WORKING PAPERS

Ngo Van Long
Gerhard Sorger

A dynamic principal-agent problem as a feedback Stackelberg differential game

April 2009

Working Paper No: 0905

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

A dynamic principal-agent problem as a feedback Stackelberg differential game

Ngo Van LONG

Department of Economics, McGill University, Montreal, Canada,

Gerhard SORGER*

Department of Economics, University of Vienna, Vienna, Austria.

*Corresponding author: Department of Economics, University of Vienna, Hohenstaufengasse 9, A-1010 Vienna, AUSTRIA. Phone: +43 1 4277 37443. Fax: +43 1 4277 9374. Email: gerhard.sorger@univie.ac.at

Abstract: We consider situations in which a principal tries to induce an agent to spend effort on accumulating a state variable that affects the well-being of both parties. The only incentive mechanism that the principal can use is a state-dependent transfer of her own utility to the agent. Formally, the model is a Stackelberg differential game in which the players use feedback strategies. Whereas in general Stackelberg differential games with feedback strategy spaces the leader's optimization problem has non-standard features that make it extremely hard to solve, in the present case this problem can be rewritten as a standard optimal control problem. Two examples are used to illustrate our approach.

Journal of Economic Literature **classification codes:** C61, C73, D82

Key words: Differential game, principal-agent model, feedback strategies, Stackelberg equilibrium.

1 Introduction

Differential games in which the players use feedback strategies are hard to solve for Stackelberg equilibria. Basar and Olsder (1982) already noted that "such decision problems cannot be solved by utilizing standard techniques of optimal control theory [...] because the reaction set of the follower cannot, in general, be determined in closed form, for all possible strategies of the leader, and hence the optimization problem faced by the leader on this reaction set becomes quite an implausible one" (page 315). For stochastic games, the difficulties are even more severe so that Basar and Olsder (1982) concluded that the "derivation of the Stackelberg solution in stochastic dynamic games meets with unsurmountable difficulties" (page 336) and "remains, today, as a challenge for the researchers" (page 337).[1]

More than 25 years have passed since the publication of the classical text by Basar and Olsder (1982) but, still, the literature on feedback Stackelberg equilibria in differential games is very sparse. The authors of one of the more recent textbooks on the subject, Dockner et al. (2000), still admit that "the analysis of such an equilibrium in a differential game may lead to considerable technical difficulties" (page 134). In order to circumvent these difficulties, they suggest to restrict the class of strategies of the leader to a finite-dimensional space. Under such a restriction, the leader solves a finite-dimensional (static) optimization problem instead of the dynamic and, hence, infinite-dimensional problem that Basar and Olsder (1982) have referred to as "quite implausible". This approach, which has been used for example by Benchekroun and Long (1998, 2002), is not unproblematic as it may lead to a solution that differs substan-

[1] Unfortunately, the literature does not use a common terminology. Basar and Olsder (1982), in particular, distinguish between the stagewise Stackelberg equilibrium and the global Stackelberg equilibrium. It is the latter that we have in mind when we speak about Stackelberg equilibria.

tially from the true feedback Stackelberg equilibrium, or it may produce a solution in cases where a true feedback Stackelberg equilibrium does not even exist. In general, non-existence of feedback Stackelberg equilibria cannot be ruled out even in simple problems satisfying standard smoothness and curvature properties. All these issues have recently been illustrated by Shimomura and Xie (2008) in the context of common-property resource games. Nevertheless, in some cases it is possible to rewrite the leader's optimization problem in such a way that it becomes amenable to standard optimal control techniques. The main purpose of the present paper is to introduce one class of models in which such a transformation works, and to illustrate the solution method by means of two examples.

The models considered in the present paper can most easily be interpreted as dynamic principal-agent situations. The standard textbook principal-agent problem is set in a static framework; see, e.g., chapter 14 in Mas-Collel et al. (1995). The principal offers a take-it-or-leave-it contract to the agent. The contract is designed so as to induce the agent to perform a task according to the preferences of the principal. In return for his performance, the agent receives a contractual payment from the principal which is conditional on some observable and verifiable outcome of his activities. The payment scheme must be sufficiently attractive to ensure that the agent accepts the contract instead of taking an outside option. In the present paper we assume that the task to be performed by the agent is to contribute to the control of an endogenous state variable over an infinite time-horizon. This adds a crucial dynamic feature to the standard (static) principal-agent model: in designing the optimal contract, the principal must take into account the endogenously evolving state of the system. In other words, the principal has to decide on how to feed the state variable back into the agent's optimal control problem throughout the infinite time horizon. It is exactly this feature that makes the dynamic optimization problem of the principal a non-standard one, as described in the above quotes from Basar and Olsder

(1982).

We start with a rather general setup for our analysis. We assume that both players (the principal and the agent) can affect the dynamics of the state variable. More specifically, the evolution of the state is described by a linear stochastic differential equation the stochastic component of which is a standard Wiener process (Brownian motion). The two utility functions by which the principal and the agent, respectively, evaluate the state of the system can be arbitrary (smooth) functions. The costs of applying effort to control the state are also assumed be general (smooth and strictly convex) functions of the respective effort levels. Finally, the two players may have different time-preference rates. There is, however, one important structural assumption that we have to impose on the model in order to obtain analytical tractability: we assume that the contractual payment of the principal to the agent takes the form of a direct transfer of utility. This assumption together with the participation constraint will be shown to render the optimization problem of the principal (leader) tractable in several important cases.

In order to cast this dynamic principal-agent model into a differential game framework with feedback strategies, we assume that the current value of the state variable can be observed by both players and by the courts, but that it cannot be memorized. As for the agent's effort, we follow the principal-agent literature by assuming that it cannot be observed or verified. Thus, contracts can only be made contingent on the current value of the state variable as well as on the commonly known structure of the model. In technical terms, the principal's contract must be formulated in terms of feedback strategies. The hierarchical decision making process according to which the principal announces the contract before the agent decides on whether to accept it or not (and, if he does accept, on how much effort to expend) is then appropriately reflected by the Stackelberg equilibrium concept.

There have been several attempts to formulate dynamic principal-agent problems. The first paper is that of Holmstrom and Milgrom (1987). They assume (p. 308) that (i) the agent evaluates wealth at a single point in time (the terminal time) after all the actions have been taken, (ii) the cost of actions can be expressed in monetary units, (iii) the utility functions of the principal and the agent are both exponential, and (iv) neither party has any private information at the time of signing the contract. The agent's strategy at any time can be dependent on the whole history of observations up to that time. The payment the principal makes to the agent takes place only in the last period, and is dependent on the whole history of observations. The costs are incurred in each period, but they are summed up (without discounting) in the last period, and this sum is valued only in the last period. Under these assumptions, they show that the optimal compensation rule is linear in profit. Their linearity result has been generalized by Schattler and Sung (1993) and Sung (1995). Clearly, the formulation chosen by Holmstrom and Milgrom (1987) requires a high degree of commitment and memory that we do not allow in our formulation.

A recent paper by Sannikov (2008) studies a problem similar to ours. In his formulation, the principal does not exercise effort and has a utility function that is linear in profit. The agent gets paid continuously and his utility is strictly concave in the payment. Both utility functions do not contain the state variable as an argument. Our model is more general in several respects: we allow the utility functions to depend on the state variable, we allow the principal to exercise effort, and we allow the rate of discount of the principal to be different from that of the agent.

