DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

ADVERSE SELECTION AND RENEGOTIATION IN PROCUREMENT

Jean-Jacques Laffont
California Institute of Technology and
GREMAQ, Université de Toulouse

Jean Tirole
Massachusetts Institute of Technology

CALIFORNIA INSTITUTE OF TECHNOLOGY · THE TRUTH SHALL MAKE YOU FREE ·

SOCIAL SCIENCE WORKING PAPER 665

January 1988

ADVERSE SELECTION AND RENEGOTIATION IN PROCUREMENT

Jean-Jacques Laffont and Jean Tirole

ABSTRACT

As was shown by Dewatripont, optimal long-term contracts are generally not sequentially optimal. The parties ex-post renegotiate them to their mutual advantage. This paper fully characterizes the equilibrium of a simple two-period procurement situation and studies the extent to which renegotiation reduces ex-ante welfare: i) A central result is that, like in the non-commitment case, the second period allocation is optimal for the principal conditionally on his posterior beliefs about the agent. ii) The first period allocation exhibits an increasing amount of pooling when the discount factor grows. iii) With a continuum of types, it is never optimal to induce full separation. The paper also analyzes whether renegotiated long-term contracts yield outcomes resembling those under either unrenegotiated long-term contracts or a sequence of short-term contracts and it links the analysis with the multiple unit durable good monopoly problem.

ADVERSE SELECTION AND RENEGOTIATION IN PROCUREMENT*

Jean-Jacques Laffont** and Jean Tirole***
California Institute of Technology
Pasadena, California 91125

1. Introduction

Long-term relationships are optimally run by long-term contracts to which all parties are committed. Commitment prevents these parties from behaving opportunistically ex-post and thus promotes efficient conducts ex-ante. Yet full commitment is an idealized case. The corresponding optimal contracts are generally not sequentially optimal or renegotiation-proof. That is, in the process of implementing a long-term contract, all parties are better off modifying the initial contract (while this renegotiation is ex-post mutually beneficial, the parties would ex-ante like to be able to commit not to renegotiate). The commitment modelling so common in economic theory at best describes an extreme case in which the physical costs of recontracting are important or in which the parties can develop a reputation for refraining from signing mutually advantageous contracts.

This paper investigates the implications of renegotiation in an adverse selection model.[1] Section 2 sets up a simple two-period model of procurement. In each period, the agent realizes a project for the principal. The project's cost in that period depends on a time-invariant adverse selection parameter or type (the agent's ability or the state of technology) and on a cost-reducing effort. The only commonly observable variable is the realized cost in the period. In a static (one-period) framework, the optimal incentive scheme for the principal trades off two conflicting concerns (extract the agent's informational rent and give the latter appropriate incentives to reduce cost), and specifies a reward that decreases with realized cost. With two types (the case considered in most of this paper), the good (efficient) type produces at his socially optimal cost while the bad type's cost exceeds his socially optimal cost in order to reduce the good type's rent. In the twice-repeated relationship, the principal would optimally commit to repeat twice the optimal static scheme. That is, he ought to commit not to alter the first-period incentive scheme in the second period. However, this optimal commitment incentive scheme is not renegotiation proof (Dewatripont (1986)). For, suppose that the agent has produced at the high cost in the first period, demonstrating that he has low ability or that the technology is unfavorable. While the initial

* Research support from the Pew Charitable Trust and the Center for Energy Policy Research at MIT is gratefully acknowledged. We are grateful to Oliver Hart for his helpful discussions.

** California Institute of Technology and GREMAQ, Université de Toulouse

*** Massachusetts Institute of Technology.

contract induces the same inefficiently high cost in the second period, it has become common knowledge that this contract can be renegotiated to benefit both parties by giving more incentives to the agent. But this renegotiation with the bad type towards higher incentives raises the rent of the good type, if the latter mascarades as a bad type in the first period. It thus makes the good type's incentive compatibility constraint in the first period harder to satisfy.

In this paper, we allow commitment in that the two parties sign a long-term contract that is enforced if any of the parties wants it to be enforced. However nothing prevents the parties from agreeing to alter the initial contract. While the optimal contract can w.l.o.g. be designed so as not to be renegotiated in the second period, the renegotiation proofness (RP) requirement restricts the set of allowable second-period contracts. Section 3 demonstrates that there are three kinds of such contracts. In all kinds, the good type produces at his socially optimal cost level. In the first kind, the second-period allocation is that of a *"sell-out" or "fixed-price" contract*; that is, the agent, whatever his type, behaves as if he were residual claimant for his cost savings and produces at this (type-contingent) socially optimal cost. The second kind is *the "conditionally optimal contract."* That is, the agent faces the same incentive contract he would face if the principal were not bound by a previous contract and offered the optimal static contract given his posterior beliefs about the agent's type. The third kind is the class of *"rent-constrained contracts"*, in which the bad agent produces at a cost that is intermediate between his socially optimal cost and his cost in the optimal static contract given the principal's posterior beliefs (the conditionally optimal contract is thus an extreme rent-constrained contract). The principal would like to increase the bad type's cost to reduce the good type's rent, but is unable to do so because he previously offered the good type some rent level.

Section 4 characterizes the optimal intertemporal contract. The second-period contract is conditionally optimal (i.e., is of the second kind). In the first period, only the good type's incentive compatibility constraint is binding (like in the unrenegotiated contract case, but unlike in the no commitment case). The good type is indifferent between revealing his type and mascarading as a bad type. The description of the optimal contracts is therefore rather simple. Incentive constraints are binding as usual and the contracts offered in period 2 are conditionally optimal, i.e., are not distorted by the principal's ability to commit to rents. However, none of these results is obvious. Limits on commitment may, as in Laffont-Tirole (1987), lead to incentive constraints binding in both directions. The ability to commit to rents to mitigate the first period incentive constraints may a priori lead to distortions in second period contracts away from conditionally optimal contracts.

The equilibrium is a separating one only if the discount factor is small (section 5). The equilibrium probability that the good type pools with the bad type increases with the discount factor and converges to one (without ever reaching this value) when the parties become very patient.

Section 6 analyzes the case of a continuum of types. It shows that fully separating the types is feasible but never optimal for the principal.

Section 7 contains a fairly extensive comparison of the findings with those for the same model when the relationship is run by a sequence of short term contracts (Laffont-Tirole (1987, 1988)). We discuss whether the outcome under commitment and renegotiation is intermediate between those under full commitment and under no commitment.

Our work has general implications for adverse selection models. In particular, we ask if the Hart-Tirole (1987) result according to which the optimal long term contract between a buyer and a seller yields the same outcome as Coase's non-commitment durable good model holds when consumption is not binary. The answer is positive if and only if the discount factor is not too high. As a by-product, we compute the multiple unit two period monopoly price discriminating allocation.

2. The Model

a) The Commitment Framework

We consider a two-period model in which a firm (the agent) must, each period, realize a project with a cost structure:

$$c_t = \beta - e_t, \qquad t = 1,2,$$

where e_t is the level of effort exerted by the firm's manager in period t, and β is a parameter known only by the manager, which can take only two values $\underline{\beta}$ and $\bar{\beta}$, with $\bar{\beta} > \underline{\beta}$. Type $\underline{\beta}$ is called the "good type," and type $\bar{\beta}$ is the "bad type."

Each period the manager's utility level is $U \equiv s - \psi(e)$, where s is the net (i.e., in addition to cost) monetary transfer he receives from the regulator and $\psi(e)$ is his disutility of effort, where $\psi(0) = 0$, $\psi' > 0$, $\psi'' > 0$, and, for technical reasons, $\psi''' \geq 0$.[2] Let e^* denote the socially optimal level of effort, defined by the equality between the marginal disutility of effort and the marginal cost savings:

$$\psi'(e^*) = 1.$$

The socially optimal cost level is type-contingent and is equal to $\beta - e^*$.

The regulator (the principal) observes cost but not the effort level or the value of the parameter β. He has a prior about β characterized by $\nu_1 = pr(\beta = \underline{\beta})$. This probability is common knowledge.

Let S be, each period, the social utility of the project, which can be viewed for simplicity as a public good, i.e., as not sold on the market. The gross payment made by the regulator to the firm is $s + c$. We assume that there is a distortionary cost λ incurred to raise each unit of money (through excise taxes for example).

Consumers' welfare in period t is

$$S - (1+\lambda)(s_t + c_t).$$

Under complete information, a utilitarian regulator would solve in each period t

$$\max_{(e_t, s_t)} \{S - (1+\lambda)(s_t + \beta - e_t) + s_t - \psi(e_t)\} \quad \text{subject to} \quad s_t - \psi(e_t) \geqq 0.$$

The individual rationality constraint, $s_t - \psi(e_t) \geqq 0$, says that the utility level of the firm's manager must be positive to obtain his participation (the complete information problem being stationary, the allocation is the same at each period).

We assume that S is large enough so that the project is always desirable.

The optimal regulation rule is then

$$e_t = e^* \text{ and } s_t = \psi(e^*) \qquad t = 1,2.$$

Welfare is

$$(1+\delta)(S - (1+\lambda)(\psi(e^*) + \beta - e^*)).$$

Because the specific form of the principal's objective function is not crucial for our results, we from now on use the general terminology "principal" and "agent" instead of "regulator" and "firm."

We now derive the optimal static incentive scheme under incomplete information. As is well-known, (Roberts (1983), Baron and Besanko (1984)), the optimal two-period incentive scheme under full commitment is the twofold repetition of this optimal static scheme.

From the revelation principle, any incentive scheme is equivalent to a revelation mechanism in which the agent truthfully announces his type and the principal imposes associated values for s and c. The mechanism can therefore be summarized by four numbers $(\underline{s}, \underline{c})$ (when the agent announces $\underline{\beta}$) and $(\bar{s}, \bar{c})$ (when the agent announces $\bar{\beta}$). The principal faces four constraints: two individual rationality (IR) constraints, guaranteeing that the two types get a non-negative utility in the relationship, and two incentive compatibility (IC) constraints, guaranteeing that the agent does not want to conceal his type. As is usual, only two of these constraints are binding: the bad type's IR constraint and the good type's IC constraint (that the other two constraints are indeed satisfied when they are ignored in the principal's optimization program can be verified ex-post). We thus impose:

$$\bar{U} = \bar{s} - \psi(\bar{\beta} - \bar{c}) \geq 0, \tag{2-1}$$

and

$$\underline{U} = \underline{s} - \psi(\underline{\beta} - \underline{c}) \geq \bar{s} - \psi(\underline{\beta} - \bar{c}), \tag{2-2}$$

where $\underline{U}$ and $\bar{U}$ denote the good and the bad type's utilities or rents. In the optimal contract, (2-1) and (2-2) are satisfied with equality:

$$\bar{U} = 0 \tag{2-3}$$

and

$$\underline{U} = \bar{U} + \Phi(\bar{c}) = \Phi(\bar{c}) \tag{2-4}$$

where $\Phi(c)$ denotes the good type's rent and is determined by the bad type's cost level:

$$\Phi(c) \equiv \psi(\bar{\beta} - c) - \psi(\underline{\beta} - c). \tag{2-5}$$

Under our assumptions, Φ is a decreasing and convex function of c. That is, the good type's rent decreases with the bad type's cost, but at a decreasing rate. The principal maximizes his expected welfare. Replacing s by $[U + \psi(e)]$ yields:

$$\text{Min } \underset{\beta}{E}[(1+\lambda)(\psi(\beta - c) + c) + \lambda U]. \tag{2-6}$$

Note that welfare is expressed in terms of *efficiency* $E[(1+\lambda)(\psi(\beta - c) + c)]$ and *rent* $E(\lambda U)$. (The reasonings in this paper aimed at improving on a given contract will either increase efficiency keeping rent constant or possibly increase both efficiency and rent.) That is, the total cost for a given type is $[\psi(e) + c]$, which has shadow cost $(1 + \lambda)$, to which must be added the shadow cost of the agent's rent. We thus solve:

$$\underset{(\underline{c},\bar{c})}{Min}\{\nu_1[(1+\lambda)(\psi(\underline{\beta} - \underline{c}) + \underline{c}) + \lambda\Phi(\bar{c})] + (1-\nu_1)(1+\lambda)(\psi(\bar{\beta} - \bar{c}) + \bar{c})\} \tag{I}$$

The good type's cost is socially optimal:

$$\underline{c} = \underline{\beta} - e^* \tag{2-7}$$

However, the bad type's cost is inflated so as to reduce the good type's rent:

$$\psi'(\bar{\beta} - \bar{c}) = 1 + \frac{\lambda\nu_1}{(1+\lambda)(1-\nu_1)}\Phi'(\bar{c}). \tag{2-8}$$

We let $\bar{c}(\nu_1)$ denote the unique solution to equation (2-8). It is easily verified that $\bar{c}(\nu_1)$ exceeds the socially optimal cost $\bar{\beta} - e^*$ (unless $\nu_1 = 0$), and that is increases with ν_1.

