Communications and Efficiency in Coordination Games

Banks, Jeffrey S. and Randall L. Calvert

Working Paper No. 196
August 1989

Communication and Efficiency in Coordination Games

Jeffrey S. Banks and Randall L. Calvert

Rochester Center for Economic Research
Working Paper No. 196

COMMUNICATION AND EFFICIENCY IN COORDINATION GAMES

Jeffrey S. Banks
and
Randall L. Calvert

Department of Political Science
University of Rochester

Working Paper No. 196

March 1988
Revised August 1989

This research was supported by NSF grant SES-8908226. The authors are grateful to workshop participants at Harvard University, the University of Illinois, the IMSSS, the University of Iowa, the Massachusetts Institute of Technology, the University of Rochester, and Washington University, St. Louis, for helpful comments.

Abstract

Communication and Efficiency in Coordination Games

Coordination games with complete information have been used to model a variety of social interactions. The present study examines in detail how improvements in coordination can be achieved through communication among the players and between players and outside actors. Its main goal is to understand the role of communication in a coordination game with asymmetric information. We characterize the set of Bayesian equilibria in a two-player coordination game with incomplete information; derive an ex ante efficient communication mechanism with a mediator; characterize all ex ante efficient symmetric decision rules with communication; and give necessary and sufficient conditions under which communication can be efficient without mediation. Among other results, we find that incomplete information without communication yields efficiency gains over complete information; and that in general unmediated communication is insufficient to achieve incentive efficiency.

Communication and Efficiency in Coordination Games

1. Introduction

In any game with multiple equilibria, the players face a problem of coordination: there is a danger that different players will use strategies corresponding to different equilibria in such a way that the result is Pareto-inferior to all those equilibria.[1] The players would prefer, if possible, to coordinate their actions so as to avoid such inferior outcomes. The problem of coordination is captured most elegantly in the "battle of the sexes" game of Luce and Raiffa (1957, p. 90-91):

		Player 2's decision	
		x	y
Player 1's decision	x	0 0	1 t
	y	t 1	0 0

where $t > 1$. In this game, player 1 prefers one of the two pure-strategy Nash equilibria, (x, y), while player 2 prefers the other, (y, x). There is one other equilibrium, a symmetric, mixed-strategy profile in which each player chooses x with probability $t/(t+1)$. To assume that the players successfully coordinate on one of the pure-strategy equilibria would beg the question of how, in an initially symmetric situation, the players arrive at an outcome that is asymmetric in terms of actions and expected payoffs (Farrell 1987). The symmetric equilibrium, then, seems the most reasonable prediction. In that case, the players will fail to coordinate with probability $(t^2+1)/(t+1)^2$, and indeed achieve an expected payoff that is Pareto-dominated by both of the pure strategy equilibria.

Coordination games with complete information have been used to model a variety of social interactions, including market entry (Dixit and Shapiro 1985), product compatibility (Farrell and Saloner 1988), networking (Katz and Shapiro 1985), political leadership (Calvert 1987), repeated prisoners' dilemmas (Hardin 1982), and bargaining (Schelling 1960). In the Dixit-Shapiro model, for example, the action "x" denotes the decision to enter a naturally monopolistic market, while "y" denotes staying out. The focus in all these studies is to determine how, in a setting in which the basic coordination game is augmented with such processes as communication and repeated play, players could increase the equilibrium probability of coordinating.

The present study examines in detail how such improvements in coordination can be achieved through communication among the players and between players and outside actors. In the case of the battle-of-the-sexes game above, our analysis extends that of Farrell (1987), who demonstrates that "cheap talk" among the players could yield efficiency improvements through increased coordination. In section 2 we show how communicating players can generally do even better than the Farrell model indicates, and can in fact achieve full Pareto efficiency.

Our main goal, however, is to understand the role of communication in a coordination game with asymmetric information. Taking again the example of the Dixit-Shapiro model, suppose that each firm's true production costs for each market are private information. Assuming that the initial situation is symmetric, then in addition to solving the original coordination problem the firms also can profit, *ex ante*, by letting the lowest-cost firm be the entrant in any market. Section 3 describes the basic two-player coordination game with incomplete information, and characterizes the efficient outcomes and the Bayesian equilibria. In section 4, we assume in addition that the players are able to communicate with one another and with outside actors before playing the game. In general, the role of the "outside actor" is to improve outcomes by serving as an impartial mediator. We derive an ex ante efficient decision rule -- the pattern of communication and behavior yielding the highest payoff achievable in equilibrium through any such pre-play communication process. Using that derivation we derive properties that must hold for any symmetric, incentive-efficient efficient decision rule and examine the importance of the mediator's (i) ability to communicate privately with the players and (ii) inability to enforce suggestions about how they should play the game. Section 5, finally, gives necessary and sufficient conditions under which bilateral communication alone, without mediation, can generate such an efficient outcome. In contrast with recent results concerning bargaining games (see for example Matthews and Postlewaite 1989), we prove that bilateral communication is not sufficient in general to achieve the payoff levels available with the assistance of a mediator.

2. Coordination with Communication under Complete Information

Farrell (1987) and others have argued that one means of overcoming such coordination failures in practice is through communication between the players prior to selecting their actions. Suppose that the players simultaneously and costlessly announce their (nonbinding) intentions to play x or y, and consider the following strategies for subsequently playing the coordination game: if they announce (x, y) they play the (x, y) equilibrium in the game; if they announce (y, x) they play (y, x); and if they announce the same action then they play the mixed strategy equilibrium. Farrell (1987) then shows that there is a symmetric equilibrium in this game in which the probability of coordination and hence the resulting payoffs are higher than they would be without communication. Farrell also shows that, as the number of such communication stages increases prior to the play of the game, the probability of coordination increases, but is bounded away from 1. Thus no matter how much of this "cheap talk" is allowed, a fully coordinated outcome cannot be achieved through symmetric behavior of the players.

One difficulty with such communication games is that there are many symmetric equilibria. Farrell (1987) selects from among these the one that is "conventional" (Palfrey and Rosenthal's (1988) term) in that, if the players announce a Nash equilibrium, they subsequently play this equilibrium. This selection argument, however, is based on the labelling of the messages the players send, so that if the messages were not labelled as the actions are labelled, some other outcome could "conventionally" be selected. Suppose that in the one-stage communication game the two messages were labelled "heads" and "tails", and that the convention to be used were that of "matching pennies:" the players truthfully announce the results of private coin flips, and if the players' announcements match, i.e. both announce "heads" or both announce "tails", the players will play (x, y), while if they do not match they play (y, x). These strategies of communicating and playing are in equilibrium and yield symmetric, ex ante efficient payoffs -- the outcome is (x, y) with probability 1/2 and (y, x) with probability 1/2. Indeed, these strategies implement the efficient symmetric correlated equilibrium (Aumann 1974), which is the best the players can achieve in the original game.[2]

Thus in the complete-information battle of the sexes with one stage of simultaneous communication using a two-element set of available, non-binding, non-verifiable messages (exactly Farrell's setup), the players can achieve efficiency in a symmetric equilibrium without the need for a joint randomizing device or for an impartial mediator to carry out randomizations and give instructions. One possible conclusion from all this is that, if the players are communicating in order to arrive at coordinated outcomes in a manner that preserves their symmetric situations, then the players themselves may be able to figure out how to interpret messages in such a way that coordination is achieved. In the environment described by Farrell, the players may realize that they are both better off playing the "matching pennies" version of the game and so rely on a different interpretation of the messages than in the "cheap talk" equilibrium. The key to understanding bilateral communication may require analyzing equilibrium behavior when the labels on the messages have been "neutralized," that is, each player simply has a set of messages M. In this framework, two communication processes differ only if the number of messages available to the players differs. Therefore, given a game such as the battle of the sexes above, we can inquire whether, if the players have enough messages, there exists an equilibrium in the resulting communication game that implements an efficient solution. For the battle of the sexes with complete information, the answer is yes; as long as the players have at least two messages, they can achieve a symmetric and efficient solution. We shall return to this method of evaluating bilateral communication for an incomplete-information environment in section 5.

3. A Coordination Game with Incomplete Information

Suppose now that the payoff t to each player in the battle-of-the-sexes game is private information. We consider in particular the following version of the game, to which we refer henceforth as G:

Player 1's decision \ Player 2's decision	x	y
x	0, 0	t_1, 1
y	1, t_2	0, 0

where $t_i \in \{a, b\} \equiv T_i$, $b > a > 1$, and $\text{Prob}(t_i = b) = p$ for $i = 1, 2$. Then the original coordination problem persists, in that the players' preferences over the "coordinated" outcomes differ, but now in addition it is unclear which of the players has the greater stake in achieving his preferred outcome. This complication presents the players with an opportunity: prior to learning their true payoffs, they could realize mutual gains in expected utility if they could agree that if only one player turns out to be in his "high-stakes" situation the resulting outcome is to be biased in this player's favor. For suppose that they managed to choose the higher-stakes player's preferred outcome with probability λ and the other's with probability 1 − λ; then each player's expected utility conditional on the players being of different types would be

$$(1-p)p[\lambda + (1-\lambda)a] + p(1-p)[\lambda b + (1-\lambda)]$$
$$= p(1-p)[\lambda(b-a) + a + 1] \ ,$$

which is maximized when $\lambda = 1$. Such distributional issues are important in many of the real world situations modeled as coordination games as cited in Section 1. The remainder of this paper concerns the possibility of achieving such a favorable distributional arrangement, along with the possibility of achieving coordination at all, given the requirement that the players behave rationally in choosing between x and y as well as (in subsequent sections) in revealing their true payoffs. We begin by establishing what happens in equilibrium when the players cannot communicate.

Let $D_i = \{x, y\}$ be player i's decision set in the coordination game, and write $D = D_1 \times D_2$, $T = T_1 \times T_2$. A strategy for player i in the game G is a function mapping each possible type into a probability distribution over actions. For convenience we write player i's strategy as

$$\sigma_i : D_i \times T_i \dashrightarrow [0, 1],$$

where $\sigma_i(d_i; t_i)$ is the probability that player i takes action $d_i \in D_i$, given type $t_i \in T_i$. For any $d \in D$ and $t \in T$, let $u_i(d, t)$ denote i's utility from the action pair $d = (d_1, d_2)$ given types $t = (t_1, t_2)$; thus $u_1(x, y; a, b) = a$, $u_2(x, y; a, b) = 1$, etc. A pair of strategies $\sigma = (\sigma_1, \sigma_2)$ induces (expected) payoffs for the players in the usual fashion:

$$U_i(\sigma; t_i) = \sum_{t_{-i} \in T_{-i}} \{Pr(t_{-i}) \sum_{d \in D} \sigma_i(d_i; t_i)\sigma_{-i}(d_{-i}; t_{-i})u_i(d, t)\}.$$

where $-i$ denotes the opponent of player i.

Definition. A Bayesian equilibrium to the game G is a strategy pair $\sigma^* = (\sigma_1^*, \sigma_2^*)$ such that for $i = 1, 2$,

$$U_i(\sigma^*; t_i) \geq U_i(\sigma_i, \sigma_{-i}^*; t_i) \qquad \forall\, t_i \in T_i, \forall\, \sigma_i.$$

Let B(G) denote the set of Bayesian equilibria of the game G.

As with the complete information version of the game, G possesses asymmetric equilibria. In particular, the following strategy pairs constitute Bayesian equilibria:

$$\sigma_1(x;\cdot) = \sigma_2(y;\cdot) = 1, \quad (1)$$
$$\sigma_1(y;\cdot) = \sigma_2(x;\cdot) = 1.$$

Thus if the players select the coordinated outcome regardless of type, such a strategy pair is an equilibrium. Of course, regardless of the players' types they still face the problem of which equilibrium to play, since the players' preferences over these asymmetric equilibria still diverge. By the same reasoning as we used in the complete-information version of the game, a symmetric equilibrium, if one exists, would yield a useful prediction of the players' choices. An equilibrium is symmetric if it satisfies

$$[d_1 = d_2 \text{ and } t_1 = t_2] \Longrightarrow \sigma_1(d_1;\ t_1) = \sigma_2(d_2;\ t_2)$$

or equivalently,

$$\sigma_1(d_1;\ t_1) = \sigma_2(d_1;\ t_1) \quad (2)$$

For every d_1, d_2, t_1, and t_2.

Our first result derives B(G) and shows that it contains exactly one symmetric equilibrium.

Proposition 1. The game G has the following Bayesian equilibria, and no others:

(1.0) the type-independent strategy pair in (1) above, which is in equilibrium for any value of p;

(1.1) for $p < \frac{a}{a+1}$, $\sigma_i(x;\ a) = \frac{(1-p)a - p}{(1-p)(a+1)}$, $\sigma_i(x;\ b) = 1$;

(1.2) for $\frac{a}{a+1} \leq p \leq \frac{b}{b+1}$,

(i) $\sigma_i(x;\ a) = 0$, $\sigma_i(x;\ b) = 1$ for $i = 1, 2$; or

(ii) $\sigma_i(x,\ a) = 0$, $\sigma_i(x,\ b) = \frac{a}{p(a+1)}$ and

$\sigma_{-i}(x,\ a) = \frac{(1-p)b - p}{(1-p)(b+1)}$, $\sigma_{-i}(x,\ b) = 1$ for $i = 1, 2$;

(1.3) for $p > \frac{b}{b+1}$, $\sigma_i(x;\ a) = 0$, $\sigma_i(x;\ b) = \frac{b}{p(b+1)}$.

