

# ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

*An International Society for the Advancement of Economic  
Theory in its Relation to Statistics and Mathematics*

<http://www.econometricsociety.org/>

*Econometrica*, Vol. 78, No. 1 (January, 2010), 285–308

## MEDIATED PARTNERSHIPS

DAVID RAHMAN

*University of Minnesota, Minneapolis, MN 55455, U.S.A.*

ICHIRO OBARA

*University of California, Los Angeles, Los Angeles, CA 90096, U.S.A.*

---

The copyright to this Article is held by the Econometric Society. It may be downloaded, printed and reproduced only for educational or research purposes, including use in course packs. No downloading or copying may be done for any commercial purpose without the explicit permission of the Econometric Society. For such commercial purposes contact the Office of the Econometric Society (contact information may be found at the website <http://www.econometricsociety.org> or in the back cover of *Econometrica*). This statement must be included on all copies of this Article that are made available electronically or in any other format.

---

## MEDIATED PARTNERSHIPS

BY DAVID RAHMAN AND ICHIRO OBARA<sup>1</sup>

This paper studies partnerships that employ a mediator to improve their contractual ability. Intuitively, profitable deviations must be attributable, that is, there must be some group behavior such that an individual can be statistically identified as innocent, to provide incentives in partnerships. Mediated partnerships add value by effectively using different behavior to attribute different deviations. As a result, mediated partnerships are necessary to provide the right incentives in a wide range of economic environments.

KEYWORDS: Mediated contracts, partnerships, private monitoring.

### 1. INTRODUCTION

PROVIDING INCENTIVES IN PARTNERSHIPS is a classic topic of economic theory.<sup>2</sup> Although it is well known that communication is a basic facet of incentive provision (Aumann (1974), Forges (1986), Myerson (1986)), this insight has not been systematically applied to partnership problems. This paper adds to the literature by asking the following question. Consider a group of individuals whose behavior is subject to moral hazard, but who have rich communication and contractual protocols: (i) a disinterested mediator who can make confidential, verifiable but nonbinding recommendations to agents, and (ii) budget-balanced payment schemes<sup>3</sup> that may depend on both the mediator's recommendations and individual reports. What outcomes can this group enforce?

Our main result (Theorem 1) shows that identifying obedient agents (IOA) is both necessary and sufficient for every outcome to be virtually enforceable<sup>4</sup> in this mediated environment, regardless of preferences. IOA means that for any profile of deviations, there is some behavior by the agents that statistically identifies an innocent individual after any unilateral deviation in the profile. IOA enjoys the following crucial property: different behavior may be used to attribute innocence after different deviations.

Let us intuitively explain this result. On the one hand, providing incentives with budget balance requires punishing some agents and rewarding others si-

<sup>1</sup>Many thanks are owed to Harold Demsetz, Larry Jones, Michihiro Kandori, Narayana Kocherlakota, David Levine, Roger Myerson, Itai Sher, Joe Ostroy, Phil Reny, Joel Sobel, Bill Zame, a co-editor, and four anonymous referees for help with previous drafts. We are also grateful to numerous seminar audiences. D. Rahman gratefully acknowledges financial support from the Spanish Ministry of Education Grant SEJ 2004-07861 while at Universidad Carlos III de Madrid and the National Science Foundation Grant SES 09-22253.

<sup>2</sup>See Alchian and Demsetz (1972), Holmström (1982), Radner, Myerson, and Maskin (1986), Legros and Matsushima (1991), Legros and Matthews (1993), d'Aspremont and Gérard-Varet (1998), and others.

<sup>3</sup>Budget balance means that the sum of payments across individuals always equals zero.

<sup>4</sup>An outcome is "virtually enforceable" if there is an enforceable outcome arbitrarily close to it.

multaneously. If, after a unilateral deviation, an innocent party cannot be identified, then the deviator could have been anyone, so the only way to discourage the deviation is to punish everyone. However, this violates budget balance. On the other hand, IOA implies that budget-balanced incentives can be provided by rewarding the innocent and punishing all others. To prove this, we establish and take advantage of the following observation. Rich contractual protocols enable the use of payments that differ after different recommended actions. We show that effectively, to reward the innocent after a given deviation profile, different behavior may be used to find such innocent parties. But this is just the definition of IOA. Without rich contractual protocols, the same payments must be made after every recommendation, and we show that as a result, the same behavior must be used to identify the innocent.

The value of mediated partnerships over ordinary ones (Theorems 2 and 4) now follows. Without payment schemes contingent on recommendations, it is possible to provide incentives by rewarding the innocent only if the same behavior is used to attribute innocence after every deviation. The difference between this requirement and the clearly less stringent IOA characterizes the value of mediated partnerships. As it turns out, mediated partnerships provide incentives in many natural environments where incentives would otherwise fail. For instance, for generic distributions of output, mediated partnerships can provide incentives<sup>5</sup> even without production complementarities,<sup>6</sup> yet ordinary ones cannot (Example 1).<sup>7</sup>

This paper adds to the literature (Section 6) in two basic ways. First, it extends the work of Legros and Matthews (1993), who derived nearly efficient partnerships in restricted environments with output-contingent contracts. Although they noted that identifying the innocent is important for budget-balanced incentives, they did not address statistical identification and did not use different behavior to identify the innocent after different deviations. Second, being necessary for Theorem 1, IOA exhausts the informational economies from identifying the innocent rather than the guilty.<sup>8</sup> This contrasts with the literature on repeated games, where restricted communication protocols were used by Kandori (2003) and others to prove the Folk theorem.<sup>9</sup> Such papers typically require a version of pairwise full rank (Fudenberg, Levine, and Maskin (1994)), which intuitively means identifying the deviator after every

<sup>5</sup>See the working paper version of this paper for a proof of genericity.

<sup>6</sup>See Legros and Matthews (1993, Example B) to enforce partnerships with complementarities.

<sup>7</sup>For example, we do not require that the distribution of output has a “moving support,” that is, the support of the distribution depends on individual behavior. This assumption, made by Legros and Matthews (1993), is not generic, so an arbitrarily small change in probabilities leads to its failure.

<sup>8</sup>Heuristically, knowing who deviated implies knowing someone who did not deviate, but knowing someone who did not deviate does not necessarily imply knowing who did.

<sup>9</sup>See Section 6 for a more detailed discussion of this literature.

deviation. This is clearly more restrictive than IOA, which only requires identifying a nondeviator.

The paper is organized as follows. Section 2 presents a motivating example where a mediated partnership is virtually enforced, yet none of the papers above applies. Section 3 presents the model and main definitions. Section 4 states our main results, discussed above. Section 5 refines our main assumptions in the specific context of public monitoring and studies participation as well as liability constraints. Section 6 reviews the literature on contract theory and repeated games, and compares it to this paper. Finally, Section 7 concludes. Proofs appear in the [Appendix](#).

## 2. EXAMPLE

We begin our analysis of mediated partnerships with an example to capture the intuition behind our main result, Theorem 1. The example suggests the following intuitive way to attain a “nearly efficient” partnership: appoint a *secret principal*.

**EXAMPLE 1:** Consider a fixed group of  $n$  individuals. Each agent  $i$  can either work ( $a_i = 1$ ) or shirk ( $a_i = 0$ ). Let  $c > 0$  be each individual’s cost of effort. Effort is not observable. Output is publicly verifiable and can be either good ( $g$ ) or bad ( $b$ ). The probability of  $g$  equals  $P(\sum_i a_i)$ , where  $P$  is a strictly increasing function of the sum of efforts. Finally, assume that each individual  $i$ ’s utility function equals  $z_i - ca_i$ , where  $z_i$  is the amount of money received by  $i$ .

[Radner, Myerson, and Maskin \(1986\)](#) introduced this partnership in the context of repeated games. They considered the problem of providing incentives for everyone to work—if not all the time, at least most of the time—without needing to inject or withdraw resources from the group as a whole. They effectively showed that in this environment there do not exist output-contingent rewards that both (i) balance the group’s budget, that is, the sum of individual payments always equals zero, and (ii) induce everyone to work most of the time, let alone all of the time. Indeed, for everyone to work at all, they must be rewarded when output is good. However, this arrangement violates budget balance, since everyone being rewarded when output is good clearly implies that the sum of payments across agents is greater when output is good than when it is bad.

