\langle \cdot, \cdot \rangle_v$) denotes the inner product associated with the Hilbert space U (respectively, V), and the measurable transformations are restricted by the further (implicit) condition that the expectations of the related expressions are well defined. With this construction, Γ_1, Γ_1^s, and Γ_2 become Hilbert spaces under the natural inner products derived from those defined on U, V, and V, respectively. Note that, to each pair (γ_1, γ_2) in $\Gamma_1 \times \Gamma_2$, there corresponds an unique element $\beta_1 \in \Gamma_1^s$, defined by $\beta_1(z) = \gamma_1[\gamma_2(z), z]$.

We now introduce two functions, $L_1 : \mathbb{R}^n \times U \times V \to \mathbb{R}$, $L_2 : \mathbb{R}^n \times U \times V \to \mathbb{R}$, as the loss functions of DM1 and DM2 respectively, and further introduce

$$J_i : \mathbb{R}^m \times \Gamma_1 \times \Gamma_2 \to \mathbb{R}, \qquad i = 1, 2,$$

as

$$J_i(z, \gamma_1, \gamma_2) = \underset{x|z}{E}\{L_1(x, u, v) \mid u = \gamma_1(v, z),\ v = \gamma_2(z)\}, \tag{4}$$

where $E_{x|z}$ denotes expectation over the statistics of x with conditioning on the observed value of z. Finally, we let $\bar{J}_i(\gamma_1, \gamma_2)$, $i = 1, 2$, defined by

$$\bar{J}_i(\gamma_1, \gamma_2) = \underset{z}{E}\{J_i(z, \gamma_1, \gamma_2)\}, \tag{5}$$

denote the expected cost of DMi, under the policy pair $\gamma_1 \in \Gamma_1$, $\gamma_2 \in \Gamma_2$. We assume at this point that the follower's loss functional $L_2(x, u, v)$ is strictly convex on $U \times V$, for every $x \in \mathbb{R}^n$.

The problem faced by the leader is to find a strategy (synonymously, an incentive scheme) which, by also taking into account rational (expected-cost minimizing) responses of the follower, leads to a most favorable performance for the leader. This performance may be defined as the global minimum value of $\bar{J}_1(\gamma_1, \gamma_2)$ over $\Gamma_1 \times \Gamma_2$, or equivalently over $\Gamma_1^s \times \Gamma_2$ [assuming that it exists]:

$$\bar{J}_1^t = \min_{(\gamma_1,\gamma_2)\in\Gamma_1^s\times\Gamma_2} \bar{J}_1(\gamma_1, \gamma_2) = \bar{J}_1(\gamma_1^t, \gamma_2^t) \tag{6}$$

which corresponds to some specific choices of $\gamma_1 \in \Gamma_1^s$ and $\gamma_2 \in \Gamma_2$ (in this case, $\gamma_1 = \gamma_1^t$ and $\gamma_2 = \gamma_2^t$); or, more generally, there may exist some pair in $\Gamma_1^s \times \Gamma_2$ (denoted again (γ_1^t, γ_2^t)) which is chosen according to some criterion and is considered to be most favorable to the leader. The question we address, in the next section, is the existence and derivation of an "optimal" incentive scheme $\gamma_1^0 \in \Gamma_1$ for the leader, under which the best ($\bar{J}_2$-minimizing) policy for the follower is $\gamma_2^t \in \Gamma_2$, and a corresponding element in Γ_1^s for the leader is $\gamma_1^t = \gamma_1^0[\gamma_2^t(z), z]$. Note that this is a meaningful problem because, given a pair $(\gamma_1^t, \gamma_2^t) \in \Gamma_1^s \times \Gamma_2$, there exists a plethora of elements γ_1 in Γ_1 with the property $\gamma_1[\gamma_2^t(z), z] = \gamma_1^t(z)$. Furthermore, $\gamma_1^0 \in \Gamma_1$, in this case, is clearly a Stackelberg strategy for the leader [12].

3. Optimal affine incentive schemes for the leader.

Let us first introduce, for each $z \in \mathbb{R}^m$, the set

$$\Omega_t(z) = \{(u, v) \in U \times V \mid \tilde{J}_2(z, u, v) \leqq \tilde{J}_2(z, u_z^t, v_z^t)\} \tag{7}$$

where

$$\tilde{J}_2(z, u, v) \triangleq \mathop{E}_{x|z} \{L_2(x, u, v) | z\} \tag{8}$$

and (u, v) are taken as (deterministic) elements in $U \times V$. Since $L_2(x, \cdot, \cdot)$ was taken to be strictly convex on $U \times V$, $\Omega_t(z)$ is also strictly convex for each $z \in \mathbb{R}^m$, with $\{u_z^t = \gamma_1^t(z), v_z^t = \gamma_2^t(z)\}$ being a boundary point. This implies that, for each $z \in \mathbb{R}^m$, there exists a hyperplane passing through (u_z^t, v_z^t), and if, further, $\tilde{J}_2(z, \cdot, \cdot)$ is Fréchet differentiable on $U \times V$, for each $z \in \mathbb{R}^m$, the equation of this supporting hyperplane can be written as