(1.4) for $p = a/(a+1)$ and $p = b/(b+1)$, a continuum of other mixed-strategy, type-dependent equilibria resembling (1.2)(ii) is also possible.

Note. The combination of (1.1), (1.2)(i), and (1.3) defines a unique symmetric equilibrium for each value of p. The extra equilibria of (1.4) are specified in Case 4 of the proof.

Proof. For notational simplicity let $q_i = \sigma_i(x;\ a)$ and $r_i = \sigma_i(x;\ b)$ be the strategy mixtures used by each type of player i. Also, let $s_i = (1-p)q_i + pr_i$ represent i's apparent mixed strategy to player $-i$. Then player i's best response to q_{-i}, r_{-i} is the pair q_i, r_i that maximizes the expected payoffs

$$(1 - q_i)\cdot s_{-i} + aq_i\cdot(1 - s_{-i}) \quad \text{and}$$
$$(1 - r_i)\cdot s_{-i} + br_i\cdot(1 - s_{-i}) ,$$

respectively. This gives the following best-response function for player i:

$$q_i \begin{cases} = 0 & \text{if} \quad s_{-i} > \dfrac{a}{a+1}, \text{ i.e., } (1-p)q_{-i} + pr_{-i} > \dfrac{a}{a+1} \\ \in [0, 1] & \text{if} \qquad = \\ = 1 & \text{if} \qquad < \end{cases}$$

and

$$r_i \begin{cases} = 0 & \text{if} \quad s_{-i} > \dfrac{b}{b+1}, \text{ i.e., } (1-p)q_{-i} + pr_{-i} > \dfrac{b}{b+1} \\ \in [0, 1] & \text{if} \qquad = \\ = 1 & \text{if} \qquad < \end{cases} .$$

It is clear from this best response correspondence that (1.0) is in equilibrium regardless of the values of a, b, and p (but recall that $1 < a < b$ and $p \in [0, 1]$). Also, looking at the best response as a function of s_{-i}, it is clear that for any equilibrium either $q_i = 0$ or $r_i = 1$ or both (see Figure 1). We use this fact to examine other all other possible Bayesian equilibria for each of the intervals for p represented in (1.1), (1.2), and (1.3).

Case 1: $p < a/(a+1)$. Suppose that $r_i < 1$. Then $q_i = 0$ so $s_i = pr_i < a/(a+1) < b/(b+1)$. Applying the best response correspondence derived above to player $-i$, then, we must have $q_{-i} = r_{-i} = 1$, which in turn implies $q_i = r_i = 1$, just the pure equilibrium (1.0).

The same reasoning holds for each player, so it remains to consider what happens when $r_1 = r_2 = 1$. First, notice that both q_1 and q_2 must lie in the open interval (0, 1): for if $q_i = 1$ then $s_i = 1 - p + p = 1 > b/(b+1)$, forcing $r_{-i} = 0$, a contradiction; and if $q_i = 0$ then $s_i = p < a/(a+1)$, implying $q_{-i} = 1$ and thus $r_i = 0$, again a contradiction. Now to have $q_i \in (0, 1)$ requires $s_{-i} = a/(a+1)$, that is, $(1-p)q_{-i} + p = a/(a+1)$, so $q_{-i} =$

$\frac{\frac{b}{b+1} - p}{1 - p} = \frac{(1-p)a - p}{(1-p)(a+1)}$. Notice that this is always in (0, 1) as required since $p < a/(a+1) < b/(b+1)$. Identical reasoning gives the same value for q_i, yielding the equilibrium given in (1.1) of the proposition.

Case 2: $p > b/(b+1)$. Reversing the roles of q_i and r_i and considering the cases $q_i > 0$ and $q_1 = q_2 = 0$, reasoning identical to that in Case 1 above gives the equilibrium in (1.3) of the proposition.

Case 3: $a/(a+1) < p < b/(b+1)$. We consider each of several (exhaustive) possibilities in turn. First, suppose $q_1 = q_2 = 0$. Then $s_i = pr_i$ for each i. Letting $r_1 = r_2 = 1$ would give $s_1 = s_2 = p$, so for Case 3 these values are consistent with equilibrium (1.2)(i). Suppose, on the other hand, that $r_i < 1$. This would require $s_{-i} \geq b/(b+1)$, that is, $r_{-i} \geq \frac{b}{p(b+1)} > 1$. Thus the only equilibrium having $q_1 = q_2 = 0$ is (1.2)(i).

Second, suppose $r_1 = r_2 = 1$, so $s_i = (1-p)q_i + p$ for each i. If $q_1 = q_2 = 0$ we have equilibrium (1.2)(i) again, just as above. Suppose $q_i > 0$; this would require $s_{-i} \leq a/(a+1)$, that is, $q_{-i} \leq \frac{\frac{a}{a+1} - p}{1 - p} < 0$. Thus the only equilibrium having $r_1 = r_2 = 1$ is (1.2)(i).

Finally, suppose that for either i = 1 or i = 2, $q_i = 0$ and $r_{-i} = 1$, so $s_i = pr_i$ and $s_{-i} = (1-p)q_{-i} + p$. Either of $r_i = 1$ or $q_{-i} = 0$ implies the other, and yields (1.2)(i) again as an equilibrium. Either of $r_i = 0$ or $q_{-i} = 1$ implies the other and yields the pure equilibrium (1.0). The remaining possibility is to let $r_i \in (0, 1)$. This requires $s_{-i} = b/(b+1)$, hence $q_{-i} = \frac{(1-p)b - p}{(1-p)(b+1)}$. In turn, this value of q_{-i} is in (0, 1) and requires $s_i = a/(a+1)$, hence $r_i = \frac{a}{p(a+1)} \in (0, 1)$, giving the asymmetric mixed equilibrium (1.2)(ii).

Case 4: boundaries. At the boundaries of Case 3, the strategies in

(1.1), (1.2)(i), and (1.3) coincide where appropriate. Equilibrium (1.2)(ii) expands somewhat, however. The argument made in Case 3 when the initial assumption was $q_1 = q_2 = 0$ still holds, except that now when $p = b/(b+1)$ it is possible to have $r_i = 1$ and $r_{-i} \in (\frac{a}{p(a+1)}, 1)$, the lower bound on r_{-i} being needed to maintain $q_i = 0$ as a best response. Likewise when the initial assumption was $r_1 = r_2 = 1$, the argument still holds with the additional possibility when $p = a/(a+1)$ that $q_{-i} = 0$ and $q_i \in (0, \frac{(1-p)b - p}{(1-p)(b+1)})$, the upper bound being needed to maintain $r_{-i} = 1$ as a best response. These calculations yield the extra equilibria mentioned in (1.4) of the proposition. □

Denote the intervals defined in Proposition 1 by $P_1 = [0, a/(a+1))$, $P_2 = [a/(a+1), b/(b+1)]$, and $P_3 = (b/(b+1), 1]$. Then we can describe the symmetric equilibrium as follows. For $p \in P_1$, b-types always take their preferred actions (x) while a-types randomize, using their preferred strategies with a probability that ranges from $a/(a+1)$ at $p = 0$ to 0 at $p = a/(a+1)$. For $p \in P_2$, b-types always take their preferred action while a-types always take their less preferred action. Finally, for $p \in P_3$, a-types always take their less-preferred action while b-types randomize, using their preferred strategy with a probability ranging from 1 at $p = b/(b+1)$ to $b/(b+1)$ at $p = 1$.

We will be concerned throughout this paper with comparing the efficiency of various equilibria. To facilitate comparisons to and among the equilibria described in Proposition 1, Table 1 shows the corresponding payoffs for players of each type, playing each role (i or -i) in each of the equilibrium strategies.

Table 1 here

Obviously some potential gains from coordination go unrealized under the equilibria described in (1.1) through (1.4), since in those the players fail with positive probability to coordinate. While the pure strategy equilibrium of (1.0) avoids this problem, it seems an unwise prediction for the same reasons as under complete information: it begs the question of how the players reach a particular asymmetric solution to this initially symmetric problem. Thus we concentrate on the symmetric equilibrium.

Recall that one of the features of the symmetric equilibrium in the complete information battle of the sexes game is that the equilibrium is Pareto-dominated by both of the asymmetric, pure-strategy equilibria; in particular both players would rather be at their least preferred of these outcomes. With incomplete information, however, the analogous statement need no longer be true. To see this, note in Table 1 that an a-type always prefers either of the type-independent, pure-strategy equilibria, since in the symmetric equilibrium an a-type's payoff is always less than 1, which is the payoff in a player's less-preferred pure-strategy equilibrium regardless of type. However if $p \leq a/(a+1)$ and $b > a+1$, then a b-type player prefers the symmetric equilibrium payoff, since this gives a payoff of $b/(a+1)$. In addition, from an ex ante perspective a player's expected utility from the strategy profile σ is is $pU_i(\sigma,b) + (1-p)U_i(\sigma,a)$ so that if, e.g., $p = a/(a+1)$ and $a(b+1) > (a+1)^2$ both players' ex ante expected utility from the symmetric equilibrium is higher than that from the less-preferred pure-strategy equilibrium. Thus for some parameter values the symmetric equilibrium is not Pareto-dominated by any other Bayesian Nash equilibrium. More generally, the values shown in Table 1 can be used to calculate that the symmetric equilibrium is efficient in this sense when $p \in P_1$ if and only if $p \leq 1/(b-a)$; and efficient when $p \in P_2$ if and only if

$$p \notin \left[\frac{1 - \sqrt{\frac{b-3}{b+1}}}{2}, \frac{1 + \sqrt{\frac{b-3}{b+1}}}{2} \right].$$

(Notice that the asymmetric, mixed-strategy equilibrium in P_2 is always worse for player i than the symmetric equilibrium, so efficiency of the symmetric equilibrium on P_2 only depends on whether it is dominated by the pure-strategy equilibrium.)

It is interesting to compare the interim payoffs from this game with those of the complete information version of G. Let $U(t,p)$ denote the expected payoff from the symmetric Bayesian equilibrium for a player of type t. If $p \in P_1$, then $U(a,p) = a/(a+1)$ and $U(b,p) = b/(a+1)$; if $p \in P_2$, then $U(a,p) = p$ and $U(b,p) = b(1-p)$; and if $p \in P_3$, then $U(a,p) = U(b,p) = b/(b+1)$. From Section 2 we know that the mixed strategy equilibrium generates a payoff of $t/(t+1)$ for both players. Suppose now that the complete information game is played where $t_1 \neq t_2$; if we continue to rely on the mixed-strategy equilibrium (although the strategies are not the same for both players), we see that the payoff for player i is $t_i/(t_i+1)$, regardless of the value of t_j, $j \neq i$. Thus, comparing the complete-information payoffs with the values of $U(t,p)$ calculated above, we see that both types receive as least as high an expected payoff in the incomplete information game as under complete information, and for any value of p at least one type receives a strictly higher payoff. Thus the players are better off under incomplete information.

This comparison is analogous to that found in Kreps, Milgrom, Roberts, and Wilson (1982), where the presence of incomplete information in a finitely repeated prisoner's dilemma game creates Pareto-improvements over the outcomes under complete information. A result in the same spirit is derived by Vickers (1986) in a model of monetary policy making. Both these models rely on the multiperiod "signaling" nature of the strategies, which provides the opportunity of reputation-building. Our result, on the other hand, shows that incomplete information can provide Pareto improvements even in a one-shot, simultaneous-move game. This occurs because the presence of differing types gives the players a device by which they can distinguish themselves from one another and, more often than was possible under complete information, achieve the necessary asymmetric outcome of the symmetric coordination game.

4. Incentive Efficient Decision Rules with Communication

Having established the Bayesian equilibria as benchmarks for the outcomes of rational play in the incomplete-information coordination game G, we can now evaluate the possibility of improving on these outcomes through communication between the players and between players and outside actors. Specifically, we now consider augmented forms of G in which the players first have opportunities to send and receive messages according to some specified extensive form, after which they play G. All payoffs accrue from the actual play of G, and none directly from the communication process. Throughout our analysis, no binding commitments can be made during the communication process about behavior in the ultimate play of G, and no verification of claims made in the communication process is possible.

In order to conduct the analysis, we first cover several essential terms and tools.

Preliminaries. A strategy profile in an augmented game of the type described above, say G^+, determines a sequence of messages, actions, and ultimately an outcome of the basic game. We can think of an equilibrium strategy profile in $B(G^+)$ as generating a decision rule for G. A decision rule is a function assigning a probability distribution over outcomes as a function of the players' types; we will write δ: T × D --> [0, 1], so that $\delta(d; t)$ is the probability of outcome $d = (d_1, d_2)$ of G when the players are of types $t = (t_1, t_2)$. As Holmstrom and Myerson (1983) note, the decision rule is the relevant object for making efficiency comparisons. In this section we characterize those decision rules that are incentive efficient; that is, they constitute the highest payoffs the players could ever achieve through equilibrium play in any arbitrary communication game ending with play of G. In so doing we will also further explore the efficiency of the Bayesian equilibrium described in Proposition 1 above.