An arrangement that still does not solve the partnership problem, but nevertheless induces most people to work, is appointing an agent to play the role of [Holmström](#)’s principal. Call this agent 1 and define output-contingent payments to individuals as follows. For  $i = 2, \dots, n$ , let  $\zeta_i(g) = \bar{z}$  and  $\zeta_i(b) = 0$  be

agent  $i$ 's output-contingent money payment for some  $\bar{z} \geq 0$ . To satisfy budget balance, agent 1's transfer equals

$$\zeta_1 = - \sum_{i=2}^n \zeta_i.$$

By construction, the budget is balanced. It is easy to see that everyone but agent 1 will work if  $\bar{z}$  is sufficiently large. However, agent 1 has the incentive to shirk.<sup>10</sup>

With mediated contracts, it is possible to induce everyone to work most of the time. Indeed, consider the following incentive scheme. For any small  $\varepsilon > 0$ , a mediator or machine asks every individual to work (call this event  $\mathbf{1}$ ) with probability  $1 - \varepsilon$ . With probability  $\varepsilon/n$ , agent  $i$  is picked (assume everyone is picked with equal probability) and secretly asked to shirk, while all others are asked to work (call this event  $\mathbf{1}_{-i}$ ). For  $i = 1, \dots, n$ , let  $\zeta_i(g|\mathbf{1}) = \zeta_i(b|\mathbf{1}) = 0$  be agent  $i$ 's contingent transfer if the mediator asked everyone to work. Otherwise, if agent  $i$  was secretly asked to shirk, for  $j \neq i$ , let  $\zeta_j(g|\mathbf{1}_{-i}) = \bar{z}$  and  $\zeta_j(b|\mathbf{1}_{-i}) = 0$  be agent  $j$ 's transfer. For agent  $i$ , let

$$\zeta_i(\mathbf{1}_{-i}) = - \sum_{j \neq i} \zeta_j(\mathbf{1}_{-i}).$$

By construction, this contract is budget-balanced. It is also incentive compatible. Indeed, it is clear that asking an agent to shirk is always incentive compatible. If agent  $i$  is recommended to work, incentive compatibility requires that

$$\frac{\varepsilon(n-1)}{n} P(n-1) \bar{z} - \left[ \frac{\varepsilon(n-1)}{n} + (1-\varepsilon) \right] c \geq \frac{\varepsilon(n-1)}{n} P(n-2) \bar{z},$$

which is satisfied if  $\bar{z}$  is sufficiently large because  $P$  is strictly increasing.<sup>11</sup> Under this contract, everyone works with probability  $1 - \varepsilon$ , for any  $\varepsilon > 0$ , by choosing  $\bar{z}$  appropriately, so everyone working is approximated with budget-balanced transfers.

The arrangement above solves the partnership problem of Radner, Myerson, and Maskin (1986) by occasionally appointing a secret principal. To induce everyone to work, this contract effectively appoints a different principal for different workers. Appointing the principals secretly allows for them to be used simultaneously. Finally, they are chosen only seldom to reduce the inherent loss from having a principal in the first place.

<sup>10</sup>This contract follows Holmström's suggestion to the letter: agent 1 is a "fixed" principal who absorbs the incentive payments to all others by "breaking" everyone else's budget constraint.

<sup>11</sup>Here,  $\frac{\varepsilon(n-1)}{n} + (1-\varepsilon)$  is the probability that an agent is asked to work and  $\frac{\varepsilon(n-1)}{n}$  is the probability that, in addition, someone else was appointed the secret principal.

Example 1 reveals the logic behind our main result, Theorem 1. If a worker deviates (i.e., shirks), then he will decrease the probability of  $g$  not only when everyone else is asked to work, but also when a principal is appointed. In this latter case, innocence can be attributed to the principal, so the deviator can be punished by having every worker pay the principal. In other words, for each worker and any deviation by the worker there is a profile of actions by others such that his deviation can be statistically distinguished from someone else's (in this case, a principal, since the principal's deviation would raise the probability of  $g$ ). This turns out to be not only necessary but also sufficient for solving any partnership problem.

### 3. MODEL

This section develops our model of mediated partnerships. It describes the environment, the timing of agents' interaction, notions of enforcement, and attribution.

Let  $I = \{1, \dots, n\}$  be a finite set of agents, let  $A_i$  be a finite set of actions available to any agent  $i \in I$ , and let  $A = \prod_i A_i$  be the (nonempty) space of action profiles. Write  $v: I \times A \rightarrow \mathbb{R}$  for the profile of agents' utility functions, where  $v_i(a)$  denotes the utility to any agent  $i \in I$  from any action profile  $a \in A$ . A *correlated strategy* is any probability measure  $\sigma \in \Delta(A)$ .<sup>12</sup> Let  $S_i$  be a finite set of *private signals* observable only by agent  $i \in I$  and let  $S_0$  be a finite set of *publicly verifiable* signals. Let  $S := \prod_{j=0}^n S_j$  be the (nonempty) space of all signal profiles. A *monitoring technology* is a measure-valued map  $\text{Pr}: A \rightarrow \Delta(S)$ , where  $\text{Pr}(s|a)$  denotes the conditional probability that signal profile  $s$  was observed given that action profile  $a$  was played.

We model rich communication protocols by introducing a disinterested *mediator* who fulfills two roles: (i) making confidential recommendations to agents over what action to take and (ii) revealing the entire profile of recommendations publicly at the end of the game. This mediator may be seen as a proxy for any preplay communication among the players (Aumann (1987)).

Incentives are provided to agents with linear transfers. An *incentive scheme* is any map  $\zeta: I \times A \times S \rightarrow \mathbb{R}$  that assigns monetary payments contingent on individuals, recommended actions, and reported signals, all of which are assumed verifiable.

DEFINITION 1: A *contract* is any pair  $(\sigma, \zeta)$ , where  $\sigma$  is a correlated strategy and  $\zeta$  is an incentive scheme. It is called *standard* if  $\zeta_i(a, s)$  is not a function of  $a$ , that is, payments do not depend on recommendations; otherwise, the contract is called *mediated*.

<sup>12</sup>If  $X$  is a finite set,  $\Delta(X) = \{\mu \in \mathbb{R}_+^X : \sum_x \mu(x) = 1\}$  is the set of probability vectors on  $X$ .

Standard contracts are a special case of mediated ones, but not otherwise. For instance, the secret principal of Section 2 is a nonstandard mediated contract, since payments depend on recommendations. The literature has mostly focused on standard contracts to study incentives, whereas this paper concentrates on mediated ones.

It is important to emphasize that a standard contract does not do away with the mediator altogether—only as regards payments. Indeed, as will be seen below and was suggested in Example 1 above, we emphasize using the mediator not so much to correlate behavior, but rather to correlate payoffs so as to provide incentives.

The timing of agents’ interaction unfolds as follows. First, agents agree on some contract  $(\sigma, \zeta)$ . A profile of recommendations is drawn according to  $\sigma$  and made to agents confidentially by some mediator. Agents then simultaneously take some action, which is neither verifiable nor directly observable. Next, agents observe unverifiable private signals and submit a verifiable report of their observations before observing the public signal (the timing of signals is not essential, just simplifying). Finally, recommendation- and report-contingent transfers are made according to  $\zeta$ .

Specifically, we assume that agents report their private signals simultaneously, and consider contracts where agents willingly behave honestly (report truthfully) and obediently (follow recommendations). In other words, strategic behavior is assumed to constitute a *communication equilibrium*, as in Myerson (1986) and Forges (1986), of the game induced by a given contract  $(\sigma, \zeta)$ .

If every agent is honest and obedient, agent  $i$ ’s expected utility from  $(\sigma, \zeta)$  is

$$\sum_{a \in A} \sigma(a) v_i(a) - \sum_{(a,s)} \sigma(a) \zeta_i(a, s) \Pr(s|a).$$

Of course, agent  $i$  may disobey his recommendation  $a_i$  to play some other action  $b_i$  and lie about his privately observed signal. A *reporting strategy* is a map  $\rho_i : S_i \rightarrow S_i$ , where  $\rho_i(s_i)$  is the reported signal when  $i$  privately observes  $s_i$ . For instance, the truthful reporting strategy is the identity map  $\tau_i : S_i \rightarrow S_i$  with  $\tau_i(s_i) = s_i$ . Let  $R_i$  be the set of all reporting strategies for agent  $i$ . For every agent  $i$  and every pair  $(b_i, \rho_i) \in A_i \times R_i$ , the conditional probability that  $s \in S$  will be reported when everyone else is honest and plays  $a_{-i} \in A_{-i}$  equals<sup>13</sup>

$$\Pr(s|a_{-i}, b_i, \rho_i) := \sum_{t_i \in \rho_i^{-1}(s_i)} \Pr(s_{-i}, t_i|a_{-i}, b_i).$$

A contract  $(\sigma, \zeta)$  is *incentive compatible* if obeying recommendations and reporting honestly is optimal for every agent when everyone else is honest and

<sup>13</sup>We use the notation  $s = (s_{-i}, s_i)$  for  $s_i \in S_i$  and  $s_{-i} \in S_{-i} = \prod_{j \neq i} S_j$ ; similarly for  $a = (a_{-i}, a_i)$ .

obedient, that is,  $\forall i \in I, a_i \in A_i, (b_i, \rho_i) \in A_i \times R_i,$

$$\begin{aligned}
 (*) \quad & \sum_{a_{-i}} \sigma(a)(v_i(a_{-i}, b_i) - v_i(a)) \\
 & \leq \sum_{(a_{-i}, s)} \sigma(a) \zeta_i(a, s) (\Pr(s|a_{-i}, b_i, \rho_i) - \Pr(s|a)).
 \end{aligned}$$

The left-hand side of (\*) reflects the *utility gain*<sup>14</sup> for an agent  $i$  from playing  $b_i$  when asked to play  $a_i$ . The right-hand side reflects his *monetary loss* from playing  $(b_i, \rho_i)$  relative to honesty and obedience. Such a loss originates from two sources. On the one hand, playing  $b_i$  instead of  $a_i$  may change conditional probabilities over signals. On the other, reporting according to  $\rho_i$  may affect conditional payments.