Another type of dynamic extension of the principal-agent problem is concerned with the ratchet-effect problem: if in period 1 an agent reveals his type, the principal may become fully informed in period 2 and thus, in the absence of pre-commitment, she will be able to press the agent to his reservation utility level in period 2. Knowing this, the agent may not want to reveal his type in period 1. This type of problem has been studied for example by Laffont and Tirole (1988). One dynamic extension that avoids the ratchet effect problem involves the situation where the agent's characteristics change every period in a random and independent way as in the papers by Gaudet, Lasserre, and Long (1995, 1996, 1998). In Gaudet, Lasserre, and Long (1995), for example, the dynamic element consists of changes in the value of an exhaustible resource stock. The principal is the government that owns the resource stock and wants to extract royalty payments from the resource-extracting firms, whose cost characteristics are private information. Gaudet, Lasserre, and Long (1995) solve for a time-dependent Stackelberg equilibrium and derive the optimal resource royalty scheme for each period.

The rest of the paper is organized as follows. In section 2 we describe the model. Section 3 considers the agent's optimization problem. Given the principal's contract, this is a standard optimal control problem that can be solved by the maximum principle or by dynamic programming. A crucial step in our analysis is to represent the optimality conditions in the form of an implementability condition for the principal. In doing this we follow the approach taken by Chang (1988) or, more recently, by Shimomura and Xie (2008). The implementability condition then becomes a constraint for the principal's optimization problem, which we discuss in section 4. The main result of that section is the reformulation of this non-standard optimization problem. There are two cases in which the reformulation yields indeed a standard optimal control problem: the case of equal time-preference rates and the deterministic case. Section 5 discusses the issues of time-consistency and Markov-perfection of the equilibrium. The equilibria that we discuss typically do not have these properties and we explain that requiring them to hold would be a very strong restriction. In sections 6 and 7 we discuss examples of our general model in which we can easily solve for feedback Stackelberg equilibria. Section 6 deals with a linear model in which both players choose to apply constant effort in equilibrium, and section 7 considers a linear-quadratic setup in which the equilibrium effort strategies turn out to be linear functions of the state. Finally, section 8 presents concluding remarks and describes directions for future research.

2 Model formulation

Time is measured continuously and extends from 0 to $+\infty$. The state of the system at time t is denoted by $x(t) \in \mathbb{R}$. The state can be interpreted as a stock that potentially affects the well-being of two players, which we call the principal and the agent, respectively. The initial stock $x(0) = x_0 \in \mathbb{R}$ is a given parameter. The state variable evolves according to the stochastic differential equation

$$\mathrm{d}x(t) = [\alpha u_P(t) + u_A(t) - \delta x(t)]\ \mathrm{d}t + \sigma(x(t))\,\mathrm{d}w(t)\,,\ x(0) = x_0, \tag{1}$$

where α and δ are non-negative constants, and where $u_P(t) \in \mathbb{R}$ and $u_A(t) \in \mathbb{R}$ are control variables of the principal and the agent, respectively. While the agent can always influence the evolution of the state variable, this is the case for the principal if and only if $\alpha > 0$. The case $\alpha = 0$, on the other hand, reflects a situation in which the principal herself has no control over the state of the system. Finally, the process w is a standard Wiener process (Brownian motion) and $\sigma : \mathbb{R} \mapsto [0, +\infty)$ is a continuously differentiable function.

Applying effort to control the state causes flow costs $C_P(u_P(t))$ and $C_A(u_A(t))$ to the principal and the agent, respectively. It is assumed that both $C_P : \mathbb{R} \mapsto \mathbb{R}$ and $C_A : \mathbb{R} \mapsto \mathbb{R}$ are twice continuously differentiable, strictly convex, and satisfy $C_P(0) = C_A(0) = C'_P(0) = C'_A(0) = 0$. The principal has the additional possibility of transferring utility to the agent, possibly with a distortion. More specifically, we assume that the principal has a second control variable $v(t)$ with the interpretation that $\beta v(t)$ units of the principal's utility must be sacrificed in order to increase the agent's utility flow by $v(t)$ units. The parameter β can be any strictly positive number. Finally, it is assumed that the principal derives utility $F_P(x(t))$ from the stock, whereas the agent's utility from the stock is $F_A(x(t))$. The utility functions $F_P : \mathbb{R} \mapsto \mathbb{R}$ and $F_A : \mathbb{R} \mapsto \mathbb{R}$ are assumed to be continuously differentiable. The two players maximize their expected life-time utility and have time-preference rates ρ_P and ρ_A, respectively, both assumed to be strictly positive. Taking all of these assumptions into account, we can write the principal's and the agent's objective functionals as

$$J_P = \mathbb{E}_0 \int_0^{+\infty} e^{-\rho_P t} \left[F_P(x(t)) - C_P(u_P(t)) - \beta v(t) \right] \, \mathrm{d}t \tag{2}$$

and

$$J_A = \mathbb{E}_0 \int_0^{+\infty} e^{-\rho_A t} \left[F_A(x(t)) - C_A(u_A(t)) + v(t) \right] \, \mathrm{d}t, \tag{3}$$

respectively, where $\mathbb{E}_0$ denotes the expectation conditional on information available at time 0.

All parameters of the model (including the initial state x_0) as well as the functions C_P, C_A, F_P, F_A, and σ are assumed to be common knowledge. The state variable $x(t)$ is observable by both players and by the courts. However, neither the players nor the courts can keep a memory of the state (except for its initial value x_0). Enforceable contracts between the two parties can therefore only be based on the commonly known structure of the model and on the current value of the state variable.

At time 0 the principal proposes a contract (U_P, V), where $U_P : \mathbb{R} \mapsto \mathbb{R}$ and $V : \mathbb{R} \mapsto \mathbb{R}$ map the state space into the respective control spaces. The interpretation of such a contract is that the principal promises to spend own effort according to $u_P(t) = U_P(x(t))$ and to transfer utility according to $v(t) = V(x(t))$ provided that the agent satisfies his contractual duties.[2] The agent is assumed to have an outside option that yields utility $\bar{J}_A$. Without loss of generality we set $\bar{J}_A = 0$. Also at time 0, but after the principal's announcement of the contract, the agent has to decide whether to accept the principal's contract or not. If he accepts the contract, he foregoes his access to the outside option and chooses a function $U_A : \mathbb{R} \mapsto \mathbb{R}$ which determines his contribution according to $u_A(t) = U_A(x(t))$. We assume that contracts can be fully enforced by the courts. This requires that the courts can observe the principal's effort $u_P(t)$, her transfer $v(t)$, and the agent's compliance with the requirement not to use the outside option. In other words, once the agent has agreed to the principal's contract, the principal is committed to her strategies (U_P, V) and the agent has no access to his outside option. Note that both the contract (U_P, V) and the agent's strategy U_A may depend on the initial state x_0. We are interested in a hierarchical Stackelberg equilibrium. This means that the agent chooses his strategy U_A as a best response to the principal's strategies (U_P, V), and that the principal takes this best response fully into account when she optimally selects her own strategies (U_P, V).

For technical reasons we have to assume that the triple of strategies (U_P, U_A, V) satisfies certain restrictions. First, we assume that U_A is twice continuously differentiable and that U_P and V are continuously differentiable. We also assume that U_P and U_A are such that the stochastic differential equation that results from (1) by replacing the control variables $u_P(t)$ and $u_A(t)$ by $U_P(x(t))$ and $U_A(x(t))$, respectively, has a unique solution and that along this solution it holds that

$$\lim_{t \to +\infty} \mathbb{E}_0 \left\{ e^{-\rho_P t} \int_{x_0}^{x(t)} C_A'(U_A(z))\,\mathrm{d}z \right\} = 0. \tag{4}$$

Condition (4) is hard to interpret as it jointly involves the equilibrium state trajectory x, the

[2]As will be explained in a moment, the agent's contractual duties consist merely of being 'on the job' or 'on site', i.e., not taking his outside option. This follows from the fact that the agent's effort is not observable.

principal's time-preference rate ρ_P, and the agent's marginal cost function C'_A and effort strategy U_A. Essentially, however, it is a boundedness assumption. For example, in the deterministic case with $\sigma(x) = 0$ for all x, a sufficient but by no means necessary condition for (4) to hold is that $x(t)$ remains bounded. This suggests that (4) will typically hold if the marginal benefits of the state $F'_P(x)$ and $F'_A(x)$ for large values of x are small compared to the marginal costs of maintaining the state at x.[3]