$$\langle \nabla_u \tilde{J}_2(z, u_z^t, v_z^t), u - u_z^t \rangle_u + \langle \nabla_v \tilde{J}_2(z, u_z^t, v_z^t), v - v_z^t \rangle_v = 0 \tag{9}$$

where $\nabla_u \tilde{J}_2(z, u_z^t, v_z^t) \in U^*$ is the Fréchet derivative of $\tilde{J}_2$ with respect to u, evaluated at the point (u_z^t, v_z^t), and U^* is the Hilbert space adjoint to U; $\nabla_v \tilde{J}_2$ is analogously defined as an element of the adjoint space V^*. Now, assuming that, for every $z \in \mathbb{R}^m$, $\nabla_u \tilde{J}_2(z, u_z^t, v_z^t) \neq 0$, it follows by utilizing the Hahn–Banach theorem [25] (see also [15, Lemma 1]) that there exists a bounded linear operator $Q_z^*: U^* \to V^*$ satisfying

$$Q_z^* \nabla_u \tilde{J}_2(z, u_z^t, v_z^t) = \nabla_v \tilde{J}_2(z, u_z^t, v_z^t), \tag{10}$$

so that

$$u_z = u_z^t - Q_z(v - v_z^t) \tag{11}$$

lies, for each $z \in \mathbb{R}^m$, on the hyperplane described by (9) and passes through the point (u_z^t, v_z^t). Here, $Q_z: V \to U$ is the bounded linear operator that is adjoint to Q_z^*, for each fixed $z \in \mathbb{R}^m$.

The next question is whether (11) is a well-defined strategy for the leader, i.e. whether it belongs to Γ_1, which requires Q_z to be a measurable function of z. We now establish an even stronger regularity property for Q_z under some regularity conditions on $\bar{J}_2$, u_z^t and v_z^t:

LEMMA 1. *Let $\nabla_u \tilde{J}_2(z, u_z^t, v_z^t)$ and $\nabla_v \tilde{J}_2(z, u_z^t, v_z^t)$ be weakly continuous*[1] *in z. Then, there exists a linear bounded operator $Q_z: V \to U$, weakly continuous in z, whose adjoint satisfies the linear equation* (10).

Proof. For brevity in notation, let $\nabla_u \tilde{J}_2(z, u_z^t, v_z^t) \triangleq \tilde{u}_z^*$, and $\nabla_v \tilde{J}_2(z, u_z^t, v_z^t) \triangleq \tilde{v}_z^*$. For each fixed $z \in \mathbb{R}^m$, introduce a bounded linear operator $P_z: U^* \to V^*$ by

$$P_z u^* = \frac{\langle \tilde{u}_z^*, u^* \rangle_{u^*}}{\|\tilde{u}_z^*\|_{u^*}^2} \tilde{v}_z^*, \quad \text{with } u^* \in U^*, \tag{12}$$

which clearly satisfies (10), when substituted for Q_z^*. Now, for any $u^* \in U^*$ and $v \in V$,

$$\langle v, P_z u^* \rangle_v = \langle \tilde{u}_z^*, u^* \rangle_{u^*} \frac{\langle v, \tilde{v}_z^* \rangle_v}{\|\tilde{u}_z^*\|_{u^*}^2},$$

and the latter expression is a continuous functional of z by virtue of the weak continuity of $\tilde{u}_z^*$ and $\tilde{v}_z^*$. This then implies that P_z is weakly continuous in z, and thereby P_z^* is also weakly continuous in z [24]. Now, taking $Q_z = P_z^*$, it readily follows that there exists a version of Q_z (satisfying (10)) that is weakly continuous in z. □

The following proposition now summarizes the solution to the incentive problem formulated in § 2.

PROPOSITION 1. *For the incentive problem of* § 2, *if* (i) $J_2(z, u, v)$ *is Fréchet differentiable on $U \times V$,* (ii) *for every $z \in \mathbb{R}^m$, $\nabla_u \tilde{J}_2(z, u_z^t, v_z^t) \neq 0$, and* (iii) $\nabla_u \tilde{J}_2(z, u_z^t, v_z^t)$ *and $\nabla_v \tilde{J}_2(z, u_z^t, v_z^t)$ are weakly continuous*[2] *in z, there exists an optimal incentive strategy $(\gamma_1^0(v, z)$ for the leader, in the form*

$$u_z^0 = \gamma_1^0(v, z) = u_z^t - Q_z(v - v_z^t), \tag{13}$$

where the linear operator $Q_z: V \to U$ is chosen according to (10) *and is weakly continuous in z.*

Remark 1. If U and V are finite-dimensional spaces, $U^* = U$ and $V^* = V$, and consequently $\nabla_u \tilde{J}_2$ and $\nabla_v \tilde{J}_2$ are (column) vectors of appropriate dimensions for each $z \in \mathbb{R}^m$. Then Q_z becomes a matrix-valued function of z, and can be chosen as

$$Q_z = \nabla_u \tilde{J}_2(z, u_z^t, v_z^t)[\nabla_v \tilde{J}_2(z, u_z^t, v_z^t)]'/\|\nabla_u \tilde{J}_2(z, u_z^t, v_z^t)\|^2. \tag{14}$$

Note that, under the hypotheses of Proposition 1, every element of Q_z will be a continuous function of z.

