

UC Berkeley

Other Recent Work

Title

Moral Hazard and Verifiability: The Effects of Renegotiation in Agency

Permalink

<https://escholarship.org/uc/item/1678w3w9>

Authors

Hermalin, Benjamin E.
Katz, Michael L.

Publication Date

1990-05-01

Peer reviewed

UNIVERSITY OF CALIFORNIA AT BERKELEY

Department of Economics

Berkeley, California 94720

Working Paper No. 90-141

**Moral Hazard and Verifiability:
The Effects of Renegotiation in Agency**

Benjamin E. Hermalin and Michael L. Katz

University of California, Berkeley

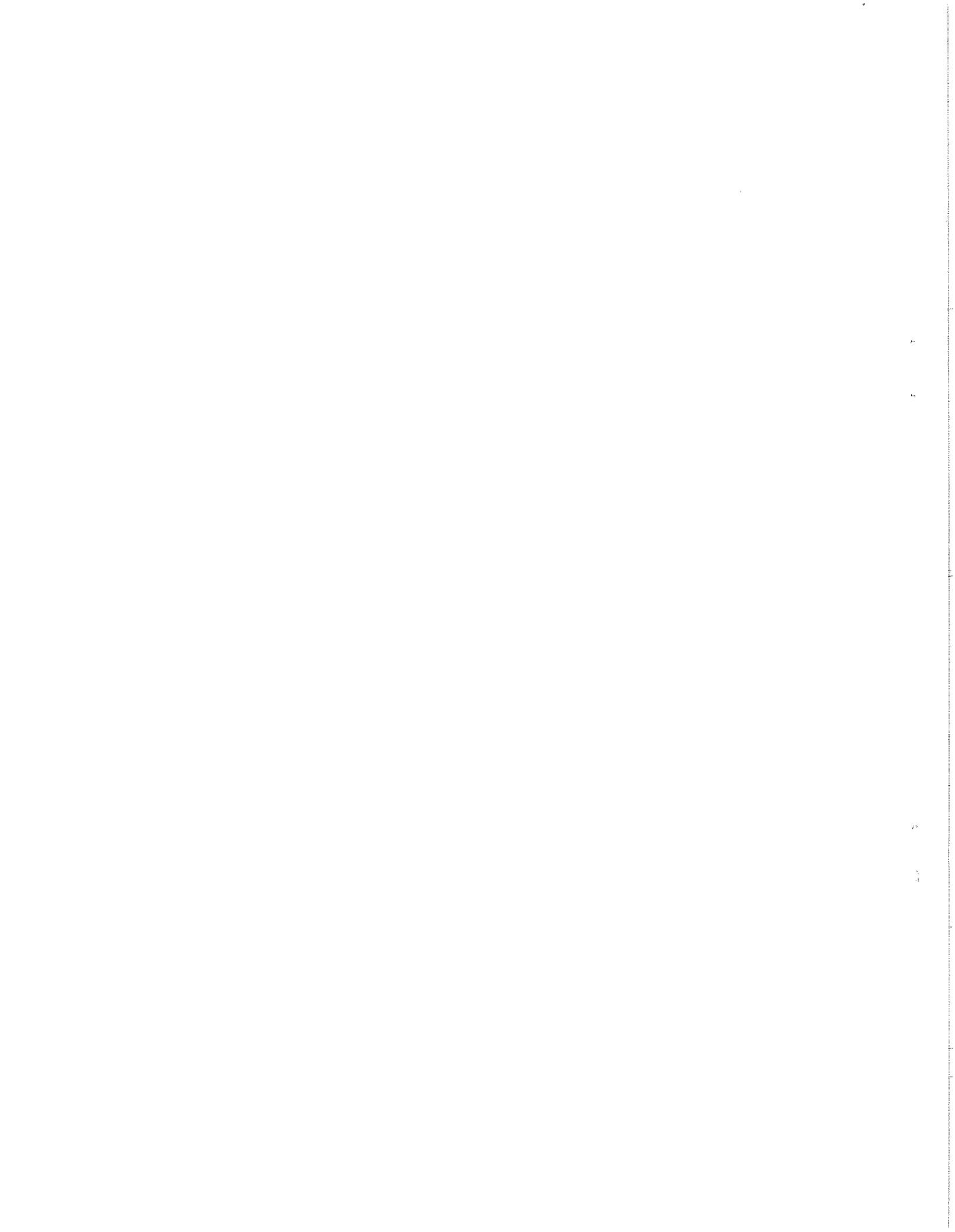
May 1990

Key words: agency, renegotiation, verifiability

Abstract

We examine the effects of the renegotiation in an agency relationship. We show how renegotiation affects: 1) the set of actions the principal can induce the agent to take; and 2) the cost of implementing a given action. We show that, when the principal receives an unverifiable signal of the agent's actions, renegotiation can improve welfare. This result stands in contrast to Fudenberg and Tirole's [1989] finding that renegotiation lowers welfare when the principal receives no signal about the agent's action prior to renegotiation.

JEL Classification: 026, 022



MORAL HAZARD AND VERIFIABILITY: THE EFFECTS OF RENEGOTIATION IN AGENCY¹

1. INTRODUCTION

An agency relationship involves three parties: a principal, an agent, and a third party (the "judge") responsible for enforcing the contract between the principal and the agent. Typically these parties are asymmetrically informed, and the distribution of information among the three parties plays an important role in determining the outcome of the agency relationship.

Consider a simple moral hazard problem in which the agent takes an action, a . Two reasonable assumptions about the distribution of information in this problem are:

- 1) the agent knows what action he took -- he observes a .
- 2) the judge sees some imperfect signal of the agent's action, x .

Because x can be seen by the judge, we refer to it as a verifiable performance measure. We also must make an assumption about how the principal's information compares with the agent's and the judge's. The classic principal-agent model (e.g., Holmstrom [1979], Shavell [1979], and Grossman and Hart [1983]) assumes that the principal has the same information as the judge -- both observe only x . Recent work (e.g., Lewis and Sappington [1989]) considers the other extreme, in which the principal and agent are equally well informed -- both observe a as well as x (what

¹ We are grateful to participants of the Northwestern University Summer Workshop, the Pew Charitable Trusts Conference on Incentives in Procurement Contracting, the Harvard University Conference on Incentives and Ex Post Constraints, and seminars at the University of California at Berkeley and the University of California at Davis. We especially wish to thank Bengt Holmström, Andreu Mas-Colell, Preston McAfee, Michael Riordan, Jean Tirole, and three anonymous referees for their insightful comments. This material is based on work supported by the National Science Foundation under grant no. SES-8722198 and by the Pew Charitable Trusts.

has become known as the case of an *observable but unverifiable* action). In many circumstances, the most reasonable assumption is that the principal's information about the agent's action is worse than the agent's but, because of the principal's expertise and relation-specific knowledge, better than the judge's.

