Working Paper Series

# Limited Commitment and Costly Enforcement

WP 90-02

Jeffrey M. Lacker
Federal Reserve Bank of Richmond

This paper can be downloaded without charge from:
http://www.richmondfed.org/publications/

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

TEXT
6520

Federal Reserve Bank
of San Francisco
JUN 11 1990
LIBRARY

Working Paper 90-2

# LIMITED COMMITMENT AND COSTLY ENFORCEMENT

Jeffrey M. Lacker

Research Department, Federal Reserve Bank of Richmond
P.O. Box 27622, Richmond, VA 23261, 804-697-8279

August 28, 1989
Revised, December 19, 1989

**ABSTRACT**: A costly "facility" has a monopoly on the ability to coerce transfers and verify all private information. If invoked, the facility reads instructions recorded *ex ante*, and carries out the contingent transfers among agents, charging agents for the cost. Agents agree *ex ante* to a set of recorded instructions to the facility, and then play a sequential game without commitment. A basic two-agent insurance environment serves as an application throughout. When both agents have full information the costs and limitations of the facility constrain the set of attainable allocations, even though the facility is never invoked in equilibrium. When there is private information, the model can be viewed as a reformulation of the standard costly-auditing model, but the incentive constraints are significantly more severe. Pure strategy optimal contracts are debt contracts, as in Townsend (1979), but mixed strategy optimal contracts cannot be ruled out in general. An extension shows that if costs vary with the realized state in a particular way, debt contracts as in Williamson (1987) can be optimal, even allowing for mixed strategies.

I would like to thank John Weinberg for many lengthy conversations, William Novshek, and Stacey Schreft for helpful comments on an earlier draft. I alone remain responsible for any errors. The views expressed are mine and do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

As part of an intertemporal exchange, one party often promises to take an action at a future date which will not then be in their immediate interest--repaying a loan, for example. Why are these promises credible? People sometimes fulfill their promises in order to maintain a reputation, valuable at a future date, or to enjoy the gains from an enduring relationship. Yet in many intertemporal arrangements between agents such considerations are absent, either because the desired relation is of inherently limited duration or because it is feasible for an agent to depart later to a location where current reputation is of no use. Why are intertemporal promises credible in one-time arrangements?

The immediate answer, of course, is that agents write contracts. Specifically, one agent, called *a*, might promise to pay a sum in the future to another agent, called *b*. The two agents create a document, called a "contract," that describes the promise. If agent *a* fails to live up to the promise, agent *b* can present the document to a facility--"the court"--and the facility will take some actions affecting both agents.[1] Agent *a*'s promise is credible because both agents know that, when the time comes, agent *a* would rather fulfill the promise than incur the action of the facility that would result from not fulfilling the promise. The facility might, for example, confiscate the sum agent *a* had promised to pay, perhaps collecting an additional sum as a penalty. In order to be credible, of course, the action of the facility must also make it credible for agent *b* to promise to invoke the facility in the event that agent *a* reneges on the contract.

This paper studies contingent contracts in a simple two-agent insurance environment when the enforcement facility is costly.[2] Agent *a* has a random endowment and is to make a payment, possibly contingent, to agent *b*, whose endowment is not random. Neither agent is able to confiscate the good from the other agent; the facility has a monopoly on the ability to coerce transfers of the good between agents.[3] The facility is capable of verifying the realized value of the random endowment and the payments actually made by agents. Agents are able to record, *ex ante*, an agreed-upon schedule specifying the facility's transfers among agents as a function of the realized endowment and payment. If invoked by either party, the facility follows this prearranged schedule automatically. The facility imposes only one condition on the contracts it enforces: the sum of the net transfers from the two agents must in all cases meet or exceed the resource cost to the facility. Any transfer schedules meeting this feasibility condition are allowed.

Given the *ex ante* agreement, agents choose subsequent actions without commitment--they play a sequential game. The payoffs embodied in the transfer schedule induce an equilibrium in the sequential game, so we can think of agents as agreeing *ex ante* to certain (possibly mixed) strategies and then specifying the transfer schedule necessary to "support" these strategies as an equilibrium of the sequential game. The transfers serve as punishments for off-equilibrium strategies, and help induce agent *a* to voluntarily make the requisite payment.

When the random endowment is publicly observed by agent *b*, we find that any payment schedule can be supported as long as it provides agent *a* with a minimum consumption equal to the cost of enforcement, even though the facility is never invoked in equilibrium. The intuition is plain: suppose agent *a* makes a payment and retains an amount just slightly less than the cost of enforcement. Then if agent *b* invokes the facility, the difference between the cost of the facility and the amount retained by agent *a* must be collected from agent *b*, making *b* worse off for having invoked. Knowing that *b* will not invoke the facility, *a* can get away with any payment that yields consumption less than or equal to the facility cost; the only credible promise is one that leaves agent *a* with consumption at least as large as the facility cost. Thus costly enforcement can affect equilibrium contracts in which enforcement never takes place. Subject to this lower bound on agent *a*'s consumption, however, any arbitrary payment schedule can be optimal.

When the random endowment is unobserved by agent *b*, contracts are more sharply constrained. Now agent *a* makes a payment and agent *b* takes a costly action, invoking the facility, that causes the state to be irrefutably revealed and a further payment to be made. Thus the facility can be interpreted as a audit service possessing enforcement powers, and the model can be viewed as a reformulation of standard costly verification or costly auditing models (see Townsend 1979, 1988, Border and Sobel 1987, Mookherjee and Png 1989, Moore 1989). Optimal contracts will in general involve enforcement with positive probability in some states. If the optimal contract involves only pure strategies for both agents, or if we restrict attention to pure strategy contracts, then the optimal contract is a simple debt contract, exactly as in Townsend (1979). Mixed strategy contracts cannot be ruled out in general. However, we can establish that although the facility would allow it, it is never optimal for agents to instruct the facility to introduce extraneous uncertainty into transfer schedules, so that deterministic facility transfers always suffice.

There are a number of important and striking differences between costly enforcement and the standard costly verification setup; these are explored in detail in Section 4. First, the sequential rationality of agent *b*'s decision to invoke the facility implies that (in the language of the costly auditing framework) the audited agent pays for the audit, the opposite of the results of Border and Sobel (1987), Townsend (1988) and Mookherjee and Png (1989).[4] Second, agent *b*'s decision is made knowing only the payment made by agent *a*. Consistency of agent *b*'s beliefs imposes significant constraints across states on equilibrium transfer schedules that are absent in the costly verification setup. Third, we treat agent *a* as selecting a payment to make to agent *b*, rather than selecting a message concerning the realized state. Thus incentive constraints on agent *a* are significantly more severe here, involving action-by-action moral hazard constraints for each state rather than just state-by-state revelation constraints. These differences arise because the standard costly verification model, like many other models of bilateral contracts under imperfect information, is an application of a Principal-Agent or mechanism design framework, whereas here both agents are treated as players in a sequential game that they themselves (in part) design. The central message of the paper is that this difference can be quite important, and that the incentive constraints implied by the latter approach can be significantly more stringent. It is worth noting, however, that the framework described in this paper is not inconsistent with the Revelation Principle and the theory of mechanism design. An appendix shows how finding optimal contracts in the private information case can be formulated as a simple--in fact, linear--programming problem, with constraints like those in Townsend (1988), Prescott and Townsend (1984) or Myerson (1986).

In principle, the framework described here could be extended in many ways. One interesting case is when the facility's cost is allowed to vary with the realized state in a particular way. We assume that the payment is public, and that the amount retained by agent *a* is costly to ascertain if it is positive but it is costless to verify if none remains. If both agents are risk neutral in this setting, debt contracts of the form described by Williamson (1987) are optimal, even though we allow mixed strategies.[5] The example is interesting because it formalizes a novel explanation for debt contracts; if it is easy to prove to a creditor that your wealth is exhausted, but hard to prove its exact value when it is not, it might be optimal to make the contract contingent in the former case and noncontingent in the latter.

Many recent papers have considered models in which the enforcement of contracts is imperfect in that some information that is known to agents is unverifiable by the enforcement facility: see, for example, Huberman and Kahn (1987, 1988), Hart and Moore (1988), Green and Laffont (1987), and Fudenberg and Tirole (1988).[6] Contracts then play a role in structuring *ex post* bargaining. Although this set of assumptions is not pursued here, the objectives are closely parallel. The imperfection in commitment studied here is that enforcement is costly; for the works just cited, the enforcement facility is poorly informed. We should note here that *ex post* contractual renegotiation, an important element in some of the papers just cited, is an opportunity implicitly available to agents in the environment of this paper but is redundant, even though agents would wish to forego the deadweight enforcement cost if they were fully informed *ex post*.