As a special class of problems, let us consider now the case when $L_1(z, u, v)$ is quadratic on $\mathbb{R}^n \times U \times V$:

$$L_2(x, u, v) = \tfrac{1}{2}\langle u, A_{11}u\rangle_u + \langle u, A_{12}v\rangle_u + \tfrac{1}{2}\langle v, A_{22}v\rangle_v + \langle u, C_1 x\rangle_u + \langle v, C_2 x\rangle_v, \tag{15}$$

where A_{ij} and C_i are linear bounded operators, A_{22} is strongly positive, and $A_{11} - A_{12}(A_{22})^{-1}A_{21}^*$ is also strongly positive. Then,

$$\tilde{J}_2(z, u, v) = \tfrac{1}{2}\langle u, A_{11}u\rangle_u + \langle u, A_{12}v\rangle_u + \tfrac{1}{2}\langle v, A_{22}v\rangle_v + \langle u, C_1\hat{x}\rangle_u + \langle v, C_2\hat{x}\rangle_v$$

[1] See [24] for a definition of weak continuity.

[2] A set of sufficient conditions for this is that i) $\tilde{J}_2(z, u, v)$ be continuously Fréchet differentiable in u and v, and be continuous in z, and ii) u_z^t and v_z^t be weakly continuous in z.

where $\hat{x}=E[x|z]$. Given a point $(u_z^t, v_z^t)\in U\times V$, for each $z\in\mathbb{R}^m$, the Fréchet derivatives at this operating point can easily be determined to be $\langle\cdot,\tilde{u}_z\rangle_u$ and $\langle\cdot,\tilde{v}_z\rangle_v$, where $\tilde{u}_z\in U$ and $\tilde{v}_z\in V$ are

$$\tilde{u}_z=A_{11}u_z^t+A_{12}v_z^t+C_1\hat{x}, \tag{16a}$$

$$\tilde{v}_z=A_{12}^*u_z^t+A_{22}v_z^t+C_2\hat{x}. \tag{16b}$$

This then leads to the following Corollary (to Proposition 1) in view of (12).

COROLLARY 1. *Let L_2 be given by* (15), *and $\tilde{u}_z$ and $\tilde{v}_z$ be defined by* (16a) *and* (16b), *respectively. If u_z^t, v_z^t are weakly continuous in z, $\hat{x}=E[x|z]$ is continuous in z, and, for every $z\in\mathbb{R}^m$, $\tilde{u}_z\neq 0$, there exists an optimal incentive strategy for the leader which is affine in v and weakly continuous in z, and is given by*

$$\gamma_1^0(v,z)=u_z^t-\frac{\tilde{u}_z}{\langle\tilde{u}_z,\tilde{u}_z\rangle_u}\langle\tilde{v}_z, v-v_z^t\rangle_v. \tag{17}$$

Proof. In view of the discussion preceding Corollary 1, the proof will be complete if we show that $Q_z\colon V\to U$ in (13) (and (10)) is given by $[\langle\tilde{v}_z,\cdot\rangle_v/\|\tilde{u}_z\|_u^2]\tilde{u}_z$. Towards this end, we first observe from (12) that a possible solution of (10) is given by

$$Q_z^*u^*=\frac{\langle\tilde{u}_z^*,u^*\rangle_{u^*}}{\|\tilde{u}_z^*\|_{u^*}^2}\tilde{v}_z^*\quad\text{with } u^*\in U^*,$$

where $\tilde{u}_z^*$ and $\tilde{v}_z^*$ are the Fréchet derivatives belonging to U^* and V^*, respectively. Since U^* and V^* are Hilbert spaces, corresponding to $\tilde{u}_z^*$ and $\tilde{v}_z^*$ there are unique elements $\tilde{u}_z\in U$ and $\tilde{v}_z\in V$, with the property $\langle\tilde{u}_z^*,u^*\rangle_{u^*}=\langle\tilde{u}_z,u^*\rangle_u$ and $\langle\hat{v}_z^*,v^*\rangle_{v^*}=\langle\tilde{v}_z,v^*\rangle_v$ for all $u^*\in U^*$, $v^*\in V^*$ and every fixed $z\in\mathbb{R}^m$ (see [25]). These elements $\tilde{u}_z$ and $\tilde{v}_z$ can explicitly be determined in our case (because of the specific structure of L_2) and are given by (16a) and (16b), respectively. Hence, we have

$$\langle v, Q_z^*u^*\rangle_v=\frac{\langle\tilde{u}_z,u^*\rangle_u}{\|\tilde{u}_z\|_u^2}\langle v,\tilde{v}_z\rangle_v=\langle Q_zv,u^*\rangle_u$$

whereby

$$Q_zv=\frac{\langle v,\tilde{v}_z\rangle_v}{\|\tilde{u}_z\|_u^2}\tilde{u}_z,$$

which establishes the desired results. □

4. A more general formulation: Leader acquires private information. We now extend the analysis and results of the previous section to a more general class of incentive problems wherein the leader observes, in addition to z, the output of a second random variable $\tilde{y}$ (taking values in $\mathbb{R}^p$). This random variable will in general be correlated with x and z; however, we assume (for technical reasons) existence of a measurable transformation $f\colon\mathbb{R}^m\times\mathbb{R}^p\to\mathbb{R}^p$, so that the random variables $y=f(\tilde{y},z)$ and z are statistically independent, and the sigma fields generated by $(\tilde{y},z)$ and (y,z) are the same. (For example, if y and z have a joint Gaussian distribution, $f(\tilde{y},z)\triangleq\tilde{y}-E[\tilde{y}|z]$.) Therefore, we henceforth assume that $u=\gamma_1(v,z,y)$, $v=\gamma_2(z)$, and z and y are statistically independent.

