On the implementability condition in agency models

Sekiguchi, Tadashi

On the Implementability Condition in Agency Models[*]

Tadashi Sekiguchi

Faculty of Economics
Kobe University
2-1 Rokkodai-cho, Nada-ku
Kobe 657-8501 Japan
e-mail: sekiguchi@econ.kobe-u.ac.jp

December 14, 2001

Abstract

This paper considers standard agency models, and provides a new necessary and sufficient condition for the agent's action to be implementable. The new condition works best when we are interested in implementability of the first-best action, which is an interesting question to ask if we introduce a possibility of renegotiation. We derive some economic sufficient conditions of implementability of the first-best action. We also show that our condition is applicable to other models than the standard agency models, such as multi-agency models and an adverse selection model where the principal can observe a signal of the agent's type ex post.

Keywords: agency, implementability, renegotiation

1

1 Introduction

In the standard agency model, the profit-maximizing principal must simultaneously decide the action she wants to induce the agent to choose, and the contract which supports it in the least costly way. It is in this respect that the decomposition proposed by Grossman and Hart (1983) facilitates analysis considerably. The essence of their approach is to fix an action arbitrarily, and then find a contract which minimizes the (expected) payment to the agent, under the constraint that the agent is willing to choose the action. The solution of this problem determines the maximum surplus of the principal, associated with the action. Then the principal chooses the action which maximizes the associated surplus. The decomposition is analytically convenient because the cost-minimization problem the principal faces given an action often becomes a tractable convex program, while the original problem of finding the optimal contract is not.[1]

It is therefore important to determine whether the above cost-minimization problem, given an action, has a solution at all; we say the action is *implementable* if the problem associated with the action is feasible. The existing condition of implementability is the one by Hermalin and Katz (1991), which is a direct application of Fan's (1956) theorem on existence of a solution in linear inequality systems. The condition, which is a necessary and sufficient condition of implementability, builds on some convex analysis. Therefore, it is most convenient when the convexity structure of the model is well understood, although this need not always be the case.

The purpose of this paper is to provide a new, alternative necessary and sufficient condition of implementability in the agency model. In terms of mathematics, our condition is not very different from the one by Hermalin and Katz (1991), but its appearance is. The condition states that an action of the agent is implementable if (and only if) the action is a solution of some maximization problem over the set of the agent's actions. As will be seen in subsequent sections, our condition is most relevant in examining implementability of the first-best action, namely, the action that the principal induces the agent to choose if the agent's action is perfectly observable. This is because the maximization problem that appears in our condition can be easily related to the maximization problem which determines the first-best action.

Using the condition, we derive the following economic sufficient conditions of implementability of the first-best action. First, if the first-best action is still first-best under the assumption that the principal can force the agent to randomize among his actions, then the first-best action is implementable. Since the first-best action is usually defined without a possibility of such randomization, being first-best even when mixed actions are allowed is an additional requirement. Second, if the first-best action is also first-best in the model where the agent is

---

[1] In addition, one can explore properties of the cost-minimization problem which hold independent of the fixed action. Those properties naturally take over to the optimal contract.

rather assumed to be risk-neutral (any other aspects of the model remain the same), then that action is implementable.

If the agent is risk-averse, the first-best action need not be the second-best action, because the cost of incomplete risk-sharing for the first-best action may be very large, in comparison with other implementable actions. Nevertheless, the issue of implementability of the first-best action is interesting per se, in relation to the literature on the role of contractual renegotiation in agency models. For example, Hermalin and Katz (1991) consider an agency model where the principal, after offering the initial contract and then observing the agent's action, makes a renegotiation offer. That is, the agent's action is assumed to be observable in their framework, but it is not verifiable or contractible. In this setting, Hermalin and Katz (1991) show that, as a consequence of renegotiation, any implementable action (when renegotiation is not allowed) is implementable under renegotiation *at the first-best cost.* Therefore, their result states that any agency cost associated with incomplete risk-sharing disappears when renegotiation is possible, *if the action is implementable.* However, the agency cost still remains positive if the first-best action is not implementable. Thus implementability of the first-best action is an interesting question to ask, because it determines when the agency cost completely vanishes, thanks to renegotiation.[2]

In this paper, we also discuss how our implementability condition applies to other related models. We consider two types of model. The first is an extension of the agency model to the case with multiple agents. As in the single-agent model, we present an economic sufficient condition of implementability of the first-best action profile. Again, the result is most relevant when renegotiation is possible. Ishiguro and Itoh (2001) extend the idea of Hermalin and Katz (1991) to the multi-agent setting, and show that any agency cost of an action that is implementable without renegotiation vanishes, if there is an opportunity of renegotiation. Together with their result, our condition is again helpful in determining when the agency cost vanishes at all.

The second model is an adverse selection model studied by Riordan and Sappington (1988), where the regulator (interpreted as the principal) observes a contractible signal about the type (cost structure, more specifically) of the regulated monopolist (interpreted as the agent), after the monopolist has chosen its outputs. In this model, a contract can depend on the ex post signal, and this feature allows us to rewrite the incentive problem of each type of the agent (that is, the incentive to report his type truthfully) into an incentive problem of a closely related moral hazard problem. Thus implementability of a particular outcome in an adverse selection model can be reduced to implementability of particular actions in the corresponding agency model, and therefore our technique is appli-

---

[2]Hermalin and Katz (1991) also consider a model in which it is the agent who offers a renegotiation contract, and show that the first-best action is always implementable at the first-best cost. From this viewpoint, our results help examine when the format of the renegotiation game, in particular, who makes the renegotiation offer, is irrelevant for the purpose of achieving the first-best.

cable. We present an economic sufficient condition for implementability of the first-best outcome, which is different from the one by Riordan and Sappington (1988).