Proposition 1. The optimal (static or dynamic) commitment solution is characterized by:

$$\underline{c}(\nu_1) = \underline{\beta} - e^*,$$

$$\bar{c}(\nu_1) > \bar{\beta} - e^*,$$

$$\frac{d\bar{c}}{d\nu_1} > 0, \quad \text{and}$$

$$\underline{U}(\nu_1) = \Phi(\bar{c}(\nu_1)).$$

We implicitly assumed in the previous analysis that the probability of the bad type is not too small; for, above some cut-off level of ν_1, the principal would choose not to let the bad type produce at all. We will henceforth assume that $1-\nu_1$ is sufficiently high so that the principal does not elect to ignore the bad type.[3]

For further reference, we also derive the optimal *pooling* allocation. To this purpose, suppose that the principal is constrained to pick a single cost target c for both types (in the commitment case, the principal would never elect to do so; see Proposition 1; but this thought experiment will be useful later, as the solution under renegotiation may involve pooling in the first period). The principal pays a transfer equal to $\psi(\bar{\beta}-c)$ so as to satisfy (2-1). The total cost is thus $[\psi(\bar{\beta}-c)+c]$, regardless of the agent's type. The good type's rent is $\Phi(c)$. Hence, the principal chooses c so as to solve

$$\underset{(c)}{\mathrm{Min}}\{(1+\lambda)\,E\ (\psi(\beta-c)+c)+\lambda\nu_1\Phi(c)=(1+\lambda)[\nu_1(\psi(\underline{\beta}-c)+c)+(1-\nu_1)(\psi(\bar{\beta}-c)+c)]$$

$$+\lambda\nu_1\Phi(c)\}.$$

The solution of this strictly convex program, $c^p(\nu_1)$, lies between the two types' socially optimal costs:

$$\underline{\beta}-e^*<c^p(\nu_1)<\bar{\beta}-e^*, \tag{2-9}$$

and decreases with the probability of the good type:

$$\frac{dc^p}{d\nu_1}<0. \tag{2-10}$$

b) The Renegotiation Game

We now assume that the parties can sign a long-term contract at date 1, but that the principal can at date 2 offer to renegotiate the initial contract. The principal puts prior beliefs ν_1 on the agent's having the good type. The two parties are initially bound by the "nul contract," which specifies no production and no transfer in either period. At the beginning of period 1, the principal offers a long-term contract $\{s_1(c_1), s_2^0(c_1,c_2)\}$. This contract is called a short-term contract if s_2^0 is the (second-period) nul contract. After observing the agent's performance c_1, the principal updates his beliefs to ν_2 and offers a new second-period contract, that the agent accepts or refuses. At any stage, the parties abide by the contract in force if the agent rejects the new contract offer. The old contract is superseded by the new one if the agent accepts the latter. Last, one can restrict the contract offered in period 1 to be renegotiation-proof in period 2, since parties have rational expectations.

3. Renegotiation Proof Second-Period Contracts

Suppose that at the beginning of date 2, beliefs are $\nu_2 = \nu$, and that the parties are bound by an initial contract that yields second-period rents $\underline{U}^0$ and $\bar{U}^0$ to the good and bad types (these utilities do not include the foregone first-period transfer and disutility of effort). In this section we ignore the second period subscript 2. Without loss of generality, let us assume that $\bar{U}^0 = 0$, by adjusting if necessary the first period transfers.

The principal offers a new contract, yielding second period costs $\{\underline{c}, \bar{c}\}$ and rents $\{\underline{U}, \bar{U}\}$ for the two types, so as to solve:

$$\text{Min}\{\nu[(1+\lambda)(\psi(\underline{\beta}-\underline{c})+\underline{c}+\lambda\underline{U}]+(1-\nu)[(1+\lambda)(\psi(\bar{\beta}-\bar{c})+\bar{c})+\lambda\bar{U}]\} \tag{II}$$

S.T.

$$\underline{U} \geq \bar{U} + \Phi(\bar{c}) \tag{3.1}$$

$$\bar{U} \geq 0 \tag{3.2}$$

$$\underline{U} \geq \underline{U}^0 \tag{3.3}$$

The levels of rent committed to, $\underline{U}^0 = 0$ and $\bar{U}^0$, are renegotiation proof if the solution to (II) involves $\underline{U} = \underline{U}^0$ and $\bar{U} = \bar{U}^0$. One can always choose to realize these levels of rent by the allocations which are the solutions to program (II).

Note that program (II) includes only the good type's IC constraint (3.1). As is usual, the ignored IC constraint for the bad type is checked ex-post. The only difference between programs (I) and (II) is the presence of the extra IR constraint (3.3). That is, the good type may have been promised a higher second period rent than program (I) (see Proposition 1) would award him.

The solution to (II) clearly involves $\bar{U} = 0$ (no new rent for the bad type) and $\underline{c} = \underline{\beta} - e^*$ (the good type's cost is socially optimal). Let us simplify consequently the optimization program to:

$$\text{Min}\{\nu[(1+\lambda)(\psi(e^*)+\underline{\beta}-e^*)+\lambda\underline{U}]+(1-\nu)[(1+\lambda)(\psi(\bar{\beta}-\bar{c})+\bar{c})]\} \tag{III}$$

S.T.

$$\underline{U} \geq \Phi(\bar{c}) \qquad (\lambda_1) \tag{3.4}$$

$$\underline{U} \geq \underline{U}^0 \qquad (\lambda_2) \tag{3.5}$$

Three cases can be distinguished according to which constraints are binding in program (III). The Lagrangian of this convex program reduces to:

$$L = \nu\lambda\underline{U} + (1-\nu)(1+\lambda)(\psi(\bar{\beta} - \bar{c}) + \bar{c}) - \lambda_1(\underline{U} - \Phi(\bar{c})) - \lambda_2(\underline{U} - \underline{U}^0) \tag{3.6}$$

with first order conditions:

$$\psi'(\bar{\beta} - \bar{c}) = 1 + \frac{\lambda_1}{(1-\nu)(1+\lambda)}\Phi'(\bar{c}) \tag{3.7}$$

$$\lambda_1 + \lambda_2 = \nu\lambda \tag{3.8}$$

Case 1 occurs when $\underline{U}^0$ is small so that (3.5) is not binding ($\lambda_2 = 0$). From (3.7) (3.8)

$$\psi'(\bar{\beta} - \bar{c}) = 1 + \frac{\nu}{1-\nu} \cdot \frac{\lambda}{1+\lambda}\Phi'(\bar{c}) \tag{3.9}$$

We obtain the same result as in Proposition 1. The solutions to (I) and (III) coincide except that ν_1 is replaced by $\nu = \nu_2$. The allocation is optimal for the principal conditionally on his posterior beliefs. The contract is called *conditionally optimal*. From Proposition 1 $\underline{U} = \psi(\bar{\beta} - \bar{c}) - \psi(\underline{\beta} - \bar{c}) = \Phi(\bar{c}(\nu))$. This case is therefore valid for $\underline{U}^0 \leq \Phi(\bar{c}(\nu))$. Actually, for the first period contract to be renegotiation proof we must have $\underline{U}^0 = \Phi(\bar{c}(\nu))$.

Case 2 occurs when $\underline{U}^0$ is increased beyond $\Phi(\bar{c}(\nu))$. Then, both constraints are binding. Then $\underline{U} = \underline{U}^0$ and $\bar{c}$ is defined by $\underline{U} = \Phi(\bar{c})$. This case ceases to be valid when $\underline{U}^0$ is so large that the incentive constraint (3.4) ceases to be binding. Then $\lambda_1 = 0$; $\psi'(\bar{\beta} - \bar{c}) = 1$ or $\bar{c} = \bar{\beta} - e^*$, the socially optimal level. Case 2 occurs for $\underline{U}^0$ between $\Phi(\bar{c}(\nu))$ and $\Phi(\bar{\beta} - e^*)$ giving a cost $\bar{c}$ between $\bar{\beta} - e^*$ and $\bar{c}(\nu)$. A contract specifying $\{\underline{c} = \underline{\beta} - e^*;\ \bar{\beta} - e^* \leq \bar{c} \leq \bar{c}(\nu);\ \underline{U} = \Phi(\bar{c})\}$ is called *rent-constrained* and is renegotiation proof for $\underline{U}^0 = \underline{U}$. The principal would wish to lower the good type's rent, but cannot do so because of the existence of the initial contract. This loss in rent is partially compensated by the fact that the cost of the bad type can be brought closer to the efficient level while still satisfying the good type's incentive constraint. Clearly, *in the second period*, the principal would prefer a lower value of $\underline{U}^0$ yielding a conditionally optimal contract. However, this does not mean that the principal is better off committing to the conditionally optimal contract, because the value of $\underline{U}^0$ affects the first period's incentive constraint.

Finally, when $\underline{U}^0$ lies between $\Phi(\bar{\beta} - e^*)$ and $\Phi(\underline{\beta} - e^*)$,[4] (3.3) is binding and (3.1) is not. As observed above, the solution is such that $\bar{c} = \bar{\beta} - e^*$, so that the two cost levels are socially optimal. The cost allocation is identical to that under a sell-out or fixed-price contract, in which the agent is the residual claimant for his cost savings. It is renegotiation proof if it corresponds to the rent $\underline{U}^0$ for the good type. Note that all sell-out contracts have the same *efficiency* $\underset{\beta}{E}(1 + \lambda)(\psi(e^*) + \beta - e^*)$. They differ only by the good type's rent. Therefore, from a second period view point, the principal prefers the one with the lowest rent.

For further reference we gather our analysis in a Proposition and four corollaries.

Proposition 2: Normalizing $\bar{U} = 0$, renegotiation-proof contracts can be indexed by a single parameter, the good type's rent $\underline{U}$.

1. For $\underline{U} = \Phi(\bar{c}(\nu))$, it is the conditionally optimal contract:

$$\underline{c} = \underline{\beta} - e^*; \quad \bar{c} = \bar{c}(\nu)$$

2. For $\Phi(\bar{c}(\nu)) < \underline{U} < \Phi(\bar{\beta} - e^*)$, it is a rent constrained contract:

$$\underline{c} = \underline{\beta} - e^*; \quad \bar{\beta} - e^* \leq \bar{c} = \Phi^{-1}(\underline{U}) \leq \bar{c}(\nu)$$

3. For $\Phi(\bar{\beta} - e^*) \leq \underline{U} \leq \Phi(\underline{\beta} - e^*)$, it is a sell-out contract:

$$\underline{c} = \underline{\beta} - e^*; \quad \bar{c} = \bar{\beta} - e^*.$$

Corollary 1: In a renegotiation-proof contract, the good type's rent exceeds that in a conditionally optimal contract: $\underline{U} \geq \underline{U}(\nu) \equiv \Phi(\bar{c}(\nu))$.

Corollary 2: The principal's welfare is strictly decreasing with the (good type's) rent $\underline{U}$ which indexes the set of renegotiation-proof contracts. The efficiency of the allocation increases with the rent on $[\Phi(\bar{c}(\nu)), \Phi(\bar{\beta} - e^*)]$ (rent constrained contracts) and does not depend on the rent on $[\Phi(\bar{\beta} - e^*), \Phi(\underline{\beta} - e^*)]$ (sell-out contracts).

Corollary 3: Consider a rent-constrained contract indexed by $\bar{U}$ which is renegotiation proof for beliefs ν. Then, it remains renegotiation proof for beliefs $\nu' > \nu$.

Proof: From Proposition 2, $\Phi^{-1}(\underline{U}) \leq \bar{c}(\nu)$. From Proposition 1, $\frac{d\bar{c}}{d\nu} \geq 0$. Therefore, we still have $\bar{\beta} - e^* \leq \Phi^{-1}(\underline{U}) \leq \bar{c}(\nu')$.

Corollary 4: For any renegotiation-proof contract that is not conditionally optimal, there exists an arbitrary close renegotiation-proof contract with a (slightly) lower rent for the good type, and a (slightly) higher welfare for the principal in period 2.

4. Characterization of the Optimal Contract.

In this section, we partially characterize the optimal contract. Section 5 completes the characterization.