3 The agent's optimization problem

Consider any given pair of the principal's strategies (U_P, V). In this section we describe the agent's optimal response to (U_P, V). The agent seeks to maximize

$$J_A = \mathbb{E}_0 \int_0^{+\infty} e^{-\rho_A t} \left[F(x(t)) - C_A(u_A(t))\right] \, \mathrm{d}t$$

subject to

$$\mathrm{d}x = \left[f(x(t)) + u_A(t)\right] \, \mathrm{d}t + \sigma(x(t)) \, \mathrm{d}w \,, x(0) = x_0,$$

where $F(x) = F_A(x) + V(x)$ and $f(x) = \alpha U_P(x) - \delta x$. This is a standard stochastic optimal control problem amenable to solution by the dynamic programming approach. Denoting the optimal value function by W_A, the Hamilton-Jacobi-Bellman equation of this problem is given by

$$\rho_A W_A(x) = \max \left\{ F(x) - C_A(u_A) + W'_A(x)[f(x) + u_A] + (1/2)W''_A(x)\sigma(x)^2 \, \middle| \, u_A \in \mathbb{R} \right\}. \quad (5)$$

Due to our assumptions on the cost function C_A, the right-hand side of this equation has a unique maximizer u_A determined by the first-order condition $C'_A(u_A) = W'_A(x)$. It follows that

[3]In section 4 below, the technical reasons for imposing equation (4) will become clear.

the agent's optimal strategy must satisfy

$$C_A'(U_A(x)) = W_A'(x). \tag{6}$$

This equation has the simple economic interpretation that the marginal cost to the agent of exerting effort is equal to the marginal contribution of this effort to his own stock value. Integrating condition (6) with respect to the state variable one obtains

$$W_A(x) = \int_{x_0}^{x} C_A'(U_A(z))\,\mathrm{d}z + W_A(x_0). \tag{7}$$

Finally, by substituting $u_A = U_A(x)$, (6), and (7) back into the Hamilton-Jacobi-Bellman equation (5) we get

$$\begin{aligned} & \rho_A \int_{x_0}^{x} C_A'(U_A(z))\,\mathrm{d}z + \rho_A W_A(x_0) \\ = \; & F(x) - C_A(U_A(x)) + C_A'(U_A(x))[f(x) + U_A(x)] + (1/2)C_A''(U_A(x))U_A'(x)\sigma(x)^2. \end{aligned} \tag{8}$$

Now recall that the agent has an outside option that yields overall utility equal to 0. The agent will therefore only accept the principal's contract if the condition $W_A(x_0) \geq 0$ is satisfied. Moreover, in equilibrium this participation constraint must be binding because, if it were not, the principal could gain from replacing V by V_ε, where $V_\varepsilon(x) = V(x) - \varepsilon$ and where ε is a sufficiently small positive number. As a matter of fact, because V and V_ε differ only by a deterministic constant, both contracts (U_P, V) and (U_P, V_ε) provide the same incentives to the agent, but the principal's expected utility under (U_P, V_ε) exceeds her expected utility under (U_P, V) by $\beta\varepsilon/\rho_P > 0$. To summarize, in equilibrium it must hold that $W_A(x_0) = 0$ and it follows therefore from (8) and the definitions of F and f that

$$\begin{aligned} V(x) \;=\; & -F_A(x) + C_A(U_A(x)) + \rho_A \int_{x_0}^{x} C_A'(U_A(z))\,\mathrm{d}z \\ & -C_A'(U_A(x))[\alpha U_P(x) + U_A(x) - \delta x] - (1/2)C_A''(U_A(x))U_A'(x)\sigma(x)^2. \end{aligned} \tag{9}$$

Equation (9) is an implementability condition very much in the spirit of Chang (1988) and Shimomura and Xie (2008). It is a condition that the principal's contract (U_P, V) must satisfy in order for the strategy U_A to be optimal for the agent. In the next section we shall use this condition to formulate the principal's optimization problem.

In order to interpret condition (9) in economic terms, we use (6) and (7) to rewrite it as

$$V(x) = C_A(U_A(x)) - F_A(x) + \rho_A W_A(x) - W_A'(x)[\alpha U_P(x) + U_A(x) - \delta x] - (1/2)W_A''(x)\sigma(x)^2.$$

This says that the transfer the agent receives (when the observed state is x) is sufficient to compensate his effort cost, $C_A(U_A(x))$, net of the current benefits that he derives from the stock, $F_A(x)$, net of the 'interest cost' as perceived by the agent, $-\rho_A W_A(x)$, and net of the 'investment value'. The latter consists of a term corresponding to the deterministic component in the state equation (1) as well as an adjustment for risk.

4 The principal's optimization problem

As has been explained in section 2, the principal takes the optimal response of the agent into account when she chooses her optimal contract (U_P, V). The optimal response of the agent, in turn, has implicitly been described by the implementability condition (9). Note in particular that, for any given pair (U_P, U_A), condition (9) determines a unique transfer strategy V. It is therefore possible to reformulate the principal's optimization problem as follows: the principal seeks to find a pair of effort strategies (U_P, U_A) such that her objective functional in (2) is maximized subject to the constraints $u_P(t) = U_P(x(t))$, $u_A(t) = U_A(x(t))$, $v(t) = V(x(t))$, (1), and (9). Unfortunately, this is not a standard optimal control problem because the implementability condition contains both derivatives and integrals of the strategy U_A. Consequently, standard optimization techniques like the maximum principle or the Hamilton-Jacobi-Bellman equation cannot be directly applied. In the rest of this section we will therefore rewrite the principal's optimization problem in a more convenient form. In particular, we shall identify two important special cases of the model in which the principal's optimization problem can indeed be written as a standard optimal control problem. In one of these case, the two players have the same time-preference rate, and in the other case there is no uncertainty.

Using the constraints $v(t) = V(x(t))$ and (9) we can write the objective functional from (2) as

$$
\begin{aligned}
J_P &= \mathbb{E}_0 \int_0^{+\infty} e^{-\rho_P t} \Big\{ F_P(x(t)) + \beta F_A(x(t)) + \beta C_A'(U_A(x(t)))[\alpha U_P(x(t)) + U_A(x(t)) - \delta x(t)] \\
&\quad + (\beta/2) C_A''(U_A(x(t))) U_A'(x(t)) \sigma(x(t))^2 - C_P(U_P(x(t))) - \beta C_A(U_A(x(t))) \Big\} \, \mathrm{d}t \qquad (10) \\
&\quad - (\beta \rho_A / \rho_P) \mathbb{E}_0 \left\{ \rho_P \int_0^{+\infty} D(x(t), t) \, \mathrm{d}t \right\},
\end{aligned}
$$

where

$$
D(x,t) = e^{-\rho_P t} \int_{x_0}^{x} C_A'(U_A(z)) \, \mathrm{d}z.
$$

Note that the right-hand side of equation (10) consists of two terms. The rate of impatience of the agent appears only in the second term. We claim that

$$
\begin{aligned}
\mathbb{E}_0 \left\{ \rho_P \int_0^{+\infty} D(x(t), t) \, \mathrm{d}t \right\} &= \mathbb{E}_0 \int_0^{+\infty} e^{-\rho_P t} \Big\{ C_A'(U_A(x(t)))[\alpha U_P(x(t)) + U_A(x(t)) - \delta x(t)] \\
&\quad + (1/2) C_A''(U_A(x(t))) U_A'(x(t)) \sigma(x(t))^2 \Big\} \, \mathrm{d}t. \qquad (11)
\end{aligned}
$$