For this problem, we now first modify the definitions of the strategy sets (2) and (3) to read

$$\Gamma_1=\{\text{measurable } \gamma_1: V\times\mathbb{R}^m\times\mathbb{R}^p\to U, \text{ such that } \underset{z,y}{E}\{\langle\gamma_1[\gamma_2(z),z,y],\gamma_1[\gamma_2(z),z,y]\rangle_u\}<\infty\ \forall\gamma_2\in\Gamma_2\} \tag{18}$$

and

$$\Gamma_1^s=\{\text{measurable } \gamma_1:\mathbb{R}^m\times\mathbb{R}^p\to U, \text{ such that } \underset{z,y}{E}\{\langle\gamma_1(z,y),\gamma_1(z,y)\rangle_u\}<\infty\}. \tag{19}$$

Furthermore, we redefine J_i and $\bar{J}_i$ as

$$J_i(z,\gamma_1,\gamma_2)=\underset{x,y|z}{E}\{L_i(x,u,v)|u=\gamma_1(v,z,y), v=\gamma_2(z)\}, \tag{20}$$

$$\bar{J}_i(\gamma_1,\gamma_2)=\underset{z}{E}\{J_i(z,\gamma_1,\gamma_2)\}, \tag{21}$$

and let $\{u_{z,y}^t=\gamma_1^t(z,y), v_z^t=\gamma_2^t(z)\}$ denote a pair in $\Gamma_1^s\times\Gamma_2$ that (globally) minimizes $\bar{J}_1$. For each fixed $z\in\mathbb{R}^m$, we let

$$B(z)=\{\text{measurable } \beta_z:\mathbb{R}^p\to U, \text{ so that } \underset{y|z}{E}\{\langle\beta_z(y),\beta_z(y)\rangle_u\}<\infty\}, \tag{22}$$

and note that to each $\gamma_1\in\Gamma_1^s$ and for fixed $z\in\mathbb{R}^m$ there will correspond a unique $\beta_z\in B(z)$ such that $\gamma_1(z,\cdot)=\beta_z(\cdot)$.

Now utilizing the statistical independence of z and y, let us introduce as a counterpart (7), and for each $z\in\mathbb{R}^m$, the set

$$\Omega_t(z)=\{(\beta,v)\in B(z)\times V|\tilde{J}_2(z,\beta,v)\leqq\tilde{J}_2(z,\beta_z^t,v_z^t)\}, \tag{23}$$

where

$$\tilde{J}_2(z,\beta,v)=\underset{x,y|z}{E}\{L_2(x,\beta(y),v)|z\}, \tag{24}$$

and β_z^t is the restriction of $\gamma_1^t\in\Gamma_1^s$ to $B(z)$. It is worth to note, at this point, that

$$\begin{aligned}\underset{z}{E}\{\tilde{J}_2(z,\beta_z^t,v_z^t)\}&=\underset{z}{E}\left\{\underset{x,y|z}{E}\{L_2(x,\beta_z^t(y),\gamma_2^t(z))|z\}\right\}\\&=\underset{x,y,z}{E}\{L_2(x,\gamma_1^t(x,\gamma_1^t(z,\gamma),\gamma_2^t(z)))\}\equiv\bar{J}_2(\gamma_1^t,\gamma_2^t).\end{aligned} \tag{25}$$

We can now proceed with the derivation of an optimal incentive scheme by following the analysis of § 3, with U replaced by $B(z)$, where the latter can be made a Hilbert space under the inner product

$$\langle\beta_1,\beta_2\rangle_\beta=\underset{y}{E}\{\langle\beta_1(y),\beta_2(y)\rangle_u\},\qquad \beta_i\in B(z).$$

It is easy to see that strict convexity of $L_2(x,\cdot,\cdot)$ on $U\times V$ implies strict convexity of $\tilde{J}_2(z,\cdot,\cdot)$ on $B(z)\times V$, for each $z\in\mathbb{R}^m$, and hence assuming that the latter is Fréchet differentiable on $B(z)\times V$, the equation of the hyperplane supporting $\Omega_t(z)$ at (β_z^t, v_z^t) is

$$\langle\nabla_\beta\tilde{J}_2(z,\beta_z^t,v_z^t),\beta-\beta_z^t\rangle_\beta+\langle\nabla_v\tilde{J}_2(z,\beta_z^t,v_z^t),v-v_z^t\rangle_v=0 \tag{26}$$

where $\nabla_\beta \tilde{J}_2(z, \beta^t_z, v^t_z) \in B(z)^*$ is the Fréchet derivative of $\tilde{J}_2$ with respect to β, and evaluated at the point (β^t_z, v^t_z). Since there is a natural counterpart of Lemma 1 in this framework, validity of the following counterpart of Proposition 1 can readily be established:

PROPOSITION 2. *For the incentive problem formulated in this section, if* (i) $\tilde{J}_2(z, \beta, v)$ *is Fréchet differentiable on* $B(z) \times V$, (ii) *for every* $z \in \mathbb{R}^m$, $\nabla_\beta \tilde{J}_2(z, \beta^t_z, v^t_z) \neq 0$, *and* (iii) $\nabla_\beta \tilde{J}_2(z, \beta^t_z, v^t_z)$ *and* $\nabla_v \tilde{J}_2(z, \beta^t_z, v^t_z)$ *are weakly continuous in* z, *there exists an optimal incentive strategy* $\gamma_1^0(v, z, y)$ *for the leader, given by*

$$u^0_{z,y} = \gamma_1^0(v, z, y) = u^t_{z,y} - Q_z(v - v^t_z)(y), \tag{27}$$

where $Q_z: V \to B(z)$ *is a linear bounded operator which is weakly continuous in* z, *and whose adjoint satisfies the linear equation*

$$Q_z^* \nabla_\beta \tilde{J}_2(z, \beta^t_z, v^t_z) = \nabla_v \tilde{J}_2(z, \beta^t_z, v^t_z), \tag{28}$$

which is defined on V^*.

For the special case when L_2 is quadratic, as given by (15), $\tilde{J}_2(z, \beta, v)$ can be written as

$$\begin{aligned}\tilde{J}_2(z, \beta, v) &= \underset{x,y|z}{E}\{\tfrac{1}{2}\langle \beta(y), A_{11}\beta(y)\rangle_u + \langle \beta(y), A_{12}v\rangle_u \\ &\qquad + \tfrac{1}{2}\langle v, A_{22}v\rangle_v + \langle \beta(y), C_1x\rangle_u + \langle v, C_2x\rangle_v\} \\ &= \tfrac{1}{2}\underset{y}{E}\{\langle \beta(y), A_{11}\beta(y)\rangle_u\} + \langle \hat{\beta}, A_{12}v\rangle_u \\ &\qquad + \tfrac{1}{2}\langle v, A_{22}v\rangle_v + \underset{x,y|z}{E}\{\langle \beta(y), C_1x\rangle_u + \langle v, C_2\hat{x}\rangle_v\}\end{aligned} \tag{29}$$

where

$$\hat{\beta}_z = \underset{y|z}{E}\{\beta_z(y)\,|\,z\} = \underset{y}{E}\{\beta_z(y)\},$$

$$\hat{x}(z) \triangleq \underset{x,y|z}{E}\{x\,|\,z\} = \underset{x|z}{E}\{x\,|\,z\}.$$

For fixed $z \in \mathbb{R}^m$, the Fréchet (or Gateaux) differential [25] of $\tilde{J}_2$ with respect to β, and at the point (β^t_z, v^t_z) is

$$\begin{aligned}\delta_\beta \tilde{J}_2(z, \beta^t_z, v^t_z; h_z) &= \underset{y}{E}\{\langle \beta^t_z(y), A_{11}h_z(y)\rangle_u\} \\ &\qquad + \langle A_{12}v^t_z, \hat{h}_z\rangle_u + \underset{x,y|z}{E}\{\langle C_1x, h_z(y)\rangle_u\},\end{aligned}$$

where $h_z \in B(z)$ is an admissible variation and $\hat{h}_z \triangleq E_y\{h_z(y)\}$. Since

$$\underset{x,y|z}{E}\{\cdot\} = \underset{y}{E}\left\{\underset{x|y,z}{E}\{\cdot\}\right\},$$

this expression can be written as

$$\delta_\beta \tilde{J}_2(z, \beta^t_z, v^t_z; h_z) = \underset{y}{E}\{\langle A_{11}\beta^t_z(y) + A_{12}v^t_z + C_1\hat{\hat{x}}, h_z(y)\rangle_u\}$$

where

$$\hat{\hat{x}}(z, y) \triangleq \underset{x|y,z}{E}\{x\,|\,y, z\},$$

and it readily follows from this expression that the Fréchet derivative of $\tilde{J}_2$ with respect to $\beta_z \in B(z)$ is $\langle \cdot, \tilde{\beta}_z \rangle_\beta$ where $\tilde{\beta}_z \in B(z)$ is given by

$$\tilde{\beta}_z = A_{11}\beta_z^t + A_{12}v_z^t + C_1\hat{\hat{x}}(z, \cdot). \tag{30}$$

The Fréchet derivative with respect to $v \in V$, on the other hand, readily follows from (29) to be $\langle \cdot, \tilde{v}_z \rangle_v$ where $\tilde{v}_z \in V$ is given by

$$\tilde{v}_z = A_{12}^* \underset{y}{E}\{\beta_z^t(y)\} + A_{22}v_z^t + C_2\hat{x}(z). \tag{31}$$

In view of these relations, a possible solution for Q_z, whose adjoint satisfies (28), is

$$Q_z(\cdot) = \frac{\langle \cdot, \tilde{v}_z \rangle_v}{\|\tilde{\beta}_z\|_\beta^2}\tilde{\beta}_z, \tag{32}$$

which follows by following the arguments used in the proof of Corollary 1 in § 3. This then leads to the following corollary (to Proposition 2):