The environment studied here is also related to the recent literature on breach of contract (Shavell 1980 and 1984, Rogerson 1983). As in the literature just cited, the enforcement facility is informationally handicapped, or contracts are for some reason incomplete. The focus is on the economic effects of alternative legal rules (definitions of "damages"). We begin at the opposite pole, examining the effects of an "ideal" civil court system in which agents design their own penalties and rewards, constrained only by the feasibility of allocations.

The next section describes the general environment to be studied. Section 2 examines the case in which agent *b* observes the random endowment of agent *a*. In Section 3 the case in which agent *b* does not observe *a*'s endowment is examined. In Section 4 the model is compared with models of costly auditing. The case of variable enforcement costs is considered briefly in Section 5. An appendix shows how in the private information case the optimal contract is the solution to a linear programming problem.

## 1. Economic Environment

There are two agents, named *a* and *b*. At time $t=1$ agent *a* will receive a quantity $x$ of a divisible consumption good. At time $t=0$, the quantity of the good is random. $x$ is an element of the finite set $\mathbf{X}$, and $f(x)$, the probability of $x$, is strictly positive for all $x$ in $\mathbf{X}$. Define $\bar{x}$ as $\text{MAX}\{x \in \mathbf{X}\}$, and $\underline{x}$ as $\text{MIN}\{x \in \mathbf{X}\}$. For simplicity, we will assume that agent *b*'s endowment is zero. This is an important simplification, since it allows us to restrict attention to contracts in which only one agent wants to invoke the facility;

if *b* has a positive endowment, equilibrium arrangements might require a payment from *b* to *a*, and thus might require *a* to invoke the facility. In the next section we will assume that the value of x is publicly observed, and in subsequent sections we will assume the value of x is private information to agent *a*, and thus unknown to *b* when it is realized. Given a realized value of x, agent *a* can transfer some amount y to agent *b*. Define a set $Y \equiv [0, \bar{x}]$, the set of payments that are feasible for some $x \in X$. We will assume that y can in no case be greater than x; so agent *a* chooses y from $Y(x) \equiv Y \times [0,x]$.

There is a "facility," to be described in detail shortly, that is capable of transferring goods to and from either or both agents. The facility acts, however, only if invoked by one of the agents. Define $\tau^a$ as the amount, possibly negative, collected from agent *a*, and $\tau^b$ as the amount, possibly negative, collected from agent *b*. Then agent *a*'s consumption is $c^a = x - y - \tau^a$, and agent *b*'s consumption is $c^b = y - \tau^b$. These provide utility of $u^a(c^a)$ and $u^b(c^b)$ to agents *a* and *b* respectively. If the facility is not invoked, consumptions are simply $c^a = x - y$ and $c^b = y$.

Agents *a* and *b* meet in an initial period and are capable of making written records then. In particular, they are capable of recording detailed instructions to the facility regarding how transfers are to be carried out in the event that the facility is invoked. For the purposes of our model, we require that the record be unalterable, durable, and verifiable. We have in mind a notarial service that makes a record of the agreement, verifies the identities of the parties involved and their consent to the agreement, and stores the record, making it available for future inspection. Note that the *ex ante* transfer schedule agreement closely resembles actual explicit contracts, more so than standard models in which contracts are implicit; the agreement specifies desired behavior, along with the remedy and punishment for noncompliance or off-equilibrium behavior.

The enforcement facility has a monopoly on the power to coerce transfers of goods between agents; otherwise transfers are strictly voluntary. One can imagine that an authority available to the facility possesses the only coercive force in the economy, and no individual agent on their own is capable of coercing any other agent. The facility is also capable of discovering the exact true values of both the random endowment, x, and the payment, y, made by agent *a*. One can interpet this as the ability of a hearing or court proceedings to uncover and verify facts, perhaps by using force or threat of penalty to compel testimony.

If invoked, the facility reads the record that was made *ex ante* by the two agents. The record specifies a schedule of transfers to be carried out by the facility. Since the facility is to be invoked, if ever, after the endowment is realized and the payment is made, the transfer can be made contingent on y and x. For the greatest generality, we assume that this means specifying a lottery, $\pi(\tau|y,x)$, giving the probability of transfers $\tau=(\tau^a,\tau^b)$ conditional on the realized values of y and x. The facility carries out the transfers automatically.[7,8]

The operation of the facility involves some costs when it is invoked, and these will be denoted $\gamma \geq 0$. For simplicity we will assume that $\gamma \leq \underline{x}$. The costs are incurred by the facility but must be recovered somewhere, so we will assume that they are recovered from the contracting parties. We neglect consideration of any costs of maintaining the facility on "standby," that is, the costs incurred having the facility in place, even if not invoked.[9] These would resemble the cost of a public good, and cost allocation would presumably involve the same welfare principles, although complicated by the manner in which cost allocation could distort contract choice. If costs are to be recovered from the contracting parties, then we must have $\tau^a + \tau^b \geq \gamma$ with probability one. A transfer schedule is *feasible* if it has this property. There is nothing about the notarial service, however, that prevents agents from writing down transfer schedules that are not feasible. We can suppose, however, that the rules governing the operation of the facility specify a mapping from the space of all possible recorded messages to the space of feasible transfer schedules. This would include, for example, a rule stating that infeasible transfer schedules, or otherwise unintelligble recordings, generate punitive transfers for each agent if the facility is invoked--a penalty for frivolous suits or suits brought on inadmissable contracts. If these rules of operation are known to all agents, we can assume that agents select directly the feasible transfer schedule that is the outcome of this mapping. We adopt this formalization without explicit proof.

## 2. Optimal Contracts With the Endowment Publicly Observed

After specifying a transfer schedule in the intitial period, agents take a sequence of actions <u>without</u> precommitment: they play a sequential game. First agent *a*'s endowment, x, is realized. Then agent a chooses an amount, possibly random, to transfer to agent *b*. Agent *b* then decides whether or not to invoke the facility. Payoffs are then determined by the transfer schedule.

More formally, for any given realized endowment x, a strategy for agent a is a lottery, $\pi^a(x)$ over possible payments, specifying the probability

$\pi^a(y|x)$ of paying the amount y in $Y(x)$. Given a realized payment, y, agent *b* chooses a probability of invoking the facility. Define a dummy variable, s, that is equal to one if *b* elects to invoke the facility, or "sue," and zero if *b* elects not to "sue." For a given realization of the endowment, x, and the payment, y, a strategy for *b* is a measure $\pi^b(y,x)$ on $\{0,1\}$: $\pi^b(s=1|y,x)$ and $\pi^b(s=0|y,x)$ give the probabilities that *b* sues or does not sue, respectively.

The expected utility of agent *a* under a given strategy $\pi^a(x)$, and for a given realization of x, depends on the strategy of agent *b* and on the following preagreed transfer schedule.

$$\int_Y \left\{\left[\int_\tau u^a(x-y-\tau^a)\pi(\tau|y,x)d\tau\right]\pi^b(s=1|y,x) \quad + \quad u^a(x-y)\pi^b(s=0|y,x)\right\} \pi^a(y|x)dy. \tag{2.1}$$

For every realized value of x, agent *a* takes as given agent *b*'s strategy, $\pi^b$, and the transfer schedule, $\pi$, and selects $\pi^a(x)$ to maximize (2.1).

We can now describe the expected utility of agent *b*. Taking the transfer schedule, $\pi$, as given, the choice of strategy $\pi^b$ by *b* provides expected utility of

$$\left[\int_\tau u^b(y-\tau^b)\pi(\tau|y,x)d\tau\right]\pi^b(s=1|y,x) \quad + \quad u^b(y)\pi^b(s=0|y,x). \tag{2.2}$$

For any given y, agent *b* takes $\pi$ as given and selects $\pi^b$ to maximize (2.2).

An *equilibrium*, for a given feasible transfer schedule $\pi(\tau|y,x)$, is a pair of strategies, $\pi^{a*}(y|x)$ and $\pi^{b*}(s|y)$, that satisfy the following conditions.

