Discussion Paper No. 591

ACCEPTABLE AND PREDOMINANT CORRELATED EQUILIBRIA

by

Roger B. Myerson

April 1984
revised February 1985

J. L. Kellogg Graduate School of Management
Northwestern University
Evanston, Illinois 60201

Abstract. For stategic-form games with communication, acceptable correlated equilibria are defined as correlated equilibria that are stable when every player has an infinitesimal probability of trembling to any of his feasible actions. A set of acceptable actions is defined for each player, and it is shown that a correlated equilibrium is acceptable if and only if all unacceptable actions have zero probability. The unacceptable actions can be found by computing certain vectors called codomination systems, which extend the concept of dominated actions. Predominant correlated equilibria are defined by iterative elimination of unacceptable actions and are shown to exist.

Acknowledgements. The author is deeply indebted to Sergiu Hart for many discussions and suggestions, without which this paper could not have been written. This research was supported by fellowships from the John Simon Guggenheim Memorial Foundation and the Alfred P. Sloan Foundation.

ACCEPTABLE AND PREDOMINANT CORRELATED EQUILIBRIA

by

Roger B. Myerson

1. Introduction

As defined by Nash [1951], an equilibrium of a game is plan of behavior for every player in the game such that each player is willing to behave as planned if he expects all other players to do so. In an equilibrium, there may be some events that have probability zero, as long as all players behave as planned. The concept of equilibrium allows players to completely ignore all the outcomes of the game in such zero-probability events. In some equilibria, however, a player's willingness to obey the equilibrium plan would disappear if he gave any consideration at all to these zero-probability events. In such equilibria, the rationality of players' behavior often seems questionable, and so game theorists have searched for refinements of the equilibrium concept which exclude these equilibria.

For example, consider a two-player game in which player 1 can choose x_1 or y_1, player 2 can choose x_2 or y_2, and their payoffs (u_1,u_2) depend on their combination of choices as follows:

	x_2	y_2
x_1	1,5	1,5
y_1	0,0	2,1

There are two equilibria of this game: (x_1,x_2) and (y_1,y_2). However, the (x_1,x_2) equilibrium is <u>imperfect</u>, because player 2 would not be willing to choose x_2 if there were any chance that player 1 might choose y_1 (perhaps as a result of some tremble or mistake). Or to look at it another way, the action x_2 is (weakly) <u>dominated</u> for player 2 because, no matter what player 1 may do,

y_2 is always at least as good as x_2 for player 2.

These two concepts, perfection and domination, exclude the same equilibrium (x_1,x_2) in this example and may seem logically very similar, but in fact they do represent two different strands of the literature on noncooperative games. Concepts of perfect equilibria were introduced by Selten [1975], and related concepts have been developed by Myerson [1978], Kreps and Wilson [1982], Kalai and Samet [1984], van Damme [1983], and Kohlberg and Mertens [1982]. In each of these papers, the basic idea is that a reasonable equilibrium ought to be stable in some sense when small probabilities of players' mistakes (or some small perturbations in the payoffs) are introduced into the game. The concept of eliminating dominated actions was developed by Luce and Raiffa [1957], and two related concepts of _inferior_ actions and _rationalizable_ actions have been proposed by Harsanyi [1975] and by Bernheim [1984] and Pearce [1984]. In these domination concepts, the basic idea is to identify some actions that would be intrinsically unreasonable (in some sense) for a player to choose, and then to consider only equilibria that do not use these actions.

Unfortunately, there have been few logical connections between these two strands of noncooperative game theory. (See McLennan [1983] for a recent paper that does make such a connection.) It has been recognized that, for two-player games only, an equilibrium is perfect if and only if it does not use any weakly dominated action. But in games with three or more players, an equilibrium may be imperfect even though there are no dominated actions. Conversely, iterative elimination of dominated actions may eliminate equilibria that are not excluded by any perfectness concept.

In this paper, we show that these two strands can be unified for games with communication, where the fundamental solution concept is _correlated_ equilibrium, as defined by Aumann [1974]. In Section 2, we define acceptable correlated equilibria, which are stable against small probabilities of players' mistakes. In Section 3, we define unacceptable actions in a way which includes all weakly dominated actions, and we show that a correlated equilibrium is acceptable if and only if it does not use any unacceptable actions. Thus, for games with communication, a natural analogue of Selten's concept of perfectness is equivalent to a generalizated concept of elimination of dominated actions. In Section 4, we consider the process of iteratively eliminating unacceptable actions, to develop a concept of predominance, which is analogous to the concept of iterative or wide domination discussed by Luce and Raiffa [1957].

Given a game in strategic form, we say that it is a game with communication if the players can communicate before each player chooses his action. Following Aumann [1974], we allow that the players may communicate either directly or through a mediator, and the communication may be either deterministic or influenced by some random variable with an objective probability distribution. However, after the communication is over, each player still controls his own action or strategy separately; that is, we are assuming that jointly binding commitments are not allowed. (Thus, as Aumann has emphasized, a game with communication is different from a cooperative game, in which such commitments are allowed.) A correlated equilibrium of the given game is any probability distribution over the possible outcomes that could be implemented by a Nash equilibrium of any such extended game with preplay communication.

As a special case of the revelation principle (see Myerson [1982, 1983]), it can be shown that there is no loss of generality in considering only communication systems of the following form: a mediator randomly selects an outcome to recommend; then the mediator confidentially tells each player only the action that is recommended for him; and the probability distribution that generates the recommendations should be designed so that it is an equilibrium for each player to plan to obey the mediator's recommendations.

It is straightforward to show that the set of correlated equilibria is a convex set and includes all Nash equlibria. However, there may exist correlated equlibria that are not convex combinations of Nash equilibria, as Aumann [1974] has shown. For example, in the following game

	x_2	y_2
x_1	5,1	0,0
y_1	4,4	1,5

there is a correlated equilibrium in which each of the outcomes (x_1,x_2), (y_1,y_2), and (y_1,x_2) gets probability 1/3. (This correlated equilibrium could not be implemented if the players could only communicate with noiseless face-to-face communication. A mediator or noisy channel is needed, as Farrell [1983] has discussed.)

It is worth remarking that correlated equilibrium is in many ways a mathematically simpler concept than Nash equilibrium.1/ For many games, it may be easier to compute the set of all correlated equilibria, which is convex, than the set of all Nash equilibria, which may not even be a connected set. Thus, it is not surprising that refinements of the equilibrium concept also become simpler in games with communication.

1/In a universe with an omnipresent deity, no observer could ever be sure that players in a game were not getting correlated guidance, through the medium of silent prayer. Thus, game theorists would have to use correlated equilibria rather than Nash equilibria in all analysis.

The analysis of this paper is limited to games in strategic form--that is, games with a given structure consisting of: a set of players, a set of possible strategies or actions for each player, and a payoff function for each player. There is no specification of any dynamic structure to the play in a strategic-form game, so we generally assume that all players choose their actions simultaneously. The advantage to studying the strategic form is that it is a structurally simple and yet very general model. Furthermore, game theorists have long argued that there is no loss of generality in studying the strategic form, since any dynamic extensive form game can be normalized to an equivalent game in strategic form. (See Kuhn [1953], Luce and Raiffa [1957], and Mertens and Kohlberg [1983], for example. Selten [1975] suggested an alternative concept of agent-normalization.) Certainly the set of Nash equilibria of an extensive-form game and its normal form are identical.[2/]

However, there are limitations to studying refinements of the equilibrium concept in strategic-form games only. First, the problem of imperfect equilibria may appear to be a "knife-edge" phenomenon in strategic form. In fact, van Damme [1981] has shown that there is an open dense set of games in strategic form that have no imperfect equilibria. For the first example in this paper, the imperfect equilibrium at (x_1,x_2) exists only because player 2 gets exactly the same payoff from (x_1,x_2) as from (x_1,y_2). If the payoff of 5 in the upper right were perturbed slightly, all else equal, then (x_1,x_2) would either become a perfect equilibrium or cease to be an equilibrium at all. However, the equality of payoffs in (x_1,x_2) and (x_1,y_2) is not a mere

---

[2/]In this paper, the terms "strategic form" and "normal form" are not used synonymously. The strategic form is a general mathematical structure for characterizing games, formally described in Section 2. The normal form of an extensive-form game is a specific game in the strategic form that is the normalized equivalent of the extensive-game form.

coincidence if this game arises as the normal form of the dynamic extensive-form game shown in Figure 1. Thus, we should recognize that the existence of imperfect equilibria may be a result of the underlying dynamic structure of the game, even though we ignore this structure when we study the normal form.

