

**Long-term, Short-term and Renegotiation: On the Value of Commitment in Contracting**



Patrick Rey; Bernard Salanie

*Econometrica*, Vol. 58, No. 3 (May, 1990), 597-619.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28199005%2958%3A3%3C597%3ALSAROT%3E2.0.CO%3B2-K>

*Econometrica* is currently published by The Econometric Society.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/econosoc.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## LONG-TERM, SHORT-TERM AND RENEGOTIATION: ON THE VALUE OF COMMITMENT IN CONTRACTING

BY PATRICK REY AND BERNARD SALANIE<sup>1</sup>

Long-term relationships are often governed by short-term contracts; this is usually explained by referring to the costs of specifying and enforcing a complete contingent contract. We focus here on the benefits usually associated with long-term commitment, namely the efficiency costs involved when long-term contracts are not available. We prove that renegotiable short-term contracts will implement the long-term optimum in a multi-period principal-agent framework when transfers are not too limited, objectives are conflicting, and there is no relevant asymmetric information at the contracting dates. This last assumption excludes adverse selection models, but not repeated moral hazard models when technologies and preferences are time separable.

KEYWORDS: Multiperiod relationships, commitment, renegotiation, moral hazard.

### 1. INTRODUCTION

IN AN IDEAL ECONOMIC WORLD, parties entering a long-run relationship would sign a complete long-term contract, taking into account all future contingencies. Such a contract would never need to be renegotiated. Many economic relationships, however, are governed by a sequence of contracts of shorter duration; this is the case for labor contracts, regulatory procedures, lender-borrower agreements, procurement contracts, etc. In addition, contracts often include renegotiation clauses, and, even if not, they are sometimes renegotiated. Long-term contracts covering the full length of the relationship are thus not the rule, but a single contract may cover a substantial fraction of the total duration of the relationship.

To explain the emergence of such short-term contracts, and to understand what determines their duration, both the costs and benefits associated with long-term commitment should of course be considered. Williamson (1985), among others, has convincingly suggested that the costs relate mostly to the difficulties of specifying and enforcing a large set of contingencies. Unfortunately, these transaction costs remain a notoriously vague category, despite some promising recent efforts.<sup>2</sup> This paper does not attempt to formalize these costs, but focuses instead on the benefits.

Among the more important advantages are: intertemporal smoothing, when the environment is not stationary (as when some cost must initially be sunk), when the participants discount future in different ways, or when some risks are

<sup>1</sup> We thank S. Weber for his encouragements at the beginning of this work and J. Tirole for helpful discussions on a previous version. We are also grateful to P. Bolton, F. Hahn, J. J. Laffont, G. Laroque, J. Malcomson, F. Spinnewyn, and three anonymous referees for their remarks and criticisms. Any remaining errors are of course ours. A very preliminary version was presented by P. Rey in January, 1986 at the European Winter workshop of the Econometric Society.

<sup>2</sup> See, for instance, Dye (1985) for a first attempt, and Hart-Holmstrom (1987, Section 3.2), for a discussion of the problems encountered.

involved;<sup>3</sup> provision of incentives, when unobservable actions affect the future outcomes of the relationship, or when the parties have some private information, which they would be more reluctant to reveal in the absence of commitment about its ulterior use.<sup>4</sup>

Contracting "on the spot" cannot resolve these intertemporal trade-offs and will thus often be inefficient.<sup>5</sup> It does not necessarily follow that contracts should cover the full duration of the relationship to be efficient: long-run efficiency may a priori be achieved through a sequence of sufficiently long contracts, each of them covering only part of the relationship. In the presence of asymmetric information, however, long-term contracts are likely to dominate shorter contracts since, as is well-known, the incentive problems created by private information are generally best overcome through ex ante commitment to ex post inefficiencies.<sup>6</sup> In other situations, just how long contracts need to be is not clear. Our purpose is to show that short-term contracting, where parties successively negotiate limited-horizon contracts, can be as efficient as long-term contracting when there is no asymmetric information at the recontracting dates.

We consider the following framework. We use a multiperiod agency model, where at the beginning of each period one of the parties (the principal) can propose a contract to the other (the agent), on a take-it-or-leave-it basis. Long-term contracts will refer to contracts signed at the beginning of the first period and covering the whole relationship. Spot contracts will refer to contracts relevant to only the current period, whereas short-term contracts will refer to contracts covering a limited number of periods.

There are two steps in our arguments. First we look for a sequence of (possibly overlapping and renegotiated) short-term contracts which replicate an optimal long-term contract. This is feasible whenever the principal is not constrained in his period-by-period utility transfers to the agent: we refer to this as the surjectivity condition. The second step is to check that this sequence of short-term contracts is dynamically consistent, in the sense of Kydland-Prescott (1977): it must be in the principal's interest to offer these successive short-term contracts, and in the agent's interest to accept them. As mentioned, dynamic consistency generally does not obtain when there is some private information at a contracting date. It turns out that in the absence of asymmetric information at the contracting dates, dynamic consistency obtains whenever the two parties' objectives are conflicting (in a sense which will be made precise). Under these

<sup>3</sup> See, for instance, Clifford-Crawford (1987), Crawford (1988).

<sup>4</sup> If the parties have to decide initially on an investment level, and this investment cannot be contracted upon, the lack of commitment generally leads to underinvestment. For an analysis of long-term contracts with private information, see Roberts (1982) and Baron-Besanko (1984). The absence of commitment is likely to give rise to a ratchet effect, as analyzed by Freixas-Guesnerie-Tirole (1985) and Laffont-Tirole (1988): information is more costly to extract in the first periods because the corresponding rent is expected to be lost afterwards.

<sup>5</sup> In the absence of such trade-offs, then spot contracts may indeed be efficient; see, for instance, Townsend (1982, Section II), Crawford (1988, Section I).

<sup>6</sup> Even if a long-term contract is signed, it may be difficult to prevent renegotiation; for an analysis of renegotiation-proof contracts, see Dewatripont (1986), Tirole (1986), and Hart-Tirole (1988).

three conditions (no private information, surjectivity and conflicting objectives), short-term contracting achieves long-term efficiency.<sup>7</sup>

The paper is organized as follows. We first focus on the case of complete information to clarify the intuition (Section 2). We set up a general, multiperiod agency problem, and present the full-commitment and no-commitment solutions. A first theorem then states the efficiency of short-term contracts, under two alternative sets of assumptions. Both sets consist of surjectivity and conflict conditions, plus a few technical assumptions needed to ensure the existence of optimal contracts; the second set contains a weaker surjectivity condition (local, as opposed to global), at the cost of strengthening the conflict and technical assumptions. Lastly, we indicate how to extend the results when there are exogenous uncertainties or when random schemes are used.

Section 3 extends the analysis to the incomplete information case. We first stress that the first theorem cannot be extended to situations where a party possesses some private information at one of the contracting dates. We then focus on a repeated moral hazard problem, and present the long-term solution; as stressed by Lambert (1983) and Rogerson (1985), it involves intertemporal risk-sharing and, therefore, could not be implemented by spot contracting. A second theorem states the efficiency of short-term contracting under conditions that replicate the previous ones in this new context.

Section 4 includes some final remarks and reviews recent contributions to the literature on repeated moral hazard, showing how to reinterpret them in the light of this analysis. It emphasizes in particular the analogy between short-term and loan contracts, and discusses the role of credit markets in repeated moral hazard problems.<sup>8</sup>

## 2. SYMMETRIC INFORMATION

### A. Basic Framework: Complete Information and Long-term Contracts

We consider a multi-period principal-agent relationship. There are  $T$  periods ( $T$  is a finite integer and  $T \geq 3$ ); at each period  $t$ , the principal and the agent must share a global basket of  $L$  goods. Let  $W(x_1, \dots, x_T)$  and  $U(x_1, \dots, x_T)$  denote the utility levels of the principal and the agent if they agree upon a partition that gives  $x_t$  to the agent at  $t = 1, \dots, T$ . The partition for period  $t$  is required to belong to some set  $X_t = \prod_{l=1}^L X_{tl}$ , where each  $X_{tl} \subset \mathbb{R}$ , and we suppose that the agent gets an outside option  $\underline{x}_t$  at period  $t$  if he refuses to deal with the principal at that period. In the following,  $x^t$  will denote a sequence of partitions  $(x_1, \dots, x_t)$ ; we similarly define  $X^t = X_1 \times \dots \times X_t$ .

<sup>7</sup> Renegotiation is useful in our model and occurs even in the case of complete information, in contrast to models where it takes place only when an exceptional event arises (Harris-Holmstrom (1987)), and to the literature on renegotiation-proof contracts where it never takes place. Renegotiation also occurs when contracts are incomplete, as in Huberman-Kahn (1988), Hart-Moore (1988), Green-Laffont (1988), Maskin-Moore (1988), and can then also be a useful tool (see Aghion-Dewatripont-Rey (1989)).

<sup>8</sup> For a more extensive discussion on this topic, see Chiappori et al. (1988).

The principal is always better off if he can commit himself over the whole duration of the relationship. We therefore first consider long-term contracts.

The principal acts as a Stackelberg leader. At the beginning of the first period, he can propose a long-term contract  $x^T$  that specifies partitions for all future periods. The agent may accept or reject this offer: if he accepts, the contract is implemented, otherwise he obtains his reservation option  $\underline{x}^T$ .

An optimal long-term contract is (one of) the best, from the principal's point of view, among those that satisfy the agent's long-term rationality constraint. Formally, an optimal long-term contract is a solution of:

$$(P^{LT}) \quad \text{Max}_{x^T} W(x^T)$$

subject to

$$x^T \in X^T,$$

$$U(x^T) \geq \underline{U} = U(\underline{x}^T).$$

We will assume in the following that there exists an optimal long-term contract, denoted by  $x^{T*}$ , and we will denote the corresponding principal's and agent's long-term optimal payoffs by  $W^*$  and  $U^*$ .

Note that we implicitly ruled out the possibility of random schemes, which may indeed be useful if the Pareto frontier is not convex: in that case, the principal may be better off by proposing a lottery over several contracts, which need only meet the agent's participation constraint in expected terms. We will indicate at the end of this section how our analysis can be extended when such random schemes are allowed.

### B. Spot Contracts

In contrast to the previous framework, we suppose in this section that the principal and the agent cannot commit themselves to any partition in future periods: therefore, at period  $t$ , the principal can only offer, and the agent can only agree upon, a spot contract specifying the partition  $x_t$  for the current period.