Legros and Matsushima (1991) use the same technique based on Fan's (1956) theorem for the issue of enforcing efficient action profiles within a partnership, and they derive a necessary and sufficient condition to implement the efficient action profiles by using an appropriate transfer scheme. Similar mathematics is used in Fudenberg, Levine and Maskin (1994), who provide a sufficient condition of sustaining the efficient actions by repeated play, rather than a transfer schedule. See d'Aspremont and Gerard-Varet (1998) for an overview which explains their results in a unified way. Thus it is interesting to examine whether our new implementability condition gives some new insights to those partnership models and provides some economically interesting sufficient conditions of implementability. The main difference between the agency model (in particular, the multi-agency model) and the partnership model is that budget-balancing enters as an additional, linear constraint in the latter model.[3] It is an interesting avenue of future research to accommodate the budget-balancing constraint into our approach and extend it to the partnership model.[4]

The rest of this paper proceeds as follows. In Section 2, we describe the model. In Section 3, the existing implementability condition is discussed, and then we present a new condition of implementability. In Section 4, we apply the new implementability condition to the standard agency model, and provide several sufficient conditions of implementability of the first-best action. In Section 5, we present an example of the agency model and see how our condition works. In section 6, we extend our method to a multi-agency model and an adverse selection model with an ex post signal. In both cases, we provide a sufficient condition of attainability of the first-best actions/outcomes.

2   The Basic Model

There are two parties, the principal (often denoted P), and the agent (often denoted A).[5] We denote the set of A's actions (or efforts) by E, which is assumed to be a finite set. The set of signals is a finite set Y, and the probability

---

[3] The idea of budget-balancing is also important in the repeated games studied by Fudenberg, Levine and Maskin (1994). In order to achieve cooperative outcomes, the players want to arrange future reward and punishment, so that future play does not entail efficiency loss. This idea is quite similar to that of budget-balancing.

[4] As d'Aspremont and Gerard-Varet (1998) point out, the methods for the partnership models are first used in some *adverse selection models* such as d'Aspremont and Gerard-Varet (1979). Therefore, as we connect our implementability condition in the agency model with the adverse selection model by Riordan and Sappington (1988), there may be some connection between some other adverse selection models and the partnership model. It is another avenue of future research to explore the connection.

[5] Throughout the paper, we assume that a female principal hires a male agent.

distribution of the signal given an action $e \in E$ is denoted by $\pi(y|e)$. We assume $\pi(y|e) > 0$ for any $y \in Y$ and $e \in E$.

The agent's utility function is separable with respect to his action and monetary income, and is given by $u(w) - c(e)$, where w is the income. We assume $u : R \to R$ is continuous, strictly increasing and unbounded above and below, so that the inverse u^{-1} exists, with the domain being R. We do not need any additional restriction such as strict concavity. We also assume that the agent's reservation utility is zero. The utility function of the principal can depend directly on both the signal she observes and the agent's action. Specifically, we assume that if A chooses e, a signal $y \in Y$ is realized and P pays w to A, then the principal's utility is $R(e, y) - w$. Thus $R(e, y)$ is her revenue under the state (e, y).

Possible dependence of R on e implies that P might be able to observe e. However, we assume e is unverifiable, and therefore the contract can depend only on y. We thus denote a contract by $w(y)$. One interpretation is that P can observe the agent's action only after a long period of time, and P must pay to A before she learns e, because A is liquidity-constrained. In this case, it is implicitly assumed that she cannot collect any penalty after learning the actual effort of the agent.

In this setting, the principal's objective is to find out the optimal contract that is incentive compatible and individually rational. Hence the problem P solves is

$$\max_{\{e, w(\cdot)\}} \sum_{y \in Y} [R(e, y) - w(y)] \pi(y|e)$$

subject to

$$\sum_{y \in Y} u(w(y)) \pi(y|e) - c(e) \geq \sum_{y \in Y} u(w(y)) \pi(y|e') - c(e') \tag{1}$$

for any $e' \in E$, and

$$\sum_{y \in Y} u(w(y)) \pi(y|e) - c(e) = 0. \tag{2}$$

We denote the above maximization problem by (P^*). Since we assume that u is unbounded below, we can formulate (P^*) so that the individual rationality condition (2) is binding, without loss of generality.

3   The Implementability Condition

The core of the decomposition by Grossman and Hart (1983) is to solve (P^*) in two steps. First, we fix $e \in E$ arbitrarily, and find a contract that maximizes P's payoff, under the constraint that A chooses e in an incentive compatible and individually rational manner. We thus solve the following minimization

problem, which we call $(P(e))$.

$$\min_{w(\cdot)} \sum_{y \in Y} w(y)\pi(y|e)$$

subject to (1) and (2). Note that $(P(e))$ is a convex problem if u is concave.

We call $e \in E$ is *implementable* iff the constraint set of the minimization problem $(P(e))$ is nonempty; that is, a contract $w(\cdot)$ exists such that (1) and (2) hold. We also say that a contract *implements* e if it satisfies (1) and (2). While implementability does not generally ensure existence of a solution of $(P(e))$, it does when we assume u is concave.[6] Let $V(e)$ be the value of $(P(e))$. Then P will implement e that maximizes $\sum_{y \in Y} R(e,y)\pi(y|e) - V(e)$ over the set of implementable actions. This is the second step of the Grossman-Hart decomposition.

Here we concentrate on the first step and discuss when a given action $e \in E$ is implementable. Hermalin and Katz (1991) offer a clearest answer to this problem, which utilizes the existence theorem for the linear inequality system by Fan (1956). We state the result without a proof, because it is well-known.