Efficency Concepts. We compare decision rules via the expected payoffs associated with those rules. Such a comparison can potentially be made at three different stages of the play of the game: (i) the "ex ante" stage, prior to the players learning their types; (ii) the "interim" stage, after they haved learned their types but prior to any moves in the game; or (iii) the "ex post" stage, with full knowledge of both players' types. We say that a decision rule δ is ex ante (respectively interim, ex post) efficient in some feasible set Γ if there does not exist another decision rule γ in Γ such that the payoffs according to γ ex ante (resp. interim, ex post) Pareto-dominate those of δ. In what follows we focus attention on the criterion of ex ante incentive efficiency. The focus on ex ante payoffs has two motivations. First, the set of ex ante efficient decision rules is a subset of the interim, as well as ex post, efficient decision rules (Holmstrom and Myerson 1983); hence ex ante efficient decision rules constitute a selection from the set of interim efficient decision rules. The second motivation concerns endogenous institutional design: the players realize that in the future they will be faced with the decision problem summarized by G. Therefore, prior to the play of G and prior to their learning of types, the players may have an incentive to influence the resulting outcomes by establishing the guidelines of their forthcoming interaction. Under such circumstances the players would prefer to choose to interact so as to achieve payoffs that are efficient from their ex ante point of view.

Symmetric Decision Rules. A further restriction in our approach is that we focus on symmetric decision rules. Given the symmetry of the original game G, it seems natural to require the associated communication games to reflect this property as well. As argued previously, then, we should concentratre on symmetric equilibria as the predictions for outcomes in the augmented game; and symmetric equilibria of G^+ generate symmetric decision rules. We say that a decision rule δ is symmetric if it is independent of the labeling of the players. More formally, we require that the following equality hold for all decisions d_1, d_2 and all types t_1, t_2:

$$\delta(d_1, d_2; t_1, t_2) = \delta(d_2, d_1; t_2, t_1) \quad (3)$$

Direct Mechanisms and the Revelation Principle. Finally, we turn to the subject of modelling communication between the players. Suppose that the augmented game G^+ begins with a communication stage, whose normal form has pure-strategy sets C_1 and C_2 for the two players. In the communication stage, the players send messages to one another or to an outside actor, and receive such messages, according to completely specified orders of moves, sets of available messages at each move, and rules by which the outside actor responds to received messages. Let $\rho_i'(c_i;\ t_i)$ be player i's probability of using communication strategy c_i when his type is t_i. The outcome of the communication stage is $c = (c_1,\ c_2)$, and it can be observed partially by the two players (only partially because, for example, some of the outside actor's messages may be private information). Let c^i be player i's observation of the outcome c. We assume here that c can be written equivalently as either $(c_1,\ c_2)$ or as $(c^1,\ c^2)$. As a notational convenience, we may write $c^1 = c^2$ whenever player 1's observation of c given his initial situation is identical to that of player 2 -- e.g., when the players have sent the same messages to one another and received the same messages from outside actors. This will prove useful in considering symmetric G^+ below.

The players proceed to play G, having observed c. Their actions in G are now chosen according to strategies $\alpha_i'(d_i;\ t_i,\ c^i)$, the probability that player i uses decision d_i in G, given that he is of type t_i and he observed c^i as the outcome of the communication stage. The decision rule generated by the strategy profiles ρ' and α' is given by

$$\delta(d;\ t) = \sum_{c \in C} \alpha_1'(d_1;\ t_1,\ c^1)\ \alpha_2'(d_2;\ t_2,\ c^2)\ \rho_1'(c_1;\ t_1)\ \rho_2'(c_2;\ t_2). \qquad (4)$$

By restricting attention to a certain class of simple communication processes, it is possible to learn all there is to know about decision rules that can be derived from any communication process. A <u>direct mechanism</u> is a communication process having the following simple form. First the players privately report their types to a single outside actor, the "mediator;" they may lie. Second, the mediator uses a rule (the "mechanism") known to the players to derive from their reports a recommended decision for playing G; he communicates each player's recommendation privately. Such a mechanism takes the form of a function $\mu : D \times T \dashrightarrow [0, 1]$ such that $\sum_{d \in D} \mu(d, t) = 1$ for all $t \in T$. Thus $\mu(d, t)$ is the probability that the arbitrator recommends the decision d_1 to player 1 and d_2 to player 2, given that the players have reported types t. The mechanism μ then generates an augmented game G^{μ} in which each player's strategy takes the following form: communciate according to $\rho_i : T_i \times T_i \dashrightarrow [0, 1]$, which describes the probability that player i reports each type given i's true type; and act according to $\alpha_i : D_i \times T_i \times D_i \dashrightarrow [0, 1]$, which describes the probability that i takes each action in D_i given type i and the mediator's recommendation. (When the meaning is clear from the context, we will sometimes refer to the decision rule generated by a mechanism μ simply as μ.)

The revelation principle (Myerson 1985) states that any decision rule resulting from an equilibrium ρ' and α' in any G^+ based on G, that is, from any preplay communication process, can also be derived from an incentive compatible direct mechanism (ICDM) based on G. A direct mechanism μ is incentive compatible if honest reporting of types and obedience of the mediator's recommendation by both players constitutes a Bayesian equilibrium in G^{μ}; i.e. if the strategies $\rho_i(t_i, t_i) = 1$ and $\alpha_i(d_i; t_i, d_i) = 1$ for each i, each t_i, and each d_i are in equilibrium. The revelation principle is proved simply by noting that, given any G^+ and given the equilibrium ρ', α' in $B(G^+)$, the resulting δ can be achieved in an ICDM by setting $\mu(d; t) = \delta(d; t)$, for then truthful reporting in G^{μ} leads to the same result as using the equilibrium strategy for a player's true type in G^+, and no type would like to report falsely or disobey the mediator given that the other player does not do so.

Since we wish to concentrate on symmetric decision rules, we need to establish a further result, namely that it will suffice for us to look at symmetric ICDMs. A symmetric ICDM $\mu(d;\ t)$ is one that obeys the same conditions as did the symmetric decision rule $\delta(d;\ t)$ in (3), namely

$$\mu(d_1,\ d_2;\ t_1,\ t_2) = \mu(d_2,\ d_1;\ t_2,\ t_1).$$

As previously in (2), we say that ρ' is symmetric if $\rho_1'(c_1;\ t_1) = \rho_2'(c_1;\ t_1)$ and that α' is symmetric if $\alpha_1'(d_1;\ t_1,\ c^1) = \alpha_2'(d_1;\ t_1,\ c^1)$; that is, the players use identical mixed strategies in identical situations. We can now state the following "symmetric revelation principle."

Lemma 1. Let ρ', α' be a symmetric equilibrium for any augmented game G^+. Then the resulting decision rule δ is symmetric. Further, any symmetric decision rule δ that results from equilibrium strategies in any augmented game can be achieved using a symmetric ICDM.

Proof. Since ρ' and α' are symmetric, from (4) we have

$$\begin{aligned} \delta(d_1,\ d_2;\ t_1,\ t_2) &= \sum_{c \in C} \alpha_1'(d_1;\ t_1,\ c^1)\ \alpha_2'(d_2;\ t_2,\ c^2)\ \rho_1'(c_1;\ t_1)\ \rho_2'(c_2;\ t_2) \\ &= \sum_{c \in C} \alpha_1'(d_2;\ t_2,\ c^2)\ \alpha_2'(d_1;\ t_1,\ c^1)\ \rho_1'(c_2;\ t_2)\ \rho_2'(c_1;\ t_1) \\ &= \delta(d_2,\ d_1;\ t_2,\ t_1)\ , \end{aligned}$$

so δ is symmetric, proving the first claim. To prove the second claim, the revelation principle ensures that δ can be implemented by some μ; since this is done by setting $\mu = \delta$, μ is symmetric. □

The set of all decision rules deriving from incentive-compatible behavior in symmetric ICDMs defines the feasible set of decision rules for our analysis of efficiency below. By the revelation principle, however, our efficiency results will apply to the feasible set of all decision rules derived from equilibrium behavior in <u>any</u> symmetric mechanism. And by Lemma 1, this is equivalent to the set of all incentive-compatible decision rules.

Finally, it is important to note that our restriction to symmetric decision rules will never lead us to consider a decision rule that is not efficient <u>overall</u>. This is made clear in the following result:

Lemma 2. Suppose δ^* is (ex ante, interim, or ex post) incentive-efficient within the class of all symmetric decision rules. Then δ^* is also incentive-efficient in the class of all decision rules.

Proof. Suppose instead that δ^{**} is a decision rule that is incentive-compatible, i.e. can be implemented in equilibrium, and that Pareto-dominates δ^*. Then δ^{**} must be asymmetric. Let δ' be defined by

$$\delta'(d_1, d_2; t_1, t_2) = \delta^{**}(d_2, d_1; t_2, t_1) ;$$

that is, δ' is identical to δ^{**} except that it treats the players in the opposite fashion from δ^{**}. Thus δ' dominates δ^* as well. Let μ^*, μ^{**}, and μ' be the ICDMs that implement δ^*, δ^{**}, and δ' respectively. Define one more ICDM, μ'', by

$$\mu''(d_1, d_2; t_1, t_2) = \frac{1}{2} [\mu^{**}(d_1, d_2; t_1, t_2) + \mu'(d_1, d_2; t_1, t_2)] ,$$

so that μ'' is the "average" of μ^{**} and its mirror image μ'. We will show that μ'' is incentive compatible and symmetric, and thus that $\delta \equiv \mu''$ is a symmetric, incentive compatible decision rule that dominates δ^*. This contradicts the assumed incentive-efficiency of δ^{*} among symmetric decision rules, so such a δ^{**} cannot exist, and thus δ^{*} is incentive efficient in the class of all decision rules.

Obedience under μ'' is immediate since the recommendation for player i in effect comes from either μ^{**} or μ', and -i's recommendation comes from the same mechanism; since both μ^{**} and μ' elicit obedience, so does μ''. To prove honesty under μ'', let $V(r_i;\ t_i,\ \mu)$ represent player i's expected payoff under mechanism μ from reporting type $r_i \in T_i$ when i's true type is t_i, given that player -i will be honest and obedient and given that player i will behave in a specified fashion (obey x only, obey y only, obey both, or obey neither) after the mediator's recommendations are made. Then we know that $V(t_i;\ t_i,\ \mu) \geq V(r_i;\ t_i,\ \mu)$ for all r_i, for $\mu = \mu^{**}$ and for $\mu = \mu'$, from the incentive compatibility of μ^{**} and μ', respectively. But $V(r_i;\ t_i,\ \mu'') = 1/2\ [V(r_i;\ t_i,\ \mu) + V(r_i;\ t_i,\ \mu)]$, so honesty holds for μ'' as well. Thus μ'' is incentive compatible. Finally, μ'' is symmetric since by the definition of μ'

$$\begin{aligned}\mu''(d_1,\ d_2;\ t_1,\ t_2) &= 1/2\ [\mu^{**}(d_1,\ d_2;\ t_1,\ t_2) + \mu'(d_1,\ d_2;\ t_1,\ t_2)] \\ &= 1/2\ [\mu'(d_2,\ d_1;\ t_2,\ t_1) + \mu^{**}(d_2,\ d_1;\ t_2,\ t_1)] \\ &= \mu''(d_2,\ d_1;\ t_2,\ t_1). \qquad \square\end{aligned}$$

Results: Incentive-Efficient Mechanisms. Our goal then is to characterize those decision rules that are ex ante efficient in the feasible set consisting of all symmetric decision rules derived from equilibrium play in a communication-augmented version of G -- i.e., the "incentive-efficient" symmetric decision rules (Holmstrom and Myerson 1983). An obvious initial question is whether the requirement that a decision rule be derived from equilibrium behavior in G^{+} is a binding constraint on payoffs. In other words, is the "classically efficient" symmetric decision rule achievable in equilibrium? Using the reasoning from the beginning of Section 3 about the optimal distribution of expected payoff between an a-type and a b-type, it is easy to see that the classically efficient symmetric decision rule is equivalent to the following mechanism:

$$\delta_c(x, y; \tau, \tau) = \delta_c(y, x; \tau, \tau) = 1/2 ;$$

$$\delta_c(x, y; b, a) = 1.$$

Thus if both players are of the same type then each player receives his preferred outcome half the time, as in the correlated equilibrium of the complete information game; if the players are of different types, then the player of type b receives his preferred outcome with certainty. From an ex ante perspective, this decision rule maximizes the payoffs of the players; yet it is clearly not incentive compatible, since in a direct mechanism with $\mu = \delta_c$ each players would have an incentive to report his type as b regardless of his true type. By Lemma 1, then, since classical efficiency cannot be achieved with a symmetric ICDM, there is no symmetric communication process that enables the players to achieve classical efficiency.

On the other hand, the following mechanism, which we label the "flat mechanism," is always incentive compatible:

$$\mu_f(x, y; \cdot, \cdot) = \mu_f(y, x; \cdot, \cdot) = 1/2.$$

This ICDM ignores the reported types and simply guarantees an ex post efficient outcome. (Note that this is the <u>unique</u> ex post efficient symmetric decision rule.) Clearly μ_f is incentive compatible, since neither type gains from dishonesty or disobedience. The interim payoffs for each player are then $(t_i+1)/2$. If we let δ_n denote the decision rule associated with the ("no communication") symmetric Bayesian equilibrium in Proposition 1, we get the following result.

Proposition 2. For all values of a, b, and p, μ_f interim (and hence ex ante) dominates δ_n.