DEFINITION 2: A correlated strategy  $\sigma$  is *exactly enforceable* (or simply *enforceable*) if there is an incentive scheme  $\zeta : I \times A \times S \rightarrow \mathbb{R}$  to satisfy (\*) for all  $(i, a_i, b_i, \rho_i)$  and

$$(**) \quad \forall (a, s), \quad \sum_{i \in I} \zeta_i(a, s) = 0.$$

Call  $\sigma$  *virtually enforceable* if there exists a sequence  $\{\sigma^m\}$  of enforceable correlated strategies such that  $\sigma^m \rightarrow \sigma$ .

A correlated strategy is enforceable if there is a budget-balanced<sup>15</sup> incentive scheme that makes it incentive compatible. It is virtually enforceable if it is the limit of enforceable correlated strategies. This requires budget balance along the way, not just asymptotically. For instance, in Example 1, everybody working is virtually enforceable, but not exactly enforceable.

We end this section by defining a key condition called identifying obedient players, which will be shown to characterize enforcement. We begin with some preliminaries.

A *strategy* for any agent  $i$  is a map  $\alpha_i : A_i \rightarrow \Delta(A_i \times R_i)$ , where  $\alpha_i(b_i, \rho_i|a_i)$  stands for the probability that  $i$  reacts by playing  $(b_i, \rho_i)$  when recommended to play  $a_i$ . For any  $\sigma$  and any  $\alpha_i$ , let  $\Pr(\sigma, \alpha_i) \in \Delta(S)$ , defined pointwise by

$$\Pr(s|\sigma, \alpha_i) = \sum_{a \in A} \sigma(a) \sum_{(b_i, \rho_i)} \Pr(s|a_{-i}, b_i, \rho_i) \alpha_i(b_i, \rho_i|a_i),$$

be the vector of report probabilities if agent  $i$  deviates from  $\sigma$  according to  $\alpha_i$ .

<sup>14</sup>Specifically, utility gain is probability-weighted, weighted by  $\sigma(a_i) = \sum_{a_{-i}} \sigma(a)$ , the probability of  $a_i$ .

<sup>15</sup>Budget balance means here that the sum of payments across individuals always equals zero. Some authors use budget balance to mean that payments add up to the value of some output. On the other hand, our model may be interpreted as using utilities that are net of profit shares.



DEFINITION 3: A strategy profile  $\alpha = (\alpha_1, \dots, \alpha_n)$  is *unattributable* if

$$\forall a \in A, \quad \Pr(a, \alpha_1) = \dots = \Pr(a, \alpha_n).^{16}$$

Call  $\alpha$  *attributable* if it is not unattributable, that is, there exist agents  $i$  and  $j$  such that  $\Pr(a, \alpha_i) \neq \Pr(a, \alpha_j)$  for some  $a \in A$ .

Intuitively, a strategy profile  $\alpha$  is unattributable if a unilateral deviation from honesty and obedience by any agent  $i$  to a strategy  $\alpha_i$  in the profile would lead to the same conditional distribution over reports. Heuristically, after a deviation (from honesty and obedience) belonging to some unattributable profile, even if the fact that someone deviated was detected, anyone could have been the culprit.

Call  $\alpha_i$  *disobedient* if  $\alpha_i(b_i, \rho_i | a_i) > 0$  for some  $a_i \neq b_i$ , that is, it disobeys some recommendation with positive probability. A disobedient strategy may be “honest,” that is,  $\rho_i$  may equal  $\tau_i$ . However, dishonesty by itself (obeying recommendations but choosing  $\rho_i \neq \tau_i$ ) is not labeled as disobedience. A *disobedient strategy profile* is any  $\alpha = (\alpha_1, \dots, \alpha_n)$  such that  $\alpha_i$  is disobedient for at least one agent  $i$ .

DEFINITION 4: A monitoring technology *identifies obedient agents* (IOA) if every disobedient strategy profile is attributable.

IOA means that for every disobedience by some arbitrary agent  $i$  and every profile of others’ strategies, an action profile exists such that  $i$ ’s unilateral deviation has a different effect on report probabilities from at least one other agent. For instance, the monitoring technology of Example 1 identifies obedient agents. There, if a worker shirks, then good news becomes *less* likely, whereas if a principal works, then good news becomes *more* likely. Hence, a strategy profile with  $i$  disobeying is attributable by just having another agent behave differently from  $i$ . This implies IOA. Intuitively, IOA holds by using different principals for different workers.

#### 4. RESULTS

This section presents the paper’s main results, characterizing enforceable outcomes in terms of the monitoring technology, with and without mediated contracts. We begin with a key lemma that provides a dual characterization of IOA.

LEMMA 1: *A monitoring technology identifies obedient agents if and only if there exists a function  $\xi: I \times A \times S \rightarrow \mathbb{R}$  such that  $\sum_i \xi_i(a, s) = 0$  for every*

<sup>16</sup>We slightly abuse notation by identifying action profiles with pure correlated strategies.

$(a, s)$  and

$$\forall (i, a_i, b_i, \rho_i), \quad 0 \leq \sum_{(a_{-i}, s)} \xi_i(a, s) (\Pr(s|a_{-i}, b_i, \rho_i) - \Pr(s|a))$$

with a strict inequality whenever  $a_i \neq b_i$ .

Intuitively, Lemma 1 shows that IOA is equivalent to the existence of budget-balanced “probability-weighted” transfers  $\xi$  such that (i) the budget is balanced, (ii) no deviation is profitable, and (iii) every disobedience incurs a strictly positive monetary cost. If every action profile is recommended with positive probability, that is, if  $\sigma \in \Delta^0(A) := \{\sigma \in \Delta(A) : \sigma(a) > 0 \forall a \in A\}$  is any completely mixed correlated strategy, then there is an incentive scheme  $\zeta$  with  $\xi_i(a, s) = \sigma(a)\zeta_i(a, s)$  for all  $(i, a, s)$ . Therefore, IOA implies that given  $\sigma \in \Delta^0(A)$  and  $\xi$  satisfying Lemma 1, for any profile  $v$  of agents’ utility functions, we may scale  $\zeta$  appropriately to overcome all incentive constraints simultaneously. Hence, the second half of Lemma 1 implies that every completely mixed correlated strategy is exactly enforceable, regardless of the utility profile. Approximating each correlated strategy with completely mixed ones establishes half of our main result, Theorem 1 below. The other half argues that if IOA fails, then there exist a profile of utility functions and a correlated strategy that is not virtually enforceable. In this sense, IOA is the weakest condition on a monitoring technology that—independently of preferences—guarantees virtual enforcement.

**THEOREM 1:** *A monitoring technology identifies obedient agents if and only if for any profile of utility functions, every correlated strategy is virtually enforceable.*

Theorem 1 characterizes monitoring technologies such that “everything” is virtually enforceable, regardless of preferences. It says that identifying obedient agents in a weak sense is not only necessary, but also sufficient for virtual enforcement. Intuitively, if, after a disobedience, some innocent agent can be statistically identified then that agent can be rewarded at the expense of everyone else, thereby punishing the deviator. Heuristically, if a strategy profile can be attributed, then there is an incentive scheme that discourages every strategy in that profile. Theorem 1 says that for every disobedient strategy profile there is a scheme that discourages it if and only if there is a scheme that discourages all disobedient strategy profiles simultaneously.