Theorem 1: The principal offers the agent a choice between two contracts in the first period. The first is picked by the good type only and yields the efficient cost in both periods. In the second contract, both types produce at the same cost level in the first period; the second-period allocation is the conditionally optimal one given posterior beliefs ν_2 in $[0,\nu_1]$.

To prove Theorem 1, we first show that, in a sense, the relevant IC constraint in the first period is the good type's. That is, only randomization by the good type can be an equilibrium behavior:

Lemma 1: The principal offers the agent a choice between two contracts, one chosen by the good type, and the other chosen by the bad type and possibly by the good type.

Proof: See Appendix 1.

The ability to commit, despite the renegotiation proofness condition, enables the principal to neglect the bad type's incentive constraint. The consequence of this major lemma is that incentive constraints are binding only for the good type as in the full commitment case and contrary to the no commitment case.

Lemma 1 implies that the overall optimal contract can be described as in Figure 1.

The top branch in Figure 1 represents the first contract and is chosen by the good type with probability x. The low and middle branches in Figure 2 represent the second contract. The middle branch is chosen by the good type with probability $1-x$. The low branch is always chosen by the bad type.

Let $\underline{U}_2$ denote the rent that the good type is promised if he chooses a_1 in period 1. From Corollary 2, $\underline{U}_2 \geq \Phi(a_2)$.

The second-period cost following cost b_1 is the socially efficient one $b_2 = \underline{\beta} - e^*$, as it has become common knowledge from the observation of the first-period cost b_1 that the agent has type $\underline{\beta}$.

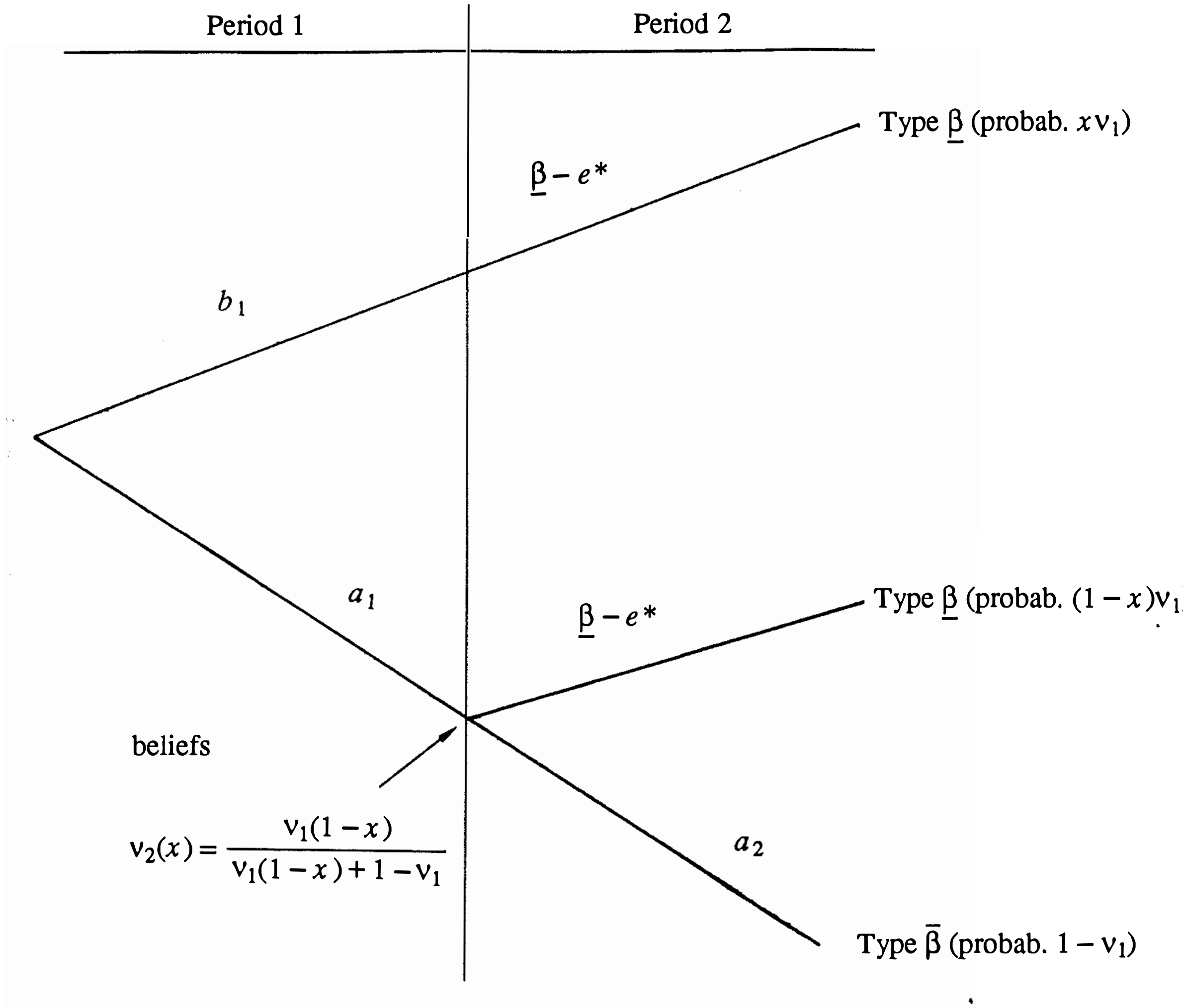
The good type must be given a rent in the first contract that is sufficient not to induce him to choose the second contract. The best way do this is to ask him to produce efficiently, $b_1 = \underline{\beta} - e^*$, and to promise him a total rent:

$$\underline{U} = \Phi(a_1) + \delta \underline{U}_2 \tag{4-1}$$

It remains to determine the optimal pooling cost a_1, the bad type second period cost a_2 in the second contract and the optimal x. The determination of the probability x that the good type reveals his type (separates) is tackled in section 5. For a given x, the principal's welfare is obtained by solving:

$$\underset{\{a_2,a_2\}}{\text{Min}} \{(1+\lambda)[\nu_1 x(\psi(e^*) + \underline{\beta} - e^*) + \nu_1(1-x)(\psi(\underline{\beta} - a_1) + a_1) \tag{4-2}$$

$$+ (1-\nu_1)[\psi(\overline{\beta} - a_1) + a_1)] + \lambda\nu_1\Phi(a_1)$$



Period 1
Period 2
Type $\underline{\beta}$ (probab. $x\nu_1$)
$\underline{\beta} - e^*$
b_1
a_1
$\underline{\beta} - e^*$
Type $\underline{\beta}$ (probab. $(1-x)\nu_1$)
beliefs
$\nu_2(x) = \frac{\nu_1(1-x)}{\nu_1(1-x)+1-\nu_1}$
a_2
Type $\bar{\beta}$ (probab. $1-\nu_1$)

Figure 1

$$+\ \delta[(1+\lambda)\nu_2(\psi(e^*)+\underline{\beta}-e^*)+(1-\nu_2)(\psi(\overline{\beta}-a_2)+a_2)\ +\ \lambda\nu_1\underline{U}_2]\}$$

S.T.

$$\underline{U}_2 \geq \Phi(a_2) \quad (4\text{-}3)$$

$$\underline{\beta}-e^* \leq a_2 \leq \overline{c}(\nu_2) \quad (4\text{-}4)$$

(4-3) is clearly binding.

Lemma 2: $\underline{U}_2 = \Phi(a_2)$.

Proof: The second period contract must be a rent-constrained contract (including the two extremes in this class). For, if the second period contract were a sell-out one with rent exceeding $\Phi(\overline{\beta}-e^*)$, the principal could specify a slightly lower rent for the good type while keeping efficiency constant and thus increase welfare (see Corollary 2).

Q.E.D.

The intuition for this important result (the fact that second period contracts are conditionally optimal) is that any increase of the rent beyond $\Phi(a_2)$) serves no purpose in period 2 and requires a further increase of the rent of the good type when he reveals his type because the incentive constraint of the good type is binding.

The optimization program (4-2) can be broken down in two separate optimizations, minimization of first period cost with respect to a_1, and minimization of second period cost with respect to a_2 subject to (4-3) and (4-4).

To complete the proof of theorem 1 we consider the second minimization which is rewritten:

$$\text{Min}\{(1-\nu_2)(\psi(\overline{\beta}-a_2)+a_2)+\lambda\nu_1\Phi(a_2)\} \quad (4\text{-}5)$$

S.T.

$$\overline{\beta}-e^* \leq a_2 \leq \overline{c}(\nu_2) \quad (4\text{-}6)$$

Lemma 3: The optimal a_2 equals $\overline{c}(\nu_2)$.

Proof: Consider first the unconstrained minimization. From the Bayesian revision of expectations $\nu_2 \leq \nu_1$. The problem is therefore analogous to a one period static problem except that the rent is more costly because $\nu_1 \geq \nu_2$. So the optimal solution of the unconstrained problem is larger that $\overline{c}(\nu_2)$ (from the first order equation). As the objective function (4-5) is strictly convex in a_2, the optimal solution of the constrained problem is $\bar{c}(\nu_2)$.

Q.E.D.

From Lemma 3, we know that the second period contract is conditionally optimal given ν_2. Minimization of (4-2) with respect to a_1 yields:

$$\frac{(1-\nu_1)\psi'(\bar{\beta}-a_1)+(1-x)\nu_1\psi'(\underline{\beta}-a_1)}{(1-\nu_1)+(1-x)\nu_1} \tag{4-7}$$

$$= 1 - \frac{\lambda}{1+\lambda}\cdot\frac{\nu_1}{1-\nu_1 x}\,[\psi'(\bar{\beta}-a_1)-\psi'(\underline{\beta}-a_1)]$$

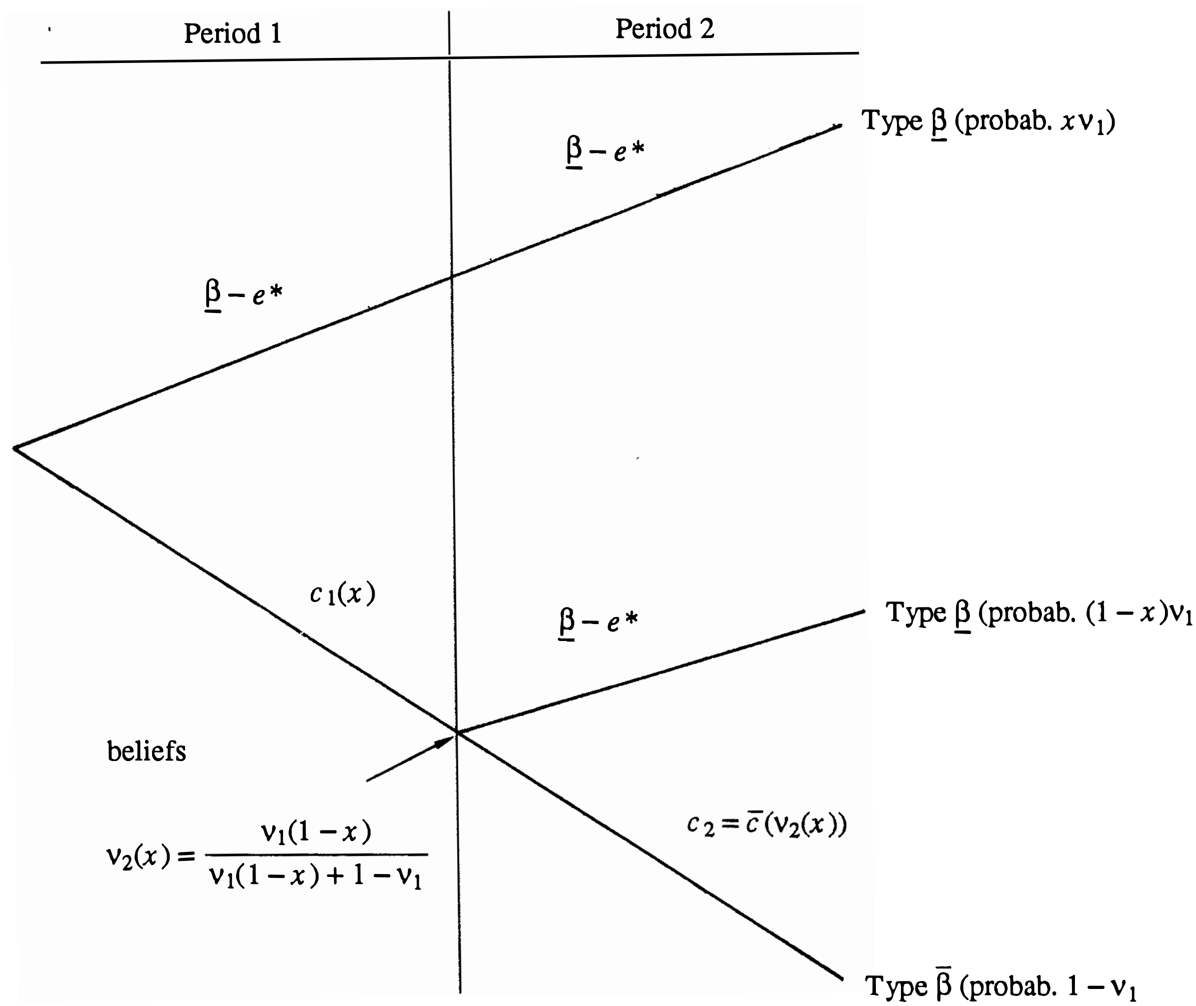
Indeed, the optimization problem is here identical to that determining the optimal pooling contract (see (2.8)), but for the fact that only a fraction $(1-x)$ of the good types produce at cost a_1. The two programs coincide when $x=0$. When $x=1$, (4-2) coincides with the commitment (separating) program. Letting $a_1 = c_1(x)$ denote the solution of (4-7), we obtain immediately:

Theorem 2: The first-period cost in the pooling branch $c_1(x)$ is independent of the discount factor (for a given x), and is an increasing function of the probability x that the good type separates in the first period. In a pooling equilibrium $c_1(0) = c^p(\nu_1)$, and in a separating equilibrium $c_1(1) = \bar{c}(\nu_1)$.

