

# NON-COOPERATIVE GAMES

MIHAI MANEA

Department of Economics, MIT

## 1. NORMAL-FORM GAMES

A *normal* (or *strategic*) *form game* is a triplet  $(N, S, u)$  with the following properties:

- $N = \{1, 2, \dots, n\}$  is a finite set of *players*
- $S_i \ni s_i$  is the set of *pure strategies* of player  $i$ ;  $S = S_1 \times \dots \times S_n \ni s = (s_1, \dots, s_n)$
- $u_i : S \rightarrow \mathbb{R}$  is the *payoff function* of player  $i$ ;  $u = (u_1, \dots, u_n)$ .

Outcomes are *interdependent*. Player  $i \in N$  receives payoff  $u_i(s_1, \dots, s_n)$  when the *pure strategy profile*  $s = (s_1, \dots, s_n) \in S$  is played. The game is *finite* if  $S$  is finite. We write  $S_{-i} = \prod_{j \neq i} S_j \ni s_{-i}$ .

The structure of the game is *common knowledge*: all players know  $(N, S, u)$ , and know that their opponents know it, and know that their opponents know that they know, and so on.

For any measurable space  $X$  we denote by  $\Delta(X)$  the set of probability measures (or distributions) on  $X$ .<sup>1</sup> A *mixed strategy* for player  $i$  is an element  $\sigma_i$  of  $\Delta(S_i)$ . A *mixed strategy profile*  $\sigma \in \Delta(S_1) \times \dots \times \Delta(S_n)$  specifies a mixed strategy for each player. A *correlated strategy profile*  $\sigma$  is an element of  $\Delta(S)$ . A mixed strategy profile can be seen as a special case of a correlated strategy profile (by taking the product distribution), in which case it is also called *independent* to emphasize the absence of correlation. A *correlated belief* for player  $i$  is an element  $\sigma_{-i}$  of  $\Delta(S_{-i})$ . The set of *independent beliefs* for  $i$  is  $\prod_{j \neq i} \Delta(S_j)$ . It is assumed that player  $i$  has von Neumann-Morgenstern preferences over  $\Delta(S)$  and  $u_i$  extends to  $\Delta(S)$  as follows

$$u_i(\sigma) = \sum_{s \in S} \sigma(s) u_i(s).$$

---

*Date:* January 19, 2017.

These notes benefitted from the proofreading and editing of Gabriel Carroll. The treatment of classic topics follows Fudenberg and Tirole's text "Game Theory" (FT). Some material is borrowed from Muhamet Yildiz.

<sup>1</sup>In most of our applications  $X$  is either finite or a subset of a Euclidean space.

## 2. DOMINATED STRATEGIES

Are there obvious predictions about how a game should be played?

**Example 1** (Prisoners' Dilemma). Two persons are arrested for a crime, but there is not enough evidence to convict either of them. Police would like the accused to testify against each other. The prisoners are put in different cells, with no possibility of communication. Each suspect can stay silent ("cooperate" with his accomplice) or testify against the other ("defect").

- If a suspect testifies against the other and the other does not, the former is released and the latter gets a harsh punishment.
- If both prisoners testify, they share the punishment.
- If neither testifies, both serve time for a smaller offense.

	C	D
C	1, 1	-1, 2
D	2, -1	0, 0*

Note that each prisoner is better off defecting regardless of what the other does. Cooperation is a strictly dominated action for each prisoner. The only outcome if each player privately optimizes is  $(D, D)$ , even though it is Pareto dominated by  $(C, C)$ .

**Example 2.** Consider the game obtained from the prisoners' dilemma by changing player 1's payoff for  $(C, D)$  from  $-1$  to  $1$ . No matter what player 1 does, player 2 still prefers

	C	D
C	1, 1	1, 2*
D	2, -1	0, 0

$D$  to  $C$ . If player 1 knows that 2 never plays  $C$ , then he prefers  $C$  to  $D$ . Unlike in the prisoners' dilemma example, we use an additional assumption to reach our prediction in this case: player 1 needs to deduce that player 2 never plays a dominated strategy.