In general, an optimal long-term contract cannot be implemented via such spot contracts. The reason is that a sequence of spot contracts does not allow for intertemporal exchanges: in the principal-agent framework, it must meet an agent's rationality constraint at each period, whereas a long-term contract only needs to satisfy one, intertemporal, rationality constraint. In the case where  $L = 1$  and the objectives are conflicting, the spot rationality constraints even completely determine the outcome of spot contracting: at the last period, the agent will refuse any partition that gives him less utility than  $\underline{x}_T$ , and the principal will not make any offer that gives himself less utility than  $\underline{x}_T$ ; going backwards, the only perfect equilibrium yields  $\underline{x}^T$ .

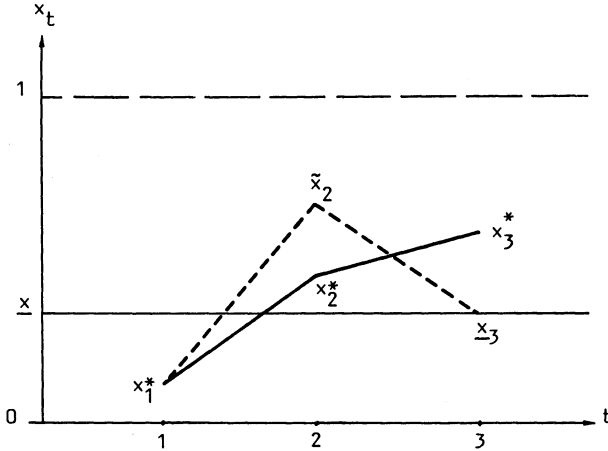


FIGURE 1

C. Short-Term Contracts and Renegotiation

Although our analysis easily extends to less limited commitment, we now assume that the principal and the agent can commit themselves to *two-period* contracts, which can be renegotiated at each period: at period 1, the principal can propose a contract  $({}_1x_1, {}_1x_2)$  that specifies partitions for the first and the second period. The agent then gets  ${}_1x_1$  in period 1 if he accepts the contract, and  $\underline{x}_1$  if he rejects it. A new contract  $({}_2x_2, {}_2x_3)$  may be proposed by the principal at the beginning of period 2; if it is accepted, the agent gets  ${}_2x_2$  in period 2; if it is rejected, the agent gets  ${}_1x_2$  if the previous contract had been accepted, and  $\underline{x}_2$  otherwise. And so on...<sup>9</sup>

The key feature of these contracts is that they overlap: the partition  ${}_t x_{t+1}$  proposed at period  $t$  for period  $(t + 1)$  may be renegotiated at period  $(t + 1)$ . This allows the principal to adjust the agent's rationality constraint so that neither party wants to withdraw from the relationship. The idea will be to look for "promises"  $\tilde{x}_t$  which can implement a long-term optimum  $x^{T*}$ , i.e. such that at each period,  $t$ ,  $({}_t x_t, {}_t x_{t+1}) = (x_t^*, \tilde{x}_{t+1})$  is proposed by the principal and accepted by the agent.

Consider the following example. There are three periods, one unit of numéraire to be shared at each period,  $\underline{x}_1 = \underline{x}_2 = \underline{x}_3 = \underline{x}$ , the objectives are conflicting and the agent only cares about the total quantity he receives:  $x_1 + x_2 + x_3$ . Suppose that the optimal long-term contract satisfies  $x_1^* < x_2^* < x_3^*$ , as in Figure 1 (e.g., the principal discounts the future). Spot contracts cannot implement this optimum: since  $x_3^*$  is larger than the reservation option  $\underline{x}$ , the principal would be a fool to make such an offer in the last period when he can avoid it. Let us now consider short-term (two-period) contracting and suppose  $(x_1^*, \tilde{x}_2)$  has been

<sup>9</sup> We could of course allow both parties to negotiate one-period contracts, in particular in the last period. But this can be anticipated and taken into account when designing the two-period contracts.

accepted in the first period; then the agent will accept the renegotiated contract  $(x_2^*, x_3^*)$  if and only if  $x_2^* + x_3^*$  is larger than  $\tilde{x}_2 + \underline{x}$ . Moreover, if  $\tilde{x}_2$  is chosen such that  $\tilde{x}_2 + \underline{x}$  is exactly equal to  $x_2^* + x_3^*$ , then the best the principal can do in the second period is to offer  $(x_2^*, x_3^*)$ , which will be accepted by the agent. Thus if the principal proposes  $(x_1^*, \tilde{x}_2)$  in the first period (with  $\tilde{x}_2 + \underline{x} = x_2^* + x_3^*$ ) and if the agent accepts this offer, the long-term optimum  $(x_1^*, x_2^*, x_3^*)$  will actually be implemented. This simple example illustrates the basic idea and stresses the two points mentioned in the introduction:

First, is it possible to find an *admissible* promise  $\tilde{x}_2$  in  $X_2$  satisfying the above equality? If, for instance, in the above example

$$X_1 = X_2 = X_3 = [0, 1]$$

(both parties must have a nonnegative consumption in each period),

$$W(x_1, x_2, x_3) = \log(1 - x_1) + (1/4)\log(1 - x_2) + (1/16)\log(1 - x_3),$$

$$U(x_1, x_2, x_3) = x_1 + x_2 + x_3 \text{ and } \underline{x} = 3/4,$$

then the long-term optimum is given by:  $(x_1^*, x_2^*, x_3^*) = (12/28, 24/28, 27/28)$ . Since the adequate promise would have to be  $\tilde{x}_2 = x_2^* + x_3^* - \underline{x} = 30/28 > 1$ , short-term contracts would fail to implement the long-term optimum.

Second, even if such a partition is admissible, will this sequence of short-term contracts implement the long-term optimum (i.e. here, will  $(x_1^*, \tilde{x}_2)$  be proposed and accepted in the first period)?

We now make the argument more formal. The natural approach is to look at the (subgame) perfect equilibria of the game played by the principal and the agent. Let us describe this game more carefully. For each period  $t = 1, \dots, T - 1$ , the principal's and the agent's *choice sets* are respectively  $X'_t = X_t \times X_{t+1}$  and  $D = \{a, r\}$ : the principal makes an offer  $({}_t x_t, {}_t x_{t+1})$ , which the agent may either decide to accept ( $d_t = a$ ) or to reject ( $d_t = r$ ). At the beginning of period  $t$ , the *history* is an element  $h_t \in H_t$ , where  $H_1 = \{\emptyset\}$  and for  $t = 2, \dots, T$ :  $h_t = ({}_1 x_1, {}_1 x_2, d_1, \dots, {}_{t-1} x_{t-1}, {}_{t-1} x_t, d_{t-1})$  and  $H_t = \prod_{\tau=1}^{t-1} (X'_\tau \times D)$ . The *outcome*  $x^t(h_t)$  associated with history  $h_t (t = 2, \dots, T)$  is defined by:

$$\cdot x_1(h_t) = \begin{cases} {}_1 x_1 & \text{if } d_1 = a, \\ \underline{x}_1 & \text{if } d_1 = r, \end{cases}$$

$$\cdot \forall \tau = 2, \dots, t - 1:$$

$$x_\tau(h_t) = \begin{cases} {}_\tau x_\tau & \text{if } d_\tau = a, \\ {}_{\tau-1} x_\tau & \text{if } d_\tau = r \quad \text{and} \quad d_{\tau-1} = a, \\ \underline{x}_\tau & \text{if } d_\tau = r \quad \text{and} \quad d_{\tau-1} = r, \end{cases}$$

$$\cdot x_t(h_t) = \begin{cases} {}_{t-1} x_t & \text{if } d_{t-1} = a, \\ \underline{x}_t & \text{if } d_{t-1} = r. \end{cases}$$

A *principal's strategy* is of the form  $\sigma = (\sigma_1, \dots, \sigma_{T-1})$ , where  $\sigma_t$  is a mapping from  $H_t$  to  $X'_t$ ; an *agent's strategy* is of the form  $\delta = (\delta_1, \dots, \delta_{T-1})$ , where  $\delta_t$  is a

mapping from  $H'_t = H_t \times X'_t$  to  $D$ . A path  $\pi$  is an element of  $H_T$ ; for any  $\pi = (\pi_1, \dots, \pi_{T-1})$  (where  $\pi_t = (x_t, x_{t+1}, d_t) \in X'_t \times D$ ), we will write  $\pi^t = (\pi_1, \dots, \pi_t)$ . The path  $\pi(\sigma, \delta)$  induced by the strategy profile  $(\sigma, \delta)$  is defined by:

$$\begin{aligned} \cdot \pi_1(\sigma, \delta) &= (\sigma_1(0), \delta_1(0, \sigma_1(0))), \\ \cdot \forall t = 1, \dots, T-2: \quad \pi_{t+1}(\sigma, \delta) &= (\sigma_{t+1}(\pi^t), \delta_{t+1}(\pi^t, \sigma_{t+1}(\pi^t))). \end{aligned}$$

Lastly, the *principal's and agent's payoffs* associated with  $(\sigma, \delta)$  are respectively  $\mathscr{W}(\sigma, \delta) = W(x^T(\pi(\sigma, \delta)))$  and  $\mathscr{U}(\sigma, \delta) = U(x^T(\pi(\sigma, \delta)))$ . We can now precisely define what we mean by short-term implementation.

**DEFINITION:** Short-term contracting *implements* the long-term optimum if there exists a subgame perfect equilibrium which implements an optimal long-term contract, and every subgame perfect equilibrium  $(\hat{\sigma}, \hat{\delta})$  gives to the principal his first-best payoff:  $\mathscr{W}(\hat{\sigma}, \hat{\delta}) = W^*$ .

Let us state a first set of assumptions that yields the efficiency of short-term contracts:

A1: The functions  $U$  and  $W$  are continuous on  $X^T$ , which is closed in  $\mathbb{R}^{L \cdot T}$  and contains  $x^T$ . For every  $x_0^T$  in  $X^T$ ,  $\{x^T \in X^T \mid U(x^T) \geq U(x_0^T)\}$  (resp.  $\{x^T \in X^T \mid W(x^T) \geq W(x_0^T)\}$ ) is bounded below (resp. bounded above).<sup>10</sup>

A2: (a) The function  $W$  (resp.  $U$ ) is strictly decreasing (resp. increasing) with respect to all of its arguments. (b) For every  $t = 1, \dots, T$ ,  $X_t$  is convex.