(i) The strategy $\pi^{a*}(y|x)$ maximizes (2.1) for all $x \in \mathbf{X}$, taking $\pi^{b*}(s|y,x)$ and $\pi(\tau|y,x)$ as given.

(ii) The strategy $\pi^{b*}(s|y)$ maximizes (2.2) for all $y \in \mathbf{Y}$, taking $\pi(\tau|y,x)$ as given.

A broad class of allocations can be supported as equilibrium contracts in this environment, as one might expect. For now, restrict attention to contracts with pure strategies for agent *a*, in which *a* is to transfer y(x) if the endowment is x. A candidate deterministic transfer schedule that might support this schedule is

$$\tau^a(y,x) = \begin{cases} x - y, & \text{if } \; y < y(x), \\ -y + \gamma, & \text{otherwise.} \end{cases}$$

$$\tau^b(y,x) = \tau^a(y,x) + \gamma. \tag{2.3}$$

If the payment is less than the preagreed amount y(x), and the facility is invoked, then agent *a* is "punished" by having all of the remaining good confiscated and handed over to agent *b*, net of the cost of the facility, $\gamma$. If the facility is invoked but the payment was at least as great as the preagreed amount, then the facility confiscates all of the good in the possession of agent *b*. An equilibrium strategy for agent *b* might be:

contract. Furthermore, any mixed strategy contract involves extraneous uncertainty and some costly enforcement with positive probability.[10] Such contracts are always strictly dominated because of the deadweight costs and (possibly) strict risk aversion. Therefore, we have the following result.

> <u>Proposition 1</u>: If the random endowment of agent *a* is publicly observed, then a contract is optimal if and only if the payment schedule is deterministic and never greater than $x - \gamma$ ($\pi^a(y(x)|x) = 1$, where $0 < y(x) < x - \gamma$); Furthermore, in an optimal contract enforcement occurs only if the payment is less than the scheduled amount ($\pi^b(s=1|y,x) > 0$ if $y < y(x)$, and 0 otherwise).

Thus, under perfect information the cost of enforcement constrains equilibrium contracts, even if the enforcement facility is never invoked. The constraints take the form of a range condition, a maximum payment for each state. Within the class of payment schedules satisfying this range condition, however, contracts are unrestricted. <u>Any</u> contingent schedule satisfying the bounds (2.4) is an optimal contract.

## <u>3. Optimal Contracts With the Endowment Observed Only by Agent *a*.</u>

A common contractual problem arises when one agent would like to insure another against an unobserved endowment shock. Many financial contracts such as partnership or equity arrangments fall in this category, as do many explicit insurance contracts as well. In our setting this means assuming that agent *b* does not observe the endowment x before enforcement occurs. As before, we will assume that the enforcement facility does uncover the value of x if invoked. To simplify matters we suppose that agent *b* is unable to verify the value of x independently.[11] This is merely an extreme version of giving the facility a cost advantage over private agents in gathering reliable information. The power of the facility to compel testimony might explain this informational advantage. Alternatively, we might interpret the facility as a audit-arbitration service. As part of the *ex ante* arrangement, agent *a* irrevocably empowers the facility to perform an audit if agent *b* so requests. When invoked, the service audits *a*, and then issues a binding decision regarding transfers.[12]

As before, a contract includes a strategy, $\pi^a$, for agent *a* specifying the probability $\pi^a(y|x)$ of paying the amount y in $Y(x)$. Given a realized payment, y, agent *b* chooses a probability of invoking the facility. For a given realized payment, y, a strategy for *b* is then probabilities, $\pi^b(s=1|y)$ and $\pi^b(s=0|y)$, that *b* sues or does not sue, respectively. These probabilities cannot depend on the realized endowment x, since it is not observed by *b*. The expected utility of agent *a* under the strategy $\pi^a(x)$, for a given realization x, is then

$$\int_Y \left\{\left[\int_\tau u^a(x-y-\tau^a)\pi(\tau|y,x)d\tau\right]\pi^b(s=1|y) \quad + \quad u^a(x-y)\pi^b(s=0|y)\right\}\pi^a(y|x)dy. \qquad (3.1)$$

For every realized value of x, agent *a* takes as given agent *b*'s strategy, $\pi^b$, and the transfer schedule, $\pi$, and selects $\pi^a(x)$ to maximize (3.1).

Upon receiving a payment, agent *b* must make some inference concerning the value of x, since the transfer executed by the facility could depend on x. Taking the strategy of agent *a* as given, agent *b* uses Bayes' Rule, wherever possible, to form beliefs about x. Let $\mu^b(y)$ denote agent *b*'s posterior beliefs about x, given the payment y, so that $\mu^b(x|y)$ is the probability $b$ assigns to the value x, given y. Then for all values of y that have positive probability

$$\mu^b(x|y) \quad = \quad \pi^a(y|x)f(x) \left[ \sum_x \pi^a(y|\hat{x})f(\hat{x}) \right]. \tag{3.2}$$

For values of y that have zero probability under the strategy $\pi^a$, we require that $\mu^b$ have support on the set $\{x \in \mathbf{X} | y \in \mathbf{Y}(x)\}$, that is, that $\mu^b(x|y)$ assign zero probability to values of x for which y is not feasible. Other than this we place no restrictions on $\mu^b$; we comment on this below.[13]

We can now describe the expected utility of agent $b$. Taking agent $a$'s strategy, $\pi^a$, and the transfer schedule, $\pi$, as given, the choice of strategy $\pi^b$ by $b$ provides expected utility of

$$\sum_x \left[ \int_\tau u^b(y-\tau^b)\pi(\tau|y,x)d\tau \right] \mu^b(x|y)\pi^b(s=1|y) \quad + \quad u^b(y)\pi^b(s=0|y). \tag{3.3}$$

This must hold for every y, even values of y which have zero probability, since agent $a$ has the opportunity to select those values. For any given y, agent $b$ takes $\pi$ and $\mu^b$ as given and selects $\pi^b$ to maximize (3.3).

An *equilibrium*, for a given feasible transfer schedule $\pi(\tau|y,x)$, is now a pair of strategies, $\pi^{a*}(y|x)$ and $\pi^{b*}(s|y)$, along with posterior beliefs $\mu^{b*}(x|y)$, that satisfy the following conditions:

(i) The strategy $\pi^{a*}(y|x)$ maximizes (3.1) for all $x \in \mathbf{X}$, taking as given the strategy $\pi^{b*}(s|y)$ and the transfer schedule $\pi(\tau|y,x)$.

(ii) The strategy $\pi^{b*}(s|y)$ maximizes (3.3) for all $y \in Y$, taking as given the beliefs $\mu^{b*}(x|y)$ and the transfer schedule $\pi(\tau|y,x)$.

(iii) The posterior beliefs $\mu^{b*}(x|y)$ satisfy Bayes' Rule, (3.2), for all values of y that occur with positive probability, but are otherwise arbitrary.

This definition of equilibrium follows closely Townsend's (1988) definition of a "Bayesian sequential Nash equilibrium." Each agent's strategy is a best-response, given the other agent's strategy, and beliefs must satisfy Bayes' Rule wherever possible. For payment quantities that agent *a* never selects in equilibrium, Bayes' Rule is undefined and the approach adopted here is that they are arbitrary. A *contract* consists of the transfer schedule, $\pi$, along with the equilibrium strategies $\pi^{a*}$ and $\pi^{b*}$. If for some equilibrium contract $(\pi,\pi^{a*},\pi^{b*})$ no other equilibrium contract provides *ex ante* expected utility that is no smaller for any agent and strictly larger for at least one agent, then $(\pi,\pi^{a*},\pi^{b*})$ is an *optimal contract*.[14]

This definition of equilibrium deserves some comment, since as Kreps and Wilson (1982), Myerson (1986) and many others have emphasized, "off-equilibrium" beliefs can play a critical role in determining the set of equilibria in sequential games. The literature on general sequential games under uncertainty takes the payoffs agents receive as exogenously specified. In the sequential game that arises here, payoffs, both on- and off-equilibrium, are in part determined by the players in advance. This makes off-equilibrium beliefs considerably less critical, as we shall see. In fact, it is quite straightforward to take any equilibrium satisfying the above definition, and show that there is a corresponding *sequential equilibrium* as defined by Kreps and Wilson (1982). A reasonable conjecture is that equilibria by the above definition would satisfy many other recently proposed "refinements" of Bayesian sequential Nash equilibrium.