**Definition 1.** A strategy  $s_i \in S_i$  is *strictly dominated* by  $\sigma_i \in \Delta(S_i)$  if

$$u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i}.$$

**Example 3.** There are situations where a strategy is not strictly dominated by any pure strategy, but is strictly dominated by a mixed one. For instance, in the game below  $B$  is

	L	R
T	3, $x$	0, $x$
M	0, $x$	3, $x$
B	1, $x$	1, $x$

strictly dominated by a 50-50 mix between  $T$  and  $M$ , but not by either  $T$  or  $M$ .

**Example 4** (A Beauty Contest). Consider an  $n$ -player game in which each player announces a number in the set  $\{1, 2, \dots, 100\}$  and a prize of \$1 is split equally between all players whose number is closest to  $2/3$  of the average of all numbers announced. Talk about the Keynesian beauty contest.

We can iteratively eliminate dominated strategies, under the assumption that “I know that you know that I know... that I know the payoffs and that no one would ever use a dominated strategy.

**Definition 2.** For all  $i \in N$ , set  $S_i^0 = S_i$  and define  $S_i^k$  recursively by

$$S_i^k = \{s_i \in S_i^{k-1} \mid \nexists \sigma_i \in \Delta(S_i^{k-1}), u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i}^{k-1}\}.$$

The set of pure strategies of player  $i$  that survive *iterated deletion of strictly dominated strategies* is  $S_i^\infty = \bigcap_{k \geq 0} S_i^k$ . The set of surviving mixed strategies is

$$\{\sigma_i \in \Delta(S_i^\infty) \mid \nexists \sigma'_i \in \Delta(S_i^\infty), u_i(\sigma'_i, s_{-i}) > u_i(\sigma_i, s_{-i}), \forall s_{-i} \in S_{-i}^\infty\}.$$

**Remark 1.** In a finite game the elimination procedure ends in a finite number of steps, so  $S^\infty$  is simply the set of surviving strategies at the last stage.

**Remark 2.** In an infinite game, if  $S$  is a compact metric space and  $u$  is continuous, then one can use Cantor’s theorem (a decreasing nested sequence of non-empty compact sets has nonempty intersection) to show that  $S^\infty \neq \emptyset$ .

**Remark 3.** The definition above assumes that at each iteration all dominated strategies of each player are deleted simultaneously. Clearly, there are many other iterative procedures

that can be used to eliminate strictly dominated strategies. However, the limit set  $S^\infty$  does not depend on the particular way deletion proceeds.<sup>2</sup> The intuition is that a strategy which is dominated at some stage is dominated at any later stage.

**Remark 4.** The outcome does not change if we eliminate strictly dominated mixed strategies at every step. The reason is that a strategy is dominated against all pure strategies of the opponents if and only if it is dominated against all their mixed strategies. Eliminating mixed strategies for player  $i$  at any stage does not affect the set of strictly dominated pure strategies for any player  $j \neq i$  at the next stage.

**2.1. Detour on common knowledge.** Common knowledge looks like an innocuous assumption, but may have strong consequences in some situations. Consider the following story. Once upon a time, there was a village with 100 married couples. The women had to pass a logic exam before being allowed to marry; thus all married women were perfect reasoners. The high priestess was not required to take that exam, but it was common knowledge that she was truthful. The village was small, so everyone would be able to hear any shot fired in the village. The women would gossip about adulterous relationships and each knew which of the other women's husbands were unfaithful. However, no one would ever inform a wife about her own cheating husband.

The high priestess knew that some husbands were unfaithful, and one day she decided that such immorality should not be tolerated any further. This was a successful religion and all women agreed with the views of the priestess.

The priestess convened all the women at the temple and *publicly announced* that the well-being of the village had been compromised—there was at least one cheating husband. She also pointed out that even though none of them knew whether her husband was faithful, each woman knew about the other unfaithful husbands. She ordered each woman to shoot her husband on the midnight of the day she was certain of his infidelity. 39 silent nights went by and on the 40<sup>th</sup> shots were heard. How many husbands were shot? Were all the unfaithful husbands caught? How did some wives learn of their husbands' infidelity after 39 nights in which *nothing* happened?

---

<sup>2</sup>This property does not hold for *weakly* dominated strategies.