COROLLARY 2. *Let L_2 be given by* (15), *and $\tilde{\beta}_z$ and $\tilde{v}_z$ be defined by* (30) *and* (31), *respectively. If $\gamma_1^t(z, y)$ is weakly continuous in z and y, $\gamma_2^t(z)$ is weakly continuous in z, $\hat{\hat{x}}(z, y)$ is continuous in z and y, $\hat{x}(z)$ is continuous in z, and, for every $z \in \mathbb{R}^m$, $\tilde{\beta}_z \neq 0$, there exists an optimal incentive strategy for the leader which is affine in v, and weakly continuous in z and y, and is given by*

$$\gamma_1^0(v, z, y) = \gamma_1^t(z, y) - \frac{\tilde{\beta}_z(y)}{\underset{y}{E}\{\|\tilde{\beta}_z(y)\|_u^2\}}\langle \tilde{v}_z, v - \gamma_2^t(z) \rangle_v. \tag{33}$$

Remark 2. An important observation that can be made from (33) is that the dynamic part of the leader's optimal policy depends not only on the common information z (about x) but also on the leader's "private" information y.

5. A scalar example. To illustrate Corollary 2, and especially the structural dependence of γ_1^0 on the common and private information (z and y), we consider in this section a structurally simple numerical example. Let $n = m = p = 1$, and $U = V = \mathbb{R}$. Let x, w_1 and w_2 be independent zero-mean Gaussian random variables with variance 1. Define $z = x + w_1$ and $\tilde{y} = x + w_2$, in which case

$$y = \tilde{y} - E[\tilde{y} | z] = \tilde{y} - \tfrac{1}{2}z.$$

Assume that $\gamma_1^t(z, y)$ and $\gamma_2^t(y)$ are in the structural form (where $\alpha_1, \alpha_2, \alpha_3$ are known scalars)

$$\gamma_1^t(z, y) = \alpha_1 z + \alpha_2 y, \qquad \gamma_2^t(z) = \alpha_3 z, \qquad \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R},$$

which would constitute a globally minimizing solution to a quadratic cost function for the leader.

Now let L_2 be given as

$$L_2(x, u, v) = \tfrac{1}{2}(u)^2 + uv + 2(v)^2 + ux + 2vx,$$

which is strictly convex in the pair (u, v). Then $\tilde{\beta}_z(y)$ and $\tilde{v}_z$ can be computed to be [from (30) and (31), respectively]

$$\tilde{\beta}_z(y) = \alpha_1 z + \alpha_2 y + \alpha_3 z + E[x | z, y] = (\alpha_1 + \alpha_3 + \tfrac{1}{2})z + (\alpha_2 + \tfrac{1}{3})y \triangleq \bar{\alpha}_1 z + \bar{\alpha}_2 y,$$

$$\tilde{v}_z = \underset{y|z}{E}[\alpha_1 z + \alpha_2 y] + 4\alpha_3 z + 2E[x | z] = (\alpha_1 + 4\alpha_3 + 1)z \triangleq \bar{\alpha}_3 z.$$

Under the parametric restriction $\alpha_2 \neq -\frac{1}{3}$, all the hypotheses of Corollary 2 are satisfied; and since

$$\underset{y|z}{E}\{[\tilde{\beta}_z(y)]^2\} = \bar{\alpha}_1^2 z^2 + \tfrac{3}{2}\bar{\alpha}_2^2,$$

an optimal incentive scheme for the leader is

$$\gamma_1^0(v, z, y) = \alpha_1 z + \alpha_2 y - \frac{(\bar{\alpha}_1 z + \bar{\alpha}_2 y)\bar{\alpha}_3 z}{\bar{\alpha}_1^2 z^2 + 3\bar{\alpha}_2^2/2}[v - \alpha_3 z]. \tag{34}$$

Note that this policy is nonlinear in the common information z, and the private information y also enters the gain term multiplying v. We show in Appendix 1, by direct verification, that (34) indeed constitutes an optimal incentive scheme for the leader, forcing the follower to the desired solution $\gamma_2^t(z) = \alpha_3 z$.

An important question that can be raised at this point is whether (34) constitutes a *unique* solution to the problem under consideration within the class of incentive schemes that are affine in v, or equivalently, whether the gain term in (34) solves (28) uniquely. We address this question in the sequel and show that the affine solution is *not* unique.

Towards this end, let us first assume that there is no private information for the leader, and $\gamma_1^t(z) = \alpha_1 z$. Then, an optimal solution can be obtained from Corollary 1 as

$$\gamma_1^0(v, z) = \alpha_1 z - \left(\frac{\bar{\alpha}_3}{\bar{\alpha}_1}\right)[v - \alpha_3 z] \tag{35}$$

which is in fact the unique one in the class of affine policies, and is linear also in the static information z, thus corroborating a result obtained in [16] for linear-quadratic problems with hierarchical decision structure. Now, if an additional information y comes in to the leader, which is statistically independent of the random variable z characterizing the common information, there seems, at the outset, no particular reason for the gain term in (35) to change, since v is measurable only with respect to the sigma field generated by z. Hence, intuitively, one expects the policy

$$\gamma_1^{00}(v, z, y) = \alpha_1 z + \alpha_2 y - \left(\frac{\bar{\alpha}_3}{\bar{\alpha}_1}\right)[v - \alpha_3 z] \tag{36}$$

to constitute an optimal incentive scheme when both z and y are acquired by the leader. This is indeed true, and the validity of this intuitive result has been established in Appendix 1 by showing that

$$\arg\min_{v} \underset{x,y|z}{E}\{L_2(x, \gamma_1^{00}(v, z, y), v)|z\} = \gamma_2^t(z).$$

Hence, the conclusion is that the scalar example of this section admits at least two affine optimal incentive schemes one of which is also linear in the static information (z, y).

