EXTENSIVE-FORM CORRELATED EQUILIBRIA

By Eilon Solan

The paper studies extensive-form correlated equilibria in stochastic games. An extensive-form correlated equilibrium is an equilibrium in an extended game, where a correlation device chooses at every stage, as a function of past signals (but independently of the actions of the players) a private signal for each player.

We define the notion of individually rational payoffs for stochastic games, and characterize the set of extensive-form correlated equilibrium payoffs using feasible and individually rational payoffs. Our result implies that extensive-form correlated equilibria and communication equilibria are payoff equivalent in stochastic games.

1. INTRODUCTION

Correlated equilibria were introduced by Aumann (1974, 1987) for one-shot games. A correlated equilibrium is a probability distribution over the set of action combinations, such that if (i) a mediator chooses an action combination according to this distribution and reveals to each player his action in the combination and (ii) each player assumes that all other players follow the recommendation of the mediator, then no player can profit by disobeying the recommendation of the mediator and choosing another action.

Correlated equilibria are an appropriate solution concept as soon as preplay communication between the players is possible, or whenever players have different information or beliefs. In the words of Aumann (1987), "if it is common knowledge that the players in a game are Bayesian utility maximizers who treat uncertainty about other players' actions like any other uncertainty, then the outcome is necessarily a correlated equilibrium."

An equivalent formulation of correlated equilibria is by introducing correlation *devices*. A correlation device chooses before the start of play a signal to each player according to some known joint distribution. Any Nash equilibrium in an extended game that includes a correlation device, where the players can base their choice of action on the signal they receive from the device, is a correlated equilibrium.

Aumann (1974) provided several examples in which all players profit by using a correlation device, i.e., games where the unique Nash equilibrium gives all the players strictly less than some of the correlated equilibria.

In the present paper we are interested in dynamic games; games that are played in stages, and the payoff for the players depends on the whole history.

Consider for example a multi-stage game with observable actions (see, e.g., Fudenberg and Levine (1983) or Fudenberg and Tirole (1991)). The game is played for some fixed number of stages. At every stage the players observe the actions that were played by all the players in all previous stages, and choose simultaneously actions for that stage. The payoff, that is given to each player at the end of the game, depends on the whole play.

When defining the notion of correlation devices in such a model, several generalizations come to mind. First, the device can choose one private signal for each player before the start of play, as in one-shot games, and reveal to each player his chosen signal (*correlation device*, Forges (1986, 1988)).

However, one can conceive of a device that chooses a private signal for each player before every stage, and reveals to each player his chosen signal before this player can choose an action. In this case, the data on which the device bases its choice may vary. In the most general case, the device may base its choice on daily messages that it receives from the players (*communication device*, Forges (1986, 1988), Myerson (1986) or Mertens (1994)). One can restrict oneself to devices that base their choice only on previous signals that they chose, and *not* on any other data (*autonomous communication device*, Forges (1986, 1988)).

In the present paper we are interested in autonomous correlated devices: devices that choose a signal for each player at every stage, and the signals depend only on previous signals, and not on the actions of the players. Thus, there is no communication between the players, and correlation is achieved by the lotteries of the device.

EXAMPLE 1 *(Myerson 1986)*

Consider the following 2-player two-stage game (in all the examples that we give player 1 is the row player and player 2 is the column player):

stage 1

T	**2**
B	$2,2$

stage 2

	L	R
T	$0,0$	$1,5$
B	$5,1$	$0,0$

Figure 1

At the first stage player 1 chooses either Top or Bottom. If he chose Top, then at the second stage player 1 chooses either Top or Bottom and player 2 chooses either Left or Right. The payoff for the players is indicated in Figure 1.

The Nash equilibrium payoffs are $(2,2)$ and $(5,1)$. The payoff $(2,2)$ can be attained by the following equilibrium profile: (i) at stage 1, player 1 plays Bottom and (ii) at stage 2, player 1 plays Top and player 2 plays Right. The payoff $(5,1)$ can be attained by the following equilibrium profile: (i) at stage 1, player 1 plays Top and (ii) at stage 2, player 1 plays Bottom and player 2 plays Left. By using a suitable correlation device the players can implement $(3,3)$ as an equilibrium payoff:

- Player 1 chooses Top at the first stage.
- A device chooses with probability 1/2 the pair (Bottom,Left) and with probability 1/2 the pair (Top,Right), and reveals his choice to the players.
- The players follow the recommendation of the device at the second stage.

Note that in order to have this mechanism an equilibrium, it is necessary that player 1 knows that a lottery will be performed before stage 2, but he must not know the outcome of this lottery before he chooses his action at stage 1.

A Nash equilibrium payoff in an extended game that includes an autonomous correlation device is an *extensive-form correlated equilibrium payoff*. Thus, the set of extensive-form correlated equilibrium payoffs is generally larger than the set of Nash equilibrium payoffs.

If the players have the same finite recall, that is, at stage n the players are told the actions that were played only at stages $n-1, n-2, \ldots, n-k$, or more generally, in a situation of symmetric incomplete information, the game includes information sets — there are different histories that are indistinguishable by the players. Thus, the behavior of the players must be the same after such indistinguishable histories. Nevertheless, the payoff for the players depends on the whole history, regardless of the information of the players.

In addition, various moves in the game can be made by nature, and the results of these moves be announced to the players (or, the players may receive some information on those moves).

A natural model that includes these two extensions of multi-stage games, as well as other models (like repeated games with incomplete information on one side (see, e.g., Aumann and Maschler (1995))), is the model of stochastic games, presented by Shapley (1953).

A stochastic game is played in stages. At every stage each player chooses an action, and a new state of the world is chosen according to a probability distribution that depends on the current state of the world and the actions chosen by the players. In contrast to standard stochastic games, the payoff is not an aggregation of some daily payoffs, but it is a function of the play.