Proof. From Table 1, the highest payoff for an a-type under δ_n is b/(b+1), which is strictly less than (a+1)/2 since a > 1; thus an a-type is better off under μ_f. Similarly, the highest payoff for a b-type is b/(a+1). Since a > 1, (b+1)(a+1) > 2b, which implies that (b+1)/2 > b/(a+1); thus a b-type too is better off under μ_f. □

We now know that a simple use of an outside mediator, namely one who employs the rule μ_f, provides improvements over the (symmetric) equilibrium play, without communication between the players, of the game G. The remaining question then concerns whether μ_f itself can be improved upon. Note that the flat mechanism μ_f and the classically efficient decision rule δ_c differ only when the types differ, in which case δ_c biases the outcome in the b-type's favor. Hence any improvements over the flat mechanism must include such a bias. However, since δ_c is not incentive compatible there must be an offsetting loss in ex post efficiency to maintain honesty, in particular the honesty of an a-type. In other words, gains in ex ante efficiency come at the expense of ex post efficiency.

Our main results are embodied in the next two propositions. They characterize incentive-efficient symmetric mechanisms and allow us to calculate the maximum payoffs achievable for comparison with the payoffs from δ_n, μ_f, and δ_c. For convenience we write $Pr(t_j)$ for the probability that player j is of type t_j, that is, 1-p or p as appropriate, and we designate the following criterion values for p:

$$\bar{p} = \frac{1}{a+1} - \frac{a-1}{b-a} \quad \text{and} \quad \bar{\bar{p}} = 1 - \frac{(b+1)(a-1)}{(a+1)(b-a)}$$

(note that neither $\bar{p}$ nor $\bar{\bar{p}}$ need lie in [0, 1]). Propositions 3 and 4 are proved simultaneously in an appendix. Proposition 3 gives the conditions under which the flat mechanism is incentive efficient, and derives some important general characteristics of an ex ante incentive efficient symmetric ICDM. Notice that by Lemma 2, "efficiency" in both propositions is relative to the set of all decision rules.

Proposition 3. If μ^* is a symmetric, ex ante efficient ICDM for the game G, then the following conditions hold:

(3.1)
$$\sum_{t_{-i}\in T_{-i}} \Pr(t_{-i}) \sum_{d\in D} \mu^*(d;\ a,\ t_{-i})\cdot u_i(d;\ a,\ t_{-i})$$
$$= \sum_{t_{-i}\in T_{-i}} \Pr(t_{-i}) \sum_{d\in D} \mu^*(d;\ b,\ t_{-i})\cdot u_i(d;\ a,\ t_{-i}), \quad \forall\ i \in \{1,\ 2\};$$

(3.2) $\mu^*(x,\ y;\ a,\ a) = \mu^*(y,\ x;\ a,\ a) = 1/2;$

(3.3) if $p > \bar{\bar{p}}$, then

$$\mu^*(x,\ y;\ t) = \mu^*(y,\ x;\ t) = 1/2, \quad \forall\ t \in T;$$

(3.4) if $p < \bar{\bar{p}}$ and

(i) $p \geq \frac{a-1}{a+1}$, then $\mu^*(x,\ y;\ b,\ a) = 1$

and $\mu^*(x,\ y;\ b,\ b) + \mu^*(y,\ x;\ b,\ b) < 1$;

(ii) $p < \frac{a-1}{a+1}$, then $\mu^*(x, y; b, a) \in (1/2, 1)$

and $\mu^*(x, y; b, b) + \mu^*(y, x; b, b) = 0.$

Proof. See Appendix.

Condition (3.1) says that in an efficient mechanism if player i is an a-type he will be indifferent between honestly and dishonestly reporting his type, conditional on the other player being honest and both players obediently following the mediator's suggested actions. Condition (3.2) says that a-types always coordinate, each coordinated outcome occurring with probability 1/2 (the latter is a result of symmetry); (3.3) says that if the prior probability that the players are b-types is sufficiently high, then the efficient mechanism is flat. Note that $\bar{\bar{p}}$ is not necessarily in the interval (0,1); thus, if a = 2 and b = 3.5, then $\bar{\bar{p}} < 0$ so that for all values of the prior p the flat mechanism μ_f is incentive efficient, while if a = 1.01 and b = 10, then $\bar{\bar{p}} > 1$ and for no value of p is μ_f incentive efficient. Finally, condition (3.4) notes that for lower values of p, when the players report different types the b-type's preferred outcome occurs more than half the time, and in some cases all of the time; and that if both players are b-types, they may fail to coordinate, and under some circumstances must always fail. This possibility of failure, "caused" by the mediator operating according to μ^*, is what keeps type-a players honest in the reporting stage.

The following straighforward application of Lemma 1 demonstrates the breadth of the characterization in Proposition 3.

Corollary 1. Let δ be any symmetric, incentive-efficient decision rule for G. Then δ obeys the same conditions as does μ^* in Proposition 3.

The interpretation of (3.1) is now rather strained: under any such δ, an a-type is indifferent between being treated as an a-type or as a b-type. The other three conditions, however, now directly provide important properties true of all members of this important class of decision rules.

A second immediate result from Proposition 3 concerns the need for privacy in the communication of the mediator's suggestions to the players in an ICDM. We can say exactly when public communication of suggestions will be sufficient.

Corollary 2. Public announcement of suggestions by the mediator is capable of providing ex ante efficiency in a symmetric mechanism if and only if $p > \bar{\bar{p}}$ (that is, if and only if the flat mechanism is efficient).

Proof. Clearly public announcement can be used to implement the flat mechanism since then any unilateral deviation by a player would lead to a coordination failure and a lower payoff. If the flat mechanism is optimal, we are done. If not, then by (3.4) the mediator must sometimes induce coordination failures to achieve ex ante efficiency; but this is impossible if suggestions are announced publicly, since players lose nothing by disobeying such a suggestion. □

In general the ex ante incentive-efficient symmetric ICDM is not unique, as shown in the Appendix. This non-uniqueness is due to the possibility of "trading off" certain types of inefficiencies to achieve the same payoff results while maintaining incentive compatibility. The following proposition, however, completely characterizes one such efficient mechanism. Using this characterization, we can calculate the maximum payoff achievable through any symmetric communication process.

Proposition 4. The following symmetric ICDM μ^* is ex ante incentive efficient: $\mu^*(x, y; a, a) = \mu^*(y, x; a, a) = 1/2$, $\mu^*(y, y; b, a) = \mu^*(y, y; b, b) = 0$, and

(4.1) if $p \geq \bar{\bar{p}}$, then $\mu^* = \mu_f$;

(4.2) if $p < \bar{\bar{p}}$ and $p \geq \dfrac{a-1}{a+1}$, then $\mu^*(x, x; b, a) = 0$, $\mu^*(x, y; b, a) = 1$, and $\mu^*(x, x; b, b) = \dfrac{a-1}{p(a+1)}$;

(4.3) if $p \in [\bar{p}, \bar{\bar{p}})$ and $p < \dfrac{a-1}{a+1}$, then $\mu^*(x, x; b, a) = 0$, $\mu^*(x, y; b, a) = \dfrac{1}{2} + \dfrac{(a+1)p}{2(a-1)}$, and $\mu^*(x, x; b, b) = 1$;

(4.4) if $p < \bar{p}$ and $p < \dfrac{a-1}{a+1}$, then $\mu^*(x, x; b, a) = \dfrac{a-1-(a+1)p}{2[a-(a+1)p]}$, $\mu^*(x, y; b, a) = 1 - \mu^*(x, x; b, a)$, and $\mu^*(x, x; b, b) = 1$.

Proof. See Appendix.

Statement (4.1) matches (3.3) above; namely, if $p > \bar{\bar{p}}$ then the only ex ante incentive efficient mechanism is flat. Statement (4.2) follows from (3.4)(i) as well as allocating all of the probability of failure if the types are (b, b) to the outcome (x, x). Under the conditions in (4.3), even if the players' types differ they still coordinate with certainty, although not always on the b-type's preferred outcome; further, if both are b-types they will certainly fail to coordinate. Finally, (4.4) identifies conditions under which, if the players are different types, the outcome is either on the b-type's preferred outcome or on the ex post inefficient outcome (x, x), but never on the a-type's preferred outcome.

The proof of Proposition 4 reveals one further important fact about the maximization of payoffs using a mediator. Suppose that the mediator, rather than merely suggesting how the players should play in G, is able costlessly to <u>enforce</u> those suggestions. It turns out that no more can be gained from such enforcement than can be had from the ICDM described in Proposition 4.

Corollary 3. Suppose that G^+ is further augmented so that the mediator's suggestions are costlessly enforced. Call the latter enforcement game G^e. Then the maximum achievable ex ante expected payoff from G^e is identical to that generated in G^+ by the ICDM described in Proposition 4.

Proof. This is the situation covered by the revelation principle in its original form (Dasgupta, Hammond, and Maskin 1979; Myerson 1979). The optimal decision rule of this type is found by maximizing ex ante expected utility subject to honesty constraints. But the proof of Proposition 4 does exactly this, having shown that the obedience constraints are not binding for this ICDM. □

Intuitively, the mediator can accomplish no more with costless enforcement than he can with an ICDM because he still faces the problem of enticing truthful reports from the players. Corollary 3 demonstrates that it is always possible to solve the truthfulness problem in such a way that obedience is already taken care of. Thus, in this setting of asymmetric information, once we have learned what decision rules and payoffs are possible using an ICDM, we have also learned all there is to know about decision rules attainable with enforcement.

Another implication of Corollary 3 bears noting. The assumption of perfect enforcement means that the mediator is free to choose any of the outcomes of G without regard to the game-theoretic structure of G. So given this set of available outcomes, and given that the players can communicate and use mediation, it is irrelevant for purposes of maximizing expected payoff whether or not the players are forced to play the battle-of-the-sexes game at all. Indeed, Proposition 4 and Corollary 3 together show that one way for the perfectly enforcing mediator to induce maximal payoffs would be for the mediator to force the players to engage in G after using the mechanism μ^*.

5. Efficiency Without a Mediator

We have derived an incentive efficient decision rule via the revelation principle by examining communication processes aided by impartial outside actors. Our next problem is to identify when such mediation is needed, that is, to characterize those situations in which the players can achieve an efficient outcome simply by communicating with one another. This is an important difference, since it may be hard to imagine such impartial mediation in some real-life situations. For the incentive-efficient direct mechanism above, however, the mediator is potentially crucial because it may be necessary to keep secret both the report of each player and the recommended action for each player. Without such secrecy, the players could never be induced to fail to coordinate, and thus the temptation for an a-type to report falsely could not be overcome. The question now is whether, or under what conditions, some alternative communication process can accomplish the same expected payoffs without such secrecy.

We concentrate here on another special form of the augmented game G^+ introduced in section 4. Assume now that each player's set C_i of pure communication strategies is just a finite set of possible <u>messages</u> that player i could send to -i. Throughout this section we maintain symmetry of the game by assuming $C_1 = C_2$, and write $C = C_1 \times C_2$. Following the sending of messages, the players simultaneously decide on actions $d_i \in D_i$ as before. Thus again $\rho_i'(c_i;\ t_i)$ is player i's probability of using message c_i when his type is t_i, and $\alpha_i'(d_i;\ t_i,\ c)$ is i's probability of taking decision d_i when i's type is t_i and the players have previously communicated messages $c = (c_1, c_2)$ to one another.[3]

In terms of the model developed thus far, then, our aim is to determine whether or not, given a basic coordination game G, there exists a symmetric, unmediated communication process C and a symmetric equilibrium in the resulting game G^+ that generates an incentive efficient decision rule for G. We restrict our attention to symmetric processes and rules since otherwise it is always possible to generate trivial incentive-efficient decision rules by allowing only one player, in effect, to communicate. Using the same notation as in Proposition 3, the answer can be derived from the following result:

Proposition 5. A necessary and sufficient condition for there to exist a message set C whose associated augmented game G^+ has a symmetric equilibrium (ρ^*, α^*) such that the resulting decision rule is ex ante incentive efficient is that $p \geq \bar{\bar{p}}$.

Proof. (Sufficiency.) If $p \geq \bar{\bar{p}}$, then by (3.3) the flat mechanism is efficient. Let $C_i = \{m', m''\}$, and for each $i \in \{1, 2\}$ and each $t_i \in T_i$ consider the following message and decision strategies:

$$\rho_i^*(m'; t_i) = \rho_i^*(m''; t_i) = 1/2 ,$$

and

$$\alpha_1^*(x; t_1, c_1, c_2,) = \alpha_2^*(y; t_2, c_1, c_2) = \begin{cases} 1 \text{ if } c_1 = c_2 \\ 0 \text{ if } c_1 \neq c_2 , \end{cases}$$

with the complementary probability weights in α_i^* being placed on the other action. This is analogous to the "heads and tails" strategies used in the complete-information game in section 2 above. It is easily seen that these strategies constitute an equilibrium to the game G^+; further, each of (x, y) and (y, x) occurs half the time regardless of the players' types, thus implementing the flat mechanism of (3.3), which is ex ante incentive efficient as required.