[Insert Figure 1 here.]

A more serious issue has been raised in another paper by this author (Myerson [1985]). For games with communication, the normal form is not an adequate representation of the dynamic extensive form, because the possibilities for communication during the game are suppressed when we normalize the game. That is, unless we assume that players can communicate only before the play begins, two extensive-form games that have the same normal form may have different sets of communication equilibria. (See Myerson [1985] for examples.)

In Myerson [1985], a concept of <u>sequential communication equilibrium</u> is defined for dynamic multistage games with communication. The relationship between this paper and Myerson [1985] can be most simply described by stating that acceptable correlated equilibria (as defined in this paper) are related to the sequential communication equilibria of Myerson [1985] as Selten's [1975] trembling-hand perfect equilibria are related to Kreps and Wilson's [1982] sequential equilibria. Kreps and Wilson define sequential equilibria for dynamic games without communication by requiring that players' beliefs and actions should be rational in all possible events, including events that have zero probability. (For a Nash equilibrium, rationality would be required only

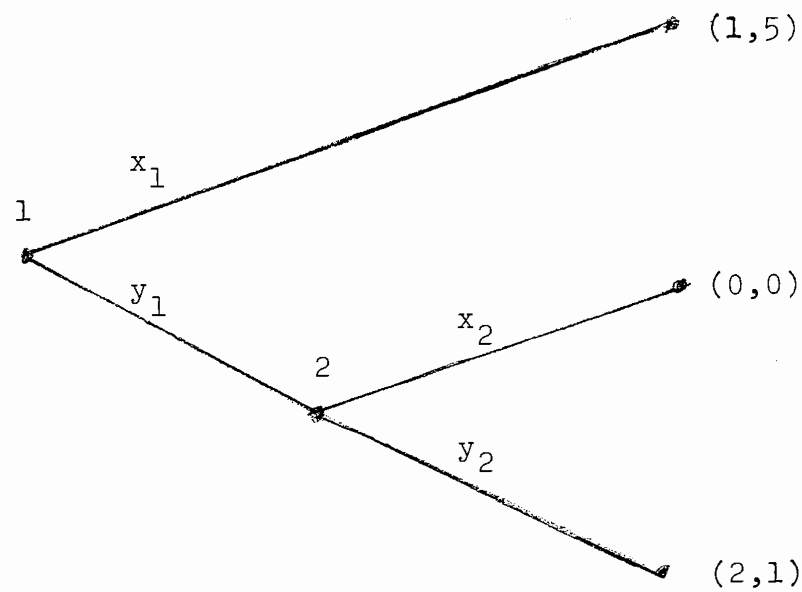

FIGURE 1.

in events that have positive probability.) Myerson [1985] defines sequential communication equilibrium by imposing a similar rationality requirement in dynamic games with communication. In both of these "sequential" concepts, the players' beliefs in zero-probability events are generated by taking the limit as small probabilities of players' trembles or mistakes go to zero. In static strategic-form games there are no problems about players' rational beliefs after zero-probability events, because all players are assumed to make all decisions simultaneously. However, one may still ask whether an equilibrium or correlated equilibrium is stable against the introduction of the small probabilities of trembles that would generate rational beliefs in a corresponding dynamic game. Selten's perfect equilibria and our acceptable correlated equilibria are respectively the equilibria and correlated equilibria that are stable in this sense.

2. Basic definitions

Let us consider a game Γ in strategic form, with the following structure

(2.1) $$\Gamma = (N, (C_i)_{i\in N}, (u_i)_{i\in N}),$$

where N is a nonempty finite set and, for each i in N, C_i is a nonempty finite set and u_i is a function from C into the real numbers $\mathbb{R}$, where

$$C = \times_{j\in N} C_j.$$

Here N denotes the set of players in the game, C_i denotes the set of actions or strategies available to player i in the game, and $u_i(c)$ denotes the payoff (measured in some von Neumann-Morgenstern utility scale) that player i would get if $c = (c_j)_{j\in N}$ were the combination of actions chosen by the players.

For any finite set X, we let $\Delta(X)$ denote the set of all probability distributions on the set X,

$$\Delta(X) = \{\mu : X \to [0,1] \mid \sum_{x \in X} \mu(x) = 1\}.$$

A <u>correlated strategy</u> for the players in Γ is any probability distribution μ in $\Delta(C)$. We may think of such a correlated strategy as being implemented by a mediator as follows. First, the mediator randomly selects some combination of actions in C, with $\mu(c)$ being the probability that he selects $c = (c_j)_{j \in N}$. Then the mediator separately and confidentially recommends to each player i that he should use the action c_i that is the i^{th} component of the selected vector c.

We do not assume that the mediator can compel the players to obey his recommendations, so he must use a correlated strategy that gives no player any incentive to disobey. Following Aumann [1974], we say that a correlated strategy μ is a <u>correlated equilibrium</u> iff it satisfies the following incentive constraints:

$$(2.2) \qquad \sum_{c_{-i} \in C_{-i}} \mu(c)(u_i(c) - u_i(c_{-i}, e_i)) \geq 0, \quad \forall i \in N, \ \forall c_i \in C_i, \ \forall e_i \in C_i.$$

(We use here the following notation:

$$C_{-i} = \times_{j \in N-i} C_j, \quad \text{where} \quad N-i = \{j \in N \mid j \neq i\};$$

(c_{-i}, e_i) is the vector in C with i^{th} component e_i and all other components the same as in c_{-i}; and $c = (c_{-i}, c_i)$.) Constraint (2.2) asserts that player i would not get higher expected utility from using e_i than from using c_i when the mediator has recommended action c_i. Thus, obedience of the mediator's recommendations by all players is a Nash equilibrium if and only if their correlated strategy is a correlated equilibrium.

In some correlated equilibria, a player may be willing to obey his recommendations only if he is absolutely sure that all players will also be obedient. It might be that player i would be unwilling to use action c_i if he believed that there was any chance of player j disobediently using e_j. A correlated equilibrium that required an absolute certainty of obedience would be less stable, in some sense. Thus, to develop a refined solution concept in which such unstable correlated equilibria are ruled out, we now consider a model of mediation with mistakes, similar to the perturbed games of Selten [1975]. In this model, we suppose that, for any player j and action e_j in C_j, there is always a small positive probability that player i might "tremble" and use d_i by mistake, no matter what the mediator and the other players do. Thus, every other player i must always take into account the possibility of player j using d_j.

First we need some more notation. For any $S \subseteq N$, if $S \neq \emptyset$ we let

$$C_S = \times_{j \in S} C_j,$$

(so $C_N = C$) and we let

$$C_\emptyset = \{\emptyset\}.$$

For any c in C and any e_S in C_S, we let (c_{-S}, e_S) be the combination of actions in which the j^{th} player's action is e_j if $j \in S$, and is c_j if $j \notin S$.