Stochastic games were applied in many contexts: resource extraction (Levhari and Mirman (1980), Amir (1987), Sundaram (1989) and Majumdar and Sundaram (1991)), Altruistic growth (Bernheim and Ray (1983) and Leininger (1986)), racing models (Winston (1978) and Harris and Vickers (1987)) and dynamic duopoly (Cyert and DeGroot (1970) and Maskin and Tirole (1988a,b)).

In the present paper we study stochastic games in a most general set-up. The state space, action spaces of the players and the payoff function may be arbitrary. Since the set-up is general, we are concerned with ϵ-Nash equilibria of the extended game that includes a correlation device, whereby players will be able to profit at most ϵ by deviating from the equilibrium path

Our main result concerns a characterization of the set of extensive-form equilibrium payoffs using feasible and individually rational payoffs.

The first question that one asks in such a characterization is, what is an individually rational payoff ? In a dynamic game, a payoff that is individually rational today may be irrational tomorrow.

For example, consider the following game, where the payoffs are those of player 2.

stage 1

T	**2**
B	-2

stage 2

L	R
0	-1

Figure 2

It is clear that -1 is an individually rational payoff for player 2, but, if player 1 plays T, then, unless there is some binding agreement, player 2 will never agree to play R.

Thus, we are led to define individually rational strategy profiles instead of individually rational payoffs. Intuitively, a strategy profile is *individually rational* if, after any history, no player prefers to play some action and then be punished forever, than to play any action that has a positive probability according to his strategy, knowing that in the future everyone will continue to follow that profile.

We then prove that for any individually rational strategy profile corresponds an autonomous correlation device and an equilibrium strategy profile in the extended game, which are payoff-equivalent, and vice-versa. By defining the set of feasible and individually rational payoffs as the set of all payoffs of individually rational strategy profiles, we conclude that the set of extensive-form correlated equilibrium payoffs coincides with the set of feasible and individually rational payoffs.

If the state space and action spaces are countable, the device can base its choice only on previous signals, and not on any other data. If they are not countable, we must assume that the device knows the moves done by nature. Alternatively, the device may base its choice on previous states of the world as well as previous signals. However, the device never bases its choice of new signal on the actions taken by the players.

Though the result may remind one of the Folk Theorem, there is a significant difference: whereas the Folk Theorem characterizes the set of equilibrium payoffs in a repeated game by means of the one-shot game, in a general stochastic game the payoff depends on the whole play.

Our result implies that communication between the players is not needed in order to achieve *any* feasible and individually rational payoff as an extensive-form correlated equilibrium payoff.

In addition, we provide an example of a multi-stage game where all the players profit by using an autonomous correlation device. The main ideas of our proofs appear in the context of the autonomous correlation device that is constructed for this game. In this example, the unique correlated equilibrium is also the unique Nash equilibrium of the game, thus, using a device that sends messages only before the start of play cannot help the players.

Our work is related to that of Myerson (1986), who studies multi-stage games, and characterizes the set of sequential communication equilibria using codominated actions. Nevertheless, there are some important differences. First, Myerson's equilibria are sequential, while in our equilibria players may be required to punish a deviator, which may be irrational for some of the players (though punishment never occurs on the equilibrium path). Second, Myerson is concerned with finite multi-stage games, the players in his set-up have asymmetric information, and they send messages to the correlation device, whereas we study general stochastic games (which means that the information is symmetric) and players cannot send messages to the device.

One can obtain similar results if one uses the notion of correlation as defined by Moulin and Vial (1978) instead of the notion defined by Aumann

(1974, 1987).

The paper does not address the question of whether the set of feasible and individually rational payoffs is empty or not. When the game lasts for finitely many stages the existence of an equilibrium (and therefore, of a feasible and individually rational payoff) is clear. In the general case, non-emptiness of this set was proved for the discounted payoff by Nowak (1991), Mertens and Parthasarathy (1987) and Solan (1998), and for the undiscounted payoff only when the state and action spaces are finite (Solan and Vieille (1998)).

The paper is arranged as follows. In Section 2 we illustrate the basic ideas underlying the correlation device that we shall use by an example, in Section 3 we present the model, various types of correlation devices, and the main results, and in Section 4 we prove the main results of the paper.

2. EXAMPLE

Consider the following two-player multi-stage game:

stage 1

	L	C	R
T	$1,1$	$1,0$	$0,2$
B	**2**	$0,0$	$1,-4$

stage 2

L	R
$3,-1$	$0,-2$

Figure 3

At stage 1, player 1 chooses a row, and player 2 independently chooses a column. If the players chose (B, L) then the game continues to stage 2, where player 2 chooses a cell. If the players chose another pair of actions at the first stage, or after the choice of player 2 at the second stage, the players receive a payoff as indicated in Figure 3.

Clearly, if the game reaches stage 2, and player 2 is rational, he will choose L, and the players will receive the payoff $(3, -1)$.

We shall now verify that the game has a unique correlated equilibrium, and therefore a unique Nash equilibrium. Assume that at the first stage the correlation device chooses the cells according to the following distribution:

	L	C	R
T	α	γ	ϵ
B	β	δ	ν

Figure 4

where $\alpha + \beta + \gamma + \delta + \epsilon + \nu = 1$ and $\alpha, \beta, \gamma, \delta, \epsilon, \nu \geq 0$. If this distribution is a correlated equilibrium, then the following inequalities must hold:

$$\begin{aligned} \alpha - \beta &\geq 0, 2\alpha - 4\beta \\ 0 &\geq \gamma - \delta, 2\gamma - 4\delta \\ 2\epsilon - 4\nu &\geq 0, \epsilon - \nu \\ \alpha + \gamma &\geq 3\alpha + \epsilon \\ 3\beta + \nu &\geq \beta + \delta. \end{aligned}$$

It follows that

$$2\beta + \nu \geq \delta \geq \gamma \geq 2\alpha + \epsilon \geq 2\beta + \epsilon \geq 2\beta + 3\nu,$$

which implies that $\epsilon = \nu = 0$ and $\gamma = \delta = 2\alpha = 2\beta$. Thus, the unique Nash equilibrium is for player 1 to play at the first stage $(1/2, 1/2)$ and for player 2 to play at the first stage $(1/3, 2/3, 0)$, whereby the players receive an expected payoff of $(1, 0)$. Moreover, the unique correlated equilibrium coincides with the distribution over the entries of the matrix induced by this Nash equilibrium.