A3:  $\forall \hat{x}_1 \in X_1, \exists \tilde{x}_2 \in X_2$  s.t.:  $U(\hat{x}_1, \tilde{x}_2, \underline{x}_3, \dots, \underline{x}_T) = \underline{U}$ ;  $\forall t = 2, \dots, T-1$ ,  $\forall \bar{x}^t \in X^t, \forall \hat{x}_t \in X_t, \exists \tilde{x}_{t+1} \in X_{t+1}$  s.t.:

$$U(\bar{x}_1, \dots, \bar{x}_{t-1}, \hat{x}_t, \tilde{x}_{t+1}, \underline{x}_{t+2}, \dots, \underline{x}_T) = U(\bar{x}_1, \dots, \bar{x}_t, \underline{x}_{t+1}, \dots, \underline{x}_T).$$

These assumptions ensure that, starting from any point in the game, the principal will always succeed to implement the long-term optimum thanks to adequate promises. More precisely, consider the following programs, for  $t = 1, \dots, T-1$  and  $\bar{x}^t \in X^t$ :

$$(Q_t(\bar{x}^t)) \quad \text{Max}_{(x_t, \dots, x_T)} W(\bar{x}^{t-1}, x_t, \dots, x_T)$$

subject to

$$\forall \tau = t, \dots, T, x_\tau \in X_\tau$$

$$U(\bar{x}^{t-1}, x_t, \dots, x_T) \geq U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T).$$

Starting from some history  $h_t$  which yields the outcome  $\bar{x}^t$ , a solution of program  $(Q_t(\bar{x}^t))$  defines a  $(\bar{x}^t$ -truncated) long-term optimum which clearly

<sup>10</sup> If  $U$  and  $W$  are quasi-concave, the last part of A1 can be replaced with a more standard condition, namely that  $\{U(x) \geq U(x_0)\}$  and  $\{W(x) \geq W(x_0)\}$  have no common direction of recession (see Rockafellar (1970, Theorem 17.3)).



characterizes the best the principal can obtain given that history. The technical assumption A1 ensures that such  $\bar{x}^t$ -optima do exist. Assumption A2(a) expresses some divergence between the principal's and the agent's objectives; together with A2(b), it guarantees that the agent's participation constraint is always binding: all  $\bar{x}^t$ -optima just give  $U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_t)$  to the agent. In other words, the agent has no hope to get more than what he "already" has, i.e. what he can guarantee for himself by refusing any further offer. Lastly, assumption A3, which is a kind of surjectivity assumption, ensures that for any  $t = 1, \dots, T - 2$  and any history  $h_t$ , it is possible to find a promise  $\tilde{x}_{t+1}$  such that the short-term contract  $(x_t, x_{t+1}) = (\hat{x}_t, \tilde{x}_{t+1})$ , where  $\hat{x}_t$  is the first component of the associated  $\bar{x}^t$ -optimum, just "gives" the agent what he already has:

$$U(\bar{x}^{t-1}, \hat{x}_t, \tilde{x}_{t+1}, \underline{x}_{t+2}, \dots, \underline{x}_t) = U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_t).$$

Using then backward induction, it is straightforward to verify that under Assumptions A1–A2–A3, all perfect equilibria are such that: (i) in every period  $t$  and for every outcome  $\bar{x}^t$ , the agent is willing to accept any short-term contract which gives him  $U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_t)$ ; (ii) the principal proposes him a contract which just meets this participation constraint and implements the first component of the associated long-term optimum, thanks to an adequate promise. This ensures that the equilibrium payoff to the principal always is  $W^*$ .

Assumption A3 requires that adequate "promises" exist for every  $t = 2, \dots, (T - 1)$  and any point  $\bar{x}^t$  in  $X^t$ . This may be judged unacceptably strong; indeed, if the agent's utility function is time-separable ( $U(x_1, \dots, x_T) = \sum_{t=1}^T u(x_t)$ ) and the sets  $X_t$  are time invariant ( $X_t = X$ ), A3 implies that the values of the utility function should be unbounded over  $X$  ( $u(X) = \mathbb{R}$ ). One may therefore wonder whether it is possible to only require the existence of such promises along a long-term optimum, as in the following assumption:

A3': *There exists a long-term optimum  $x^{T*}$  such that:  $\forall t = 2, \dots, T - 1, \exists \tilde{x}_t \in X_t$  s.t.*

$$U(x_1^*, \dots, x_{t-1}^*, \tilde{x}_t, \underline{x}_{t+1}, \dots, \underline{x}_T) = \underline{U}.$$

While this "local" form of surjectivity assumption is much weaker than the "global" form A3, it leads to certain technical difficulties: namely, it is no longer possible to exhibit the strategies that support the perfect equilibria of the game out of the equilibrium path; it is even possible, under A1–A2–A3', that no perfect equilibrium exists.<sup>11</sup> To preclude this, we need to slightly strengthen A1 and A2, as follows:

A1': *A1 holds and  $X^T$  is bounded above in  $\mathbb{R}^{LT}$ .*

A2': *A2 holds and: (i)  $U$  is (ordinally) time-separable, or  
(ii)  $\forall t = 2, \dots, T - 1, X_t$  is unbounded below.*

<sup>11</sup> Such a badly behaved example is available upon request.

We can now state a first theorem, whose proof is relegated to the Appendix:

**THEOREM 1:** *Assume A1–A2–A3 or A1'–A2'–A3'; then short-term contracting implements the long-term optimum.*

Note that our analysis rests on the overlapping of the short-term contracts, which will not be possible if one of the contractants (a government, say) changes at some point of the relationship, unless it is possible to credibly commit one's successor.

*D. On the Assumptions*

The two assumptions of conflict and surjectivity are crucial to the result; we will discuss them in turn.

Conflict of interest (Assumption A2 or A2') obtains in most principal-agent models, because of the presence of monetary transfers. To see why it is crucial, assume that it is not satisfied, that there are three periods, and that the program  $(Q_1(x_1))$  has a unique solution  $(\bar{x}_2, \bar{x}_3)$ , and  $U(\bar{x}_1, \bar{x}_2, \bar{x}_3) > U^*$ . Then it would be profitable for the agent to refuse the contract  $(x_1^*, \bar{x}_2)$  at period 1, as the latter would lead him to  $U^*$ . In other words, without the conflict assumption, the absence of long-term commitment may give the agent more negotiation power.<sup>12</sup> It may even lead to inefficiencies. Take the following example:  $T = 4$ ,<sup>13</sup>  $X_t = X = \{H, L\}$ ,  $\underline{x}_t = \underline{x} = L$ , and the payoffs  $(W, U)$  are given by:

<i>HHHH</i> : (0, 0)	<i>HLHH</i> : (0, 1)	<i>LHHH</i> : (0, 1)	<i>LLHH</i> : (0, 0)
<i>HHHL</i> : (0, 1)	<i>HLHL</i> : (3, 0)	<i>LHHL</i> : (0, 0)	<i>LLHL</i> : (0, 1)
<i>HHLH</i> : (1, 1)	<i>HLLH</i> : (0, 0)	<i>LHLH</i> : (0, 0)	<i>LLLH</i> : (0, 1)
<i>HHLL</i> : (0, 0)	<i>HLLL</i> : (0, 1)	<i>LHLL</i> : (0, 1)	<i>LLLL</i> : (2, 0)

A very strong form of surjectivity condition (which implies A3) is satisfied here: all possible utility levels for the agent (i.e., 0 and 1) can be achieved starting from any  $(x_1, x_2, x_3, x_4)$  and changing the “partition” of any single period. However, the outcome of two-period contracting (*LLLL*, which gives 2 to the principal and 0 to the agent) is Pareto dominated by the long-term optimum *HLHL*. The reason the latter cannot be implemented is that as the objectives are “not completely conflicting,” any first period offer  $(H, x_2)$  would guarantee to the agent a utility level strictly greater than 0. If the principal proposes *HL*, the agent can guarantee 1 for himself by refusing any further contract (which yields *HLLL*). If the principal proposes *HH*, the agent can only guarantee 0 for himself

<sup>12</sup> Without the conflict assumption, the outcome of long-term contracting may depend on whether the principal can commit himself not to propose any other contract in the future in case of agent's refusal in the first period. When referring to long-term commitment, we implicitly assumed that the principal can commit not to propose any new contract in periods  $t = 2, \dots, T$ .

<sup>13</sup> For  $T = 3$ , surjectivity alone guarantees the Pareto efficiency of short-term contracting.

by refusing any other contract; he can however obtain 1 by refusing any contract in the second period and accepting to renegotiate in the third period from *HLL*, which gives 0 to both parties, to *HLLH*, which gives 1 to both: there is convergence of interests at this point.

Surjectivity (Assumption A3 or A3') is more problematic. A3 will be satisfied if the agent's utility function goes to infinity at the boundary of the admissible sets  $X_t$ . If the weaker assumption A3' is not satisfied, short-term contracting is inefficient. However one can show, by adapting the proof of Theorem 1, that under the conflictuality assumption A2', the outcome of short-term contracting is efficient with respect to the set of available promises; it will correspond to the program:

$$\text{Max}_{(x^T, \tilde{x}_2, \dots, \tilde{x}_{T-1})} W(x^T)$$

subject to

$$\begin{aligned} x^T &\in X^T; \\ \forall \tau = 2, \dots, T-1: \tilde{x}_\tau &\in X_\tau, \quad U(\underline{x}^T) = U(x_1, \tilde{x}_2, \underline{x}_3, \dots, \underline{x}_T) \\ &= \dots \\ &= U(x_1, \dots, x_{T-2}, \tilde{x}_{T-1}, \underline{x}_T) \\ &= U(x_1, \dots, x_{T-2}, x_{T-1}, x_T). \end{aligned}$$

Therefore, when objectives are conflicting, the efficiency of short-term contracting will only depend on the set of admissible transfers between the principal and the agent (and of course on the way the agent values these transfers: intuitively (this can be made more precise), short-term contracting will be all the more efficient as the time-preferences of both parties are more similar). If transfers are bounded (for example, because of limited liability or wealth constraints—see our discussion of the example at the beginning of Section 2.C), and if the long-term optimum is constrained by the boundary of  $X$ , then even the local assumption A3' may not be satisfied. If the long-term optimum is not constrained, then increasing the length of short-term contracts relative to the time between renegotiation dates will bring the outcome closer to efficiency.<sup>14</sup>

Note that the surjectivity condition is very asymmetric, as it only concerns the party whose bargaining power is zero, i.e. the agent. Obviously, switching roles would mean that the condition must now be satisfied by the principal. Also, a more symmetric condition would be needed if both parties had some bargaining power.

<sup>14</sup> We implicitly assumed that accepting a contract at any period does not affect the agent's outside options in future periods. This is not always the case; in wage contracts for instance, the worker may acquire experience while employed. However, this would not alter the results: the principal only has to take this new dependence into account when designing promises.

*E. Random Schemes and Exogenous Uncertainty*

The analysis can be extended to take into account the possibility of random schemes, i.e. lotteries over partitions. The only difference is that adequate promises have to be designed for every partition in the support of an optimal lottery, and that these promises need give  $\underline{U}$  to the agent only in expected terms. The surjectivity conditions can easily be modified along these lines.

