Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Performance-Based Incentives in a Dynamic Principal-Agent Model

Erica L. Plambeck, Stefanos A. Zenios,

Performance-Based Incentives in a Dynamic Principal-Agent Model

Erica L. Plambeck • Stefanos A. Zenios
Department of EES&OR, Stanford University, Stanford, California 94305
Graduate School of Business, Stanford University, Stanford, California 94305
elp@stanford.edu • stefzen@leland.stanford.edu

The principal-agent paradigm, in which a principal has a primary stake in the performance of some system but delegates operational control of that system to an agent, has many natural applications in operations management (OM). However, existing principal-agent models are of limited use to OM researchers because they cannot represent the rich dynamic structure required of OM models. This paper formulates a novel dynamic model that overcomes these limitations by combining the principal-agent framework with the physical structure of a Markov decision process. In this model one has a system moving from state to state as time passes, with transition probabilities depending on actions chosen by an agent, and a principal who pays the agent based on state transitions observed. The principal seeks an optimal payment scheme, striving to induce the actions that will maximize her expected discounted profits over a finite planning horizon. Although dynamic principal-agent models similar to the one proposed here are considered intractable, a set of assumptions are introduced that enable a systematic analysis. These assumptions involve the "economic structure" of the model but not its "physical structure." Under these assumptions, the paper establishes that one can use a dynamic-programming recursion to derive an optimal payment scheme. This scheme is memoryless and satisfies a generalization of Bellman's principle of optimality. Important managerial insights are highlighted in the context of a two-state example called "the maintenance problem".
(*Dynamic Principal Agent Problem; Incentives In Operations Management; Maintenance*)

1. Introduction

Operations Management (OM) is a natural area of application for the principal-agent paradigm, in which one party (the principal) has a primary stake in the performance of some system but delegates operational control of that system to another party (the agent). Most of the problems that we study in OM involve such delegated control, although classical OM models often suppress this feature. Unfortunately, the classical economic models for the principal-agent problem are of limited use to OM researchers, because they focus either on one-shot static problems or else on "repeated" problems involving a simple kind of multi-period replications, whereas even stylized OM models typically require a richer dynamic structure. The first contribution of this paper is to formulate a general dynamic model that combines the principal-agent framework with the physical structure of a Markov decision process (MDP): In this model a system moves from state to state as time passes, with transition probabilities depending on actions chosen by an agent, and a principal who pays the agent based on the state transitions observed. The principal seeks an optimal compensation scheme, one that induces the agent to adopt actions which maximize the principal's finite-horizon expected discounted profits.

1523-4614/00/0203/0240$05.00
1526-5498 electronic ISSN

Generally speaking, the design of an optimal compensation scheme in the dynamic principal-agent context is considered an intractable problem. In fact, even in the simpler repeated principal-agent setting, the analysis of optimal schemes is formidable and involves complex and subtle economic reasoning. Part of the complexity arises because the best schemes are history dependent, paying the agent according to the complete history of the underlying system, and because they require *strategic commitment* in which both parties initially agree to be bound to terms for future payments that would not be adopted at a future time. This implies that the optimal compensation scheme may not represent the principal's best choice in each period and state, and thus Bellman's principle of optimality, the driving force behind the analysis of dynamic decision models, does not hold. Therefore, a second contribution of this paper is to identify a set of assumptions that alleviate the analytic complexity associated with history dependence and strategic commitment. These assumptions, which specialize conditions first articulated by Fudenberg et al. (1990), involve the information sets and utility functions of the two parties and their ability to transfer wealth between time periods through borrowing and lending, but they do not involve the state space and transition structure, or the reward structure of the model. In other words, they involve the "economic structure" of the problem but not its "physical structure."

The paper's third contribution is foundational in nature. It will be shown that under the conditions referred to immediately above, the dynamic principal-agent model is amenable to a two-step dynamic-programming analysis. The first step formulates a series of constrained nonlinear optimization problems, described by Equations (22)–(24) in §4, whose solution characterizes the optimal compensation scheme for the static analogue of the dynamic principal-agent model. The second step solves a dynamic-programming recursion, described by Equation (28) in §4, that uses for its inputs the optimal value of the optimization problems from the first step. The solution to this recursion gives a *memoryless* (history-independent) optimal compensation scheme for the dynamic model that satisfies a generalization of Bellman's principle of optimality: It is optimal not only in the first period and initial state but in every subsequent period and state.

The remainder of this paper is organized as follows: Section 2 provides a preliminary and incomplete description of a simple example in order to motivate our model formulation and analysis. It also presents a guide to the literature on principal-agent models. Section 3 presents a mathematical description of the dynamic principal-agent model, and §4 provides the main theorems that justify the two-step dynamic-programming analysis. Section 5 provides the mathematical analysis of the example introduced in §2. Concluding remarks appear in §6.

2. Motivating Example

A simple example, which will be referred to as the "maintenance problem," will provide context for our formulation. It will also justify the key ingredients of our dynamic principal-agent model and motivate the main steps in its analysis. The example represents a stylized abstraction of a multiperiod machine maintenance problem with delegated control. In its description we abandon the generic principal-agent language in favor of a more specialized vocabulary that reflects our interest in OM applications. The vocabulary restriction does not, however, limit the applicability of our model and results. A guide to the literature linking our study to existing models and results is provided at the end of this section.

Preliminary Formulation. Consider a critical machine, which can be in one of two possible states: an operational state, and a nonoperational one. The machine generates profit for its owner (whom we call "her") only in the operational state, and the owner delegates repair responsibilities to a manager (whom we call "him"). The manager repairs the machine when it is in the nonoperational state, and he can either exercise high effort or low effort. The strategy used to choose his effort level at each period is referred to as the *maintenance strategy.* The cost-of-effort and probability of transition back to the operational state are higher when high effort is exercised. In the operational state the manager undertakes nondiscretionary preventive maintenance activities. The owner cannot monitor the manager's efforts in the nonoperational state but she observes the state of the machine. Thus, she faces the following problem: Design a performance-based compensation scheme that will reward the manager based on the machine's observed states, and will motivate him to take actions that will maximize the owner's expected profits over a finite horizon.

Before we can formally formulate and analyze the owner's problem, we first specify a model for the preferences of the two parties. The manager is assumed to choose his maintenance strategy by solving a *consumption-effort* problem. This problem is analogous to the consumption-investment problem of financial economics (see Ingersoll 1987), and its main features are as follows: In each period, the manager acts and consumes based on the information available to him, striving to maximize his expected utility. The manager's utility is exponential and additively separable over time, increasing in consumption and decreasing in repair effort. Furthermore, the manager is able to transfer consumption between time periods by saving his income in a bank account and withdrawing a discretionary amount from the same account for consumption. The interest rate in this account is fixed and the balance in each period can either be negative, reflecting a loan, or positive. For the owner's preferences, the model assumes that she is a classical expected profit maximizer.

Explicit representation of the agent's consumption and investment activities might seem to be unnecessary. However, the agent is exposed to a risky income stream because his payment is linked to machine performance (a stochastic signal of his effort), and Smith (1998) shows that an accurate representation of the agent's preferences over these risky income streams requires a consumption model with access to perfect capital markets. In addition, Fudenberg et al. (1990) show that the *frictionless ideal* assumption of perfect capital markets eliminates the problem of strategic commitment described in §1, and thus, validates a dynamic-programming analysis of the principal's problem; the contributions of Fudenberg et al. (1990) will be discussed in the guide to the literature.

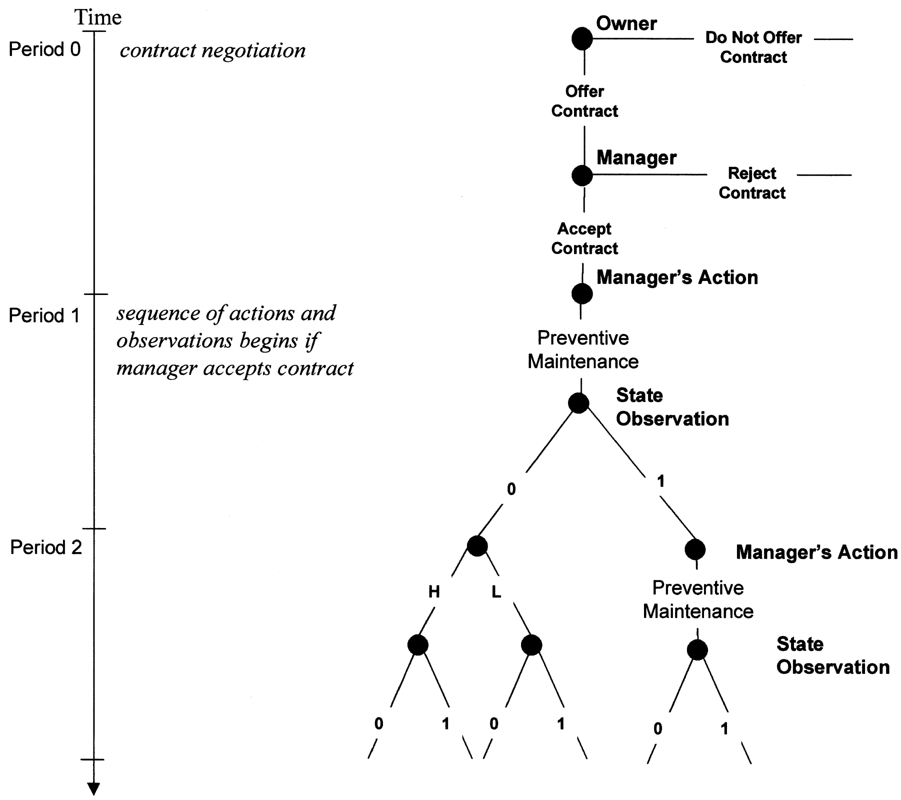
This completes a formal outline of the model for the "maintenance problem". Next, we provide a formal description of the owner's problem, which will be referred to as the *optimal contracting problem.*

The Optimal Contracting Problem. A precise mathematical description is given in §3, along with the formulation of the dynamic principal-agent model. Figure 1 presents a graphical description, and demonstrates that the problem can be represented as a two-stage Stackelberg game. The owner moves first and offers a long-term compensation scheme (contract) to the manager. The compensation scheme will reward the manager based on the observed states of the machine and can be history dependent. The manager decides whether to accept this compensation scheme, and if he does, he then determines the maintenance strategy that he will adopt.

The precise ingredients of this problem can be described more effectively backwards. Given the compensation scheme, the manager has to make two decisions. He must determine the maintenance strategy that maximizes his expected utility. This gives rise to the *incentive compatibility* constraint which states that given the compensation scheme, the manager will select the expected-utility-maximizing maintenance strategy. Then, moving backwards in time, he must decide whether or not to accept the compensation scheme. Assuming that the market for managers is competitive and it offers an alternative employment opportunity that guarantees a reservation utility level, this decision generates the *participation constraint:* The manager will accept the compensation scheme if his expected utility from accepting it is no less than his reservation utility. Next, we turn to the problem analyzed by the owner: Among all possible compensation schemes that satisfy the incentive compatibility and participation constraint, she will choose the one that maximizes her expected discounted profits. The analysis of the owner's problem produces an optimal *compensation scheme* and an optimal *maintenance strategy* that will voluntarily be adopted by the manager. The tuple of the resulting compensation scheme and maintenance strategy will be referred to as the *optimal long-term contract.*

Overview of Analysis and Main Results. The analysis of the owner's problem proceeds in a two-step

Figure 1 Graphical Representation of the "Maintenance Problem"



backward logic that imitates the two stages of the Stackelberg game description. This logic was first advocated by Grossman and Hart (1983). In the first step, one considers each possible maintenance strategy that can be adopted by the manager, and derives a minimum-cost compensation scheme for which this strategy represents the manager's best response. In the second step, one chooses the strategy, which when induced using the compensation schemes derived in Step 1, will maximize the owner's expected profit. The mathematical details are described in §4.

Analysis of the first step shows that it is sufficient to consider memoryless compensation schemes, and thus one need not worry about the complexities arising from history-dependent schemes. The memoryless schemes are derived from the analysis of simple static principal-agent models, described by Equations (22)–(24) in §4, and they have the following structure: When the strategy prescribes high effort in a given period, then the compensation scheme reimburses the manager for a fraction of his cost of effort, and promises an incentive bonus to be paid when the machine is repaired. The incentive bonus motivates the manager to adopt high effort, and his expected total compensation per period exceeds his cost of effort. The promised incentive bonus is the same each time the strategy prescribes high effort, and it is independent of the time it takes for the machine to be repaired. On the other hand, when the strategy prescribes low effort, then the compensation scheme reimburses the manager for his cost of effort only, without any incentive bonus.