To put Theorem 1 in perspective, consider the scope of enforcement with standard contracts. By Example 1, IOA is generally not enough for enforcement with standard contracts, but the following strengthening is. Given a subset  $B \subset A$  of action profiles and an agent  $i$ , let  $B_i := \{b_i \in A_i : \exists b_{-i} \in A_{-i} \text{ s.t. } b \in B\}$  be the projection of  $B$  on  $A_i$ . Call a strategy  $\alpha_i$  *B-disobedient* if it is disobedient at some  $a_i \in B_i$ , that is, if  $\alpha_i(b_i, \rho_i|a_i) > 0$  for some  $b_i \neq a_i \in B_i$ . A *B-disobedient strategy profile* is any  $\alpha = (\alpha_1, \dots, \alpha_n)$  such that  $\alpha_i$  is *B-disobedient* for some agent  $i$ . Given  $\sigma \in \Delta(A)$ ,  $\alpha$  is *attributable at  $\sigma$*  if there exist agents  $i$

and  $j$  such that  $\Pr(\sigma, \alpha_i) \neq \Pr(\sigma, \alpha_j)$ , and say  $\Pr$  identifies obedient agents at  $\sigma$  (IOA- $\sigma$ ) if every  $\text{supp } \sigma$ -disobedient<sup>16</sup> strategy profile is attributable at  $\sigma$ . Intuitively, IOA- $\sigma$  differs from IOA in that IOA allows for different  $\alpha$ 's to be attributed at different  $\sigma$ 's, whereas IOA- $\sigma$  does not.

**THEOREM 2:** *A monitoring technology identifies obedient agents at  $\sigma$  if and only if for any profile of utility functions,  $\sigma$  is exactly enforceable with a standard contract.*

Theorem 2 characterizes enforceability with standard contracts of any correlated strategy  $\sigma$  in terms of IOA- $\sigma$ . Intuitively, it says that enforcement with standard contracts requires that every  $\alpha$  be attributable at the same  $\sigma$ .<sup>17</sup> Theorem 2 also sheds light onto the value of mediated contracts. Indeed, the proof of Theorem 1 shows that enforcing a completely mixed correlated strategy (i.e., such that  $\sigma(a) > 0$  for all  $a$ ) only requires IOA, by allowing for different strategy profiles to be attributable at different action profiles. This condition is clearly weaker than IOA- $\sigma$ . On the other hand, IOA is generally not enough to enforce a given pure-strategy profile  $a$ , as Example 1 shows with  $a = \mathbf{1}$  there. Since agents receive only one recommendation under  $a$ , there is no use for mediated contracts, so by Theorem 2, IOA- $a$  characterizes exact enforcement of  $a$  with both standard and mediated contracts.<sup>18</sup>

Now consider the intermediate case where  $\sigma$  has arbitrary support. Fix a subset of action profiles  $B \subset A$ . A strategy profile  $\alpha = (\alpha_1, \dots, \alpha_n)$  is  $B$ -attributable if there exist agents  $i$  and  $j$  such that  $\Pr(a, \alpha_i) \neq \Pr(a, \alpha_j)$  for some  $a \in B$ . Otherwise,  $\alpha$  is called  $B$ -unattributable. For instance,  $A$ -attribution is just attribution. Say  $\Pr$   $B$ -identifies obedient agents ( $B$ -IOA) if every  $B$ -disobedient strategy profile is  $B$ -attributable. For instance,  $A$ -IOA is just IOA and  $\{a\}$ -IOA equals IOA- $a$ .

**THEOREM 3:** *For any subset  $B \subset A$ , the following statements are equivalent:*  
 (i) *The monitoring technology  $B$ -identifies obedient agents.* (ii) *Every correlated strategy with support equal to  $B$  is enforceable for any profile of utility functions.*  
 (iii) *Some fixed correlated strategy with support equal to  $B$  is enforceable for any profile of utility functions.*

Theorem 3 characterizes enforcement with mediated contracts of any correlated strategy  $\sigma$  with  $\text{supp } \sigma$ -IOA. Hence, only the support of a correlated strategy matters for its enforcement for all preferences. Moreover, any other correlated strategy with support contained in  $\text{supp } \sigma$  becomes virtually en-

<sup>16</sup>By definition,  $\text{supp } \sigma = \{a \in A : \sigma(a) > 0\}$  is the support of  $\sigma$ .

<sup>17</sup>Even for virtual enforcement with standard contracts, the same  $\sigma$  must attribute all  $\alpha$ 's. For example, in Example 1 there is no sequence  $\{\sigma^m\}$  with  $\sigma^m \rightarrow \mathbf{1}$  and  $\Pr$  satisfying IOA- $\sigma^m$  for all  $m$ .

<sup>18</sup>Again, we abuse notation by labeling  $a$  as both an action profile and a pure correlated strategy.

forceable, just as with Theorem 1. Intuitively, mediated contracts allow for different actions in the support of a correlated strategy to attribute different strategy profiles, unlike standard contracts, as shown above. Therefore, clearly IOA- $\sigma$  is more restrictive than  $\text{supp } \sigma$ -IOA.

Although the results above focused on enforcement for all utility profiles, restricting attention to fixed preferences does not introduce additional complications and yields similar results. Indeed, fix a profile  $v: I \times A \rightarrow \mathbb{R}$  of utility functions. A natural weakening of IOA involves allowing unprofitable strategy profiles to be unattributable. A strategy profile  $\alpha$  is called  $\sigma$ -profitable if

$$\sum_{(i,a,b_i,p_i)} \sigma(a)\alpha_i(b_i, p_i|a_i)(v_i(a_{-i}, b_i) - v_i(a)) > 0.$$

Intuitively, the profile  $\alpha$  is  $\sigma$ -profitable if the sum of each agent’s utility gains from a unilateral deviation in the profile is positive. Enforcement now amounts to the following declarations.

**THEOREM 4:** (i) *Every  $\sigma$ -profitable strategy profile is  $\text{supp } \sigma$ -attributable if and only if  $\sigma$  is enforceable.* (ii) *Every  $\sigma$ -profitable strategy profile is attributable at  $\sigma$  if and only if  $\sigma$  is enforceable with a standard contract.*

Theorem 4 characterizes enforceability with and without mediated contracts. It describes how mediated contracts add value by relaxing the burden of attribution: Every profile  $\alpha$  that is attributable at  $\sigma$  is  $\text{supp } \sigma$ -attributable, but not conversely. For instance, in Example 1, let  $\sigma(S)$  be the probability that  $S \subset I$  are asked to work, and suppose that  $\sigma(I) > 0$ . Let  $\alpha$  be the strategy profile where every agent  $i$  shirks with probability  $p_i$  if asked to work (and obeys if asked to shirk), with  $p_i = \sigma(I)[P(n - 1) - P(n)] / \sum_{S \ni i} \sigma(S)[P(|S| - 1) - P(|S|)] \in (0, 1]$ . By construction, the probability of good output equals  $\sigma(I)P(n - 1) + \sum_{S \neq I} \sigma(S)P(|S|)$ , which is independent of  $i$ . Therefore,  $\alpha$  is not attributable at any  $\sigma$  with  $\sigma(I) > 0$ . However,  $\alpha$  is attributable, since the monitoring technology identifies obedient agents.

## 5. DISCUSSION

In this section we decompose IOA to understand it better under the assumption of public monitoring. We also consider participation and liability constraints.

### 5.1. Public Monitoring

To help understand IOA, let us temporarily restrict attention to *publicly verifiable* monitoring technologies, that is, such that  $|S_i| = 1$  for all  $i \neq 0$ . In this case, IOA can be naturally decomposed into two parts. We formalize this decomposition next.

A strategy  $\alpha_i$  for any agent  $i$  is *detectable* if  $\Pr(a, \alpha_i) \neq \Pr(a)$  at some  $a \in A$ . Say  $\Pr$  *detects unilateral disobedience* (DUD) if every disobedient strategy is detectable,<sup>19</sup> where different action profiles may be used to detect different strategies. Say *detection implies attribution* (DIA) if for every detectable strategy  $\alpha_i$  and every strategy profile  $\alpha_{-i}$ ,  $\alpha = (\alpha_{-i}, \alpha_i)$  is attributable. Intuitively, DIA says that if a strategy is detected, someone can be (statistically) ruled out as innocent.

**THEOREM 5:** *A publicly verifiable monitoring technology identifies obedient agents if and only if (i) it detects unilateral disobedience and (ii) detection implies attribution.*

An immediate example of DIA is [Holmström’s \(1982\)](#) principal, that is, an individual  $i_0$  with no actions to take or signals to observe (both  $A_{i_0}$  and  $S_{i_0}$  are singletons). The principal is automatically obedient, so every detectable strategy can be discouraged with budget balance by rewarding him and punishing everyone else. DIA isolates this idea and finds when the principal’s role can be fulfilled internally. It helps to provide budget-balanced incentives by identifying innocent individuals to be rewarded and punishing all others (if necessary) when a disobedience is detected.