We also observe that the rent given to the good type in the case of renegotiation proof commitment is strictly higher than in the case of commitment (i.e., $\Phi(c_1(x)) + \delta\Phi(\bar{c}(\nu_2)) \geq \Phi(\bar{c}(\nu_1)) + \delta\Phi(\bar{c}(\nu_1))$. This results from the fact that $\Phi(c_1(x)) \geq \Phi(\bar{c}(\nu_1))$ (since $c_1(x) \leq \bar{c}(\nu_1)$) and $\Phi(\bar{c}(\nu_2)) > \Phi(\bar{c}(\nu_1)$ (since $\nu_2 < \nu_1 \Rightarrow \bar{c}(\nu_2) < \bar{c}(\nu_1)$). (The last inequality is strict because of Theorem 3 below).

Theorems 1 and 2 reduce the computation of the optimal contract to the choice of a single number x in $[0,1]$ and are summarized in Figure 2.

Remark: The principal's behavior is equivalent to the offering of a choice between a long-term and a short-term contract. The acceptance of the short-term contract (to produce at the cost target $c_1(x)$ in the first period) is followed by the second-period conditionally optimal contract. As we will see in Section 6, the main difference with the non-commitment case is the possibility for the principal to sign a long-term contract with the good type to which the bad type would be committed if he were to sign it.



Period 1
Period 2
Type $\underline{\beta}$ (probab. $x\nu_1$)
$\underline{\beta} - e*$
$\underline{\beta} - e*$
$c_1(x)$
$\underline{\beta} - e*$
Type $\underline{\beta}$ (probab. $(1-x)\nu_1$
beliefs
$\nu_2(x) = \frac{\nu_1(1-x)}{\nu_1(1-x)+1-\nu_1}$
$c_2 = \bar{c}(\nu_2(x))$
Type $\bar{\beta}$ (probab. $1-\nu_1$

Figure 2

5. How Much Pooling?

This section completes the derivation of the optimal contract by determining the probability x that the good type separates in the first period as a function of the discount factor. The principal's optimization program over x may not be concave, as we have little information about the curvature of the functions $c_1(x)$ and $c_2(x) \equiv \bar{c}(\nu_2(x))$. If the solution is not unique, the following properties hold for any optimizing value. For notational simplicity, we will write $x(\delta)$ as if it were unique. So, for instance, "$x(\delta)$ is non-increasing with δ" means "if x an optimum for δ and $\tilde{x}$ is an optimum for $\tilde{\delta} > \delta$, then $x \geq \tilde{x}$."

Theorem 3:

(i) The good type's probability of separation x is non-increasing with the discount factor δ.

(ii) There exists $\delta_0 > 0$ such that for all $\delta \leq \delta_0$, the optimal contract is a separating one ($x = 1$).

(iii) When δ becomes large, the optimal contract tends towards a pooling contract ($x \to 0$). However, a pooling contract is never optimal ($x > 0$ for all δ).

Thus, when the discount factor increases, the optimal allocation moves from full revelation to full pooling. While full separation is optimal for small discount factors, full pooling is optimal only in the limit of large discount factors. Note that large discount factors (above 1) need not be absurd, as the discount factor reflects the relative lengths of the accounting periods (or the relative importance of the first- and second-period projects).

Proof of Theorem 3:

(i) Let $W(x,\delta,c_1,c_2)$ denote the principal's welfare, where c_1 denotes the first-period cost in the pooling branch, and c_2 the second-period cost of the bad type. At the optimum, c_1 and c_2 are functions of x, but not of δ: Theorem 1 implies that $c_2 = c_2(x) \equiv \bar{c}(\nu_2(x))$, and Theorem 2 yields $c_1 = c_1(x)$. One has:

$$W(x,\delta,c_1,c_2) = G(x,c_1) + \delta H(x,c_2), \tag{5-1}$$

where

$$G(x,c_1) \equiv S - (1+\lambda)[\nu_1 x(\psi(e^*) + \underline{\beta} - e^*) + \nu_1(1-x)(\psi(\underline{\beta} - c_1) + c_1) \tag{5-2}$$

$$+ (1-\nu_1)(\psi(\bar{\beta} - c_1) + c_1)]$$

$$- \lambda\nu_1\Phi(c_1)$$

is the "first-period welfare," and

$$H(x,c_2) \equiv S - (1+\lambda)[\nu_1(\psi(e^*) + \underline{\beta} - e^*) + (1-\nu_1)(\psi(\bar{\beta} - c_2) + c_2)] \tag{5-3}$$

$$- \lambda\nu_1\Phi(c_2)$$

is the "second-period welfare."

Consider two discount factors $\delta < \tilde{\delta}$ and let $\{x, c_1 = c_1(x), c_2 = c_2(x)\}$ and $\{\tilde{x}, \tilde{c}_1 = c_1(\tilde{x}), \tilde{c}_2 = c_2(\tilde{x})\}$ denote associated optimal allocations. Because renegotiation proofness depends only on the separating probability and the second-period cost, and not on the discount factor, the principal could have chosen the allocation $\{\tilde{x}, \tilde{c}_1, \tilde{c}_2\}$ when facing discount factor δ. Hence:

$$W(x, \delta, c_1, c_2) \geq W(\tilde{x}, \delta, \tilde{c}_1, \tilde{c}_2). \tag{5-4}$$

Similarly,

$$W(\tilde{x}, \tilde{\delta}, \tilde{c}_1, \tilde{c}_2) \geq W(x, \tilde{\delta}, c_1, c_2). \tag{5-5}$$

Adding (5-4) and (5-5), and using (5-2) and (5-3) yields:

$$(\tilde{\delta} - \delta)\{[(1+\lambda)(1-\nu_1)(\psi(\bar{\beta} - c_2) + c_2) + \lambda\nu_1\Phi(c_2)] \tag{5-6}$$

$$- [(1+\lambda)(1-\nu_1)(\psi(\bar{\beta} - \tilde{c}_2) + \tilde{c}_2) + \lambda\nu_1\Phi(\tilde{c}_2)]\} \geq 0.$$

Recall that the function $[(1+\lambda)(1-\nu_1)(\psi(\bar{\beta} - c) + c) + \lambda\nu_1\Phi(c)]$, which is nothing but the objective function under commitment, is convex in c and takes its minimum value at $c = \bar{c}(\nu_1)$ by definition of $\bar{c}(\nu_1)$. Recall further that $c_2 = \bar{c}(\nu_2(x))$ and $\tilde{c}_2 = \bar{c}(\nu_2(\tilde{x}))$, where $\nu_2(x)$ and $\nu_2(\tilde{x})$ are lower than ν_1, implying that c_2 and $\tilde{c}_2$ are lower than $\bar{c}(\nu_1)$ (Proposition 1). Equation (5-6), together with $\tilde{\delta} > \delta$, implies that $c_2 \leq \tilde{c}_2$, which (again from Proposition 1) implies that $\nu_2(x) \leq \nu_2(\tilde{x})$ or $x \geq \tilde{x}$.

(ii) Let us first show that when δ tends to 0, $x(\delta)$ tends towards 1. If it does not, there exists a subsequence of discount factors tending to 0 (and associated values of $x(\delta)$) such that $1 - x(\delta) \geq \alpha > 0$. Along this subsequence, the good type produces, with probability α at least, at a first-period cost exceeding $c^p(\nu_1)$ (from Theorem 2) and thus bounded away from $\underline{\beta} - e^*$. Thus the welfare loss relative to the commitment solution does not converge to 0. But choosing $\{x = 1, c_1 = \bar{c}(\nu_1), c_2 = \bar{c}(0) = \bar{\beta} - e^*\}$ yields a welfare $W(x, \delta, c_1, c_2)$ that converges to the welfare under commitment when δ tends to 0 (see (5-1)), a contradiction.

Second, at $\delta = 0$, the optimum is the static optimum and thus involves full separation $(x = 1)$. Furthermore,

$$\frac{d}{dx}(W(x, \delta, c_1(x), c_2(x)))\bigg|_{\{\delta = 0, x = 1\}} = \frac{d}{dx}(G(x, c_1(x))) \tag{5-7}$$

$$= (1+\lambda)[(\psi(\underline{\beta} - \bar{c}(\nu_1)) + \bar{c}(\nu_1))$$

$$- (\psi(e^*) + \underline{\beta} - e^*)]$$

$$> 0,$$

where use is made of the envelope theorem. Hence $W(1,\delta,c_1(1),c_2(1)) > W(x,\delta,c_1(x),c_2(x))$ for all x close to (but lower than) 1 and all δ close to 0.

The intuition behind this proof is that if $\varepsilon (= 1 - x)$ is the probability of pooling, the first-period loss in welfare due to pooling is proportional to ε, while the second-period gain due to a reduction in the good type's rent is proportional to $\delta\varepsilon$.

(iii) When δ tends to $+\infty$, the welfare under pooling $W(0,\delta,c_1(0),c_2(0))$ tends to the welfare under commitment. So must the optimal welfare. From (5-3), $c_2(x(\delta))$ must converge to $c_2(0) = \bar{c}(\nu_1)$, which implies that $\nu_2(x(\delta))$ converges to ν_1 or $x(\delta)$ converges to 0 (for δ large, G becomes negligible relative to δH).

Next, fix δ. Let us show that $x = 0$ cannot be optimal:

$$\frac{d}{dx}(W(x,\delta,c_1(x),c_2(x)))\bigg|_{x=0} = \frac{\partial}{\partial x}(W(x,\delta,c_1(x),c_2(x))\bigg|_{x=0} \tag{5-8}$$

using the envelope theorem: $\partial W/\partial c_1 = 0$ for all x; and $\partial W/\partial c_2 = 0$ for $x = 0$, as the second-period cost $c_2(0)$ is the commitment one $\bar{c}(\nu_1)$ (note that for $x > 0$, $\partial W/\partial c_2 > 0$: the principal is constrained by renegotiation proofness in his choice of c_2). Hence,

$$\frac{d}{dx}(W(x,\delta,c_1(x),c_2(x)))\bigg|_{x=0} = \frac{\partial G}{\partial x}\bigg|_{x=0} > 0. \tag{5-9}$$

Thus full pooling cannot be optimal.

The intuition here is that at the full pooling allocation, small changes in c_2 have only second-order effects because the second-period allocation is the commitment one. A small decrease in c_2 allows x to become positive without violating renegotiation proofness, and the first-period allocation is improved to the first order in x.

Q.E.D.

Last, it is instructive to consider the case of small uncertainty ($\Delta\beta = \bar{\beta} - \underline{\beta}$ small). Under non-commitment (see our 1987 paper), the welfare distortion relative to commitment is of the first-order in $\Delta\beta$ (i.e., proportional to $\Delta\beta$) for the best pooling contract. In contrast, it remains finite (i.e., does not converge to 0 with $\Delta\beta$) for the best separating contract (so that full pooling always dominates full separation for $\Delta\beta$ small).