The results also show that although the owner can always create compensation schemes that entice the manager to provide low effort, doing the same for high effort is not always feasible. In particular, if transitions to the operational state are not informative about the manager's effort, then it is impossible to design incentive schemes that motivate the manager to take high effort. Theorem 1 in §4 provides a precise mathematical specification of these results for the general dynamic principal-agent model, and Proposition 3 in §5 specializes the results to the "maintenance problem".

The analysis from the first step generates expressions for the per-period cost to the owner of payments that motivate the manager to take a desirable action. This cost is the same each time the owner wishes to induce a certain action, and it implies that one can use a standard dynamic-programming recursion to specify the strategy that will be adopted by the manager in response to the optimal compensation scheme. The compensation scheme that will induce this strategy is then constructed using the memoryless payments derived in Step 1. The derived long-term contract is *immune to renegotiation:* Neither party will be better off by renegotiating the contract at any period. The mathematical details for the general dynamic principal-agent model are proved in Theorem 2 of §4, and Proposition 4 of §5 presents the application to the "maintenance problem".

Guide to the Literature. The principal-agent paradigm has a long history of applications in several fields of managerial relevance. To provide a perspective on the vast existing literature on agency models, and its relationship to our study, we now provide a summary of several important papers. We consider papers which contributed to the theoretical development of the principal-agent paradigm, as well as papers that introduced fruitful applications or empirical results. Hart and Holmstrom (1987) provide an excellent review of agency theory, and Gibbons (1998) discusses recent theoretical developments and open research questions. Prendergast (1998) surveys empirical studies on the provision of incentives. Mas-Colell et al. (1995) provide an introductory textbook treatment of principal-agent models, and the text by Salanie (1997) introduces broader issues in the theory of contracts, including strategic commitment and renegotiation in a dynamic setting.

One of the first formulations of a principal-agent model is due to Mirrlees (1974), who has considered a static one-shot relationship between a principal and an agent. Holmstrom (1979) explored Mirlees' simple static model further and demonstrated that there is loss of efficiency, known as agency loss, when the agent is risk-averse and his actions are unobservable, and that the agency loss depends on how informative the observed outcomes are about the agent's unobserved action. Grossman and Hart (1983) further explored the role of information in agency models and introduced the two-step solution approach that we adopt in our analysis here.

Recognizing that dynamic extensions of the static principal-agent model would be interesting, Radner (1985) considered an infinitely repeated version of the static model. He demonstrated that agency losses disappear if future utilities are not discounted, and that an optimal compensation scheme is history dependent and operates according to "control charts", punishing the agent when his performance statistics are inconsistent with the desirable actions. These results require that the principal and the agent will not negotiate midstream to shift to a Pareto superior compensation scheme and action strategy, an assumption that is implausible without strategic commitment to a long-term contract.

Fudenberg et al. (1990) demonstrated that strategic commitment to a long-term contract adds value for the principal *only* if the agent cannot borrow and save to smooth intertemporal fluctuations in his income. In their model the principal and agent have common knowledge of each other's utility structure and the physical system dynamics. Furthermore, the agent has access to perfect financial markets so that his incentives depend only on the net present value of the compensation scheme and not the timing of payments. The authors' key insight is that in any long-term contract the principal can rearrange the compensation scheme without altering its net present value along any sample path, so that at the beginning of every period the agent will voluntarily continue with the contract. Therefore, an optimal contract can be executed as a series of single-period contracts, and long-term strategic commitment is unnecessary. Our study adopts their economic framework and develops a complete characterization, using dynamic-programming techniques, of a memoryless optimal long-term contract when the system dynamics are Markovian.

In a related study, Spear and Srivastava (1987) provide a dynamic-programming characterization of the optimal contract when the system dynamics do not have state structure and the agent does not have access to capital markets. Holmstrom and Milgrom (1987) consider a principal-agent problem in which the agent controls the drift rate of a Brownian motion, and his utility is an exponential function of (undiscounted) income less effort-cost. They show that, although the principal observes the full history of the process, the optimal payment is simply a linear function of the aggregate displacement.

The applications of the principal-agent paradigm are numerous. Admati and Pfleiderer (1997) provide an application in investment management. The textbook by Lazear (1995) describes applications in personnel economics. Lambert (1987) considers agency problems in managerial accounting, and Wolfson (1985) describes an empirical analysis of agency problems in taxation and oil drilling. However, agency models are underepresented in the OM literature. Atkinson (1979) is, to the best of our knowledge, the first paper to introduce agency problems in operations management. A more recent study by Porteus and Whang (1991) uses agency models to study manufacturing/marketing incentives.

The principal-agent paradigm is also relevant for contracting problems in supply-chain management. Tsay et al. (1998) give an overview of the state-of-the-art in this research arena. Because the conventional wisdom is that analysis of contractual arrangements is "often too complex to be amenable to multi-period analysis" (see Tsay et al. 1998), very few supply chain studies analyze multi-period contracts. Notable exceptions are Bassok and Anupindi (1999), Tsay and Lovejoy (1999), and Cohen and Agrawal (1998). Our analysis suggests that the dynamic principal-agent model can provide a framework that will make the analysis of multiperiod supply-chain contracting problems tractable.

Our analysis also builds on the extensive literature on MDPs and replaces the traditional risk-neutral objective with a risk-sensitive one. For a textbook treatment of MDPs the reader is referred to Bertsekas (1987) or Puterman (1994), and for a discussion of risk-sensitivity in MDPs see Whittle (1990). MDPs with multiple decision makers have also attracted considerable attention. Filar and Vrieze (1996) provide a comprehensive treatment of the competitive situation where the transitions of a Markov chain are influenced by the simultaneous moves of two players. However, their model does not capture systems with delegated control that are the focus of our dynamic principal-agent model.

3. The Dynamic Principal-Agent Model

In this section we formulate the dynamic principal-agent model and present the main questions that will be addressed. The model is structured along the lines outlined in §2, and it assumes that the owner of a production system delegates operational control to a manager. The manager controls a finite state, finite action Markov decisions process, and his actions are unobservable by the owner.

The model assumes that there are T epochs, or periods, indexed by $t = 1, \ldots, T$. Let X_{t-1} denote the state of the Markov chain at the beginning of period t and let W_{t-1} denote the manager's wealth at the beginning of the same period; the manager's wealth represents the balance in his bank account and can be negative. X_0 and W_0 denote the initial state and manager's wealth respectively. The state at the beginning of period t is denoted by X_{t-1} instead of X_t because X_t will be used to denote the state observed at the end of that period, after the manager takes any actions that may affect the transition to a new state. The state set is assumed to be finite and is denoted by $X := \{1, \ldots, N\}$. The states of the Markov chain are observable both by the owner and the manager, but the manager's wealth is only observable by the manager; Lemma 1 in the Appendix shows that the analysis and results would not change if the manager's wealth was also observable by the owner. The state history of the Markov chain until the beginning of period t is denoted by $X^{t-1} = (X_0, X_1, \ldots, X_{t-1})$.

Within each period, the following sequence of events takes place: First, the manager selects an action $a_t(X^{t-1})$ from the set of all permissible actions in state X_{t-1}, $A(X_{t-1})$, and a consumption level $c_t(X^{t-1})$ that reduces his wealth to $W_{t-1} - c_t(X^{t-1})$; both the action and consumption decisions may depend on the manager's wealth but this is not reflected in the notation. The strategies used to choose the action and consumption decisions in each period and state will be referred to as the action and consumption strategies, respectively. As a result of his actions, the manager incurs a cost-of-effort $g_{X_{t-1}}(a_t(X^{t-1}))$. This cost-of-effort is an inconvenience cost that reduces his immediate utility but not his wealth. Next, his wealth is multiplied by a fraction $1/\alpha$, where α is the discount rate, and a transition to a new state X_t is observed. This transition is determined according to the stationary transition matrix $P_{X_{t-1},X_t}(a_t(X^{t-1})) = P(X_t \mid X_{t-1}, a_t(X^{t-1}))$. Then the owner, who cannot observe the manager's actions, receives a state-dependent reward $\pi(X_t)$ and makes a transfer payment to the manager according to the history-dependent compensation scheme $s_t(X^t)$. This changes the manager's wealth to

$$W_t = \frac{1}{\alpha}\,[W_{t-1} - c_t(X^{t-1})] + s_t(X^t), \tag{1}$$

and terminates the sequence of events for period t. In the last period T, the same sequence of events occurs and then the relationship terminates.

Our notation assumes that the manager's compensation in each period t may depend arbitrarily on the observed state history up to that period, X^t. The rationale behind this assumption is simple: The only signal the owner has about the manager's actions is the observed state history, and therefore she may wish to fully utilize this signal when designing the compensation scheme. A consequence of adopting a history-dependent compensation scheme is that the manager's optimal actions and consumption decisions in each period may also be history dependent. These decisions can also depend on the manager's wealth but it can be assumed that they will only depend on his current wealth in each period and not on wealth history; Ingersoll (1987) shows that this does not affect the generality of our conclusions as long as the manager's utility is exponential.

The manager is assumed to be risk-averse and effort-averse with a constant coefficient of risk-aversion $r > 0$. His utility is additively separable over time; it depends on his consumption decisions $c = (c_t(X^{t-1}))_{t=1,\dots,T}$ his actions $a = (a_t(X^{t-1}))_{t=1,\dots,T}$ and his terminal wealth W_T, and is expressed as follows:

$$U(a, c) = -\sum_{t=1}^{T} \alpha^{t-1} \exp[-r(c_t(X^{t-1}) - g_{X_{t-1}}(a_t(X^{t-1})))] - \frac{\alpha^T}{(1-\alpha)} \exp[-r(w + (1-\alpha)W_T)]; \tag{2}$$

w is a fixed wage per period received by the manager after the last period T. The expression $-\exp[-r(c_t(X^{t-1}) - g_{X_{t-1}}(a_t(X^{t-1}))]$ represents the manager's immediate utility in period t, and is increasing in his consumption and decreasing in the cost-of-effort. The terminal utility expression $-\alpha^T/(1-\alpha)\exp[-r(w + (1-\alpha)W_T)]$ captures the manager's expected future utility after the termination of the contract. Assuming that the manager consumes optimally over an infinite horizon but cannot borrow after time T, this expression gives his optimal expected future utility from period $T+1$ onwards. It relies on the fact that his optimal consumption per period after time T is equal to the interest from his terminal wealth, $(1-\alpha)W_T$, and the fixed wage, w. This special form of the terminal utility expression is also related to the participation constraint that will be formally introduced later in this section, and streamlines the analysis. The manager's initial wealth appears only implicitly in the utility function through its impact on the manager's consumption decision and terminal wealth. The reader will also notice that the following convention is adopted in (2): the letters a and c represent a multiperiod strategy of action and consumption decisions. This represents a slight abuse of notation because a and c are also used to denote single-period decisions. However, it should be clear from the context whether a and c represent single-period decisions or else multiperiod strategies.

The owner is assumed to be risk-neutral and evaluates his reward and compensation scheme with their net present value,

$$\sum_{t=1}^{T} \alpha^t\,[\pi(X_t) - s_t(X^t)] + \alpha^{T+1}\,\Pi(X_T); \tag{3}$$

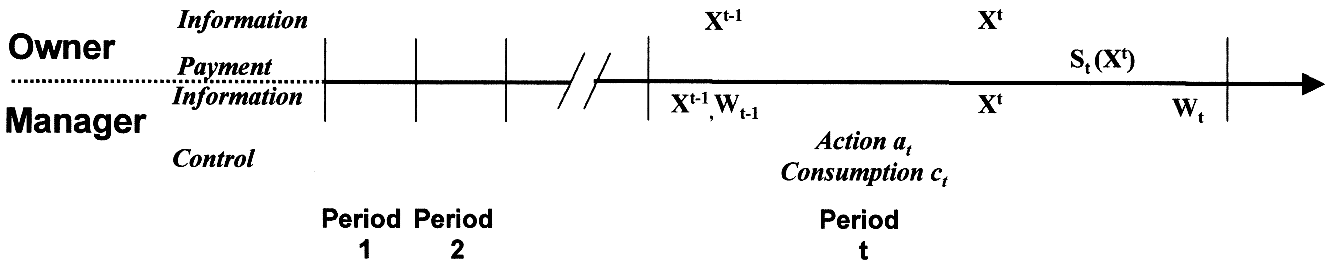
where $\Pi(X_T)$ is the terminal reward for the owner.