The existance of an optimal contract is assured by the analysis of the appendix, where it is shown that finding an optimal contract can be formulated as a linear programming problem.

Our first result is that optimal transfer schedules are deterministic. Thus contracts will not request that the facility perform some unnecesary randomization to determine the punishment or reward to either agent. The logic behind this result is that any random conditional transfer can be feasibly replaced with a certainty equivalent level of consumption. Such a replacement does not affect the sequential rationality constraints because there are separate transfer schedules for on- and off-equilibrium actions. Therefore, we have the following result.

Proposition 2: The optimal contract never requires randomness in the conditional transfer schedule for the facility: for any optimal contract there is an equilibrium contract that provides identical *ex ante* expected utlity to both agents and in which the distribution $\pi(\tau|y,x)$ is deterministic.

Proof: A typical term in the braces in (3.1) is of the form:

$$\left[\int_{\tau} u^a(c^a)\pi(\tau|\hat{y},x)d\tau\right]\pi^{b*}(s=1|\hat{y})\pi^{a*}(y|x) \quad + \quad u^a(x-\hat{y})\pi^{b*}(s=0|y)\pi^{a*}(y|x) \qquad (3.4)$$

The expression that agent *b* maximizes is

$$\sum_{x}\left[\int_{\tau} u^b(c^b)\pi(\tau|y,x)d\tau\right]\mu^{b*}(x|y)\pi^{b}(s=1|y) \quad + \quad u^b(y)\pi^b(s=0|y). \qquad (3.5)$$

For some proposed equilibrium transfer schedule, $\pi(\tau|y,x)$, construct deterministic transfer schedules, $\tau^{a*}(y,x)$, and $\tau^{b*}(y,x)$, that satisfy:

$$u^a(x-y-\tau^{a*}(y,x)) \quad = \quad \int_{\tau} u^a(x-y-\tau^a)\pi(\tau|y,x), \text{ and}$$

$$u^b(y-\tau^{b*}(y,x)) \quad = \quad \int_{\tau} u^b(y-\tau^b)\pi(\tau|y,x),$$

for all y and x. Because $u^a$ and $u^b$ are both concave, and because the original transfer schedule is feasible, we know that $\tau^{a*}(y,x)$ and $\tau^{b*}(y,x)$ are feasible. It can be directly verified that the expressions (3.4) and (3.5) are unaffected, and thus incentive constraints are unaffected. So if $(\pi,\pi^{a*},\pi^{b*})$ is an equilibrium contract, then $(\tau^*,\pi^{a*},\pi^{b*})$ is an equilibrium contract, where $\tau^*\equiv(\tau^{a*},\tau^{b*})$. Furthermore, the objective function is unaffected. If $u^a$ and $u^b$ are linear, and $(\pi,\pi^{a*},\pi^{b*})$ is an optimal contract then so is $(\tau^*,\pi^{a*},\pi^{b*})$. Q.E.D.

Although there are always optimal deterministic transfer schedules, there are two remaining sources of randomness in equilibrium allocations--the strategies of agents *a* and *b*. In general, a mixed strategy for agent *b* will be useful because the set of $b$'s pure strategies is nonconvex. Since the set of pure strategies for agent $a$ is convex, one might suspect mixed strategies for $a$ are unnecessary. However, examples in Myerson (1986) and Townsend (1988) show that in sequential games with imperfect information it might be optimal to "scramble" information by randomization. In our environment having agent $a$ play mixed strategies scrambles the information set of agent $b$, so it would seem impossible to eliminate mixed strategies for agent $a$ in general.

Transfer schedules are easy to characterize for "off-equilibrium" values of the payment. More precisely, we can choose off-equilibrium values for the transfer schedule without loss of generality. Define A(x) as $\{y|\pi^{a^*}(y|x)>0\}$, the set of payments that occur with positive probability when the realized endowment is x, and define A as $\{y|\pi^{a^*}(y|x)>0, x\in\mathbf{X}\}$, the set of payments that occur with positive probability for some realized endowment. Then writing out the incentive constraints for agent $a$:

$$\begin{aligned} & u^a(x-y-\tau^a(y,x))\pi^{b^*}(s=1|y) \quad + \quad u^a(x-y)\pi^{b^*}(s=0|y) \\ \geq \quad & u^a(x-y-\tau^a(\hat{y},x))\pi^{b^*}(s=1|\hat{y}) \quad + \quad u^a(x-\hat{y})\pi^{b^*}(s=0|\hat{y}) \\ & \qquad \text{for all } y\in A(x), \text{ for all } \hat{y}, \text{ and for all } x\in\mathbf{X}. \end{aligned} \tag{3.6}$$

For some fixed x, consider a value of $\hat{y}$ that never occurs in equilibrium, that is $\hat{y}\notin A$. The transfer from agent $a$, $\tau^a(x,\hat{y})$, appears only in the first term on the right side of (3.6) for various values of $y\in A(x)$. For any given contract we can make $\tau^a(\hat{y},x)$ as large as possible without violating (3.6), and without affecting equilibrium payoffs. Agent $b$'s strategy for $\hat{y}$ need not be affected since $\tau^b(\hat{y},x)$ is still feasible. Thus, without loss of generality we can presume that $\tau^a(\hat{y},x) = x - \hat{y}$: if agent $a$ is ever caught making an off-equilibrium action, a's consumption is set to zero.

Another simple property of optimal contracts is also immediately available. Consider agent $b$'s sequential rationality constraints for any given value of y, and rewrite them as follows:

$$\pi^{b*}(s=1|y) > 0 \quad \Rightarrow \quad \sum_x u^b(y-\tau^b(y,x))\mu^{b*}(x|y) \geq u^b(y);$$

$$\pi^{b*}(s=1|y) < 1 \quad \Rightarrow \quad \sum_x u^b(y-\tau^b(y,x))\mu^{b*}(x|y) \leq u^b(y);$$

$$\sum_x u^b(y-\tau^b(y,x))\mu^{b*}(x|y) > u^b(y) \quad \Rightarrow \quad \pi^{b*}(s=1|y) = 1; \text{ and,}$$

$$\sum_x u^b(y-\tau^b(y,x))\mu^{b*}(x|y) < u^b(y) \quad \Rightarrow \quad \pi^{b*}(s=1|y) = 0. \tag{3.7}$$

If the equilibrium strategy of agent $b$ is to invoke the facility with positive probability for some given payment, y, then it must be in $b$'s interest to do so. Because $u^b$ is concave, this implies that the average value of $\tau^b$ across the relevant values of x is negative. In other words, if the facility is invoked in equilibrium agent $b$ must receive a net transfer ($-\tau^b \geq 0$) that is nonnegative on average. Together with the feasibility condition on net transfers, this implies that the average value of $\tau^a$, the transfer from agent $a$, is greater than or equal to $\gamma$. This condition is even simpler if a given payment y is made for only one realization of x. In this case agent $b$ is fully informed--$\mu^{b*}(x|y) = 1$--and sequential rationality along with feasibility implies that $-\tau^b \geq 0$, and $\tau^a \geq \gamma$. Thus, we have the following result.

> <u>Proposition 3</u>: Suppose that agent *b* is fully informed for some payment, y ($\mu^{b*}(x|y) = 1$ for some x). If the probability of invoking the facility is strictly positive ($\pi^{b*}(s=1|y)>0$), then the consumption of agent *a* is <u>smaller</u> when the facility is invoked than when it is not invoked by at least the amount $\gamma$:
>
> $$\begin{aligned} c^a(s=1,y,x) &= x - y - \tau^a(y,x) \\ &\leq x - y - \gamma = c^a(s=0,y,x) - \gamma. \end{aligned} \qquad (3.8)$$
>
> The feasibility constraint is met with equality ($\tau^a + \tau^b = \gamma$) and agent *b* receives a nonegative transfer ($-\tau^b \geq 0$). If the probability of invoking the facility is strictly between zero and one, then (3.8) holds with equality, $\tau^a(y,x) = \gamma$, and $-\tau^b = 0$.

Agent *a* is <u>worse off</u> when *b* invokes the facility; *a* pays at least the cost of the facility. If *b* invokes with a probability less than one, agent *a* pays <u>exactly</u> the cost of the facility.