(Necessity.) Let $p < \bar{\bar{p}}$ and suppose that an incentive-efficient decision rule is generated by message set C and symmetric equilibrium (ρ^*, α^*). For $j \in \{a, b\}$ define $C^j = \{m' \in C_i: \rho_i^*(m'; j) > 0\}$, so that C^a are the messages sent by a-types and C^b are those sent by b-types. By (3.2), $\forall$ $(c_1, c_2) \in C^a \times C^a$, either $\alpha_1^*(x; a, c_1, c_2) = \alpha_1^*(y; a, c_1, c_2) = 1$ or both = 0. Thus whenever a message pair from $C^a \times C^a$ is sent, the players coordinate with certainty. Further, by the symmetry assumption, for any $c_{-i} \in C^a$ player i's message strategy $\rho_i^*(\cdot; a)$ is such that the outcome (x, y) occurs with probability 1/2. This leads to the following result, of which we will make later use:

$$C^b \setminus C^a \neq \phi. \qquad (5)$$

To prove this, suppose the contrary; then for every $c_1 \in C^b$ the probability that (x, y) occurs, if 1 is a b-type and 2 is an a-type, is 1/2. From (3.4), however, incentive efficiency requires that this probability be strictly greater than 1/2, a contradiction.

Using (5), then, let $m' \in C^b \setminus C^a$, and suppose that $p < (a-1)/(a+1)$. Then if (m', m') is observed (which occurs with positive probability), it becomes common knowledge that both players are b-types. Thus the expected utility to a b-type player, conditional on the opponent being a b-type, is strictly positive, since following (m', m') the players must play an equilibrium, any of which (see Proposition 1) would yield positive expected payoffs to both players. But this contradicts (3.4)(ii). Thus for $p < (a-1)/(a+1)$ this Proposition is proved.

Suppose finally that $p \geq (a-1)/(a+1)$. To complete the proof we will use the fact that now

$$C^a \cap C^b = \phi. \qquad (6)$$

To prove (6), suppose that there were a message $m' \in C^a \cap C^b$. Then from the proof of (5) we know that the probability of (x, y) occurring if either player is type b and sends m' is 1/2. But this would contradict (3.4)(i), so there can be no such m'. By (6), then, if an incentive efficient equilibrium exists in this case it must be separating. Further, since the efficient mechanism is not flat, there is a message pair $c \in C^b \times C^b$ such that upon observing c the players adopt the complete information mixed strategy: $\alpha_i^*(x;b,m) = b/(b+1)$.

By (3.1), an a-type is indifferent in an optimal mechanism between honesty and dishonesty, given honesty by the other player and given obedience. In the posited incentive-efficient equilibrium (ρ^*, α^*) to the game G^+, this implies that an a-type is indifferent between using the strategies $(\rho_i^*(\cdot; a), \alpha_i^*(\cdot;a, \cdot))$ and $(\rho_i^*(\cdot, b), \alpha_i^*(\cdot; b, \cdot))$. However, for the case $p \geq (a-1)/(a+1)$ we can construct a strategy $\alpha_i'(\cdot;a, \cdot)$ which generates a strictly higher payoff than α_1^* if player 1 is an a-type, contradicting the assumption that (ρ^*, α^*) constituted an equilibrium in G^+. The new strategy is as follows: 1 plays according to α_1^* except when $t_1 = a$ and the message pair (m', m'), where $m' \in C^a \cap C^b$, whose existence is guaranteed by (6), is observed. In that case, use $\alpha_1'(x; a, m') = 0$. This gives an a-type player 1 a strictly higher payoff in this event than following the b-type's mixed strategy, so (ρ^*, α^*) cannot have been an equilibrium. □

Proposition 5 shows that the only situation in which a mediator is not required in order to achieve efficiency is when the mediator's role would be in some sense trivial; when $p \geq \bar{\bar{p}}$, the mediator ignores the reporting of types and simply assures an ex post efficient outcome. This result can be contrasted with that of Mathews and Postlewaite (1989), who study the role of bilateral communication in two-person sealed-bid double auctions. They show that any incentive compatible decision rule that is in addition ex post individually rational can be achieved as an equilibrium of some bilateral communication game; no mediation is ever required. Their sufficiency argument is similar to ours, in that behavior in the communication stage is "mixed" and chosen so as to correlate on particular outcomes with the required probabilities. The key feature differentiating our result from that of Matthews and Postlewaite is their assumption that the players' payoffs are in effect linear in money: the utility derived by the seller is just the payment by the buyer minus the value of the object sold, and analogously for the buyer. Total payoff is just the difference between the players' true valuations, times the probability of trade. In our model, on the other hand, the "total" payoff when the players are of different types depends on which coordinated outcome is chosen, and is higher when the b-type is favored. Thus while distributional features in our mechanisms have an effect on ex ante expected utility, analogous features of the Matthews-Postlewaite model have no such consequences. In our model, no mediator is needed just in order to achieve a maximal amount of coordination regardless of distribution, just as none is needed to assure maximal realization of gains from trade in their model. However, if each player in the bargaining model could vary in the (nonlinear) rate of increase of utility in "money," a mediator might be required to achieve incentive efficiency in that model as well.

6. Discussion

Summary. Using the standard "battle of the sexes" game as a basic model, we have analyzed the effects of incomplete information and communication upon the payoffs in coordination games. In the process we have discovered interesting features of optimal communication processes. We learned first of all that, from an ex-ante standpoint, expected payoffs from the coordination game with incomplete information are higher than those that would result if the types were revealed and the appropriate complete-information symmetric equilibrium played. Incomplete information about payoffs, however, introduces a new efficiency consideration to the problem: by agreeing ex ante to favor the player facing the higher stakes, the players could in principle receive higher expected payoffs than if they were treated symmetrically regardless of their types. To achieve such efficiency gains, however, the players must somehow reveal their private information about payoff values. Thus communication serves two purposes: the coordination of actions and the revelation of private information about payoff values.

To understand how communication could best be used for these purposes, we applied a version of the revelation principle to construct a mechanism that generates the highest payoffs achievable through <u>any</u> communication process. Generally this optimal communication mechanism requires the presence of an impartial mediator, but under certain conditions joint randomization alone (the "flat mechanism") constitutes such an optimal mechanism. More generally, we demonstrated the tradeoffs between, on the one hand, the ability to achieve ex ante distributional efficiency by using information about true payoffs and, on the other hand, degrading coordination in order to provide the incentive for players to reveal that information. The presence of these tradeoffs implies that no communication process can enable the players to achieve all the available gains from coordination.

Once we had derived an optimal mechanism, we were able to state several important properties of symmetric efficient decision rules and of the abilities of the mediator. Proposition 3 and Corollary 1 showed, among other things, that any symmetric, incentive-efficient decision rule must fail to coordinate the players' actions with certain probability in certain situations. Corollary 2 showed that the mediator's ability to communicate privately the suggested actions to the players is critical exactly when the flat mechanism is not optimal; in light of Proposition 5, this means that whenever a mediator is required for efficiency, privacy of communication is required also. Corollary 3 showed that the mediator need not have any ability to enforce those suggestions -- even costless enforcement would yield no further gain in ex ante expected payoffs.

The impartial mediator in this optimal communication mechanism seems inappropriate for some applications. At the opposite extreme, then, we examined what could be accomplished through communication between the players only, with no mediation of any sort. In the case of a simple coordination problem under complete information, we saw that by communicating without the use of an arbitrator or any jointly verified randomizing device, the players could coordinate fully. This ability depends on the players being able to label the available messages, that is, to associate them with ultimate outcomes, in any fashion they wish, contrary to the usual approach in "cheap talk" models.

When there is incomplete information about payoffs, the mediator is dispensible only in the situation in which the mediator's function would be, in effect, merely to flip a coin and dictate a coordinated outcome. Such joint randomization can be achieved through bilateral communication alone. If instead the optimal mediation mechanism takes the players' messages into account in any way, then the mediator is necessary in order to achieve incentive-efficient expected payoffs. In general then, if the players rely on bilateral communication alone, the highest possible expected payoff will be only a "third best" solution to their coordination problem.

Effects of Communication on Payoffs. We can gain some insight into the effects of incomplete information and communication by comparing the graphs of ex ante payoffs from various decision rules as functions of p for particular values of a and b (Figures 1-3). Several relationships hold true in all these graphs. First, as noted above the ex ante payoffs from δ_n, the decision rule generated by the symmetric Bayesian equilibrium in Proposition 1, exceed those from δ_o, the decision rule generated by mixed-strategy Nash equilibrium under complete information with varying payoffs. Second, Proposition 2 shows that the flat mechanism μ_f interim (and ex ante) dominates δ_n. Finally, Proposition 4 shows that the optimal mechanism μ^* will sometimes be identical to μ_f, and will never achieve as high a payoff level as δ_c, that is, will never be classically efficient (for $0 < p < 1$). In what follows, we examine the graphs of all these payoffs to explore the relative sizes of these differences, and to see the degree to which the flat mechanism can be improved.

Figures 1, 2, and 3 here

For large values of a, mediated communication yields very little improvement over the flat mechanism, for which mediation is unnecessary, relative to the shortfall between μ_f and classical efficiency. Figure 1 shows the ex ante expected payoffs from symmetric equilibria for various values of b when a = 5. (Notice that the vertical scales in Figures 1a, 1b, and 1c differ.) When b = 10, $\bar{\bar{p}} < 0$ so that the flat mechanism is always optimal anyway. When b = 50 or 100, the graph for μ^* is barely distinguishable from that for μ_f for smaller values of p, and $\bar{\bar{p}}$ itself is not very large -- it approaches an upper limit of 1/3 as $b \to \infty$. Figure 2 shows that even for an a-value of 2, the potential gains from mediation are still relatively small; again the gain is restricted by the fact that $\bar{\bar{p}} < 2/3$ for a = 2. For smaller values of a, however, as shown by Figure 3 where a = 1.1, mediation comes close to classical efficiency and improves greatly on the flat mechanism, especially for large values of b and moderately small values of p. Qualitatively, all this suggests the following generalization: mediated communication is most profitable when the importance of achieving coordination at all outweighs the importance of coordinating on one's own favorite outcome (a is not much bigger than 1) much of the time (p not too large). This is especially true if, for the occasional exception when the player is a b-type, the value of obtaining one's favorite outcome becomes extremely high (b is large).

Notice, secondly, that when both a and b are small there is a big difference between the payoffs from any kind of communication (μ^* and μ_f) and the no-communication payoffs (from δ_n and δ_o). This is illustrated in Figure 3 and, to a lesser extent, in Figure 2. As b grows, however, the payoff under incomplete information with no communication increases far above that under complete information, approaching that of the flat mechanism for values of p in the low interval P_1, which itself increases in size. For small a and small-to-moderate values of p, then, we can make the following generalization: if b is only slightly greater than a, communication is important; but if b is much larger than a, the important thing is that the types of the players not be <u>verifiably</u> revealed -- the self-discrimination possible under incomplete information may be more important than communication.

Application to Entry Problems. As an illustration of how our results bear on coordination problems in the real world, consider a two-firm, market entry decision similar to those modeled by Dixit and Shapiro (1985) and Farrell (1987), described in the Section 1 above. A market is to become available for entry by either of two firms; the market will support only one firm profitably, and due to startup costs the entry decisions of the firms must in effect be made simultaneously. At the time the market becomes available, each firm learns its true production costs for operating in that market in the long run. These production costs are drawn randomly and independently by the two firms; they may be either "high" (with probability 1-p) or "low" (probability p), and are private information to the firm. Finally, if only one firm enters the market, the other firm will be able to make some small profits by selling a complementary good produced at known cost. The possible outcomes are: neither firm enters, giving each a profit of 0; both firms enter, and neither makes any profit (again, for simplicity, 0); and one firm enters, making a profit of π_l if costs are low or π_h if costs are high, and the other firm makes a profit of π, where $\pi_l > \pi_h > \pi$. These three payoffs correspond to the values b, a, and 1, respectively, in our game G.

In advance of learning the costs of the market to be contended, the firms may contemplate some form of regulation or communication to improve their ex ante expected profits. Note first of all that, if the firms have the option of verifiably revealing their true costs to one another before making the entry decision, they should nevertheless prefer <u>not</u> to do so --

their expected profits are higher in any case from playing the game under incomplete information than under complete information about costs, because in using the symmetric equilibrium strategies in Proposition 1 they will successfully coordinate more often than if they played the full-information mixed-strategy equilibrium.

Assuming that the firms do operate under incomplete information, there remains the opportunity to design a communication scheme that will maximize expected profits. If π_h is much greater than π, then the flat mechanism is optimal (Proposition 3 and Figure 1) -- the firms are best advised to engage in a simple joint randomization and ignore the distributional problem of who would profit more from being the monopolist.

However, if π_h is not much greater than π then by Proposition 3 the flat mechanism is not always best (as in Figure 3). If in addition π_l is relatively large and p is not too large, then considerable improvements in ex ante expected profits are avaiable (as in Figure 3c) and a more complex communication scheme is in order. If the scheme is to maximize ex ante expected profits, it must involve the use of an outside mediator in some capacity (Proposition 5). Further, if the resulting decision rule is to be symmetric, it must obey the conditions of Proposition 3 and Corollary 1: in particular, the probabilitiy of an uncoordinated outcome (dual entry or no entry) must be positive whenever either firm has low costs, although not when both firms have high costs. Indeed, if $p < \frac{a-1}{a+1}$ then even when exactly one firm has low costs the probability of an uncoordinated outcome must be positive.