The model also assumes that both parties have complete knowledge of each other's utility functions and cost structure. Furthermore, they both know how the manager's actions affect the transition probabilities. A schematic description of the model and the information available to each party is given in Figure 2.

Having completed an outline of the model, we can

Figure 2 **Sequence of Events and Information Structure of the Markovian Principal-Agent Problem.**

Information Structure		
	Owner	Manager
System State	Y	Y
Manager's Wealth	N	Y
Owner's Objective	Y	Y
Manager's Objective	Y	Y
Transition Probabilities	Y	Y
Manager's Actions	N	Y
Manager's Consumption	N	Y
Compensation Scheme	Y	Y



now provide a formal description of the optimal contracting problem. As was explained in §2, this problem is more conveniently analyzed backwards, and our presentation will follow a two-step backwards logic.

Step 1: The Manager's Problem. Given a fixed compensation scheme $s = (s_t(X^t))_{t=1,\ldots,T}$, the manager must select at the beginning of Period 1 a production strategy $a = (a_t(X^t))_{t=1,\ldots,T}$ and consumption strategy $c = (c_t(X^t))_{t=1,\ldots,T}$ to solve the following problem:

$$\max_{a,c} \quad E_{a,c}[U(a, c) \mid X_0, W_0]; \tag{4}$$

$$\text{subject to} \quad W_t = \frac{1}{\alpha}[W_{t-1} - c_t(X^{t-1})] + s_t(X^t);\ t = 1, \ldots, T. \tag{5}$$

The notation $E_{a,c}[\cdot]$ indicates that the expectation is taken with respect to the transitions induced by the production strategy a and consumption strategy c. This problem can be equivalently formulated as follows: The manager selects a production strategy a and consumption strategy c that satisfy the *incentive compatibility constraints,*

$$E_{a,c}[U(a, c) \mid X_0, W_0] \geq E_{\tilde{a},\tilde{c}}[U(\tilde{a}, \tilde{c}) \mid X_0, W_0];$$
$$\text{for each admissible } \tilde{a} = (\tilde{a}_t(X^{t-1}))_{t=1,\ldots,T}$$
$$\text{and } \tilde{c} = (\tilde{c}_t(X^{t-1}))_{t=1,\ldots,T}. \tag{6}$$

The incentive compatibility constraint is only imposed at the beginning of Period 1. Nowhere is it required that this constraint will be satisfied in every subsequent period and achievable state.

In addition to selecting the production and consumption strategies that solve Problem (6), the manager has to decide whether to accept or reject the compensation scheme. The market for managers is assumed to be competitive, so if he rejects the contract he has an alternative employment opportunity offering him a fixed wage w per period. His *reservation utility* in this case will be

$$-\frac{1}{1-\alpha}\exp[-r(w + (1-\alpha)W_0)];$$

this follows because in this alternative the optimal strategy is to consume the wage, w, and interest from his wealth, $(1-\alpha)W_0$, in each period. Therefore, the manager will agree to work for the owner only if the *participation constraint* holds:

$$E_{a,c}[U(a, c) \mid X_0, W_0] \geq -\frac{1}{1-\alpha}\exp[-r(w + (1-\alpha)W_0)]. \tag{7}$$

The attentive reader will notice that the right hand side in (7) is the same as the expression for the terminal utility in (2), but with W_0 replacing W_T. The alternative employment opportunity with a wage w per period is assumed to be available to the manager whenever he is not under contractual obligation to the owner.

Step 2: The Owner's Problem. Working our way backwards, we now describe the owner's problem. The previous discussion shows that given any compensation scheme, the manager solves Problem (6)–(7) to obtain an incentive- and participation-compatible production strategy $a = (a_t(X^t))_{t=1,\ldots,T}$, and consumption strategy $c = (c_t(X^t))_{t=1,\ldots,T}$. This production strategy induces a Markov chain that generates state-dependent rewards for the owner. Intuitively, therefore, the owner wishes to select the compensation scheme $s = (s_t(X^t))_{t=1,\ldots,T}$ and instructions for the manager in the form of a production strategy $a = (a_t(X^t))_{t=1,\ldots,T}$ and consumption strategy $c = (c_t(X^t))_{t=1,\ldots,T}$ that will maximize her expected profit

$$E_{a,c}\left[\sum_{t=1}^{T} \alpha^t[\pi(X_t) - s_t(X^t)] + \alpha^{T+1}\,\Pi(X_T) \mid X_0\right] \quad (8)$$

subject to the incentive compatibility Constraint (6) and participation Constraint (7).

The analysis of Problem (6)–(8) will determine an optimal long-term contract which consists of an optimal compensation scheme $s^* = (s_t^*(X^t))_{t=1,\ldots,T}$, and the optimal production and consumption strategies $a^* = (a_t^*(X^t))_{t=1,\ldots,T}$ and $c^* = (c_t^*(X^t))_{t=1,\ldots,T}$ adopted by the manager in response to this compensation scheme. A triple of a compensation scheme s, production strategy a, and consumption strategy c will be referred to as a *long-term contract* (a, c, s). The production strategy a and consumption strategy c are instructions or suggestions for the manager. These instructions cannot be enforced, but the manager will voluntarily adopt them if they represent his optimal response to the compensation scheme s; that is, he will adopt them if they satisfy Constraints (6)–(7). In that case, we will say that s *implements* a and c.

It should be emphasized here that in addition to determining the optimal compensation scheme, the owner will also have to decide whether to offer this compensation scheme to the manager. Intuitively, she will choose to offer it only if her expected profit is nonnegative. Incorporating this decision in the analysis is trivial and will be omitted for purposes of clarity.

Before moving on, it will be necessary to make a brief digression into the relation between the optimal contract and the manager's wealth. Naturally, one would expect that the manager's optimal consumption and production strategy would depend on his wealth in each period. However, because the manager's wealth is unobservable by the owner, the owner will not be able to determine the manager's optimal response to any compensation scheme that she offers, and thus she will be unable to determine the optimal contract. Fortunately, this problem disappears because the manager's exponential utility implies that the optimal production strategy chosen by the manager in response to any compensation scheme does not depend on his wealth; this phenomenon is known as *absence-of-wealth-effects* and is presented in Lemma 1 in the Appendix. As the owner's profit depends on the manager's actions only through the production strategy, it follows that she will be able to determine the optimal contract even when she does not know the manager's wealth.

It is convenient to provide a more compact description of the optimal contracting Problem (6)–(8) using the concept of a value function from dynamic programming. This will facilitate the description of the main results in §4 and of the proofs in the Appendix. Let $V_t^O(X^{t-1}; (a, c, s))$ denote the value function for the owner in period t. This gives the owner's future expected discounted profit from period t to period T when the state history until period t is X^{t-1} and the long-term contract (a, c, s) is in effect; the value function does not depend on the manager's wealth W_{t-1} because of the absence-of-wealth-effects described above. In addition, let $V_t^M(X^{t-1}, W_{t-1}; (a, c, s))$ denote the value function for the manager in period t which gives his future expected utility from period t to period T when the state history until period t is X^{t-1}, his current wealth is W_{t-1}, and the long-term contract (a, c, s) is in effect; Ingersoll (1987) shows that the manager's value function in period t does not depend on his past wealth $\{W_\tau\}_{\tau=1,\ldots,t-2}$. For brevity of notation, when a and c represent the manager's optimal response to s, the notation $V_t^O(X^{t-1}; (a, c, s))$ and $V_t^M(X^{t-1}, W_{t-1}; (a, c, s))$ will be replaced by $V_t^O(X^{t-1}; s)$ and $V_t^M(X^{t-1}, W_{t-1}; s)$, respectively.

One can then use the value function notation to provide the following, more compact, representation of the optimal contracting Problem (6)–(8). Determine an optimal contract (a^*, c^*, s^*) that solves the following problem:

$$\max_{(a,c,s)} \quad V_1^O(X^0; (a, c, s)) \tag{9}$$

$$\text{subject to} \quad V_1^M(X^0, W_0; (a, c, s)) \geq -\frac{1}{1-\alpha} \exp[-r(w + (1-\alpha)W_0)], \tag{10}$$

$$V_1^M(X^0, W_0; (a, c, s)) \geq V_1^M(X^0, W_0; (\tilde{a}, \tilde{c}, s)); \tag{11}$$

$$\text{for every admissible } \tilde{a} = (\tilde{a}_t(X^{t-1}))_{t=1,\ldots,T} \text{ and } \tilde{c} = (\tilde{c}_t(X^{t-1}))_{t=1,\ldots,T}.$$

Constraint (10) represents the participation Constraint (7), and Constraint (11) represents the incentive compatibility Constraint (6).

Having completed the description of the optimal contracting problem, we now proceed to outline two important questions that provide a framework for the subsequent study of the optimal contract.

The Optimal Contract When Actions Are Observable. To start at the very beginning, the first question is the following: *Is there a tension between the objectives of the owner and the manager in Problem* (9)–(11)? The answer to this question is now well understood and is explained in Mas-Colell et al. (1995). Briefly, tension arises only when the manager's actions cannot be observed by the owner *and* the manager is risk-averse.

To better understand the subtlety of this answer, it is convenient to consider a simplified version of the problem where the manager's actions are observable. This represents a best-case scenario for the owner and identifies her best possible payoff. Furthermore, this problem provides a benchmark to which the solution of (9)–(11) will be compared. Formally, the simplified problem is defined as follows. Determine a compensation scheme s, which rewards the manager based on his *observed* actions, a production strategy a, and a consumption strategy c to solve the following problem:

$$\max_{a,c,s} \quad V_1^O(X^0; (a, c, s)) \tag{12}$$

$$\text{subject to} \quad V_1^M(X^0; (a, c, s)) \geq -\frac{1}{1-\alpha} \exp[-r(w + (1-\alpha)W_0)]; \tag{13}$$

$$V_1^M(X^0, W_0; (a, c, s)) \geq V_1^M(X^0, W_0; (a, \tilde{c}, s)); \tag{14}$$

$$\text{for every admissible } \tilde{c} = (\tilde{c}_t(X^{t-1}))_{t=1,\ldots,T}.$$

Unlike Problem (9)–(11), the incentive compatibility Constraint (14) for this problem does not involve the production strategy a; the owner does not need to motivate the manager to adopt the desirable production strategy because she can actually monitor his actions. She simply threatens to impose an extreme penalty if the manager deviates from the prescribed actions. The participation Constraint (13) reflects the presence of alternative employment opportunities as in (10). The optimal contract for this problem will be referred to as the *first-best contract* (Mas-Colell et al. 1995) and will be denoted by (a^{fb}, c^{fb}, s^{fb}). Furthermore, the owner's value function under the first-best contract, $V_t^O(X^{t-1}; (a^{fb}, c^{fb}, s^{fb}))$, will be denoted by $V_t^{fb}(X^{t-1})$.

It is instructive to consider how the optimal contract that solves Problem (9)–(11), which is referred to as the *second-best contract* (Mas-Colell et al. 1995), will differ from the first-best contract that solves (12)–(14). Because in Problem (9)–(11) the owner cannot observe the manager's actions, she adopts a compensation scheme that links the manager's reward to observed system outcomes (states). However, this scheme exposes the manager to some risk, and therefore its effectiveness will depend on the manager's risk-aversion. When the manager is risk-neutral, there is no conflict between the owner's desire to compensate the manager using a risky payment scheme and the manager's willingness to accept such a scheme. As a result, first-best can be achieved: The second-best production strategy coincides with the first-best and the owner's expected profit also coincides with her expected profits under first-best. In contrast, when the manager is risk-averse, the owner must pay a *risk premium* in order to entice the manager to accept the risky payments in the performance-based compensation scheme. The risk premium is the amount by which the expected payment to the manager exceeds his expected costs. The final result is a second-best production strategy that may be different than the first-best, and an expected profit for the owner that is lower than her first-best expected profit.