We claim that any point on the interval $(3/2, 1/2)$-$(2, 0)$ is an equilibrium payoff when the players use an appropriate mechanism. Consider the following mechanism:

(i) Before the start of play, a mediator chooses randomly either T with probability x or B with probability $1 - x$, where $1/2 \leq x \leq 3/4$.

(ii) The mediator tells player 1 his choice.

(iii) At stage 1, player 1 follows the choice of the mediator, and player 2 plays L.

(iv) After stage 1, but before stage 2, the mediator reveals his choice to player 2 (and thus, his *recommendation* to player 1 at the first stage).

(v) At the second stage, player 2 plays L or R, according to whether the choice of the mediator was B or T. Thus, if player 1 did not follow the recommendation of the mediator, he is "punished".

If the players follow this mechanism then the expected payoff is $(3-2x, 2x-1)$. Moreover, no player has any profitable deviation.

In order for this mechanism to be an equilibrium, it is necessary that player 2 not know the choice of the mediator before stage 1 (otherwise, he has a profitable deviation), and it is necessary that player 1 knows that the choice will be revealed to player 2 before stage 2 (so that not following the recommendation of the mediator is not profitable).

3. THE MODEL AND THE MAIN RESULTS

For every measurable space Y we denote by $\mathcal{P}(Y)$ the space of probability measures over Y. If $\mu \in \mathcal{P}(Y)$ and $C \subset Y$ is a measurable set, then $\mu[C]$ is the measure of C according to μ. A function $f : X \to \mathcal{P}(Y)$ is measurable if for every measurable subset $C \subset Y$ the function $g : X \to [0,1]$ defined by $g(x) = f_x[C]$ is measurable. A product (*resp.* union) of measurable spaces is always endowed with the product (*resp.* union) σ-algebra. Finally, a correspondence is a set-valued function, and a correspondence $\phi : X \to Y$ is measurable if the set $\{x \in X \mid \phi(x) \cap C \neq \emptyset\}$ is X-measurable for every Y-measurable set C.

A *stochastic game* G is given by:

(i) A finite set of players N.

(ii) A measurable space of states S.

(iii) An initial state $s_\star \in S$.

(iv) For every player $i \in N$, a σ-compact metric space of pure actions A_0^i. We denote $A_0 = \times_{i \in N} A_0^i$.

(v) For every player $i \in N$, a measurable correspondence $A^i : S \to A_0^i$ with closed values. $A^i(s)$ is the set of actions available for player i in state s. We denote $A(s) = \times_{i \in N} A^i(s)$. The space of *infinite histories* is denoted by H_∞:

$$H_\infty = \{(s_1, a_1, s_2, a_2, \ldots) \mid s_1 = s_\star, a_n \in A(s_n), s_n \in S \qquad \forall n \in \mathbf{N}\}.$$

We endow H_∞ with the σ-algebra generated by all the finite cylinders.

(vi) A measurable transition rule q that assigns for each $(s,a) \in \mathrm{Gr}(A)$ a probability measure in $\mathcal{P}(S)$.

(vii) For every player $i \in N$, a measurable bounded utility function $u^i : H_\infty \to [-R, R]$, where $R \in \mathbf{R}$.

The game is played in stages. The initial state of the game is $s_1 = s_\star$. At stage n each player is informed of the whole history (including the current state s_n), and chooses an action $a_n^i \in A^i(s_n)$, independently of the other players. The action combination $a_n = (a_n^i)$ that was chosen and the current state s_n determine a new state s_{n+1} according to the probability measure $q(s_n, a_n)$.

The payoff for each player i is determined by the *infinite* path that has occurred, and is equal to $u^i(s_1, a_1, s_2, a_2, \ldots)$.

Note that our definition of a utility function is more general than the standard approach of using daily payoffs. The utility function can be the discounted sum, the lim sup, lim inf or any Banach limit of some daily payoffs.

3.1 ON STRATEGIES

A history of length n is a sequence $h = (s_1, a_1, \ldots, a_n, s_n)$ where $s_1 = s_\star$, $s_k \in S$ and $a_k \in A(s_k)$ for $k = 1, \ldots, n$. The last state of a history h is denoted by $s_L(h)$. The history $(s_\star)$ is denoted by $s_\star$. The space of all finite histories is denoted by H.

Definition 3.1: A *strategy* of player i is a measurable function $\sigma^i : H \to \mathcal{P}(A_0^i)$ such that $\sigma^i(h) \in \mathcal{P}(A^i(s_L(h)))$. A *profile* is a vector of strategies $\sigma = (\sigma^i)_{i \in N}$. A *correlated profile* is a measurable function $\sigma : H \to \mathcal{P}(A_0)$ such that $\sigma(h) \in \mathcal{P}(A(s_L(h)))$.

Note that every profile is a correlated profile. We denote by Σ^i the space of profiles of player i, by $\Sigma_\star$ the space of correlated profiles, and by $\Sigma_\star^{-i}$ the space of correlated profiles of players $N \setminus \{i\}$; that is, the space of measurable functions $\sigma^{-i} : H \to \mathcal{P}(A_0^{-i})$ such that $\sigma^{-i}(h) \in \mathcal{P}(A^{-i}(s_L(h)))$ for every $h \in H$, where $A^{-i}(s_L(h)) = \times_{j \neq i} A^j(s_L(h))$.