Contracts that involve only pure strategies are of interest for two reasons. First, it might turn out that the optimal contract is a pure strategy contract. In this case we would like to know something about the characteristics of such a contract, even if we are unable to find conditions under which they occur. Second, pure strategy contracts are quite simple and easy to characterize. Even though a restriction to pure strategies can be fairly severe in terms of foregone welfare, as Townsend (1988) shows, there are settings in which certain qualitative characteristics of equilibria are unaltered and tractability is dramatically improved [see, for example, Williamson (1987), or Lacker (1989)].

A *pure strategy contract* is one in which both $\pi^a$ and $\pi^b$ are deterministic. There is a transfer schedule, $y(x)$, and an enforcement region, $S \subset Y$, over which $\pi^b(s=1|y) = 1$. Restricting attention to pure strategy contracts then, we have one main result.

<u>Proposition 4</u>: Suppose only pure strategy contracts are allowed, agent *a* is strictly risk averse, and agent *b* is risk neutral. Then the optimal contract is a debt contract. Specifically, for some $x' \in X$, and some $\bar{y} \in Y$:

(a) $S = [0, \bar{y})$;

(b) $y(x) = \bar{y}$ for all $x \geq x'$, and $y(x) < \bar{y}$ for all $x < x'$;

(c) $c^a(x) = x - y(x) - \tau^a(y(x),x) = \bar{c}^a$, for all $x < x'$, where $\bar{c}^a$ is a constant satisfying $\bar{c}^a \leq \underline{x} - \gamma$, and $c^a(x) = x - \bar{y}$ for all $x \geq x'$;

(d) if $y < \bar{y}$ and $x < x'$, then $\tau^a(y,x) = x - y - \bar{c}^a$, and $\tau^b(y,x) = -(x - y - \bar{c}^a - \gamma)$;

(e) without loss of generality, for $x \geq x'$, if $y < \bar{y}$ then $\tau^a(y,x) = x - y$, and $\tau^b(y,x) = -(x - y - \gamma)$, and if $y \geq \bar{y}$ then $\tau^a(y,x) = -\bar{y} + \gamma$ and $\tau^b = \bar{y}$.

<u>Proof</u>: Claim: If $y(x) \in S$, then $\tau^a(y,x) + \tau^b(y,x) = \gamma$. Suppose not and $\tau^a(y,x) + \tau^b(y,x) > \gamma$. Then decrease $\tau^a(y,x)$ by a small amount. This affects no rationality conditions and makes agent *a* strictly better off, so the original contract could not have been optimal, proving the claim.

By a familiar argument, if $y(x) \in \mathbf{S}'$ then $y(x) = \bar{y}$, where $\bar{y} \le y'$ for all $y' \in \mathbf{S}'$, and where $\mathbf{S}'$ is the complement of $\mathbf{S}$ in $\mathbf{Y}$. For $y < \bar{y}$, $y \neq y(x)$, set $\tau^a(y,x) = x - y$, and $\tau^b(y,x) = -(x - y - \gamma)$. Feasibility is easily checked. If the contract was sequentially rational, it is still sequentially rational after this substitution. Since only off-equilibrium payoffs are affected, this substitution is without loss of generality. Similarly, for $y > \bar{y}$, $y \in \mathbf{S}'$, set $\tau^a(y,x) = -\bar{y} + \gamma$, and $\tau^b(y,x) = \bar{y}$, without loss of generality, establishing (e).

Define $\mathbf{S_x} = \{x | y(x) \in \mathbf{S}\}$, and $\mathbf{S'_x}$ as the complement of $\mathbf{S_x}$ in $\mathbf{X}$. Define $z(x) = y(x) + \tau^a(y(x),x)$. Sequentially rational implies that $\gamma \le z(x) \le \mathrm{MIN}[x,\bar{y}]$ for $x \in \mathbf{S'_x}$. Consider an arbitrary $z(x)$ for $x \in \mathbf{S_x}$, that satisfies $\gamma \le z(x) \le \mathrm{MIN}[x,\bar{y}]$ for each $x \in \mathbf{S_x}$. Then $c^a(x) = x - z(x) \ge 0$, and $c^b(x) = y(x) - \tau^b(y(x),x) = z(x) - \gamma \ge 0$, so the implied consumptions are feasible. Since $z(x) \ge \gamma$, we can always select $y(x)$ and $\tau^a(y(x),x)$ so that $y(x) \ge 0$ and $-\tau^b(y(x),x) = \tau^a(y(x),x) - \gamma \ge 0$, implying that (ii) is satisfied. Since $z(x) \le \bar{y}$ for all $x \in \mathbf{S_x}$, we know that $x - z(x) \ge x - \bar{y}$ for all $x \in \mathbf{S_x}$ such that $x \ge \bar{y}$. Therefore, (i) is satisfied. We have proven that finding $y(x)$, $\mathbf{S}$, $\tau^a(y,x)$, and $\tau^b(y,x))$ that satisfy sequential rationality constraints is equivalent to finding $\mathbf{S_x}$, $z(x)$ on $\mathbf{S_x}$, and $\bar{y}$ that satisfy

$$\gamma \le z(x) \le \mathrm{MIN}[x,\bar{y}], \text{ and} \tag{3.9}$$

$$0 \le \bar{y} \le \mathrm{MIN}\ \mathbf{S'_x}. \tag{3.10}$$

We can now find an optimal contract by choosing $\mathbf{S_x}$, $z(x)$, and $\bar{y}$ to

$$\text{MAX} \sum_{x \in S_x} u^a(x-z(x))f(x) + \sum_{x \in S'_x} u^a(x-\bar{y})f(x)$$

$$\text{s.t.} \sum_{x \in S_x} u^b(z(x)-\gamma)f(x) + \sum_{x \in S'_x} u^b(\bar{y})f(x) \geq K. \tag{3.11}$$

$$(3.9) \text{ and } (3.10),$$

for some nonegative parameter K. This problem is identical to Townsend's (1979) Problem 3.1, in the special case that his $y_1$, the endowment of our agent *b*, is zero. Consequently, his Lemma 3.1 and Proposition 3.2 apply directly, implying (a) and (c). Q.E.D.

Beyond these simple properties, general results are difficult to obtain. In the Appendix a linear programming problem is derived that yields optimal contracts as solutions. The derivation relies on the approach of Prescott and Townsend (1984), formulating moral hazard incentive compatibility conditions as linear constraints in an appropriate commodity space. In principle one could compute optimal contracts for various example environments, as in Townsend (1988). Unfortunately, the dimensions of this problem are far larger than the problems computed by Townsend, and exceed currently available computing capacity.

## 4. A Comparison With Costly Auditing

The costly enforcement environment we have described closely parallels the "costly state verification" or "costly auditing" models of Townsend (1979, 1988), Mookherjee and Png (1987), and Border and Sobel (1987), among others.[15] Invoking the facility in our environment automatically reveals the state and triggers state-contingent transfers, and thus closely resembles an audit. In this section we compare and contrast the two models.

The central difference is that costly verification is typically modeled as a Principal-Agent problem or as a revelation game. The person with the random endowment--agent *a* in our environment--is an "Agent" and the other--agent *b*--is a "Principal." The Agent receives a realization of the endowment, and then sends a message to the Principal concerning its value. Depending on the message, an allocation "results" according to some allocation rule. An allocation is a specification of a payment from the Agent, a probability of auditing, and a further payment (or fine). The mapping from messages to allocations is the mechanism, or contract. The Revelation Principal implies that we can restrict attention to "direct revelation mechanisms" that satisfy the incentive constraints that agent *a* chooses to send a truthful message for all realizations. The implementation of the allocation rule or contract depends on the ability of the Principal to precommit to following the rule. In the costly enforcement model however, precommitment is only available to the facility, and allocations are determined by agent's actions rather than the messages they send. Thus the incentive constraints take the form appropriate to moral hazard environments, as in Myerson (1986) or Prescott and Townsend's (1984) economy $E_3$. The central differences between costly enforcement and costly verification all stem from this distinction. But note that the definition of optimal contracts in the costly enforcement environment here is fully consistent with the principles of mechanism design, as the derivation of the programming problem in the appendix shows.