For any $\varepsilon > 0$, η is an ε-correlated strategy iff η is a probability distribution in $\Delta(C \times (\bigcup_{S \subseteq N} C_S))$ such that:

$$(2.3) \qquad \varepsilon\ \eta(c,e_S) \geq (1-\varepsilon) \sum_{e_i \in C_i} \eta(c,e_{S\cup\{i\}})$$

$$\forall c \in C,\ \forall i \in N,\ \forall S \subseteq N-i,\ \forall e_S \in C_S;$$

and

$$(2.4) \qquad \text{if} \quad \eta(c,e_S) > 0 \quad \text{then} \quad \eta(c,e_{S\cup\{i\}}) > 0,$$

$$\forall c \in C,\ \forall i \in N,\ \forall S \subseteq N-i,\ \forall e_{S\cup\{i\}} \in C_{S\cup\{i\}}.$$

(Here, the i^{th} component of $e_{S\cup\{i\}}$ is e_i, and all other components of $e_{S\cup\{i\}}$ form the vector e_S.) We interpret $\eta(c,e_S)$ as the probability that the mediator will recommended c, and all players not in the set S will choose their actions rationally, but the players in S will tremble and accidentally use the actions in the vector e_S. Thus, $\eta(c,\emptyset)$ is the probability that the mediator will recommend c and all players will choose rationally. Condition (2.4) asserts that every possible tremble for player i must have strictly positive conditional probability, given any recommendation vector and any vector of trembles among other players. Condition (2.3) asserts that, given any vector of recommendations selected by the mediator and any vector of trembles among players other than i, the conditional probability of player i also trembling is not greater than ε. Thus, in any limit of ε-correlated strategies, as ε goes to zero, the conditional probability of any player trembling would be always equal to zero, independently of any given information about the other players and the mediator. Notice also that if η is an ε-correlated strategy and $\hat{\varepsilon} > \varepsilon$, then η is also an $\hat{\varepsilon}$-correlated strategy.

An <u>ε-correlated equilibrium</u> is defined to be an ε-correlated strategy η

that satisfies the following incentive constraints

$$(2.5)\qquad \sum_{c_{-i}\in C_{-i}} \sum_{S\subseteq N-i} \sum_{e_S\in C_S} \eta(c,e_S)\left(u_i(c_{-S},e_S)-u_i(c_{-(S\cup\{i\})},e_{S\cup\{i\}})\right) \geq 0,$$

$$\forall i\in N,\ \forall c_i\in C_i,\ \forall e_i\in C_i.$$

That is, player i should not expect to gain by using action e_i when he is told to use c_i and he is not "trembling" or out of control.

We say that μ is an <u>acceptable correlated equilibrium</u> iff, for every strictly positive ε, there exists some ε-correlated equilibrium η^ε such that

$$(2.6)\qquad \lim_{\varepsilon\to 0} \eta^\varepsilon(c,\emptyset) = \mu(c), \quad \forall c\in C.$$

That is, an acceptable correlated equilibrium is any limit of ε-correlated equilibria as ε goes to zero. We may think of an acceptable correlated equilibrium as a correlated equilibrium in which obedient behavior by every player could still be rational when each player always has a positive but infinitesimal probability of trembling.

Our first result is that acceptable correlated equilibria are, in fact, correlated equilibria, and do exist. (All proofs are in Section 5.)

<u>Theorem 1</u>. The set of acceptable correlated equilibria of the game Γ is a nonempty subset of the set of correlated equilibria. Furthermore, any perfect equilibrium in the sense of Selten [1975] is an acceptable correlated equilibrium.

3. Acceptable actions

We say that c_i is an acceptable action for player i iff, for every $\varepsilon > 0$, there exists some ε-correlated equilibrium η such that

$$\sum_{c_{-i} \in C_{-i}} \eta(c,\emptyset) > 0.$$

That is, c_i is acceptable iff it can be rationally used by player i when the probabilities of trembling are arbitrarily small. We let E_i denote the set of acceptable actions for player i, and we let

$$E = \times_{i \in N} E_i.$$

The following useful lemma is proven is Section 5.

Lemma 1. A combination of actions c is in E if and only if, for every $\varepsilon > 0$, there exists some ε-correlated equilibrium η such that $\eta(c,\emptyset) > 0$.

We can now state the main result of this paper.

Theorem 2. μ is an acceptable correlated equilibrium if and only if μ is a correlated equilibrium and

$$\sum_{c \in E} \mu(c) = 1. \tag{3.1}$$

Proof. The proof is deferred to Section 5.

This result offers a much simpler characterization of the set of acceptable equilibria. Once the set E is known, the set of acceptable correlated equilibria is just the set of μ in $\Delta(C)$ that satisfy (3.1) and the incentive constraints (2.2). Notice that these are a finite collection of linear constraints on μ, so the set of acceptable correlated equilibria is a convex compact polyhedron (as is the set of correlated equilibria).

For this characterization to be useful, we need a practical way to verify whether an action is acceptable or not. The definition of acceptability provides a direct way to show that that an action is acceptable. We now need to develop a criterion that can be used to show when an action is not acceptable.

We let

$$A = \underset{i \in N}{\times} \mathbb{R}_+^{C_i \times C_i}.$$

That is, if $\alpha \in A$ then, for each i in N and each c_i in C_i, $\alpha_i(e_i|c_i)$ is a nonnegative number, which one may interpret as a shadow price for the incentive constraint (2.2) that player i should not expect to gain by using e_i instead of c_i when c_i is recommended to him.

For any c in C, any α in A, and any S that is a subset of N, we define

$$V_S(c,\alpha) = \sum_{i \in S} \sum_{e_i \in C_i} \alpha_i(e_i|c_i)\big(u_i(c) - u_i(c_{-i},e_i)\big). \tag{3.3}$$

$V_S(c,\alpha)$ may be called the <u>aggregate incentive value</u> of c for S, with respect to α. Notice that it is a weighted sum of the contributions of c to the incentive constraints for the members of S.

When the mediator recommends c and the players in S tremble to e_S, the aggregate incentive value for the nontrembling players, with respect to α, is $V_{N\setminus S}((c_{-S},e_S),\alpha)$. For any c in C and α in A, we let $V_*(c,\alpha)$ denote the vector of all such aggregate incentive values for the nontrembling players, with all possible combinations of trembles from c (including $e_S = \emptyset$); that is,

$$V_*(c,\alpha) = \big(V_{N\setminus S}((c_{-S},e_S),\alpha)\big)_{S\subseteq N, e_S \in C_S}.$$

We may write $V_*(c,\alpha) = 0$ iff

$$V_{N\setminus S}((c_{-S},e_S),\alpha) = 0, \quad \forall S \subseteq N, \quad \forall e_S \in C_S.$$

We may write $V_*(c,\alpha) <_\ell 0$ iff

$$\exists S \subseteq N, \ \exists e_S \in C_S \text{ such that } V_{N\backslash S}((c_{-S},e_S),\alpha) \neq 0,$$

and, for each $T \subseteq N$ and each d_T in C_T,

$$\text{if } V_{N\backslash T}((c_{-T},d_T),\alpha) > 0 \text{ then } \exists Q \subset T \text{ such that } V_{N\backslash Q}((c_{-Q},d_Q),\alpha) < 0.$$

(In this definition, d_Q denotes the subvector of d_T consisting only of those components indexed on the members of Q. Read $<_\ell$ as "is lexicographically less than.") Thus, if $V_*(c,\alpha) = 0$ then the aggregate incentive value (with respect to α) for the nontrembling players is always zero when the mediator recommends c. If $V_*(c,\alpha) <_\ell 0$ then, when c is recommended, for any set of trembles that leaves the nontrembling players with a positive aggregate incentive value, there must exist a smaller (and therefore much more likely) set of trembles that leaves the nontrembling players with negative aggregate incentive value; furthermore, there is some set of trembles that actually does leave the nontrembling players with negative aggregate incentive value.