By Ionescu-Tuclea Theorem (see, e.g., Neveu (1965), Proposition V.1.1), every finite history $h \in H$ and every correlated profile σ induce a probability measure $\mathbf{P}_{h,\sigma}$ over H_∞; that is, the probability measure induced by σ in the subgame beginning with h. We denote expectation w.r.t. this probability measure by $\mathbf{E}_{h,\sigma}$.

3.2 ON PAYOFFS

For every correlated profile σ and every finite history h we denote

$$\gamma^i(h, \sigma) = \mathbf{E}_{h,\sigma} u^i(s_1, a_1, \ldots).$$

The *payoff* of a correlated profile σ is defined by $\gamma(\sigma) = (\gamma^i(s_\star, \sigma))_{i \in N}$.

For every finite history $h \in H$ we define the *punishment level* by:

$$v_h^i = \inf_{\sigma^{-i} \in \Sigma_\star^{-i}} \sup_{\sigma^i \in \Sigma^i} \gamma^i(h, \sigma).$$

v_h^i is the punishment level that players $N \setminus i$ can inflict on player i when using a correlation device.

We assume that for every $\epsilon > 0$ and every player $i \in N$ there exists a correlated profile $\tilde{\sigma}_\epsilon^{-i} \in \Sigma_\star^{-i}$ such that

$$|v_h^i - \sup_{\sigma^i \in \Sigma^i} \gamma^i(h, (\tilde{\sigma}_\epsilon^{-i}, \sigma^i))| < \epsilon.$$

That is, there exists a measurable ϵ-punishment profile. In general, such a correlated profile need not exist. However, in various special cases such a correlated profile is known to exist: (i) if the state and action spaces are countable, then there are no measurability issues, and (ii) if the utility function is the discounted sum or the limsup of daily payoffs, then existence was proved in general set-ups (see, e.g., Mertens, Sorin and Zamir (1994) for the discounted sum, and Maitra and Sudderth (1993) for the limsup).

3.3 THE MEASURE OF IRRATIONALITY OF STRATEGIES

In this section we assign to each correlated profile σ and every player i a non-negative number $U^i(\sigma)$, that measures how much player i can profit by deviating from σ^i, provided his deviation is followed by an indefinite punishment. In other words, $U^i(\sigma)$ is the measure of irrationality of following σ for player i.

Let $h \in H$ be a finite history of length n and σ be a correlated profile. For every player i we define

$$V_h^i(\sigma) = \sup_{b^i \in A^i(s_L(h))} \mathbf{E}_{\sigma^{-i}(h)} v^i_{h,(a_n^{-i},b^i),s_{n+1}}.$$

Given that the history h has occurred, today players $N \setminus \{i\}$ follow σ, but tomorrow they start to punish player i, $V_h^i(\sigma)$ is the maximal payoff that player i can guarantee.

For every action b^i of player i we define

$$U^i(h,\sigma,b^i) = \max\{0, V_h^i(\sigma) - \mathbf{E}_{\sigma^{-i}(h),b^i}\gamma^i((h,(a_n^{-i},b^i),s_{n+1}),\sigma)\}.$$

That is, given that player i should play the action b^i, $U^i(h,\sigma,b^i)$ is his loss compared to deviating and being punished from the next stage on. Denote

$$U^i(h,\sigma) = \mathbf{E}_{\sigma^i(h)} U^i(h,\sigma,b^i),$$

the expected loss of player i if he follows σ, given that the history h has occurred.

Any measurable stopping time $t : H_\infty \to \mathbf{N}$ and every correlated profile σ induce, by Ionescu-Tuclea Theorem, a probability measure over H. Denote expectation w.r.t. this measure by $\mathbf{E}_{t,\sigma}$. Define the *measure of irrationality of σ* for player i by

$$U^i(\sigma) = \sup_t \mathbf{E}_{t,\sigma} U^i(h,\sigma)$$

where the supremum is over all measurable stopping times. In other words, given that the players should follow σ, player i may stop following σ whenever he chooses. However, one stage afterwards, he is being punished with his punishment level. $U^i(\sigma)$ measures the maximal amount that player i can profit by such a process, where the profit is measured w.r.t. following σ indefinitely.

DEFINITION 3.2: The *set of all feasible and individually rational payoffs* is the set E_1 of all vectors $v \in \mathbf{R}^N$ such that for every $\epsilon > 0$ there exists a correlated profile σ that satisfies:

(i) $|\gamma^i(\sigma) - v^i| < \epsilon$ for every player $i \in N$.

(ii) $U^i(\sigma) < \epsilon$ for every player $i \in N$.

3.4 ON CORRELATION

DEFINITION 3.3: An *autonomous correlation device* is a pair $\mathcal{E} = ((M^i)_{i\in N}, (\mathcal{E}_n)_{n\in\mathbf{N}})$ where

- M^i is a measurable space of *signals* for player i. We denote $M = M(\mathcal{E}) = \times_{i\in N} M^i$.
- $\mathcal{E}_n : (S \times M)^{n-1} \times S \to \mathcal{P}(M)$ is a measurable function.

An autonomous correlation device is a *correlation device* if $\mathcal{E}_n(m_1, s_1, \ldots, m_{n-1}, s_{n-1})$ is an atom for every $n > 1$ (i.e. the players can receive an "informative" signal only before the beginning of the game).

Given an autonomous correlation device $\mathcal{E} = ((M^i)_{i\in N}, (\mathcal{E}_n)_{n\in\mathbf{N}})$ we define a new game $G(\mathcal{E})$ which is played like the game G, but at every stage n, *before* the players choose actions, the device chooses a signal $m_n = (m_n^i)_{i\in N}$ according to the probability measure $\mathcal{E}_n(s_1, m_1, \ldots, m_{n-1}, s_n)$, and sends to each player i the signal m_n^i. Each player can base his choice of an action on all the signal that he has received from the device.