Formally, we can think of the principal and the agent, but not the judge, as observing s , an additional signal of the agent's action. We can think of this unverifiable signal as ranging from perfectly informative to completely uninformative. One of the goals of this paper is to begin to explore the important intermediate cases.

The asymmetry of information across parties raises a number of important questions about how the distribution of information among the three parties affects the outcome of the principal-agent problem. For instance, to what extent does the quality of the judge's information determine the equilibrium outcome? We can begin to answer this question by supposing that the contract can depend solely on x . How informative does x have to be to induce the agent to take a particular action? This is, in effect, the question answered by Riordan and Sappington [1988] in their analysis of a buyer-seller relationship in which the seller has private information about a cost parameter. Although the buyer and (implicitly) the judge in their model never observe the cost parameter, they do observe a signal that is correlated with it. Riordan and Sappington showed that the buyer can purchase the efficient quantity of the good at an expected price equal to cost if, for each value of the cost parameter, the density over signals conditional on that value is distinct (in a sense to be made precise below) from the densities conditional on the other values.

Riordan and Sappington were concerned with a hidden state problem in which the parties are asymmetrically informed at the time of contracting. In this paper, we derive the analogous condition for a hidden action problem (Proposition 2): If the

density of x conditional on \hat{a} is distinct from the densities conditional on the agent's other actions in the same sense as in Riordan and Sappington, then there exists a contract $\hat{w}(x)$ that induces the agent to take an action \hat{a} . As we will discuss below, this condition states that \hat{a} is implementable if the verifiable performance measure x is informative about whether \hat{a} was chosen.

In addition to deriving conditions for feasibility, we ask about cost and efficiency. In Riordan and Sappington's model, both parties are risk neutral, so implementability of the first-best action assures the feasibility of attaining the full first-best outcome. While risk neutrality is a sensible assumption in their application, it is not an interesting one for a pure moral hazard problem; with a risk-neutral agent, the agency problem has the trivial solution of "selling the store" to the agent. When the judge's signal is a noisy measure of performance, the incentive scheme used to implement \hat{a} will typically require the agent to bear risk, which raises the cost of implementing \hat{a} when the agent is risk averse.

When the principal's information is better than the judge's, it is natural to ask whether the principal can use this information to increase the set of implementable actions or to lower the costs of implementing a given action. The common view in the contract theory literature is that she cannot because contracts based on unverifiable information are unenforceable (see Hart [1987] for a survey). In our opinion, this view is misleading. It is correct to argue that a judge cannot enforce contractual contingencies based on information unknown to him. But under reasonable conditions, the parties can write a contract that *in effect* makes the agent's compensation contingent on information that is unverifiable.

One approach is to set up a direct revelation mechanism in which the principal and agent reveal their private information to the judge. When the information is common knowledge between the principal and the agent, the contract could specify that

they go to court and simultaneously announce that information to the judge. The judge would then "shoot" the two if their reports disagreed. As is well known, there is a Nash equilibrium in which both parties report truthfully. This suggests that the extent to which information is verifiable before the judge is irrelevant. But, it also well known that this mechanism has difficulties, including: 1) possibly more compelling equilibria exist; 2) the mechanism is unbalanced; and 3) the mechanism is not robust to situations in which the principal and the agent have different private information.

Several authors have tried to find better-behaved revelation mechanisms. For instance, Gibbons and Murphy [1987] developed balanced-budget mechanisms that induce truthful revelation of a common signal as one of their (possibly many) equilibria when players are risk-neutral. Cremer and McLean [1988] and McAfee and Reny [1988] examined the closely related problem of an auction game in which the seller (our judge) wants to induce buyers (our principal and agent) to reveal their imperfectly correlated valuations to him. Cremer and McLean derived conditions under which the seller can induce revelation at the same cost as in a world of full information if all parties are risk neutral.²

In our opinion, agency relationships are not governed by contracts that use the courts as direct revelation mechanisms. Presenting evidence in court is much different from simply sending messages and being rewarded according to some preset payment schedule. Here, we pursue an alternative approach which does not rely on messages, and which is a natural one for the agency setting. We examine contract

² These conditions turn out to be quite similar in structure to those derived by Riordan and Sappington because one party's report to the judge essentially serves as the "performance measure" for the other party. But note that in Cremer and McLean the informativeness condition concerns the principal's information about the agent's information and *vice versa*. No use is made of the judge's information.

renegotiation.³ Suppose, for example, that the principal and the agent both observe the agent's action a , but the judge does not. The two parties sign an incentive contract under which the agent's compensation is contingent solely on the judge's signal, x . After the agent has taken his action, but before the realization of x , the principal and agent form a posterior distribution for realizations of x and, hence, the agent's compensation. Based on that distribution, the principal buys out the agent's contract at the certainty equivalent of the agent's compensation under the initial contract.

What effect does renegotiation have when the agent's action is observable, but unverifiable, and the principal makes take-it-or-leave-it offers in renegotiation? We first show that renegotiation has no effect on the set of implementable actions (Proposition 1). In this case, the entire effect of renegotiation is to reduce the cost of inducing any implementable action. In fact, renegotiation reduces the cost all the way down to its first-best level. In the standard agency model, there is a contract-design tradeoff between the provision of incentives and insurance, and consequently the cost of implementing an action exceeds its first-best cost. Renegotiation obviates this tradeoff by providing a second instrument for achieving the two goals. Loosely speaking, the initial contract creates incentives, while the renegotiated contract provides insurance.

The idea that renegotiation is a means to contract effectively on unverifiable information builds on the work of Huberman and Kahn [1987 and 1988]. They examined a sequential move game in which the first player's actions are observable but unverifiable, while the second player's actions are verifiable. Huberman and Kahn

³ Of course, by the revelation principle, this procedure could be replicated by some artificial direct revelation mechanism.

showed that, by renegotiating contracts written contingent on the verifiable actions, it is possible in some cases to achieve the same outcome as would be achieved if all actions were verifiable. In their model, however, the initial contract does not allow the verifiable action to be contingent on any verifiable measure of the unverifiable action; that is, the initial contract simply fixes the verifiable action.⁴ In contrast, we allow the initial contract to make the second player's verifiable action (the principal's payment to the agent) contingent on a verifiable measure of the first player's unverifiable action. This is an important generalization for a principal-agent problem. A further generalization is that, while Huberman and Kahn require the second player to observe perfectly the first player's action, we allow the second player to receive only an imperfect signal of the first player's action.