The incentive constraints on agent *a* are very different in the two models. In the costly verification setup the allocation, including the payment made by agent *a*, is "delivered" by the Principal. Agent *a* merely chooses between alternative announcements, each of which results in a particular allocation. The incentive constraint is that agent *a* (weakly) prefer the allocation resulting from announcing that the true state was x to the allocation that *a* would receive if *a* pretended that the endowment was $\hat{x}$. In our model, this corresponds to agent *a* (weakly) preferring the allocation resulting from strategy $\pi^{a*}(y|x)$ to the allocation resulting from strategy $\pi^{a*}(y|\hat{x})$ for some $\hat{x}$. This constraint can be written

$$\int_{Y} \left\{ \left[ \int_{\tau} u^a(x-y-\tau^a)\pi(\tau|y,x)d\tau \right] \pi^b(s=1|y) \quad + \quad u^a(x-y)\pi^b(s=0|y) \right\} \pi^a(y|x)dy$$

$$\geq \quad \int_{Y} \left\{ \left[ \int_{\tau} u^a(x-y-\tau^a)\pi(\tau|y,x)d\tau \right] \pi^b(s=1|y) \quad + \quad u^a(x-y)\pi^b(s=0|y) \right\} \pi^a(y|\hat{x})dy,$$

$$\text{for all } x\in\mathbf{X}, \text{ and for all } \hat{x}\in\mathbf{X},\ \hat{x}\neq x. \tag{4.1}$$

These closely parallel the incentive constraints in Townsend (1988). These are the constraints that would apply if, in the environment of this paper, agent *a* were to announce a state, and then the payment is made <u>for</u> agent *a*, say by some third party. Here $\pi^{b*}(s=1|y)$ plays the role of the Principal's conditional probability of audit, and $\tau^a(y,x)$ plays the role of the post-audit transfer. Clearly, the constraints (4.1) are implied by the sequential rationality equilibrium condition (i), since the strategy $\pi^{a*}(\cdot|\hat{x})$ could have been chosen at x. But it should also be clear that the latter constraints are considerable stronger than (4.1). For some arbitrary contract, consider two payments, $y_1$ and $y_2$, both in the support of $\pi^{a*}(\cdot|x)$. The equilibrium condition (i) implies that the value of the term in braces on the left side of (4.1) must be identical for both $y_1$ and $y_2$. For example, if $y_1 < y_2$, then either the transfers $\tau^a(y_1,x)$ and $\tau^a(y_2,x)$ are different, or the enforcement probabilities $\pi^{b*}(s=1|y_1)$ and $\pi^{b*}(s=1|y_1)$ are different, or both. But (4.1) does not necessarily ensure that this condition holds.

Townsend (1988) allows for consumption lotteries that are contingent on the message sent by agent *a* and on whether or not verification takes place. Allowing mixed strategies for agent *a* effectively allows for random consumption for agent *a* contingent on not being verified or, in our environment, not having the facility invoked. But when verification takes place in Townsend's setup, the agent receives a lottery that depends on both the announced and the actual state. In the costly enforcement model the transfer schedule depends only on the payment y and the state x, not on the announced value of x. For example, consider a value of y that is in the support of both $\pi^{a*}(\cdot|x)$ and $\pi^{a*}(\cdot|\hat{x})$. The transfer $\tau^a(y,x)$ does not depend on $\hat{x}$, and so it cannot be used to make consumption zero when verification turns up a counterfactual announcement. This is why Proposition 2 holds here and does not in Townsend's setup. To get the same result in Townsend's model, imagine that agent *a* announces $\hat{x}$, y is then determined according to $\pi^{a*}(y|\hat{x})$, and then if verification takes place, the consumption lottery can depend only on y and x, the true state. Thus, the costly enforcement model implies further constraints on allocations, beyond those of the costly verification model.

Perhaps the most striking difference is that agent *b*'s action, for a given payment, must be sequentially rational. In the costly verification setup the Principal receives a report on the state, a report that is always truthful in equilibrium. Even if two different states result in the same nonverification payment, the Principal knows the state, and can apply different audit probabilities. In the costly enforcement model, the strategy of agent *b* must respect agent *b*'s information sets. In particular, two states that give rise to the same payment must elicit the same strategy for *b*. Thus the constraints in (3.7) involve *b*'s beliefs.

Furthermore, the sequential rationality of agent *b*'s action implies that the contract must motivate agent *b*, and thus must penalize agent *a* when the facility is invoked, as shown in Proposition 3 above. This is exactly the opposite of the results of Border and Sobel (1987), Mookherjee and Png (1989), and Townsend (1988) that for any given state the agent has *higher consumption* when audited than when auditing does not occur, and the difference is due precisely to the way in which commitment is modelled. By treating the person invoking the facility as a Principal, incentives to invoke need not be provided. This points up a potential pitfall in applying the Principal-Agent model, in which commitment capabilities are asymetric, to private bilateral contractual arrangements.[16]

In independent work Robert Moore (1989) has explored a costly auditing model with a sequential rationality constraint imposed on the auditing agent--agent *b* in the present paper. Moore's setup is essentially an extension of the model that appears in the appendix of Bernanke and Gertler (1989). Both agents are risk neutral, and stochastic auditing is allowed, but unlike Townsend (1988) consumption lotteries are not allowed. Thus, Moore's results can be viewed as applying to a special case of the model of the present paper, in which agent *a* is restricted to pure strategies.[17]

5. Variable Enforcement Costs: An Example

In this section we allow the costs of the facility to vary with the realized value of x. In our basic setup let us assume that the payment, y, is public and costlessly known to the facility. Now suppose the facility cost depends on $x - y$, the amount remaining in the possession of agent *a*, according to the function $\Gamma(x-y)$. For example, auditing inventories, assessing the value of collateral, or compiling catalogues of dispersed assets might easily incur costs that vary positively with the verified quantity. In particular, suppose that verification costs are zero when $x - y = 0$, and strictly positive when $x - y > 0$. This amounts to assuming that it is costless to verify that nothing remains in the possession of agent *a*, but costly to verify the quantity of any positive amount. For example, the act of transfering the last unit of the good to the lender might inextricably reveal itself, or it might be costless to falsify any nonzero quantity of the good but prohibitively costly to falsify the absence of the good [see Lacker and Weinberg (1989)].

Under these assumptions on $\Gamma$, and under the additional assumption that both agents are risk neutral, the optimal contract is a debt contract of the type described by Williamson (1987): $y^*(x) = \mathrm{MIN}[x,\bar{y}]$, where $\bar{y}$ is a constant. Enforcement occurs with probability one for all $y < \bar{y}$. In a "Williamson debt contract" consumption of the "borrower" is zero over the contingent region; under a "Townsend debt contract" consumption is in general positive. Note that whenever enforcement occurs in equilibrium, $y^*(x) = x$ and $\Gamma(x-y^*(x)) =$

0. Thus enforcement costs are never incurred. The optimality of this contract relies heavily on the risk neutrality of both agents: under this specification maximizing the weighted average of expected utilities is equivalent to minimizing the expected value of enforcement costs. Note that, unlike Williamson's setup, mixed strategies are allowed; debt contracts are not generally optimal in his setup when random verification is allowed. Note also that "credit rationing" will not arise here, because the nonmonotonicity of the utility frontier in Williamson's setup is absent here.

Two important caveats are worth noting. First, the example is not robust to perturbations in the cost function, $\Gamma$. If there is any positive cost of enforcement for $x - y = 0$, debt contracts are not necessarily optimal. Second, if agents have the opportunity to make transfers and invoke the facility again after already invoking the facility once, then any contract can be supported: simply have agent *a* give everything to agent *b*, have *b* invoke, costlessly verifying that $x - y = 0$. Now that the state is common knowledge, have *b* give some quantity back to agent *a* according to any arbitrary schedule, supported by the threat of enforcement as in Section 2.

Nevertheless, some insight emerges. A debt contract might be attractive because it is easy to prove penury to your creditor but hard to prove your exact status for more favorable outcomes. One thinks of costly but widely observed acts associated with "bankruptcy": liquidating assets, closing facilities--"turning out one's pockets" as it were. Further investigation might be warranted into models in which an enforcement facility interacts with private agent's ability to manipulate appearances as in Lacker and Weinberg (1989).