One can imagine a regulatory agency playing the role of this outside mediator. The firms would privately report their cost information to the agency, which subsequently advises each firm on whether or not to enter the market[4]. By employing an incentive compatible mechanism the agency leads the firms to report cost information truthfully as well as to carry out the agency's suggestions obediently, improving expected profits for both firms. By Corollary 3, of course, this improvement is the same as could be expected if the agency had enforcement powers. We can view Proposition 5, then, as identifying a source of the demand for regulation, differing from that suggested by "capture theories" of regulation (Bernstein 1955; Stigler 1971). In the latter, the demand for regulation follows when regulation permits firms to cartelize markets, whereas here we predict that, even in the absence of this incentive, firms may prefer to create a regulatory system when such a system would facilitate the efficient coordination of the firms' behavior. Note that in the present context this regulation is not necessarily bad from the consumer's point of view, since the contested market is assumed to be naturally monopolistic. Indeed, by reducing the probability of the "no entry" outcome, the regulation might increase consumer surplus.

Conclusion. Substantively, our analysis provides elements of a theory of institutions and institutional design when a group of actors faces a coordination problem. The design of institutions takes place in an ex ante process in which individuals attempt to determine a communication mechanism through which coordination will be consistent with rational behavior by the participants. Such a mechanism becomes an "institution" only if the players have no incentive to change the mechanism. Further development of such a theory of institutions will require specification of the process in which the mechanism is created and changed, including considerations of "durability" (Holmstrom and Myerson 1983; Crawford 1985) and the possible role of an arbitrator who, unlike ours, is not impartial (such as the "principal" in Myerson 1982). The work of Myerson (1982) on "optimal coordination mechanisms" (which does not focus on coordination as we define it) is similar to such a model of institutional creation, except that here we want the entire group of "agents" to produce the mechanism, with no exogenously designated "principal." Such a model would offer a wealth of important applications in social science, in areas as diverse as the evolution of economic institutions (North 1981), the development of social norms (Taylor 1987), and the creation of political constitutions (Buchanan 1975).

Appendix

Proof of Propositions 3 and 4

Symmetric mechanisms can be described by seven variables, $v_1, \ldots, v_7$, defined as follows:

$\mu(x, x; a, a) = v_1$

$\mu(y, y; a, a) = v_2$

$\mu(x, y; a, a) = \mu(y, x; a, a) = 1/2(1-v_1-v_2)$

$\mu(x, y; b, a) = \mu(y, x; a, b) = v_3$

$\mu(x, x; b, a) = \mu(x, x; a, b) = v_4$

$\mu(y, y; b, a) = \mu(y, y; a, b) = v_5$

$\mu(y, x; b, a) = \mu(x, y; a, b) = 1-v_3-v_4-v_5$

$\mu(x, x; b, b) = v_6$

$\mu(y, y; b, b) = v_7$

$\mu(x, y; b, b) = \mu(y, x; b, b) = 1/2(1-v_6-v_7)$.

The <u>ex</u> <u>ante</u> expected payoff, given honesty and obedience, from a mechanism $\mu = \langle v_1, \ldots, v_7\rangle$ can then be written as

$$\pi(\mu) = (1-p)(a+1)[(1-p)(1-v_1-v_2)/2 + p(1-v_3-v_4-v_5)] + p(b+1)[p(1-v_6-v_7)/2 + (1-p)v_3].$$

For an a-type to report honestly, given that the other player will be honest and that both will be obedient, μ must satisfy the constraint

$$(a+1)p(v_6+v_7)/2 - (a+1)(1-p)(v_1+v_2)/2$$
$$- (a-1)(v_3-1/2) + [1-(a+1)p](v_4+v_5) \geq 0. \qquad \text{(C1)}$$

This condition is derived as follows: if an a-type reports honestly, his expected payoff is $(1-p)(a+1)(1-v_1-v_2)/2 + p[a(1-v_3-v_4-v_5) + v_3]$, while if he reports b his payoff is $(1-p)(av_3+1-v_3-v_4-v_5) + p(a+1)(1-v_6-v_7)/2$. The difference between these two expressions gives the left-hand side of (C1). Similarly we can derive the honesty condition for b-types, assuming obedience:

$$(v_3-1/2)(b-1) + (b+1)(1-p)(v_1+v_2)/2$$
$$- (b+1)p(v_6+v_7)/2 - [1-(b+1)p](v_4+v_5) \geq 0. \qquad \text{(C2)}$$

For an a-type player i to obey a recommendation of $d_i = x$, given honesty and given that the other player, j, will accept his recommended move d_j, we need to calculate the probabilities that the arbitrator will suggest $d_j = x$ and $d_j = y$, given that $d_i = x$. In general this conditional probability is given by the expression:

$$\Pr\{d_j=x \mid d_i=x\} = \frac{\Pr\{d_i=x \text{ and } d_j=x\}}{\Pr\{d_i=x\}}$$

and similarly for $\Pr\{d_j=y \mid d_i=x\}$. Ignoring the denominators (which are identical throughout the inequality below), the obedience condition for an a-type with $d_i = x$ becomes

$$a[(1-p)(1-v_1-v_2)/2 + p(1-v_3-v_4-v_5)] - (1-p)v_1 - pv_4 \geq 0. \qquad \text{(C3)}$$

Similarly, when $d_i = y$,

$$(1-p)(1-v_1-v_2)/2 + pv_3 - a[(1-p)v_2 + pv_5] \geq 0. \tag{C4}$$

For b-types the analogous obedience conditions are

$$b[(1-p)v_3 + p(1-v_6-v_7)/2] - (1-p)v_4 - pv_6 \geq 0 \tag{C5}$$

$$(1-p)(1-v_3-v_4-v_5) + p(1-v_6-v_7)/2 - b[(1-p)v_5 + pv_7] \geq 0. \tag{C6}$$

Thus constraints (C1) through (C6) give incentive compatibility conditions on feasible mechanisms μ. There should also be constraints so that neither type would prefer to be both dishonest <u>and</u> disobedient. These would add six more constraints to the problem: three for an a-type specifying that honesty and obedience is preferred to dishonesty combined with disobedience only when action x is suggested, with disobedience only when action y is suggested, or with both; similarly for a b-type. In what follows we ignore these constraints and solve for the relatively unconstrained optimal mechanism, and later show that at this optimum it is the case that both types will prefer to be obedient even if they are dishonest. Thus the ignored constraints will also be satisfied at the relatively unconstained optimum, so that this is a global optimum as well. For example, if (C3) and (C4) are altered by replacing "a" with "b", then we have the condition that a b-type will want to be obedient after reporting dishonestly; similar conditions are derived by replacing "b" with "a" in (C5) and (C6). Label these constraints (C3') - (C6'). Since b > a, if (C6) holds then (C6') holds as well, and if (C3) holds then (C3') holds also. So all we will need to show below is that (C4') and (C5') hold.

The final constraints on feasible mechanisms require that the relevant parameters generate well defined probabilities:

$$v_3 + v_4 + v_5 \leq 1 \tag{C7}$$

$$v_1 + v_2 \leq 1 \tag{C8}$$

$$v_6 + v_7 \leq 1 \tag{C9}$$

$$v_k \geq 0, \; k = 1,..,7. \tag{C10}$$

An <u>ex ante</u> efficient ICDM thus is one that solves the following program:

$$\text{P:} \quad \max_{\mu} \pi \quad \text{subject to (C1) through (C10).}$$

Denote a solution to program P as $\mu^* = \langle v_1^*, \ldots, v_7^* \rangle$.

From the definition of the ex ante payoff π we have that $\partial\pi/\partial v_3 > 0$, while $\partial\pi/\partial v_k < 0$, $k \neq 3$. These last signs follow from the fact that the variables v_k, $k \neq 3$ summarize the probability of ending up "off-diagonal", and as such are associated with lower payoffs. Further, since the flat mechanism is always incentive compatible, an efficient mechanism will have $v_3^* \geq 1/2$. Note also that, since v_k, $k \neq 3$ are probabilities of <u>failing</u> to coordinate, they do not contribute to a player's willingness to follow the suggested action; thus the obedience constraints (C3) - (C6) remain satisfied if we lower any of these variables. In particular, if the efficient mechanism is such that $\sum_{k\neq 3} v_k > 0$ it must be that either (C1) or (C2) holds with equality, but not both (both hold in the flat mechanism where $\sum_{k\neq 3} v_k = 0$).

Further, it is immediate that if either constraint holds with equality, the other will hold with strict inequality if $\sum_{k\neq 3} v_k > 0$ and $v_3 \geq 1/2$. We claim that in an efficient mechanism, (C1) is the constraint that holds with equality. To see this, suppose that (C2) holds with equality in an efficient mechanism; then it must be that $v_6^* = v_7^* = 0$, since othewise we could lower either variable by a sufficiently small amount, still satisfy (C1) - (C6), and thereby increase π. Since $v_3^* \geq 1/2$, (C2) will hold with equality only if the mechanism is flat or if $(b+1)p < 1$ and $v_4+v_5 > 0$. In the latter case, though, we could lower v_4 or v_5 by a sufficiently small amount, satisfy the constraints, and improve the payoff π. Thus, the ex ante efficient mechanism will be such that (C1) holds with equality, proving statement (3.1), so that by deriving the efficient mechanism on the hyperplane defined by (C1) we will have derived the efficient mechanism in general. Consequently we can ignore constraint (C2) in solving for the efficient mechanism, since (C2) holds when (C1) is satisfied with equality.

Constraint (C1) holding with equality also implies that $v_1^* = v_2^* = 0$, since if not we could lower these variables by a small amount, still satisfy (C1) - (C6), and increase the expected payoff; thus statement (3.2) follows. Hence in any efficient mechanism, if both players are a-types they will always coordinate, and do so with equal probability on either player's preferred outcome. Since the optimal values of v_1 and v_2 are fixed at zero, we can ignore these variables in solving for the efficient mechanism.

Along the hyperplane determined by (C1), we can write v_3 as a function of $(v_4, \ldots, v_7)$:

$$v_3(v_4, \ldots, v_7) = 1/2 + \frac{1}{a-1} [\frac{a+1}{2} p(v_6+v_7) + (1-(a+1)p)(v_4+v_5)].$$

Let $\Omega(p)$ denote the set of values $(v_4, \ldots, v_7)$ such that $v_3(v_4, \ldots, v_7) \geq 1/2$ and $v_k \geq 0$, $k = 4,\ldots,7$. For convenience define $\pi'(v_4, \ldots, v_7)$ as $\pi(0, 0, v_3(v_4, \ldots, v_7), v_4, \ldots, v_7)$. Then any solution to the following program will also be a solution to program P above:

$$P': \max_{\Omega(p)} \pi'(v_4, \ldots, v_7) \quad \text{subject to (C3)-(C6), (C7), (C9).}$$

Note that π' is additively separable and linear in all of its arguments. In what follows we will ignore constraints (C3) and (C4) as well, so that after characterizing a candidate for an efficient mechanism it will remain to show that (C3) and (C4), as well as (C4') and (C5'), are satisfied.

The solution of the program P' can be thought of as a search for profitable deviations from the flat mechanism. From our analysis of <u>ex ante</u> classical efficiency, we know that any such improvement must be obtained by increasing the frequency of coordination on the b-type's preferred outcome when the players are of different types; that is, by increasing v_3 above 1/2. To satisfy constraint (C1), however, any increase in v_3 must be offset by an occasional failure to coordinate. Since we know that $v_1 = v_2 = 0$, this failure must occur in cases where at least one of the players is a b-type. The question, then, is where to induce the failure -- for instance, if we increase both v_6 and v_3 along the constraint (C1), does the resulting increase in payoff when the players are of different types offset the loss when both are b-types? The answer will depend on the value of p given (a,b); the separability of π' will allow us to examine individually the tradeoffs from increasing each of $v_4,\ldots,v_7$.

Differentiating π',

$$\frac{\partial\pi'}{\partial v_4} = \frac{\partial\pi'}{\partial v_5} = \frac{1-(a+1)p}{(a-1)} [p(1-p)(b+1) - p(1-p)(a+1)] - p(1-p)(a+1)$$

$$= p(1-p) \left[\frac{1-(a+1)p}{(a-1)} (b-a) - (a+1)\right].$$

Hence our value $\bar{p} = 1/(a+1) - (a-1)/(b-a)$ is that for which $\partial\pi'/\partial v_4 \gtrless 0$ as $p \lessgtr \bar{p}$; likewise for $\partial\pi'/\partial v_5$. Similarly,

$$\frac{\partial\pi'}{\partial v_6} = \frac{\partial\pi'}{\partial v_7} = \left[\frac{(a+1)p}{(a-1)} \right] [p(1-p)(b+1) - p(1-p)(a+1)] - p^2(b+1)$$

$$= p(1-p) \left[\frac{(a+1)p}{(a-1)} \right] (b-a) - p^2(b+1).$$

We have set $\bar{\bar{p}} = 1 - (b+1)(a-1)/(a+1)(b-a)$, so that $\partial\pi'/\partial v_6 \gtrless 0$ as $p \lessgtr \bar{\bar{p}}$; similarly for $\partial\pi'/\partial v_7$. It is immediate that $\bar{p} < \bar{\bar{p}}$; this plus the fact that $\Omega(p)$ includes the origin then establishes statement (3.3); namely if $p > \bar{\bar{p}}$, the efficient mechanism is flat: $\mu^* = \mu_f$. It also establishes statement (4.1) which says the same for $p \geq \bar{\bar{p}}$. (This is the first of several points at which strict and weak inequalities have been inserted arbitrarily in Proposition 4; the boundaries between statements (4.1) and (4.2), between (4.1) and (4.3), and between (4.3) and (4.4) are sources of nonuniqueness of the optimal mechanism, as the values of some of the v_k do not coincide there. The boundaries between (4.2) and (4.3) and between (4.2) and (4.4) are assigned completely arbitrarily, as the v_k-values do coincide.)