Next, we give a dual characterization of DIA that sheds light onto the role it plays in [Theorem 1](#). A publicly verifiable monitoring technology  $\Pr$  satisfies *incentive compatibility implies enforcement* (ICE) if for every  $K: A \times S \rightarrow \mathbb{R}$ , there exists  $\xi: I \times A \times S \rightarrow \mathbb{R}$  such that

$$\forall(a, s), \quad \sum_{i \in I} \xi_i(a, s) = K(a, s),$$

$$\forall(i, a_i, b_i), \quad 0 \leq \sum_{(a_{-i}, s)} \xi_i(a, s)(\Pr(s|a_{-i}, b_i) - \Pr(s|a)).$$

The function  $K(a, s)$  may be regarded as a budgetary surplus or deficit for each combination of recommended action and realized signal. Intuitively, ICE means that any budget can be attained by some payment scheme that avoids disrupting any incentive compatibility constraints. As it turns out, this is equivalent to DIA.

**THEOREM 6:** *Given any publicly verifiable monitoring technology, detection implies attribution if and only if incentive compatibility implies enforcement.*

This result helps to clarify the roles of DUD and DIA in [Theorem 1](#). [Rahman \(2008\)](#) showed that DUD characterizes virtual enforcement without budget

<sup>19</sup>This condition on a monitoring technology was introduced and analyzed by [Rahman \(2008\)](#).

balance of any correlated strategy  $\sigma$ , regardless of preferences. ICE guarantees the existence of a further contract to absorb any budgetary deficit or surplus of the original contract without violating any incentive constraints. Therefore, the original contract plus this further contract can now virtually enforce  $\sigma$  with a balanced budget.<sup>20</sup>

If the monitoring technology is not publicly verifiable, DUD plus DIA is sufficient but unnecessary for IOA. Necessity fails in general because there may exist dishonest but obedient strategies that IOA allows to remain unattributable even if detectable, as the next example shows.<sup>21</sup>

EXAMPLE 2: There are three agents and  $A_i$  is a singleton for every agent  $i$ , so IOA is automatically satisfied. There are no public signals and each agent observes a binary private signal:  $S_i = \{0, 1\}$  for all  $i$ . The monitoring technology is

$$\Pr(s) := \begin{cases} \frac{6}{25}, & \text{if } \sum_i s_i = 3, \\ \frac{3}{25}, & \text{if } \sum_i s_i = 1 \text{ or } 2, \\ \frac{1}{25}, & \text{if } \sum_i s_i = 0. \end{cases}$$

The following equation is a profile of (trivially obedient) unattributable strategies that are also detectable, violating DIA. Suppose that agent  $i$  deviates by lying with probability  $2/5$  after observing  $s_i = 1$  and lying with probability  $3/5$  after observing  $s_i = 0$ . For every agent  $i$ , the joint distribution of reported private signals becomes

$$\Pr(s) = \begin{cases} \frac{27}{125}, & \text{if } \sum_i s_i = 3, \\ \frac{18}{125}, & \text{if } \sum_i s_i = 2, \\ \frac{12}{125}, & \text{if } \sum_i s_i = 1, \\ \frac{8}{125}, & \text{if } \sum_i s_i = 0. \end{cases}$$

<sup>20</sup>A comparable argument was provided by d’Aspremont, Cremer, and Gérard-Varet (2004) for Bayesian mechanisms.

<sup>21</sup>Without a publicly verifiable monitoring technology, IOA is equivalent to DUD plus “disobedient detection implies attribution,” that is, every disobedient and detectable strategy is attributable. However, this latter condition lacks an easily interpreted dual version as in Theorem 6.

5.2. Participation and Liability

Individual rationality—or participation—constraints are easily incorporated into the present study of incentives by imposing the family of inequalities

$$\forall i \in I, \quad \sum_{a \in A} \sigma(a)v_i(a) - \sum_{(a,s)} \sigma(a)\zeta_i(a, s) \Pr(s|a) \geq 0.$$

**THEOREM 7:** *Participation is not a binding constraint if  $\sum_i v_i(a) \geq 0$  for all  $a \in A$ .*

Theorem 7 generalizes standard results (e.g., d’Aspremont and Gérard-Varet (1998, Lemma 1)) to our setting.

Next, we study limited liability given  $z \in \mathbb{R}_+^I$ , by imposing constraints of the form  $\zeta_i(a, s) \geq -z_i$ . Intuitively, an agent can never pay any more than  $z_i$ . Call  $z_i$  agent  $i$ ’s liability, and call  $z$  the distribution of liability. A group’s total liability is defined by  $\hat{z} = \sum_i z_i$ . Without participation constraints, Theorem 5 of Legros and Matsushima (1991) and Theorem 4 of Legros and Matthews (1993) easily generalize to this setting.

**THEOREM 8:** *In the absence of participation constraints, only total liability affects the set of enforceable outcomes, not the distribution of liability.*

Including participation constraints leads to the following characterization.

**THEOREM 9:** *The correlated strategy  $\sigma$  is enforceable with individual rationality and liability limited by  $z$  if and only if*

$$\begin{aligned} & \sum_{(a,i,b_i,\rho_i)} \sigma(a)\alpha_i(b_i, \rho_i|a_i)(v_i(a_{-i}, b_i) - v_i(a)) \\ & \leq \sum_{i \in I} \pi_i(v_i(\sigma) - z_i) + \hat{\eta} \sum_{i \in I} z_i \end{aligned}$$

for every  $(\alpha, \pi)$  such that  $\alpha$  is a strategy profile and  $\pi = (\pi_1, \dots, \pi_n) \geq 0$ , where  $\hat{\eta} := \sum_{(a,s)} \min_i \{\Pr(s|a, \alpha_i) - (1 + \pi_i) \Pr(s|a)\}$  and  $v_i(\sigma) = \sum_a \sigma(a)v_i(a)$ .

Theorem 9 generalizes Theorems 7 and 8, as the next result shows.

**COROLLARY 1:** *Suppose that  $\sigma$  is enforceable with individual rationality and liability limited by  $z$ . (i) If  $v_i(\sigma) \geq z_i$ , then agent  $i$ ’s participation is not a binding constraint. (ii) The distribution of liability does not matter within the subset  $t$  of agents whose participation constraint is not binding, that is,  $\sigma$  is also enforceable with individual rationality and liability limited by any  $z'$  with  $z_j = z'_j$  for  $j \in I \setminus t$  and  $\sum_{i \in t} z_i = \sum_{i \in t} z'_i$ .*

## 6. LITERATURE

To help identify this paper's contribution, let us now compare its results with the literature. Broadly, the paper contributes (i) a systematic analysis of partnerships that fully exploit internal communication, and (ii) results that show that attribution and IOA yield the weakest requirements on a monitoring technology for enforcement and virtual enforcement. IOA enjoys the key property that different action profiles can be used to attribute different disobedient strategy profiles, in contrast with the literature, which we discuss below.

In contract theory, Legros and Matsushima (1991) characterized exact enforcement with standard contracts and publicly verifiable signals, but they did not interpret their results in terms of attribution, nor did they consider virtual enforcement. Another related paper is d'Aspremont and Gérard-Varet (1998). In the same context as Legros and Matsushima (1991), they derived intuitive sufficient conditions for enforcement. A closer paper to ours is by Legros and Matthews (1993), who studied virtual enforcement with standard contracts and deterministic output. They proposed a contract that uses mixed strategies to identify nonshirkers whenever possible,<sup>22</sup> but the same correlated strategy must identify nonshirkers after every deviation, unlike mediated contracts. Their contract fails to provide the right incentives if output given efforts is stochastic and its distribution does not have a "moving support," that is, the support does not depend on efforts. The key difference between their contract and ours is that mediated partnerships correlate agents' payoffs not just to output, but also to others' mixed strategies. As a result, mediated partnerships can virtually enforce efficient behavior even without a moving support, as Example 1 and Theorem 1 show.<sup>23</sup>

In the context of repeated games, the closest papers to ours may be Kandori (2003), Aoyagi (2005), and Tomala (2009). They establish versions of the Folk theorem by interpreting players' continuation values as linear transfers. Kandori allowed agents to play mixed strategies and reported on the realization of such mixtures after observing a public signal. He considered contracts contingent on the signals and these reports.<sup>24</sup> Although his contracts are non-standard, they fail to fully employ communication. For instance, they fail to provide incentives in Example 1. Aoyagi used dynamic mediated strategies that rely on " $\epsilon$ -perfect" monitoring and fail if monitoring is costly or one-sided. Our results accommodate these issues. Finally, Tomala studied a class of recursive communication equilibria.

<sup>22</sup>A (stronger) form of identifying nonshirkers was suggested in mechanism design by Kosenok and Severinov (2008). However, they characterized full surplus extraction rather than enforcement.

<sup>23</sup>Fudenberg, Levine, and Maskin (1994) considered a form of virtual enforcement without a moving support. However, they required much stronger assumptions than ours, discussed momentarily.