The remainder of our paper is organized as follows. The formal set up is given in the next section. The following section analyzes the model under the assumptions that the agent's action is observable, but unverifiable, and that the principal makes a take-it-or-leave-it offer in renegotiation. In Section 4, we generalize the renegotiation bargaining game. In Section 5, we examine what happens when the principal's signal, s , is imperfect but is a sufficient statistic for the agent's action. We show that, although the first best is generally unattainable, the principal is better off with renegotiation than without. This result should be contrasted with Fudenberg and Tirole's [1989] finding that renegotiation reduces welfare when the principal's unverifiable signal is completely uninformative.⁵ In

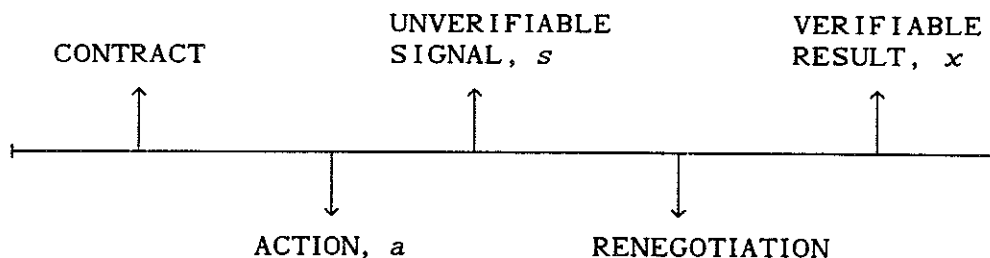
⁴ Alternatively, one can interpret Huberman and Kahn as assuming there is no verifiable measure of the unverifiable action.

⁵ Fudenberg and Tirole do not explicitly model signals received by the principal. Implicitly, however, they assume the principal receives some signal that notifies her

their model, renegotiation undermines commitment, and, hence, serves solely as an added constraint in the contract design problem.⁶ In Section 6, we examine what happens if, besides a signal of the agent's action, the parties also receive a signal that provides no additional information about the agent's action, but which does provide additional information about the verifiable performance measure prior to renegotiation. The paper closes with some brief remarks on directions for generalizing our analysis.

2. THE MODEL

We consider a principal-agent model with the following timing:



In the *contract* stage, the principal offers the agent an incentive contract on a take-it-or-leave-it basis. If the agent leaves it, the game ends and the players receive their reservation utility levels. Without loss of generality, we normalize the agent's reservation utility level to be 0. If the agent takes it, the contract becomes binding on both players.

In the *action* stage, the agent chooses an action, a , from the finite set \mathcal{A} ,

that *an* action has been taken (thus, she knows when to commence renegotiation), but which is completely uninformative about *which* action has been taken.

⁶ In this respect, the Fudenberg and Tirole model of renegotiation in agency is part of a larger literature that has tended to view renegotiation as undermining commitment. See, e.g., Hart and Moore [1988], Hart and Tirole [1988], Green and Laffont [1988], and Dewatripont and Maskin [1990].

which has $m + 1$ elements. Only the agent observes his choice of action at this stage.

In the *signal* stage, the principal and agent observe a common signal, s , about the agent's action. The signal is unverifiable; that is, no outside authority who might be called upon to enforce a contract can observe the signal directly.

In the *renegotiation* stage, the principal offers a new contract to the agent. Initially, we assume the offer is made on a take-it-or-leave-it basis. If the agent leaves the principal's offer, the original contract remains binding. If the agent takes it, the original contract is torn up and the new contract becomes binding.

Finally, in the *result* stage, there is the realization of some result, x , which is a stochastic function of the agent's action. The set of possible realizations, \mathcal{X} , is finite and contains n elements indexed by i . Denote the probability of x_i conditional on action a by $\pi_i(a)$, and let $\pi(a) = (\pi_1(a), \dots, \pi_n(a))'$ denote the probability vector (density) over the result conditional on action a .⁷

The result can be interpreted in many ways. It can represent: the result of production by the agent (e.g., bushels of peanuts produced by a tenant farmer); a financial measure of the agent's performance (e.g., the share price of a company the agent manages for the principal); or the outcome of some dispute-resolution mechanism (e.g., the judge's ruling itself). A maintained assumption is that the result is verifiable; that is, in the case of a dispute, the judge (alternatively, an arbitrator or a jury) can verify what the value of x is.

The principal is risk neutral with utility function $b(a) - y$, where $b(a)$ is either a deterministic function of a or the expected value of a stochastic function of a , and y is the compensation ultimately paid the agent. In the classic agency

⁷ The prime, $'$, denotes vector transpose.

model, x is the financial return on the asset owned by the principal and managed by the agent. In this interpretation, $b(a) = \pi(a)'x$, where $x = (x_1, \dots, x_n)'$. Here, we allow for the possibility that the financial return is not perfectly correlated with the result.

The agent's utility is additively separable in income and action: $u(y) - c(a)$, where $c(a)$ represents the disutility, or cost, of taking action a and $u(\cdot)$ is the agent's utility of money. We assume $u(\cdot)$ is continuous, strictly increasing, concave, and unbounded. Note that these assumptions imply that the inverse function $u^{-1}(\cdot)$ exists and that, for any $u \in \mathbb{R}$, there exists a $y \in \mathbb{R}$ such that $u(y) = u$.

For reasons given in the introduction, we rule out direct revelation mechanisms, and we assume that contracts are written contingent on the result only. Consequently, we may consider a contract to be an n -dimensional vector of contingent payments, $w = (w_1, \dots, w_n)'$, where w_1 is to be paid if the result is x_1 . We can also express a contract as the n -dimensional vector of contingent utilities that it induces, $u = (u_1, \dots, u_n)'$, where $u_1 = u(w_1)$.

3. THE BASIC CASE

In this section, we assume that the agent's action is observable, but not verifiable and the principal makes a take-it-or-leave-it offer in renegotiation. The principal's problem is to choose a contract and an action that maximize her expected utility subject to the constraints that the agent accept the contract and that the contract provide the agent with an incentive to take the chosen action. Throughout this paper, our solution concept is perfect Bayesian equilibrium.

As a benchmark, suppose that the agent's choice of action were verifiable. In this case, the principal could induce the agent to take an arbitrary action \hat{a} (i.e., implement \hat{a}) simply by offering a contract that specifies a punishment for taking

$a \neq \hat{a}$ sufficient to induce the agent to take action \hat{a} . Furthermore, as the principal has all the bargaining power, she can implement \hat{a} at a cost of $u^{-1}(c(\hat{a}))$ -- the agent is paid his reservation wage for that action. Following Grossman and Hart [1983], we say that $u^{-1}(c(\hat{a}))$ is the *first-best cost* of inducing \hat{a} .

When the agent's action is verifiable the principal's problem is

$$\max_{a \in \mathcal{A}} b(a) - u^{-1}(c(a)). \quad (1)$$

Since \mathcal{A} is finite, at least one solution to (1) exists. Let a^* denote the solution to (1).⁸ We refer to the outcome in which the principal is able to implement a^* at first-best cost as the *first-best outcome*. We refer to the subset of actions that minimize $c(a)$ as the *least-cost actions*. So that the agency problem is not trivial, we assume that a^* is *not* a least-cost action.