Single-Period vs. Multiperiod Contracts: Commitment and Renegotiation. Problem (9)–(11) focuses on the selection of an optimal long-term contract and assumes that the terms of the contract can only be negotiated at the outset of the relationship. A second question that arises is *whether the initial optimal contract would remain in effect if both the manager and the owner have the right to demand a renegotiation of the contract in a later period t.*

In order to provide an answer, it is first necessary to formally describe the *renegotiation problem.* Suppose that in period t, after both parties have observed the state history X^{t-1} and the manager has learned his wealth W_{t-1}, either party may demand to renegotiate the terms of the contract. In renegotiation, the owner offers a new contract specifying the production strategy, consumption strategy, and compensation scheme that will be in effect from period t onwards. The manager chooses whether to accept this contract or opt for the alternative employment opportunity with a fixed wage w per period. Formally, the owner's period t renegotiation problem is to select a contract $(\hat{a}, \hat{c}, \hat{s})$ to

$$\max_{(a,c,s)} V_t^O(X^{t-1}; (a, c, s)) \tag{15}$$

subject to the participation constraint

$$V_t^M(X^{t-1}, W_{t-1}; (a, c, s)) \geq -\frac{1}{1-\alpha} \exp[-r(w + (1-\alpha)W_{t-1})] \tag{16}$$

and the incentive compatibility constraint

$$V_t^M(X^{t-1}, W_{t-1}; (a, c, s)) \geq V_t^M(X^{t-1}, W_{t-1}; (\tilde{a}, \tilde{c}, s)); \tag{17}$$

$$\text{for every admissible } \tilde{a} = (\tilde{a}_t(X^{t-1}))_{t=1,\ldots,T} \text{ and } \tilde{c} = (\tilde{c}_t(X^t))_{t=1,\ldots,T}.$$

The attentive reader will notice that the renegotiation Problem (15)–(17) is equivalent to the optimal contracting Problem (9)–(11) but with period t replacing Period 1. The owner will wish to renegotiate if the solution to the renegotiation problem improves her net future profits. On the other hand, the manager will wish to renegotiate if, given the current state history X^{t-1} and wealth W_{t-1}, his expected future utility with the original contract will be less than the reservation utility $-1/1 - \alpha \exp[-r\,(w + (1-\alpha)W_{t-1})]$ achieved by the solution to the renegotiation problem; in this case, the manager will wish to default unless the contract is renegotiated. If neither the manager nor the owner wish to renegotiate, then the contract is *immune to renegotiation;* immunity to renegotiation is analogous to the subgame perfection criterion for repeated games in which neither party has an incentive to deviate from his or her equilibrium strategy at any period (Gibbons 1992). The solution method that will be developed in §4 constructs an optimal contract that is immune to renegotiation in each period t and possible history X^{t-1}. Thus, the answer to the question posed earlier is affirmative.

General Remarks. The model presented here is quite general and reflects problems of delegated control that arise in OM. Because our motivating example is the relationship between the owner of a production facility and its manager, in the manager's utility, (2), the cost-of-effort is not a monetary cost that depletes the manager's capital but an inconvenience cost that reduces his leisure time. Nevertheless, the utility Function (2) can be modified to reflect systems where the cost-of-effort is monetary. To do that, one can simply eliminate the cost-of-effort term from the utility Function (2) and subtract it from the wealth Equation (1). A simple change of variables shows that this modification does not affect the analysis; see Plambeck (2000).

An important feature of our model is the banking assumption. Not only does it underlie the analysis to follow, but it is also representative of several applications. For example, when studying the relationship between two parties in the supply chain, it is natural to assume that the impact of the relationship on the performance of the supply chain will depend on the availability of capital that allows each party to take the appropriate actions. The banking assumption allows the interested parties to have access to capital. Implicit in this assumption is the secondary assumption that the manager's access to banking is unlimited, on equal terms with the owner, and that the interest rate is the same for negative and positive balances. This is a simplification made purely for analytical tractability.

4. Overview of Main Results

In this section we present an overview of our main results. We start with a characterization of the first-best contract, then develop a method that identifies a second-best contract, and conclude with general properties of the second-best contract.

4.1. The First-Best Contract

It is convenient to start with the first-best contract because the results will motivate the subsequent developments for the second-best contract. The main finding is that the first-best compensation scheme completely protects the manager from risk. In each period, it offers a fixed wage w, which guarantees that the manager receives his reservation utility and a reimbursement, $g_j(a)$, for the manager's cost-of-effort under the first-best production strategy. The first-best production strategy can be derived using a dynamic-programming recursion. The following proposition provides a formal statement.

Proposition 1. *The owner's value function under the first-best production strategy a^{fb} is derived from the following dynamic-programming recursion:*

$$V_{T+1}^{fb}(j) = \Pi_{T+1}(j),$$
$$V_t^{fb}(j) = \max_{a \in A(j)} \{-g_j(a) - w + \alpha \sum_{k=1}^{N} P_{jk}(a)\ (\pi(k) + V_{t+1}^{fb}(k))\}; \quad (18)$$

for $j = 1, \ldots, N$ and $t = 1, \ldots, T$. Furthermore, the first-best production strategy is given by the actions that maximize the right-hand side of (18), and the first-best compensation scheme s^{fb} is as follows:

$$s_t^{fb}(X^{t-1}) = \alpha^{-1}[w + g_{X_{t-1}}\ (a_t^{fb}(X_{t-1}))]. \quad (19)$$

In the first-best consumption strategy c^{fb}, the manager consumes his Income (19) and interest from his wealth in each period.

The proof is a straightforward generalization of the analogous result for the single-period principal-agent problem (see Grossman and Hart 1983) and is omitted for brevity; see Plambeck (2000) for details.

The dynamic-programming Equation (18) has an intuitive interpretation: In each period t, the owner determines the action a that would maximize the expected future payoffs generated by taking that action *minus* the immediate cost of the compensation provided to the manager in that period. The immediate compensation cost is given by (19). Motivated by this interpretation, one would hypothesize that a similar dynamic-programming recursion would hold for the second-best optimal production strategy, with the single-period compensation cost for action $a \in A(j)$, $g_j(a) + w$, in (18) replaced by the single-period compensation cost of inducing action $a \in A\ (j)$ using a performance-based compensation scheme. This logic is compelling and the following subsection shows that it is valid. However, the validity of this logic depends critically on the assumption that the manager can borrow and lend freely, and that his utility is additively separable and exponential.

4.2. The Second-Best Contract

In order to study the second-best contract that solves (9)–(11), we adopt a two-stage approach that imitates the backward logic of §3. First, the manager's problem is analyzed from the perspective of the owner. Specifically, all production strategies that are incentive- and participation-compatible for some compensation scheme are determined, and then for each of those strategies the compensation scheme that *implements* them at a minimum cost to the owner is determined. In the second step, the analysis determines the best implementable production strategy and through that, the compensation scheme that would implement it at a minimum cost.

This two-step approach leads into an analytically tractable two-step solution technique: In the first-step, a series of single-period nonlinear optimization problems are solved to derive an expression for the single-period cost of implementing each possible production strategy. In the second step, a dynamic-programming recursion analogous to (18) is identified. In the remainder of this section, we present more details about these two steps, outline the main results, and discuss their implications. The proofs can be found in the Appendix.

Step 1: The Manager's Problem. The problem to be analyzed in this step is formally described as follows. Given a production strategy $a = (a_t(X^{t-1}))_{t=1,\ldots,T}$, determine the compensation scheme s and consumption strategy c to maximize the owner's value function:

$$\max_{(c,s)} V_1^O(X^0; (a, c, s)) \quad (20)$$

subject to the participation Constraint (10) and incentive compatibility Constraint (11). The solution to this problem will identify the compensation scheme that implements production strategy a at minimum cost to the owner.

The main finding is that the multiperiod Problem (20), (10)–(11) can be decomposed into a sequence of memoryless single-period problems, and the solution to each single-period problem gives the compensation that should be offered to the manager in each period in order to motivate him to take the action prescribed by strategy a in that period. This compensation depends only on the observed transitions and prescribed actions in each period. By combining the solutions to the single-period problems together, the solution to the multiperiod problem is derived.

A crucial observation that was first formally proved in Fudenberg et al. (1990) makes this finding possible. In essence, because the manager's utility is exponential and the manager has unlimited access to banking, the owner can redistribute and rearrange all the payments so that the participation Constraint (10) is binding both in the first period and in every subsequent period. This implies that, without loss of generality, one can concentrate on compensation schemes with the following property:

$$V_t^M(X^{t-1}, W_{t-1}; s) = -\frac{1}{1-\alpha}\exp\left[-r(w + (1-\alpha)W_{t-1})\right]; \text{ for each } t,\ X^{t-1} \text{ and } W_{t-1}; \tag{21}$$

recall that $V_t^M(X^{t-1}, W_{t-1}; s)$ denotes the manager's optimal value function in period t when he is compensated using scheme s. Expression (21) leads into a streamlined formulation of the incentive compatibility Constraint (11) because it provides closed-form expressions for the value function in a dynamic-programming representation of this constraint. Furthermore, standard dynamic-programming arguments show that Constraint (11) decomposes into a set of independent single-period constraints, one for each period, and thus, Problem (20), (10)–(11) can be reformulated as a series of memoryless single-period problems. A formal proof of this statement is presented in Proposition 5 in the Appendix.

We are now in a position to formalize this general overview. The single-period problem is formulated as follows. For each state $j = 1, \ldots, N$ and action $a \in A(j)$, determine a set of compensation rates $s_{jk}(a)$: $k = 1, \ldots, N$ that solve the following convex optimization problem:

$$\min \sum_{k=1}^{N} P_{jk}(a)s_{jk}(a); \tag{22}$$

$$\sum_{k=1}^{N} P_{jk}(a)\left(-\exp[-r(1-\alpha)(s_{jk}(a) - \alpha^{-1}g_j(a))]\right) = -\exp[-r\alpha^{-1}(1-\alpha)w]; \tag{23}$$

$$\sum_{k=1}^{N} P_{jk}(a')\left(-\exp[-r(1-\alpha)(s_{jk}(a) - \alpha^{-1}g_j(a'))]\right) \leq -\exp[-r\alpha^{-1}(1-\alpha)w] \text{ for each } a' \in A(j); \tag{24}$$

the compensation rates $s_{jk}(a)$ represent a single-period payment to the manager when a transition to state k is observed. The solution to this optimization problem, which will be referred to as the Single-Period Optimization Problem and will be denoted by $SP(j, a)$, gives the compensation scheme that implements action $a \in A(j)$ at a minimum cost to the owner when $T = 1$ and the initial state is j. Constraint (23) is the participation constraint, which states that the if the manager takes action a, then his expected single-period utility will be equal to his reservation utility; this constraint is binding as indicated by (21). Constraint (24) is the incentive compatibility constraint which states that the manager's expected single-period utility under each alternative action $a' \in A(j)$ is lower than his reservation utility, which is attained under action a. Thus, action a represent's the manager's best single-period response to the compensation rates $s_{jk}(a)$. The technical details about the derivation of the constraints are provided in Equations (75)–(77) in the Appendix. The optimal value of this optimization problem is denoted by $z_j^*(a)$, and the set of compensation rates that attain the optimum is denoted $s_{jk}^*(a)$: $k = 1, \ldots, N$. Although the optimum may not always be attainable, the following proposition characterizes conditions under which it is.

PROPOSITION 2. *If the optimization problem $SP(j, a)$ has a nonempty feasible set, then there exists a single-period compensation scheme that implements action $a \in A(j)$ in the single-period problem at a minimum cost to the owner. This compensation scheme solves the optimization problem $SP(j, a)$ and makes a transfer payment $s_{jk}^*(a)$ to the manager* *whenever a transition to state k is observed. The expected single-period cost of compensation is* $z_j^*(a)$.

At this point, it is natural to define the set of implementable policies $A^*(j)$ as the set of all policies $a \in A(j)$ such that the single-period optimization problem $SP(j, a)$ has a feasible solution.