For any positive integer K, we say that $(\alpha^1,\ldots,\alpha^K)$ is a (<u>weak)</u> <u>codomination system</u> iff, for every k in $\{1,\ldots K\}$, α^k is in A, and for every c in C, either

$$(3.4) \qquad V_*(c,\alpha^j) = 0 \quad \forall j \in \{1,\ldots K\},$$

or

$$(3.5) \qquad \exists m \in \{1,\ldots K\} \text{ such that } V_*(c,\alpha^m) <_\ell 0 \text{ and, } \forall j < m, \ V_*(c,\alpha^j) = 0.$$

(A related concept of <u>sequential codomination</u> is introduced elsewhere, by Myerson [1985]. The concept defined here may be called a <u>weak codomination</u>

system whenever it is necessary to distinguish it from the concept of sequential codomination.)

It can be shown that, if $(\alpha^1,\ldots,\alpha^K)$ is a codomination system, then, for sufficiently small ε and any ε-correlated strategy η, the expected aggregate incentive value for the nontrembling players,

$$(3.6) \qquad \sum_{c\in C} \sum_{S\subseteq N} \sum_{d_S\in C_S} \eta(c,d_S)\, V_{N\setminus S}((c_{-S},d_S),\alpha^k),$$

is either zero for all k, or else it is negative for the lowest k such that it is nonzero. Furthermore, (3.6) will be negative for some k whenever there is positive probability under η that the mediator will recommend some action c such that (3.5) holds. This is important because the expected aggregate incentive value for the nontrembling players (with respect to any α) must be nonnegative if η is an ε-correlated equilibrium, since the incentive constraints are all satisfied. Thus, if $(\alpha^1,\ldots,\alpha^K)$ is a codomination system and (3.5) holds then, for all sufficiently small ε, c cannot be used in any ε-correlated equilibrium. This observation motivates the following criterion for identifying unacceptable actions.

Theorem 3. A combination of actions c is in E if and only if, for every codomination system, $(\alpha^1,\ldots,\alpha^K)$,

$$V_*(c,\alpha^k) = 0, \quad \forall k \in \{1,\ldots,K\}.$$

Proof. See Section 5.

Thus, to show that $c \in E$, it suffices to show that, for every positive ε there exists some ε-correlated equilibrium under which the probability of recommending c is positive. Conversely, to show that $c \notin E$, it suffices to find some codomination system with some k such that $V_*(c,\alpha^k) <_\ell 0$. Once we have used these two tests to identify the set E, we can use Theorem 2 to identify the set of all acceptable correlated equilibria.

As the name suggests, codomination systems are logically related to the more familiar concept of domination of actions in strategic form games (see Luce and Raiffa [1957]). To see how, suppose that $\sigma = (\sigma(e_i))_{e_i \in C_i}$ is a randomized strategy for player i (so $\sigma \in \Delta(C_i)$) that weakly dominates some action c_i, in the sense that, for every c_{-i} in C_{-i},

$$u_i(c) \leq \sum_{e_i \in C_i} \sigma(e_i) u_i(c_{-i},e_i),$$

with strict inequality for at least one c_{-i}. Then we can construct a codomination system with $K = 1$ by letting

$$\alpha_i^1(e_i|c_i) = \sigma(e_i), \quad \forall e_i \in C_i,$$

and

$$\alpha_j^1(e_j|d_j) = 0 \text{ if } j \neq i \text{ or } d_j \neq c_i.$$

Then it is straightforward to check that $V_*(c,\alpha^1) <_\ell 0$ for every c_{-i}, and $V_*(d,\alpha^1) = 0$ for every d such that $d_i \neq c_i$. Thus any weakly dominated action is unacceptable or codominated.

However, there also exist games in which there are unacceptable actions and unacceptable equilibria but there are not any weakly dominated actions. For example, consider the three-person game with $C_1 = \{x_1,y_1\}$, $C_2 = \{x_2,y_2,z_2\}$, $C_3 = \{x_3,y_3,z_3\}$, and utility functions (u_1,u_2,u_3) as follows:

x_1	x_3	y_3	z_3
x_2	2,1,1	0,2,0	0,2,0
y_2	0,0,2	0,3,0	0,0,3
z_2	0,0,2	0,0,3	0,3,0

y_1	x_3	y_3	z_3
x_2	1,3,3	1,3,3	1,3,3
y_2	1,3,3	1,3,3	1,3,3
z_2	1,3,3	1,3,3	1,3,3

(This example is derived from a similar game suggested by J. Farrell.) Each action of each player is the unique best response to some combination of other players' actions, so there are indeed no dominated actions. However, y_2, z_2, y_3, z_3, and y_1 are all unacceptable actions. To prove this unacceptability, consider the codomination system (α^1,α^2) such that

$$\alpha_i^1(x_i|y_i) = \alpha_i^1(x_i|z_i) = 1, \qquad \forall i \in \{2,3\},$$

$$\alpha_1^2(x_1|y_1) = 1,$$

and all other $\alpha_i^k(e_i|c_i)$ are equal to zero. So $V_*(c,\alpha^1) <_\ell 0$ if $c_2 \in \{y_2,z_2\}$ or $c_3 \in \{y_3,z_3\}$, and $V_*(c_1,x_2,x_3) = 0$ $\forall c_1$, which proves that y_2, z_2, y_3, and z_3 are unacceptable. Then y_1 is unacceptable because $V_*((y_1,x_2,x_3),\alpha^2) <_\ell 0$ and $V_*((x_1,x_2,x_3),\alpha^2) = 0$.

Thus, $E = \{(x_1,x_2,x_3)\}$, and so the unique acceptable correlated equilibrium is the Nash equilibrium at (x_1,x_2,x_3). There are other correlated equilibria which are unacceptable, including the Nash equilibrium at (y_1,y_2,y_3).

---

4. Predominant actions and equilibria.

The theory of acceptable correlated equilibria has led us to the conclusion that players should be unwilling to ever use their unacceptable actions. This conclusion suggests that elimination of all the unacceptable actions from a game should not change the outcome. Thus, given any strategic-form game Γ as above, we define the acceptable residue of Γ, denoted $R(\Gamma)$, to be the game that differs from Γ only in that each player's set of feasible actions in $R(\Gamma)$ is his set of acceptable actions in Γ. That is, the acceptable residue of Γ is

$$R(\Gamma) = (N, (E_i)_{i\in N}, (u_i)_{i\in N}).$$

where E_i is as defined in Section 3. A correlated strategy μ in $\Delta(E)$ is a correlated equilibrium of the game $R(\Gamma)$ iff it satisfies the following incentive constraints:

$$(4.1) \qquad \sum_{e_{-i}\in E_{-i}} \mu(e)\big(u_i(e) - u_i(e_{-i},d_i)\big) \geq 0, \quad \forall i\in N, \quad \forall e_i\in E_i, \quad \forall d_i\in E_i.$$

Any correlated strategy for the game $R(\Gamma)$ can be considered to be a correlated strategy for Γ as well, by simply assigning zero probabilities to all combinations of actions that are in C but not in E. (That is, we may think of $\Delta(E)$ as a subset of $\Delta(C)$.) To verify that a correlated equilibrium of $R(\Gamma)$ is also a correlated equilibrium of Γ, it suffices to check that the above incentive constraints also hold for all d_i in C_i, not just for d_i in E_i as (4.1) requires. That is, we must check that no player could gain by disobediently choosing an unacceptable action. In fact, this is always true.