Let $H^i(M)$ be the space of all finite histories that player i can observe in $G(\mathcal{E})$; that is, the space of all sequences $(s_1, m_1^i, a_1, \ldots, s_{n-1}, m_{n-1}^i, a_{n-1}, s_n, m_n^i)$ such that $s_1 = s_\star$, $a_k \in A(s_k)$ and $m_k^i \in M^i$. Note that, since the signals are private, each player observes a different history. Let $H(M)$ be the space of all finite histories that an outside observer, who observes *both* the actions of the players and the signals sent to *all* the players, can observe. Let $H_\infty(M)$ be the space of all infinite histories that this outside observer can observe. We endow $H_\infty(M)$ with the σ-algebra generated by all the finite cylinders. Note that the spaces $(H^i(M))_{i\in N}$, $H(M)$ and $H_\infty(M)$ are independent of $(\mathcal{E}_n)_{n\in\mathbf{N}}$.

A *strategy* for player i in $G(\mathcal{E})$ is a measurable function $\tau^i : H^i(M) \to \mathcal{P}(A^i)$ such that $\tau^i(h) \in \mathcal{P}(A^i(s_L(h)))$ for every $h \in H(M)$. A *profile* $\tau = (\tau^i)_{i\in N}$ is a vector of strategies, one for each player.

In the sequel, σ always refer to correlated profiles in the game G, and τ refers to profiles in the extended game $G(\mathcal{E})$.

For every history $(s_1, m_1, a_1, \ldots, s_n, m_n) \in H(M)$ we denote

$$\tau(s_1, m_1, a_1, \ldots, s_n, m_n) = (\tau^i(s_1, m_1^i, a_1, \ldots, s_n, m_n^i))_{i\in N}.$$

By Ionescu-Tuclea Theorem, every autonomous correlation device $\mathcal{E}$, every profile τ in $G(\mathcal{E})$ and every finite history $h \in H$ induce a probability measure over $H_\infty(M)$. We denote expectation w.r.t. this measure by $\mathbf{E}_{h,\mathcal{E},\tau}$. Define for every finite history $h \in H(M)$, the expected payoff w.r.t. τ by

$$u^i_{\mathcal{E}}(h, \tau) = \mathbf{E}_{h,\mathcal{E},\tau} u^i(s_1, a_1, \ldots).$$

Define $\mathbf{R}^N_{++} = \{r \in \mathbf{R}^N \mid r^i > 0 \quad \forall i \in N\}$. Addition in $\mathbf{R}^N$ is defined coordinate-wise.

Definition 3.4: Let $\epsilon \in \mathbf{R}^N_{++}$. A profile τ in $G(\mathcal{E})$ is an *ϵ-equilibrium* if for every player $i \in N$ and every strategy τ'^i of player i in $G(\mathcal{E})$

$$u^i_{\mathcal{E}}(s_\star, \tau) > u^i_{\mathcal{E}}(s_\star, (\tau^{-i}, \tau'^i)) - \epsilon^i.$$

Definition 3.5: The *set of extensive-form correlated equilibrium payoffs* is the set E_2 of all vectors $v \in \mathbf{R}^N$ such that for every $\epsilon \in \mathbf{R}^N_{++}$ there exists an autonomous correlation device $\mathcal{E}$ and an ϵ-equilibrium profile τ in $G(\mathcal{E})$ such that $|u^i_{\mathcal{E}}(s_\star, \tau) - v^i| < \epsilon^i$ for every player $i \in N$.

3.5 THE MAIN RESULTS

The main result of the paper is that the set of feasible and individually rational payoffs coincides with the set of extensive-form correlated equilibrium payoffs.

Theorem 3.6: $E_1 = E_2$

This result follows from two propositions. Proposition 3.7 claims that for every correlated profile and every $\epsilon > 0$, there exists an autonomous correlation device and a strategy profile in the extended game, such that if all the players follow this strategy profile in the extended game, then each player i can profit by deviating at most his measure of irrationality plus ϵ. Moreover, the two strategy profiles are payoff-equivalent.

Proposition 3.8 claims that for every ϵ-equilibrium strategy profile in an extended game there exists a correlated profile in the original game, such that its measure of irrationality is smaller than ϵ.

PROPOSITION 3.7: *For every correlated profile σ and every $\epsilon \in \mathbf{R}^N_{++}$ there exists an autonomous correlation device $\mathcal{E}$ and a $U(\sigma)+\epsilon$-equilibrium profile τ in $G(\mathcal{E})$ such that $u_{\mathcal{E}}(s_\star, \tau) = u(s_\star, \sigma)$. If (i) the profile σ in* not *correlated, or (ii) the state and action spaces are countable, then the autonomous correlation device can depend only on previous signals, and not on previous states.*

PROPOSITION 3.8: *Let $\epsilon \in \mathbf{R}^N_{++}$. For every autonomous correlation device $\mathcal{E}$ and ϵ-equilibrium profile τ in $G(\mathcal{E})$ there exists a correlated profile σ such that $u_{\mathcal{E}}(s_\star, \tau) = u(s_\star, \sigma)$ and $U^i(\sigma) \leq \epsilon^i$ for every $i \in N$.*

When the max-min level of each player is constant over the space of finite histories, a stronger result holds. In such a case, no correlation is needed *along* the play in order to sustain any feasible and individually rational payoff as a correlated equilibrium payoff.

THEOREM 3.9: *If v^i_h is independent of h for every fixed player $i \in N$, then the set of feasible and individually rational payoffs coincides with the set of correlated equilibrium payoffs.*

Remark: Though uniform equilibrium payoffs (see, e.g., Mertens, Sorin and Zamir (1994)) are not in the scope of our model (since the uniform equilibrium payoff cannot be defined as a limit of ϵ-equilibrium payoffs using some utility function) similar results can be obtained, with analogous proofs.

4. THE PROOFS

Whenever we refer to the unit interval, we mean the interval $[0, 1)$, equipped with the σ-algebra of Borel sets and with the Lebesgue measure λ.

First of all we prove that for every correlated profile σ there exists a correlation device $\mathcal{E}$ and a profile τ in $G(\mathcal{E})$ that mimic σ — they both induce the same probability measure over H_∞ (and therefore yield the same payoff).