The solutions to the single-period Problem (22)–(24) can be used to construct compensation schemes that implement a given multiperiod production strategy at a minimum cost to the owner. This is explained in the following theorem:

THEOREM 1. *The owner can implement an admissible production strategy* $a = (a_t(X^{t-1}))_{t=1,\ldots,T}$ *if and only if* $a_t(X^{t-1}) \in A^*(X_{t-1})$ *almost surely. If the latter condition holds, then the following compensation scheme implements this production strategy at a minimum-cost to the owner:*

$$s_t(X^t) = s^*_{X_{t-1},X_t}(a_t(X^t)). \tag{25}$$

The expected present value in period $t = 1$ *of all the payments made to the manager under this compensation scheme is*

$$E_a\left[\sum_{t=1}^{T} \alpha^t z^*_{X_{t-1}}(a_t(X^{t-1})) \mid X^0\right]. \tag{26}$$

Step 2: The Owner's Problem. The next step is to determine an implementable production strategy that maximizes the owner's expected profit among all feasible implementable strategies. This involves substituting Expression (26) into the owner's value function $V_1^O(X^0; (a, c, s))$, and solving for the optimal production strategy a. Simple algebra shows that this is equivalent to the following problem: Determine the production strategy a that maximizes

$$E_a\left[\sum_{t=1}^{T} \alpha^t [\pi(X_t) - z^*_{X_{t-1}}(a_t(X^{t-1}))] + \alpha^{T+1}\Pi(X_T) \mid X_0\right]. \tag{27}$$

From this it follows that the second-best production strategy can be derived using a dynamic programming recursion and it is history independent:

THEOREM 2. *(a) Consider the following dynamic-programming recursion:*

$$V^*_{T+1}(j) = \Pi_{T+1}(j)$$

$$V^*_t(j) = \alpha \max_{a\in A^*(j)} \{-z^*_j(a) + \textstyle\sum_{k=1}^N P_{jk}(a)(\pi(k) + V^*_{t+1}(k))\} \tag{28}$$

A memoryless second-best production strategy $a^* = (a^*_t(X_{t-1}))_{t=1,\ldots,T}$ *is given by the actions that maximize the right-hand side of (28). The second-best compensation scheme that implements this production strategy is derived from the single-period optimal compensation schemes* $s^*_{jk}(a)$ *as follows:*

$$s^*_t(X^t) = s_{X_{t-1},X_t}(a^*_t(X_{t-1})). \tag{29}$$

(b) The value function $V^*_t(j)$ *gives the expected future discounted profit for the principal from period t to period T when the system is in state j at the beginning of period t, and the second-best contract solving (28)–(29) is adopted.*

Properties of the Second-Best Contract. The first issue that needs to be addressed is that of existence and uniqueness. In general, the existence of a second-best contract that solves Problem (9)–(11) is not guaranteed. However, the constructive method outlined in Theorems 1 and 2 implies that a second-best contract exists for the dynamic principal-agent model because the action and state space are finite. However, the second-best contract of Theorem 2 is not necessarily unique, and other second-best contracts also exist.

Theorems 1 and 2 specify the compensation scheme s^* and production strategy a^* in a second-best long-term contract triple (a^*, c^*, s^*), but they do not specify the consumption strategy c^* that is implicit in this contract. Because the primary focus of our analysis is on the second-best compensation scheme, an explicit description of the consumption strategy in the second-best contract will not be provided in the main body of the paper. Lemma 1 in the Appendix shows how to construct the consumption strategy c^*.

In contrast to the first-best contract, the second-best contract does not reimburse the manager for his cost-of-effort and thus, exposes him to some risk. In this contract, the payment scheme rewards the manager based on the observed transitions of the Markov chain. Because the owner cannot observe the manager's actions directly, she uses the observed transitions as an *indirect* observation of the hidden actions and bases the payments on them.

A striking feature of the second-best contract derived here is that it can be implemented as a sequence of memoryless single-period contracts. This finding appears to be counterintuitive. Since the owner cannot observe the manager's actions, it is natural to expect that she could use a history-dependent contract to infer his actions, and thus eliminate the information asymmetry that is the major source of the tension in Problem (9)–(11). Intuitively, history-dependent contracts that involve deferred compensation offer an additional leverage to the owner and should increase her expected profits. Although this logic is compelling, it fails in the context of our problem because the manager has access to banking. Banking allows him to borrow in anticipation of any deferred future payments from a history-dependent payment scheme and eliminates this scheme's power to implement actions that would increase the owner's profit compared to her profits under the best history-independent scheme.

Renegotiation. Returning to the issue of single-period versus multiperiod contracts, an important corollary of Theorems 1 and 2 is that the optimal contract (a^*, c^*, s^*) is immune to renegotiation. A formal statement follows:

Corollary 1 (Fudenberg et al. 1990). *The optimal contract* (a^*, c^*, s^*) *of Theorem 2 is immune to renegotiation. That is, for each period t, possible history X^{t-1} and possible wealth W_{t-1}, and for any solution ($\hat{a}$, $\hat{c}$, $\hat{s}$) to the period t renegotiation problem (15)–(17), the following conditions hold:*

$$V_t^M(X^{t-1},\ W_{t-1};\ a^*,\ c^*,\ s^*) = V_t^M(X^{t-1},\ W_{t-1};\ \hat{a},\ \hat{c},\ \hat{s}), \tag{30}$$

and

$$V^O(X^{t-1};\ a^*,\ c^*,\ s^*) = V_t^O(X^{t-1};\ \hat{a},\ \hat{c},\ \hat{s}). \tag{31}$$

This implies that the second-best optimal contract derived in Theorem 2 is optimal for the owner in each period and state; according to the contract theory language this contract is *sequentially optimal.* And, neither the manager nor the owner need to commit to a long-term relationship, as such commitment can be voluntarily enforced by the derived second-best contract; but the two parties must commit to the single-period relationships. In particular, the manager will not default at the beginning of any period because (21) implies that he is indifferent between the alternative employment opportunity and continuing with the initial contract.

Corollary 1 also represents a generalization of Bellman's principle of optimality. Paraphrasing an early statement which appeared in Bellman (1957),

> There exists an optimal second-best contract with the following property: Whatever the initial state and initial decisions for both parties are, the remaining decisions must constitute an optimal contract with regard to the state resulting from the first contract.

5. A Two-State Example: The Maintenance Problem

In this section we provide a precise mathematical analysis of the "maintenance problem" described in §2.

Let 1 denote the operational state and 0 the non-operational one. The profit per period for the owner is π_1 in State 1 and is zero in State 0. The manager is responsible for repairing the machine when in State 0. He can either exercise high effort, denoted by H, or low effort, denoted by L. The cost-of-effort and probability of transition to State 1 are higher when high effort is exercised. In State 1, the manager undertakes routine preventive maintenance activities, denoted by M.

Step 1: The Manager's Problem. The first step in the analysis is to solve three single-period optimization problems: $SP(0, L)$, $SP(0, H)$ and $SP(1, M)$. The solutions to these problems provide the basic ingredients for the second-best compensation scheme. The main results are summarized in the following proposition; the proof involves standard Lagrangian analysis of Problem (22)–(24) and is described in Plambeck (2000).

Proposition 3. *(a) Problems $SP(0, L)$ and $SP(1, M)$ always have a feasible solution. On the other hand, $SP(0, H)$ has a feasible solution if and only if the following condition holds:*

$$\exp\left[-\frac{r(1-\alpha)}{\alpha}\,(g_0(H) - g_0(L))\right] > \frac{P_{00}(H)}{P_{00}(L)}; \tag{32}$$

for brevity of notation let ν denote the left-hand side of (32).

(b) The following table (Table 1) summarizes the solution to the three single-period problems $SP(0, L)$, $SP(0, H)$ and $SP(1, M)$.

Table 1

Initial State j	Terminal State k	Action a	Single-Period Compensation $s^*_{jk}(a)$	Expected Compensation Cost $z^*_j(a)$	
1	*1*	*M*	$\alpha^{-1}[w + g_1(M)]$	$\alpha^{-1}[w + g_1(M)]$	
1	*0*	*M*	$\alpha^{-1}[w + g_1(M)]$		
0	*1*	*L*	$\alpha^{-1}[w + g_0(L)]$	$\alpha^{-1}[w + g_0(L)]$	
0	*0*	*L*	$\alpha^{-1}[w + g_0(L)]$		(33)
0	*1*	*H*	$\alpha^{-1}[w + g_0(H)] + [r(1 - \alpha)]^{-1} \ln\left(\nu \frac{P_{00}(L) - P_{00}(H)}{\nu P_{00}(L) - P_{00}(H)}\right)$	$\alpha^{-1}[w + g_0(H)] + [r(1 - \alpha)]^{-1} \times \left\{P_{00}(H) \ln\left(\nu \frac{P_{01}(L) - P_{01}(H)}{\nu P_{01}(L) - P_{01}(H)}\right)\right.$	
0	*0*	*H*	$\alpha^{-1}[w + g_0(H)] + [r(1 - \alpha)]^{1} \ln\left(\nu \frac{P_{01}(L) - P_{01}(H)}{\nu P_{01}(L) - P_{01}(H)}\right)$	$\left. + P_{01}(H) \ln\left(\nu \frac{P_{00}(L) - P_{00}(H)}{\nu P_{00}(L) - P_{00}(H)}\right)\right\}$	

Part (a) of the Proposition states that it is always possible to implement low effort in state 0 and at the same time offer the manager his reservation utility. Furthermore, Part (b) of the Proposition shows that this can be achieved by a compensation scheme that offers a fixed wage w in each period plus a reimbursement, $g_0(L)$, for the cost-of-effort. The reader will notice that this payment scheme is the same as the payment scheme that would be adopted if the manager's actions were observable; see Expression (19). On the other hand, Part (a) also implies that the owner can implement high effort in State 0 if and only if Expression (32) holds, and Part (b) shows that this is achieved using a performance-based compensation scheme that bases the manager's remuneration on the observed transitions of the Markov chain.

The intuition behind the last two observations is the following: The ability of the owner to design a compensation scheme that implements high effort depends on the manager's aversion to risk, the discount rate, and the informativeness of the observed transitions of the Markov chain. One would expect that as the manager becomes more risk-averse, as the observed transitions become less informative, or as the discount rate increases, then it becomes more difficult for the owner to implement action H. This is reflected in (32). The likelihood ratio $P_{00}(H)/P_{00}(L)$ indicates whether transitions from 0 to 1 are informative about the manager's unobserved actions. If this ratio is close to 0 it follows that transitions from 0 to 1 are a reliable indicator of the manager's unobservable actions and thus, the owner can compensate the manager based on these transitions. If, on the other hand, this ratio is close to 1, these transitions do not contain any information and thus, they cannot be used effectively in a performance-based compensation scheme and high effort cannot be implemented. Expression (32) also captures the effect of the risk-aversion coefficient r and discount rate α.

A more careful examination of the compensation that implements high effort reveals some interesting observations. First, the compensation rates $s^*_{jk}(H)$ described in Table 1 include two components: a deterministic component, α^{-1} $(g_0(H) + w)$, and a performance-based component. The deterministic component coincides with the compensation that would be provided if the manager's actions were observable; see Expression (19). On the other hand, the performance-based component links the manager's compensation to system transitions: If the observed transition is to State 0, then this component is a penalty and thus the total payment is only a fraction of the manager's cost of effort. However, if the transition is to State 1, then the total payment exceeds the cost of effort and motivates the manager to adopt high effort. In addition, the expected cost of the performance-based component is given by the expression

$$[r(1 - \alpha)]^{-1}\left[P_{00}(H)\ \ln\left(\nu\,\frac{P_{01}(L) - P_{01}(H)}{\nu P_{01}(L) - P_{01}(H)}\right) + P_{01}(H)\ \ln\left(\nu\,\frac{P_{00}(L) - P_{00}(H)}{\nu P_{00}(L) - P_{00}(H)}\right)\right], \quad (34)$$

which is positive. This is the risk premium that the owner must pay the manager in order to persuade him to accept the risk in the performance-based compensation scheme. The risk premium converges to zero as the manager becomes risk-neutral ($r \to 0$) or the discount rate becomes large ($\alpha \to 1$). When the discount rate is large, the manager's differential cost for high effort is small compared to his discounted future earnings. Therefore, because the manager has access to banking to smooth his monetary consumption over time, he is effectively neutral to the risk associated with adopting high effort in the current period. The relationship between discounting and risk aversion is explored in greater detail in Fudenberg et al. (1990).

Step 2: The Owner's Problem. With the solution to the single-period Problems (22)–(24) at hand, we now proceed to solve the dynamic-programming equation for the second-best optimal maintenance strategy (28). The main results are summarized in the following proposition.