Theorem 1 *Fix $e \in E$ and let $E' = E \setminus \{e\}$. e is implementable if and only if there exists no $\{\lambda(e_i)\}_{e_i \in E'}$, where $\lambda(e_i) \geq 0$ for any $e_i \in E'$, such that*

$$\sum_{e_i \in E'} \lambda(e_i)\pi(y|e_i) = \pi(y|e) \tag{3}$$

for any $y \in Y$, and

$$\sum_{e_i \in E'} \lambda(e_i)c(e_i) < c(e). \tag{4}$$

Theorem 1 has a clear economic interpretation. To see that, note first that if there exist positive weights, $\{\lambda(e_i)\}_{e_i \in E'}$, such that (3) and (4) hold, we have $\sum_{e_i \in E'} \lambda(e_i) = 1$, by summing (3) up over y. Hence $\{\lambda(e_i)\}_{e_i \in E'}$ can be interpreted as a mixed strategy of the agent. (3) implies that the mixed strategy induces the same probability distribution of the signal as e, and therefore results in the same expected payment from P, given any possible contract. Since (4) holds, this outcome-equivalent deviation costs less to the agent than e, and therefore A is better off by choosing the mixed strategy given any contract. Thus no contract gives A an incentive to choose e, so that e is not implementable. What makes this result surprising is the only-if part; namely, no existence of such outcome-equivalent and cost-saving deviation is sufficient for implementability.

Using Theorem 1, Hermalin and Katz (1991) provide the following simple sufficient condition of implementability, which they call the "convex-hull condition." We omit the proof, since it is trivial.

---

[6]Furthermore, strict concavity of u implies uniqueness of the solution.

Proposition 1 *Fix $e \in E$ and let $E' = E \setminus \{e\}$. e is implementable if $\pi(\cdot|e)$ is not in the convex hull of $\{\pi(\cdot|e')\}_{e' \in E'}$.*

Using Proposition 1, we can prove, for example, that e is implementable if there exists $y \in Y$ such that e uniquely maximizes (or uniquely minimizes) $\pi(y|\cdot)$. This observation is in particular useful when E and Y satisfy the monotone likelihood ratio property (MLRP). Then, for the "best" signal y, $\pi(y|\cdot)$ is maximized at the maximum level of efforts, which implies implementability of that level of efforts.

As the above argument indicates, Theorem 1 is a convenient tool to determine implementability of a certain action, but it requires some knowledge about the convexity structure of the model. If the signal space is very rich, it may be difficult to determine what the convex hull of $\{\pi(\cdot|e')\}_{e' \in E'}$ is and how it is related to $\pi(\cdot|e)$. This is why we believe introducing an alternative implementability condition is useful. The following result, Theorem 2, provides such an alternative. As will be seen in the subsequent sections, this condition can be conveniently applied to various agency models, especially when we are interested in attainability of first-best outcomes.

Theorem 2 *Fix $e \in E$ and let M be the set of probability distributions on E. e is implementable if and only if there exist two functions, $g : M \to R$ and $h : R \to R$, such that*

(i) For any $m \in M$ such that $\sum_{e_i \in E} m(e_i)\pi(y|e_i) = \pi(y|e)$ for any $y \in Y$, where $m(e_i)$ is the probability of e_i under m, we have $g(e) = g(m)$.[7]

(ii) h is strictly increasing, and

(iii) e solves

$$\max_{m \in M}[g(m) - h(\sum_{e_i \in E} m(e_i)c(e_i))]. \tag{5}$$

Proof. Suppose e is implementable, and let $\{w(y)\}_{y \in Y}$ be a contract that satisfies (1) and (2). For $m \in M$, define $g(m) = \sum_{y \in Y} \sum_{e_i \in E} w(y)m(e_i)\pi(y|e_i)$. It is easy to see that g satisfies (i). Hence if we set $h(x) = x$, for which (ii) is trivially satisfied, then (1) immediately proves that (iii) is also satisfied. This completes the proof of the "only-if" part.

Next, fix $e \in E$ and suppose that there exist g and h such that (i)-(iii) hold. Define $E' = E \setminus \{e\}$ and suppose that there exists $\lambda = \{\lambda(e_i)\}_{e_i \in E'}$, where $\lambda(e_i) \geq 0$ for any $e_i \in E'$, such that (3) and (4) hold. Then $m \in M$ exists such that $m(e) = 0$, $\sum_{e_i \in E'} m(e_i)\pi(y|e_i) = \pi(y|e)$ for any $y \in Y$, and that $\sum_{e_i \in E'} m(e_i)c(e_i) < c(e)$. Hence by (i) and (ii), e cannot be a solution of (5), a contradiction. Consequently, there cannot exist any $\lambda \geq 0$ satisfying (3) and (4). Hence by Theorem 1, e is implementable.

---

[7]We abuse notation so that e means a particular element of M.

As the above proof suggests, the idea of Theorem 2 is not very different from that of Theorem 1. However, its apparent statement is quite different, because it is stated in terms of a maximization problem. While the main difficulty in applying Theorem 2 to implementability problems is to find out functions g and h with the above properties, we will see in the subsequent sections that the task might not be so difficult, especially when we are interested in implementability of the first-best actions in agency and other related models.

4   The Standard Agency Model

In this section, we demonstrate how the new implementability condition, Theorem 2, can be applied to the standard agency model. Throughout this section, we make an additional assumption that the ex post revenue of the principal depends solely on y. Thus we write $R(e, y) = R(y)$. Therefore, the agent's action has an influence on the principal's payoff only through $\pi(\cdot|e)$.