As a matter of fact, by Ito's lemma we have

$$
\begin{aligned}
\rho_P D(x(t), t) \, \mathrm{d}t &= -\, \mathrm{d}D(x(t), t) + e^{-\rho_P t} C_A'(U_A(x(t))) \sigma(x(t)) \, \mathrm{d}w(t) \qquad (12) \\
&\quad + e^{-\rho_P t} \Big\{ C_A'(U_A(x(t)))[\alpha U_P(x(t)) + U_A(x(t)) - \delta x(t)] \\
&\quad + (1/2) C_A''(U_A(x(t))) U_A'(x(t)) \sigma(x(t))^2 \Big\} \, \mathrm{d}t.
\end{aligned}
$$

Because of $D(x_0,0)=0$ and (4) we obtain

$$-\mathbb{E}_0\int_0^{+\infty} \mathrm{d}D(x(t),t) = -\lim_{t\to+\infty}\mathbb{E}_0\left\{e^{-\rho_P t}\int_{x_0}^{x(t)} C_A'(U_A(z))\,\mathrm{d}z\right\} + D(x_0,0) = 0.$$

Because w is a standard Wiener process we have

$$\mathbb{E}_0\int_0^{+\infty} e^{-\rho_P t}C_A'(U_A(x(t)))\sigma(x(t))\,\mathrm{d}w(t) = 0.$$

Using the last two identities, integrating equation (12) from $t=0$ to $+\infty$, and forming expectations $\mathbb{E}_0$ we get (11).

Finally, we substitute (11) into (10). After some rearrangement and using $U_P(x(t)) = u_P(t)$ and $U_A(x(t)) = u_A(t)$, this yields

$$\begin{aligned} J_P \;=\; & \mathbb{E}_0\int_0^{+\infty} e^{-\rho_P t}\Big\{F_P(x(t)) + \beta F_A(x(t)) - C_P(u_P(t)) - \beta C_A(u_A(t)) \\ & +[\beta(\rho_P-\rho_A)/\rho_P]C_A'(u_A(t))[\alpha u_P(t) + u_A(t) - \delta x(t)] \\ & +[\beta(\rho_P-\rho_A)/(2\rho_P)]C_A''(u_A(t))U_A'(x(t))\sigma(x(t))^2\Big\}\,\mathrm{d}t. \end{aligned} \tag{13}$$

Notice that the principal's objective function, after taking into account the implementability constraint, consists of three terms: (i) the first term is the weighted sum of utilities of the principal and the agent (with weights 1 and β, respectively), (ii) the second term is non-zero only if the two rates of impatience, ρ_P and ρ_A, differ from each other, and (iii) the third term is non-zero only if in addition to $\rho_P \neq \rho_A$ the function σ is not identically equal to 0 . The only non-standard term in this objective functional is the third one as it contains the derivative of the strategy U_A. We summarize our findings in the following result.

Theorem 1 *(a) If $\rho_P = \rho_A = \rho$, then the principal's optimization problem is to maximize*

$$J_P = \mathbb{E}_0\int_0^{+\infty} e^{-\rho t}\left\{F_P(x(t)) + \beta F_A(x(t)) - C_P(u_P(t)) - \beta C_A(u_A(t))\right\}\,\mathrm{d}t$$

subject to the state equation (1).

(b) If $\sigma(x) = 0$ holds for all x, then the principal's optimization problem is to maximize

$$J_P = \int_0^{+\infty} e^{-\rho_P t} G(x(t), u_P(t), u_A(t))\,\mathrm{d}t$$

subject to the state equation

$$\dot{x}(t) = \alpha u_P(t) + u_A(t) - \delta x(t)\,, x(0) = x_0,$$

where

$$\begin{aligned} & G(x, u_P, u_A) \qquad (14) \\ = \;& F_P(x) + \beta F_A(x) - C_P(u_P) - \beta C_A(u_A) + [\beta(\rho_P - \rho_A)/\rho_P] C'_A(u_A)[\alpha u_P + u_A - \delta x]. \end{aligned}$$

(c) In all other cases, the principal's optimization problem is to maximize J_P given in (13) subject to $u_A(t) = U_A(x(t))$ and the state equation (1).

PROOF: The result follows immediately from (13). □

We shall illustrate the application of theorem 1 by means of two examples in sections 6 and 7 below. In this regard we note that the integrand in part (a) of the theorem is strictly concave with respect to (u_P, u_A) due to our assumptions on the cost functions C_P and C_A. In the case of part (b), the integrand is not necessarily strictly concave unless one imposes additional assumptions on the cost functions. This will also be illustrated below.

The case of equal time-preference rates $\rho_P = \rho_A$ (case (a) of the above theorem) has the interesting feature that the principal's objective functional is simply a weighted average of the two players' net utility functions, whereby 'net' refers to the absence of the transfer. Thus, the principal selects a Pareto efficient pair of effort strategies (U_P, U_A) and designs the transfer V

in such a way that the agent is induced to implement U_A. Which Pareto-efficient pair (U_P, U_A) is chosen by the principal depends on the size of the distortion β. The more costly it is to the principal to transfer one unit of utility to the agent, the more weight the principal gives to the agent's preferences.

In the general case $\rho_P \neq \rho_A$ the principal's objective functional is no longer a weighted average of the two players' net utility functions, as can be seen, for example, from equation (14) in part (b) of the theorem.

5 Markov-perfection and time-consistency

In the discussion so far we have assumed that both players are bound by the contract once they have signed it: the principal must satisfy her contractual agreement (U_P, V) and the agent has no access to his outside option. In other words, there is no possibility of renegotiation at any point in time $t > 0$ or in any state $x \neq x_0$. In this section we shall discuss if and when the players have incentives to renegotiate the contract, and we shall analyze the implications of allowing renegotiations to take place.

To discuss this point recall that the rationality of the principal together with the participation constraint for the agent requires $W_A(x_0) = 0$. If for some $t > 0$ it holds that $W_A(x(t)) < 0$, then it follows that the agent can benefit from terminating the contract with the principal at time t and switching to his outside option that gives him utility 0. If, on the other hand, $W_A(x(t)) > 0$ is satisfied for some t, then the principal can benefit from terminating the contract and setting up a new one that reduces the agent's continuation value to 0. In the examples presented in sections 6 and 7 below, both of these situations will be seen to occur generically (which one occurs depends on the parameter values as well as on the realizations of the stochastic disturbances). This shows that the equilibrium is generically not time-consistent.

The incentives for termination or renegotiation of the contract that have been discussed in the previous paragraph are ruled out if it can be ensured that along the equilibrium path $W_A(x(t)) = 0$ holds for all t. In this situation we call the equilibrium time-consistent. An even stronger requirement would be that $W_A(x) = 0$ holds for all $x \in \mathbb{R}$. This requirement, which ensures that neither party has an incentive to terminate the present arrangement in any possible state of the system (even off the equilibrium path), is called Markov-perfection. In what follows we shall first discuss Markov-perfection for the general stochastic case and then consider time-consistency for the special case of no uncertainty.

Theorem 2 *In a Markov-perfect equilibrium, it holds that $U_A(x) = 0$ and $V(x) = -F_A(x)$ for all $x \in \mathbb{R}$. Furthermore, the strategy U_P is determined as the optimal policy function of the following stochastic optimal control problem: maximize*

$$J_P = \mathbb{E}_0 \int_0^{+\infty} e^{-\rho_P t} \left[F_P(x(t)) + \beta F_A(x(t)) - C_P(u_P(t))\right] \mathrm{d}t$$

subject to

$$\mathrm{d}x(t) = [\alpha u_P(t) - \delta x(t)]\,\mathrm{d}t + \sigma(x(t))\,\mathrm{d}w(t).$$

Proof: Because of our assumptions on the cost function C_A, the condition $W_A(x) = 0$ for all $x \in \mathbb{R}$ translates directly into $U_A(x) = 0$ and $V(x) = -F_A(x)$; see equations (6) and (9). Using these observations it is easily seen from (13) and (1) that the principal faces the stochastic optimal control problem stated in the present theorem. This completes the proof. $\square$

It can be seen from the above theorem that the requirement of Markov-perfection imposes very strong restrictions on the principal. More specifically, she has to neutralize the agent's preferences for the state, i.e., $V(x) = -F_A(x)$. This makes the agent completely indifferent to the value of the state and, consequently, the agent has no incentive whatsoever to spend any effort, i.e., $U_A(x) = 0$. This is the case even if, absent the transfer, the agent derives positive utility from the state. Thus, in a Markov-perfect equilibrium only the principal controls the state and in doing so she takes into account her own preferences and the necessary compensation to the agent.