We now return to our assumption that the signal is unverifiable. As is well known, *without* renegotiation, a contract u implements action \hat{a} if and only if it is incentive compatible and individually rational:

$$\pi(\hat{a})'u - c(\hat{a}) \geq \pi(a)'u - c(a), \quad \forall a \neq \hat{a} \quad (\text{IC})$$

and

$$\pi(\hat{a})'u - c(\hat{a}) \geq 0. \quad (\text{IR})$$

Perhaps surprisingly, these are also the conditions for u to implement \hat{a} *with* renegotiation. To see this, consider what offer the principal would make to the agent at the renegotiation stage. Because the agent is risk averse, there are gains to be had from insuring the agent fully. Since the principal has the bargaining power, she captures all these gains. If the agent has taken action a , then in

⁸ Although there may exist multiple solutions to (1), we will speak of a^* as if it were the unique solution. It will be clear how the results generalize to the multiple solution case.

renegotiation the principal offers $u^{-1}(\pi(a)'u)$, the certainty equivalent of the agent's compensation under the initial contract. Hence, if the agent chooses action a under initial contract u , he knows his utility will be $\pi(a)'u - c(a)$, whether or not there is a renegotiation stage. Thus, the (IC) and (IR) constraints are unaffected by the possibility of renegotiation. It follows that the set of implementable actions is unaffected as well.

The benefit to renegotiation, therefore, is not that it expands the set of implementable actions. Rather, renegotiation is valuable because it reduces the cost of implementing any action to its first-best cost. To see this, let \hat{a} denote the action to be implemented. The (IR) constraint must be binding in equilibrium, since otherwise there would exist a \hat{u} and $\varepsilon > 0$, such that $\hat{u}_i = u_i - \varepsilon \forall i$ and \hat{u} solves (IR) and (IC) at a lower cost to the principal. In equilibrium, the agent undertakes \hat{a} and the principal's offer in renegotiation is $u^{-1}(\pi(\hat{a})'u) = u^{-1}(c(\hat{a}))$, since the (IR) constraint is binding.

To summarize, we have shown the following for the basic case:

Proposition 1: *When the agent's action is observable, but unverifiable, and the principal makes a take-it-or-leave-it offer in renegotiation, the set of implementable actions is the same with or without renegotiation. Moreover, with renegotiation, if an action is implementable, it is implementable at first-best cost.*

Notice that, with renegotiation, the ultimate compensation is *not* paid according to the initial contract. Nonetheless, the initial contract is important because it establishes the threat points for bargaining in renegotiation. That is, the renegotiated compensation is determined by the *expected* outcome should the parties carry out the initial contract. This is what distinguishes this analysis from the classic principal-agent model. In the classic model, compensation is contingent on

the *realization* of x (which is uncertain given a), whereas here compensation is contingent on the *distribution* of x (which is certain given a). Thus, in the classic model, the cost of implementing any action is greater than the first-best cost because the agent must be compensated for bearing risk.

Proposition 1 clearly implies that the principal is no worse off with renegotiation than without renegotiation: any action she could implement without renegotiation, she can implement with renegotiation; moreover, she implements that action at minimal (first-best) cost. Grossman and Hart [1983] showed that, if there is no shifting support (i.e., if $\pi_i(a) > 0, \forall i, \forall a \in \mathcal{A}$), then the without-renegotiation cost of implementing any action other than a least-cost action is strictly greater than the first-best cost of implementation. In summary, we have:

Corollary 1: *The principal is no worse off with renegotiation than without it. Moreover, in the absence of a shifting support, she is strictly better off when implementing all but the least-cost actions.*

We turn now to establishing the conditions under which a given action \hat{a} is implementable.

Proposition 2: *Suppose either*

- i) there is no renegotiation; or*
- ii) the agent's action is observable, but unverifiable, and the principal makes a take-it-or-leave-it offer in renegotiation.*

Then action \hat{a} is implementable if and only if there is no strategy for the agent that induces the same density over results as \hat{a} and which costs less, in terms of expected

disutility, than \hat{a} .⁹

Proof: Since we can always add or subtract the appropriate constant from each element of \mathbf{u} to satisfy the (IR) constraint without changing the (IC) constraints, it is clear that \hat{a} is implementable if and only if there exists a \mathbf{u} that solves the (IC) constraints. Rewriting the (IC) constraints as

$$\left(\pi(a_k) - \pi(\hat{a}) \right)' \mathbf{u} < c(a_k) - c(\hat{a}), \forall k,$$

where $k = 1, \dots, m$ indexes the elements in \mathcal{A} other than \hat{a} , we see that the (IC) constraints constitute a system of m linear inequalities. Theorem 22.1 of Rockafellar [1970] states that a solution to this system exists if and only if there is *no* vector $(\lambda_1, \dots, \lambda_m)' \geq \mathbf{0}$ (where $\mathbf{0}$ is an m -dimensional vector of zeros) such that

$$\sum_{k=1}^m \lambda_k \left(\pi(a_k) - \pi(\hat{a}) \right) = 0$$

and

$$\sum_{k=1}^m \lambda_k \left(c(a_k) - c(\hat{a}) \right) < 0.$$

Since this system is homogeneous, we are free to interpret the vector $(\lambda_1, \dots, \lambda_m)'$ as a strategy (i.e., $\sum \lambda_k = 1$). The result follows.¹⁰

■

Intuitively, if the density induced by \hat{a} is distinct from the density induced by

⁹ Here, and throughout the paper, "strategy" refers to mixed, as well as, pure strategies.

¹⁰ As noted in the introduction, similar mathematics underlies Theorem 2 of Cremer and McLean [1988] and the main proposition of Riordan and Sappington [1988]. It also underlies Proposition 1 of Melumad and Reichelstein's [1989] analysis of the value of communication in an agency relationship.

any other strategy, then the result is informative with respect to determining whether \hat{a} was chosen. Since the range of $u(\cdot)$ is unbounded, even a small amount of information can be used to implement \hat{a} by rewarding the agent for choosing \hat{a} and punishing him for choosing $a \neq \hat{a}$. The following formalizes this notion of "distinct":

Corollary 2: *Action \hat{a} is implementable if $\pi(\hat{a})$ is not an element of the convex hull of $\{\pi(a) | \forall a \neq \hat{a}\}$.*¹¹

Of course, even if there are other strategies that induce the same density as \hat{a} , \hat{a} is still implementable if the agent finds these other strategies more costly than \hat{a} .

To our minds, the "convex hull" condition is a weak one. Except when $n = 2$, it can be met without regard to m or n ; i.e., it is not a "dimensionality" condition. Furthermore, it is straightforward to show that it is met in many circumstances. It is met if the mode of $\pi(\hat{a})$ is different than the mode of $\pi(a)$ for all other a . It is met if there is any n -dimensional vector q such that $\pi(\hat{a})'q > \pi(a)'q, \forall a \neq \hat{a}$ (i.e., the expectation of q is greatest given $\pi(\hat{a})$). And it is met if there is a non-empty subset \hat{X} of X such that $\text{Prob}\{\hat{X}|\hat{a}\} = 0$ and $\text{Prob}\{\hat{X}|a\} > 0, \forall a \neq \hat{a}$ (i.e., there is a "shifting support").¹²

From the above analysis, it is clear that

Corollary 3: *With renegotiation, the first-best action is attainable at first-best cost if and only if there is no strategy for the agent that induces the same density*

¹¹ We thank Preston McAfee for suggesting this interpretation of our results.