<sup>24</sup>Obara (2008) extended Kandori's contracts to study full surplus extraction with moral hazard and adverse selection in the spirit of Cremer and McLean (1988), ignoring budget balance.



There are several differences between these papers and ours. One especially noteworthy difference is that to prove the Folk theorem they make much more restrictive assumptions than IOA, structurally similar to the *pairwise full rank* (PFR) of Fudenberg, Levine, and Maskin (1994). Intuitively, PFR-like conditions ask to identify deviators instead of just nondeviators. To see this, let us focus for simplicity on public monitoring and recall the decomposition of IOA into DUD and DIA (Theorem 5).

For every  $i$ , let  $C_i$  (called the *cone* of agent  $i$ ) be the set of all  $\eta \in \mathbb{R}^{A \times S}$  with,

$$\forall(a, s), \quad \eta(a, s) = \sum_{b_i \in A_i} \alpha_i(b_i | a_i) (\Pr(s | a_{-i}, b_i) - \Pr(s | a))$$

for some  $\alpha_i : A_i \rightarrow \Delta(A_i)$ . DIA imposes on agents' cones the restriction

$$\bigcap_{i \in I} C_i = \{\mathbf{0}\},$$

where  $\mathbf{0}$  stands for the origin of  $\mathbb{R}^{A \times S}$ .

In other words, agents' cones do not overlap. PFR implies that for *every pair* of agents, their cones do not overlap. Intuitively, this means that upon any deviation, it is possible to identify the deviator's identity. On the other hand, DIA only requires that *all* agents' cones fail to overlap simultaneously. Thus, it is possible to provide budget-balanced incentives even if there are two agents whose cones overlap (i.e., their intersection is larger than just the origin), so PFR fails. In general, DIA does not even require that there exist two agents whose cones fail to overlap, in contrast with *local compatibility* of d'Aspremont and Gérard-Varet (1998). Figure 1 illustrates this point.<sup>25</sup>

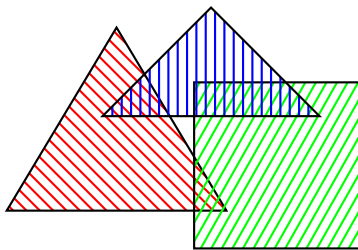


FIGURE 1.—A cross section of three nonoverlapping cones in  $\mathbb{R}^3$  (pointed at the origin behind the page) such that every pair of cones overlaps.

<sup>25</sup>Figure 1 is not pathological. Indeed, Example 1 may be viewed as a version of Figure 1.

7. CONCLUSION

Mediated partnerships embrace the idea that—as part of an economic organization—it may be beneficial for private information to be allocated differently across agents to provide the right incentives. As Example 1 illustrates, mediated partnerships can enforce outcomes that standard ones simply cannot. Indeed, mediated contracts can provide the right incentives in partnerships with stochastic output whose distribution fails to exhibit a “moving support” (i.e., the support is independent of effort), even without complementarities in production. Standard contracts cannot.

In general, mediated partnerships are enforceable if and only if it is possible to identify obedient agents. This means that after any unilateral deviation, innocence is statistically attributable to someone, although different actions may be used to attribute innocence after different deviations.<sup>26</sup> Informationally, this is clearly less costly than attempting to attribute guilt, as well as using the same actions to attribute innocence after every deviation. This latter difference exactly captures the value of mediated partnerships.

APPENDIX: PROOFS

PROOF OF LEMMA 1: With only finitely many actions and finitely many agents, the second half of the lemma holds if and only if there exists  $\xi$  such that  $\sum_i \xi_i(a, s) = 0$  for all  $(a, s)$  and

$$\forall(i, a_i, b_i, \rho_i), \quad \Delta_i(a_i, b_i) \leq \sum_{(a_{-i}, s)} \xi_i(a, s)(\Pr(s|a_{-i}, b_i, \rho_i) - \Pr(s|a)),$$

where  $\Delta_i(a_i, b_i) = 1$  if  $a_i \neq b_i$  and 0 otherwise. Consider the linear program that consists of choosing  $\xi$  to minimize 0 subject to the above constraints. The dual problem is to choose a vector  $(\lambda, \eta)$  such that  $\lambda \geq 0$  to maximize  $\sum_{(i, a_i, b_i, \rho_i)} \lambda_i(a_i, b_i, \rho_i) \Delta_i(a_i, b_i)$  subject to

$$\forall(i, a, s), \quad \sum_{(b_i, \rho_i)} \lambda_i(a_i, b_i, \rho_i)(\Pr(s|a_{-i}, b_i, \rho_i) - \Pr(s|a)) = \eta(a, s).$$

Here, the vector  $\lambda \geq 0$  collects the multipliers on incentive constraints and  $\eta$  collects those of the budget balance constraints. Since the dual is feasible (with  $(\lambda, \eta) = 0$ ), by the strong duality theorem (see, e.g., Schrijver (1986, p. 92)), the condition on  $\xi$  above fails if and only if there exists a dual feasible solution

<sup>26</sup>Although identifying obedient agents is impossible with only two agents and public monitoring, it holds generically in richer environments, even with just three agents or a minimal amount of public information. See the working paper version of this article for these results.

$(\lambda, \eta)$  such that  $\lambda_i(a_i, b_i, \rho_i) > 0$  for some  $(i, a_i, b_i, \rho_i)$  with  $a_i \neq b_i$ . Let  $\Lambda = \max_{(i, a_i)} \sum_{(b_i, \rho_i)} \lambda_i(a_i, b_i, \rho_i) > 0$  and define

$$\alpha_i(b_i, \rho_i | a_i) := \begin{cases} \lambda_i(a_i, b_i, \rho_i) / \Lambda, & \text{if } (b_i, \rho_i) \neq (a_i, \tau_i), \\ 1 - \sum_{(b_i, \rho_i) \neq (a_i, \tau_i)} \lambda_i(a_i, b_i, \rho_i) / \Lambda, & \text{otherwise.} \end{cases}$$

By construction,  $\alpha_i$  is disobedient and unattributable (using  $\alpha_{-i}$ ): IOA fails. *Q.E.D.*

**PROOF OF THEOREM 1:** Sufficiency follows from the paragraph preceding the statement of the theorem. For necessity, suppose that IOA fails, that is, there is a disobedient profile  $\alpha = (\alpha_1, \dots, \alpha_n)$  that is also unattributable. Let  $a^* \in A$  be an action profile where  $\alpha$  is disobedient, that is, there exists an agent  $i^*$  such that  $\alpha_{i^*}(b_{i^*}, \rho_{i^*} | a_{i^*}^*) > 0$  for some  $b_{i^*} \neq a_{i^*}^*$ . Let  $v_i(a) = 0$  for all  $(i, a)$  except for  $v_{i^*}(b_{i^*}, a_{-i}^*) = 1$ . Consider any correlated strategy  $\sigma$  that places positive probability on  $a^*$ . For a contradiction, assume that there is a payment scheme  $\zeta$  that enforces  $\sigma$ . Summing the incentive constraints at  $a^*$  across agents and using budget balance together with the definition of  $v$ , we obtain

$$0 < \sigma(a^*) = \sum_{(i, b_i, a_{-i})} \sigma(a_i^*, a_{-i}) (v_i(b_i, a_{-i}) - v_i(a_i^*, a_{-i})) \leq 0.$$

Therefore,  $\sigma$  is not enforceable. Finally, this implies that  $a^*$  is not virtually enforceable. *Q.E.D.*

**PROOF OF THEOREM 2:** The proof follows that of Lemma 1. By the strong duality theorem,  $\text{Pr}$  satisfies IOA- $\sigma$  if and only if there exists a payment scheme  $\zeta: I \times S \rightarrow \mathbb{R}$  that only depends on reported signals for each agent such that  $\sum_i \zeta_i(s) = 0$  for all  $s$  and

$$\begin{aligned} \forall i \in I, a_i \in B_i, (b_i, \rho_i) \in A_i \times R_i, \\ 0 \leq \sum_{(a_{-i}, s)} \sigma(a) \zeta_i(s) (\text{Pr}(s | a_{-i}, b_i, \rho_i) - \text{Pr}(s | a)), \end{aligned}$$

with a strict inequality if  $a_i \neq b_i$ , where  $B_i = \{a_i \in A_i : \exists a_{-i} \text{ s.t. } \sigma(a) > 0\}$ . Call this dual condition IOA\*- $\sigma$ . By scaling  $\zeta$  as necessary, IOA\*- $\sigma$  clearly implies that any deviation gains can be outweighed by monetary losses. Conversely, if IOA- $\sigma$  fails, then there is a profile of deviation plans  $\alpha$  such that  $\text{Pr}(\sigma, \alpha_i) = \text{Pr}(\sigma, \alpha_j)$  for all  $(i, j)$  and there is an agent  $i^*$  such that  $\alpha_{i^*}$  satisfies  $\alpha_{i^*}(b_{i^*}, \rho_{i^*} | a_{i^*}^*) > 0$  for some  $a_{i^*} \in B_{i^*}$  and  $b_{i^*} \neq a_{i^*}^*$ . For all  $a_{-i^*}$ , let  $0 = v_{i^*}(a) < v_{i^*}(a_{-i^*}, b_{i^*}) = 1$  and  $v_j(a) = v_j(a_{-i^*}, b_{i^*}) = 0$  for all  $j \neq i^*$ . Now  $\sigma$

cannot be enforced by any  $\zeta : I \times S \rightarrow \mathbb{R}$  such that  $\sum_i \zeta_i(s) = 0$  for all  $s$ , since  $\sum_{(i,b_i,\rho_i)} \alpha_i(b_i, \rho_i|a_i) \sum_{a_{-i}} \sigma(a)(v_i(a_{-i}, b_i) - v_i(a)) > \sum_{(i,s)} \zeta_i(s)(\Pr(s|\sigma, \alpha_i) - \Pr(s|\sigma)) = 0$ , being a nonnegative linear combination of incentive constraints, will violate at least one. Q.E.D.