Let us now turn to the weaker requirement of time-consistency, i.e., to the condition $W_A(x(t)) = 0$ for all t. In a non-degenerate stochastic setting, this condition is not much different from Markov-perfection because, due to the stochastic disturbances, the equilibrium state trajectory will reach any point in the state space with positive probability. Therefore, in the rest of this section we restrict attention to the deterministic model with $\sigma(x) = 0$ for all x.

It is obvious that the time-consistency condition $W_A(x(t)) = 0$ for all t holds if and only if the utility flow to the agent is constant and equal to 0 for all t, that is, if and only if $F_A(x(t)) - C_A(U_A(x(t))) + V(x(t)) = 0$ holds for all t. Substituting this into the implementability constraint (9) and noting that $\sigma(x) = 0$ holds for all x we obtain

$$0 = \rho_A \int_{x_0}^{x(t)} C_A'(U_A(z))\,\mathrm{d}z - C_A'(U_A(x(t)))[\alpha U_P(x(t)) + U_A(x(t)) - \delta x(t)].$$

Note that this equation can also be written as

$$C_A'(U_A(x(t)))\dot{x}(t) = \rho_A \int_{x_0}^{x(t)} C_A'(U_A(z))\,\mathrm{d}z.$$

Using the variable transformation $z = x(s)$, which implies $\mathrm{d}z = \dot{x}(s)\,\mathrm{d}s$, and defining $y(t) = C_A'(U_A(x(t)))\dot{x}(t)$, the above equation can be expressed as

$$y(t) = \rho_A \int_0^t y(s)\,\mathrm{d}s.$$

This shows that $y(0) = 0$ and $\dot{y}(t) = \rho_A y(t)$ must hold. It follows that $y(t) = 0$ for all t and, hence, for every t, it must either be the case that $C_A'(U_A(x(t))) = 0$ or $\dot{x}(t) = 0$. Since $C_A'(u_A) = 0$ holds if and only if $u_A = 0$, the condition $y(t) = 0$ for all t is therefore equivalent to

$$U_A(x(t))\dot{x}(t) = U_A(x(t))[\alpha U_P(x(t)) + U_A(x(t)) - \delta x(t)] = 0 \tag{15}$$

for all t. What we have just shown is that in a time-consistent equilibrium of the deterministic model the implementability condition (9) is given by (15). This allows us to formulate the following result.

Theorem 3 *Suppose that $\sigma(x) = 0$ holds for all $x \in \mathbb{R}$. The principal's optimization problem in a time-consistent feedback equilibrium is to maximize*

$$J_P = \int_0^{+\infty} e^{-\rho_P t} \left\{ F_P(x(t)) + \beta F_A(x(t)) - C_P(u_P(t)) - \beta C_A(u_A(t)) \right\} \mathrm{d}t$$

subject to the the state equation

$$\dot{x}(t) = \alpha u_P(t) + u_A(t) - \delta x(t)\,,\ x(0) = x_0$$

and the control constraint

$$u_A(t)[\alpha u_P(t) + u_A(t) - \delta x(t)] = 0. \tag{16}$$

PROOF: It has been shown above that $F_A(x(t)) - C_A(U_A(x(t))) + V(x(t)) = 0$ must hold which implies $v(t) = V(x(t)) = -F_A(x(t)) + C_A(u_A(t))$. Substituting this into (2) we obtain the objective functional stated in the theorem. The control constraint (16) is just a reiteration of (15). $\square$

Although the instantaneous utility of the principal in theorem 3 is a strictly concave function of the controls and the state equation is linear in the controls, the principal's optimization problem is likely to have no optimal solution. This is the case because the set of admissible controls defined by (16) is not convex (more specifically, it is the union of two straight lines in the control space). Except perhaps for borderline cases we therefore cannot expect the game under consideration to have a time-consistent feedback Stackelberg equilibrium. For this reason we shall not discuss this case any further.

6 A linear example

In this section we assume that $F_P(x) = \gamma_P x$, $F_A(x) = \gamma_A x$, and $\delta = 0$. We shall discuss the equilibrium under these assumptions following the case distinction from theorem 1, that is, we will first assume that the players have equal time-preference rates, then we will look at the deterministic case, and finally we will also briefly study the general case.

So, for the moment suppose that $\rho_P = \rho_A = \rho$. According to theorem 1(a), the Hamiltonian-Jacobi-Bellman equation of the principal's optimal control problem is

$$\rho W_P(x) = \max \left\{ \gamma x - C_P(u_P) - \beta C_A(u_A) + W_P'(x)(\alpha u_P + u_A) + [\sigma(x)^2/2] W_P''(x) \right\},$$

where $\gamma = \gamma_P + \beta\gamma_A$ and where the maximization is with respect to $(u_P, u_A) \in \mathbb{R}^2$. It is straightforward to verify that the linear function $W_P(x) = (\gamma/\rho)x + B$ satisfies this equation and that the corresponding optimal policy functions are

$$U_P(x) = \bar{u}_P := (C_P')^{-1}(\alpha\gamma/\rho) \text{ and } U_A(x) = \bar{u}_A := (C_A')^{-1}(\gamma/(\beta\rho)). \tag{17}$$

Here, the constant B is defined by

$$B = (1/\rho)\left[(\alpha\gamma/\rho)\bar{u}_P + (\gamma/\rho)\bar{u}_A - C_P(\bar{u}_P) - \beta C_A(\bar{u}_A)\right].$$

Substituting these results into the implementability condition (9) we obtain

$$V(x) = (\gamma_P x - \gamma x_0)/\beta + C_A(\bar{u}_A) - (\gamma/\beta)(\alpha\bar{u}_P + \bar{u}_A). \tag{18}$$

We summarize these findings in the following lemma.

Lemma 1 *Suppose that $F_P(x) = \gamma_P x$, $F_A(x) = \gamma_A x$, $\delta = 0$, and $\rho_P = \rho_A = \rho$ hold. Then it follows that the unique feedback Stackelberg equilibrium of the game is given by (17)-(18).*

Note that, under the assumptions of this lemma, equations (7), (17), and $W(x_0) = 0$ imply that $W_A(x) = \gamma(x - x_0)/(\beta\rho)$. This shows that the sign of $W_A(x(t))$ coincides with the sign of $\gamma[x(t) - x_0]$ and it follows therefore that, due to the stochastic evolution of the state variable, both cases $W_A(x(t)) > 0$ and $W_A(x(t)) < 0$ will typically occur. Recalling our discussion from section 5 this implies that there will typically exist incentives to terminate the contract for both players (although at different times). Note, however, that the strategies and value functions in the present case are independent of the form of the function σ.

Now let us turn to the deterministic model in which $\sigma(x) = 0$ holds for all x but in which the time-preference rates of the two players may differ from each other. According to theorem 1(b), the Hamiltonian of the principal's optimization problem is given by

$$H(x, u_P, u_A) = \gamma x + \Phi(u_P, u_A) + \mu(\alpha u_P + u_A),$$

where μ denotes the adjoint variable and where

$$\Phi(u_P, u_A) = [\beta(\rho_P - \rho_A)/\rho_P]C_A'(u_A)(\alpha u_P + u_A) - C_P(u_P) - \beta C_A(u_A). \tag{19}$$

From our assumptions imposed on the cost function C_P it follows that Φ is strictly concave with respect to u_P. For the remainder of this section we assume in addition that Φ is strictly concave with respect to (u_P, u_A). Note that this assumption trivially implies that the Hamiltonian function is jointly concave in (x, u_P, u_A).