¹² Indeed, as a referee pointed out to us, one can view the convex hull condition as a generalization of the notion that an action is always implementable with a shifting support: an action is also implementable with a "shifting" density.

over results as a^* and which costs less, in terms of expected disutility, than a^* .

Of course, despite the weakness of the convex hull condition, it is possible that a^* is not implementable. In this case, with renegotiation, the principal chooses to implement the action that solves

$$\max_{a \in \mathcal{A}^I} b(a) - u^{-1}(c(a)), \quad (2)$$

where \mathcal{A}^I denotes the set of implementable actions (characterized by Proposition 2). Since the least-cost actions are elements of \mathcal{A}^I , it is certainly non-empty, and (2) has a solution. Note that with renegotiation, even if the principal can implement only a second-best action, she does so at a first-best cost.

4. GENERALIZED RENEGOTIATION

So far, we have considered renegotiation in which the principal makes a take-it-or-leave-it offer. Now consider the opposite extreme in which the agent makes a take-it-or-leave-it offer in renegotiation. It is easy to show that the convex hull condition remains a *sufficient* condition for an action to be implementable at first-best cost when the agent makes a take-it-or-leave-it offer in renegotiation. The necessary condition, however, is slightly altered: *action \hat{a} is implementable with renegotiation only if there is no strategy for the agent that induces the same density over results as \hat{a} and which costs less, in terms of expected compensation, than \hat{a} .* By the "expected compensation" of strategy $(\lambda_1, \dots, \lambda_m)'$, we mean

$$\sum_{k=1}^m \lambda_k u^{-1}(c(a_k)).$$

Since $u^{-1}(\cdot)$ is a convex function, it follows straightforwardly from Proposition 2

that any action that is implementable without renegotiation is implementable when the agent makes a take-it-or-leave-it offer in renegotiation.

Recall that, in the classic agency model, $b(a) = \pi(a)'x$. Generalizing this condition, we obtain the following result:

Proposition 3: *Suppose the agent's action is observable, but unverifiable, and the agent makes a take-it-or-leave-it offer in renegotiation. If there exists a vector β that solves $\pi(a)'\beta = b(a)$, $\forall a$, then the first-best outcome is attainable. In particular, the first-best is attainable when the return on the principal's asset is verifiable (i.e., $\beta = x$).¹³*

Proof: Suppose the principal offers the initial contract $w^* = (\beta_1 - F^*, \dots, \beta_n - F^*)'$, where $F^* = b(a^*) - u^{-1}\{c(a^*)\}$. In renegotiation, the agent will offer to sell out his contract for $\pi(a)'\bar{w}^* = b(a) - b(a^*) + u^{-1}\{c(a^*)\}$, given he took action a , and the principal will accept. It is straightforward to verify that this contract supports the first-best outcome. ■

While at first surprising, the intuition behind this result is familiar. In renegotiation, the agent will receive a fixed amount in order to share risk optimally. That amount is determined by the principal's threat point, which depends solely on the expected value of the financial return from the project undertaken by the agent and the expected value of the wage payments under the initial contract. Hence, although he is risk averse, the agent makes his choice of action solely to maximize a function of expected values -- he behaves as if he were risk neutral.

¹³ As the proof makes clear, even this is stronger than we need if the asset can literally be sold. All that is, then, required is that the judge can verify whether the asset ("the store") has been sold.

Since the agent is risk neutral with respect to the outcome of his action, the principal can generate incentives by "selling the store" to the agent under the initial contract for a lump sum. After the agent has taken his action, the contract no longer has an incentive role to play, and the principal buys the store back from the agent for a lump sum (which depends on the choice of a) in order to provide insurance.

Relinquishing bargaining power at the renegotiation stage (weakly) raises the principal's expected surplus, given that she retains all of the bargaining power at the time that the initial contract is signed. Since the principal can anticipate what is going to happen in renegotiation, she simply adjusts the original selling price to compensate for the agent's later bargaining power in renegotiation. Moreover, the principal can gain from precommitting not to exploit any sunk investment (i.e., choice of a) made by the agent: in addition to reducing the cost of implementing actions, renegotiation can expand the set of implementable actions

Take-it-or-leave-it bargaining is extreme, particularly since only one side's threat point matters for the outcome. In general, we would expect bargaining in which the outcome is a function of *both* sides' threat points. Under certain regularity conditions on the properties of equilibrium and assumptions on the informational content of the result, we can extend our analysis to more general bargaining games:

Proposition 4: *Suppose that*

- 1) *the agent's utility in the unique subgame perfect equilibrium of the renegotiation bargaining game is given by $h(u,v)$, where u (v) is the agent's (principal's) expected utility should the initial contract remain in force,*

$h_1 > 0$, $h_2 < 0$, $h(u,v) \rightarrow -\infty$ as $u \rightarrow -\infty$ and $v \rightarrow \infty$, and $h(u,v) \rightarrow \infty$ as $u \rightarrow \infty$ and $v \rightarrow -\infty$; and

2) there exists an index set I over elements of \mathcal{X} such that

$$\sum_{i \in I} \pi_i(\hat{a}) > \sum_{i \in I} \pi_i(a), \quad \forall a \neq \hat{a};$$

then action \hat{a} is implementable at first-best cost with renegotiation.

Proof: Let $\bar{\pi}_1(a) = \sum_{i \in I} \pi_i(a)$. Consider the family of compensation contracts under which the agent is paid w_1 if x_j is realized, $j \in I$, and he is paid $w_2(w_1)$ otherwise, where $w_2(w_1)$ is defined by

$$u(w_1) - u(w_2) = \frac{c(\hat{a}) - \min_{a \neq \hat{a}} c(a)}{\bar{\pi}_1(\hat{a}) - \max_{a \neq \hat{a}} \bar{\pi}_1(a)}.$$

For all w_1 , the contract $(w_1, w_2(w_1))$ satisfies the (IC) constraints for \hat{a} . Define

$$\gamma(w_1) \equiv h \left[\pi(\hat{a})u(w_1) + [1-\pi(\hat{a})]u(w_2(w_1)), -\pi(\hat{a})w_1 - [1-\pi(\hat{a})]w_2(w_1) \right].$$

The function $\gamma(w_1)$ is strictly increasing, continuous, and unbounded. Therefore, there exists a w_1^* such that $\gamma(w_1^*) = c(\hat{a})$. The contract $(w_1^*, w_2(w_1^*))$ implements \hat{a} at first-best cost. ■

Note that when there are three or more elements in \mathcal{X} , condition (2) implies, but is not implied by, the convex hull condition. On the other hand, condition (2) is implied by many of the conditions that imply the convex hull condition, including the mode condition and the shifting support condition discussed earlier. Finally, note that Proposition 4 gives sufficient, but not necessary, conditions for \hat{a} to be implementable at first-best cost.