Under commitment and renegotiation, the welfare loss relative to commitment under both the best full pooling and the best full separating contracts (as well as contracts corresponding to intermediate x's) turns out to be of the second order in $\Delta\beta$. To see this, note first that for $x = 1$ (separating contract), the allocation differs from the commitment one only with respect to the bad type's second-period cost, which is equal to $\bar{\beta} - e^*$ instead of $\bar{c}(\nu_1)$. So the welfare loss under the best separating equilibrium is

$$L^s \equiv W^c - W(1,\delta,\bar{c}(\nu_1),\bar{\beta} - e^*) = \delta\{[(1+\lambda)(1-\nu_1)(\psi(\bar{\beta} - \bar{c}(\nu_1)) + \bar{c}(\nu_1))$$

$$+ \lambda\nu_1\Phi(\bar{c}(\nu_1))]$$

$$- ((1+\lambda)(1-\nu_1)(\psi(e^*) + \bar{\beta} - e^*)$$

$$+ \lambda\nu_1\Phi(\bar{\beta} - e^*)]\}.$$

But, from (2-8), the difference between $\bar{c}(\nu_1)$ and $(\bar{\beta} - e^*)$ is proportional to $\Delta\beta$ for $\Delta\beta$ small. Furthermore, $\bar{c}(\nu_1)$ minimizes the commitment cost, so that small variations around $\bar{c}(\nu_1)$ have only second-order effects. Hence, L^S is proportional to $(\Delta\beta)^2$.

The proof that $L^P \equiv W^c - W(0,\delta,c^P(\nu_1),\bar{c}(\nu_1))$ is proportional to $(\Delta\beta)^2$ as well is similar. It suffices to note that the best pooling contract differs from the commitment allocation only with respect to the first period cost, which is equal to $c^P(\nu_1)$ instead of $\underline{\beta} - e^*$ for type $\underline{\beta}$ and $\bar{c}(\nu_1)$ for type $\bar{\beta}$.[5] We conclude that the best pooling contract and the best separating contract involve little loss for $\Delta\beta$ small contrary to the non-commitment one.[6]

6. Continuum of Types

We now assume that the agent's type β belongs to an interval $[\underline{\beta},\bar{\beta}]$, and is distributed according to the cumulative distribution function $F(\cdot)$ (such that $F(\underline{\beta}) = 0$, $F(\bar{\beta}) = 1$) with continuous density $f(\cdot)$. We make the classic monotone hazard rate assumption: $F(\beta)/f(\beta)$ is a non-decreasing function of β.

Our 1988 paper studies this continuum model under the non-commitment assumption (the relationship is run by two consecutive short-term contracts). A main result there is that separation is not feasible, let alone desirable. That is, there exists no separating first-period incentive scheme $s_1(c_1)$ (even a suboptimal one); for any $s_1(\cdot)$, the equilibrium function $c_1(\beta)$ does not fully reveal the agent's type. We investigate whether separation is feasible and desirable under renegotiable commitment. The answer if found in:

Theorem 4

(i) There exist separating (first-period) incentive schemes. The optimal contract in the class of separating schemes yields the commitment allocation in period 1, and the socially efficient cost in period 2.

(ii) A separating contract is never optimal for the principal.

Proof of Theorem 4

(i) In a separating equilibrium, the agent's type is common knowledge at the beginning of period 2. The possibility of renegotiation implies that the agent's second-period effort is socially optimal: $e_2(\beta) = e^*$. Hence the agent's second-period rent $U_2(\beta)$ grows one-for-one with the agent's efficiency: $\dot{U}_2(\beta) = -1$ or $U_2(\beta) - U_2(\bar{\beta}) = \bar{\beta} - \beta$.[12] Thus, fixing $U_2(\bar{\beta}) = 0$ w.l.o.g., both the agent's effort and his rent, and therefore the principal's second-period welfare are the same in all separating contracts. We call the second-period contract the sell-out contract.

The principal, if constrained to choose a separating contract, thus maximizes his first-period welfare. But, by definition, the welfare optimal scheme is the commitment scheme. The commitment scheme is computed for a continuous distribution in Laffont-Tirole (1986). Under the monotone hazard rate assumption, the agent produces at cost $c_1(\beta) = c^*(\beta)$, where $c^*(\beta) \geq \beta - e^*$ (with strict inequality except at $\beta = \underline{\beta}$) and $c^*(\beta)$ is a strictly increasing function of β.

Conversely, suppose that the principal offers the following contract: "The agent can choose first-period cost in the interval $[c^*(\underline{\beta}), c^*(\bar{\beta})]$. If he has produced at cost c_1 in the first period, he must produce at cost $(c^{*-1}(c_1) - e^*)$ in the second, and receives intertemporal transfer $[\psi(c^{*-1}(c_1) - c_1) + \delta\psi(e^*) + (\int_{c^{*-1}(c_1)}^{\bar{\beta}} \psi'(\beta - c^*(\beta))d\beta + \delta(\bar{\beta} - c^{*-1}(c_1))]$." He thus asks for the efficient effort e^* in period 2. The first part of the transfer is the compensation for the intertemporal disutility of effort. The second part corresponds to the rent in the commitment contract, plus the second-period rent. By construction, this contract yields the commitment welfare in the first period, and the sell-out welfare in the second. The agent's local incentive compatibility constraint is satisfied by construction; checking that the global incentive compatibility constraint holds as well is routine.

(ii) The non-separation result is proved in Appendix 2. The intuition is the following. In the best separating equilibrium (characterized in (i)), the first-period allocation is the commitment one. That is, it maximizes ex-ante welfare subject to the informational constraints. This implies that any change in the first-period allocation has second-order effect. In contrast, the second-period allocation is not optimal from the point of view of the *ex-ante* informational structure. This implies that changes in the corresponding allocation have first-order effects on welfare.

From (i), we know that the only way to change the second-period allocation is to create some pooling in the first period. Our proof shows that, starting from the best separating contract, the principal can force the less efficient types to pool in the first period and increase his intertemporal welfare. More precisely, suppose that he penalizes the agent heavily if the latter's cost exceeds $c^*(\bar{\beta} - \varepsilon)$, where ε is positive and small, and that he keeps the same transfers for costs in $[c^*(\underline{\beta}), c^*(\bar{\beta} - \varepsilon)]$ as in the commitment solution. The "bad types," i.e., those in $[\bar{\beta} - \varepsilon, \bar{\beta}]$, now pool at cost $c^*(\bar{\beta} - \varepsilon)$. This increases the bad types' efficiency (because c_1 is brought closer to its efficient level for those types. Recall that $c^*(\beta) > \beta - e^*$), but it increases all types' rent (because $\dot{U}_1(\beta) = -\psi'(\beta - c_1(\beta))$ and $U_1(\bar{\beta}) = 0$). Overall, the change decreases welfare only to the third order in ε: to the second order times the length ε

over which the change operates. In contrast, in period 2, the pooling of the bad types goes in the right direction from an ex-ante point of view. Because the principal offers the conditionally optimal contract given truncated beliefs on $[\bar{\beta} - \varepsilon, \bar{\beta}]$, the cost of each bad type (but type $\bar{\beta} - \varepsilon$) is raised a bit (in a credible way), which moves the allocation if the direction of the commitment solution. The welfare gain is second order in ε: first order times the length ε over which the change operates.

Q.E.D.

Theorem 4 shows that renegotiable commitment is intermediate between full commitment (for which separation is optimal) and non-commitment (for which separation is not feasible). Here separating contracts exist, but are not optimal.

7. Commitment, Renegotiation and Non-Commitment.

In Laffont-Tirole (1987, 1988), we studied the model of this paper under the assumption that the relationship is run by a sequence of two short-run contracts (the non-commitment case). That is, the principal offers a first-period incentive scheme $s_1(c_1)$, observes c_1, and offers in period 2 the contract $s_2(c_2,c_1)$ that is conditionally optimal given posterior beliefs. We view the explorations of commitment and renegotiation and of non-commitment as complementary. The first refers to a complete contract situation and the second to a situation in which the parties cannot commit, either because of legal constraints (as may be the case for public procurement) or because the second-period contingencies are hard to foresee or costly to include in the initial contract. Alternatively, when complete contracts can be signed, the comparison between the two yields a measure of the value of commitment. Figure 3 gathers some results from the three papers and compares commitment, commitment and renegotiation, and non-commitment.

Nature of equilibrium \ Nature of Commitment	Full commitment (c)	Commitment and renegotiation (r)	No commitment (nc)
Two types			
Binding IC Constraint in first period	Good type's	Good type's	Good type's, bad type's, or both
First period revelation	Full separation	Randomization[a] by good type.	Randomization[a] by one or the two types
Equilibrium for small δ	Full separation	Full separation	Full separation
Equilibrium for large δ	Full separation	Tends to full pooling	Tends to full pooling[b]
Second-period contract conditionally optimal?	No	Yes	Yes
Good type's rent (U^c, U^r, U^{nc})	U^c	$U^r > U^c$ for δ small $\frac{U^r - U^c}{\delta} \to 0$ as $\delta \to +\infty$[c]	$U^{nc} = U^r > U^c$ for δ small $U^{nc} \gtrless U^c$ in general[d]
Principal's expected welfare (W^c, W^r, W^{nc})	W^c	$W^r < W^c$	$W^{nc} = W^r$ for δ small $W^{nc} < W^r$ otherwise[e]
Continuum of types			
Full separation feasible ?[f]	Yes	Yes[g]	No
Full separation desirable ?[f]	Yes	No	No ("much pooling")

Figure 3

Notes on Figure 3

a) The "randomization" can be degenerate, as in the case of full separation.

b) Only the weaker property that full pooling is preferred to full separation is proved in our 1987 paper. However, it is easily shown that the equilibrium allocation is essentially the one obtained under full pooling.

c) $U^r = \Phi(c_1(x)) + \delta\Phi(\overline{c}(\nu_2(x)))$ is equal to $\Phi(\overline{c}(\nu_1)) + \delta\Phi(\overline{\beta} - e^*) > (1+\delta)\Phi(\overline{c}(\nu_1)) = U^c$ for δ small. When δ tends to infinity, $U^r/\delta \simeq \Phi(\overline{c}(\nu_1)) = U^c/\delta$.

d) For δ small, the non-commitment equilibrium is separating and the rent is $U^{nc} = \Phi(\overline{c}(\nu_1)) + \delta\Phi(\overline{\beta} - e^*) = U^r$.

e) In general, $W^{nc} \leq W^r$, because under commitment and renegotiation, the principal can always offer a short-term contract in the first-period and thus duplicate the non-commitment solution. The two welfares coincide only when the bad type's IC constraint is not binding in the non-commitment case, i.e., when δ is small. See also the comments below.

f) "Full separation" means that the principal learns the agent's type at the end of the first period. "Feasibility" refers to the existence of a (not necessarily optimal) contract that separates the types. "Desirability" refers to the optimal contract.

g) The principal can fully separate the types by offering a sell-out contract from date 1 on (i.e., offering $s_1(c) = s_2(c) = (\psi(e^*) + \overline{\beta} - e^*) - c$, where $\overline{\beta}$ is the upper bound of the interval of types.

The renegotiation case technically resembles the commitment case in that the IC constraints are well-behaved: Only the good type's IC constraint is binding. In contrast, under non-commitment, the good type must receive a high first-period reward to reveal his information, because ratcheting makes such revelation costly to him. The bad type may then be tempted to "take-the-money-and-run," i.e., to mimic the good type in the first period, get the high reward and refuse to produce in the second period (this strategy is particularly tempting if δ is high, because the good type values the future much and therefore must be bribed more to reveal his type). This possibility makes the bad type's IC constraint binding if the discount factor is not too small. The take-the-money-and-run strategy can be prevented under commitment (even with renegotiation) by forcing the agent to repeat his first-period performance if the latter was excellent (i.e., equal to $\underline{\beta} - e^*$).

In both the renegotiation and non-commitment cases, the first-period contract involves pooling if the discount factor is not too small. Furthermore, the second-period contract is conditionally optimal. In a sense, the main difference between these two cases is the possibility for the principal under commitment and renegotiation to prevent the take-the-money-and-run strategy. This power allows him to give the good type more incentives to separate without having the bad type mimic the good type in the first period. Because the take-the money-and-run strategy is not optimal for the bad type for small, discount factors it is not surprising that the renegotiation and non-commitment solutions coincide for small discount factors.