PROOF OF THEOREM 3: (i)  $\Leftrightarrow$  (ii) follows by applying a version of the proof of Lemma 1 and Theorem 1 after replacing  $B$  with  $A$ . (i)  $\Leftrightarrow$  (iii) follows similarly, after fixing any correlated strategy  $\sigma$  with support equal to  $B$ . Q.E.D.

PROOF OF THEOREM 4: (i) follows by applying the proof of Lemma 1 with both  $\sigma$  and  $v$  fixed to the incentive compatibility constraints (\*). (ii) follows by a similar version of the proof of Theorem 2, again with both  $\sigma$  and  $v$  fixed. Q.E.D.

PROOF OF THEOREM 5: IOA clearly implies DUD (just replace  $\alpha_{-i}$  with honesty and obedience for every  $\alpha_i$  in the definition of attribution). By IOA, if a profile  $\alpha$  is unattributable, then it is obedient, hence every deviation plan in the profile is undetectable (since the monitoring technology is publicly verifiable) and DIA follows. Conversely, DIA implies that every unattributable  $\alpha_i$  is undetectable, and by DUD, every undetectable  $\alpha_i$  is obedient. Q.E.D.

PROOF OF THEOREM 6: Consider the following primal problem: Find a feasible  $\xi$  to solve,

$$\forall(i, a_i, b_i), \quad 0 \leq \sum_{(a_{-i},s)} \xi_i(a, s)(\Pr(s|a_{-i}, b_i) - \Pr(s|a)),$$

and

$$\forall(a, s), \quad \sum_{i \in I} \xi_i(a, s) = K(a, s).$$

The dual of this problem is given by

$$\inf_{\lambda \geq 0, \eta} \sum_{(a,s)} \eta(a, s)K(a, s) \quad \text{s.t.}$$

$$\forall(i, a, s), \quad \sum_{b_i \in A_i} \lambda_i(a_i, b_i)(\Pr(s|a_{-i}, b_i) - \Pr(s|a)) = \eta(a, s).$$

If ICE is satisfied, then the value of the primal equals 0 for any  $K : A \times S \rightarrow \mathbb{R}$ . By the strong duality theorem, the value of the dual is also 0 for any  $K : A \times S \rightarrow \mathbb{R}$ . Therefore, any  $\eta$  satisfying the constraint for some  $\lambda$  must be 0 for all  $(a, s)$ , so DIA is satisfied.

For sufficiency, if DIA holds, then the value of the dual is always 0 for any  $K : A \times S \rightarrow \mathbb{R}$ . By strong duality, the value of the primal is also 0 for any  $K$ .

Therefore, given  $K$ , there is a feasible primal solution  $\xi_i(a, s)$  that satisfies all primal constraints, and ICE holds. *Q.E.D.*

**PROOF OF THEOREM 7:** We use the following notation. Given a correlated strategy  $\sigma$  and a deviation plan  $\alpha_i$ , let  $\Delta v_i(\sigma, \alpha_i) = \sum_{(a, b_i, \rho_i)} \sigma(a) \alpha_i(b_i, \rho_i | a_i) \times (v_i(a_{-i}, b_i) - v_i(a))$  be the utility gain from  $\alpha_i$  at  $\sigma$  and let  $\Delta \Pr(s|a, \alpha_i) = \sum_{(a, b_i, \rho_i)} \alpha_i(b_i, \rho_i | a_i) (\Pr(s|a_{-i}, b_i, \rho_i) - \Pr(s|a))$  be the change in the probability that  $s$  is reported from  $\alpha_i$  at  $a$ . Enforcing an arbitrary correlated strategy  $\sigma$  subject to participation constraints reduces to finding transfers  $\zeta$  to solve the family of linear inequalities

$$\begin{aligned} \forall(i, a_i, b_i, \rho_i), \quad & \sum_{a_{-i}} \sigma(a) (v_i(a_{-i}, b_i) - v_i(a)) \\ & \leq \sum_{(a_{-i}, s)} \sigma(a) \zeta_i(a, s) (\Pr(s|a_{-i}, b_i, \rho_i) - \Pr(s|a)), \\ \forall(a, s), \quad & \sum_{i=1}^n \zeta_i(a, s) = 0, \\ \forall i \in I, \quad & \sum_{a \in A} \sigma(a) v_i(a) - \sum_{(a, s)} \sigma(a) \zeta_i(a, s) \Pr(s|a) \geq 0. \end{aligned}$$

The dual of this problem subject to participation is

$$\begin{aligned} \max_{\lambda, \pi \geq 0, \eta} \quad & \sum_{i \in I} \Delta v_i(\sigma, \lambda_i) - \pi_i v_i(\sigma) \quad \text{s.t.} \\ \forall(i, a, s), \quad & \sigma(a) \Delta \Pr(s|a, \lambda_i) = \eta(a, s) + \pi_i \sigma \Pr(s|a), \end{aligned}$$

where  $\pi_i$  is a multiplier for agent  $i$ 's participation constraint and  $v_i(\sigma) = \sum_a \sigma(a) v_i(a)$ . Adding the dual constraints with respect to  $s \in S$ , it follows that  $\pi_i = \pi$  does not depend on  $i$ . Redefining  $\eta(a, s)$  as  $\eta(a, s) + \pi \Pr(s|a)$ , the set of feasible  $\lambda \geq 0$  is the same as without participation constraints. Since  $\sum_i v_i(a) \geq 0$  for all  $a$ , the dual is maximized by  $\pi = 0$ . *Q.E.D.*

**PROOF OF THEOREM 8:** We use the same notation as in the [proof](#) of Theorem 7. Let  $z = (z_1, \dots, z_n)$  be a vector of liability limits for each agent. Enforcing  $\sigma$  subject to limited liability reduces to finding  $\zeta$  such that

$$\begin{aligned} \forall(i, a_i, b_i, \rho_i), \quad & \sum_{a_{-i}} \sigma(a) (v_i(a_{-i}, b_i) - v_i(a)) \\ & \leq \sum_{(a_{-i}, s)} \sigma(a) \zeta_i(a, s) (\Pr(s|a_{-i}, b_i, \rho_i) - \Pr(s|a)), \end{aligned}$$

$$\forall(a, s), \quad \sum_{i=1}^n \zeta_i(a, s) = 0,$$

$$\forall(i, a, s), \quad \zeta_i(a, s) \leq z_i.$$

The dual of this metering problem subject to one-sided limited liability is given by

$$\max_{\lambda, \beta \geq 0, \eta} \sum_{i \in I} \Delta v_i(\sigma, \lambda_i) - \sum_{(i, a, s)} \beta_i(a, s) z_i \quad \text{s.t.}$$

$$\forall(i, a, s), \quad \sigma(a) \Delta \Pr(s|a, \lambda_i) = \eta(a, s) + \beta_i(a, s),$$

where  $\beta_i(a, s)$  is a multiplier on the liability constraint for agent  $i$  at  $(a, s)$ . Adding the dual equations with respect to  $s$  implies  $-\sum_s \beta_i(a, s) = \sum_s \eta(a, s)$  for all  $(i, a)$ . Therefore,

$$-\sum_{(i, s)} \beta_i(a, s) z_i = \sum_{(i, s)} \eta(a, s) z_i = \widehat{z} \sum_{s \in S} \eta(a, s),$$

where  $\widehat{z} = \sum_i z_i$ , so we may eliminate  $\beta_i(a, s)$  from the dual and get the equivalent problem:

$$\max_{\lambda \geq 0, \eta} \sum_{i \in I} \Delta v_i(\sigma, \lambda_i) + \widehat{z} \sum_{(a, s)} \eta(a, s) \quad \text{s.t.}$$

$$\forall(i, a, s), \quad \sigma(a) \Delta \Pr(s|a, \lambda_i) \geq \eta(a, s).$$