LEMMA 4.1: *For every correlated profile σ there exists a correlation device $\mathcal{E}$ and a profile τ in $G(\mathcal{E})$ such that $\mathbf{P}_{s_\star,\sigma}$ is equal to the marginal probability measure induced by $\mathbf{P}_{s_\star,\mathcal{E},\tau}$ over H_∞.*

Proof: We define a correlation device that (i) chooses before the game a sequence of numbers $(z_1, z_2, \ldots)$ in the unit interval such that each number is chosen independently from the others with the uniform distribution, and (ii) reveals the whole sequence to all the players. The players, who observe the history, should choose after each history h, an action combination according to $\sigma(h)$. They use the nth number in the sequence (where n is the length of the history) in order to choose this action.

Since the set-up is general, the only point that should be verified is whether this choice can be made using a single number, and can it be made in a measurable way.

First we shall assume that A_0^i is finite for each player $i \in N$. In this case, $A(s)$ is finite for every state s. Let h be the finite history up to stage n. Then at stage n, the players should perform the correlated lottery $\sigma(h)$ over $A(s_L(h))$, using the number z_n.

Denote

$$\mathrm{supp}(\sigma(h)) = \{a_1, \ldots, a_R\}$$

and $\mu_r = \sigma(h)[a_r]$ for every $r = 1, \ldots, R$. Each player i will play at stage n the action a_r^i, where $r \in \{1, \ldots, R\}$ is the unique integer that satisfies

$$\sum_{r' < r} \mu_{r'} < z_n \leq \sum_{r' \leq r} \mu_{r'}.$$

For the general case, we need the following lemma, which is proved in the Appendix.

LEMMA 4.2: *Let H be a measurable space and A_0 a σ-compact metric space. Let $f : H \to \mathcal{P}(A_0)$ be a measurable function. There exists a measurable correspondence $g : H \times A_0 \to [0, 1)$ that satisfies:*

(i) $(g_h(a))_{a\in A_0}$ *is a partition of* $[0,1)$ *for every fixed* $h \in H$*.*

(ii) $\lambda[g_h(C)] = f_h[C]$ *for every* $h \in H$ *and every measurable subset* $C \subset A_0$*, where* $g_h(C) = \cup_{a\in C} g_h(a)$ *(recall that* λ *is the Lebesgue measure).*

Apply Lemma 4.2 for $f = \sigma$ to get a measurable correspondence $g : H \times A_0 \to [0,1)$. At stage n, given that history h has occurred, the players play the action combination $g_h^{-1}(z_n)$.

The measurability of the profile follows from the measurability of g and σ. ■

By applying Lemma 4.1 to the punishment profile $\tilde{\sigma}^i_\epsilon$ we have: COROLLARY 4.3: *For every player* i *and every* $\epsilon > 0$ *there exists a correlation device* $\mathcal{E}$ *and a profile* τ^{-i} *for players* $N \setminus \{i\}$ *in* $G(\mathcal{E})$ *such that for every strategy* τ^i *of player* i *in* $G(\mathcal{E})$ *we have*

$$u^i_{\mathcal{E}}(h,\tau) < v^i_h + \epsilon \qquad \forall h \in H(M).$$

Any profile τ in $G(\mathcal{E})$ defines a natural correlated profile σ_τ — the behavior observed by an outside observer that does not notice the signals:

$$\sigma_\tau(h) = \mathbf{E}_{h,\mathcal{E}}\tau(s_1, m_1, a_1, \ldots, s_n, m_n),$$

where $h = (s_1, a_1, \ldots, s_n)$. It follows that τ and σ_τ both induce the same probability measure over H_∞. Therefore we have the following result.

LEMMA 4.4 *For every player* i *and every finite history* $h = (s_1, a_1, \ldots, s_n) \in H$

$$\mathbf{E}_{h,\mathcal{E},\tau} u^i_{\mathcal{E}}((s_1, m_1, a_1, \ldots, s_n, m_n), \tau) = u^i(h, \sigma_\tau).$$

Proof of Proposition 3.8:
Let $\mathcal{E} = ((M^i)_{i\in N}, (\mathcal{E}_n)_{n\in\mathbf{N}})$ be an autonomous correlation device, τ be an ϵ-equilibrium profile in $G(\mathcal{E})$ and $\epsilon \in \mathbf{R}^N_{++}$.

Fix a player $i \in N$. Since τ is an ϵ-equilibrium profile, it follows that for every strategy τ'^i of player i in $G(\mathcal{E})$

$$u^i(s_\star, \sigma_\tau) = u^i_{\mathcal{E}}(s_\star, \tau) > u^i_{\mathcal{E}}(s_\star, (\tau^{-i}, \tau'^i)) - \epsilon^i. \qquad (1)$$

Every stopping time t and every strategy $\tilde{\tau}^i$ of player i in $G(\mathcal{E})$ define a strategy τ'^i for player i as follows: follow τ^i as long as the stopping time does not call the process to a halt, and afterwards follow $\tilde{\tau}^i$.

Since (1) holds for every strategy τ'^i of player i, it follows that $U^i(\sigma_\tau) \leq \epsilon^i$, as desired. ∎

Proof of Proposition 3.7:
Let σ be any profile in G, and $\epsilon \in \mathbf{R}^N_{++}$. Choose $\delta = \frac{1}{2} \min_{i \in N} \epsilon^i$.

We shall define an autonomous correlation device that mimics the behavior of σ. At every stage the device chooses a recommended action for each player i according to σ, and sends to each player two signals: (i) the recommended action for him to play at the current stage and (ii) the actions that were recommended for *all* the players at the *previous* stage. Thus, each player observes whether all the players followed the recommendation of the device at the previous stage. If any player has deviated, he is punished with his punishment level (if several players deviate at the same stage, then the deviator with the minimal index is punished). To punish effectively, the device chooses for every player i, *before* the start of play, a sequence of numbers in the unit interval, and sends these numbers to all the players *except* player i. This sequence is used by players $N \setminus \{i\}$ in order to execute the δ-punishment profile $\tilde{\sigma}^{-i}_\delta$ against player i if the need arises.