Specifically, the first line in A3 should be replaced with:

$$\forall \hat{x}^T \in X^T, \exists \tilde{x}_2 \text{ s.t. } U(\hat{x}_1, \tilde{x}_2, \underline{x}_3, \dots, \underline{x}_T) = U(\hat{x}^T).$$

As to A3', it should be replaced with:  $\exists \pi^*$ , an optimal lottery over  $X^T$  s.t. for all  $t = 2, \dots, T - 1$ , for all  $\hat{x}^t$  in the support of  $\pi^*$  at  $t$ ,

$$\exists \tilde{x}_t \text{ s.t. } U(\hat{x}^{t-1}, \tilde{x}_t, \underline{x}_{t+1}, \dots, \underline{x}_T) = E_{\pi^*}[U(x^T) | x^{t-1} = \hat{x}^{t-1}].$$

Note also that random schemes, by convexifying the utility frontier, may help to make it decreasing and thus yield conflicting objectives.<sup>15</sup>

The case of exogenous uncertainty can be dealt with similarly, as long as information is symmetric; adequate promises then have to be defined for each state of nature.<sup>16</sup> Note however that risk sharing may add a new motive for consumption smoothing, and thus a new reason for the inefficiency of spot contracts.

3. ASYMMETRIC INFORMATION

Two main types of asymmetric information patterns are usually distinguished: adverse selection, which refers to agent's private information, and moral hazard, which refers to agent's unobservable actions.

It is well-known by now that in the case of private information, long-term optima cannot in general be reached without full commitment. The basic reason is that once some information has been revealed, the contractants would like to change the contract. In particular, both parties might be willing ex-ante to commit to ex-post inefficiencies, thus reducing revelation costs. In the absence of long-term commitment, such a contract would be renegotiated once information has been revealed.

We will therefore concentrate on moral hazard problems. As emphasized, for example, by Lambert (1983) and Rogerson (1985), this situation involves intertemporal risk-sharing, and thus memory: whenever some outcome affects the current wage in a Pareto-optimal contract, it also affects the future wages.

<sup>15</sup> We thank a referee for motivating this development.

<sup>16</sup> See also the theorem in the next section, which obviously applies to exogenous uncertainty (as the special case when the agent cannot choose his actions).

Accordingly, long-term contracts will in general dominate sequences of spot contracts.<sup>17</sup> However, our previous analysis can be extended to the case of pure moral hazard: in the absence of relevant private information at any contracting date, conflict and surjectivity assumptions ensure that long-run optimality can be reached via short-term contracting. Note that the repetition of a moral hazard problem may generate private information patterns, if the agent's hidden action affects the future conditions (probabilities and preferences) of the relationship. Such effects will have to be avoided when introducing moral hazard.

### A. The Model

Let us introduce moral hazard in the previous model. Now at each period  $t$ , an outcome  $r_t$  is jointly determined by an action  $a_t$ , chosen by the agent from some set  $A_t$ , and a random disturbance. There are  $n_t$  different possible outcomes; we will denote by  $R_t$  the set of outcomes. The principal's and the agent's Von Neumann-Morgenstern utility functions are given by  $W(r^T, x^T)$  and  $U(x^T, a^T)$ , where  $r^T \in R^T = R_1 \times \dots \times R_T$ ,  $a^T \in A^T = A_1 \times \dots \times A_T$ , and  $x^T \in X^T$  denotes the vector of transfers from the principal to the agent.

In order to avoid any private information problem at any date in the (re)negotiation game, we need to introduce some separability of probabilities and preferences over time; we will focus on the following case:

(i) random disturbances are serially uncorrelated; for  $t = 1, \dots, T$ ,  $r_t \in R_t$ , and  $a_t \in A_t$ , we will denote by  $p_t(r_t; a_t)$  the probability of obtaining outcome  $r_t$  in period  $t$ , conditional on action  $a_t$ .

(ii) The agent's VNM utility function is additively time separable (multiplicative time-separability would do as well); for  $a^T \in A^T$  and  $x^T \in X^T$ , we will write:  $U(x^T, a^T) = \sum_{t=1}^T U_t(x_t, a_t)$ . At each period  $t$ , the agent's outside option is given by a (nonrandom) payment  $\underline{x}_t$  and the corresponding best action  $\underline{a}_t = \text{Arg max}_{a_t \in A_t} U_t(\underline{x}_t, a_t)$ .

All the above characteristics of the model are supposed to be common knowledge, except the agent's actions which are only observable by the agent and not by the principal; outcomes and payments are publicly observable by both agents and verifiable by some third party.<sup>18</sup> These assumptions are sufficient to exclude private information. Clearly, alternative assumptions might be used to the same effect (see the discussion in Fudenberg-Holmstrom-Milgrom (1990)); they would however lead to a different statement of the conditions which in the following ensure conflict and surjectivity.

<sup>17</sup> Spot contracts might also involve memory. However, this will not be the case for the optimal sequence of spot contracts in a stationary model, in which Rogerson's memory result nevertheless applies to the optimal long-term contracts.

<sup>18</sup> We shall neglect the difference between observable and contractible variables in the following, although this difference may be important in practice (see Chiappori-Macho-Rey-Salanié (1988) for an extended discussion).

*B. Long-term Contracts*

When long-term contracting is available, the principal offers  $x = (x_1(\cdot), \dots, x_T(\cdot))$ , a vector of contingent payments, where  $x_t(\cdot) = \{x_t(r^t) \mid r^t \in R^t\}$  specifies the payments associated in period  $t$  with all sequences of outcome in  $R^t = R_1 \times \dots \times R_t$  (and thus  $x_t(\cdot) \in X_t^{n_1 \dots n_t}$ ).

The agent either accepts or refuses the principal's offer, and chooses  $a = (a_1, \dots, a_T(\cdot))$ , a vector of contingent actions, where

$$a_t(\cdot) = \{a_t(r^{t-1}) \mid r^{t-1} \in R^{t-1}\}$$

represents the choices of actions associated in period  $t$  with all sequences of outcomes  $r^{t-1} (a_t(\cdot) \in A_t^{n_1 \dots n_{t-1}})$ .

We denote  $X^T = X_1^{n_1} \times \dots \times X_T^{n_1 \dots n_T}$  and  $A^T = A_1 \times \dots \times A_T^{n_1 \dots n_{T-1}}$ .

Let us denote respectively by  $U(x, a)$  and  $W(x, a)$  the agent's and the principal's expected utilities associated with a pair  $(x, a)$ :

$$U(x, a) = \sum_{r^T \in R^T} \left( \prod_{t=1}^T p_t(r_t; a_t(r^{t-1})) \right) U(x^T(r^T), a^T(r^{T-1})),$$

$$W(x, a) = \sum_{r^T \in R^T} \left( \prod_{t=1}^T p_t(r_t; a_t(r^{t-1})) \right) W(r^T, x^T(r^T)).$$

The principal's problem is therefore:<sup>19</sup>

$$\text{Max}_{(x, a)} \mathcal{W}(x, a)$$

subject to

$$(x, a) \in X^T \times A^T$$

$$a \in \text{Arg max}_{a' \in A^T} \mathcal{U}(x, a'),$$

$$(P^{LT}) \quad \mathcal{U}(x, a) \geq \underline{U} = U(\underline{x}^T, \underline{a}^T).$$

Let  $(x^*, a^*)$  be a long-term optimum and  $U^*$  and  $W^*$  denote the associated agent's and principal's expected utilities. As previously, this long-term optimum neglects "spot" individual rationality constraints, and it generally involves intertemporal risk-sharing, which makes it impossible to be implemented via "on the spot" negotiations. As we will show, limited commitment may still suffice to achieve long-run efficiency.

*C. Short-term Contracts*

Let us now assume that long-term contracts are no longer available, but that two-period contracts can still be signed at the beginning of each period and renegotiated in the next period. The principal's period  $t$  decision set is thus  $X_t' = X_t^{n_t} \times X_{t+1}^{n_{t+1}}$ ; he makes an offer  $({}_t x_t(r_t), {}_{t+1} x_{t+1}(r_t, r_{t+1}))_{(r_t, r_{t+1}) \in R_t \times R_{t+1}}$  which

<sup>19</sup> We do not assume that whenever the agent is indifferent between actions  $a$  and  $a'$ , he always chooses the most favorable to the principal: this is here a feature of all perfect equilibria.

the agent may either accept or refuse. The agent's decision set is  $D_t = \{\text{accept, refuse}\} \times A_t$ .<sup>20</sup> Definitions of history, strategies, payoffs, and perfect equilibria then follow straightforwardly.

The argument of Section 2 can be transposed immediately: if a surjectivity assumption holds, the principal can at each period devise promises that will, for every state of nature, give to the agent the level of utility he would get under the continuation of the optimal long-term contract. Therefore the agent will behave as he would under the latter contract, and in particular he will choose the right actions. If a conflict assumption also holds, then such a sequence of short-term contracts will indeed be optimal, and the long-term optimum will be implemented. The only difference is that it is now more difficult to give a precise form to the conflict and surjectivity assumptions:

(i) As for conflict, no reasonably weak assumption implies that the individual rationality constraint is binding; only separability of the agent's utility function into action and wages is known to have this property, so we will assume it in A''2 (in an additive form; a multiplicative form would do just as well). We must stress that this is needlessly strong.<sup>21</sup> Note, however, that the participation constraint is assumed not to be binding in efficiency wage models, so that our theorem cannot apply in these cases.

(ii) It would be possible to prove a result under either a local or a global surjectivity assumption; because these would be quite cumbersome to enunciate, we make the "superglobal" assumption in A''3 that the agent's utility function takes all real values.<sup>22</sup> (This also subsumes Assumption A2 in Grossman-Hart (1983), that ensures that the constraint set is never empty.)

We stress, again, that the precise form of these assumptions is not related to the essence of the problem: the important thing is that conflict and surjectivity be satisfied.

Finally, the technical assumptions needed to ensure the existence of an optimal contract under moral hazard are rather strong, even in the static or long-term commitment cases; we adopt those of Grossman-Hart (1983) in A''1.

The following theorem, whose proof can be found in the Appendix, extends the previous result to this moral hazard framework:

**THEOREM 2:** *Assume that for every  $t = 1, \dots, T$ :*

A''1: (i)  $X_t$  is a real open interval  $]m_t, M_t[$  ( $-\infty \leq m_t < M_t \leq +\infty$ ); (ii)  $A_t$  is a compact subset of  $\mathbb{R}^k$  ( $k \in \mathbb{N}$ ); (iii)  $W$  is continuous w.r.t.  $x_t$ ; (iv)  $U_t$  is continuous w.r.t.  $(x_t, a_t)$ ; (v) for every  $r_t$  in  $R_t$ ,  $p_t(r_t, \cdot)$  is continuous and positive.