Suppose now that $p < \bar{\bar{p}}$, so that an increase in v_6 or v_7 increases the <u>ex</u>

ante expected payoff π'. Let $C_5(v_4, v_6, v_7)$ denote the LHS of constraint (C5); then

$$\frac{\partial C_5}{\partial v_6} = pb[(1-p)(a+1)/(a-1) - 1]/2 - p,$$

so that $\partial C_5/\partial v_6 \gtrless 0$ as

$$b[(1-p)(a+1)/(a-1) - 1] \gtrless 2.$$

Since $p < \bar{\bar{p}}$, $(1-p) > [(b+1)(a-1)]/[(b-a)(a+1)]$, so that

$$b[(1-p)(a+1)/(a-1) - 1] > b[(b+1)/(b-a) - 1] = b(a+1)/(b-a) > 2.$$

Thus for $p < \bar{\bar{p}}$, $\partial C_5/\partial v_6 > 0$; a similar result holds for $\partial C_5/\partial v_7$. Since (C5) holds in $\Omega(p)$ when $v_6 = v_7 = 0$ (because $v_3 \geq 1/2$, $v_3 \geq v_4$, and $b > 1$), (C5) will be satisfied for all parameter values in $\Omega(p)$, so that we can ignore this constraint. Thus in determining the optimal values of $v_4, \ldots, v_7$ when $p < \bar{\bar{p}}$ the relevant constraints are (C6), (C7), and (C9) (we return subsequently to (C4)). We make the following claim: if there exists an efficient mechanism with $v_7 > 0$, then there exists another mechanism yielding the same payoffs with $v_7 = 0$. To see this, note that v_6 and v_7 enter symmetrically in constraints (C7) and (C9) through $v_3(\cdot)$ as well as entering symmetrically in the payoff π'. Hence a "shift" in probability from v_7 to v_6 will not effect these. Further, if (C6) is satisfied when $v_7 > 0$, then $v_7' = v_7 - \epsilon$ and $v_6' = v_6 + \epsilon$ satisfy (C6) as well, since the marginal impact of v_7 on (C6) is greater than that of v_6; this then proves the claim, and shows why in general the ex ante incentive efficient mechanism may not be unique. Thus for all $p < \bar{\bar{p}}$ we will set $v_7^* = 0$ in order to compute an optimal mechanism as in Proposition 4. An analogous argument allows us, without loss of generality, to set $v_5^* = 0$ as well. Thus constraint (C6) will always be satisfied, implying that the only relevant variables when $p < \bar{\bar{p}}$ are v_4 and v_6 and the only relevant constraints are (C7) and (C9).

Suppose now that $p \in [\bar{p}, \bar{\bar{p}})$. Since π' is then nonincreasing in v_4, a mechanism for which there does not exist $v_4' < v_4^*$ such that the resulting mechanism is feasible would be an optimal mechanism. Let $v_6^7(v_4)$ denote the value of v_6 such that constraint (C7) holds with equality given v_4. Since $v_6^7(v_4)$ is not necessarily decreasing in v_4, the feasible set need not be convex to the v_6-axis. Thus it is not immediately obvious that the optimal value of v_4 will be zero. To see that this will in fact be the case, we establish the effect on π' of changes in v_4 along the boundary determined by $v_6^7(v_4)$. Implicit differentiation of the equation where (C7) holds with equality gives us that

$$\frac{\partial v_6^7(v_4)}{\partial v_4} = \frac{-\partial v_3/\partial v_4 + 1}{\partial v_3/\partial v_6}; \text{ thus}$$

$$\left.\frac{\partial \pi'}{\partial v_4}\right|_{v_6 = v_6^7(v_4)} = p(b+1)\ \{(p/2)\ \frac{-\partial v_3/\partial v_4 + 1}{\partial v_3/\partial v_6}$$

$$+ (1-p)[\ \frac{\partial v_3}{\partial v_4} - \frac{\partial v_3}{\partial v_6}\ \frac{\partial v_3/\partial v_4 + 1}{\partial v_3/\partial v_6}\ ]\}$$

$$= p(b+1)\{[\ \frac{a-(a+1)p}{(a+1)}\ ] - (1-p)\}$$

$$= [p(b+1)/(a+1)][a - (a+1)] < 0.$$

Thus π' is decreasing in v_4 along the boundary determined by $v_6^7(v_4)$. This implies that for $p \in (\bar{p}, \bar{\bar{p}})$, $v_4^* = 0$.

The optimal value of v_6 for the program P' will then be equal to $\min\{1, (a-1)/[p(a+1)]\}$, depending on where the function v_6^7 intersects the v_6 axis. This then establishes statement (3.4)(i), statement (4.2) for $p \in [\bar{p}, \bar{\bar{p}})$, and statement (4.3), provided we can satisfy the "ignored" constraints. To see that these hold as well, note that $v_4^* = v_5^* = 0$ implies that (C3), (C4), and (C4') are all satisfied. The remaining constraint to check is (C5'), which we can rewrite as

$$a[(1-p)v_3 + p(1-v_6)/2)] \geq pv_6.$$

If $p < (a-1)/(a+1)$, then $v_6^* = 1$ and $v_3^* = \frac{1}{2} + \frac{(a+1)p}{2(a-1)}$, so that (C5') becomes

$$a(1-p)(a-1+(a+1)p) \geq 2p(a-1).$$

Now $p < \bar{p}$ implies that $(1-p) > (a-1)/(a+1)$, which in turn is greater than p, so it follows that $a(1-p)(a-1) > p(a-1)$ and $a(1-p)(a+1)p > p(a-1)$, and (C5') holds. If $p > (a-1)/(a+1)$ then $v_6 = \frac{a-1}{(a+1)p}$, and (C5') becomes

$$a[(1-p) + \frac{p}{2}(1 - \frac{a-1}{(a+1)p})] \geq \frac{a-1}{a+1}.$$

Cancelling terms, we get

$$a(a + 1 - p) \geq a-1,$$

so that (C5') holds for $p > (a-1)/(a+1)$ as well.

Finally, let $p < \bar{p}$, so that the positive indirect effect (through v_3) on π' of an increase in v_4 outweighs the negative direct effect. As above we can without loss of generality let $v_5^* = v_7^* = 0$. Further, since $\partial\pi'/\partial v_4$ is decreasing along the constraint defined by $v_6^7(v_4)$, if the function v_6^7 intersects the v_6 axis below $v_6 = 1$ then the optimal mechanism will be such that $v_4^* = 0$; otherwise $v_4^* > 0$. This establishes statements (3.4)(ii) and (4.4), and statement (4.2) for $p < \bar{p}$, again provided the "ignored" constraints are satisfied. As above, constraints (C3), (C4), and (C4') follow immediately; if $p < (a-1)/(a+1)$ constraint (C3) follows from the fact that for $p < \bar{p}$, $(1-p) > p$. What remains is to show that (C5') holds as well. Suppose $p < (a-1)/(a+1)$. Since $v_6^* = 1$ and $v_4^* = 1-v_3^*$, we can write (C5') as

$$(1-p)(a+1)v_3 \geq 1.$$

Substituting $v_3(v_4, \ldots, v_7)$ for v_3,

$$(1-p)(a+1)\left[1 + \frac{(a+1)p}{a-1} + \frac{1-(a+1)p}{a-1}\,\frac{a-1-(a+1)p}{a-(a-1)p}\right] \geq 2$$

Cancelling terms, we get

$$(1-p)(a+1)\ [(a-1)(a-(a+1)p) + a - 1] \geq 2(a-1)\ [a - (a+1)p],$$

or

$$[(1-p)(a+1)]^2 \geq 2[a - (a+1)p].$$

At $p = 0$ this holds as a strict inequality, while at $p = (a-1)/(a+1)$ this holds with equality. Further, the graph of the left-hand side crosses that of the right-hand side from above at $p = (a-1)/(a+1)$, so that $\forall\ p < (a-1)/(a+1)$ the above inequality holds. If $p > (a-1)/(a+1)$, then the argument for when $p \in (\overline{p},\overline{\overline{p}})$ shows that (C5') holds in this case as well.

Notes

1. Provided that the equilibria are not themselves strictly Pareto ranked, in which case an obvious focal solution is at hand. To prove that such a coordination problem always exists in a two-player game with multiple equilibria, suppose that there are two equilibria in which both players use different strategies. Without loss of generality, say that each player uses strategy 1 in the first equilibrium, and strategy 2 in the second. Let a_{ij} and b_{ij} be the payoffs to players A and B, respectively, when player A uses strategy i and player B uses strategy j. Since (1, 1) is an equilibrium, it must be that $a_{11} \geq a_{i1}$ for all i, and likewise $b_{11} \geq b_{1j}$. Since (2, 2) is an equilibrium, $a_{22} \geq a_{i2}$ and $b_{22} \geq b_{2j}$. If there is no strict Pareto relation between the two equilibria, we may also assume WLOG that $a_{11} \geq a_{22}$ and $b_{22} \geq b_{11}$. But then it is immediate that $a_{12} \leq a_{11}$ and $b_{12} \leq b_{22}$, and so the outcome (1, 2) is weakly Pareto-inferior to both (1, 1) and (2, 2). If the original Nash equilibrium conditions are satisfied with strict inequality for a player, then the inferiority of (1, 2) is strict for that player. Extension of this result to an arbitrary finite number of players is straightforward.

2. This is the same method as used by Matthews and Postlewaite (1989) to allow players to randomize over efficient allocations without the use of a mediator or of a jointly observable random device. Notice that the choice of this convention itself entails a coordination problem: do the players choose (x, y) when they match and (y, x) when they do not, or vice versa? At least

this derived coordination problem is of the no-conflict sort, where t = 1. With communication, such problems are relatively easy to solve, so it does not seem worrisome in the present context.

3. Since we focus on the Bayesian equilibria of G^+, this specification could cover as well any communication process in which the players have a sequence of opportunities to send such messages, by considering each C_i as a set of normal-form strategies as before. Notice that this would not be true of some equilibrium refinements, such as sequential equilibrium.

4. This raises the issues of (1) providing the agency with incentives to behave in an optimal manner, and (2) the ability of the agency to commit itself to carry out the designed mechanism. We postpone consideration of these important topics.

References

Aumann, R. (1974): "Subjectivity and Correlation in Randomized Strategies." Journal of Mathematical Economics 1: 67-96.

Bernstein, Marver (1955): Regulating Business by Independent Commission. Princeton: Princeton University Press.

Calvert, R. (1987): "Coordination and Power: The Foundation of Leadership among Rational Legislators." Photocopy, University of Rochester.

Crawford, Vincent P. (1985): "Efficient and Durable Decision Rules: A Reformulation." Econometrica 53: 817-835.

Crawford, Vincent P., and Hans Haller (1989): "Learning How to Cooperate: Optimal Play in Repeated Coordination Games." Photocopy, University of California, San Diego.

Dasgupta, P., P. Hammond, and E. Maskin (1979): "The Implementation of Social Choice Rules: Some Results on Incentive Compatibility." Review of Economic Studies 46: 185-216.

Dixit, A. and C. Shapiro (1985): "Entry Dynamics with Mixed Strategies." In L.G. Thomas, ed., The Economics of Strategic Planning. Lexington MA: Lexington Books.

Farrell, J. (1987): "Cheap Talk, Coordination, and Entry." Rand Journal of Economics 18: 34-39.

Farrell, J. and G. Saloner (1988): "Coordination through Committees and Markets." Rand Journal of Economics 19: 235-52.

Hardin, R. (1982): Collective Action. Baltimore: Johns Hopkins University Press.

Holmstrom, B. and R. Myerson (1983): "Efficient and Durable Decision Rules." Econometrica 51: 1799-1819.

Katz, M.L., and C. Shapiro (1985): "Network Externalities, Competition, and Compatibility." American Economic Review 75: 424-40.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." Journal of Economic Theory 27: 245-252.

Luce, R. Duncan, and Howard Raiffa (1957): Games and Decisions: Introduction and Critical Survey. New York: John Wiley and Sons.

Matthews, S. and A. Postlewaite (1989): "Pre-Play Communication in Two-Person Sealed-Bid Double Auctions." Journal of Economic Theory 48: 238-263.

Myerson, R. (1979): "Incentive Compatibility and the Bargaining Problem." Econometrica 47: 61-74.

Myerson, R. (1985): "Bayesian Equilibrium and Incentive Compatibility: An Introduction. In L. Hurwicz et. al., eds., Social Goals and Social Organization: Essays in Memory of Elisha Pazner. Cambridge: Cambridge University Press.

Myerson, R. (1986): "Multistage Games with Communication." Econometrica 54: 323-58.

Palfrey, T., and H. Rosenthal (1988): "Testing for Effects of Cheap Talk in a Public Goods Game with Private Information." Photocopy, California Institute of Technology.

Schelling, T. (1960): The Strategy of Conflict. Cambridge MA: Harvard Unviersity Press.

Stigler, George (1971): "The Economic Theory of Regulation." Bell Journal of Economics 2: 3-21.