Formally, define an autonomous correlation device $\mathcal{E}$ as follows:

- Before the beginning of play, the device chooses for each player i a sequence $(z^i_1, z^i_2, \ldots)$ of independent uniformly distributed numbers in the unit interval, and sends each player i the sequences $\{(z^j_1, z^j_2, \ldots)\}_{j \neq i}$. In addition, the device chooses a sequence $(z_1, z_2, \ldots)$ of independent uniformly distributed numbers in the unit interval.

- At stage 1 the device sends to each player the signal a^i_1, where $a_1 = (a^i_1)_i = g^{-1}_{s_1}(z_1)$, and g is the function defined by Lemma 4.2 w.r.t. σ.

- At every stage n ($n > 1$) the device sends to each player i the signal (a^i_n, a_{n-1}), where $a_n = (a^i_n)_i = g^{-1}_{(s_1,a_1,s_2,a_2,\ldots,s_n)}(z_n)$.

We now define a profile τ in $G(\mathcal{E})$. The definition is divided into two parts: for histories where no deviation was detected (players then follow the recommendation of the device), and for histories where a deviation was detected (players then punish the deviator forever).

Formally, each player remembers if a deviation was ever detected, and if so, who the deviator was.

If no deviation was detected before stage n, denote by (b_n^i, b_{n-1}) the signal that player i receives at stage n. If $b_{n-1} \neq a_{n-1}$, let j be the minimal index such that $b_{n-1}^j \neq a_{n-1}^j$, and mark player j as the deviator. Note that all the players observe such a deviation, and, in particular, mark the same player as the deviator (except maybe the deviator himself).

If $b_{n-1} = a_{n-1}$, play a_n^i.

If a deviation of player j was ever detected, then players $N \setminus \{j\}$ play $\tilde{\sigma}_\epsilon^{-j}(h)$, where h is the history of the game up to stage n, using the number z_n^j, as described in Lemma 4.1.

It is clear that no player i can deviate from τ in $G(\mathcal{E})$ and profit more than $U^i(\sigma) + \epsilon$.

The measurability of the device follows from the measurability of σ, $(\tilde{\sigma}_\epsilon^j)_{j \in N}$ and g. ■

Remark: If σ is not correlated, then the device may be independent of previous states. Indeed, the device can choose at each stage n for every player i a uniformly distributed number y_n^i in the unit interval, and send to each player i the pair $(y_n^i, (y_{n-1}^j)_{j \in N})$. Player i then calculates the actions that each player should have played in the previous stage using g, and observes if anyone deviated. He then calculates the action he should play at the current stage, provided no deviation has been detected, and plays it.

Remark: If the state and action spaces are countable, then the space of all histories of length n is countable. Thus, at stage n, the device may calculate a recommended device for *every* possible history of length n, and send each player a vector of signals, one for each such history. The players, who know the realized history, know which signal should be taken into account, and which should be ignored.

Remark: If the transition is deterministic, then the device need not base its choice on previous states, since by knowing the previous state and the recommended actions, it can calculate the new state. If any player deviates (the device then no longer knows the correct state of nature), then this player is punished, and the players ignore the recommendations of the device anyway.

Proof of Theorem 3.9:
Assume now that for every fixed player $i \in N$, v_h^i is independent of h, and denote this value by v^i.

Fix $\epsilon > 0$. We denote by P^i the space of *pure* strategies of player i, and $P = \times_{i \in N} P^i$. Every correlated profile σ induces a probability measure over P. This probability measure is also denoted by σ.

Let σ be a correlated profile such that $U^i(\sigma) < \epsilon$ for each player $i \in N$. Denote by H_∞^δ the set of all histories $h_\infty \in H_\infty$ such that $u^i(h, \sigma) < v_h^i - \delta$ for some beginning h of h_∞.

Since $U^i(\sigma) < \epsilon$ it follows that $\mathbf{P}_{s_\star,\sigma} H_\infty^{\sqrt{\epsilon}} \leq \sqrt{\epsilon}$.

Define a correlation device $\mathcal{E}$ with a signal space $M^i = P$ for each player i. The device chooses a *pure* profile according to σ, and reveals to all the players the profile that was chosen. The players are then requested to follow the pure profile that was chosen by the device. A deviator, who will be noticed immediately, will be punished with his punishment level, which is independent of the history.

With probability greater than $1 - \sqrt{\epsilon}$ no player can profit more than $\sqrt{\epsilon}$; hence this profile is a $(1+R)\sqrt{\epsilon}$-equilibrium in $G(\mathcal{E})$ (recall that R is a bound of u). ■

MEDS Department, Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208, U.S.A.; e-solan@nwu.edu

APPENDIX

Proof of Lemma 4.2:
The proof follows similar lines as the proof of Theorem C, p. 172 in Halmos (1950). We are going to define g as a partition of the unit interval into subsets, which are either single points or sub-intervals.

For every $n \in \mathbf{N}$, let $\{B_i^n\}_i$ be a finite partition of A_0 such that $\{B_i^n\}_{n,i}$ generates the σ-algebra of A_0. Moreover, these sets are chosen such that each set B_i^n is a proper subset of $B_{j(n,i)}^{n-1}$ for some $j(n,i)$, and that if $i_1 \leq i_2$ then $j(n,i_1) \leq j(n,i_2)$.