PROPOSITION 4. *(a) The second-best optimal maintenance strategy $a^* = (a_t^*(X_{t-1}))$ implements high effort in State 0 if and only if the following condition holds:*

$$V_{t+1}^*(1) - V_{t+1}^*(0) \geq \frac{z_0^*(H) - z_0^*(L)}{P_{01}(H) - P_{01}(L)} - \pi_1. \quad (35)$$

(b) As the planning horizon $T \to \infty$, the second-best maintenance strategy converges to a stationary strategy. This strategy adopts high effort in State 0 if and only if the following condition holds,

$$\frac{z_0^*(H)[1 - \alpha(P_{11}(M) - P_{01}(L))]}{P_{01}(H) - P_{01}(L)} - \frac{z_0^*(L)[1 - \alpha(P_{11}(M) - P_{01}(H))]}{P_{01}(H) - P_{01}(L)} \leq \pi_1 - g_1(M) - w. \quad (36)$$

To put these results into perspective, it is convenient to contrast them to the corresponding results for the first-best contract. One can show that in the first-best case, it is optimal to implement high effort in State 0 as the planning horizon $T \to \infty$, if and only if the following condition holds:

$$g_0(H)\ \frac{[1 - \alpha(P_{11}(M) - P_{01}(L))]}{P_{01}(H) - P_{01}(L)} - g_0(L)\ \frac{[1 - \alpha(P_{11}(M) - P_{01}(H))]}{P_{01}(H) - P_{01}(L)} \leq \alpha(\pi_1 - g_1(M)). \quad (37)$$

Comparing (37) to the analogous result for the second-best, (36), the following theme emerges: The conditions under which it is optimal to implement high effort are more stringent when the manager's efforts are unobservable. The intuition is simple and has been reiterated in several instances so far: To implement high effort when the manager's effort is not observable, it is necessary to employ a performance-based compensation scheme that exposes the manager to some risk. Because the manager is risk-averse, he can only be enticed to accept such a scheme if it involves a risk premium. However, the cost of this risk premium makes high effort less desirable for the owner. If this cost is substantial, then high effort may not be optimal in the second-best case, even if it is optimal in the first-best.

To conclude the discussion, we now focus on the additional cost incurred by the owner because the manager's actions are unobservable. This cost, which is referred to as the *agency cost* in the contract theory literature, can be derived by subtracting the owner's second-best value function from her first-best, and a closed-form expression can be derived as the planning horizon $T \to \infty$. The details are straightforward and are omitted. The following expressions give the agency cost when both the first-best and second-best contracts implement high effort in State 0 (i.e., both (32) and (36) hold); similar expressions can be derived for other cases,

$$\lim_{T\to\infty} [V_{T-t}^{fb}(0) - V_{T-t}^*(0)] = \frac{(1 - \alpha P_{11}(M))[\alpha z_0^*(H) - (g_0(H) + w)]}{(1 - \alpha)[1 - \alpha(P_{11}(M) - P_{01}(H))]} \quad (38)$$

$$\lim_{T\to\infty}[V^{fb}_{T-t}(1) - V^{*}_{T-t}(1)] = \frac{\alpha(1 - P_{11}(M))[\alpha z^{*}_{0}(H) - (g_0(H) + w)]}{(1 - \alpha)[1 - \alpha(P_{11}(M) - P_{01}(H))]}$$

$$\text{for } t \text{ fixed.} \tag{39}$$

Three observations are in place here. First, (38)–(39) show that the magnitude of the agency cost depends critically on the cost differential

$$\alpha z^{*}_{0}(H) - (g_0(H) + w) = \frac{\alpha}{r(1 - \alpha)} \times \left\{P_{00}(H)\ \ln\left[\nu\,\frac{(P_{01}(L) - P_{01}(H))}{\nu P_{01}(L) - P_{01}(H)}\right] + P_{01}(H)\ \ln\left[\nu\,\frac{(P_{00}(L) - P_{00}(H))}{\nu P_{00}(L) - P_{00}(H)}\right]\right\}, \tag{40}$$

which gives the risk premium paid by the owner in State 0. Second, the agency cost also depends critically on the amount of time the system spends in states where agency costs are incurred. In our example, agency costs are only incurred in State 0, and thus the agency costs decrease as the probability $P_{11}(M)$ increases; as $P_{11}(M)$ increases, the time spent in the least-costly State 1 also increases. Third, the agency cost becomes zero when the manager becomes risk-neutral ($r \to 0$), which agrees with our earlier assertion that risk-aversion causes the tension in Problem (9)–(11).

6. Concluding Remarks

The principal-agent paradigm has many applications in OM. However, existing principal-agent models are of limited use to OM researchers because they suppress the rich physical structure that is common in OM problems. In this paper, we have developed a dynamic principal-agent model that embeds the physical structure of Markov decision processes into the principal-agent framework. It is our premise that this model can provide the starting point for the analysis of operational systems with delegated control.

Our analysis has established that the proposed model is analytically tractable and that the optimal contracting problem can be solved using a two-step dynamic-programming method. This method identifies an incentive-payment scheme that aligns the objectives of the principal (owner) and the agent (manager), and leads to incentive-efficient outcomes; that is, outcomes that maximize the principal's expected profit and satisfy the agent's incentive compatibility constraints. An application has demonstrated the use of this method. In future studies, we intend to further explore the dynamic principal-agent model in the context of manufacturing and service systems with delegated controls; see Fuloria and Zenios (1999) and Plambeck and Zenios (1999).

The mathematical analysis of operational systems with delegated control is complex and involves several simplifications. In that respect, our modeling framework suffers from the same simplifications that plague the traditional principal-agent models with moral hazard: It assumes that both the principal and the agent have perfect knowledge of each other's preferences and cost structure, and perfect knowledge of the transition probabilities in the Markovian system. In reality, the information available to each party about the other party's preferences and characteristics may be imperfect and thus, learning or sorting phenomena will appear. The conventional wisdom is that dynamic principal-agent models with imperfect information are analytically intractable. However, provided that the principal and agent do have symmetric information, our analysis can be extended to a more general setting in which transition probabilities, the agent's coefficient of risk aversion, and his cost of effort are time and history-dependent. This simply involves solving a larger set of single-period problems (one for each action, time period, and state history) and using the full state-history in the dynamic-programming Recursion (28).

Our modeling framework makes some strong assumptions about the agent's access to banking and his utility. These assumptions can only be defended in light of analytical tractability and because they provide a crude, yet reasonable, approximation of reality. Nevertheless, because of the numerous simplifying assumptions, any conclusions derived from the dynamic principal-agent model should be interpreted cautiously. In general, analyses based on this model should not attempt to provide a robust and complete "solution" to the problem of designing optimal incentive contracts for systems with delegated control. Rather, they should only focus on generalizable and defensible insights.

In particular, the assumptions which enable a dynamic-programming solution for the optimal contract also imply that the principal is indifferent between a long-term contract with one agent and a succession of single-period contracts with a series of identical agents. For example, assuming that the agent's cost-of-effort depends only on the current state of the system precludes (unobservable) capital investment by the agent to reduce his cost of production in subsequent periods. Furthermore, the assumption of common knowledge avoids the issue of the principal learning about the agent's capabilities by observing his performance over time. Therefore, our modeling framework is not applicable to problems in which longevity in the principal-agent relationship adds significant value.

Several important managerial insights can be extracted from the analysis of the dynamic principal-agent model and the two-state example. The unobservability of the agent's actions, combined with his risk aversion, reduce the principal's profits compared to the first-best case. Monitoring the agent's actions can eliminate this "agency loss", and our model provides a means for estimating the value of such a monitoring systems. The set of actions that the principal can implement is also reduced relative to the first-best scenario, because payments to the agent are necessarily based on noisy indicators of his actions. If the agent is risk-averse, the quality of the performance signal is poor, and the incremental effort required for a particular action is sufficiently high, then the principal will not be able to design a compensation scheme that implements this action. This shows that the flexibility of a system to respond to changing environmental conditions is limited when control is delegated.

This last observation supports our basic premise that the dynamic principal-agent model provides richer and more relevant insights than its simpler static counterpart. This is because it demonstrates that agency losses are expected to be higher in systems where flexibility and dynamic control are critical for superior performance. Additional evidence supporting this hypothesis is provided in the companion paper by Plambeck and Zenios (1999), in which a problem of dynamic control in a make-to-stock production system with delegated control is analyzed. The loss of responsiveness caused by delegated control also has implications for the optimal design of dynamic systems. Specifically, our analysis identifies states where the loss of responsiveness is more profound, and thus motivates the design of systems where such states will rarely be occupied.

Our results establish a two-level hierarchy of insights which pertain to the management of dynamic systems with delegated control. In the first level, one establishes insights that are well-known in the agency literature and which are based on the analysis of the static models (22)–(24). In the second level, which is linked to the dynamic-programming Recursion (28), one obtains insights about the impact of control delegation on dynamic operational procedures. These insights cannot be extracted from the analysis of the simpler static agency models.

In conclusion, the proposed model and companion analysis have demonstrated that the performance and operations of dynamic systems depend crucially on whether control is centralized or delegated. Systems with delegated control involve subtle trade-offs that are absent from traditional models for centralized systems. The dynamic principal-agent model can capture these trade-offs, provide insights about strategies that optimally balance them, and quantify the effect of such strategies on system performance. Much of the insight to be gained from the dynamic principal-agent model can also be gained from the simpler static models. Thus, our model justifies the validity of the existing static agency model. However, only the dynamic principal-agent model appears to be relevant for systems where dynamic control is critical for superior performance.[1]

Appendix: Main Proofs

The purpose of this Appendix is to prove Theorems 1 and 2. To that end, we will solve Problem (20), (10)–(11). The analysis will establish the validity of Theorem 1 and will proceed in three steps. In the first

[1]The authors are grateful to Mike Harrison for pointing out the problem of delegation of control in OM, and for his numerous detailed suggestions that improved the exposition in this article. They have also benefited from numerous stimulating discussions with Larry Wein, Prashant Fuloria, and Eran Liron. The paper has also benefited considerably from the comments of two anonymous reviewers.

step, the focus will be on the incentive compatibility Constraint (11). A major obstacle in analyzing this constraint is that the incentive compatible production strategy may depend on the manager's wealth, which is not observable by the owner. Therefore, when solving Problem (20), (10)–(11) the owner may not be able to infer whether the production strategy a is incentive compatible for some compensation scheme s. Lemma 1 will show that given a compensation scheme, the incentive compatible production strategy is wealth-independent. This will be done by showing that the manager's optimal value function for a given compensation scheme decomposes into a wealth-dependent and a wealth-independent component. The wealth-independent component can be solved using a dynamic-programming recursion whose solution determines the incentive compatible production strategy. This implies that the owner can solve Problem (20), (10)–(11). Lemma 1 is a fundamental building block in the subsequent analysis.

In the second step, it will be shown that when analyzing Problem (20), (10)–(11), one can concentrate on compensation schemes satisfying Property (21). This result is stated formally in Proposition 5, and its proof relies on Lemma 2. This lemma shows that the incentives created by any compensation scheme are not affected by changes in the compensation scheme that *either* increase the net present value of the payments made from each period onwards by a constant that depends only on the system history up to that period, *or* redistribute the payments, leaving their net present value unchanged.

Finally, it will be shown that Property (21) streamlines the analysis of Constraints (10)–(11), and decomposes Problem (20), (10)–(11) into a series of T single-period problems described in the statement of Theorem 1. Proposition 2, which characterizes the single-period problems of Theorem 1, is proved as an extension of the proof of Theorem 1.

Lemma 1. *Given a compensation scheme s, the manager's optimal value function decomposes into a wealth-dependent and a wealth-independent component as follows:*

$$V_t^M(X^{t-1}, W_{t-1}; s) = \exp[-r(1-\alpha)W_{t-1}]U_t^M(X^{t-1}; s). \quad \text{(A1)}$$

The wealth-independent component is derived recursively from the dynamic-programming equations:

$$U_{T+1}(X^T; s) = -\frac{1}{(1-\alpha)}\exp(-rw) \quad \text{(A2)}$$

$$U_t(X^{t-1}; s) = \max_{b\in R, a\in A(X_{t-1})}\Bigg\{-\exp[-r(b - g_{X_{t-1}}(a))] + \alpha\sum_{k=1}^N P_{X_{t-1},k}(a)\exp\left[-r(1-\alpha)\left(s_t(X^{t-1},k) - \frac{b}{\alpha}\right)\right] U_{t+1}((X^{t-1},k); s)\Bigg\} \text{ for } t = 1,\ldots,T. \quad \text{(A3)}$$

If $a_t(X^{t-1}; s)$ and $b_t(X^{t-1}; s)$ maximize the right-hand side of the dynamic-programming Recursion (A3), then the production strategy $a = (a_t(X^{t-1}; s))_{t=1,\ldots,T}$ and consumption strategy $c = ((1-\alpha)W_{t-1} + b_t(X^{t-1}; s))_{t=1,\ldots,T}$ are implemented by the compensation scheme s.

Proof. The proof will proceed inductively.