A''2: (i)  $W$  is strictly decreasing w.r.t.  $x_t$ ; (ii)  $U_t$  is separable into action and wages:  $U_t(x_t, a_t) = u_t(x_t) - v_t(a_t)$ ; (iii)  $u_t$  is strictly increasing and concave.

<sup>20</sup> For  $t = 1, \dots, T-2$ ; for  $t = T-1$ ,  $D_{T-1} = \{\text{accept, refuse}\} \times A_{T-1} \times A_T$ .

<sup>21</sup> In particular, adding a last period where only transfers and consumption take place (i.e. there is no agent's choice of action), as in Malcomson-Spinnewyn (1988) or Fudenberg-Holmstrom-Milgrom (1990), would also ensure conflict.

<sup>22</sup> The strength of this assumption is also enough to ensure that our analysis extends to ex-ante random schemes, where the contract is drawn in a lottery before the agent chooses his action.

A''3:  $u_t(X_t) = \mathbb{R}$ .

Then short-term contracting implements the long-term optimum.

4. CONCLUSION

Our analysis sheds some light on the role of commitment in long-term relationships.

In the complete information case, commitment essentially serves to solve intertemporal trade-offs. It will thus be useful when the contractants cannot solve such trade-offs outside the relationship, for instance through access to credit markets. Indeed, if both parties have free access to perfect credit markets, so that they only care about the discounted sum of their revenues, and not about their temporal profile, then there exists an infinity of optimal contracts, all of which yield the same discounted sum of revenues in every state of nature; at least one of these contracts meets all "spot" rationality constraints and can be implemented via spot contracting.

On the contrary, if at least one party does not have free access to a credit market, then some commitment will generally be necessary to achieve long-term efficiency. However, when objectives are conflicting and in the absence of bounds on utility transfers, limited commitment (commitment for a "near" future) suffices: short-term contracting yields the long-term optimum.

Short-term contracts may be interpreted as loan contracts which enable the principal, acting as a banker, to implement the optimal long-term contract by ignoring the spot rationality constraints. Suppose for instance that there is only one good and that at each period  $t$  the agent has an exogenous endowment  $\underline{x}_t$ . Then short-term contracts  $(x_t^*, \tilde{x}_{t+1})$  can be mimicked by a sequence of one-period loan contracts where at each period  $t$  the principal lends  $(x_t^* - \tilde{x}_t)$ <sup>23</sup> and the agent commits himself to reimburse in the next period  $\underline{x}_{t+1} - \tilde{x}_{t+1}$ . When the agent has access to a credit market (which presupposes some commitment on the agent's side) and this access can be contracted upon, a short-term contract can also be described as a "spot" contract that specifies both consumption and savings or income.<sup>24</sup> Thus in this case "spot contracting" is sufficient for long-term optimality.

The problem is a priori more complex when intertemporal trade-offs in resource allocation interfere with asymmetric information patterns. In particular, in the presence of private information at some contracting date, long-term commitment generally strictly dominates any form of limited commitment. Note that there may still exist short-term contracts which, if they were to be proposed, would lead the agent to act and reveal his information as in the long-term

<sup>23</sup> With the convention  $\tilde{x}_1 = \underline{x}_1$  and  $\tilde{x}_T = x_T^*$ .

<sup>24</sup> For example, the consumption profile  $(x_1^*, \tilde{x}_2)$  may be considered as a "spot" contract including a transfer  $x_1 = (x_1^* - \underline{x}_1) + (1/(1+i))(\tilde{x}_2 - \underline{x}_2)$ , and constrained savings  $(1/(1+i))(\tilde{x}_2 - \underline{x}_2)$ , where  $i$  denotes the interest rate.



optimum; but it will usually not be in the principal's interest to propose these contracts, as they will typically not make a complete use of the information already revealed by the agent.<sup>25</sup>

When there is no relevant private information at any of the contracting dates, we show that conflictuality and the absence of bounds on utility transfers still ensure the long-term optimality of short-term contracting. This result can be related to two other contributions in the repeated moral hazard literature.

Malcomson-Spinnewyn (1988) analyzes a repeated moral hazard problem where the principal has access to a perfect credit market and acts as a banker for the agent. In a model where particular forms of conflict and surjectivity condition are satisfied,<sup>26</sup> they show that a sequence of loan contracts can implement the long-term optimum. Given the analogy mentioned above between short-term contracts and loan contracts, our analysis again may provide a reinterpretation of their result and show how to extend it to the case when the principal does not have access to a credit market.

Fudenberg-Holmstrom-Milgrom (1990) consider situations where both agents have access to a perfect credit market and where the conflict and surjectivity conditions are satisfied, and they focus on sequentially optimal contracts, defined as optimal contracts, the continuation of which is still efficient at the beginning of each period and is equivalent to the outside option for the agent. They show that whenever there is no relevant private information at any of the contracting dates (common knowledge of technology and preferences), there exist optimal long-term contracts which are sequentially optimal and thus, in this sense, may be implemented by spot contracts.<sup>27</sup> When the agent's borrowings and savings cannot be contracted upon, they thus generalize the argument given above for the symmetric information case to any situation where the nonobservability of the agent's consumption does not introduce relevant private information (this in particular excludes wealth effects in moral hazard problems). When the agent's borrowings and savings can be contracted upon, spot and short-term implementation coincide and, clearly, both imply sequential optimality; our analysis thus also provides a reinterpretation of their results, and again shows how they might be extended to situations where credit markets are not available.

*INSEE, 18 blvd. A. Pinard, 75675 Paris Cedex 14, France*

*Manuscript received July, 1988; final revision received July, 1989.*

<sup>25</sup> This argument also implies that optimal long-term contracts would call for renegotiation, which may be difficult to prevent. A possible conjecture is that even where there is asymmetric information, short-term contracting could implement optimal renegotiation-proof contracts.

<sup>26</sup> In Malcomson-Spinnewyn (1988) as in Fudenberg-Holmstrom-Milgrom (1990), the roles of principal and agent are reversed, so that the surjectivity condition bears on the principal, who moreover is supposed to have free access to a perfect credit market. A strong form of surjectivity condition is actually satisfied. As to conflict, see footnote 21.

<sup>27</sup> Short-term implementation implies sequential efficiency; the difference between the two concepts is illustrated by the four-period example presented in Section 2.D, where the long-term optimal contract (with payoffs (3,0)) is sequentially optimal (it is renegotiation-proof and gives  $\underline{U}$  to the agent) but cannot be implemented by short-term contracts.

APPENDIX

PROOF OF THEOREM 1: Because the proofs under the two sets of conditions (A or A') use the same strategy, we will carry them simultaneously, pausing to stress the differences when they arise. The proof centers around the study of the properties of the programs  $(P_t(\bar{x}^t))$  and the functions  $\hat{W}_t: X^t \rightarrow \mathbf{R}$  and  $\hat{U}_t: X^t \rightarrow \mathbf{R}$  recursively defined by:

- (i)  $\hat{W}_T = W, \hat{U}_T = U,$
- (ii)  $\forall t = 1, \dots, T-1, \forall \bar{x}^t \in X^t,$   
 (a)  $(P_t(\bar{x}^t))$  is defined by:

$$\text{Max}_{(x_t, x_{t+1})} \hat{W}_{t+1}(\bar{x}^{t-1}, x_t, x_{t+1})$$

subject to

$$(x_t, x_{t+1}) \in X'_t$$

$$\hat{U}_{t+1}(\bar{x}^{t-1}, x_t, x_{t+1}) \geq \hat{U}_{t+1}(\bar{x}^t, \underline{x}_{t+1}),$$

(b)  $\hat{W}_t(\bar{x}^t)$  is the value of  $(P_t(\bar{x}^t))$ , and

(c)  $\hat{U}_t(\bar{x}^t)$  is the value of  $(P_t(\bar{x}^t))$  for the agent.

Clearly, using backward induction, the family of programs  $(P_t(\bar{x}^t))$ , if they are well-defined (i.e. if solutions to  $(P_t(\bar{x}^t))$  always exist and yield all exactly the same utility level to the agent), characterizes the strategies and payoffs associated to those perfect equilibria where the agent, when indifferent between accepting and rejecting an offer, always accepts (note that the principal may "offer"  $(\bar{x}_t, \underline{x}_{t+1})$ ). More precisely, any pair of strategies  $(\hat{\sigma}, \hat{\delta})$  such that, for every  $t = 1, \dots, (T-1)$  and any  $(h_t, x_t, x_{t+1})$  in  $H'_t$ :

- (i)  $\hat{\sigma}_t(h_t)$  is a solution to  $(P_t(x^t(h_t)))$ ;
  - (ii)  $\hat{\delta}_t(h_t, x_t, x_{t+1}) = a$  if and only if  $\hat{U}_{t+1}(x^{t-1}(h_t), x_t, x_{t+1}) \geq U_{t+1}(x^t(h_t), \underline{x}_{t+1})$ ;
- supports such a perfect equilibrium.

We will now prove that: (1) the  $(P_t)$ ,  $\hat{U}_t$ , and  $\hat{W}_t$  are well-defined, and thus there exists such a perfect equilibrium; (2)  $W_t(\underline{x}_t) = W^*$ , i.e. all perfect equilibria of the above type yield the long-term optimum; (3) no perfect equilibrium involves the agent breaking an indifference situation by refusing the offer presented by the principal: thus all perfect equilibria will yield  $W^*$ .

(1) Validation of the definitions of  $(P_t)$ ,  $\hat{U}_t$ , and  $\hat{W}_t$

Let us first denote by  $W_t^*(\bar{x}^t)$  the value of program  $(Q_t(\bar{x}^t))$ , which defines the  $\bar{x}^t$ -optima.  $W_t^*(\bar{x}^t)$  thus represents the maximal value of the principal's objective, given that the agent can guarantee for himself the equivalent of  $(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_t)$ .

These definitions are obviously non-void under A1: the optimization set is nonempty (it contains  $(\bar{x}_t, \underline{x}_{t+1}, \dots, \underline{x}_T)$ ) and compact, as it may be restricted to

$$\{(x_t, \dots, x_T) \in X_t \times \dots \times X_T \mid U(\bar{x}^{t-1}, x_t, \dots, x_T) \geq U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T) \text{ and}$$

$$W(\bar{x}^{t-1}, x_t, \dots, x_T) \geq W(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_t)\},$$

which from A1 is a compact subset of the closed set  $X_t \times \dots \times X_T$ , and the function  $W$  is continuous. Using Weierstrass's theorem, the functions  $W_t^*$  are therefore well-defined, and they are continuous, by the Maximum Principle.