For every $n \in \mathbf{N}$ and $i \in \{1,\ldots,n_i\}$ define the measurable function $\alpha_i^n : H \to [0,1]$ by:

$$\alpha_i^n(h) = \sum_{j<i} f_h[B_i^n].$$

For every $a \in A$, let $i(a,n)$ be the unique index such that $a \in B_{i(n,a)}^n$. Define now the measurable functions

$$G(h,a) = \lim_{n\to\infty} \alpha_{i(n,a)}^n(h)$$

and

$$K(h,a) = \sup_{b \ \mid \ G(h,b)<G(h,a)} G(h,b).$$

Finally we define $g_h(a)$ to be equal to the interval $[G(h,a),G(h,a)]$ if $G(h,a) = K(h,a)$, to the interval $(K(h,a),G(h,a)]$ if $G(h,a) > K(h,a)$ and $K(h,b) = \max_{b \ \mid \ G(h,b)<G(h,a)}$, and to the closed interval $[K(h,a),G(h,a)]$ otherwise.

It is easy to verify that g is measurable. Moreover

$$\lambda[g_h(B_i^n)] = \alpha_i^n(h) - \alpha_{i-1}^n(h) = f_h[B_i^n].$$

It follows that requirement (ii) is satisfied for every measurable subset C. ■

REFERENCES

Amir R. (1987): "Games of Resource Extraction: Existence of Nash Equilibrium," Cowles Foundation D.P. #825

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

————— (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1-18.

Aumann R.J., and B.M. Maschler B.M. (1995): *Repeated Games with Incomplete Information*, The MIT Press.

Bernheim D., and D. Ray (1987): "Altruistic Growth with Intergenerational Altruism," *Review of Economics Studies*, 54, 227-242

Cyert R. M., and M. H. DeGroot (1970): "Multiperiod Decision Models with Alternating Choice as a Solution to the Duopoly Problem," *Quarterly Journal of Economics*, 84, 410-429

Forges F. (1985): "Correlated Equilibria in a Class of Repeated Games with Incomplete Information," *International Journal of Game Theory*, 14, 129-150.

————— (1986): "An Approach to Communication Equilibria," *Econometrica*, 54, 1375-1385.

————— (1988): "Communication Equilibria in Repeated Games with Incomplete Information," *Mathematics of Operation Research*, 13, V. 2, 77-117.

Fudenberg D., and D. Levine (1983): "Subgame-Perfect Equilibria of Finite- and Infinite-Horizon Games," *Journal of Economic Theory*, 31, 251-268.

Fudenberg D., and J. Tirole (1991): *Game Theory*, The MIT Press.

Halmos P.R. (1950): *Measure Theory*, Van Nostrand Reinhold Company.

Harris C., and J. Vickers (1987): "Racing with Uncertainty," *Review of Economics Studies*, 54, 1-21

Leininger W. (1986): "The Existence of Perfect Equilibria in a Model of Growth with Altruism between Generations," *Review of Economics Studies*, 53, 349-367

Levhari D. and Mirman L. (1980): "The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution," *Bell Journal of Economics*, 11, 322-334

Maitra A., and W. Sudderth (1993): "Borel Stochastic Games with Limsup Payoff," *Annals of Probability*, 21, 861-885.

Majumdar M., and R. Sundaram (1991): "Symmetric Stochastic Games of Resource Extraction: Existence of Nonrandomized Stationary Equilibrium," in *Stochastic Games and Related Topics*, ed. by T.E.S. Raghavan et al., Dordrecht, Kluwer

Maskin E. and Tirole J. (1988a): "A Theory of Dynamic Oligopoly. I. Overview and Quantity Competition with Large Fixed Costs," *Econometrica*, 56, 549-569

——————— (1988b): "A Theory of Dynamic Oligopoly. II. Price Competition, Kinked Demand Curves and Edgeworth Cycles," *Econometrica*, 56, 571-599

Mertens J.F. (1994): "Correlated and Communication Equilibria," in *Game Theoretic Methods in General Equilibrium Analysis*, ed. by J.F. Mertens and S. Sorin, Kluwer Academic Press.

Mertens J.F., and T. Parthasarathy (1987) "Equilibria for Discounted Stochastic Games," CORE Discussion Paper No. 8750.

Mertens J.F., and Sorin S., and S. Zamir (1994): "Repeated Games," CORE

Discussion Paper 9421.

Moulin H., and J.P. Vial (1978): "Strategically Zero-Sum Games: The Class of Games whose Completely Mixed Equilibria Cannot Be Improved Upon," *International Journal of Game Theory* , 7, 201-221.

Myerson R.B. (1986): "Multistage Games with Communication," *Econometrica*, 54, 323-358.

Neveu J. (1965): *Mathematical Foundations of the Calculus of Probability*, Holden-Day.

Nowak A.S. (1991): "Existence of Correlated Weak Equilibria in Discounted Stochastic Games with General State Space," in *Stochastic Games and Related Topics*, ed. by T.E.S. Raghavan et al., Kluwer Academic Press.

Shapley L.S. (1953): "Stochastic Games," *Proceedings of the National Academy of Sciences of U.S.A.*, 39, 1095-1100.

Solan E. (1998): "Discounted Stochastic Games," *Mathematics of Operation Research*, to appear.

Solan E., and N. Vieille (1998a): "Correlated Equilibrium in Stochastic Games," D.P. # 1226, The Center for Mathematical Studies in Economics and Management Science, Northwestern University *and* Cahiers du CEREMADE #9848

Sundaram R. K. (1989): "Perfect Equilibrium in Non-Randomized Strategies in a Class of Stochastic Dynamic Games," *Journal of Economics Theory*, 47, 153-177

Winston W. (1978): "A Stochastic Game Model of a Weapons Development Competition," *Siam Journal of Control and Optimization*, 16, 411-419

stage 1

T	**2**
B	$2, 2$

stage 2

	L	R
T	$0, 0$	$1, 5$
B	$5, 1$	$0, 0$

	stage 1
T	**2**
B	-2

stage 2	
L	R
0	-1

stage 1

	L	C	R
T	$1, 1$	$1, 0$	$0, 2$
B	**2**	$0, 0$	$1, -4$

stage 2

L	R
$3, -1$	$0, -2$

	L	C	R
T	α	γ	ϵ
B	β	δ	ν