An apparent lesson of our three papers and of Figure 3 is that the renegotiation case is somewhat intermediate between the commitment and non-commitment paradigms.[7]

8. Application to Intertemporal Price Discrimination.

The conclusions obtained in this paper apply to alternative adverse selection models. An obvious candidate for this transposition is the repeated version of the monopoly price discrimination paradigm. Consider the following static model (see, e.g., Maskin-Riley (1984)). A monopolist produces a good at marginal cost c, and supplies an amount q to a buyer, who derives a surplus $V(q,b)$ from its consumption, where $V_q > 0, V_{qq} < 0, V_b > 0, V_{qb} > 0$. The taste parameter b is private information to the buyer and can take two values: $\underline{b}$ ("bad type" or "low valuation buyer") with probability $1 - \nu_1$ and $\bar{b}$ ("good type" or "high valuation buyer") with probability ν_1. Let $\bar{q}^*$ and $\underline{q}^*$ denote the complete information or socially optimal consumptions: $V_q(\bar{q}^*,\bar{b}) = V_q(\underline{q}^*,\underline{b}) = c$ (with $\bar{q}^* > \underline{q}^*$). Let $\Phi(q) \equiv V(q,\bar{b}) - V(q,\underline{b})$ with $\Phi' > 0$. The monopolist chooses an optimal non-linear price subject to the buyer's IR and IC constraints. In a single period context, the good type's consumption is socially optimal: $\bar{q} = \bar{q}^*$ while the bad type's consumption $\underline{q} = \underline{q}(\nu_1)$, which is lower than $\underline{q}^*$, maximizes the social surplus for this type minus the good type's rent:

$$\underline{q}(\nu_1) = \arg\max_q \{(1-\nu_1)(V(q,\underline{b}) - cq) - \nu_1\Phi(q)\}. \tag{8-1}$$

This price discrimination model is analogous to ours (b corresponds to minus β, q to minus c, etc.). Hence, we can apply our results to its twice repeated version. $V(\cdot,\cdot)$ and c are then per-period surplus and marginal cost. The solution will be called the "LT contracting solution" (where LT stands for "long-term," and the possibility of renegotiation under LT contracting is implicit). In the first period, the seller offers the buyer a choice between two consumption levels: $\bar{q}^*$, which is chosen by the good type only, and is followed by the same consumption in period 2; and $q_1(x)$, which is chosen by the bad type and possibly by the good type and is given by:

$$q_1(x) = \arg\max_q \{\nu_1(1-x)(V(q,\bar{b}) - cq) + (1-\nu_1)(V(q,\underline{b}) - cq) - \nu_1\Phi(q)\}, \tag{8-2}$$

where $(1-x)$ is the probability that the good type pools with the bad type. This pooling consumption is followed by the conditionally optimal price discrimination scheme, yielding consumptions $\bar{q}^*$ and $\underline{q}(\nu_2(x))$ to the good and bad types (where $\nu_2(x) \equiv \nu_1(1-x)/(\nu_1(1-x) + 1 - \nu_1))$.

Hart and Tirole (1987) solved this model in a T-period framework for the case of dichotomous consumption.[8] A main result of their paper is that the equilibrium LT contract is equivalent to the Coasian durable-good equilibrium. One may wonder whether an analogous result holds in the multiunit case.[9] Before tackling this problem, we make three remarks. First, the durable-good model has not yet been studied with non-binary consumption, to the best of our knowledge. Second, if such an equivalence result is to hold, we must consider non-linear pricing in each period in the durable-good model. Third, to make things comparable, we assume that supplying in period 1 a good that lasts for two periods costs $(1+\delta)c$, i.e., $(1+\delta)$ as much as supplying a single-period lived product.

It is straightforward to show that the seller cannot obtain more in the durable-good model than under a LT contract. For, in the LT contract framework, the seller can offer the consumption pattern corresponding to the durable-good equilibrium. In period 2, the buyer's consumption pattern is conditionally optimal for the seller (because the durable-good model has no commitment, the seller optimizes in the second period), and is thus renegotiation proof.[10]

Conversely, the LT contract outcome can be achieved by the durable-good monopolist subject to the caveat described below. For, a central result of our paper (transposed to price discrimination) is that following the pooling consumption, the seller uses the conditionally optimal price discrimination (see Theorem 1). So, consider the following strategies in the durable-good model: "In period 1, the seller offers for sale the quantities $q_1(x)$, at price $V(q_1(x),\underline{b})(1+\delta)$, and $\bar{q}^*$, at price $V(\bar{q}^*,\bar{b})(1+\delta) - \Phi(q_1(x)) - \delta\Phi(\underline{q}(\nu_2(x)))$ (where x is the equilibrium probability under LT contracting, and $q_1(x)$ is given by (8-2)). In period 2, no further offer is made if the buyer has purchased $\bar{q}^*$ in period 1. If the buyer has bought $q_1(x)$ in period 1, the seller offers quantities $(\bar{q}^* - q_1(x))$, at price $V(\bar{q}^*,\bar{b}) - V(q_1(x),\underline{b}) - \Phi(\underline{q}(\nu_2(x)))$, and $(\underline{q}(\nu_2(x)) - q_1(x))$, at price $V(\underline{q}(\nu_2(x)),\underline{b}) - V(q_1(x),\underline{b})$. The low-valuation buyer purchases $q_1(x)$ in the first period. The high-valuation buyer purchases $\bar{q}^*$ with probability x and $q_1(x)$ with probability $1-x$ in the first period." Given the first-period sale offers, the seller's and the buyer's behaviors clearly form a continuation equilibrium of the durable-good game. Furthermore, the first-period sale offers are optimal for the seller, because from our earlier result, the seller's profit in the durable-good model cannot exceed that for the optimal LT contract.

The caveat is apparent in the previous proof. For the equivalence result to hold, the buyer's consumption under LT contracting must be non-decreasing. This amounts to the condition: $q_1(x) \le \underline{q}(\nu_2(x))$. This condition holds for discount factors under some threshold level from Theorem 3.[11] For instance, for small discount factors, the equilibrium is separating ($x = 1$) so that $q_1(x) = \underline{q}(\nu_1) < \underline{q}(\nu_2(x)) = \underline{q}^*$. But for discount factors above the threshold level, $q_1(x)$ exceeds $\underline{q}(\nu_2(x))$, the durable-good monopolist's profit is strictly lower than the profit under LT contracting (because LT contracting allows decreasing consumption paths).

To summarize our study of the two-period framework, the equivalence between Coasian durable-good dynamics and LT contracting holds as long as the discount factor is lower than some threshold value, i.e., as long as the low valuation buyer's consumption under long-term contracting is time-increasing. Alternatively, our work can be viewed as generalizing the durable-good model to, and deriving the equilibrium for, multi-unit consumption.

Appendix 1: Proof of Lemma 1

We assume that the principal offers two contracts A and B in the first period. We shall later show that the use of more than two contracts does not increase the principal's welfare. Without loss of generality, we assume that the bad type's intertemporal utility is equal to zero (if it were equal to a strictly positive number, the principal could uniformly reduce all rents by this number and reach a higher welfare without perturbing any of the IR, IC and RP constraints). Furthermore we can choose the intertemporal structure of transfers to put the bad type's utility equal to zero in each period.

Let a_1 and b_1 denote the first-period costs in these two contracts, and a_2 and b_2 the corresponding bad type's second-period costs (Proposition 2 implies that the good type's second-period cost in both contracts is $\underline{\beta} - e^*$). Let A_1 and A_2, and B_1 and B_2 denote the good type's first- and second-period rents in contracts A and B.

Let x (respectively $1-x$) denote the probability that the good type chooses contract B (respectively A). Similarly y is the probability that the bad type chooses contract A. We assume that $1 > x,y > 0$, so that we have "double randomization." The good type randomizes between the two contracts only if he obtains the same intertemporal rent in both:

$$A_1 + \delta A_2 = B_1 + \delta B_2. \tag{A-1}$$

(A-1) will often be called the (first-period) incentive compatibility constraint.

Last let γ_2 denote the posterior probability that the agent has type $\underline{\beta}$ given that first-period cost was b_1 (i.e., contract B was chosen). Similarly μ_2 is the posterior probability following cost a_1.

Figure 4 summarizes the situation.

From our normalization (the rent of the bad type is zero in each period), the rent of the good type in period 1 is the static rent $\Phi(b_1)$ for contract B and $\Phi(a_1)$ for contract A. So we obtain:

Claim 1: $A_1 = \Phi(a_1)$ and $B_1 = \Phi(b_1)$.

From Corollary 1, we know that $A_2 \geq \underline{U}(\mu_2)$ and $B_2 \geq \underline{U}(\gamma_2)$. We next show that both second period contracts are rent constrained contracts and that one of the two is a conditionally optimal contract or:

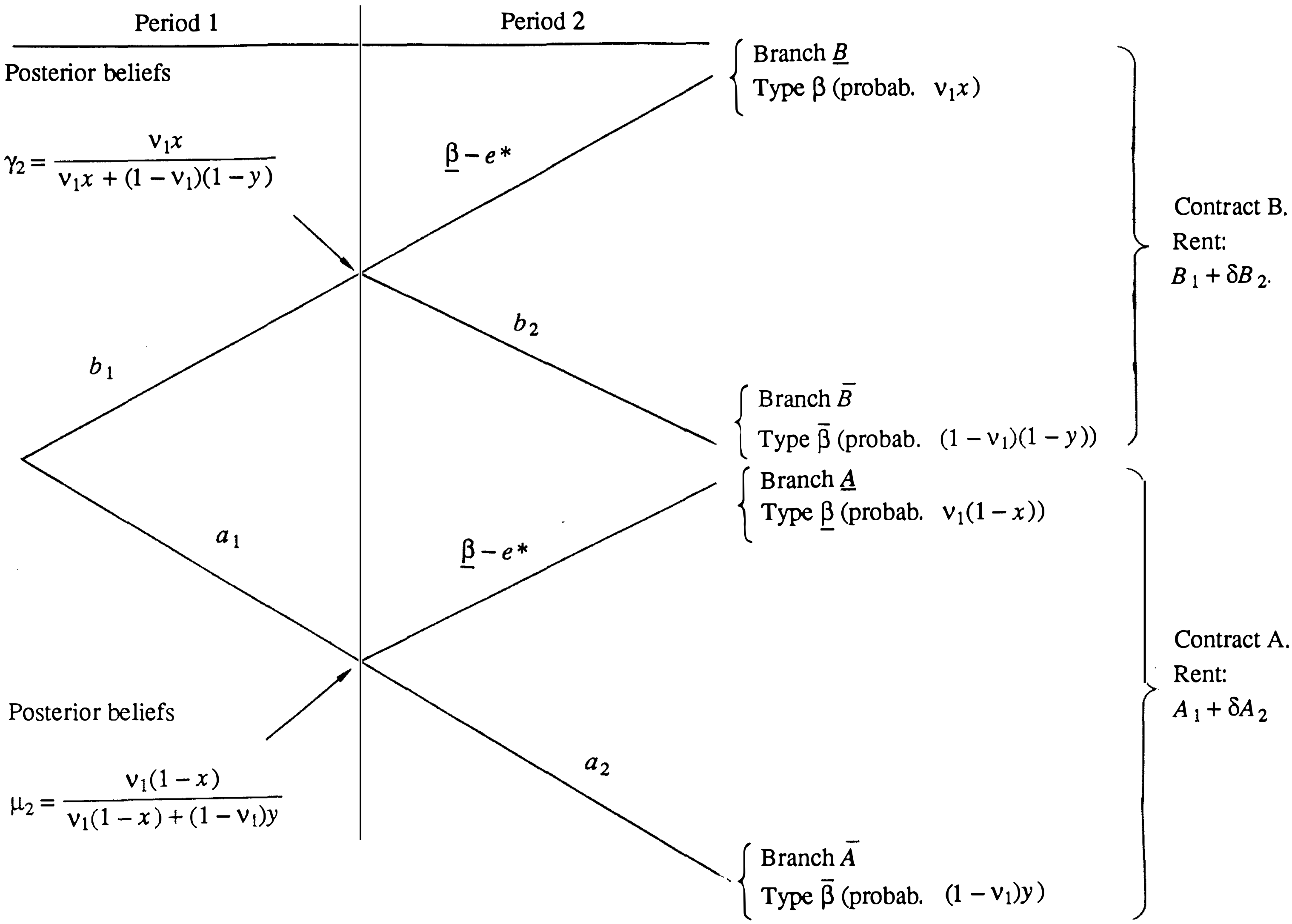
Claim 2:

(i) Either $A_2 = \underline{U}(\mu_2)$ or $B_2 = \underline{U}(\gamma_2)$

(ii) $A_2 = \Phi(a_2)$ and $B_2 = \Phi(b_2)$

Proof:

(i) Suppose that $A_2 > \underline{U}(\mu_2)$ and $B_2 > \underline{U}(\gamma_2)$. From Corollary 4, the principal could in the first period offer contracts that reduce A_2 and B_2 slightly and increase welfare. If A_2 and B_2 are reduced in equal amounts (which is feasible because they can be lowered continuously), the IC constraint (A.1) is kept satisfied and the randomizing probability and the first-period allocations can be kept the same.