As usual, the first-best action is defined as the solution of the following problem.

$$\max_{e \in E}[\sum_{y \in Y} R(y)\pi(y|e) - u^{-1}(c(e))]. \tag{6}$$

The underlying idea is that in the first-best world, P can write a contract that is directly dependent on e. Then P would pay the amount to the agent that leaves him no surplus if A chooses e, and would collect a large amount of fine otherwise. By so doing, P obtains exactly the value of the objective function (6).

For the later purpose, it is instructive to compare (6) with the following problem.

$$\max_{m \in M}[\sum_{y \in Y} \sum_{e \in E} R(y)\pi(y|e)m(e) - u^{-1}(\sum_{e \in E} m(e)c(e))]. \tag{7}$$

Both (6) and (7) have the same objective function, although only pure actions are allowed in (6). Thus the value of (6) is no greater than that of (7). Indeed, it is often the case that the value of (7) is strictly greater than that of (6). To see that, suppose u is strictly concave and two actions maximize (6). Then the objective function of (7) is also strictly concave, and therefore any (non-degenerate) convex combination of the two first-best actions achieves a greater value.

Although the literature treats (6) as the description of the first-best world, the problem (7) would be more relevant if P can force A to play mixed strategies. In this case, the implicit scenario is that P has the ability to monitor not only the agent's action, but also realization of a randomization device on which the agent's behavior can depend. Therefore P can induce a mixed strategy m by writing a contract which pays $u^{-1}(\sum_{e \in E} m(e)c(e))$ to A if he conforms to m, and collects a large amount of fine otherwise. Before randomization, the expected

payoff of A when he follows m is zero, so that A is willing to accept the contract. Consequently, P obtains the value of the objective function (7).

We do not argue here which framework is a more relevant description of the first-best world. We just invoke the new problem (7), because it helps us understand when the first-best action (the solution of (6)) is implementable.

Proposition 2 *Let e be a first-best action. If e also solves (7), then it is implementable.*

Proof. Define $g(m) = \sum_{y \in Y} \sum_{e \in E} R(y)\pi(y|e)m(e)$ and $h = u^{-1}$, and let us apply Theorem 2. It is easily seen that g and h satisfy conditions (i) and (ii), respectively. Since e solves (7), condition (iii) also follows, and therefore e is implementable.

Proposition 2 states that if the first-best action is "super first-best" in the sense that it is still optimal if P can implement any mixed strategy, then the super first-best action is implementable. Section 5 will give an example which shows the difference between being first-best and being super first-best.

If u is strictly concave and differentiable, u^{-1} is also differentiable. Therefore, e solves (7) if the first-order condition is valid at e, because the objective function of (7) is strictly concave. If we write $v \equiv u^{-1}$, e solves (7) if and only if, for any $e' \in E \setminus \{e\}$,

$$\sum_{y \in Y} R(y)[\pi(y|e') - \pi(y|e)] - v'(c(e))[c(e') - c(e)] \leq 0. \tag{8}$$

One advantage of (8) is that we need not know the exact structure of π, as far as we know the structure of expected revenues of the principal as a function of the agent's action. Thus, if we define $B(e) = \sum_{y \in Y} R(y)\pi(y|e)$, all we have to know about the model in order to apply Proposition 2 is about B, u and c.

We do not claim that Proposition 2 is always a better tool to check implementability of the first-best action than Proposition 1. Indeed, as the argument just after Proposition 1 shows, if the MLRP holds and if the maximum level of efforts is first-best, then Proposition 1 immediately proves that the first-best action is implementable. Our point is that Proposition 2, or any other result that stems from Theorem 2, offers an alternative tool to determine implementability of (possibly first-best) actions.

The following result provides another sufficient condition of implementability of the first-best action.

Proposition 3 *Suppose the first-best action e is also the first-best action if the von-Neumann Morgenstern utility function of the agent, u, is replaced with the identity mapping of R. Then e is implementable.*

Proof. Define $g(m) = \sum_{y \in Y} \sum_{e \in E} R(y)\pi(y|e)m(e)$ and $h(x) = x$, and let us apply Theorem 2. It is easy to see that g and h satisfy conditions (i) and (ii), respectively. By assumption, condition (iii) also follows, and therefore e is implementable.

The assumption of Proposition 3 is most likely to be satisfied if the agent is almost risk-neutral, because then the first-best action is likely to remain the same, if we rather assume that the agent is risk-neutral.[8] Note that the above proof does not need the fact that e is first-best in the original agency problem. Thus it follows that the action that would be optimal if the agent were risk-neutral is implementable, whether it is also first-best in the original model or not.[9]

As was stated in the introduction, Propositions 2 and 3 are important when combined with the analysis of the role of renegotiation by Hermalin and Katz (1991). Assuming that the agent's action is observable to P but unverifiable, Hermalin and Katz (1991) consider two cases in which a renegotiation contract is offered after P observes e, either by P or by A, respectively. In the framework where P makes a renegotiation offer, Hermalin and Katz (1991) show first that the set of implementable actions is not affected by introduction of renegotiation, and then show that any action that is implementable without renegotiation is implementable at the first-best cost under renegotiation. Therefore, the principal's payoff is equal to the objective function of (6).

Hence the agency cost completely disappears if the first-best action is implementable. While Hermalin and Katz (1991) present the convex hull condition (Proposition 1) as a sufficient condition of implementability of an (possibly first-best) action, our results are considered as alternative sufficient conditions of implementability of the first-best action.

Another important result of Hermalin and Katz (1991) is that if A makes the renegotiation offer, the first-best outcome is always achieved. Thus, from a welfare viewpoint, it is better to let the agent make the renegotiation offer, in a weak sense. However, to the extent that the first-best action is implementable, the same outcome is attainable when P makes the offer, and therefore the format of the renegotiation game as to who proposes renegotiation does not matter. From this viewpoint, Propositions 2 and 3 also tell us when such an irrelevance result obtains.