Moreover, under A2 the individual rationality constraint is always binding at the optimum in  $(Q_t(\bar{x}^t))$ . To see this, suppose it is not and let  $(\hat{x}_t, \dots, \hat{x}_T)$  be a solution of  $(Q_t(\bar{x}^t))$  such that  $U(\bar{x}^{t-1}, \hat{x}_t, \dots, \hat{x}_T) > U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T)$ . From the monotonicity assumption (A2(a)), one has for some  $l \in \{1, \dots, L\}$ , either  $\hat{x}_{lt} > \bar{x}_{lt}$  or, for some  $\tau \geq t+1$ ,  $\hat{x}_{l\tau} > \bar{x}_{l\tau}$ ; but then in either case, from the convexity of the  $x_t$ 's (A2(b)), the principal can decrease the adequate coordinate, thus improving his objective function and leaving the agent with more than his reservation option  $U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T)$ . The constraint is therefore binding, which, because of A2(a), implies that  $W_t^*(\bar{x}^t)$  is strictly decreasing with respect to  $\bar{x}_t$ .

We will now prove recursively on  $t = 1, \dots, T-1$  that under either set of assumptions: the functions  $\hat{W}_t$  and  $\hat{U}_t$  are well-defined and continuous; for any  $\bar{x}^t$  in  $X^t$ ,  $\hat{U}_t(\bar{x}^t) = U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T)$ ;  $\hat{W}_t$  strictly decreases with respect to each component  $\bar{x}_{it}$  of its last argument  $\bar{x}_t$ ; under A1-A2-A3,  $\hat{W}_t(\cdot) = W_t^*(\cdot)$ ; under A2'(i),  $\hat{W}_t$  decreases with respect to all of its arguments.

(a) For every  $\bar{x}^{T-1}$  in  $X^{T-1}$ , the programs  $(P_{T-1}(x^{T-1}))$  and  $(Q_{T-1}(\bar{x}_{T-1}))$  are identical. Thus obviously  $\hat{W}_{T-1}(\bar{x}^{T-1}) = W_{T-1}^*(\bar{x}^{T-1})$  and, as the participation constraint is binding,  $\hat{U}_{T-1}(\bar{x}^{T-1}) = U(\bar{x}^{T-1}, \underline{x}_T)$ . Transferring the properties of  $W_{T-1}^*$  to  $\hat{W}_{T-1}$ , the property holds for  $\tau = T-1$ . (The only new feature is that  $W_{T-1}^*$  is decreasing with respect to all of its arguments when  $U$  is

time-separable, since in  $(Q_{T-1}(x^{T-1}), \bar{x}^{T-2})$  then only affects the objective function, and not the optimization set.)

(b) Assume the above property holds for  $\tau = t + 1, \dots, T - 1$ ; we now prove it also holds for  $\tau = t$ . Here the proof depends on the surjectivity assumption: with the "global" form A3, we will be able to exhibit a solution to  $(P_t)$ , whereas the proof is less direct under the "local" form A3'.

Under A1-A2-A3,  $(P_t(\bar{x}^t))$  can be written as:

$$\text{Max}_{(x_t, x_{t+1})} W_{t+1}^*(\bar{x}^{t-1}, x_t, x_{t+1})$$

subject to

$$(x_t, x_{t+1}) \in X'_t,$$

$$U(\bar{x}^{t-1}, x_t, x_{t+1}, \underline{x}_{t+2}, \dots, \underline{x}_T) \geq U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T).$$

The value  $\hat{W}_t(x')$  of this program cannot exceed  $W_t^*(x')$ . Now consider an  $x'$ -optimum  $(\hat{x}_t, \dots, \hat{x}_T)$ , and a "promise"  $\tilde{x}_{t+1}$  in  $X_{t+1}$  (which exists from A3) such that:

$$U(\bar{x}^{t-1}, \hat{x}_t, \tilde{x}_{t+1}, \underline{x}_{t+2}, \dots, \underline{x}_T) = U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T).$$

One has by the definition of  $W_{t+1}^*$ :

$$W_{t+1}^*(\bar{x}^{t-1}, \hat{x}_t, \tilde{x}_{t+1}) = \text{Max}_{(x_{t+1}, \dots, x_T)} W(\bar{x}^{t-1}, \hat{x}_t, x_{t+1}, \dots, x_T)$$

subject to

$$\forall \tau = t + 1, \dots, T, x_\tau \in X_\tau,$$

$$U(\bar{x}_{t-1}, \hat{x}_t, x_{t+1}, \dots, x_T) \geq U(\bar{x}^{t-1}, \hat{x}_t, \tilde{x}_{t+1}, \underline{x}_{t+2}, \dots, \underline{x}_T)$$

$$= U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T)$$

$$= W_t^*(\bar{x}^t)$$

since  $\hat{x}_t$  is the first component of a  $\bar{x}^t$ -optimum. As  $(\hat{x}_t, \tilde{x}_{t+1})$  satisfies all constraints in  $(P_t(\bar{x}^t))$ ,  $(\hat{x}_t, \tilde{x}_{t+1})$  is thus a solution of  $(P_t(\bar{x}^t))$ , the value of which must therefore be  $\hat{W}_t(x') = W_t^*(\bar{x}^t)$ . Moreover, the agent's participation constraint is binding for any optimum of  $(P_t(\bar{x}^t))$ : since  $W_{t+1}^*$  (resp.  $\hat{U}_{t+1}$ ) is strictly decreasing (resp. increasing) with respect to  $x_{t+1}$ , the same argument we used for  $(Q_t(\bar{x}^t))$  can be applied here. Therefore  $\hat{U}_t(\bar{x}^t)$  is well-defined and equals  $U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T)$ .

Under A1'-A2'-A3', the existence of a solution to  $(P_t(\bar{x}^t))$  easily follows from the continuity of  $\hat{U}_{t+1}$  and the compactness of the optimization set, which is bounded below since it is included in  $\{x^t \in X^T \mid U(x^t) \geq U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T)\}$ , closed and bounded above since the  $X'_\tau$ 's possess these properties. This validates the definition of  $\hat{U}_t$  and  $\hat{W}_t$ , and the continuity of  $\hat{W}_t$  obtains by the standard argument.

Let us prove that the agent's participation constraint is binding at the optimum of  $(P_t(\bar{x}^t))$ :

(i) Under A2'(i),  $\hat{W}_{t+1}$  decreases with respect to all of its arguments; as  $\hat{U}_{t+1}(\bar{x}^{t-1}, x_t, x_{t+1}) = U(\bar{x}^{t-1}, x_t, x_{t+1}, \underline{x}_{t+2}, \dots, \underline{x}_T)$  strictly increases with respect to  $(x_t, x_{t+1})$ , the standard reasoning (based on A2) gives the result.

(ii) Under A2'(ii) as  $\hat{W}_{t+1}(\bar{x}^{t-1}, x_t, x_{t+1})$  decreases with respect to  $x_{t+1}$  and  $X_{t+1}$  is unbounded below, decreasing slightly a component of  $x_{t+1}$  improves the principal's objective and deteriorates that of the agent, so that again the participation constraint must be binding at the optimum.

The tightness of the participation constraint in  $(P_t)$  implies that  $\hat{U}_t$  is continuous and

$$\hat{U}_t(\bar{x}^t) = U(\bar{x}^t, \underline{x}_{t+1}, \dots, \underline{x}_T)$$

and  $\hat{W}_t(\bar{x}^t)$  is a decreasing function of  $\bar{x}_t$ . It remains to be proved that under A2'(i),  $\hat{W}_t$  decreases with respect to all of its arguments. But again, the time-separability of  $U$  implies that  $\bar{x}^{t-1}$  only affects the objective function in  $(P_t)$ , and not the optimization set.

This completes the proof of the property for  $\tau = t$  and thus the validation of  $(P_t)$ ,  $\hat{W}_t$ , and  $\hat{U}_t$  and the proof of the existence of a perfect equilibrium of the short-term contracting game.

$$(2) \hat{W}_1(x_1) = W^*$$

Under A1-A2-A3, this is obvious, since for all  $t$   $\hat{W}_t(\cdot) = W_t^*(\cdot)$ . Under A1'-A2'-A3', let us choose a long-term optimum  $(x_1^*, \dots, x_T^*)$  that satisfies A3' and let the associated promises be  $(x_2, \dots, x_T)$ . It is easy to check that for all  $t = 1, \dots, (T-1)$ ,  $(x_t^*, \bar{x}_{t+1})$  is a solution of  $(P_t(x_t^*, \dots, x_{t-1}^*, \bar{x}_t))$ , so that:

$$\hat{W}_t(x_t^*, \dots, x_{t-1}^*, \bar{x}_t) = \hat{W}_{t+1}(x_t^*, \dots, x_t^*, \bar{x}_{t+1})$$

and thus, since  $W_T(\cdot) = W(\cdot)$ :  $\hat{W}_1(x_1) = W^*$ .

(3) All perfect equilibria implement  $W^*$

Let  $(\hat{\sigma}, \hat{\delta})$  be a perfect equilibrium such that at some point the agent, although indifferent between accepting and refusing an offer which would improve the principal's utility, refuses it. Let  $t$  be the last period where such a refusal occurs, and  $\bar{x}^t$  and  $(x_t, x_{t+1})$  be the corresponding outcome and offer. As in the corresponding subgame the agent never breaks an indifference situation by refusing the offer presented by the principal, the value functions are given by  $\hat{U}_{t+1}$  and  $\hat{W}_{t+1}$ .