The claim in the lemma is true in period $T+1$: In that period, the manager's value function is equal to his terminal utility, and therefore (2) implies that

$$V_{T+1}^M(X^T, W_T; (a, c, s)) = -\frac{1}{(1-\alpha)}\exp[-r(w + (1-\alpha)W_T)] \quad \text{(A4)}$$

$$= -\frac{1}{(1-\alpha)}\exp(-rw)\exp[-(1-\alpha)W_T)] \quad \text{(A5)}$$

$$= -\exp[-(1-\alpha)W_T]U_{T+1}(X^T; s); \quad \text{(A6)}$$

where $U_{T+1}(X^T; s) = -1/(1-\alpha)\exp(-rw)$.

The induction hypothesis will be that the statement in the lemma holds in period $t+1$, and the proof will establish that the statement also holds in period t. Since the manager's value function $V_t^M(X^{t-1}, W_{t-1}; s)$ solves a stochastic dynamic-programming problem, it can be derived using the dynamic-programming equation:

$$V_t^M(X^{t-1}, W_{t-1}; s) = \max_{c,a\in A(X_{t-1})}\Bigg\{-\exp[-r(c - g_{X_{t-1}}(a))] + \alpha\sum_{k=1}^N P_{X_{t-1},k}(a)V_{t+1}^M\left((X^{t-1},k), s_t(X^{t-1},k) + \frac{1}{\alpha}(W_{t-1} - c); s\right)\Bigg\}; \quad \text{(A7)}$$

the notation $s_t(X^{t-1}, k)$ indicates the compensation payment in period t when the path history $X^t = (X^{t-1}, k)$. By the induction hypothesis,

$$V_{t+1}^M\left((X^{t-1},k), s_t(X^{t-1},k) + \frac{1}{\alpha}(W_{t-1}-c)\right) = \exp\left[-(1-\alpha)\left(s_t(X^{t-1},k) + \frac{1}{\alpha}(W_{t-1}-c)\right)\right] \times U_{t+1}^M((X^{t-1},k); s). \quad \text{(A8)}$$

Substituting (A8) into the dynamic-programming Recursion (A7) implies that

$$V_t^M(X^{t-1}, W_{t-1}; (a, c, s)) = \max_{c,a\in A(X_{t-1})}\Bigg\{-\exp[-r(c - g_{X_{t-1}}(a))] + \alpha\sum_{k=1}^N P_{X_{t-1},k}(a)\exp\left[-(1-\alpha)\left(s_t(X^{t-1},k) + \frac{1}{\alpha}(W_{t-1}-c)\right)\right] U_{t+1}^M((X^{t-1},k); s)\Bigg\} \quad \text{(A9)}$$

Next, if we express the consumption variable c as the as the sum of interest on existing wealth and baseline consumption b:

$$c = b + (1-\alpha)W_{t-1}, \quad \text{(A10)}$$

reformulate the maximization problem in the right-hand side of (A9) in terms of b, and factor out the constant $e^{-r(1-\alpha)W_{t-1}}$, we conclude that

$$V_t^M(X^{t-1}, W_{t-1}; (a, c, s)) = \exp[-r(1-\alpha)W_{t-1}] \max_{b,a\in A(X_{t-1})} \left\{ -\exp[-r(b - g_{X_{t-1}}(a))] + \alpha \sum_{k=1}^{N} P_{X_{t-1},k}(a) \exp\left[-r(1-\alpha)\left(s_t(X^{t-1},k) - \frac{b}{\alpha}\right)\right] U_{t+1}((X^{t-1},k); s)) \right\}. \quad \text{(A11)}$$

This completes the proof. □

Next, we show that the incentives created by a compensation scheme are not affected if the compensation scheme is modified in one of two ways: either by increasing the total payments by a constant, or by redistributing the payments without changing their net present value.

Lemma 2. *Let (a, c, s) be a long-term contract, where the production strategy a and consumption strategy c are assumed to satisfy the usual participation Constraint (10) and incentive compatibility Constraints (11). Consider an alternative compensation scheme that satisfies the following relationship for some period $t \in \{1, \ldots, T\}$ and function $\delta_t(X^{t-1})$:*

$$\sum_{r=t}^{T} \alpha^{\tau-t} s'_\tau(X^\tau) = \sum_{\tau=t}^{T} \alpha^{\tau-t} s_\tau(X^\tau) + \frac{1}{\alpha}\delta_t(X^{t-1}); \text{ almost surely.} \quad \text{(A12)}$$

Then, there exists a consumption strategy c' such that contract (a, c', s') is incentive compatible for the period t renegotiation problem, and

$$U_t(X^{t-1}; s') = \exp[-r(1-\alpha)\delta_t(X^{t-1})]U_t(X^{t-1}; s). \quad \text{(A13)}$$

Proof. The proof relies on the observation that the manager's expected utility depends on the compensation scheme only through his terminal wealth. More specifically, the manager's expected utility from period t to T under the contract (a, c, s) is

$$E_{a,c}\left[-\sum_{\tau=t}^{T} \alpha^{\tau-t} \exp[-rc_\tau(X^{\tau-1}) - g_{X_{\tau-1}}(a_\tau(X^{\tau-1}))] - \frac{\alpha^T}{(1-\alpha)} \exp[-r(w + (1-\alpha)W_T)] \mid X^{t-1}, W_{t-1}\right];$$

where the banking Assumption (1) implies that the manager's terminal wealth under the contract (a, c, s) is related to his wealth, W_{t-1}, at the beginning of period t through the relationship

$$W_T = \alpha^{-(T-t+1)} W_{t-1} + \sum_{\tau=t}^{T} \alpha^{-(T-\tau)} \left(s_\tau(X^{\tau-1}) - \frac{1}{\alpha} c_\tau(X^{\tau-1})\right). \quad \text{(A14)}$$

Consider now the alternative compensation scheme s'. The manager's terminal wealth under this scheme is

$$W'_T = \alpha^{-(T-t+1)} W_{t-1} + \sum_{\tau=t}^{T} \alpha^{-(T-\tau)} \left(s'_\tau(X^{\tau-1}) - \frac{1}{\alpha} c'_\tau (X^{\tau-1})\right); \quad \text{(A15)}$$

where c' is the consumption strategy implemented by this compensation scheme. Now, Equation (A15) implies that

$$W'_T = \alpha^{-(T-t+1)} W_{t-1} + \alpha^{-(T-t)} \left[\sum_{\tau=t}^{T} \alpha^{\tau-t} s'_\tau(X^{\tau-1}) - \frac{1}{\alpha}\sum_{\tau=t}^{T} \alpha^{\tau-t} c'_\tau (X^{t-1})\right] \quad \text{(A16)}$$

$$= \alpha^{-(T-t+1)} (W_{t-1} + \delta_t(X^{t-1})) + \sum_{t=1}^{T} \alpha^{\tau-t} \left(s_\tau(X^{\tau-1}) - \frac{1}{\alpha} c'_\tau(X^{t-1})\right); \quad \text{(A17)}$$

where (A17) follows from (A12). Comparing (A17) to (A14), one notices that the problem facing the manager in period t under contract s' is equivalent to the problem facing him at the same period under contract s, but with his wealth W_{t-1} changed to $W_{t-1} + \delta_t(X^{t-1})$. Lemma 1 implies that the production strategy a which is optimal for the manager under scheme s remains optimal under the modified scheme s', and the only effect of the new compensation scheme is to change the manager's wealth function as follows:

$$V_t^M(X^{t-1}, W_{t-1}; (a, c', s')) = V_t^M(X^{t-1}, W_{t-1} + \delta_t(X^{t-1}); (a, c, s)); \quad \text{(A18)}$$

the consumption strategy c' implemented by the compensation scheme s' is not specified here but can be easily derived from Lemma 1.

Substituting (A1) into the right-hand side and left-hand side of (A18) shows that

$$U_t(X^{t-1}; s') = \exp[-r(1-\alpha)\delta_t(X^{t-1})]\, U_t(X^{t-1}; s). \quad \square$$

We are now in a position to show that when analyzing Problem (20), (10)–(11), it is sufficient to focus on compensation schemes satisfying (21).

Proposition 5. *Suppose that the compensation scheme s implements a. Then there exists an alternative compensation scheme s' that implements the same production strategy a and satisfies the following two conditions:*

(a) The manager's participation constraint is binding in every period t and possible history X^{t-1}

$$V_t^M(X^{t-1}, W_{t-1}; s') = -\frac{1}{1-\alpha} \exp[-r(w + (1-\alpha)W_{t-1})] \quad \text{(A19)}$$

(b) The net present value of the payments made under s' is smaller than the one under s in a sample path sense

$$\sum_{t=1}^{T} \alpha^t s'_t(X^t) \le \sum_{t=1}^{T} \alpha^t s_t(X^t), \text{ almost surely} \quad \text{(A20)}$$

Proof. The logic of the proof is as follows. For each period t and history X^{t-1} we define the *termination payment* $Y_t(X^{t-1}; s)$ to be a transfer payment made from the owner to the manager if and only if the manager decides to terminate the relationship and retire in that period. The termination payment is designed so that the manager is indifferent between retiring in each period and receiving the termination payment, or continuing with the relationship and compensation scheme s. We will use the termination payment to create an alternative compensation scheme s' that maintains incentive compatibility and participation with the production strategy a, and also satisfies Conditions (A19) and (A20). The logic is similar to the one employed in Theorem 3 of Fudenberg et al. (1990).

This logic will be pursued in three steps. First, we will show how to construct the termination payments from the wealth-independent component of the manager's value function. Next, we use these termination payments to construct an alternative compensation scheme s' that satisfies (A12). This compensation scheme will force the manager's value function to satisfy Equation (A19), and it reduces the total payments made by the owner according to Equation (A20).

Let us start with the construction of the termination payments. In period t the manager will be indifferent between continuing or terminating the relationship and receiving the termination payment if his value function is equal to his utility with the termination option. Therefore, the termination payment $Y_t(X^{t-1}; s)$ must satisfy the following equality:

$$V_t^M(X^{t-1}, W_{t-1}; s) = -\frac{1}{(1-\alpha)} \exp(-rw) \exp[-r(1-\alpha)(W_{t-1} + Y_t(X^{t-1}; s))]; \quad \text{(A21)}$$

the right-hand side of (A11) gives the manager's expected utility if he retires and receives the termination payment, and the left-hand side gives his expected utility if he does not retire. Lemma 1 states that

$$V_t^M(X^{t-1}, W_{t-1}; s) = \exp[-r(1-\alpha)W_{t-1}]\, U_t(X^{t-1}; s). \quad \text{(A22)}$$

Substituting (A22) into (A21) gives the following expression for the termination payment,

$$Y_t(X^{t-1}; s) = -\frac{1}{r(1-\alpha)}[\ln(-(1-\alpha)U_t(X^{t-1}; s)) + rw]. \quad \text{(A23)}$$

Notice that this payment is positive when the wealth-independent component of the manager's value function in period t is greater than the reservation level offered by the fixed wage alternative, $-e^{-rw}/(1-\alpha)$, and otherwise it is negative. Furthermore, because the production strategy and consumption strategy implemented by s satisfy the participation Constraint (10), it follows that the termination payment in Period 1 must be

$$Y_1(X^0, s) \geq 0. \quad \text{(A24)}$$

Similarly, because the terminal utility after T is equal to the reservation utility, the termination payment in period $T+1$ is zero.

Now, construct the alternative compensation scheme

$$s_t'(X^t) = s_t(X^t) - \frac{1}{\alpha} Y_t(X^{t-1}; s) + Y_{t+1}(X^t; s). \quad \text{(A25)}$$

This is an admissible compensation scheme since the compensation in each period t depends only on the history of the process up to and including period t. Furthermore, for each period t and process history X^{t-1}, the net present value of the total payments from period t to T under the alternative payment scheme s' satisfies the following relationship (by telescoping sums)

$$\sum_{\tau=t}^{T} \alpha^{\tau-t} s_\tau'(X^\tau) = \sum_{\tau=t}^{T} \alpha^{\tau-t} s_\tau(X^\tau) - \frac{1}{\alpha} Y_t(X^{t-1}; s). \quad \text{(A26)}$$

Expression (A26) satisfies Condition (A12) in Lemma 2, and therefore it follows from the same lemma that there exists a consumption strategy c' such that (a, c', s') is incentive compatible and

$$U_t(X^{t-1}; s') = U_t(x; s) \exp[r(1-\alpha)\, Y_t(X^{t-1}; s)] \quad \text{(A27)}$$

$$= -\frac{1}{1-\alpha} e^{-rw}. \quad \text{(A28)}$$

This, combined with (A1), implies that the alternative payment scheme s' satisfies (A19). To complete the proof, notice that evaluating expression (A26) in Period 1 implies that

$$\sum_{\tau=1}^{T} \alpha^\tau s_\tau'(X^\tau) = \sum_{\tau=1}^{T} \alpha^\tau s_\tau(X^\tau) - \frac{1}{\alpha} Y_1(X^0; s) \quad \text{(A29)}$$

$$\leq \sum_{\tau=1}^{T} \alpha^\tau s_\tau(X^\tau), \quad \text{(A30)}$$

in a sample-path sense; Inequality (A30) follows from (A24). This establishes (A20) and completes the proof. □

The main implication of Proposition 5 is the following: To determine a compensation scheme s that implements an arbitrary production strategy a at a minimum cost to the owner, it is sufficient to focus on compensation schemes that satisfy Condition (21). This is the main idea behind the proof of Theorem 1.