A recent paper of Edlin and Hermalin (2001) considers the intermediate case where the renegotiation surplus is divided according to exogenously given nonzero bargaining powers of the parties. Then Edlin and Hermalin (2001) show that the first-best action is implementable at the first-best cost under

---

[8]However, the second-best contract that implements the first-best action would be affected by the exact shape of u.

[9]An alternative, more direct proof of Proposition 3 goes as follows. Since e is optimal for a risk-neutral agent, the "sell-the-firm" contract satisfies both (1) and (2). Hence the contract obtained by mapping the sell-the-firm contract by u^{-1}, implements e for the agent with u.

any allocation of bargaining powers at the renegotiation stage, if the first-best action is implementable without renegotiation and if it satisfies an additional requirement, which is closely related to the convex-hull condition. Our result is not directly applicable to their framework, because Propositions 2 and 3 alone will not tell us whether the additional condition is satisfied. Nonetheless, our method seems to complement the stronger irrelevance result by Edlin and Hermalin (2001) at least in some cases.

5 An Example

In this section, we present a specific example of the agency model, and examine how our condition of implementability of the first-best action, Proposition 2, works in comparison with the convex-hull condition. Indeed, the example has such simple structure that a direct application of Theorem 1 gives a complete characterization as to when the first-best action is implementable. Nevertheless, the example is helpful in understanding how our condition is used, and how it is different from the convex-hull condition.[10] In this respect, it is worthwhile to work with a simple example.

We have only three actions of the agent and two signals, which we write $E = \{e_1, e_2, e_3\}$ and $Y = \{y_1, y_2\}$. We assume $\pi(y_2|e_1) = 0.1$, $\pi(y_2|e_2) = 0.5$ and $\pi(y_2|e_3) = 0.9$; $u(x) = \sqrt{x}$; $c(e_1) = 0$, $c(e_2) = 1$ and $c(e_3) = \mu > 1$; and

$$R(e, y) = R(y) = \begin{cases} 5 & \text{if } y = y_2 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

Thus y_2 may be considered as a good signal, and a more costly action makes y_2 more likely. Note that the MLRP is satisfied in this example.

By solving (6), it is easy to show that the first-best action is e_2 if $\mu \geq \sqrt{3}$ and e_3 if $\mu \leq \sqrt{3}$. Since the convex-hull condition in Proposition 1 is satisfied for e_3, implementability of e_3 immediately follows, whether it is first-best or not. Thus Proposition 2 is not particularly useful when $\mu \leq \sqrt{3}$.

Hence let us assume $\mu > \sqrt{3}$, which implies that e_2 is the unique first-best action. Since

$$\pi(\cdot|e_2) = \frac{1}{2}\pi(\cdot|e_1) + \frac{1}{2}\pi(\cdot|e_3),$$

the convex-hull condition is not satisfied for e_2. By Theorem 1, however, we know that e_2 is implementable if

$$c(e_2) \leq \frac{1}{2}c(e_1) + \frac{1}{2}c(e_3).$$

As a result, e_2 is implementable in our example iff $\mu \geq 2$. It also follows that the first-best action is not implementable when $\mu \in (\sqrt{3}, 2)$.

---

[10]In fact, the example will tell us that the two conditions are mutually independent.

Let us compare this observation with what can be said about implementability of e_2 by applying Proposition 2. In order to apply the proposition, e_2 must solve (7). Since u is strictly concave and differentiable, we can use (8). If we set $e = e_2$ and $e' = e_1$, (8) reduces to $0 \leq 0$, which is always true. If we set $e = e_2$ and $e' = e_3$, then (8) reduces to $\mu \geq 2$, which is equivalent to implementability of e_2. Thus this example shows that Proposition 2 can be a powerful tool to determine when the first-best action is implementable.

However, we must admit that Proposition 2 somewhat works too well in this example. To see this, let us replace (9) in the above example with

$$ R(y) = \begin{cases} 10 & \text{if } y = y_2 \\ 0 & \text{otherwise} \end{cases} \tag{10} $$

Although this change does not have an effect on when e_2 is implementable, it does have an effect on when it is first-best. In the new example with (10), e_2 is first-best iff $\mu \geq \sqrt{5}$. Again, using (8), we can show that e_2 solves (7) iff $\mu \geq 3$. To sum up, while the first-best action e_2 is indeed implementable if $\mu \geq \sqrt{5}$, Proposition 2 covers only the case where $\mu \geq 3$.

6   Extensions

So far, we are concerned with applying our implementability condition to the standard agency model. In contrast, this section shows that it is applicable to other closely related models. In the following two subsections, we first consider an agency model with multiple agents, and then an adverse selection model studied by Riordan and Sappington (1988). As before, much emphasis is put on attainability of the first-best actions/outcomes in both models.

6.1   A Multi-Agency Model

We extend the basic model to the following multi-agency model, where there are only two agents (Agent 1 and Agent 2). Let E_i be the set of Agent i's actions, which is finite. Now we define $E = E_1 \times E_2$, which is the set of action profiles. Y is the (finite) set of signals, and the probability distribution of the signal given $e = (e_1, e_2) \in E$ is $\pi(y|e)$. We assume $\pi(y|e) > 0$ for any $y \in Y$ and $e \in E$.

Agent i's utility is $u_i(w_i) - c_i(e_i)$, where w_i is her income. Each u_i is continuous, strictly increasing and unbounded above and below, so that u_i^{-1} exists. Here P's revenue depends entirely on y, denoted by $R(y)$. Since there are two agents, now the contract is $\{w(y)\}_{y \in Y}$, where $w(y) = (w_1(y), w_2(y))$.