By assumption,  $\hat{U}_{t+1}(\bar{x}^{t-1}, x_t, x_{t+1}) = \hat{U}_{t+1}(\bar{x}^t, \bar{x}_{t+1})$  and  $(x_t, x_{t+1}) \neq (\bar{x}_t, \bar{x}_{t+1})$ . As  $\hat{U}_{t+1}$  is strictly increasing, there is at least one component of  $(x_t, x_{t+1})$  that is lower than the corresponding component of  $(\bar{x}_t, \bar{x}_{t+1})$ ; but then the principal should slightly increase this component of the offer, which is possible from the convexity of  $X_t \times X_{t+1}$ : the new offer would be accepted by the agent, which, as  $\hat{W}_{t+1}$  is continuous, would make the principal better off than with the refusal of  $(x_t, x_{t+1})$ .  
*Q.E.D.*

**PROOF OF THEOREM 2:** For every period  $t = 1, \dots, T-1$ , we will denote by  $\bar{\Pi}_t = (\bar{x}_t, \bar{a}_t, \bar{r}_t) \in X_t \times A_t \times R_t$  the realized outcome at period  $t$ , and by  $\bar{\Pi}^t = (\bar{\Pi}_1, \dots, \bar{\Pi}_t)$  the corresponding path; let  ${}^t x \in {}^t X$  and  ${}^t a \in {}^t A$  be the possible "truncated" long-term strategies ( ${}^t X = X_t^{n_t} \times \dots \times X_T^{n_T}$  and  ${}^t A = A_t \times \dots \times A_T^{n_T}$ ) and  $\mathcal{Q}_t({}^t x, {}^t a; \bar{\Pi}^{t-1})$  and  $\mathcal{W}_t(\bar{x}, {}^t a; \bar{\Pi}^{t-1})$  the corresponding expected utility levels, conditional on  $\bar{\Pi}^{t-1}$  (note that  $\bar{\Pi}^{t-1}$  indeed summarizes the relevant information for computing these expected utility levels).

As in the perfect information case, let us define the  $({}_{t-1}x_t(\cdot), \bar{\pi}^{t-1})$ -optima as the solutions of program  $(Q_t({}_{t-1}x_t(\cdot); \bar{\Pi}^{t-1}))$ :

$$\text{Max}_{({}^t x, {}^t a) \in {}^t X \times {}^t A} \mathcal{W}_t({}^t x, {}^t a; \bar{\Pi}^{t-1})$$

subject to

$${}^t a \in \text{Arg max}_{{}^t a \in {}^t A} \mathcal{Q}_t({}^t x, {}^t a; \bar{\Pi}^{t-1}),$$

$$\mathcal{Q}_t({}^t x, {}^t a; \bar{\Pi}^{t-1}) \geq \bar{U}_t({}_{t-1}\bar{x}_t(\cdot); \bar{\Pi}^{t-1}),$$

where

$$\bar{U}_t({}_{t-1}\bar{x}_t(\cdot); \bar{\Pi}^{t-1}) \stackrel{\text{def}}{=} \text{Max}_{a_t \in A_t, r_t \in R_t} \sum p_t(r_t; a_t) U(\bar{x}^{t-1}, {}_{t-1}\bar{x}_t(r_t),$$

$$x_{t+1}, \dots, x_T, a^{t-1}, a_t, a_{t+1}, \dots, a_T).$$

The assumptions of Theorem 2 ensure that all programs  $(Q_t(\cdot))$  have a solution and moreover that the agent's participation constraint is always binding (see Grossman-Hart (1983, Propositions 1 and 2 and Section 6)). The associated value functions for the principal and the agent are thus well-defined; we will denote them by  $W_t^*({}_{t-1}\bar{x}_t(\cdot); \bar{\Pi}^{t-1})$  and  $U_t^*({}_{t-1}\bar{x}_t(\cdot); \bar{\Pi}^{t-1})$  ( $= \bar{U}_t({}_{t-1}\bar{x}_t(\cdot); \bar{\Pi}^{t-1})$ ).

Let us now define the sequence of programs  $(P_{t(t-1)\bar{x}_t(\cdot); \bar{\Pi}^{t-1}})$ , the solutions of which generate the perfect equilibria of the game:

For  $t = T - 1$ :

$$\text{Max}_{(T^{-1}x, T^{-1}a) \in T^{-1}X \times T^{-1}A} W_{T-1}(T^{-1}x, T^{-1}a; \bar{\Pi}^{T-2})$$

subject to

$$T^{-1}a \in \text{Arg max}_{T^{-1}a' \in T^{-1}A} U_{T-1}(T^{-1}x, T^{-1}a'; \bar{\Pi}^{T-2})$$

$$U_{T-1}(T^{-1}x, T^{-1}a; \bar{\Pi}^{T-2}) \geq \text{Max}_{T^{-1}a' \in T^{-1}A} U_{T-1}(T^{-2}\bar{x}_{T-1}, \underline{x}_T, T^{-1}a'; \bar{\Pi}^{T-2}).$$

Let  $\hat{W}_{T-1}(T^{-2}\bar{x}_{T-1}(\cdot); \bar{\Pi}^{T-2})$  and  $\hat{U}_{T-1}(T^{-2}\bar{x}_{T-1}(\cdot); \bar{\Pi}^{T-2})$  denote the corresponding value functions for the principal and for the agent.

For  $t = 1, \dots, T - 2$ :

$$\text{Max}_{(t, x_t, x_{t+1}, a_t) \in (X_t' \times A_t)} \sum_{r_t \in R_t} p_t(r_t; a_t) \hat{W}_{t+1}(x_{t+1}(r_t); \bar{\Pi}^{t-1}, x_t(r_t), a_t, r_t)$$

subject to

$$a_t \in \text{Arg max}_{a_t' \in A_t} \sum_{r_t \in R_t} p_t(r_t; a_t') \hat{U}_{t+1}(x_{t+1}(r_t); \bar{\Pi}^{t-1}, x_t(r_t), a_t', r_t)$$

$$\sum_{r_t \in R_t} p_t(r_t; a_t) \hat{U}_{t+1}(x_{t+1}(r_t); \bar{\Pi}^{t-1}, x_t(r_t), a_t, r_t)$$

$$\geq \text{Max}_{a_t' \in A_t} \sum_{r_t \in R_t} p_t(r_t; a_t') \hat{U}_{t+1}(x_{t+1}; \bar{\Pi}^{t-1}, x_{t-1}\bar{x}_t(r_{t-1}, r_t), a_t', r_t).$$

Again,  $\hat{W}_{t(t-1)\bar{x}_t(\cdot); \bar{\Pi}^{t-1}}$  and  $\hat{U}_{t(t-1)\bar{x}_t(\cdot); \bar{\Pi}^{t-1}}$  will denote the associated value functions.

We know from the discussion opening the proof of Theorem 1 that:

(1) If all functions  $\hat{U}_t$  and  $\hat{W}_t$  are well-defined (and thus all programs  $(P_t(\cdot))$  have solutions, all of them yielding  $\hat{U}_t(\cdot)$  to the agent), then there exists a perfect equilibrium.

(2) If in addition  $\hat{W}_1 = W^*$ , then one perfect equilibrium implements the long-term optimum.

(3) Finally, short-term implementation will hold if no perfect equilibrium involves the agent refusing an offer, favorable to the principal, that leaves him indifferent.

The first two points follow from Lemma 1, proved recursively:

LEMMA 1: For every period  $t = 1, \dots, T - 1$ : (i) all functions  $\hat{U}_t(\cdot)$  and  $\hat{W}_t(\cdot)$  are well-defined, (ii)  $\hat{W}_t(\cdot) = W_t^*(\cdot)$ , (iii)  $\forall \bar{\Pi}^{t-1}, \forall_{t-1}\bar{x}_t(\cdot)$ :

$$\hat{U}_t(x_{t-1}\bar{x}_t(\cdot); \bar{\Pi}^{t-1}) = U_t^*(x_{t-1}\bar{x}_t(\cdot); \bar{\Pi}^{t-1}) (= \bar{U}_t(x_{t-1}\bar{x}_t(\cdot); \bar{\Pi}^{t-1})).$$

PROOF OF LEMMA 1: (a) Conditions (i), (ii), (iii) clearly hold for  $t = T - 1$ , since programs  $(P_{T-1}(\cdot))$  and  $(Q_{T-1}(\cdot))$  coincide.

(b) Let us now assume that the property holds for  $\tau = t + 1, \dots, T - 1$  ( $t \leq T - 2$ ); we show that it then holds for  $\tau = t - 1$ . Let us rewrite  $(P_{t(t-1)x_t; \bar{\Pi}^{t-1}})$  as:

$$\text{Max}_{(t, x_t(\cdot), x_{t+1}(\cdot), a_t) \in X_t' \times A_t} \sum_{r_t \in R_t} p_t(r_t; a_t) W_{t+1}^*(x_{t+1}(r_t); \bar{\Pi}^{t-1}, x_t(r_t), a_t, r_t)$$

subject to

$$\begin{aligned} a_t &\in \text{Arg max}_{a'_t \in A_t} \left\{ \sum_{r_t \in R_t} p_t(r_t; a'_t) \bar{U}_{t+1}(x_{t+1}(r_t, \cdot); \bar{\Pi}^{t-1}, x_t(r_t), a'_t, r_t) \right\}, \\ &\sum_{r_t \in R_t} p_t(r_t; a_t) \bar{U}_{t+1}(x_{t+1}(r_t, \cdot), \bar{\Pi}^{t-1}, x_t(r_t), a_t, r_t) \\ &\geq \text{Max}_{a'_t \in A_t} \left\{ \sum_{r_t \in R_t} p_t(r_t; a'_t) \bar{U}_{t+1}(x_{t+1}; \bar{\Pi}^{t-1}, x_t(\bar{r}_{t-1}, r_t), a'_t, r_t) \right\} \\ &= \bar{U}_t(x_{t-1} \bar{x}_t(\cdot); \bar{\Pi}^{t-1}), \end{aligned}$$

where the last equality follows from the fact that since  $a_{t+1}$  is the agent's choice under  $x_{t+1}$ ,

$$\begin{aligned} &\bar{U}_{t+1}(x_{t+1}; \bar{\Pi}^{t-1}, x_t(\bar{r}_{t-1}, r_t), a'_t, r_t) \\ &= U(\bar{x}^{t-1}, x_{t-1} \bar{x}_t(\bar{r}_{t-1}, r_t), x_{t+1}, \dots, x_T, \bar{a}^{t-1}, a'_t, a_{t+1}, \dots, a_T). \end{aligned}$$

The value of this program (if it exists) cannot exceed  $W_t^*(x_{t-1} \bar{x}_t(\cdot); \bar{\Pi}^{t-1})$ ; we will now exhibit a 3-tuple  $(x_t, x_{t+1}(\cdot), a_t)$  which indeed satisfies all constraints and yields  $W_t^*(x_{t-1} \bar{x}_t(\cdot); \bar{\Pi}^{t-1})$ .