<u>Theorem 4.</u> If μ is a correlated equilibrium of $R(\Gamma)$ (that is, μ satisfies (3.1) and (4.1)) then μ is a correlated equilibrium of Γ.

<u>Proof.</u> The proof is given in Section 5.

Theorems 2 and 4 together imply that the set of acceptable correlated equilibria of Γ is exactly equal to the set of correlated equilibria of $R(\Gamma)$. However, the set of acceptable correlated equilibria of $R(\Gamma)$ may be smaller than the set of acceptable correlated equilibria of Γ, because ε-correlated strategies in the game $R(\Gamma)$ will not assign positive probability to any combinations of actions that are outside of E, since such combinations of actions are simply not part of the game $R(\Gamma)$. Thus the set of acceptable actions for player i in $R(\Gamma)$ may be a proper subset of E_i.

In this way, we can analyze a sequence of successively smaller games by iteratively eliminating all unacceptable actions from the game that remains. Formally, for every player i and every positive integer m, we define $R^m(\Gamma)$ and E_i^m by induction, as follows. To begin, let $R^1(\Gamma) = R(\Gamma)$ and let $E_i^1 = E_i$ for all i. Then, for every $m > 1$, let E_i^m be the set of all acceptable actions for player i in the game $R^{m-1}(\Gamma)$, and let

$$R^m(\Gamma) = (N, (E_i^m)_{i\in N}, (u_i)_{i\in N}).$$

That is, each $R^m(\Gamma)$ is the acceptable residue of $R^{m-1}(\Gamma)$. Thus we may call $R^m(\Gamma)$ the <u>m-iterative residue</u> of Γ. Clearly,

$$C_i \supseteq E_i^1 \supseteq E_i^2 \supseteq E_i^3 \supseteq \dots, \quad \forall i\in N.$$

Since there are only finitely many actions in the original game Γ, there must exist some number M and some sets of actions E_i^* such that

$$E_i^M = E_i^{M+1} = E_i^{M+2} = \ldots = E_i^*, \quad \forall i \in N.$$

That is, E_i^* is the set of all actions that are acceptable for player i in all iterative residues of Γ. Let

$$E^* = \underset{i \in N}{\times} E_i^*.$$

We say that c_i is a (weakly) predominant action for player i in Γ iff $c_i \in E_i^*$. Similarly, we say that μ is a (weakly) predominant correlated equilibrium of Γ iff μ is a correlated equilibrium of Γ and

$$(4.2) \qquad \sum_{e \in E^*} \mu(e) = 1.$$

Thus, a correlated equilibrium of Γ predominant iff it is acceptable in all iterative residues of Γ. (A related concept of sequential predominance is defined by Myerson [1985]. The concept of predominance introduced here may be called weak predominance whenever it is necessary to distinguish it from sequential predominance.)

Because the set of correlated equilibria of the acceptable residue is a nonempty subset of the set of correlated equilibria of the original game, it follows inductively that the set of correlated equilibria of every iterative residue is a nonempty subset of the correlated equilibria of the original game. Of course, every iterative residue has at least one Nash (uncorrelated) equilibrium among its correlated equilibria, by Nash's [1951] general existence theorem. But the set of predominant correlated equilibria of Γ is just the set of all correlated equilibria of an M-iterative residue of Γ, for some sufficiently large M. Thus, we have derived the following general existence theorem for predominant equilibria.

Theorem 5. The set of predominant correlated equilibria of Γ is nonempty and includes at least one Nash equilibrium.

For a simple example, consider the following two-person game.

	w_2	x_2	y_2	z_2
w_1	2,2	1,1	0,0	0,0
x_1	1,1	1,1	2,0	2,0
y_1	0,0	0,2	3,0	0,3
z_1	0,0	0,2	0,3	3,0

There are no dominated actions in this game, but the only predominant equilibrium is at (w_1,w_2). The equilibrium at (x_1,x_2) is acceptable but is not predominant. It is straightforward to check that, for each player i, $E_i = E_i^1 = \{w_i,x_i\}$ and $E_i^* = E_i^2 = \{w_i\}$ in this example.

5. Proofs

Proof of Theorem 1. To check that an acceptable correlated equilibrium μ is a correlated equilibrium, observe that condition (2.5) for $\eta = \eta^\varepsilon$ converges to condition (2.2) for μ as ε goes to zero, by (2.6). (Notice that (2.3) implies that $\lim_{\varepsilon\to 0} \eta^\varepsilon(c,e_S) = 0$ for every c and e_S such that $S \neq \emptyset$.)

Now, suppose that σ in $\underset{i\in N}{\times} \Delta(C_i)$ is a perfect equilibrium. Then there exist sequences $\{\varepsilon^k\}_{k=1}^\infty$, $\{\sigma^k\}_{k=1}^\infty$, and $\{q^k\}_{k=1}^\infty$ such that

$$0 < \varepsilon^k < 1, \quad \sigma^k \in \underset{i \in N}{\times} \Delta(C_i), \quad q^k \in \underset{i \in N}{\times} \Delta(C_i), \quad \forall k;$$

$$q_i^k(c_i) > 0, \quad \forall k, \quad \forall i \in N, \quad \forall c_i \in C_i;$$

$$\lim_{k\to\infty} \varepsilon^k = 0, \quad \lim_{k\to\infty} \sigma^k = \sigma;$$

and

$$0 \leqslant \sum_{c_{-i}\in C_{-i}} \Big(\prod_{j\in N-i} \big((1-\varepsilon^k)\sigma_j^k(c_j) + \varepsilon^k q_j(c_j)\big)\Big)\sigma_i^k(c_i)\big(u_i(c) - u_i(c_{-i},e_i)\big)$$

$$\forall k, \ \forall i \in N, \ \forall c_i \in C_i, \ \forall e_i \in C_i.$$

That is, σ_i^k is a best response for player i when every other player j is independently using either his σ_j^k strategy, with probability $(1-\varepsilon^k)$, or his q_j^k "trembling-hand" strategy, with probability ε^k. Now, for any $\varepsilon > 0$, let ε^k be the largest number in the sequence such that $\varepsilon^k < \varepsilon$, and let

$$\eta^\varepsilon(c,e_S) = \Big(\prod_{i\in N} \sigma_i^k(c_i)\Big)\Big(\prod_{j\in S} \varepsilon^k q_j^k(e_j)\Big)(1-\varepsilon^k)^{|N\backslash S|}$$

$$\forall c \in C, \ \forall S \subseteq N, \ \forall e_S \in C_S,$$

and let

$$\mu(c) = \prod_{i\in N} \sigma_i(c_i), \quad \forall c \in C.$$

Then η^ε is the ε-correlated strategy that simulates the equilibrium σ^k to Selten's perturbed game with $\varepsilon^k q^k$ trembles, and μ is the correlated strategy that simulates the perfect equilibrium σ. It is straightforward to check that the conditions for perfect equilibrium imply that each η^ε is indeed an ε-correlated equilibrium, and that

$$\lim_{\varepsilon \to 0} \eta^{\varepsilon}(c,\emptyset) = \mu(c), \quad \forall c \in C,$$

so that μ is an acceptable correlated equilibrium.

It is well-known that any finite game in strategic form has at least one perfect equilibrium. (The proof is that each of Selten's perturbed games has at least one equilibrium, and $\times_{i \in N} \Delta(C_i)$ is compact, so a limit of perturbed game equilibria exists.) Thus, the set of acceptable correlated equilibria is nonempty. Q.E.D.

Proof of Lemma 1. If, for every $\varepsilon > 0$, there exists some ε-correlated equilibrium η such that $\eta(c) > 0$ then each $c_i \in E_i$ and so $c \in E = \times_{i \in N} E_i$.