An action profile $e = (e_1, e_2)$ is *implementable* if there exists $\{w(y)\}_{y \in Y}$ such that

$$ \sum_{y \in Y} u_i(w_i(y))\pi(y|e) - c_i(e_i) \geq \sum_{y \in Y} u_i(w_i(y))\pi(y|(e_i', e_j)) - c_i(e_i') \tag{11} $$

for any i and any $e'_i \in E_i$.[11]

(11) reveals that implementability of an action profile $\hat{e}$ is equivalent to implementability of two actions in two different single-agency models. Namely, for each $i = 1, 2$, let $A(i; \hat{e}_j)$ be the agency model with the action set E_i, the signal set Y, the probability distribution $\pi(y|(e_i, \hat{e}_j))$ and the utility function $u_i(w_i) - c_i(e_i)$.[12] Then $\hat{e}$ is implementable iff each $\hat{e}_i$ is implementable in $A(i; \hat{e}_j)$.

Let $e^* = (e_1^*, e_2^*)$ be a first-best action profile. That is, it solves

$$\max_{e \in E} [\sum_{y \in Y} R(y)\pi(y|e) - u_1^{-1}(c_1(e_1)) - u_2^{-1}(c_2(e_2))].$$

We want to prove a counterpart of Proposition 2 in the model with two agents. Let M_i be the set of probability distributions on E_i, whose generic element is $m_i = \{m_i(e_i)\}_{e_i \in E_i}$. Let us consider the following problem.

$$\max_{m_1 \in M_1, m_2 \in M_2} [\sum_{y \in Y} \sum_{e \in E} R(y)\pi(y|e)m(e) - \sum_{i=1}^{2} u_i^{-1}(\sum_{e_i \in E_i} m_i(e_i)c_i(e_i))], \quad (12)$$

where $m(e) = m_1(e_1)m_2(e_2)$. For the same reason discussed in Section 4, e^* need not solve (12) in general.

Proposition 4 *Let e^* be a first-best action profile. If it also solves (12), then it is implementable.*

Proof. By assumption, each e_i^* solves

$$\max_{m_i \in M_i} [\sum_{y \in Y} \sum_{e \in E} R(y)\pi(y|(e_i, e_j^*))m_i(e_i)m_j(e_j^*) - u_i^{-1}(\sum_{e_i \in E_i} m_i(e_i)c_i(e_i))].$$

If we set

$$g(m_i) = \sum_{y \in Y} \sum_{e \in E} R(y)\pi(y|(e_i, e_j^*))m_i(e_i)m_j(e_j^*)$$

and $h = u_i^{-1}$ in each $A(i; e_j^*)$, Theorem 2 implies that each e_i^* is implementable. Therefore e^* is implementable.

As in the model with a single agent, Proposition 4 is important when we introduce renegotiation. Indeed, this is the subject of Ishiguro and Itoh (2001), who analyze the same type of two-agent model with a particular renegotiation format. Ishiguro and Itoh (2001) assume the agents can observe their actions before realization of the signal, while P cannot. Immediately after observing the other agent's action, one agent makes a renegotiation offer, which is accepted

---

[11]Here we ignore the individual rationality condition because we can always adjust w so that it is satisfied.

[12]We do not specify the revenue of P, because it is irrelevant for our purpose.

or rejected by the other agent and then by the principal. In a similar vein to Hermalin and Katz (1991), Ishiguro and Itoh (2001) show that the agency cost for an action profile which is implementable under no renegotiation vanishes if such renegotiation is possible.[13] Thus if the first-best action profile is implementable without renegotiation, the agency cost completely vanishes. Proposition 4 gives an answer as to when this is the case.

As we derive Proposition 4 as an analog of Proposition 2, we can easily prove an analog of Proposition 3 in this multi-agency setting. The extension is very easy, so we omit it here.

6.2   An Adverse Selection Model

The purpose of this subsection is to apply our implementability condition to some adverse selection models. Generally speaking, the application is not easy because the adverse selection model is very different from the moral hazard model in structure. An important exception exists, however, which is a model studied by Riordan and Sappington (1988), where the principal receives some signal about the agent's actual type after the agent chooses efforts. We introduce their model and discuss attainability of the first-best outcome, which is also the subject of Riordan and Sappington (1988), using the method developed in this paper.

Let us start from description of an adapted version of the model by Riordan and Sappington (1988). P hires A, whose type is unknown to P. Let Θ be the set of the agent's types, assumed to have n elements and denoted by $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$. The agent, given his type and a contract (to be defined later), chooses an action from the set $X = \{x_1, x_2, \cdots, x_n\}$. We assume $x \in X$ is observable and verifiable. It costs the agent $c(x; \theta)$ if his type is $\theta \in \Theta$ and his action is $x \in X$. Note that both X and Θ have exactly n elements. The idea is that X is a subset of a larger action set, each element of which corresponds to the action P wants a certain type of A to choose from the larger set.[14] Since x is verifiable, P can prevent A from choosing an action outside X, by setting a severe penalty for such a deviation.

After choosing the action, and before the payment is made, P observes a signal about A's type. As in Section 2, let Y be the set of signals, which is assumed to be finite. We denote the probability of signal $y \in Y$ given that the agent is of type θ by $\mu(y|\theta)$.[15] We assume $\mu(y|\theta) > 0$ for any $y \in Y$ and $\theta \in \Theta$. The signal is also verifiable, so that P can write a contract contingent on both

---

[13]Precisely speaking, their result is based on the assumptions that the signal is two-dimensional ($y = (y_1, y_2)$), y_1 and y_2 are independent, and the marginal distribution of y_i depends entirely on e_i. Hence the signals are independent performance signals in their model.