The adjoint equation is

$$\dot{\mu}(t) = \rho_P \mu(t) - \gamma$$

and admits a constant solution $\mu(t) = \gamma/\rho_P$. We will see shortly that for this constant solution the transversality condition is satisfied so that we do not need to consider other solutions of the adjoint equation. From the strict concavity assumptions imposed above, it follows that the first-order conditions for the maximization of the Hamiltonian function with respect to (u_P, u_A) (where μ is evaluated at the constant solution of the adjoint equation) are

$$\alpha\beta(\rho_P - \rho_A)C'_A(u_A) - \rho_P C'_P(u_P) + \alpha\gamma = 0, \tag{20}$$

$$\beta(\rho_P - \rho_A)[C''_A(u_A)(\alpha u_P + u_A) + C'_A(u_A)] - \beta\rho_P C'_A(u_A) + \gamma = 0, \tag{21}$$

and that these conditions have a unique solution $(\tilde{u}_P, \tilde{u}_A)$. Substituting these results into (9) we obtain a linear strategy for the transfers given by

$$V(x) = [C'_A(\tilde{u}_A)\rho_A - \gamma_A]x + C_A(\tilde{u}_A) - C'_A(\tilde{u}_A)(\alpha\tilde{u}_P + \tilde{u}_A + \rho_A x_0). \tag{22}$$

Because the optimal controls $(\tilde{u}_P, \tilde{u}_A)$ are constant, it is obvious that in this equilibrium the state variable grows linearly and, hence, both the transversality condition of the principal's optimization problem as well as condition (4) are satisfied. Thus, we get the following result.

Lemma 2 *Suppose that $F_P(x) = \gamma_P x$, $F_A(x) = \gamma_A x$, $\sigma(x) = 0$, and $\delta = 0$. Assume furthermore that the function Φ specified in (19) is strictly concave. Then it follows that the unique feedback Stackelberg equilibrium of the game is given by $U_P(x) = \tilde{u}_p$, $U_A(x) = \tilde{u}_A$, and V, where $(\tilde{u}_P, \tilde{u}_A)$ is the unique solution of equations (20)-(21), and where V is given by (22).*

Proof: It has been shown above that the stated strategies qualify as an equilibrium. To show that it is unique, it suffices to rule out that there are other solutions to the principal's optimization problem. These would have to correspond to solutions of the adjoint equation that are not constant. It can be shown that these solutions do not satisfy (4). □

Let us also check for the deterministic case which of the two players has an incentive to terminate the contract. From lemma 2 we know that in equilibrium $\dot{x}(t) = \alpha\tilde{u}_P + \tilde{u}_A$ and, hence, $x(t) = (\alpha\tilde{u}_P + \tilde{u}_A)t + x_0$. From (7) and $W(x_0) = 0$ it follows that $W_A(x) = C'_A(\tilde{u}_A)(x - x_0)$. Combining these observations yields $W_A(x(t)) = C'_A(\tilde{u}_A)(\alpha\tilde{u}_P + \tilde{u}_A)t$. Together with the fact that the sign of $C'(\tilde{u}_A)$ coincides with the sign of $\tilde{u}_A$ this implies that the sign of $W_A(x(t))$ coincides with the sign of $\tilde{u}_A(\alpha\tilde{u}_P + \tilde{u}_A)$. In general, there are parameter constellations that lead to $\tilde{u}_A(\alpha\tilde{u}_P + \tilde{u}_A) < 0$ in which case the agent has an incentive to terminate the contract. If, however, the parameter constellation is such that both agents use positive effort, $\tilde{u}_P > 0$ and $\tilde{u}_A > 0$, then it follows that only the principal can benefit from renegotiations.

We conclude this section by briefly commenting on the general (non-deterministic) case with $\rho_P \neq \rho_A$. Theorem 1 does not give us a standard form of the principal's optimization problem. Instead of trying to develop necessary optimality conditions for this non-standard optimization problem, let us just look for an optimum in a restricted class of strategies.[4] The results presented in lemmas 1 and 2 show that, for the two special cases covered by these lemmas, the optimal effort strategy of the agent, U_A, is a constant function. Let us therefore seek an equilibrium of the stochastic model under the additional restriction that U_A is constant. In this case it holds for all $x \in \mathbb{R}$ that $U'_A(x) = 0$ and it follows from theorem 1 that the principal maximizes

$$\mathbb{E}_0 \int_0^{+\infty} e^{-\rho_P t} G(x(t), u_P(t), u_A(t))\, \mathrm{d}t$$

[4]See our introductory section 1 for a discussion of this approach and some references.

subject to the state equation (1), where the function G is defined in (14). Based on the previous results of this section, we conjecture that the optimal value function W_P is linear. In that case, $W_P''(x) = 0$ so that the corresponding term in the Hamilton-Jacobi-Bellman equation vanishes. It follows that the optimal values for u_P and u_A as well as the optimal value function W_P coincide with those found in the deterministic case. We can therefore conclude that in the stochastic model with the additional restriction of U_A being a constant function, the results from lemma 2 remain true. We conjecture that lemma 2 remains true also in the general stochastic model (i.e., without the restriction of a constant strategy U_A) but we have no proof for that conjecture.

7 A linear-quadratic example

In this section we consider the case in which the functions F_P, F_A, C_P, and C_A are quadratic polynomials and σ is a linear polynomial. In this case and under the assumptions of theorem 1(a-b), the principal's optimization problem turns out to be a linear-quadratic optimal control problem for which solution methods and characterization results are readily available. Instead of treating the general case here, we focus on a very specific example. This example is characterized by $F_P(x) = -(\gamma/2)x^2$, $F_A(x) = 0$, $C_P(u) = C_A(u) = (1/2)u^2$, $\sigma(x) = \bar{\sigma}x$, and $\alpha = 0$, where γ is strictly positive and $\bar{\sigma}$ is non-negative. Note that our maintained assumptions on the cost functions C_P and C_A are satisfied by these specifications. The assumption $\alpha = 0$ means that the principal is unable to affect the state of the system herself. She will therefore optimally choose $u_P(t) = 0$ for all t. The assumption $F_A(x) = 0$, on the other hand, implies that the agent does not care about the state of the system. To summarize, the only way how the principal can affect the state is by inducing the agent via an appropriate transfer mechanism V to spend effort. And the only reason why the agent applies some effort $u_A(t)$ is because he is induced to do so by the principal's transfers. As in the previous section we proceed along the lines suggested by the case distinction in theorem 1.

If both players have the same time-preference rate $\rho_P = \rho_A = \rho$, then we obtain from theorem 1 that the Hamilton-Jacobi-Bellman equation of the principal's problem is

$$\rho W_P(x) = \max \left\{ -(\gamma/2)x^2 - (1/2)u_P^2 - (\beta/2)u_A^2 + W_P'(x)(u_A - \delta x) + (\bar{\sigma}^2/2)W_P''(x)x^2 \right\}.$$

Maximizing with respect to u_P and u_A yields $u_P = 0$ (thereby confirming our earlier claim that $U_P(x) = 0$ holds for all x) and $u_A = (1/\beta)W_P'(x)$. Substituting these findings back into the Hamilton-Jacobi-Bellman equation we get

$$\rho W_P(x) = -(\gamma/2)x^2 + [1/(2\beta)][W_P'(x)]^2 - \delta W_P'(x)x + (\bar{\sigma}^2/2)W''(x)x^2.$$

Because of the linear-quadratic structure of the model we conjecture a quadratic optimal value function of the form $W_P(x) = (K/2)x^2$. Substituting this conjecture into the above equation, we see that the conjecture is true provided that K satisfies the quadratic equation

$$K^2 - \beta(\rho - \bar{\sigma}^2 + 2\delta)K - \beta\gamma = 0. \tag{23}$$

This equation has two real roots, one negative and one positive. Only the negative root can lead to the correct optimal value function as the latter must be concave. Hence, we conclude that $K = \bar{K}$, where $\bar{K}$ is the unique negative solution of (23). This implies that $U_A(x) = (\bar{K}/\beta)x$. Substituting this into (9) leads after simplifications (using (23)) to $V(x) = -(\gamma x^2 + \rho \bar{K} x_0^2)/(2\beta)$. We summarize our findings in the following lemma.