Conversely, suppose that $c \in E$, and let ε be a positive number. We want to show that there exists some ε-correlated equilibrium η such that $\eta(c) > 0$.

By definition of E, for every player i there exists an ε-correlated equilibrium λ^i and an action vector b^i in C such that $b^i_i = c_i$ and $\lambda^i(b^i,\emptyset) > 0$. Let

$$\lambda = \frac{1}{|N|} \sum_{i \in N} \lambda^i.$$

Then λ is also an ε-correlated equilibrium, and $\lambda(b^i,\emptyset) > 0$ for every i.

Let ν be a correlated strategy with ε-trembles defined as follows:

$$\nu(d,e_S) = 0 \quad \text{if} \quad d \neq c,$$

$$\nu(c,e_S) = (1-\varepsilon)^{|N \setminus S|} (\varepsilon)^{|S|} / |C_S|.$$

That is, the mediator in ν always recommends c, each player independently trembles with probability ε, and every action for a trembling player is equally likely. For any d in C, let $\nu_i^*(d)$ be the probability under ν that the players use the actions in d and player i does not tremble; that is

$$\nu_i^*(d) = \sum_{S\subseteq N-i} \sum_{\{e_S|(c_{-S},e_S)=d\}} \nu(d,e_S).$$

Notice that $\nu_i^*(d) = 0$ if $d_i \neq c_i$, and $\nu_i^*(d) > 0$ if $d_i = c_i$.

Now choose δ so that

$$\delta = \min_i \min_{e_{N-i}} \lambda(b^i, e_{N-i}).$$

Notice that $\delta > 0$, because $\lambda(b^i,\emptyset) > 0$ and λ is an ε-correlated equilibrium. For any player i and any e in C such that $e_i = c_i$, let

$$\hat{\eta}(b^i,e_{N-i}) = \lambda(b^i,e_{N-i}) + (\delta/2)(\nu(b^i,e_{N-i}) - \nu_i^*(e)).$$

For any (d,e_S) not covered by the preceding sentence, define $\hat{\eta}(d,e_S)$ by

$$\hat{\eta}(d,e_S) = \lambda(d,e_S) + (\delta/2)\,\nu(d,e_S).$$

This $\hat{\eta}$ satisfies the conditions (2.3) and (2.4) that are required of an ε-correlated strategy, because λ and ν satisfy them. Condition (2.5) is also satisfied by $\hat{\eta}$ because, for any player i and any vector d in C, the probability that the players do d (after trembles) when player i does not tremble is the same in both $\hat{\eta}$ and λ; that is

$$\sum_{a\in C} \sum_{S\subseteq N-i} \sum_{\{e_S|(a_{-S},e_S)=d\}} (\lambda(a,e_S) - \hat{\eta}(a,e_S)) = 0.$$

Let $\eta = \hat{\eta}/\|\hat{\eta}\|$, where

$$\|\hat{\eta}\| = \sum_{d\in C} \sum_{S\subseteq N} \sum_{e_S\in C_S} \hat{\eta}(d,e_S).$$

Then η also satisfies conditions (2.3)-(2.5), and so η is an ε-correlated equilibrium. Furthermore, $\eta(c,\emptyset) > 0$ because $\nu(c,\emptyset) > 0$.

Q.E.D.

<u>Proof of Theorem 2.</u> By Lemma 1, for every e in E and every $\varepsilon > 0$, there exists some ε-correlated equilibrium $\nu^{e,\varepsilon}$ such that $\nu^{e,\varepsilon}(e,\emptyset) > 0$. Let

$$\lambda^{\varepsilon} = \frac{1}{|E|} \sum_{e \in E} \nu^{e,\varepsilon}.$$

Then λ^{ε} is an ε-correlated equilibrium and $\lambda^{\varepsilon}(e,\emptyset) > 0$ for every e in E.

Now, suppose that μ is a correlated equilibrium and $\sum_{e \in E} \mu(e) = 1$. Let η^{ε} be defined so that

$$\eta^{\varepsilon}(c,\emptyset) = (1-\varepsilon)\ \mu(c) + \varepsilon\ \lambda^{\varepsilon}(c,\emptyset), \qquad \forall c \in C,$$

and

$$\eta^{\varepsilon}(c,e_S) = \varepsilon\ \lambda^{\varepsilon}(c,e_S), \quad \forall c \in C,\ \forall S \neq \emptyset,\ \forall e_S \in C_S.$$

That is, η^{ε} differs from μ in that we have mixed in an ε probability of λ^{ε}. The incentive constraints (2.5) are satisfied by η^{ε} because they satisfied by λ^{ε} and because μ satisfies (2.2). The conditions (2.3) and (2.4) are also satisfied by η^{ε} and because they are satisfied by λ^{ε} and because $\lambda^{\varepsilon}(c,\emptyset) > 0$ for any c such that $\mu(c) > 0$ (since such c must be in E). Thus, η^{ε} is an ε-correlated equilibrium. Furthermore, (2.6) is satisfied, so μ is an acceptable correlated equilibrium.

Conversely, suppose that μ is an acceptable correlated equilibrium. If $\mu(e) > 0$ then for any $\varepsilon > 0$, there exists some ε-correlated equilibrium η^{ε} such that $\eta^{\varepsilon}(e,\emptyset) > 0$, by (2.6). So $e \in E$ if $\mu(e) > 0$, and therefore

$$\sum_{c \in E} \mu(c) = 1.$$

Q.E.D.

Before proving Theorem 3, we need to prove the following mathematical fact.

Lemma 2. Let H be a convex subset of $\mathbb{R}^L$, and let $\mathbb{R}^L_+$ denote the nonnegative orthant of $\mathbb{R}^L$. Then $H \cap \mathbb{R}^L_+ = \emptyset$ if and only if there exists some K, where $K \leq L$, and some finite sequence of vectors $(\alpha^1,\ldots,\alpha^K)$ such that each $\alpha^k \in \mathbb{R}^L_+$ and, for each θ in H, there is some k in $\{1,\ldots,K\}$ such that $\alpha^k \cdot \theta < 0$ and $\alpha^j \cdot \theta = 0$ for every j such that $j < k$.

Proof of Lemma 2. To prove the "if" part, notice that if $\theta \in H \cap \mathbb{R}^L_+ \neq \emptyset$ then $\alpha \cdot \theta \geq 0$ for every α in $\mathbb{R}^L_+$, so no such vectors $(\alpha^1,\ldots,\alpha^K)$ could exist.

To prove the "only if" part, suppose that $H \cap \mathbb{R}^L_+ = \emptyset$. By the separating hyperplane theorem (see Rockafeller [1970]), there exists a nonzero vector α^1 such that $\alpha^1 \cdot \theta \leq 0$ for every θ in H and $\alpha^1 \cdot \omega \geq 0$ for every ω in $\mathbb{R}^L_+$. Clearly, this $\alpha^1 \in \mathbb{R}^L_+$.