6. Concluding remarks. By adopting a functional analytic framework, we have obtained optimal incentive strategies (for the leader) in a general class of hierarchical two-agent stochastic Stackelberg problems in which the leader has access to the follower's decision, to some common information, and also to some private information. The main conclusion of this analysis is that, under some fairly mild structural restrictions, there exists an optimal incentive policy for the leader, which is affine in the dynamic information and generally nonlinear in the static (common and private) information.

Even though we have used a general Hilbert space setting for the control (decision) spaces, we have assumed the random variables to take values in finite-dimensional spaces. We have chosen this framework in order to display salient features of the derivation without being distracted by the additional technical restrictions that would be required otherwise. However, our results (embodied in Propositions 1 and 2) are valid in a more general framework which allows the random variables to be weak random variables (cf. [24]) defined on (infinite-dimensional) Hilbert spaces, which includes, for example, the case of stochastic processes.

Appendix. In this appendix we show, by direct verification, that both γ_1^0 and γ_1^{00}, given by (34) and (35), respectively, solve the stochastic incentive problem of § 5.

Starting with the functional form

$$u=\alpha_1 z+\alpha_2 y-Q(z,y)[v-\alpha_3 z],$$

which is clearly a dynamic representation of the static policy $\gamma_1^t(z,y)$ at the desired equilibrium (γ_1^t,γ_2^t), we substitute this into $L_2(x,u,v)$ and take the expected value conditioned on z, with Q being an arbitrary function measurable in z and y. The result is the function

$$\begin{aligned} J(v,z)&=\underset{x,y|z}{E}\{\tfrac{1}{2}[u^t-Q(v-v^t)]^2+[u^t-Q(v-v^t)](v+x)+2v^2+2vx|z\}\\ &=\tfrac{1}{2}E\{Q^2|z\}(v-v^t)^2-E[u^tQ|z](v-v^t)\\ &\quad+\tfrac{1}{2}E\{u_t^2|z\}+\alpha_1 zv+\frac{\alpha_1}{2}z^2+E\{\alpha_2 yx|z\}\\ &\quad-E\{Q|z\}(v-v^t)v-E\{Qx|z\}(v-v^t)+2v^2+vz, \end{aligned}$$

where $u^t=\gamma_1^t(z,y)$, $v^t=\gamma_2^t(z)$.

Since $\frac{1}{2}E\{Q^2|z\}-E\{Q|z\}+2>0$ a.e. $\mathscr{P}_z$, $J(v,z)$ is strictly convex in v a.e. $\mathscr{P}_z$, and hence $v=v^t$ constitutes the unique minimizing solution to J if and only if $\partial J(v^t,z)/\partial v=0$ a.e. $\mathscr{P}_z$. This leads to the following equation to be satisfied by $Q(y,z)$:

$$\text{(A1)}\qquad \begin{aligned} &[\alpha_1-E[Q(y,z)|z](\alpha_3+\alpha_1)+4\alpha_3+1]z\\ &\qquad -\alpha_2 E[yQ(y,z)|z]-E[xQ(y,z)|z]=0. \end{aligned}$$

Let us now consider the following two choices for Q:

1) $Q(y,z)=\bar\alpha_3/\bar\alpha_1$;

2) $Q(y,z)=(\bar\alpha_1 z+\bar\alpha_2 y)\bar\alpha_3 z/[\bar\alpha_1^2 z^2+3\bar\alpha_2^2/2]$.

In the former case, (A1) reads

$$\left[\alpha_1-\frac{\bar\alpha_3(\alpha_3+\alpha_1)}{\bar\alpha_1}+4\alpha_3+1-\frac{\bar\alpha_3}{2\bar\alpha_1}\right]z=0,$$

which can easily be shown to be an identity, by making use of the definitions of $\bar\alpha_1$ and $\bar\alpha_3$. Hence γ_1^{00} given by (35) is indeed an optimal incentive scheme.

In the latter case, (A-1) reads

$$\left[\bar\alpha_3-(\bar\alpha_1-\tfrac{1}{2})\frac{\bar\alpha_1\bar\alpha_3 z^2}{\bar\alpha_1^2 z^2+3\bar\alpha_2^2/2}\right]z-\frac{3\alpha_2\bar\alpha_2\bar\alpha_3 z}{2\bar\alpha_1^2 z^2+3\bar\alpha_2^2}-E[xQ(y,z)\,|\,z]=0$$

$$\Leftrightarrow\frac{\bar\alpha_3[\bar\alpha_2 z+\bar\alpha_1 z^3]}{2\bar\alpha_1^2 z^2+3\bar\alpha_2^2}-\frac{\bar\alpha_3\bar\alpha_1 z^3}{2\bar\alpha_1^2 z^2+3\bar\alpha_2^2}-\frac{\bar\alpha_2\bar\alpha_3 z}{2\bar\alpha_1^2 z^2+3\bar\alpha_2^2}=0$$

since $E[x\,|\,z]=\frac{1}{2}z$ and $E[xy\,|\,z]=\frac{1}{2}$. The latter equation is an identity, thus corroborating the optimality of the incentive strategy γ_1^0 given by (34).