Let  $(\hat{x}, \hat{a}) = (\hat{x}_t, \dots, \hat{x}_T, \hat{a}_t, \dots, \hat{a}_T)$  denote (one of) the solution(s) of  $(Q_t(x_{t-1} x_t(\cdot); \bar{\Pi}^{t-1}))$ . The idea is to find a "promise"  $\bar{x}_{t+1}(\cdot)$  that leads the agent to accept the offer  $(x_t, x_{t+1}(\cdot)) = (x_t(\cdot), \bar{x}_{t+1}(\cdot))$  and to choose the action  $\hat{a}_t$ . Consider the following equation, with unknown  $\bar{x}_{t+1}(\cdot)$ : for all  $r_t \in R_t$  and all  $a_t \in A_t$ ,

$$(E) \quad \bar{U}_{t+1}(\bar{x}_{t+1}; \bar{\Pi}^{t-1}, x_t(r_t), a_t, r_t) = \mathcal{U}_{t+1}(x^{t+1} \hat{x}, x^{t+1} \hat{a}; \bar{\Pi}^{t-1}, \hat{x}_t(r_t), a_t, r_t).$$

Let  $\bar{x}_{t+1}(\cdot)$  be a solution to (E)—we will show later that such a solution exists.

(a) The offer  $(x_t(\cdot), \bar{x}_{t+1}(\cdot))$  induces, if it is accepted, the choice of action  $\hat{a}_t$ , since the incentive constraint in  $(P_t)$  becomes by construction:

$$\begin{aligned} a_t &\in \text{Arg max}_{a'_t \in A_t} \left\{ \sum_{r_t \in R_t} p_t(r_t; a'_t) \mathcal{U}_{t+1}(x^{t+1} \hat{x}, x^{t+1} \hat{a}; \bar{\Pi}^{t-1}, \hat{x}_t(r_t), a'_t, r_t) \right\} \\ &= \text{Arg max}_{a'_t \in A_t} \mathcal{U}_t(x^t \hat{x}, a'_t, x^{t+1} \hat{a}; \bar{\Pi}^{t-1}) \end{aligned}$$

and therefore coincides with the incentive constraint in  $(Q_t(\cdot))$ . Thus  $a_t = \hat{a}_t$  is (one of) the best choice(s) for the agent.

(b) It is indeed accepted by the agent, since (again by construction):

$$\begin{aligned} &\sum_{r_t \in R_t} p_t(r_t; \hat{a}_t) \bar{U}_{t+1}(x_{t+1}; \bar{\Pi}^{t-1}, \hat{x}_t(r_t), \hat{a}_t, r_t) \\ &= \sum_{r_t \in R_t} p_t(r_t; \hat{a}_t) \mathcal{U}_{t+1}(x^{t+1} \hat{x}, x^{t+1} \hat{a}; \bar{\Pi}^{t-1}, \hat{x}_t(r_t), \hat{a}_t, r_t). \end{aligned}$$

As  $(\hat{x}_t, \hat{a}_t)$  is the "first component" of a  $\bar{\Pi}^{t-1}$ -optimum, the right-hand side is just  $\hat{U}_t(x^t \hat{x}, x^{t+1} \hat{a}; \bar{\Pi}^{t-1})$ ; and as the participation constraint is binding in  $(Q_t)$ , it is equal to  $\bar{U}_t(x_{t-1} \bar{x}_t(\cdot); \bar{\Pi}^{t-1})$ , as announced.

(c) Moreover,  $(x^{t+1} \hat{x}(r^t, \cdot), x^{t+1} \hat{a}(r^t, \cdot))$  clearly is a solution of  $(Q_{t+1}(\bar{x}_{t+1}(r^t, \cdot); \bar{\Pi}^{t-1}, \hat{x}_t(r^t), \hat{a}_t(r^{t-1}), r_t))$ , and thus:

$$\begin{aligned} &\sum_{r_t \in R_t} p_t(r_t; \hat{a}_t(r^{t-1})) W_{t+1}^*(\bar{x}_{t+1}(\cdot); \bar{\Pi}^{t-1}, \hat{x}_t(r^t), \hat{a}_t(r^{t-1}), r_t) \\ &= \sum_{r_t \in R_t} p_t(r_t; \hat{a}_t(r^{t-1})) \hat{W}_t(x^{t+1} \hat{x}(r^t, \cdot), x^{t+1} \hat{a}(r^t, \cdot); \bar{\Pi}^{t-1}, \hat{x}_t(r^t), \hat{a}_t(r^{t-1}), r_t) \\ &= \mathcal{W}_t(x^t \hat{x}, x^{t+1} \hat{a}; \bar{\Pi}^{t-1}) \\ &= W_t^*(x_{t-1} \bar{x}_t(\cdot); \bar{\Pi}^{t-1}), \end{aligned}$$

so that  $(\hat{x}_t, \hat{x}_{t+1}, \hat{a}_t)$  fulfills all the constraints in  $(P_t)$  and yields the principal the maximum value  $W_t^*(\bar{r}_{t-1}, \bar{x}_t(\cdot); \bar{\Pi}^{t-1})$ .

Lastly, the existence of a  $\tilde{x}_{t+1}(\cdot)$  solving (E) results from our separability assumptions on  $U(\cdot)$  and from the fact that for every  $t$ ,  $u_t(X_t) = \mathbf{R}$ . It suffices to choose  $\tilde{x}_{t+1}$  depending only on its first argument; (E) then reduces to:

$$\text{for all } r_t \in R_t, \quad K(r_t) + u_{t+1}(\tilde{x}_{t+1}(r_t)) - v_{t+1}(a_{t+1}) = \text{right-hand side,}$$

where  $K(r_t)$  and the right-hand side do not depend upon the unknown  $\tilde{x}_{t+1}(r_t)$ .

This completes the proof of Lemma 1: Since we have exhibited a solution to  $(P_t)$  that attains the value  $W_t^*$ ,  $\hat{W}_t$  is well-defined and equal to  $W_t^*$ ; the standard argument would give us that the participation constraint is binding at the optimum in  $(P_t)$ , so that  $\hat{U}_t$  is well-defined and  $\hat{U}_t(\bar{r}_{t-1}, \bar{x}_t(\cdot); \bar{\Pi}^{t-1}) = \bar{U}_t(\bar{r}_{t-1}, \bar{x}_t(\cdot); \bar{\Pi}^{t-1})$ . Q.E.D.

The conclusion follows as in Theorem 1:

$\hat{W}_t = W_t^*$  for every  $t$  implies that  $W_1 = \hat{W}^*$ , and thus one perfect equilibrium implements the long-term optimum  $W^*$ .

Short-term implementation then follows from arguments analogous to those in the corresponding part of the proof of Theorem 1:

First, suppose that at some point in the game the agent is indifferent between two actions in  $A_t$  and chooses the less favorable to the principal; because the sets  $X_t$  are open, the principal can then slightly perturb the wage schedule so that the agent still accepts the offer (because of payoff-action separability), and chooses an action close to the one the principal preferred most (because of the continuity of his payoff function). As the principal's preferences are continuous, this change improves his payoff in the subgame.

Second, suppose the agent decides to refuse an offer that leaves him indifferent. The same argument as in the proof of Theorem 1, plus the upper semi-continuous dependence of the agent's choice of action with respect to the wage schedule, then ensures that the principal can again break the indifference situation to his advantage. Q.E.D.

## REFERENCES

- AGHION, P., M. DEWATRIPONT, AND P. REY (1989): "Renegotiation Design Under Symmetric Information," mimeo, MIT.
- BARON, D. AND D. BESANKO (1984): "Regulation and Information in a Continuing Relationship," *Information, Economics, and Policy*, 1, 267-302.
- CHIAPPORI, P. A., I. MACHO, P. REY, AND B. SALANIE (1988): "Repeated Moral Hazard: Memory, Commitment, and the Access to Credit Markets," mimeo.
- CLIFFORD, N., AND V. P. CRAWFORD (1987): "Short-term Contracting and Strategic Oil Reserves," *Review of Economic Studies*, 54, 311-23.
- CRAWFORD, V. P. (1988): "Long-term Relationships Governed by Short-Term Contracts," *American Economic Review*, 78, 485-99.
- DEWATRIPONT, M. (1986): "Renegotiation and Information Revelation over Time in Optimal Labor Contracts," mimeo, Harvard University.
- DYE, R. (1985): "Costly Contract Contingencies," *International Economic Review*, 26, 233-250.
- FREEMAN, Y., R. GUBENBERG, AND J. TROLES (1985): "Planning under Incomplete Information and the Ratchet Effect," *Review of Economic Studies*, 52, 173-192.
- FUDENBERG, D., B. HOLMSTROM, AND P. MILGROM (1990): "Short-term Contracts and Long-term Agency Relationships," *Journal of Economic Theory*, forthcoming.
- GREEN, J., AND J. J. LAFFONT (1988): "Contract Renegotiation and the Underinvestment Effect," mimeo, Harvard University.
- GROSSMAN, S., AND O. HART (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, 7-45.
- HARRIS, M., AND B. HOLMSTROM (1987): "On the Duration of Agreements," *International Economic Review*, 28, 389-405.
- HART, O., AND B. HOLMSTROM (1987): "The Theory of Contracts," in *Advances in Economic Theory, Fifth World Congress*, ed. by T. Bewley. Cambridge: Cambridge University Press.
- HART, O., AND J. MOORE (1988): "Incomplete Contracts and Renegotiation," *Econometrica*, 56, 755-785.

- HART, O., AND J. TIROLE (1988): "Contract, Renegotiation and Coasian Dynamics," *Review of Economic Studies*, 55, 509-540.
- HUBERMAN, G., AND C. KAHN (1988): "Limited Contract Enforcement and Strategic Renegotiation," *American Economic Review*, 78, 471-484.
- KYDLAND, F., AND E. PRESCOTT (1977): "Rules Rather than Discretion: the Inconsistency of Optimal Plans," *Journal of Political Economy*, 85, 473-492.
- LAFFONT, J. J., AND J. TIROLE (1988): "The Dynamics of Incentive Contracts," *Econometrica*, 56, 1153-1175.
- LAMBERT, R. (1983): "Long-Term Contracting and Moral Hazard," *Bell Journal of Economics*, 14, 441-452.
- MALCOMSON, J. M., AND F. SPINNEWYN (1988): "The Multiperiod Principal-Agent Problem," *Review of Economic Studies*, 55, 391-407.
- MASKIN, E., AND J. MOORE (1988): "Implementation and Renegotiation," mimeo, Harvard University.
- ROBERTS, K. (1982): "Long-term Contracts," mimeo, University of Warwick, July.
- ROCKAFELLAR, R. T. (1978): *Convex Analysis*. Princeton: Princeton University Press.
- ROGERSON, W. P. (1985): "Repeated Moral Hazard," *Econometrica*, 53, 69-76.
- TIROLE, J. (1986): "Procurement and Renegotiation," *Journal of Political Economy*, 94, 235-259.
- TOWNSEND, R. (1982): "Optimal Multiperiod Contracts and the Gain from Enduring Relationships under Private Information," *Journal of Political Economy*, 90, 1166-1186.
- WILLIAMSON, O. (1985): *The Economic Institutions of Capitalism*. New York: Free Press.