5. THE IMPERFECT SIGNAL CASE

So far we have assumed that the principal's signal of the agent's action is perfect. In many situations, however, the principal cannot directly observe the agent's action. In this section, we explore what happens when s is an imperfect signal of a . For convenience, we return to our assumption that the principal makes a take-it-or-leave-it offer in renegotiation.

Both the principal and the agent observe s , but only the agent observes his action. The signal is drawn from a finite set \mathcal{S} , whose elements are indexed by $j = 1, \dots, J$. Let $p_j(a)$ denote the probability of receiving signal s_j conditional on the action having been a , and let $p(a) = (p_1(a), \dots, p_J(a))'$.

Denote the probability of x_1 conditional on action a and signal s by $\sigma_1(a,s)$ and let $\sigma(a,s) = (\sigma_1(a,s), \dots, \sigma_n(a,s))'$. Obviously,

$$\pi_1(a) = \sum_{j=1}^J \sigma_1(a,s_j) p_j(a).$$

A. Analysis Under the Sufficient Statistic Condition

Since the agent knows a and s , while the principal only knows s , renegotiation potentially takes place under asymmetric information. When s is a sufficient statistic for both x and a , however, the principal can predict x as well as the agent can, even though the agent knows a .

Sufficient Statistic Condition: For any actions a and \hat{a} ,

$$\sigma(a,s) = \sigma(\hat{a},s), \forall s \in \mathcal{S}.$$

When this condition is satisfied, we write $\sigma(s)$ for $\sigma(a,s)$.

The sufficient statistic condition may be a reasonable assumption in some settings. Suppose, for example, that a product's quality is stochastically related

to the agent's quality-control effort. Moreover, suppose the principal and the agent both observe a test of product quality and s is their expert interpretation of the test results. Because of a lack of expertise, a judge's interpretation (x) is noisier than the principal and agent's. Conditional on s , however, the noise in the judge's interpretation is independent of the agent's quality-control effort -- s is a sufficient statistic for x .

Given signal s_j , both the principal and the agent know that the density over results is $\sigma(s_j)$. Thus, if the initial contract is u , the principal will offer $u^{-1}(\sigma(s_j)'u)$ in renegotiation. Consequently, if the agent chooses action a , his expected utility is

$$\sum_{j=1}^J p_j(a) u \left[u^{-1}(\sigma(s_j)'u) \right] - c(a) = \sum_{j=1}^J p_j(a) (\sigma(s_j)'u) - c(a) = \pi(a)'u - c(a).$$

Therefore, under the sufficient statistic assumption, the same incentive compatibility and individual rationality constraints hold with and without renegotiation. Just as in Section 3, the value of renegotiation is that it reduces the cost of implementing an action by improving the equilibrium pattern of risk sharing. In summary,

Proposition 5: *If the unverifiable signal is a sufficient statistic and the principal makes a take-it-or-leave-it offer in renegotiation, then the set of implementable actions is the same with or without renegotiation. Moreover, the cost of implementing any given action is no greater with renegotiation than without. If there is no shifting support (i.e., if $\sigma_i(s_j) > 0 \forall i,j$), the cost saving is strictly positive for all but the least-cost actions.*

It follows from Proposition 5 that, if s is a sufficient statistic for a , the principal is better off with renegotiation than without. It also follows that, if

$\pi(a)$ satisfies the convex hull condition, then a is implementable even when s is only a sufficient statistic for a .

Although, the above analysis shows that renegotiation dominates no renegotiation, it does not say how well the parties can do with renegotiation. In particular, we have not yet addressed the issue of whether the principal suffers a loss because her signal is unverifiable. Holmstrom [1979] has shown that, given the sufficient statistic condition, were it possible to contract on the signal, the optimal contract would be contingent on the signal *only*. Can the principal do as well contracting on the result as she would do if she could base a contract on her signal? The answer to this question is given by

Proposition 6: *Assume the sufficient statistic condition is satisfied and the principal makes a take-it-or-leave-it offer in renegotiation. Let Σ be the $J \times n$ matrix whose j th row is $\sigma(s_j)'$. Consider any incentive contract that could be written contingent on the principal's signal if the signal were verifiable. If Σ has full rank, then there exists an incentive contract written contingent solely on the result that induces the same action at the same expected cost as would the contract written contingent on the signal were the signal verifiable.*

Proof: Let z be the contract written contingent on the signal (i.e., the agent would receive utility z_j if signal s_j were realized). As Σ has full rank, there exists a contract u written contingent on the result that solves the equations

$$\sigma(s_j)' u = z_j$$

for $j = 1, \dots, J$. Clearly, u and z induce the same action. Furthermore, because of renegotiation, the principal expects to pay $u^{-1}(\sigma(s_j)' u)$ if signal s_j is realized; which is what she would pay if z were the contract.

■

Unlike the perfect signal case, here there is a dimensionality restriction. For the principal to do as well when the signal is unverifiable, as she could do if it were verifiable, the number of possible results must be at least as great as the number of possible signals, i.e., $n \geq J$.

B. Analysis without the Sufficient Statistic Assumption

We now turn to signals that do not satisfy the sufficient statistic condition. Our focus is on finding conditions under which a given action \hat{a} is implementable.

Suppose that $p_j(\hat{a}) > 0, \forall j$, and, in equilibrium, the agent chooses action \hat{a} with certainty. Then, at the renegotiation stage, Bayesian consistency requires that the principal believe that the agent chose action \hat{a} . Thus, she must believe the density over results is $\sigma(\hat{a}, s)$, where s is the signal she received. The principal's best response to her beliefs is to offer the agent $u^{-1}(\sigma(\hat{a}, s)'u)$ in renegotiation. If the agent has taken action \hat{a} (which he will in equilibrium), then he will accept. Of course, we must also consider what happens if the agent deviates; if he has taken some other action a , then he will accept if and only if

$$\sigma(\hat{a}, s)'u \geq \sigma(a, s)'u.$$

Consequently, with an imperfect signal and renegotiation, the incentive compatibility constraints for \hat{a} to be chosen are

$$\sum_{j=1}^J p_j(\hat{a}) \sigma(\hat{a}, s_j)'u - c(\hat{a}) \geq \sum_{j=1}^J p_j(a) \cdot \max[\sigma(a, s_j)'u, \sigma(\hat{a}, s_j)'u] - c(a), \quad \forall a \neq \hat{a}. \quad (\text{IC}'')$$

Since

$$\max[\sigma(a, s_j)'u, \sigma(\hat{a}, s_j)'u] \geq \sigma(a, s_j)'u,$$

it is clear that any u that implements \hat{a} with renegotiation must also implement \hat{a} without renegotiation. Thus, any action that can be implemented with renegotiation

can be implemented without renegotiation:

Proposition 7: *Suppose that $p_j(a) > 0, \forall j$ and $\forall a \in \mathcal{A}$, and that the principal makes a take-it-or-leave-it offer in renegotiation. With an imperfect signal, the set of implementable actions with renegotiation is a subset of the set of implementable actions without renegotiation.*

Fudenberg and Tirole [1989] showed the converse is not true; when the signal is uninformative, renegotiation can strictly reduce the set of implementable actions. Fudenberg and Tirole's finding that, when the signal is pure noise, only the least-cost actions are implementable, is one consequence of the following general result:

Proposition 8: *Suppose that $p_j(\hat{a}) > 0, \forall j$, and that the principal makes a take-it-or-leave-it offer in renegotiation. Action \hat{a} is implementable under renegotiation only if there is no strategy for the agent that induces the same density over signals as \hat{a} and which costs less, in terms of expected disutility, than \hat{a} .*

Proof: Suppose such a strategy $(\lambda_1, \dots, \lambda_m)$ exists. Hence,

$$\sum_{k=1}^m \lambda_k \left(p(a_k) - p(\hat{a}) \right) = 0$$

and

$$\sum_{k=1}^m \lambda_k \left(c(a_k) - c(\hat{a}) \right) < 0$$

(where, again, k indexes the elements in \mathcal{A} other than \hat{a}). From Theorem 22.1 of Rockafellar [1970], there is then *no* u that solves all of the following m inequalities:

$$\sum_{j=1}^J p_j(\hat{a}) \left\{ \sigma(\hat{a}, s_j)' u \right\} - c(\hat{a}) \geq \sum_{j=1}^J p_j(a_k) \left\{ \sigma(\hat{a}, s_j)' u \right\} - c(a_k).$$

As

$$\max \left[\sigma(a_k, s_j)' u, \sigma(\hat{a}, s_j)' u \right] \geq \sigma(\hat{a}, s_j)' u,$$

this means there is no u that satisfies (IC''). Therefore \hat{a} is not implementable. ■

Hence, if the unverifiable signal is uninformative (in terms of the convex hull condition), then renegotiation may limit the set of implementable actions by undermining commitment and, thus, weakening incentives.

We rule out the possibility of a shifting support (over signals) in Propositions 7 and 8 to avoid the complications inherent in specifying out-of-equilibrium beliefs for the principal. Although it is straightforward to conduct the analysis for *arbitrary* out-of-equilibrium beliefs (e.g., the principal believes the agent took the lowest-cost action consistent with an out-of-equilibrium signal), a satisfactory analysis would require the principal to hold *reasonable* out-of-equilibrium beliefs, and we do not want to get into a discussion here of what constitutes "reasonable" out-of-equilibrium beliefs in this context.

6. LEAKAGE

We have seen that renegotiation can benefit the principal. But these benefits might be undone if, between the time when the agent takes his action and when renegotiation occurs, there is "information leakage"; that is, if the principal and the agent learn information that helps them predict the result. For example, if the result itself is leaked to the principal and the agent, then there is nothing to renegotiate and there is no scope for reducing the risk the agent must bear.

Similar effects might arise in a less extreme case. Consider, for example, a

procurement relationship in which the department of defense (DoD) can observe a weapon contractor's actions (e.g., research, design, and manufacture). The contract between the DoD and contractor makes the contractor's payment contingent on various properties of the weapon (e.g., accuracy, speed, and reliability). After the contractor has taken his actions, but prior to renegotiation, certain aspects of the weapon's quality become apparent to both the DoD and the contractor. The DoD and the contractor then make predictions about how a judge would enforce their contract from their observations of the quality, as well as their knowledge of the contractor's actions. Clearly, these predictions will affect renegotiation. When quality and, thus, the predictions are stochastic functions of the contractor's actions, intuition suggests the contractor's compensation will be a stochastic function of the contractor's actions.

Fortunately, this intuition is incomplete. It ignores the fact that, although the density is a stochastic function of the agent's actions, this need not mean that the resulting threat points for renegotiation are stochastic. Given sufficient degrees of freedom in specifying the compensation contingent on the result, an action can still be implemented at first-best cost even with information leakage.

To explore these issues formally, we return to our assumption that the action is observable but unverifiable (i.e., the principal receives a perfect signal of a). In addition to the action, the parties also observe the leakage. Let ω denote the leakage. The leakage ω is drawn from the set Ω , which has L elements indexed by ℓ . Let $\psi(a, \omega)$ denote the probability vector over results conditional on a and ω . We allow for the possibility that the leakage depends stochastically on the action taken, thus we write $\phi_\ell(a)$ for the probability that ω_ℓ is the leakage conditional on the action having been a . Clearly,

$$\pi(a) = \sum_{\ell=1}^L \phi_{\ell}(a) \psi(a, \omega_{\ell}).$$

The leakage itself may be either verifiable or unverifiable.

We first note that leakage never makes renegotiation undesirable when the agent's action is observable but unverifiable. Consider unverifiable leakage first. In this case, the agent's expected utility from taking action \hat{a} given contract \mathbf{u} is

$$\sum_{\ell=1}^L \phi_{\ell}(\hat{a}) u \left[u^{-1} \left(\psi(\hat{a}, \omega_{\ell})' \mathbf{u} \right) \right] - c(\hat{a}) = \pi(\hat{a})' \mathbf{u} - c(\hat{a}),$$

and the set of implementable actions is the same with or without renegotiation. Furthermore, since the agent is risk averse, the cost of implementing \hat{a} is no greater with renegotiation than it would be without;

$$u^{-1} \left(\psi(\hat{a}, \omega_{\ell})' \mathbf{u} \right) \leq \sum_{i=1}^n \psi_i(\hat{a}, \omega_{\ell}) u^{-1}(u_i).$$

With verifiable leakage, a contract specifies a utility level contingent on both the result and the leakage; i.e., $u_{i\ell}$ is paid if ω_{ℓ} and x_i obtain. Define $\mathbf{u}^{\ell} = (u_{1\ell}, \dots, u_{n\ell})'$. Then $\{\mathbf{u}^1, \dots, \mathbf{u}^L\}$ implements \hat{a} with or without renegotiation, since

$$\sum_{\ell=1}^L \phi_{\ell}(\hat{a}) u \left[u^{-1} \left(\psi(\hat{a}, \omega_{\ell})' \mathbf{u}^{\ell} \right) \right] - c(\hat{a}) = \sum_{\ell=1}^L \phi_{\ell}(\hat{a}) \psi(\hat{a}, \omega_{\ell})' \mathbf{u}^{\ell} - c(\hat{a}).$$

Furthermore, since the agent is risk averse, the cost of implementing \hat{a} is no greater with renegotiation than without;

$$u^{-1} \left(\psi(\hat{a}, \omega_{\ell})' \mathbf{u}^{\ell} \right) \leq \sum_{i=1}^n \psi_i(\hat{a}, \omega_{\ell}) u^{-1}(u_{i\ell}).$$

We now turn to finding conditions under which \hat{a} can be implemented at first-best cost. A sufficient condition with verifiable leakage is a straightforward extension of our earlier convex hull condition: if

$$\forall l, \psi(\hat{a}, \omega_l) \notin \text{co}\left(\{\psi(a, \omega_l) \mid \forall a \neq \hat{a}\}\right), \quad (3)$$

then \hat{a} is implementable at first-best cost (where $\text{co}(\cdot)$ denotes "convex hull of"). Intuitively, when ω is verifiable, the parties can sign a "grand" contract that consists of a menu of contracts, one of which is put in force by the realization of ω . Hence, in effect, there is a separate contract for each value of ω , and the parties effectively "undo" the leakage.