Vickers, J. (1986): "Signaling in a Model of Monetary Policy with Incomplete Information." Oxford Economic Papers 38: 443-455.

Table 1. **Expected payoffs from G for each player-type, and the relations among them.**

For $p \in P_1$:

Player	Type	expected payoff in equilibrium (1.0)		(1.1)
i	a	a	<	a/(a+1)
i	b	b	<	b/(a+1)
-i	a	1	<	a/(a+1)
-i	b	1	*	b/(a+1)

For $p \in P_2$:

Player	Type	expected payoff in equilibrium (1.0)		(1.2)(i)		(1.2)(ii)†
i	a	a	>	p	≤	b/(b+1)
i	b	b	>	b(1-p)	≥	b/(b+1)
-i	a	1	>	p	>	1/(a+1)
-i	b	1	**	b(1-p)	≤	b/(a+1)

for $p \in P_3$:

Player	Type	expected payoff in equilibrium (1.0)		(1.3)
i	a	a	>	b/(b+1)
i	b	b	>	b/(b+1)
-i	a	1	>	b/(b+1)
-i	b	1	>	b/(b+1)

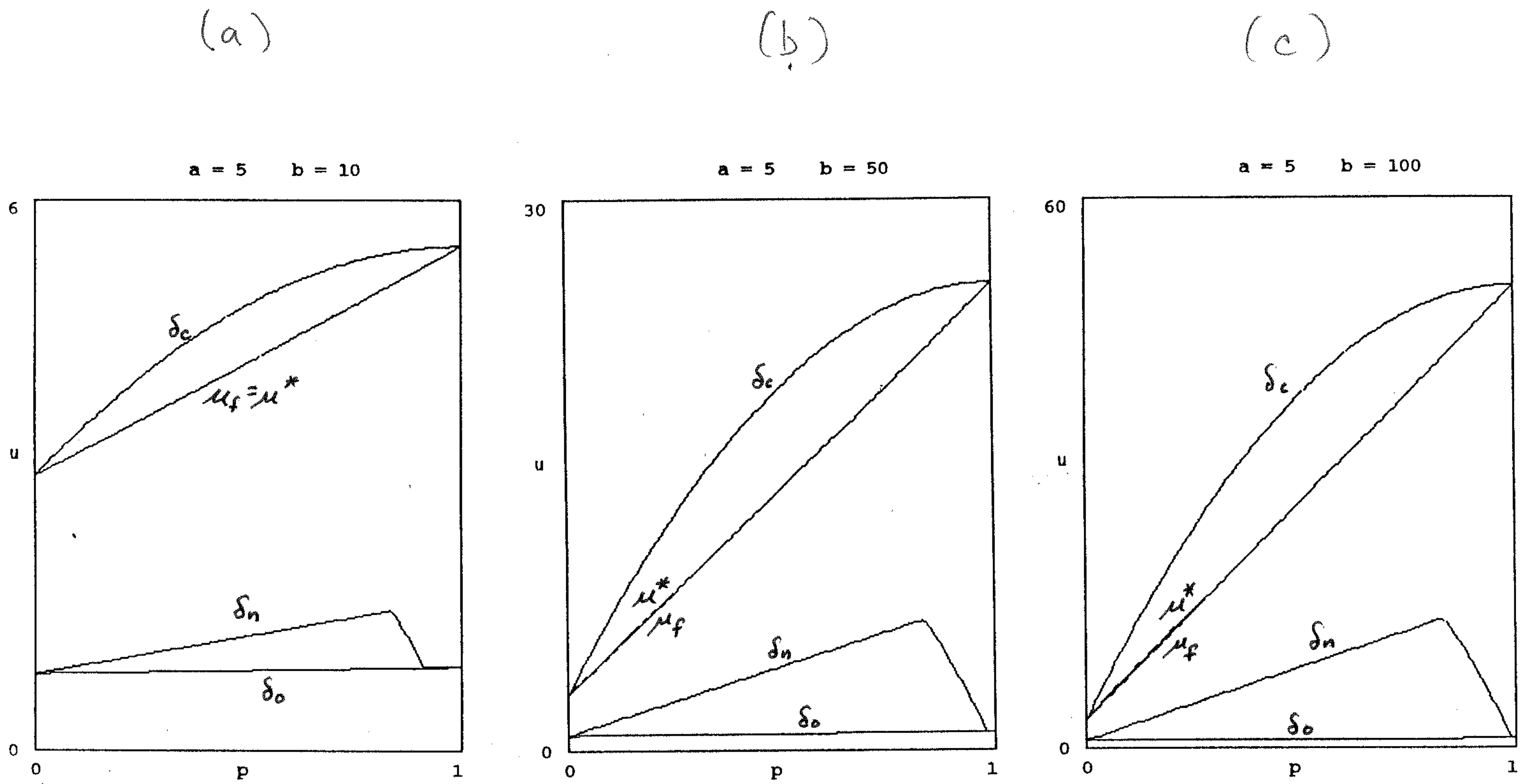
Notes:

† The payoffs for equilibria described in (1.4) are identical to those for (1.2)(ii).

* This relation is "≥" if and only if $b-a \leq 1$.

** This relation is "≥" if and only if $p \geq (b-1)/b$.

Figure 1. Payoffs from Various Decision Rules when $a=5$.



Key:

δ_c : classically efficient symmetric decision rule

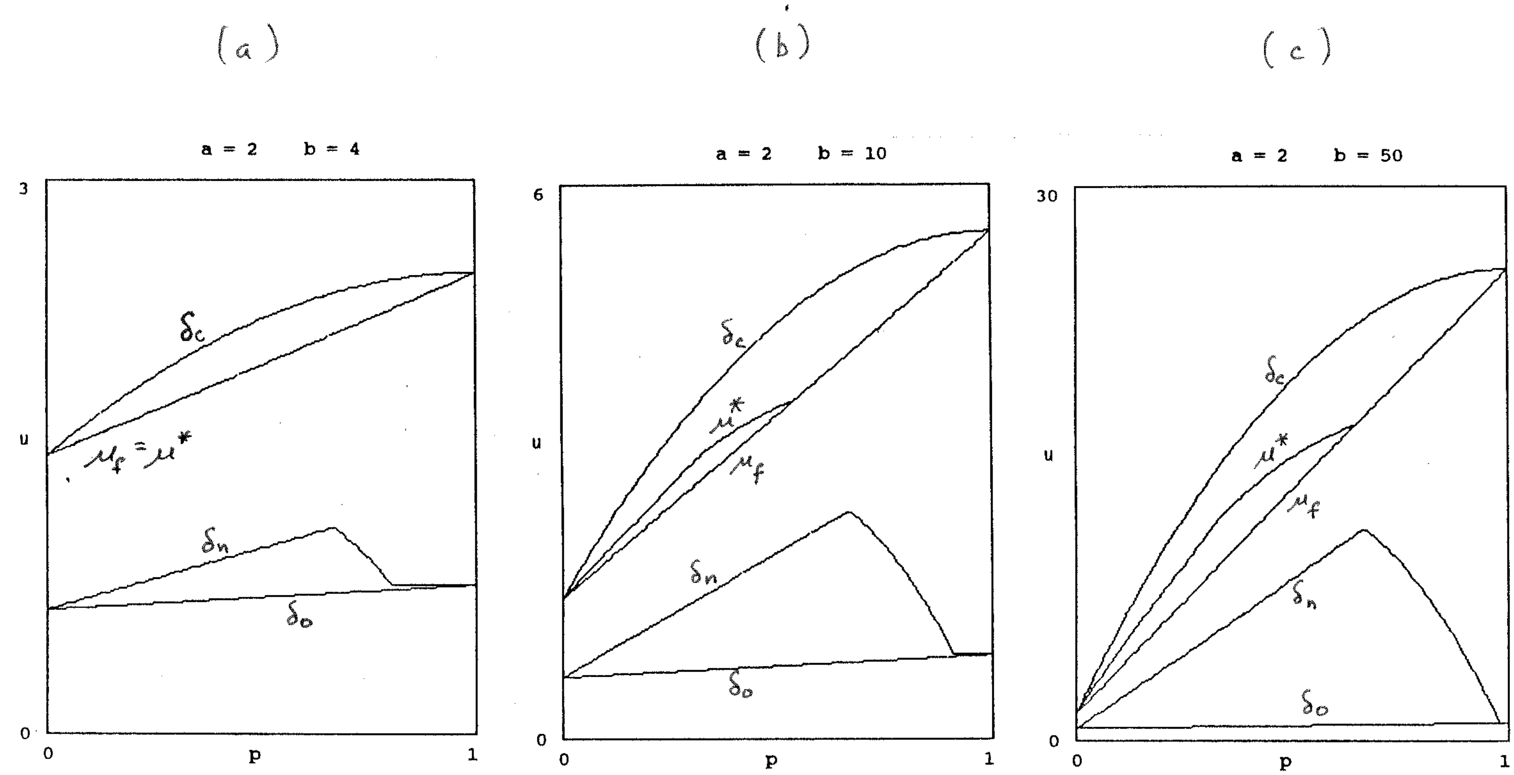
μ^* : optimal symmetric mechanism

μ_f : flat mechanism

δ_n : rule generated by symmetric Bayesian equilib. under incomplete information

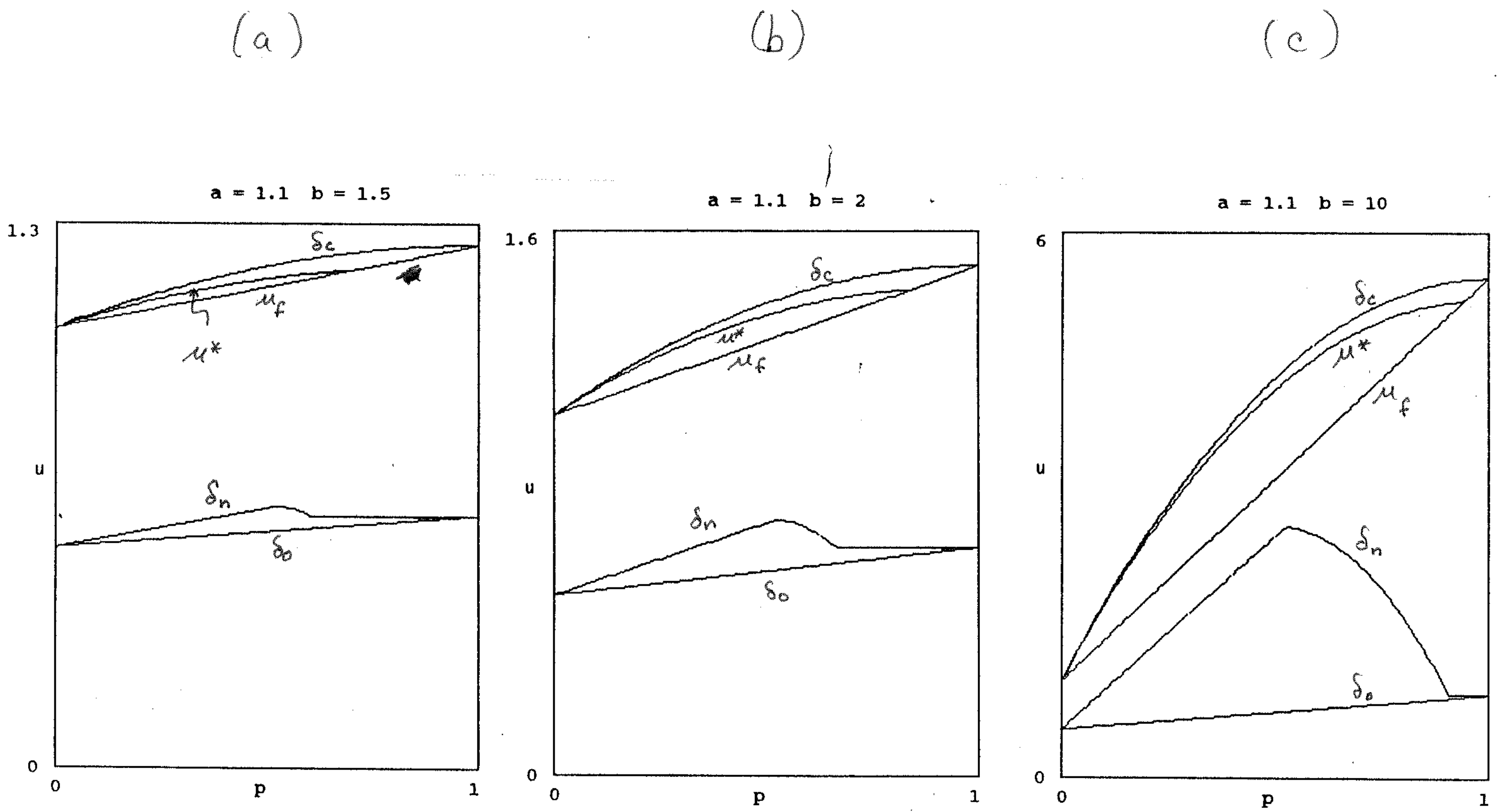
δ_o : rule generated by mixed-strategy Nash equilib under complete info. (ex ante).

Figure 2. Payoffs from Various Decision Rules when $a = 2$.



Key: as in Figure 1

Figure 3. Payoffs from Various Decision Rules when $a = 1.1$.



Key: as in Figure 1.