Any two liability profiles  $z$  and  $z'$  with  $\widehat{z} = \widehat{z}'$  lead to this dual with the same value. *Q.E.D.*

**PROOF OF THEOREM 9:** We use the same notation as in the proof of Theorem 7. Enforcing  $\sigma$  subject to participation and liability is equivalent to the value of the following problem being zero:

$$\min_{\zeta} \sum_{(i, a_i)} \varepsilon_i(a_i) \quad \text{s.t.}$$

$$\forall(i, a, s), \quad \zeta_i(a, s) \leq z_i, \quad \forall(i, a_i, b_i, \rho_i),$$

$$\sum_{a_{-i}} \sigma(a) (v_i(a_{-i}, b_i) - v_i(a))$$

$$\leq \sum_{(a_{-i}, s)} \sigma(a) \zeta_i(a, s) (\Pr(s|a_{-i}, b_i, \rho_i) - \Pr(s|a)) + \varepsilon_i(a_i),$$

$$\forall(a, s), \quad \sum_{i \in I} \zeta_i(a, s) = 0,$$

$$\forall i \in I, \quad \sum_{a \in A} \sigma(a)v_i(a) - \sum_{(a,s)} \sigma(a)\zeta_i(a, s) \Pr(s|a) \geq 0.$$

The first family of constraints imposes incentive compatibility, the second family imposes budget balance, the third family imposes individual rationality, and the last family corresponds to one-sided limited liability. The dual of this metering problem is given by the following program, where  $\lambda$ ,  $\eta$ ,  $\pi$ , and  $\beta$  represent the respective multipliers on each of the primal constraints:

$$\begin{aligned} & \max_{\alpha, \pi, \beta \geq 0, \eta} \sum_{i \in I} \Delta v_i(\sigma, \alpha_i) - \sum_{i \in I} \pi_i v_i(\sigma) - \sum_{(i,a,s)} \beta_i(a, s) z_i \quad \text{s.t.} \\ & \forall(i, a_i), \quad \sum_{(b_i, \rho_i)} \alpha_i(b_i, \rho_i | a_i) = 1, \\ & \forall(i, a, s), \quad \sigma(a) \Delta \Pr(s|a, \alpha_i) = \eta(a, s) + \pi_i \sigma(a) \Pr(s|a) + \beta_i(a, s). \end{aligned}$$

Adding the dual constraints with respect to  $s \in S$ , it follows that

$$-\sum_{(a,s)} \beta_i(a, s) = \sum_{(a,s)} \eta(a, s) + \pi_i = \widehat{\eta} + \pi_i,$$

where  $\widehat{\eta} := \sum_{(a,s)} \eta(a, s)$ . After substituting and eliminating  $\beta$ , the dual is equivalent to

$$\begin{aligned} V & := \max_{\alpha, \pi \geq 0, \eta} \sum_{i \in I} \Delta v_i(\sigma, \alpha_i) - \sum_{i \in I} \pi_i (v_i(\sigma) - z_i) + \widehat{\eta} \widehat{z} \quad \text{s.t.} \\ & \forall(i, a, s), \quad \sigma(a) \Delta \Pr(s|a, \alpha_i) \geq \eta(a, s) + \pi_i \sigma(a) \Pr(s|a). \end{aligned}$$

Now,  $\sigma$  is enforceable if and only if  $V = 0$ , that is, if and only if for any dual-feasible  $(\alpha, \pi, \eta)$  such that  $\sum_i \Delta v_i(\sigma, \alpha_i) > 0$ , we have that

$$\sum_{i \in I} \Delta v_i(\sigma, \alpha_i) \leq \sum_{i \in I} \pi_i (v_i(\sigma) - z_i) + \widehat{\eta} \widehat{z}.$$

Finally, since the dual objective is increasing in  $\eta$ , an optimal solution for  $\eta$  must solve

$$\eta(a, s) = \min_{i \in I} \{ \Delta \Pr(s|a, \alpha_i) - \pi_i \Pr(s|a) \}.$$

This completes the proof.

*Q.E.D.*

**PROOF OF COROLLARY 1:** Given the dual problem from the [proof](#) of Theorem 9, the first statement follows because if  $v_i(\sigma) \geq z_i$ , then the objective

function is decreasing in  $\pi_i$  and reducing  $\pi_i$  relaxes the dual constraints. The second statement follows by rewriting the objective as

$$\sum_{i \in I} \Delta v_i(\sigma, \alpha_i) - \sum_{i \in I \setminus t} \pi_i(v_i(\sigma) - z_i) + \widehat{\eta} \sum_{i \in I} z_i,$$

where  $t$  is the set of agents whose participation constraint will not bind ( $\pi_i^* = 0$  for  $i \in t$ ). Q.E.D.

REFERENCES

ALCHIAN, A., AND H. DEMSETZ (1972): “Production, Information Costs, and Economic Organization,” *American Economic Review*, 62, 777–795. [285]

AOYAGI, M. (2005): “Collusion Through Mediated Communication in Repeated Games With Imperfect Private Monitoring,” *Economic Theory*, 25, 455–475. [299]

AUMANN, R. (1974): “Subjectivity and Correlation in Randomized Strategies,” *Journal of Mathematical Economics*, 1, 67–96. [285]

——— (1987): “Correlated Equilibrium as an Expression of Bayesian Rationality,” *Econometrica*, 55, 1–18. [289]

CREMER, J., AND R. MCLEAN (1988): “Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions,” *Econometrica*, 56, 1247–1257. [299]

D’ASPROMONT, C., AND L.-A. GÉRARD-VARET (1998): “Linear Inequality Methods to Enforce Partnerships Under Uncertainty: An Overview,” *Games and Economic Behavior*, 25, 311–336. [285,298-300]

D’ASPROMONT, C., J. CREMER, AND L.-A. GÉRARD-VARET (2004): “Balanced Bayesian Mechanisms,” *Journal of Economic Theory*, 115, 385–396. [297]

FORGES, F. (1986): “An Approach to Communication Equilibria,” *Econometrica*, 54, 1375–1385. [285,290]

FUDENBERG, D., D. LEVINE, AND E. MASKIN (1994): “The Folk Theorem With Imperfect Public Information,” *Econometrica*, 62, 997–1039. [286,299,300]

HOLMSTRÖM, B. (1982): “Moral Hazard in Teams,” *Bell Journal of Economics*, 13, 324–340. [285, 287,288,296]

KANDORI, M. (2003): “Randomization, Communication, and Efficiency in Repeated Games With Imperfect Public Monitoring,” *Econometrica*, 71, 345–353. [286,299]

KOSENOK, G., AND S. SEVERINOV (2008): “Individually Rational, Balanced-Budget Bayesian Mechanisms and the Allocation of Surplus,” *Journal of Economic Theory*, 140, 126–261. [299]

LEGROS, P., AND H. MATSUSHIMA (1991): “Efficiency in Partnerships,” *Journal of Economic Theory*, 55, 296–322. [285,298,299]

LEGROS, P., AND S. MATTHEWS (1993): “Efficient and Nearly Efficient Partnerships,” *Review of Economic Studies*, 60, 599–611. [285,286,298,299]

MYERSON, R. (1986): “Multistage Games With Communication,” *Econometrica*, 54, 323–358. [285,290]

OBARA, I. (2008): “The Full Surplus Extraction Theorem With Hidden Actions,” *The B.E. Journal of Theoretical Economics*, 8, 1–26. [299]

RADNER, R., R. MYERSON, AND E. MASKIN (1986): “An Example of a Repeated Partnership Game With Discounting and With Uniformly Inefficient Equilibria,” *Review of Economic Studies*, 53, 59–69. [285,287,288]

RAHMAN, D. (2008): “But Who Will Monitor the Monitor?” Working Paper, University of Minnesota. [296]

SCHRIJVER, A. (1986): *Theory of Linear and Integer Programming*. New York: Wiley-Interscience. [301]



TOMALA, T. (2009): "Perfect Communication Equilibria in Repeated Games With Imperfect Monitoring," *Games and Economic Behavior*, 67, 682–694. [299]

*Dept. of Economics, University of Minnesota, 4-101 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN 55455, U.S.A.; dmr@umn.edu*

*and*

*Dept. of Economics, University of California, Los Angeles, 8283 Bunche Hall, Los Angeles, CA 90096, U.S.A.; iobara@econ.ucla.edu.*

*Manuscript received October, 2005; final revision received October, 2009.*