Lemma 3 *Suppose that $F_P(x) = -\gamma x^2/2$, $F_A(x) = 0$, $\sigma(x) = \bar{\sigma}x$, $\alpha = 0$, $\gamma > 0$, and $\rho_P = \rho_A = \rho$. Then it follows that the unique feedback Stackelberg equilibrium of the game is given* *by $U_P(x) = 0$, $U_A(x) = (\bar{K}/\beta)x$, and $V(x) = -(\gamma x^2 + \rho \bar{K} x_0^2)/(2\beta)$, where $\bar{K}$ is the unique negative solution of (23).*

Using this lemma, (7), and $W(x_0) = 0$ it is easy to calculate the optimal value function of the agent. It is given by $W_A(x) = [\bar{K}/(2\beta)](x^2 - x_0^2)$. If $x_0 = 0$, then this shows that $W_A(x) \leq 0$ holds for all $x \in \mathbb{R}$. In this situation the principal can therefore never have an incentive to terminate the contract, whereas the agent does want to renegotiate. However, if $x_0 \neq 0$, both parties can have incentives for renegotiation depending on the realization of the stochastic noise process.

It is easy to analyze how the model parameters influence the equilibrium strategies. As an example consider the parameter $\bar{\sigma}$. It can be seen from (23) that $\bar{K}$ is strictly decreasing in $\bar{\sigma}$. This implies that the absolute value of the slope of U_A is strictly increasing with respect to $\bar{\sigma}$, which means that the agent reacts stronger to changes in x in a model with high uncertainty than in a model with low noise. Analogously, one finds that, for any given x, the transfer $V(x)$ is non-decreasing with respect to $\bar{\sigma}$.

Let us now consider the deterministic version of this example. Using theorem 1(b) we see that the Hamiltonian of the principal's optimal control problem is given by

$$H(x, u_P, u_A) = -\frac{\gamma}{2}x^2 + \frac{\beta\delta(\rho_A - \rho_P)}{\rho_P}u_A x - \frac{\beta(2\rho_A - \rho_P)}{2\rho_P}u_A^2 - \frac{1}{2}u_P^2 + \mu(u_A - \delta x),$$

where μ denotes again the adjoint variable. This function is strictly concave with respect to (u_P, u_A) if and only if $\rho_P < 2\rho_A$, a condition that we assume to hold for the rest of this section.

Maximization of the Hamiltonian with respect to u_P yields $u_P = 0$. We shall therefore omit the control variable $u_P(t)$ from the following discussion and consider $u_A(t)$ as the only relevant control variable. The necessary optimality conditions for the problem under consideration are therefore

$$\frac{\beta\delta(\rho_A-\rho_P)}{\rho_P}x(t)-\frac{\beta(2\rho_A-\rho_P)}{\rho_P}u_A(t)+\mu(t)=0,$$
$$\dot{\mu}(t)=(\delta+\rho_P)\mu(t)+\gamma x(t)-\frac{\beta\delta(\rho_A-\rho_P)}{\rho_P}u_A(t).$$

Together with (1) these two equations form a set of linear equations that can be reduced to the following two-dimensional system of linear differential equations in $x(t)$ and $u_A(t)$.

$$\begin{pmatrix} \dot{x}(t) \\ \dot{u}_A(t) \end{pmatrix} = \begin{pmatrix} -\delta & 1 \\ \frac{\beta\delta(2\delta+\rho_P)(\rho_P-\rho_A)+\gamma\rho_P}{\beta(2\rho_A-\rho_P)} & \delta+\rho_P \end{pmatrix} \begin{pmatrix} x(t) \\ u_A(t) \end{pmatrix}. \tag{24}$$

The determinant of the system matrix of (24) is equal to

$$D=-\frac{[\gamma+\beta\delta(\delta+\rho_A)]\rho_P}{\beta(2\rho_A-\rho_P)}<0.$$

It follows that the matrix has one positive and one negative eigenvalue. Since the positive eigenvalue must be larger than ρ_P (which is the trace of the system matrix), the corresponding solution does not satisfy condition (4). Hence, only the stable solution is relevant. We therefore obtain the following result.

Lemma 4 *Suppose that $F_P(x)=-\gamma x^2/2$, $F_A(x)=0$, $\sigma(x)=0$, $\alpha=0$, $\gamma>0$, and $\rho_P<2\rho_A$. Then it follows that the unique feedback Stackelberg equilibrium of the game is given by $U_P(x)=0$, $U_A(x)=(\delta+z)x$, and $V(x)=vx^2/2+v_0$, where z is the negative eigenvalue of the system matrix in (24) and where $v=(\rho_A+\delta-z)(\delta+z)$ and $v_0=-\rho_A(\delta+z)x_0^2/2$.*

Proof: It has been shown above that $U_P(x)=0$. An eigenvector of the system matrix in (24) that corresponds to the negative eigenvalue z is given by $(1,\delta+z)^T$. This shows that the stable solution of (24) is characterized by $u_A(t)=(\delta+z)x(t)=U_A(x(t))$. Finally, the form of V can be derived by substituting the results already derived into the implementability constraint (9). This proves the necessity of the conditions stated in the lemma. For sufficiency just note that the principal's optimal control problem has a strictly concave Hamiltonian function and that, because of the stability of the solution of (24), the transversality condition is satisfied. This completes the proof of the lemma. □

From lemma 4 we know that $\dot{x}(t) = zx(t)$ and, hence, $x(t) = x_0 e^{zt}$. Moreover, together with (7) and $W(x_0) = 0$ the lemma implies that $W_A(x) = [(\delta + z)/2](x^2 - x_0^2)$. Taking these results together, we obtain $W_A(x(t)) = [(\delta + z)x_0^2/2](e^{2zt} - 1)$. Since $z < 0$ by construction, we see that the sign of $W_A(x(t))$ coincides with the sign of $-(\delta + z)$. If $z < -\delta$, then this expression is positive and it is the principal who can benefit from a termination of the contract. To find out under which parameter constellations it holds that $z < -\delta$, one can evaluate the characteristic polynomial of the system matrix in (24) at $-\delta$ and check whether this value is negative. It is straightforward to see that this is the case if and only if

$$\gamma > \beta\delta(2\delta + \rho_P)(\rho_A - \rho_P)/\rho_P.$$

Note that this condition is automatically satisfied if $\rho_P \geq \rho_A$. In the converse case $\rho_P < \rho_A$, on the other hand, the condition is only satisfied if γ is sufficiently large. In other words, the agent has an incentive to terminate the contract if he is less patient than the principal and if the principal's cost of a non-zero stock is small.

An interesting feature of this example is that, when the agent is more impatient than the principal (i.e. $\rho_P < \rho_A$), it may be the case that $\delta + z > 0$, which implies that $U_A(x)$ is strictly positive when $x > 0$. This means that the principal induces the agent to add to the stock when the stock is already positive even though the principal ideally would want a zero stock. At first sight, this result may seem counter-intuitive: why would the principal ask the agent to contribute to the growth of something that she does not want? Note, however, that it is still true that $\dot{x}(t) < 0$ when $x(t) > 0$, but that the agent is asked to make the convergence to zero slower. Upon reflection, this case arises only because the agent discounts the future more heavily than the principal. Starting with some $x_0 > 0$, the agent receives positive transfers early in the program and receives negative transfers as $x(t)$ gets close to zero. So what happens is that the principal sends utility to the agent early in the game and reduces the agent's utility later in the game. In other words, the principal is benefiting from intertemporal lending and borrowing. The more patient person (the principal) lends to the less patient one (the agent).