Period 1
Period 2
Posterior beliefs
$\gamma_2 = \frac{\nu_1 x}{\nu_1 x + (1-\nu_1)(1-y)}$
$\underline{\beta} - e^*$
b_2
b_1
a_1
$\underline{\beta} - e^*$
a_2
Posterior beliefs
$\mu_2 = \frac{\nu_1(1-x)}{\nu_1(1-x) + (1-\nu_1)y}$
Branch $\underline{B}$
Type $\underline{\beta}$ (probab. $\nu_1 x$)
Branch $\bar{B}$
Type $\bar{\beta}$ (probab. $(1-\nu_1)(1-y)$)
Branch $\underline{A}$
Type $\underline{\beta}$ (probab. $\nu_1(1-x)$)
Branch $\bar{A}$
Type $\bar{\beta}$ (probab. $(1-\nu_1)y$)
Contract B.
Rent:
$B_1 + \delta B_2$.
Contract A.
Rent:
$A_1 + \delta A_2$

Figure 4

(ii) Suppose without loss of generality that $A_2 = \underline{U}(\mu_2)$ and that B specifies a sell-out contract in period 2. From Proposition 2, the sell-out contract is renegotiation-proof for any posterior γ_2. This implies that we can change the probabilities x and y without perturbing the renegotiation proofness of contract B.

Corollary 3 implies that renegotiation proofness of contract A is preserved if the new probabilities $\tilde{x}$ and $\tilde{y}$ are chosen so that the induced posterior $\tilde{\mu}_2$ exceeds μ_2, i.e.,

$$\frac{(1-\tilde{x})\nu_1}{\nu_1(1-\tilde{x})+(1-\nu_1)\tilde{y}} > \mu_2 \tag{A-2}$$

or

$$\nu_1(1-\mu_2)(1-\tilde{x}) \geq (1-\nu_1)\mu_2\tilde{y} \tag{A-3}$$

The principal's welfare $W(\tilde{x},\tilde{y})$ is linear in $\tilde{x}$ and $\tilde{y}$, keeping contracts (i.e., a_1, a_2, b_1 and b_2) constant. Its maximization with respect to $\tilde{x}$, $\tilde{y}$ under (A.3) and $0 \leq \tilde{x} \leq 1$ and $0 \leq \tilde{y} \leq 1$ yields corner solutions. Consequently, at least one of the x and y is 0 or 1 and the maximum of the principal's welfare can be reached without double randomization by the agent, a contradiction.

Q.E.D.

Claim 2 implies that (A-1) can be rewritten in the following way:

$$\Phi(a_1) + \delta\Phi(a_2) = \Phi(b_1) + \delta\Phi(b_2) \tag{A-4}$$

Let us assume w.l.o.g. that $a_1 \geq b_1$. Then, $\Phi(a_1) \leq \Phi(b_1)$ and therefore $a_2 \leq b_2$ from (A--4).

Claim 3: $c^P(\gamma_2) \leq c^P(\mu_2)$.

Proof: From (2-10), this amounts to showing that $\gamma_2 \geq \mu_2$. From claim 2, we have two cases to consider.

Case a: $a_2 = \bar{c}(\mu_2)$ and $b_2 \leq \bar{c}(\gamma_2)$.

The inequality $a_2 \leq b_2$ implies that $\bar{c}(\mu_2) \leq \bar{c}(\gamma_2)$, which from Proposition 1, yields $\mu_2 \leq \gamma_2$.

Case b: $b_2 = \bar{c}(\gamma_2)$ and $a_2 < \bar{c}(\mu_2)$.

From the strict concavity of the objective function in the commitment case, raising a_2 slightly strictly increases welfare. But to keep (A-4) satisfied, a_1 must be reduced slightly. This also increases welfare (or has a second-order effect) if $a_1 \geq c^P(\mu_2)$. Hence we have $a_1 < c^P(\mu_2)$.

Next, because b_2 is conditionally optimal, a small reduction in b_2 has only a second-order effect on second period welfare, and, from Proposition 2, preserves renegotiation proofness of contract B. So it must be the case that a slight increase in b_1 (so as to keep (A-4) satisfied) does not raise first period welfare. Hence $b_1 \geq c^p(\gamma_2)$. But since $a_1 \geq b_1$, $c^p(\mu_2) > c^p(\gamma_2)$ and $\mu_2 < \gamma_2$.

Q.E.D.

We are led to consider two cases through the next result.

Claim 4: Either $c^p(\gamma_2) \leq b_1 \leq a_1 \leq c^p(\mu_2)$ (Case 1)
Or $b_1 < c^p(\gamma_2) \leq c^p(\mu_2) < a_1$ (Case 2)

Proof: Suppose first that $b_1 \leq c^p(\gamma_2) \leq a_1 \leq c^p(\mu_2)$ with either $b_1 < c^p(\gamma_2)$ or $a_1 < c^p(\mu_2)$ (or both). Raising slightly a_1 and b_1 so as to keep $[\Phi(b_1) - \Phi(a_1)]$ constant (and thus (A-4) satisfied) raises the principal's welfare to the first order by bringing the first-period costs towards the optimal pooling cost corresponding to the mix of types associated with each contract. [This again results from the strict concavity of the pooling objective function. When $b_1 = c^p(\gamma_2)$, say, a slight increase in b_1 has only second order effects on the principal's welfare]. The proof is identical when $c^p(\gamma_2) \leq b_1 \leq c^p(\mu_2) \leq a_1$, with either $c^p(\gamma_2) < b_1$ or $c^p(\mu_2) < a_1$ (or both). It then suffices to reduce b_1 and a_1 slightly keeping (A.4) satisfied.

Q.E.D.

From Claim 4 we distinguish two cases.

Case 1. Let us show that a slight increase in x raises welfare. An increase in x amounts to a displacement of the "good type population" from branch $\underline{A}$ to branch $\underline{B}$. The good type's rent is unaffected; so is the second-period efficiency (because the good type produces at $\underline{\beta} - e^*$ in both cases). The first-period efficiency strictly increases if $b_1 < a_1$. [The case $b_1 = a_1$ is uninteresting as (A-4) then implies that $b_2 = a_2$, and thus the two contracts are identical and can be merged. Renegotiation proofness is preserved in the merger because the new posterior beliefs, equal to ν_1, are a convex combination of μ_2 and γ_2 and because of the fact that $\bar{c}(\nu_1) \leq \bar{c}(\gamma_2)$ and because of Corollary 3.] For, from (2-9),

$$\underline{\beta} - e^* < c^p(\gamma_2) \leq b_1 < a_1.$$

By strict concavity of the objective function under commitment, a reduction in the good type's cost above $\underline{\beta} - e^*$ raises welfare.

The next question is whether the increase in x maintains renegotiation proofness. It clearly does for contract B from Corollary 3. It also does for contract A unless $A_2 = \underline{U}(\mu_2)$ (also from Corollary 3). So assume that $A_2 = \underline{U}(\mu_2)$. A small increase in x requires a slight upward adjustment in A_2 (i.e., a slight downward adjustment in a_2) to preserve renegotiation proofness. But this increase in A_2 has only a second-order welfare effect, because the initial contract is conditionally optimal. Next, this decrease in a_2 requires a small increase in a_1 to keep (A-4) satisfied. But $a_1 \leq c^p(\mu_2)$ implies that an increase in a_1 raises first-period welfare (or does not affect it to the first-order).

So we conclude that a slight increase in x, together with small changes in a_1 and a_2 so as to keep (RP) and (A-4) satisfied, strictly increases welfare, a contradiction.

Case 2. First suppose that $B_2 = \underline{U}(\gamma_2)$. Then any small reduction in b_2 has a second-order effect on welfare and preserves renegotiation proofness by Proposition 2. A small increase in b_1 to keep (A-4) satisfied strictly increases welfare because $b_1 < c^p(\gamma_2)$. Hence $B_2 > \underline{U}(\gamma_2)$ (and therefore $A_2 = \underline{U}(\mu_2)$).

Keeping everything else (costs) constant, let $W(\tilde{x}, \tilde{y})$ denote the principal's welfare when the randomizing probabilities are $\tilde{x}$ and $\tilde{y}$. W is linear in $\tilde{x}$ and $\tilde{y}$. From Corollary 3, any $(\tilde{x}, \tilde{y})$ satisfying

$$\nu_1(1-\mu_2)(1-\tilde{x}) \geq \mu_2(1-\nu_1)\tilde{y} \tag{A-5}$$

yields posterior beliefs $\tilde{\mu}_2 \geq \mu_2$ in contract A and thus preserves renegotiation proofness in this contract. In the $(\tilde{x}, \tilde{y})$ space, the solution of the maximization of the linear objective function W over the half-space defined by (A-5) and over the constraints that $\tilde{x}$ and $\tilde{y}$ belong to [0,1] and that $B_2 \geq \underline{U}(\gamma_2(\tilde{x}, \tilde{y}))$ (renegotiation proofness of contract B) is a corner solution. Either $B_2 = \underline{U}(\gamma_2(\tilde{x}, \tilde{y}))$ and our previous condition is violated, or either $\tilde{x}$ or $\tilde{y}$ (or both) is equal to 0 or 1, and the double randomization assumption is violated.

We thus conclude that in both cases, maximal welfare can be reached without double randomization. That is, there exists a renegotiation proof contract that yields the same intertemporal rent to the good type, and at least as much welfare to the principal, and that involves randomization by at most a single type. Note in passing that this shows also that there is no point considering more than two contracts. With more than two contracts, one can apply the above reasoning to any pair of pooling contracts. Because it is possible to keep the agent's rent constant in the inductive reduction process, this shows that there is at most one pooling contract.

The next step in the proof of Lemma 1 consists in showing that randomization by the bad type only cannot be optimal for the principal. Suppose that $x = 1$ (the case $x = 0$ is treated identically). Then $a_2 = \bar{\beta} - e^*$ because, following a_1, it is common knowledge that the agent's type is $\bar{\beta}$.

Suppose first that

$$A_1 + \delta A_2 < B_1 + \delta B_2. \tag{A-6}$$

Then $a_1 = \bar{\beta} - e^*$ (moving a_1 towards $\bar{\beta} - e^*$ raises efficiency and affects neither the incentive constraint (A-6) nor the good type's rent. Because branch $\bar{A}$ is efficient (the bad type produces at the efficient cost in each period), an increase in y raises efficiency and preserves renegotiation proofness of contract B by raising γ_2 (from Corollary 3). Thus there exists a dominating separating equilibrium (with $y = 1$).

Second, suppose that

$$A_1 + \delta A_2 = B_1 + \delta B_2. \qquad \text{(A-7)}$$

Let $W(y)$ denote the principal's welfare when y varies, everything else being kept constant. It is linear in y. If $W_y \geq 0$, one can increase y without reducing welfare, and keep renegotiation proofness in contract B. If $W_y < 0$, a slight decrease in y strictly raises welfare. However, it lowers γ_2, and to preserve renegotiation proofness in contract B, the principal must increase B_2 (i.e., lower b_2) slightly. Because the second-period contract following b_1 is conditionally optimal, this adjustment has only a second-order effect on the principal's welfare. Hence the upper bound cannot be reached by having only the bad types randomize, which completes the proof of Lemma 1.

Q.E.D.