## Appendix

In this appendix we derive a programming problem for the determination of optimal contracts for the environment in section 3, where the endowment of agent *a* ia not observed by agent *b*. To aid clarity, and to make the commodity space finite, we assume that the set Y of possible payments is finite, as if the consumption good comes in indivisible units. Thus $Y(x) = Y \cap [0,x]$. To begin we can simplify notation by defining a measure $\pi(s,y,x)$ directly over consumption; $\pi(c|s,y,x)$ gives the probability of $c=(c^a,c^b)$, given s, y, and x. To be consistent with the fact that the facility cannot perform transfers if it has not been invoked, we require

$$\pi(c|s=0,y,x) = \begin{cases} 1, & \text{for } c^a = x - y,\ c^b = y, \\ 0, & \text{otherwise} \end{cases}$$

$$\pi(c|s=1,y,x) = \begin{cases} \pi(\tau|y,x), & \text{for } c^a = x - y - \tau^a,\ c^b = y - \tau^b, \\ 0, & \text{otherwise} \end{cases} \tag{A.1}$$

Using this notation, we can rewrite equilibrium conditions compactly. It is straightforward to show that a contract satisfies (i) and (ii) in Section 3 if and only if

$$\sum_s \sum_c u^a(c^a)\pi(c|s,y,x)\pi^{b*}(s|y)\pi^{a*}(y|x)$$

$$\geq \quad \sum_s \sum_c u^a(c^a)\pi(c|s,\hat{y},x)\pi^{b*}(s|\hat{y})\pi^{a*}(y|x)$$

for all $x \in X$, and for all $y,\hat{y} \in Y(x)$, and (A.2)

$$\sum_x \sum_c u^b(c^b)\pi(c|s,y,x)\mu^{b*}(x|y)\pi^{b*}(s|y)$$

$$\geq \quad \sum_x \sum_c u^b(c^b)\pi(c|\hat{s},y,x)\mu^{b*}(x|y)\pi^{b*}(s|y)$$

for all $y \in Y$, and for all $s,\hat{s}$. (A.3)

[See Prescott and Townsend (1984) p. 28.]

For values of y that occur with positive probability a simple substitution eliminates $\mu^{b*}(y)$ from the problem, while for off-equilibrium values of y, beliefs are arbitrary and can be ignored. For y in A, $\mu^{b*}(y)$ is determined by Bayes' Rule, (3.2). Note that the denominator of the right side of (3.2) depends only on y. Therefore we can multiply both sides of (A.3) by this denominator, and then use (3.2) to substitute the numerator of the right side of (3.2) into (A.3). This gives us a version of (A.3) that holds for all "on-equilibrium" values of y:

$$\sum_x \sum_c u^b(c^b)\pi(c|s,y,x)\pi^{a*}(y|x)f(x)\pi^{b*}(s|y)$$

$$\geq \quad \sum_x \sum_c u^b(c^b)\pi(c|\hat{s},y,x)\pi^{a*}(y|x)f(x)\pi^{b*}(s|y)$$

for all $y \in A$, and for all $s,\hat{s}$. (A.4)

Agent $b$'s beliefs have been eliminated from the problem for on-equilibrium values of y.

It remains to consider the beliefs of agent $b$ for off-equilbrium values of y. For arbitrary equilibrium strategies, and for any given y not in A, we can arbitrariliy select transfers and beliefs as long as

$$\pi^{b*}(s=1|y) = 0, \qquad \text{for all } y > \bar{x} - \gamma. \qquad (A.5)$$

The final step in the transition to a programming problem is to rewrite the constraints in terms of allocations--joint distributions over consumptions and actions--rather than strategies. An allocation for our economy is a joint probability distribution over consumption, $c = (c^a, c^b)$, and the two actions, invoking the facility, s, and making the payment, y, conditional on the realized endowment. The form of the constraints (A.2), (A.4), and (A.5) suggest defining

$$\pi^*(c,s,y|\hat{s},\hat{y},x) \equiv \pi(c|\hat{s},\hat{y},x)\pi^{b*}(s|\hat{y})\pi^{a*}(y|x). \qquad (A.6)$$

The probabilities $\pi^*(c,s,y|\hat{s},\hat{y},x)$ are the choice variables in our programming problem. Our problem is to maximize the weighted average of the *ex ante* expected utilities of the two agents, for some Pareto-weights, $\lambda^a$, $\lambda^b$; $\lambda^a + \lambda^b = 1$, $\lambda^a, \lambda^b > 0$. Formally, (P1) Programming Problem 1: Maximize, by choice of probabilities $\pi^*(c,s,y|\hat{s},\hat{y},x)$,

$$\lambda^a \sum_x \sum_y \sum_s \sum_c u^a(c^a)\pi^*(c,s,y|s,y,x)f(x)$$

$$+ \quad \lambda^b \sum_x \sum_y \sum_s \sum_c u^b(c^b)\pi^*(c,s,y|s,y,x)f(x)$$

subject to the constraints

$$\sum_c \sum_s \sum_y \pi^*(c,s,y|\hat{s},\hat{y},x) \quad = \quad 1, \qquad \text{for all } (\hat{s},\hat{y},x), \qquad (A.7)$$

$$\pi^*(c,s,y|\hat{s},\hat{y},x) \quad \geq \quad 0, \qquad \text{for all } (c,s,y,\hat{s},\hat{y},x). \qquad (A.8)$$

$$\sum_c \sum_s u^a(c^a)\pi^*(c,s,y|s,y,x) \quad \geq \quad \sum_c \sum_s u^a(c^a)\pi^*(c,s,y|s,\hat{y},x),$$

$$\text{for all } x \in X, \text{ and for all } y,\hat{y} \in Y(x) \qquad (A.9)$$

$$\sum_x \sum_c u^b(c^b)\pi^*(c,s,y|s,y,x)f(x) \quad \geq \quad \sum_x \sum_c u^b(c^b)\pi^*(c,s,y|\hat{s},y,x)f(x)$$

$$\text{for all } y \in Y, \text{ and for all } s,\hat{s}. \qquad (A.10)$$

$$\sum_c \sum_y \pi^*(c,s,y|\hat{s},\hat{y},x) \quad = \quad 0, \quad \text{for all } \hat{y} \in Y, \text{ s.t. } \hat{y} > \overline{x} - \gamma. \qquad (A.11)$$

$$\sum_c \sum_s \pi^*(c,s,y|\hat{s},\hat{y},x) \quad = \quad \sum_c \sum_s \pi^*(c,s,y|\hat{\hat{s}},\hat{\hat{y}},x)$$

$$\text{for all } x, \in X, \text{ and for all } \hat{s},\hat{\hat{s}},\hat{y},\hat{\hat{y}}. \qquad (A.12)$$

$$\sum_{c}\sum_{y} \pi^*(c,s,y|\hat{s},\hat{y},x) = \sum_{c}\sum_{y} \pi^*(c,s,y|\hat{s},\hat{y},\hat{x})$$

$$\text{for all } s,\ \hat{s},\ \hat{\hat{s}},\ \hat{y},\ x,\ \hat{x}. \tag{A.13}$$

$$\sum_{s}\sum_{y} \pi^*(c^a{=}x-\hat{y},c^b{=}\hat{y},y,s|\hat{s}{=}0,\hat{y},x) = 1,$$

$$\text{for all } x\in\mathbf{X}. \tag{A.14}$$

Constraints (A.7) and (A.8) ensure that the choice variables are actually probabilities. Constraints (A.9) and (A.10) are the sequential rationality constraints for agents *a* and *b* respectively. Constraints (A.11) ensure that we can drop consideration of off-equilibrium sequential rationality for agent *b*. Constraints (A.12) allow us to recover $\pi^{a*}$, ensuring that $\pi^*(y|\hat{s},\hat{y},x)$ is independent of $\hat{s}$ and $\hat{y}$. Similarly, (A.13) allows us to recover $\pi^{b*}(y)$, ensuring that $\pi^*(s|\hat{s},\hat{y},x)$ is independent of $\hat{s}$ and x. Finally, the constraint (A.14) ensures that the probabilities are consistent with (A.1).

The problem (P1) has an objective function and constraint functions that are linear in the choice variables, and so a solution exists and, in principle, a numerical solution can be computed for given specifications of the primitives of the economy. Unfortunately, this problem is large. Consider a simple example patterned on a costly verification example of Townsend's (1988, p. 430): $\mathbf{X} = \{4,6,8,10\}$, $\gamma = 2$, and the good comes in units of 0.1. Then the commodity space has a dimension of 3,387,136, and there are 9,379,124 constraint equations.

**NOTES**

1. The use of courts to enforce contracts seems inextricably intertwined with commercial activity as we know it. In fact, courts were ubiquitous even at the emergence of commercial activity in Western Europe a millenium ago. See, for example, Gross (1906), especially p. 244.