In the rest of this section we consider the general case covered by theorem 1(c). As in the previous section, however, we restrict the set of functions from which the agent's strategy U_A can be chosen. Due to the linear-quadratic structure of the game we assume that feasible strategies U_A must be linear, i.e., of the form $U_A(x) = \psi x$ where ψ is a deterministic coefficient. Theorem 1(c) tells us that in this case the principal tries to maximize

$$J_P = \mathbb{E}_0 \int_0^{+\infty} e^{-\rho_P t} \left\{ -[L(\psi)/2]x(t)^2 - (1/2)u_p(t)^2 \right\} \,\mathrm{d}t$$

subject to

$$\mathrm{d}x(t) = (\psi - \delta)x(t)\,\mathrm{d}t + \bar{\sigma}x(t)\,\mathrm{d}w(t)\,,\ x(0) = x_0,$$

where

$$L(\psi) = \beta(2\rho_A - \rho_P)\psi^2/\rho_P + \beta(\rho_P - \rho_A)(2\delta - \bar{\sigma}^2)\psi/\rho_P + \gamma.$$

Note that $L(\psi) > 0$ holds for all $\psi \in \mathbb{R}$ provided that

$$4\gamma\rho_P(2\rho_A - \rho_P) > \beta(2\delta - \bar{\sigma}^2)^2(\rho_A - \rho_P)^2. \tag{25}$$

For the rest of this section we assume this condition to hold. Obviously, it is still optimal for the principal to choose $u_P(t) = 0$ for all t. Furthermore, the state equation displayed above implies that x is a geometric Brownian motion given by

$$x(t) = x_0 \exp\left[(\psi - \delta - \bar{\sigma}^2/2)t + \bar{\sigma}w(t)\right].$$

Substituting these results back into the objective functional we obtain

$$J_P = \begin{cases} -L(\psi)x_0^2/[2(\rho_P + 2\delta - \bar{\sigma}^2 - 2\psi)] & \text{if } \psi < (\rho_P + 2\delta - \bar{\sigma}^2)/2, \\ \\ -\infty & \text{otherwise.} \end{cases}$$

The optimal value for ψ can therefore be found by maximizing $-L(\psi)/(\rho_P + 2\delta - \bar{\sigma}^2 - 2\psi)$ subject to $\psi < (\rho_P + 2\delta - \bar{\sigma}^2)/2$. This yields

$$\psi = \frac{1}{2}\left[\rho_P + 2\delta - \bar{\sigma}^2 - \sqrt{\rho_P \frac{4\gamma + \beta(2\delta + 2\rho_A - \rho_P - \bar{\sigma}^2)(2\delta + \rho_P - \bar{\sigma}^2)}{\beta(2\rho_A - \rho_P)}}\right]$$

Having found the optimal effort strategy $U_A(x) = \psi x$, we obtain the optimal transfer from (9), namely

$$V(x) = (\psi/2)[(\rho_A + 2\delta - \bar{\sigma}^2 - \psi)x^2 - \rho_A x_0^2].$$

8 Concluding remarks

We have shown how the leader's optimization problem in a certain class of Stackelberg differential games with feedback strategies can be rewritten as a standard optimal control problem. This is a noticeable property because, in general, the structure of this problem is so complex that standard solution techniques are not applicable. Our approach is feasible because the class of models under consideration in this paper has two crucial features: (i) the leader (principal) can directly transfer utility to the follower (agent) and (ii) the follower has an outside option. Feature (i) makes the implementability constraint for the leader very simple and feature (ii) allows us to pin down the optimal value for the follower.

There are a number of interesting directions for future research. First of all, in the present paper we have restricted ourselves to the demonstration of the approach along with an illustration by two simple examples. What is still missing is to apply this approach to an interesting real-life problem and to derive economic insights from such an application. In this respect it may be worthwhile to extend the applicability of the approach to slightly modified versions of the model considered here. For example, one could try to get rid of some of the additive separability that we have assumed for the utility functions of the two players (e.g., the additive separability between the benefits of the stock and the costs of effort). In a similar vein, we believe that the linearity of the state equation is an assumption that may be relaxed. Finally, one could imagine that the approach can be generalized to situations in which there are either multiple agents or multiple principals. The latter situation is know as the common agency problem, a static version of which has been studied for example by Grossman and Helpman (1994).

As for the two crucial model features pointed out above, we suppose that it will be much harder to relax them. In a differential game in which the follower does not have an outside option, the follower's optimal value is not determined by the outside option and, hence, it remains an endogenous variable in the leader's optimization problem. This seems to be a severe obstacle to the application of standard optimal control techniques to the leader's problem. Similarly, if we allow for other incentive mechanisms than a direct transfer of utility, it is much harder to integrate the first-order conditions for the follower's problem in order to obtain a tractable implementability constraint for the leader. Despite these difficulties we plan to look into these issues in future research.

References

[1] T. Basar and G.J. Olsder (1982): *Dynamic Noncooperative Game Theory*, Academic Press, London.

[2] H. Benchekroun and N.V. Long (1998): "Efficiency-inducing taxation for polluting oligopolists", *Journal of Public Economics* **70**, 325-342.

[3] H. Benchekroun and N.V. Long (2002): "On the multiplicity of efficiency-inducing tax rules", *Economics Letters* **76**, 331-336.

[4] F.-R. Chang (1988): "The inverse optimal problem: a dynamic programming approach", *Econometrica* **56**, 147-172.

[5] E. Dockner, S. Jorgensen, N.V. Long, and G. Sorger (2000): *Differential Games in Economics and Management Science*, Cambridge University Press, Cambridge.

[6] G. Gaudet, P. Lasserre, and N.V. Long (1995): "Optimal resource royalties with unknown and temporally independent extraction cost structure", *International Economic Review* **36**, 715-749.

[7] G. Gaudet, P. Lasserre, and N.V. Long (1996): "Dynamic incentive contracts with uncorrelated private information and history-dependent outcomes", *The Japanese Economic Review* **47**, 321-334.

[8] G. Gaudet, P. Lasserre, and N.V. Long (1998): "Real investment decisions under adjustment costs and asymmetric information", *Journal of Economic Dynamics and Control* **23**, 71-95.

[9] G. Grossman and E. Helpman (1994): "Protection for sale", *American Economic Review* **84**, 833-850.

[10] B. Holmstrom and P. Milgrom (1987): "Aggregation and linearity in the provivion of intertemporal incentives", *Econometrica* **55**, 303-328.

[11] J.-J. Laffont and J. Tirole (1988): "The dynamics of incentive contracts", *Econometrica* **56**, 1153-1175.

[12] A. Mas-Colell, M.D. Whinston, and J.R. Green (1995): *Microeconomic Theory*, Oxford University Press, New York.

[13] Y. Sannikov (2008): "A continuous-time version of the principal-agent problem" *Review of Economic Studies* **75**, 957-984.

[14] H. Schattler and J. Sung (1993): "The first-order approach to the continuous-time principal agent problem with exponential utility", *Journal of Economic Theory* **61**, 331-371.

[15] K. Shimomura and D. Xie (2008): "Advances on Stackelberg open-loop and feedback strategies", *International Journal of Economic Theory* **4**, 115-133.

[16] J. Sung (1995): "Linearity with project selection and controllable diffusion rate in continuous-time principal-agent problems", *RAND Journal of Economics* **26**, 720-743.