We now construct α^k inductively, for $k = 2,\ldots,K$. Let

$$\Omega^k = \{\omega \in \mathbb{R}^L \,|\, \alpha^j \cdot \omega = 0,\ \forall j \in \{1,\ldots,k-1\}\}.$$

Since $H \cap \Omega^k$ and $\mathbb{R}^L_+ \cap \Omega^k$ are disjoint convex sets in the finite-dimensional vector space Ω^k, there exists some nonzero vector β^k in Ω^k such that $\beta^k \cdot \theta \leq 0$ for every θ in $H \cap \Omega^k$, and $\beta^k \cdot \omega \geq 0$ for every ω in $\mathbb{R}^L_+ \cap \Omega^k$. This vector $\beta^k = (\beta^k_1,\ldots,\beta^k_L)$ may have some negative components, but only where $\alpha^1 + \ldots + \alpha^{k-1}$ is strictly positive. That is, if $\beta^k_j < 0$, and $\alpha^1_j = 0,\ldots,\alpha^{k-1}_j = 0$, then we could construct a vector ω in $\mathbb{R}^L_+ \cap \Omega^k$ such that $\beta^k \cdot \omega < 0$, by letting $\omega_j = 1$ and all other $\omega_i = 0$. Thus, for sufficiently large M, $\beta^k + M(\alpha^1 + \ldots + \alpha^{k-1}) \in \mathbb{R}^L_+$. Let

$$\alpha^k = \beta^k + M(\alpha^1 + \ldots + \alpha^{k-1}).$$

By definition of Ω^k, $\alpha^k \cdot \theta = \beta^k \cdot \theta \leq 0$ for every θ in $H \cap \Omega^k$. Notice also that α^k is linearly independent of $\{\alpha^1,\ldots,\alpha^{k-1}\}$, because β^k is not zero and is orthogonal to these vectors.

The above construction of $\alpha^1,\alpha^2,\ldots$ terminates when $H \cap \Omega^k = \emptyset$, in which case $K = k-1$ and $(\alpha^1,\ldots,\alpha^K)$ satisfy the conditions of Lemma 2. The construction must terminate for some $k \leq L+1$, because $k = L+1$ would imply that $\Omega^k = \{0\}$ (since $\{\alpha^1,\ldots,\alpha^L\}$ would form a basis for $\mathbb{R}^L$) and $0 \notin H$. Q.E.D.

Proof of Theorem 3. By Lemma 1, for any c in $C\backslash E$, there exists some ε such that there is no ε-correlated equilibrium η with $\eta(c,\emptyset) > 0$. Since C is a finite set, there exists a number $\varepsilon^* > 0$ such that, for every c in $C\backslash E$, there is no ε^*-correlated equilibrium η with $\eta(c,\emptyset) > 0$. That is, $c \in E$ if and only if there exists some ε^*-correlated equilibrium η such that $\eta(c,\emptyset) > 0$. Let

$$\theta(\eta) = \left(\theta_i(\eta,d_i|c_i)\right)_{d_i\in C_i,\, c_i\in C_i,\, i\in N}$$

where

$$\theta_i(\eta,d_i|c_i) = \sum_{c_{-i}} \sum_{S\subseteq N-i} \sum_{d_S} \eta(c,d_S)(u_i(c_{-S},d_S) - u_i(c_{-S\cup\{i\}},d_{S\cup\{i\}})).$$

Thus, an ε^*-correlated strategy is an ε^*-correlated equilibrium if and only if all components of $\theta(\eta)$ are nonnegative.

Let $\Phi(c) = \{\eta|\ \eta$ is an ε^*-correlated strategy and $\eta(c,\emptyset) > 0\}$.

Notice that $\{\theta(\eta)|\ \eta \in \Phi(c)\}$ is a convex subset of $\underset{i\in N}{\times} \mathbb{R}^{C_i\times C_i}$, and that A (as defined in Section 3) is the nonnegative orthant of this vector space. Thus, $c \notin E$ if and only if

$$\{\theta(\eta) \mid \eta \in \Phi(c)\} \cap A = \emptyset.$$

Then, by Lemma 2, $c \notin E$ if and only if there exists some sequence of vectors $(\alpha^1,\ldots,\alpha^K)$ such that each $\alpha^k \in A$ and, for every η in $\Delta(c)$, there exists some k in $\{1,\ldots,K\}$ such that

(5.1) $$\alpha^k\cdot\theta(\eta) < 0, \quad \text{and} \quad \alpha^j\cdot\theta(\eta) = 0 \quad \forall j < k,$$

where

(5.2) $$\alpha^k\cdot\theta(\eta) = \sum_i \sum_{d_i} \sum_{e_i} \alpha_i^k(e_i|d_i)\ \theta_i(\eta,e_i|d_i)$$
$$= \sum_d \sum_{S\subseteq N} \sum_{e_S} \eta(d,e_S)\ V_{N\backslash S}((d_{-S},e_S),\alpha^k).$$

Suppose now that $c \notin E$ and $(\alpha^1,\ldots,\alpha^K)$ satisfies (5.1) for all η in $\Phi(c)$. Without loss of generality, we can assume that there exists some $\bar{\eta}$ in $\Phi(c)$ such that (5.1) is satisfied for $k = K$ (otherwise, we could reduce K).

We show first that $(\alpha^1,\ldots,\alpha^K)$ is a codomination system. If it were not then there would exist some (b,a_R) and some k such that $V_*(b,\alpha^j) = 0$ for every j less than k, $V_{N\backslash R}((b_{-R},a_R),\alpha^k) > 0$, and $V_{N\backslash S}((b_{-S},a_S),\alpha^k) = 0$ for every $S \subset R$. Now define η^1 and η^2 as follows:

$$\hat{\eta}^1(b,e_S) = (1-\varepsilon^*)^{|N\backslash S|}(\varepsilon^*)^{|S|}/|C_S|, \quad \forall S\subseteq N, \quad \forall e_S\in C_S;$$

$$\hat{\eta}^1(d,e_S) = 0 \quad \text{if} \quad d \neq b;$$

$$\hat{\eta}^2(b,a_S) = (1-\varepsilon^*)^{|N\backslash S|}(\varepsilon^*)^{|S|}, \quad \forall S\subseteq R;$$

$$\hat{\eta}^2(d,e_S) = 0 \quad \text{if} \quad d \neq b \quad \text{or} \quad S \not\subseteq R \quad \text{or} \quad e_S \neq a_S;$$

and $$\eta^1 = \hat{\eta}^1/\|\hat{\eta}^1\|; \quad \eta^2 = \hat{\eta}^2/\|\hat{\eta}^2\|.$$

Then η^1 is an ε^*-correlated strategy, although it may not be in $\Phi(c)$ if $b \neq c$ (since $\eta^1(c,\emptyset) > 0$ would fail). η^2 satisfies (2.3) but does not satisfy the positivity condition (2.4) required of an ε^*-correlated strategy. However, for any number γ such that $0 < \gamma < .5$, when we let

$$\eta = (1 - 2\gamma)\eta^2 + \gamma\eta^1 + \gamma\bar{\eta}$$

then we have $\eta \in \Phi(c)$. Notice that

$$0 = \alpha^j \cdot \theta(\eta^1) = \alpha^j \cdot \theta(\eta^2) = \alpha^j \cdot \theta(\bar{\eta}), \quad \forall j < k,$$

and $\alpha^k \cdot \theta(\eta^2) > 0$. Thus, for sufficiently small positive γ, we get

$$\alpha^\ell \cdot \theta(\eta) = 0 \quad \forall \ell < k, \text{ and } \alpha^k \cdot \theta(\eta) > 0,$$

which contradicts (5.1). Thus, $(\alpha^1,\ldots,\alpha^K)$ must be a codomination system.

Furthermore, if $V_*(c,\alpha^k) = 0$ for all k, then the above-constructed η^1 for $b = c$ would be in $\Phi(c)$ and would violate (5.1), since $\alpha^k \cdot \theta(\eta^1)$ would equal zero for all k. Thus, if $c \notin E$ then we can find a codomination system $(\alpha^1,\ldots,\alpha^K)$ such that $V_*(c,\alpha^k) <_\ell 0$ for some k.

Conversely, let us now suppose that $(\alpha^1,\ldots,\alpha^K)$ is a codomination system and $V_*(c,\alpha^k) <_\ell 0$ for some k. We need to show that this implies that $c \notin E$.