REFERENCES

[1] H. VON STACKELBERG, *Markform und Gleichgewicht*, Springer, Vienna, 1934: or *The Theory of the Market Economy*, Oxford Univ. Press, Oxford, 1952.

[2] C. I. CHEN AND J. B. CRUZ, JR., *Stackelberg solution for two-person games with biased information patterns*, IEEE Trans. Automat. Control, AC-17 (1972), pp. 791–798.

[3] M. SIMAAN AND J. B. CRUZ, JR., *On the Stackelberg strategy in nonzero-sum games*, J. Optim. Theory Appl., 11 (1973), pp. 533–555.

[4] ———, *Additional aspects of the Stackelberg strategy in nonzero-sum games*, J. Optim. Theory Appl., 11 (1973), pp. 613–626.

[5] J. B. CRUZ, JR., *Leader-follower strategies for multilevel systems*, IEEE Trans. Automat. Control, AC-23 (1978), pp. 244–255.

[6] T. BAŞAR AND H. SELBUZ, *Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems*, IEEE Trans. Automat. Control, AC-24 (1979), pp. 166–179.

[7] Y. C. HO, P. B. LUH AND R. MURALIDHARAN, *Information structure, Stackelberg games and incentive controllability*, IEEE Trans. Automat. Control, AC-26 (1981), pp. 454–460.

[8] Y. C. HO, P. B. LUH AND G. J. OLSDER, *A control theoretic view on incentives*, Automatica, 18 (1982), pp. 167–179.

[9] G. P. PAPAVASSILOPOULOS AND J. B. CRUZ, JR., *Nonclassical control problems and Stackelberg games*, IEEE Trans. Automat. Control, 24 (1979), pp. 155–166.

[10] B. TOLWINSKI, *Closed-loop Stackelberg solution to multi-stage linear-quadratic game*, J. Optim. Theory Appl., 34 (1981), pp. 485–501.

[11] G. P. PAPAVASSILOPOULOS AND J. B. CRUZ, JR., *Sufficient conditions for Stackelberg and Nash strategies with memory*, J. Optim. Theory Appl., 31 (1980), pp. 233–260.

[12] T. BAŞAR, *A general theory for Stackelberg games with partial state information*, Large Scale Systems, 3 (1982), pp. 47–56. An earlier version appeared in Proc. 4th International Conference on the Analysis and Optimization of Systems, Springer, 1980, New York.

[13] ———, *Equilibrium strategies in dynamic games with multilevels of hierarchy*, Automatica, 17 (1981), pp. 749–754.

[14] ———, *Performance bounds for hierarchical systems under partial dynamic information*, J. Optim. Theory Appl., 39 (1983), pp. 67–87.

[15] Y. P. ZHENG AND T. BAŞAR, *Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach*, Int. J. Control, 35 (1982), pp. 997–1011.

[16] T. BAŞAR, *Hierarchical decision making under uncertainty*, in Dynamic Optimization and Mathematical Economics, P. T. Liu, ed., Plenum, New York, 1980, pp. 205–221.

[17] ———, *Stochastic multicriteria decision problems with multi levels of hierarchy*, IEEE Trans. Automat. Control, AC-26 (1981), pp. 549–553.

[18] T. S. CHANG AND Y. C. HO, *Incentive problems: A class of Stackelberg closed-loop dynamic games*, Systems and Control Letters (1981), pp. 16–21.

[19] T. BAŞAR AND J. B. CRUZ, JR., *Concepts and methods in multiperson coordination and control*, in Optimization and Control of Dynamic Operational Research Models, S. G. Tzafestas, ed., North-Holland, Amsterdam, 1982, Ch. 11, pp. 351–387.

[20] T. GROVES AND M. LOEB, *Incentives in a divisionalized firm*, Management Sci., 25 (1979), pp. 221–230.

[21] L. HURWICZ AND L. SHAPIRO, *Incentive structures maximizing residual gain under incomplete information*, Bell J. Economics, 9 (1978), pp. 180–191.

[22] L. P. JENNERGREN, *On the design of incentives in business firms—A survey of some recent research*, Management Sci., 26 (1980), pp. 180–201.

[23] M. SALMAN AND J. B. CRUZ, JR., *An incentive model of duopoly with government coordination*, Automatica, 17 (1981), pp. 821–830.

[24] A. V. BALAKRISHNAN, *Applied Functional Analysis*, Springer, New York, 1976.

[25] D. G. LUENBERGER, *Optimization by Vector Space Methods*, John Wiley, New York, 1969.

[26] T. BAŞAR AND G. J. OLSDER, *Dynamic Noncooperative Game Theory*, Academic Press, London, 1982.