Surprisingly, there are conditions under which we can also implement \hat{a} at first-best cost even when the leakage is completely *unverifiable*. Action \hat{a} is implementable at first-best cost with unverifiable leakage if there exists a vector u that solves

$$\psi(\hat{a}, \omega_l)' u = c(\hat{a}), \quad \forall l, \quad (\text{IR}^*)$$

and

$$\pi(a_k)' u - c(a_k) \leq \pi(\hat{a})' u - c(\hat{a}), \quad \forall k. \quad (\text{IC}^*)$$

The (IR^{*}) conditions come from the fact that \hat{a} can be implemented at first-best cost only if the agent's compensation equals $u^{-1}\left(c(\hat{a})\right)$.

Using the same techniques as above, it is straightforward to show that a sufficient condition for (IR^{*}) and (IC^{*}) to have a solution is that the space spanned by $\{\psi(\hat{a}, \omega_l), l = 1, \dots, L\}$ not intersect the convex hull generated by $\{\pi(a_k), k = 1, \dots, m\}$. Note that this condition implies, but is not implied by, the sufficient condition for \hat{a} to be implementable at first-best cost with *verifiable* leakage.

Even though a given action \hat{a} may still be implementable at first-best cost with unverifiable leakage, the conditions under which this occurs are more stringent with

May 14, 1990

**Working Paper Series
Department of Economics
University of California, Berkeley**

This is a list of recent working papers in the series. Individual copies are available for \$3.50 in the USA/Canada, \$6.00 Europe/South America, and \$7.00 Pacific Rim/Middle East. Papers may be obtained from the Institute of Business and Economic Research: send requests to IBER, 156 Barrows Hall, University of California, Berkeley CA 94720. Prepayment is required. Make checks or money orders payable to "The Regents of the University of California."

- 89-112 "Simultaneous Offers and the Inefficiency of Bargaining: A Two-Period Example." Eddie Dekel. May 1989.
- 89-113 "Factor Prices in Egypt from 1900 to World War II with International Comparisons." Bent Hansen. June 1989.
- 89-114 "Micro-Level Data Sets Suitable for Investigation of Macroeconomic Issues Extracted from Reports of the State Bureau of Labor Statistics, Circa 1890." Susan B. Carter and Richard Sutch. July 1989.
- 89-115 "Staggered Price Setting with Endogenous Frequency of Adjustment." David Romer. July 1989.
- 89-116 "Phases in the Development of the International Monetary System." Barry Eichengreen. August 1989.
- 89-117 "The Capital Levy in Theory and Practice." Barry Eichengreen. August 1989.
- 89-118 "International Monetary Instability Between the Wars: Structural Flaws or Misguided Policies?" Barry Eichengreen. August 1989.
- 89-119 "The Comparative Performance of Fixed and Flexible Exchange Rate Regimes: Interwar Evidence." Barry Eichengreen. August 1989.
- 89-120 "The Gold Standard Since Alec Ford." Barry Eichengreen. August 1989.
- 89-121 "The Role of Outside Considerations in the Design of Compensation Schemes." Bernard Caillaud and Benjamin Hermalin. July 1989.
- 89-122 "Moral Hazard and Verifiability." Benjamin E. Hermalin and Michael L. Katz. October 1989.
- 89-123 "Airline Deregulation and Market Performance: The Economic Basis for Regulatory Reform and Lessons from the U.S. Experience." Theodore E. Keeler. October 1989.
- 89-124 "Credibility and Cooperation under the Gold Standard." Barry Eichengreen. October 1989.

- 89-125 "The Labor Market in the 1890s: Evidence from Connecticut Manufacturing." Susan B. Carter and Richard Sutch. December 1989.
- 89-126 "Quantifying International Capital Mobility in the 1980s." Jeffrey A. Frankel. December 1989.
- 89-127 "Topological Social Choice." Nick Baigent and Peter Huang. December 1989.
- 89-128 "Equilibrium of Incomplete European Option Markets." Peter Huang and Ho-Mou Wu. December 1989.
- 89-129 "The Impact of Corporate Restructuring on Industrial Research and Development." Bronwyn H. Hall. December 1989.
- 89-130 "Converters, Compatibility, and the Control of Interfaces." Joseph Farrell and Garth Saloner. December 1989.
- 90-131 "And Now Won/Dollar Negotiations? Lessons From the Yen/Dollar Agreement of 1984." Jeffrey A. Frankel. January 1990.
- 90-132 "One Money for Europe? Lessons from the US Currency Union." Barry Eichengreen. January 1990.
- 90-133 "Does Foreign Exchange Intervention Matter? Disentangling the Portfolio and Expectations Effects for the Mark." Kathryn Dominguez and Jeffrey A. Frankel. January 1990.
- 90-134 "The Constrained Asset Share Estimation (CASE) Method: Testing Mean-Variance Efficiency of the U.S. Stock Market." Charles Engel, Jeffrey A. Frankel, Kenneth A. Froot, and Anthony Rodrigues. January 1990.
- 90-135 "International Nominal Targeting (INT): A Proposal for Overcoming Obstacles to Policy Coordination." Jeffrey A. Frankel. January 1990.
- 90-136 "One-Sided Patience With One-Sided Communication Does Not Justify Stackelberg Equilibrium." Eddie Dekel and Joseph Farrell. February 1990.
- 90-137 "Japanese Finance: A Survey." Jeffrey A. Frankel. March 1990.
- 90-138 "Government Policy, Individual Actions, and Safety: Some Evidence for Motor Vehicles." Theodore E. Keeler. April, 1990.
- 90-139 "Adverse Selection, Short-Term Contracting, and the Underprovision of On-the-Job Training." Benjamin Hermalin. February 1990.
- 90-140 "Why Legal Restrictions on Private Contracts Can Enhance Efficiency." Philippe Aghion and Benjamin Hermalin. April 1990.
- 90-141 "Moral Hazard and Verifiability: The Effects of Renegotiation in Agency." Benjamin E. Hermalin and Michael L. Katz. May 1990.