2. Casual observation suggests that contract enforcement via the judicial system is quite costly, both in absolute amount and relative to rewards. The average annual cost of the U. S. civil litigation system was estimated at between $10 billion and $27 billion for the year 1984 (Aschkenasy 1986).

3. Lane (1958) argues that "violence-controlling enterprise" might be a natural monopoly.

4. Moore (1989) obtains the same result in a costly auditing model by imposing sequential rationality on agent *b*, but his model differs from the present paper in that mixed strategies for agent *a* are not allowed and both agents are risk neutral.

5. In Williamson (1987) debt contracts are not necessarily optimal if one allows for random verification.

6. Note that the debt contracts in Diamond (1984) rely on an enforcement facility that is unable to observe the value of the endowment.

7. Admittedly, this merely removes the issue of commitment one step, since we assume here that the facility has the ability to precommit to follow *ex ante* instructions. Nonetheless, positing a disinterested facility seems a necessary prelude to allowing bribery or investigating the motivation of facility personnel.

8. We neglect the possibility of uninformed enforcement--the possibility that the facility could enforce a contract, perhaps at a lower cost, without discovering the value of x.

9. The costs incurred by agents *a* and *b* are assumed to be zero, but this is without loss of generality; any nonzero costs would appear in equilibrium conditions added to $\gamma$ where appropriate.

10. To see this note that if $\pi^a(y'|x) > 0$ and $\pi^b(s=1|y',x)$, then (i) requires enforcement for all $y < y'$ for which $\pi^a(y|x) > 0$.

11. If agent *b* can independently verify x at a cost before deciding whether to invoke the facility, and transfer schedules can be made contingent on whether or not *b* has verified, then in optimal contracts *b* will in

any given state either verify or invoke the facility, but not both. The intuition is that verification places the agents are in a perfect information environment in which any schedule satisfying (2.4) can be supported with no enforcement in equilibrium. And if agent $b$ will invoke in some state, then private verification before hand will serve no purpose.

12. Of course, under this interpretation a legal system must exist in the background to provide commitment capability (and immunity to bribes) for the service.

13. We are assuming that agent $a$ sends no "messages" to agent $b$.

14. There is some indeterminacy concerning the content of the explicit *ex ante* document. It could consist of the full specification $(\pi, \pi^{a*}, \pi^{b*})$. Alternatively, it could just consist of the transfer schedule, $\pi$, without any specification of the desired equilibrium strategies, since these are induced by $\pi$ (although not necessarily uniquely). More realistically, it could merely include the equilibrium strategies $\pi^{a*}$ and $\pi^{b*}$. For this to suffice, the facility would have to commit itself to a mapping from equilibrium strategies $(\pi^{a*}, \pi^{b*})$ to transfer schedules $\pi$. But if the strategies $(\pi^{a*}, \pi^{b*})$ constitute an equilibrium, then there must exist a corresponding transfer schedule. As long as the facility's mapping selects a transfer schedule for which the strategies constitute an equilibrium, allocations are unaffected. In fact, a general principle here is that as long as the facility's mapping from documents to transfer schedules is known, documents need not specify contingencies exhaustively; the statutes and precedents pertaining to the contract substitute for the (potentially costly) effort of lengthening the document. Thus evidence that actual contracts are, by themselves, incomplete is not evidence that equilibrium allocations have been incompletely specified. John Weinberg pointed this out to me.

15. Williamson (1987) and Bernanke and Gertler (1989), and Moore (1989) study risk neutral versions. Baiman and Demski (1980) and Dye (1986) study versions with private actions rather than private endowments. Other applications include Gale and Hellwig (1985), Chang (1987), and Reinganum and Wilde (1985).

16. The precommitment ability of the agent identified as the Principal might arise from the agent's need to maintain a reputation in a series of repeated encounters, but this deserves explicit treatment. Melamad and Mookherjee (1989) allow the Principal to delegate responsibility for audits and commit instead to a compensation schedule for the auditor. They show that the full-commitment contract can be acheived. In the present paper the Principal has no ability to precommit (and thus is not really a Principal after all).

## REFERENCES


Aschkenasy, Janet, 1986, AIG/NYU Study: Tort Costs, National Underwriter, Property and Casualty Insurance Edition 90(18), 4.

Baiman, Stanley, and Joel Demski, 1980, Economically Optimal Performance Evaluation and Control Systems, Journal of Accounting Research 18, supplement, 184-220.

Bernanke, Ben, and Mark Gertler, 1989, Agency Costs, Net Worth, and Business Fluctuations, American Economic Review 79 (March), 14-31.

Border, Kim C., and Joel Sobel, 1987, Samurai Accountant: A Theory of Auditing and Plunder, Review of Economic Studies LIV, 525-540.

Chang, Chun, 1987, Capital Structure as Optimal Contracts, MS, Northwestern University, May.

Diamond, Douglas W., 1984, Financial Intermediation and Delegated Monitoring, Review of Economic Studies LI, 393-414.

Dye, Ronald A., 1986, Optimal Monitoring Policies in Agencies, Rand Journal of Economics 17 (Autumn), 339-350.

Fudenberg, Drew, and Jean Tirole, 1988, Moral Hazard and Renegotiation in Agency Contracts, MIT Working Paper Number 494.

Gale, Douglas, and Martin Hellwig, 1985, Incentive-Compatible Debt Contracts: The One Period Problem, Review of Economic Studies 52, 627-645.

Green, Jerry, and Jean-Jacques Laffont, 1987, Renegotiation and the Form of Efficient Contracts, Harvard Institute of Economic Research Discussion Paper Number 1338.

Gross, Charles, 1906, The Court of Piepowder, Quarterly Journal of Economics XX, 231-49.

Hart, Oliver, and John Moore, 1988, Incomplete Contracts and Renegotiation, Econometrica 56, 755-785.

Huberman, Gur, and Charles Kahn, 1987, Strategic Renegotiation and Contractual Simplicity, The Hoover Institute Working Papers in Economics E-87-31.

Huberman, Gur, and Charles Kahn, 1988, Limited Contract Enforcement and Strategic Renegotiation, American Economic Review 78, 471-484.

Kreps, David M., and Robert Wilson, 1982, Sequential Equilibria, Econometrica 50, 863-894.

Lacker, Jeffrey M., 1989, Financial Intermediation, Optimality and Efficiency, Canadian Journal of Economics XXII (May), 364-382.

Lacker, Jeffrey M., and John A. Weinberg, 1989, Optimal Contracts under Costly State Falsification, Journal of Political Economy 97 (December), 1345-1363.

Lane, Frederic C., 1958, Economic Consequences of Organized Violence, Journal of Economic History XVIII (December), 401-417.

Melumad, Nahum, and Dilip Mookherjee, 1987, Delegation as a Commitment Device: The Case of Income Tax Audits, Stanford University, Graduate School of Business, Research Paper No. 933.

Mookerjee, Dilip, and Ivan Png, 1989, Optimal Auditing, Insurance and Redistribution, Quarterly Journal of Economics CIV, 399-415.

Moore, Robert R., 1989, Costly Observation, Stochastic Auditing, and Time Consistency, MS, Tulane University, November.

Myerson, Roger B., 1986, Multistage Games with Communication, Econometrica 54, 323-358.

Reinganum, J, and L. Wilde, 1985, Income Tax Compliance in a Principal-Agent Framework, Journal of Public Economics 26, 1-18.

Rogerson, William P., 1984, Efficient Reliance and Damage Measures for Breach of Contract, The Rand Journal of Economics 15, 39-53.

Prescott, Edward C., and Robert M. Townsend, 1984, Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard, Econometrica 52, 21-46.

Shavell, Steven, 1980, Damage Measures for Breach of Contract, Bell Journal of Economics XI, 466-90.

Shavell, Steven, 1984, The Design of Contracts and Remedies for Breach, Quarterly Journal of Economics IC, 121-148.

Townsend, Robert M., 1979, Optimal Contracts and Competitive Markets with Costly State Verification, Journal of Economic Theory XXI, 265-293.

Townsend, Robert M., 1988, Information Constrained Insurance: The Revelation Principle Extended, Journal of Monetary Economics 21, 411-450.

Williamson, Stephen D., 1987, Costly Monitoring, Loan Contracts, and Equilibrium Credit Rationing, Quarterly Journal of Economics CII, 135-145.