Proof of Theorem 1. The proof will proceed in three steps. In the first step, Proposition 5 will be used to show that the incentive compatibility Constraint (11) is equivalent to the simple inequality constraints in (24), and the participation Constraint (10) is equivalent to the equality constraint in (23). Next, it will be shown that the objective in (20) is equivalent to the total discounted cost of the compensation scheme. And lastly, it will be shown that Problem (20), (10)–(11) decomposes into a series of memoryless single-period problems of the form $SP(j, a)$.

Step 1. According to Proposition 5, it is sufficient to focus on compensation schemes s that satisfy the equality

$$U_t(X^{t-1}; s) = -\frac{\exp(-rw)}{(1-\alpha)} \quad \text{(A31)}$$

for each period t and possible process history X^{t-1}. Therefore, the participation Constraint (10) can be reformulated as

$$U_1(X^0; s) = -\frac{\exp(-rw)}{(1-\alpha)}. \quad \text{(A32)}$$

Let us now turn to the incentive compatibility Constraint (11). According to the dynamic-programming Recursion (A3), this constraint is equivalent to the following statement: The compensation scheme s must be such that the production strategy a solves the following dynamic programming recursion,

$$U_{T+1}(X^T; s) = -\frac{1}{(1-\alpha)} \exp(-rw) \quad \text{(A33)}$$

$$U_t(X^{t-1}; s) = \max_{b \in R, a' \in A(X_{t-1})} \Bigg\{ \exp[-r(b - g_{X_{t-1}}(a'))] + \alpha \sum_{k=1}^{N} P_{X_{t-1},k}(a') \exp\left[-r(1-\alpha)\left(s_t(X^{t-1}, k) - \frac{b}{\alpha}\right)\right] U_{t+1}((X^{t-1}, k); s)\Bigg\},\ t = 1, \ldots, T. \tag{A34}$$

Furthermore, since it is sufficient to focus on compensation schemes that satisfy (A31), Equations (A33)–(A34) can be reformulated as follows. The compensation scheme s must satisfy the following inequalities for each period t and possible process history X^{t-1}:

$$-\frac{1}{(1-\alpha)} \exp(-rw) = \max_{b \in R} \Bigg\{ \exp[-r(b - g_{X_{t-1}}(a_t(X^{t-1})))] - \frac{\alpha \exp(-rw)}{(1-\alpha)} \times \tag{A35}$$

$$\sum_{k=1}^{N} P_{X_{t-1},k}(a_t(X^{t-1})) \exp\left[-r(1-\alpha)\left(s_t(X^{t-1}, k) - \frac{b}{\alpha}\right)\right]\Bigg\}, \tag{A36}$$

and

$$-\frac{1}{(1-\alpha)} \exp(-rw) \geq \max_{b \in R} \Bigg\{ \exp[-r(b - g_{X_{t-1}}(a')] - \frac{\alpha \exp(-rw)}{(1-\alpha)} \times \tag{A37}$$

$$\sum_{k=1}^{N} P_{X_{t-1},k}(a') \exp[-r(1-\alpha)(s_t(X^{t-1}, k) - b/\alpha)]\Bigg\},\quad \text{for every } a' \in A(X_{t-1}). \tag{A38}$$

Straightforward but tedious algebra shows that (A36) reduces to

$$\sum_{k=1}^{N} P_{X_{t-1},k}(a_t(X^{t-1})) \exp[-r(1-\alpha)s_t(X^{t-1}, k)] = \exp\left[-r\frac{(1-\alpha)}{\alpha}(w + g_{X_{t-1}}(a_t(X^{t-1})))\right], \tag{A39}$$

and (A38) reduces to

$$\sum_{k=1}^{N} P_{X_{t-1},k}(a') \exp[-r(1-\alpha)s_t(X^{t-1}, k)] \geq \exp\left[-r\frac{(1-\alpha)}{\alpha}(w + g_{X_{t-1}}(a'))\right], \text{ for every } a' \in A(X_{t-1}). \tag{A40}$$

Constraints (A39)–(A40) provide a streamlined representation of Constraints (10)–(11). This completes Step 1.

Step 2. Next, we reformulate the Objective (20). Since the production strategy a is assumed to be fixed, this objective is equivalent to minimizing the cost of the compensation scheme,

$$E_{c,a}\left[\sum_{t=1}^{T} \alpha^t s_t(X^t) \mid X_0\right]. \tag{A41}$$

Step 3. This implies that Problem (20), (10)–(11) is equivalent to the following: Determine the compensation scheme $s\cdot = (s_t(X^t))_{t=1,\ldots,T}$ to minimize (A41) subject to (A39)–(A40). This problem decomposes into one problem for each period and state-history, and the optimization problem to be solved for each period belongs to the class of single-period problems described in (22)–(24); to see this, notice that for each period t and history X^{t-1}, the Constraints (A39)–(A40) are equivalent to the constraints for the single-period problem $SP(X_{t-1}, a_t(X^{t-1}))$. This completes the proof. □

It now remains to prove Proposition 2.

Proof of Proposition 2. The statement that the solution to $SP(j, a)$ solves the single-period problem follows directly from the proof of Theorem 1. To prove the existence we proceed as follows: First, we show that the optimization problem $SP(j, a)$ is bounded below and it is equivalent to a convex optimization problem. The existence result follows from standard convex optimization arguments.

Step 1. Assume that $SP(j, a)$ is feasible, and consider its objective value for a feasible solution $s_{jk}(a)$. Jensen's inequality implies that the following bound holds:

$$\sum_{k=1}^{N} P_{jk}(a)s_{jk}(a) \geq -\frac{1}{r(1-\alpha)} \ln\left[\sum_{k=1}^{K} P_{jk}(a) \exp(-r(1-\alpha)s_{jk}(a))\right] \tag{A42}$$

$$= -\frac{1}{r(1-\alpha)} \ln\left[\exp\left(-\frac{r(1-\alpha)}{\alpha} w + g_j(a)\right)\right] \tag{A43}$$

$$= \frac{w + g_j(a)}{\alpha}; \tag{A44}$$

where (A43) follows because $s_{jk}(a)$ is feasible and thus satisfies (23). This establishes the that optimization problem $SP(j, a)$ is bounded below.

Step 2. The second step is to reformulate $SP(j, a)$ as a convex program. To do that, we use the equality Constraint (23) to eliminate one of the decision variables $s_{jk}(a)$; we choose a variable $s_{jk}(a)$ such that the $P_{jk}(a) \neq 0$. The resulting problem is convex and satisfies the conditions in Theorem 27.3 of Rockafeller (1972), therefore there exists an optimal solution $s^*_{jk}(a)$ that attains the infimum. □

Proof of Theorem 2. This is a straightforward application of the dynamic programming algorithm and is omitted. □

References

Admati, A. R., P. Pfleiderer. 1997. Does it all add up? Benchmarks and the compensation of active portfolio managers. *J. Bus.* **70** 323–350.

Atkinson, A. A. 1979. Incentives, uncertainty and risk in the newsboy problem. *Decision Sci.* **10**341–353.

Bassok, Y., R. Anupindi. 1999. Analysis of supply contracts with commitments and flexibility. *Manufacturing & Ser. Oper. Management,* forthcoming.

Bellman, R. E. 1957. *Dynamic Programming.* Princeton University Press, Princeton, NJ. 83.

Bertsekas, D. P. 1987. *Dynamic Programming: Deterministic and Stochastic Models.* Prentice Hall, Englewood Cliffs, NJ.

Cohen, M. A., N. Agrawal. 1998. An analytical comparison of long and short term contracts. *IIE Trans.*

Filar, J., K. Vrieze. 1996. *Competitive Markov Decision Processes.* Springer-Verlag, New York.

Fudenberg, D., B. Holmstrom, P. Milgrom. 1990. Short-term contracts and long-term agency relationships. *J. Econom. Theory* **51** 1–31.

Fuloria, P. C., S. A. Zenios. 1999. Optimal outcomes-based reimbursement in a health care delivery system. Working Paper, Graduate School of Business, Stanford University, Stanford, CA.

Gibbons, R. 1992. *Game Theory for Applied Economists.* Princeton University Press, Princeton, NJ.

——. 1998. Incentives in organizations. *J. Econom. Perspectives* **12** 115–132.

Grossman, S. J., O. D. Hart. 1983. An analysis of the principal-agent problem. *Econometrica* **51** 7–45.

Hart, O., B. Holmstrom. 1987. The theory of contracts. T. F. Bewley, ed. *Advances in Economic Theory Fifth World Congress.* Cambridge University Press, Cambridge, UK. 97–103.

Holmstrom, B. 1979. Moral hazard and observability. *Bell J. Econom.* **10** 74–91.

——, P. Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* **55** 302–328.

Ingersoll, J. E. 1987. *Theory of Financial Decision Making.* Rowman & Littlefield, Savage, MD.

Lambert, R. 1987. An analysis of the use of accounting and market measures of performance in executive compensation contracts. *J. Accounting Res.* Supplement **25** 85–129.

Lazear, E. P. 1995. *Personnel Economics.* MIT Press, Cambridge, MA.

Mas-Colell, A., M. D. Whinston, J. R. Green. 1995. *Microeconomic Theory.* Oxford University Press, New York.

Mirrlees, J. 1974. Notes on welfare economics, information and uncertainty. M. Balch, D. McFadden, S. Wu, eds. *Essays in Economic Behavior Under Uncertainty.* North-Holland, Amsterdam, Netherlands. 243–58.

Plambeck, E. L. 2000. Essays in Operations Management. Ph.D Thesis, Department of Engineering Economic Systems and Operations Research, Stanford University, Stanford, CA. In preparation.

——, S. A. Zenios. 1999. Incentive efficient control of a make-to-stock production system. Working Paper, Graduate School of Business, Stanford University, Stanford, CA.

Porteus, E. L., S. Whang. 1991. On manufacturing/marketing incentives. *Management Sci.* **37** 1166–1181.

Prendergast, C. 1998. The provision of incentives in firms. Working Paper, The University of Chicago Graduate School of Business, Chicago, IL.

Puterman, M. L. 1994. *Markov Decisions Processes: Discrete Stochastic Dynamic Programming.* John Wiley & Sons, New York.

Radner, R. 1985. Repeated principal-agent games with discounting. *Econometrica* **53** 1173–1198.

Rockafellar, R. T. 1972. *Convex Analysis*. Princeton University Press, Princeton, NJ.

Salanie, B. 1997. *The Economics of Contracts: A Primer.* MIT Press, Cambridge, MA.

Smith, J. E. 1998. Evaluating income streams: A decision analysis approach. *Management Sci.* **44** 1690–1708.

Spear, S. E., S. Srivastava. 1987. On repeated moral hazard with discounting. *Rev. Econom. Stud.* **54** 599–617.

Tsay, A. A., S. Nahmias, N. Agrawal. 1998. Modeling supply chain contracts: A review. S. Tayur, M. Magazine, R. Ganeshan, eds. *Quantitative Models for Supply Chain Management*. Kluwer Academic Publishers, Norwell, MA. 299–336.

——, W. S. Lovejoy. 1999. Quantity flexibility contracts and supply chain performance. *Manufacturing & Service Operations Management* **1**(2) 89–111.

Whittle, P. 1990. *Risk Sensitive Optimal Control*. Wiley, Chichester, UK.

Wolfson, M. 1985. Incentive problems in oil and gas shelter programs. J. W. Pratt and R. J. Zeckhauser eds. *Principals and Agents: The Structure of Business.* Harvard Business School Press, Boston, MA.

The consulting Senior Editor for this manuscript was Mike Harrison. This manuscript was received on April 19, 1999, and was with the authors 126 days for 3 revisions. The average review cycle time was 59.3 days.