[14]Thus we deliberately exclude the case in which P wants some types of the agent to choose the same action.

[15]The implicit assumption is that A cannot change the distribution of the signal by changing his action. Thus we do not allow any garbling of the signal by A.

$x \in X$ and $y \in Y$, which is the innovation of Riordan and Sappington (1988). Thus the contract is given by $w(x, y)$. Given a contract w, if x is chosen and y is observed, the payoff of the agent of type θ is $w(x, y) - c(x; \theta)$, and the payoff of the principal is $R(x, y) - w(x, y)$. Note that P's revenue depends on both x and y, but not on θ, while A's cost depends on both x and θ, but not on y. Finally, both parties are assumed to be risk-neutral (hence no risk issue in this model), and the agent's reservation utility is zero, irrespective of his type.

We are interested in whether P can induce A with type θ_i to choose x_i, for any i. The answer hinges on whether there exists a contract w such that

$$\sum_{y \in Y} \mu(y|\theta_i) w(x_i, y) - c(x_i; \theta_i) \geq \sum_{y \in Y} \mu(y|\theta_i) w(x, y) - c(x; \theta_i) \qquad (13)$$

for any i and $x \in X$, and

$$\sum_{y \in Y} \mu(y|\theta_i) w(x_i, y) - c(x_i; \theta_i) \geq 0 \qquad (14)$$

for any i. Unlike (2), here we cannot set the individual rationality condition to be always binding, because some type of the agent may enjoy information rents. We say w is *feasible* if it satisfies both (13) and (14).

Fix w, and define

$$v_i^* = \sum_{y \in Y} \mu(y|\theta_i) w(x_i, y) - c(x_i; \theta_i) \qquad (15)$$

for each i. Using (15), we define the following agency problem for each i, denoted by $A(i; w)$, such that

$$E = \Theta, \qquad (16)$$

$$\pi(y|\theta) = \mu(y|\theta), \qquad (17)$$

$$c(\theta_j) = c(x_i; \theta_j) + v_j^* \qquad (18)$$

and $u(x) = x$.[16] By (16), the expressions in the left-hand-sides of (17) and (18) are relevant.

The following result establishes an important link between feasibility of the given contract w and implementability of some actions in the agency problems, $A(i; w)$'s.

Proposition 5 *Let w be given, and consider $A(i; w)$ for each i. Then w is feasible iff for each i, the contract defined by $w^i = \{w(x_i, y)\}_{y \in Y}$ implements the action θ_i in $A(i; w)$.*

---

[16] Again we do not specify the revenue of P, since it is irrelevant.

Proof. Note that implementability of θ_i by w^i in each $A(i; w)$ is equivalent to

$$\sum_{y \in Y} w(x_i, y)\pi(y|\theta_i) - c(\theta_i) \geq \sum_{y \in Y} w(x_i, y)\pi(y|\theta_j) - c(\theta_j) \tag{19}$$

for any i and any j, and

$$\sum_{y \in Y} w(x_i, y)\pi(y|\theta_i) - c(\theta_i) = 0 \tag{20}$$

for any i. By (15), (17) and (18), (19) and (20) are equivalent to (13) and (14). Hence feasibility of w is equivalent to implementability of θ_i by w^i in each $A(i; w)$, which completes the proof.

Now we turn to a special case of the above adverse selection problem, in which each x_i is a solution of

$$\max_{x \in X}[\sum_{y \in Y} R(x, y)\mu(y|\theta_i) - c(x; \theta_i)]. \tag{21}$$

Thus, x_i is the efficient action given that the type of the agent is θ_i, in the sense that it maximizes the joint surplus.[17] We say *the first-best outcome is implementable* if there exists a contract w which is feasible and satisfies (14) with equality for any i. Thus implementability of the first-best outcome requires not only efficiency, but also full extraction of surplus.

Let $A^*(i)$ be the agency problem obtained by setting $v_j^* = 0$ for any j in $A(i; w)$ (any other aspect of $A(i; w)$ remains the same). The analysis of those moral hazard problems, $A^*(i)$'s, turns out to be crucial in examining implementability of the first-best outcome in our adverse selection model. To see this "duality," we give a new proof of the main result of Riordan and Sappington (1988), which provides a necessary and sufficient condition of implementability of the first-best outcome. While Riordan and Sappington (1988) prove the same result with help of the mathematics Hermalin and Katz (1991) used for the proof of Theorem 1, the idea of our proof is to utilize the relationship with $A^*(i)$.

Theorem 3 *The first-best outcome is implementable iff for each i, there does not exist an $(n+1)$-dimensional vector $\lambda \geq 0$ such that*

$$\lambda_{n+1}\mu(y|\theta_i) = \sum_{k=1}^{n} \lambda_k \mu(y|\theta_k) \tag{22}$$

for any $y \in Y$, and

$$\lambda_{n+1}c(x_i; \theta_i) - \sum_{k=1}^{n} \lambda_k c(x_i; \theta_k) = 1. \tag{23}$$

---

[17]If we instead consider that A can choose from a larger set $X' \supset X$, then each x_i would be the solution of (21) over the larger set X'. Again, we exclude the case where the same action solves (21) for two different types.

Proof. Fix i. If some $\lambda \geq 0$ satisfies (22) and (23), then summing up (22) over $y \in Y$ yields $\lambda_{n+1} \geq \lambda_i$. If $\lambda_{n+1} = \lambda_i$, then $\lambda_j = 0$ for any $j \neq i$, and therefore (23) cannot be true. Thus we have $\lambda_{n+1} > \lambda_i$.