APPENDIX 2

Proof of Theorem 4 — No Separation for a Continuum of Types

Let us recall the commitment solution (see our 1986 paper). The optimal effort is given by:

$$\psi'(e^*(\beta)) = 1 - \frac{\lambda}{1+\lambda}\frac{F(\beta)}{f(\beta)}\psi''(e^*(\beta)), \tag{A-7}$$

and the agent's rent is

$$U(\beta) = \int_{\beta}^{\bar{\beta}} \psi'(e^*(x))dx. \tag{A-8}$$

Replacing $e^*(\beta)$ by $\beta - c^*(\beta)$, the commitment cost $c^*(\beta)$ is also given by (A-7).
Differentiating (A-7) yields:

$$\frac{dc^*}{d\beta} \equiv A(\beta) = 1 + \frac{\frac{\lambda}{1+\lambda}\psi''(e^*(\beta))\frac{d}{d\beta}\left[\frac{F(\beta)}{f(\beta)}\right]}{\psi''(e^*(\beta)) + \frac{\lambda}{1+\lambda}\frac{F(\beta)}{f(\beta)}\psi'''(e^*(\beta))}. \tag{A-9}$$

Now consider the small change described in the text. The types in $[\bar{\beta}-\varepsilon, \bar{\beta}]$ pool at cost $c^*(\bar{\beta}-\varepsilon)$ in the first period. Following $c^*(\bar{\beta}-\varepsilon)$, the principal offers the commitment contract for the truncated distribution $(F(\beta) - F(\bar{\beta}-\varepsilon))/(1 - F(\bar{\beta}-\varepsilon))$ for $\beta \geq \bar{\beta}-\varepsilon$. It is straightforward to check that the new allocation is incentive compatible (this is due to the fact that the first- and second-period efforts of type $\bar{\beta}-\varepsilon$ are unchanged and that, by concavity, the types in $[\beta, \bar{\beta}-\varepsilon]$ would pool with type $\bar{\beta}-\varepsilon$ if they were forced to pool with a type in $[\bar{\beta}-\varepsilon, \bar{\beta}]$. The change in first-period welfare

ΔW_1 is given by

$$\Delta W_1 \equiv G_1 - L_1, \tag{A-10}$$

where G_1 is the gain in efficiency and L_1 the loss due to the increase in the agent's rent. We have:

$$G_1 = \int_{\bar{\beta}-\varepsilon}^{\bar{\beta}} (1+\lambda)[\psi(\beta - c^*(\beta)) + c^*(\beta) - \psi(\beta - c^*(\bar{\beta}-\varepsilon)) - c^*(\bar{\beta}-\varepsilon)]f(\beta)d\beta \tag{A-11}$$

$$\simeq \int_{\bar{\beta}-\varepsilon}^{\bar{\beta}} (1+\lambda)(c^*(\beta) - c^*(\bar{\beta}-\varepsilon))(1 - \psi'(\bar{\beta} - c^*(\bar{\beta})))f(\beta)d\beta.$$

But, from (A-7), and $F(\bar{\beta}) = 1$,

$$1 - \psi'(\bar{\beta} - c^*(\bar{\beta})) = \frac{\lambda}{(1+\lambda)} \frac{\psi''(e^*(\bar{\beta}))}{f(\bar{\beta})}, \quad \text{(A-12)}$$

and from (A-9):

$$c^*(\beta) - c^*(\bar{\beta} - \varepsilon) \simeq A(\bar{\beta})(\beta - \bar{\beta} + \varepsilon). \quad \text{(A-13)}$$

Substituting (A-12) and (A-13) into (A-11) yields:

$$G_1 \simeq \int_{\bar{\beta}-\varepsilon}^{\bar{\beta}} (1+\lambda) A(\bar{\beta})(\beta - \bar{\beta} + \varepsilon) \frac{\lambda}{1+\lambda} \psi''(e^*(\bar{\beta})) d\beta \quad \text{(A-14)}$$

or

$$G_1 = \lambda A(\bar{\beta}) \psi''(e^*(\bar{\beta})) \frac{\varepsilon^2}{2} + 0(\varepsilon^3). \quad \text{(A-15)}$$

Next we compute L_1. Because $e_1(\beta)$ is unchanged for $\beta \leq \bar{\beta} - \varepsilon$, the rent of each type $\beta \leq \bar{\beta} - \varepsilon$ increases by the same amount as that of type $\bar{\beta} - \varepsilon$ (the increase in the rents of types $\beta > \bar{\beta} - \varepsilon$ is socially negligible relative to that of types $\beta \leq \bar{\beta} - \varepsilon$, because the former types have negligible weight relative to the latter types for ε small). The increase in the rent of type $\bar{\beta} - \varepsilon$ is given by:

$$\delta U(\bar{\beta} - \varepsilon) = \int_{\bar{\beta}-\varepsilon}^{\bar{\beta}} [\psi'(\beta - c^*(\bar{\beta} - \varepsilon)) - \psi'(\beta - c^*(\beta))] d\beta \quad \text{(A-16)}$$

$$\simeq \int_{\bar{\beta}-\varepsilon}^{\bar{\beta}} \psi''(\bar{\beta} - c^*(\bar{\beta}))(c^*(\beta) - c^*(\bar{\beta} - \varepsilon)) d\beta$$

$$\simeq \int_{\bar{\beta}-\varepsilon}^{\bar{\beta}} \psi''(\bar{\beta} - c^*(\bar{\beta})) A(\bar{\beta})(\beta - \bar{\beta} + \varepsilon) d\beta$$

$$= A(\bar{\beta}) \psi''(\bar{\beta} - c^*(\bar{\beta})) \frac{\varepsilon^2}{2}.$$

But

$$L_1 = \lambda \delta U(\bar{\beta} - \varepsilon) = \lambda A(\bar{\beta}) \psi''(\bar{\beta} - c^*(\bar{\beta})) \frac{\varepsilon^2}{2} + 0(\varepsilon^3) = G_1 + 0(\varepsilon^3). \quad \text{(A-17)}$$

As claimed in the text, we have

$$\Delta W_1 = 0(\varepsilon^3). \tag{A-18}$$

Let us now consider the second period. The change in welfare is given by

$\delta \Delta W_2$, where

$$\Delta W_2 = G_2 - L_2, \tag{A-19}$$

G_2 is the gain coming from the reduction in the agent's rent and L_2 is the loss in efficiency. The computation of G_2 is identical to that of L_1, except that the effort of the high type is in the second period e^*, and not $e^*(\bar{\beta})$ like in period 1. As can easily be checked, this implies that the new $A(\bar{\beta})$, computed from the new effort e^* and from the truncated distribution, is equal to 1. Hence:

$$G_2 = \lambda \psi''(e^*)\frac{\varepsilon^2}{2} + 0(\varepsilon^3). \tag{A-20}$$

In contrast, L_2 is of the third order in ε, because the initial allocation is cost efficient. More formally:

$$L_2 = \int_{\bar{\beta}-\varepsilon}^{\bar{\beta}} (1+\lambda)(\psi(\beta - \tilde{c}(\beta)) + \tilde{c}(\beta) - \psi(e^*) - \beta + e^*)f(\beta)d\beta, \tag{A-21}$$

where $\tilde{c}(\beta)$ is the commitment solution for the truncated distribution:

$$\psi'(\beta - \tilde{c}(\beta)) = 1 - \frac{\lambda}{1+\lambda}\frac{F(\beta) - F(\bar{\beta}-\varepsilon)}{f(\beta)}\psi''(\beta - \tilde{c}(\beta)). \tag{A-22}$$

Note that for ε small,

$$\psi(\beta - \tilde{c}(\beta)) - \psi(e^*) = (\beta - \tilde{c}(\beta) - e^*) + \frac{1}{2}\psi''(e^*)(\beta - \tilde{c}(\beta) - e^*)^2, \tag{A-23}$$

using $\psi'(e^*) = 1$. Hence, (A-21) can be rewritten as:

$$L_2 = \int_{\bar{\beta}-\varepsilon}^{\bar{\beta}} \frac{(1+\lambda)}{2}\psi''(e^*)(\beta - \tilde{c}(\beta) - e^*)^2 f(\beta)d\beta. \tag{A-24}$$

But, from (A-22) and $1 = \psi'(e^*)$:

$$\beta - \tilde{c}(\beta) - e^* \simeq \frac{\lambda}{1+\lambda}(\beta - \bar{\beta} + \varepsilon). \tag{A-25}$$

(A-24) and (A-25) yield:

$$L_2 \simeq \frac{\lambda^2}{6(1+\lambda)} \psi''(e^*) f(\bar{\beta}) \varepsilon^3 = 0(\varepsilon^3). \qquad \text{(A-26)}$$

We thus conclude that

$$\Delta W_1 + \delta \Delta W_2 \simeq \delta G_2 > 0. \qquad \text{(A-27)}$$

Q.E.D.

FOOTNOTES

1. See Fudenberg-Tirole (1988) for a study of renegotiation in a moral-hazard framework, and Green and Laffont (1987) (1988) for the case of symmetric but nonverifiable information. These papers as well as the present work use a principal agent framework (i.e., concern renegotiation in contract theory). For the study of noncontractual renegotiation in game theoretic contexts see in particular Pearce (1987) and Farrell and Maskin (1987).
2. This assumption in particular ensures that the optimal incentive scheme under commitment is deterministic. More generally, our results would hold as long as ψ''' is "not too negative".
3. The reader may be worried that no such assumption can be made in a dynamic model. Indeed the second-period beliefs ν_2 might be close to 1 even though the prior beliefs ν_2 are assumed not to. This possibility fortunately does not arise in our model as the equilibrium path will involve either $\nu_2 = 1$ or $\nu_2 \leq \nu_1$. It can indeed be shown that for any ν_1 under some cut-off level, the equilibrium is as described in this paper.
4. Giving up a rent higher than $\Phi(\underline{\beta} - e^*)$ to the good type clearly serves no purpose since the allocation is already optimal in period 2 and the good type reveals himself in period 1 for the optimal static mechanism.
5. The best pooling contract dominates the pooling contract specifying $c_1 = \underline{\beta} - e^*$ for both types. But, because $\bar{c}(\nu_1) - (\underline{\beta} - e^*)$ is proportional to $\Delta\beta$ for $\Delta\beta$ small, the welfare distortion of this alternative pooling contract relative to commitment is itself of the second order.
6. The best separating contract dominates the best pooling contract for δ small, and the converse holds for δ larger (by the same reasoning as in the proof of Theorem 3, but restricting the choice of x between two values; 0 and 1).
7. Baron and Besanko (1987) study a particular form of limited commitment. The firm promises to produce in period 2 and the principal commits to use in period 2 a mechanism which is "fair," i.e., which leaves to the firm nonnegative profits given the information transmitted in period 1.
8. The binary nature of consumption simplifies matters in many respects. First, the socially optimal consumption is not type-contingent (which, for instance, implies that there exists a single sell-out contract instead of a continuum of them). Second, although some continuous consumption choice is introduced into the binary model by considering a probability that the buyer consumes in each period the nature of the proof has a simple bang-bang flavor (for instance, the critical beliefs for a socially inefficient contract to be renegotiation-proof in period 2 are independent of the bad type's probability of consumption, while they depend on c_2 in our paper).
9. We are grateful to Oliver Hart for suggesting this question.

10. This simple reasoning holds only in the two-period model. With more than two periods, a more elaborate argument is needed. See Hart-Tirole (1987) for the binary consumption case.

11. Theorem 3 implies that x is a nonincreasing function of δ. Furthermore, $\underline{q}(\nu_2(x))$ is increasing in x while $q_1(x)$ decreases with x.

12. For any contract the rent grows at rate $\dot{U}(\beta) = -\psi'(e_2(\beta))$ (see our 1986 paper. This is the differentiable version of equation (2-4)).

REFERENCES

Baron, D. and D. Besanko. (1984). "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 447-470.

Baron, D. and D. Besanko. (1987). "Commitment and Fairness in a Continuing Relationship," *Review of Economic Studies*, 54, 413-436.

Dewatripont, M. (1986). "Renegotiation and Information Revelation Over Time in Optimal Labor Contracts," Ch. 1 in *On the Theory of Commitment, with Applications to the Labor Market*, PhD thesis, Harvard University.

Farrell, J. and E. Maskin. (1987). "Renegotiation in Repeated Games," *Harvard Institute of Economic Research* 1335.

Fudenberg, D. and J. Tirole. (1988). "Moral Hazard and Renegotiation in Agency Contracts," mimeo, MIT.

Green, J. and J. J. Laffont. (1987). "Renegotiation and the Form of Efficient Contracts," HIER DP no. 1338.

____________. (1988). "Nonverifiable Information Renegotiation and Opportunism," mimeo.

Hart, O. and J. Tirole. (1987). "Contract Renegotiation and Coasian Dynamics," MIT DP 442.

Laffont, J. J. and J. Tirole. (1986). "Using Cost Observation to Regulate Firms," *Journal of Political Economy*, 94, 614-641.

____________. (1987). "Comparative Statics of the Optimal Dynamic Incentive Contract," *European Economic Review*, 31, 901-926.

____________. (1988). "The Dynamics of Incentive Contracts," *Econometrica* (forthcoming).

Maskin, E. and J. Riley. (1984). "Monopoly with Incomplete Information," *Rand Journal of Economics*, 15, 171-196.

Pearce, D. (1987). "Renegotiation-Proof Equilibria: Collective Rationality and Intertemporal Cooperation," *Cowles Foundation DP* 855.

Roberts, K. (1983). "Long Term Contracts," University of Warwick mimeo.