Let $\varepsilon > 0$ be chosen small enough so that

$$\left|V_{N\setminus S}((d_{-S},e_S),\alpha^k)\right| > \sum_{R\supset S} \sum_{\substack{b_R \in C_R \\ b_S = e_S}} \varepsilon \left|V_{N\setminus S}((d_{-R},b_R),\alpha^k)\right|,$$

for every k and (d,e_S) such that $V_{N\setminus S}((d_{-S},e_S),\alpha^k) < 0$. ($|\cdot|$ here denotes absolute value, and the second summation is over all b_R whose components for players in S are the same as in e_S.) If c were in E, then we could also find some ε-correlated equilibrium η such that $\eta(c,\emptyset) > 0$. Let m then be the smallest number such that there exists some d such that $\eta(d,\emptyset) > 0$ and $V_*(d,\alpha^m) <_\ell 0$. Then the definitions of codomination system and ε-correlated strategy imply that

$$\sum_d \sum_{S\subseteq N} \sum_{e_S} \eta(d,e_S)\, V_{N\setminus S}((d_{-S},e_S),\alpha^m) < 0.$$

But this means that $\alpha^m \cdot \theta(\eta) < 0$, and so there must exist some i, d_i, and e_i such that $\theta_i(\eta,e_i|d_i) < 0$, by (5.2). Thus, η could not be an ε-correlated equilibrium, and so $c \notin E$. Q.E.D.

<u>Proof of Theorem 4.</u> Let us pick an arbitrary number ε such that $0 < \varepsilon < \varepsilon^*$, where ε^* is as in the proof of Theorem 3. Let $\lambda = \lambda^\varepsilon$, where λ^ε is as in the proof of Theorem 2. That is, λ is an ε-correlated equilibrium and

$$\lambda(c,\emptyset) > 0 \quad \text{if} \quad c \in E,$$

$$\lambda(c,\emptyset) = 0 \quad \text{if} \quad c \in C \setminus E.$$

Suppose that, contrary to the theorem, there is a correlated strategy μ that satisfies (3.1) and (4.1) but is not a correlated equilibrium of Γ. We pick a small positive number δ and let $\nu = \delta\,\lambda + (1-\delta)\,\mu$, so that

$$\nu(c,b_S) = \begin{cases} \delta\,\lambda(c,b_S) & \text{if} \quad b_S \neq \emptyset, \\ \delta\,\lambda(c,\emptyset) + (1-\delta)\,\mu(c) & \text{if} \quad b_S = \emptyset. \end{cases}$$

For any δ between 0 and 1, ν is an ε-correlated strategy. We choose δ small enough so that ν violates the same incentive constraints that μ violates. Since λ is an ε-correlated equilibrium and μ satisfies (4.1), the only violated incentive constraints for ν involve a player gaining by disobediently choosing an unacceptable action when some acceptable action is recommended.

So let player i and actions e_i and d_i be chosen so that $e_i \in E_i$ $d_i \in C_i \setminus E_i$, and d_i is the optimal disobedience when e_i is recommended to player i under the ε-correlated strategy ν.

Choosing a very small positive number γ, we now define π so that

$$\pi(c,b_S) =$$

$$= \begin{cases} \gamma\ \nu((c_{-i},e_i),b_S), & \text{if } c_i = d_i, \\ \lambda(c,b_S) - \gamma\ \nu(c,b_S), & \text{if } c_i = e_i,\ i \in S, \text{ and } b_i \neq d_i \\ \lambda(c,b_S) - \gamma\ \nu(c,b_S) - \gamma\ \nu(c,b_{S-i}), & \text{if } c_i = e_i,\ i \in S, \text{ and } b_i = d_i, \\ \lambda(c,b_S) & \text{otherwise.} \end{cases}$$

Choosing γ sufficiently small, we can assure that π is an ε^*-correlated strategy. (Nonnegativity of $\pi(c,b_S)$ can be assured, for sufficiently small γ, because $\nu(c,b_S)$ is only positive if c is in E, in which case $\lambda(c,b_S)$ is also positive. Relaxing from ε to ε^* assures that the bounds on tremble probabilities for an ε^*-correlated equilibrium are satisfied, for all sufficiently small γ.) In effect, π differs from λ only in that the mediator is sometimes recommending d_i under π when he would have recommended e_i to a trambling i under λ; and player i only obeys a recommendation to use d_i under π when he would have trembled to d_i under λ. Thus, any recommendation to any player other than i, and any recommendation other than d_i to player i, must be incentive compatible under π, because it is so under λ. Furthermore, a recommendation to use d_i is incentive compatible under π, because i's conditional probability distribution over others' recommendations and actions when d_i is recommended to him under π is the same as when e_i is recommended to him under ν, and d_i is an optimal action for i when e_i is recommended to him under ν. Thus, π is an ε^*-correlated equilibrium, and d_i is recommended with positive probability under π. But by the way ε^* was chosen, this implies that d_i must be in E_i. This result contradicts our initial assumption that d_i was an unacceptable action preferred by i over e_i when e_i is recommended under μ. This contradiction proves the theorem. Q.E.D.

References

R. J. Aumann [1974], "Subjectivity and Correlation in Randomized Strategies," Journal of Mathematical Economics 1, 67-96.

D. Bernheim [1984], "Rationalizable Strategic Behavior," Econometrica 52, 1007-1028.

E.E. C. van Damme [1981], "Regular Equilibrium Points of n-Person Games in Normal Form," Eindhoven Institute of Technology discussion paper. To appear in Mathematics of Operations Research.

E. E. C. van Damme [1983], Refinements of the Nash Equilibrium Concept. Berlin: Springer-Verlag.

J. Farrell [1983], "Communication in Games," M.I.T. discussion paper.

J. C. Harsanyi [1975], "The Tracing Procedure: A Bayesian Approach to Defining a Solution for n-Person Noncooperative Games," International Journal of Game Theory 4, 61-94.

E. Kalai and D. Samet [1984], "Persistent Equilibria in Strategic Games," International Journal of Game Theory 13, 129-144.

E. Kohlberg and J. F. Mertens [1982], "On the Strategic Stability of Equilibria," CORE Discussion Paper No. 8248, Université Catholique de Louvain.

D. M. Kreps and R. Wilson, "Sequential Equlibria," Econometrica 50, 863-894.

H. Kuhn [1953], "Extensive Games and the Problem of Information," in Contributions to the Theory of Games 2, H. Kuhn and A. Tucker, eds. Princeton: Princeton University Press, 193-216.

R. D. Luce and H. Raiffa [1957], Games and Decisions. New York: John Wiley and Sons.

A. McLennan [1983], "Justifiable Beliefs in Sequential Equilibrium," University of Toronto discussion paper. To appear in Econometrica.

R. B. Myerson [1978], "Refinements of the Nash Equilibrium Concept," International Journal of Game Theory 1, 73-80.

R. B. Myerson [1982], "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," Journal of Mathematical Economics 10, 67-81.

R. B. Myerson [1983], "Bayesian Equilibrium and Incentive Compatibility: An Introduction," Northwestern University discussion paper. To appear in Social Goals and Social Organization, L. Hurwicz, D. Schmeidler, and H. Sonnenschein, eds.

R. B. Myerson [1985], "Multistage Games with Communication," Center for Mathematical Studies discussion paper No. 590, Northwestern University.

J. F. Nash [1951], "Non-cooperative Games," Annals of Mathematics 54, 286-295.

D. G. Pearce [1984], "Rationalizable Strategic Behavior and the Problem of Perfection," Econometrica 52, 1029-1050.

R. T. Rockafellar [1970], Convex Analysis. Princeton: Princeton University Press.

R. Selten [1975], "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," International Journal of Game Theory 4, 25-55.