Hence no existence of such $\lambda \geq 0$ is equivalent to no existence of $\eta = \{\eta_j\}_{j \neq i}$, where $\eta_j \geq 0$ for any $j \neq i$, such that

$$\mu(y|\theta_i) = \sum_{j \neq i} \eta_j \mu(y|\theta_j) \tag{24}$$

for any $y \in Y$, and

$$c(x_i; \theta_i) > \sum_{j \neq i} \eta_j c(x_i; \theta_j). \tag{25}$$

By Theorem 1, no existence of $\eta \geq 0$ such that (24) and (25) hold is equivalent to implementability of θ_i in $A^*(i)$.

The proof is complete if we show that implementability of θ_i in $A^*(i)$ for any i is equivalent to implementability of the first-best outcome. Suppose each θ_i is implementable in $A^*(i)$, and let $\{w(x_i, y)\}_{y \in Y}$ be a contract that implements it. Then, proceeding in the same way as the proof of Proposition 5, we can show that the contract $w(x, y)$ satisfies (13) and (14), with (14) being equality. Hence the first-best outcome is implementable. In contrast, suppose the first-best outcome is implementable, and let $w(x, y)$ be a corresponding contract. Since (14) holds with equality for any i, we have $v_i^* = 0$ for any i. Thus by Proposition 5, for each i, θ_i is implementable in $A^*(i)$, which completes the proof.

The above proof also establishes the following result as a corollary, which is another necessary and sufficient condition of implementability of the first-best outcome. We will use it in the subsequent analysis.

Proposition 6 *The first-best outcome is implementable iff for any i, θ_i is implementable in $A^*(i)$.*

Now we explore an economic sufficient condition of implementability of the first-best outcome, relying on our method. The following condition is essential in the analysis.

Condition 1 Let us define

$$B(x; \theta) = \sum_{y \in Y} R(x, y)\mu(y|\theta) - c(x; \theta). \tag{26}$$

Then we have

$$\min_i B(x_i; \theta_i) \geq \max_{i, j \neq i} B(x_i; \theta_j). \tag{27}$$

17

Roughly speaking, Condition 1 is satisfied if the situation is a coordination problem. Namely, P (or the regulator, in the original model of Riordan and Sappington (1988)) does not know which action fits the state of the world (known only to the agent, which is the regulated monopolist in their framework) best, but she knows that any coordination failure results in the worse social surplus than successful coordination at the worst state.[18]

Proposition 7 *The first-best outcome is implementable if Condition 1 holds.*

Proof. Fix i, and consider $A^*(i)$. Let M be the set of probability distributions on Θ. If we set $g(m) = \sum_{y \in Y} \sum_{\theta \in \Theta} R(x_i, y)m(\theta)\mu(y|\theta)$ and $h(z) = z$, then it is easy to see that conditions (i) and (ii) of Theorem 2 are satisfied. Since θ_i maximizes $B(x_i; \theta)$ over Θ by (27), (26) implies that θ_i maximizes $g(m) - h(\sum_{\theta \in \Theta} m(\theta)c(x_i; \theta))$ on M. Hence condition (iii) of Theorem 2 is also satisfied. Therefore, by Theorem 2, θ_i is implementable in $A^*(i)$. Since i is arbitrary, implementability of the first-best outcome immediately follows from Proposition 6.

It is interesting to compare this result with the economic sufficient conditions of implementability of the first-best outcome obtained by Riordan and Sappington (1988). Their sufficient conditions (Corollaries 1.4 and 1.5, Riordan and Sappington (1988)) are closely related to the convexity structure of the model, such as convexity of $c(x; \theta)$ with respect to θ, and so on. This is not surprising because their analysis is based on the same mathematics that derives Theorem 1. In contrast, we are able to provide another sufficient condition that is much more related to the underlying welfare maximization problem.

---

[18]Condition 1 is hard to interpret in the original framework by Riordan and Sappington (1988), where the type of the agent (the monopolist) corresponds to the degree of production efficiency. Thus Condition 1 presumes a different situation, but note that nothing in the description of the model limits us to the issue of regulating the monopolist. Indeed, even in the context of regulation, Condition 1 is relevant if the variable the government wants to control is coordination-oriented, such as the location of its plant and/or outlet. While the government wants the firm to locate where products are much needed from a social point of view, the firm would choose the location based on private incentives.

References

d'Aspremont, C., and L.-A. Gerard-Varet (1979): "Incentives and Incomplete Information," *Journal of Public Economics*, 11, 25-45.

d'Aspremont, C., and L.-A. Gerard-Varet (1998): "Linear Inequality Methods to Enforce Partnerships under Uncertainty: An Overview," *Games and Economic Behavior*, 25, 311-336.

Edlin, A., and B. Hermalin (2001): "Implementing the First Best in an Agency Relationship with Renegotiation: A Corrigendum," *Econometrica*, 69, 1391-1395.

Fan, K. (1956): "On Systems of Linear Inequalities," in *Linear Inequalities and Related Systems*, ed. by H. Kuhn and A. Tucker. Princeton, NJ: Princeton University Press.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1039.

Grossman, S., and O. Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, 7-45.

Hermalin, B., and M. Katz (1991): "Moral Hazard and Verifiability: The Effects of Renegotiation in Agency," *Econometrica*, 59, 1735-1753.

Ishiguro, S., and H. Itoh (2001): "Moral Hazard and Renegotiation with Multiple Agents," *Review of Economic Studies*, 68, 1-20.

Legros, P., and H. Matsushima (1991): "Efficiency in Partnerships," *Journal of Economic Theory*, 55, 296-322.

Riordan, M., and D. Sappington (1988): "Optimal Contracts with Public ex post Information," *Journal of Economic Theory*, 45, 189-199.