Since the priestess was truthful, there must have been at least one unfaithful husband in the village. How would events have unfolded if there was exactly one unfaithful husband? His wife, upon hearing the priestess' statement and realizing that she does not know of any unfaithful husband, would have concluded that her own marriage must be the only adulterous one and would have shot her husband on the midnight of the first day. Clearly, there must have been more than one unfaithful husband. If there had been exactly two unfaithful husbands, then each of the two cheated wives would have initially known of exactly one unfaithful husband, and after the first silent night would infer that there were exactly two cheaters and her husband is one of them. (Recall that the wives were all perfect logicians.) The unfaithful husbands would thus both be shot on the second night. As no shots were heard on the first two nights, all women concluded that there were at least three cheating husbands. . . Since shootings were heard on the 40<sup>th</sup> night, it must be that exactly 40 husbands were unfaithful and they were all exposed and killed simultaneously.

### 3. RATIONALIZABILITY

Rationalizability is a solution concept introduced independently by Bernheim (1984) and Pearce (1984). Like iterated strict dominance, rationalizability derives restrictions on play from common knowledge of the payoffs and of the fact that players are "reasonable" in a certain way. Dominance: it is not reasonable to use a strategy that is strictly dominated. Rationalizability: it is not rational for a player to choose a strategy that is not a best response to some beliefs about his opponents' strategies.

What is a "belief"? In Bernheim (1984) and Pearce (1984) each player  $i$ 's beliefs  $\sigma_{-i}$  about the play of  $j \neq i$  must be independent, i.e.,  $\sigma_{-i} \in \prod_{j \neq i} \Delta(S_j)$ . Alternatively, we may allow player  $i$  to believe that the actions of his opponents are correlated, i.e., any  $\sigma_{-i} \in \Delta(S_{-i})$  is a possibility. The two definitions have different implications for  $n \geq 3$ . We focus on the case with correlated beliefs. It should be emphasized that such beliefs represent a player's uncertainty about his opponents' actions and not his theory about their deliberate randomization and coordination. For instance,  $i$  may place equal probability on two scenarios: either both  $j$  and  $k$  pick action  $A$  or they both play  $B$ . If  $i$  is not sure which theory is true, then his beliefs are correlated even though he knows that  $j$  and  $k$  are acting independently.

**Definition 3.** A strategy  $\sigma_i \in S_i$  is a *best response* to a belief  $\sigma_{-i} \in \Delta(S_{-i})$  if

$$u_i(\sigma_i, \sigma_{-i}) \geq u_i(s_i, \sigma_{-i}), \forall s_i \in S_i.$$

We can again iteratively develop restrictions imposed by common knowledge of the payoffs and rationality to obtain the definition of rationalizability.

**Definition 4.** Set  $S^0 = S$  and let  $S^k$  be given recursively by

$$S_i^k = \{s_i \in S_i^{k-1} | \exists \sigma_{-i} \in \Delta(S_{-i}^{k-1}), u_i(s_i, \sigma_{-i}) \geq u_i(s'_i, \sigma_{-i}), \forall s'_i \in S_i^{k-1}\}.$$

The set of *correlated rationalizable strategies* for player  $i$  is  $S_i^\infty = \bigcap_{k \geq 0} S_i^k$ . A mixed strategy  $\sigma_i \in \Delta(S_i)$  is rationalizable if there is a belief  $\sigma_{-i} \in \Delta(S_{-i}^\infty)$  s.t.  $u_i(\sigma_i, \sigma_{-i}) \geq u_i(s_i, \sigma_{-i})$  for all  $s_i \in S_i^\infty$ .

The definition of *independent rationalizability* replaces  $\Delta(S_{-i}^{k-1})$  and  $\Delta(S_{-i}^\infty)$  above with  $\prod_{j \neq i} \Delta(S_j^{k-1})$  and  $\prod_{j \neq i} \Delta(S_j^\infty)$ , respectively.

**Example 5** (Rationalizability in Cournot duopoly). Two firms compete on the market for a divisible homogeneous good. Each firm  $i = 1, 2$  has zero marginal cost and simultaneously decides to produce an amount of output  $q_i \geq 0$ . The resulting price is  $p = 1 - q_1 - q_2$ . Hence the profit of firm  $i$  is given by  $q_i(1 - q_1 - q_2)$ . The best response correspondence of firm  $i$  is  $B_i(q_j) = \max(0, (1 - q_j)/2)$  ( $j = 3 - i$ ). If  $i$  knows that  $q_j \leq q$  then  $B_i(q_j) \geq (1 - q)/2$ .

We know that  $q_i \geq q^0 = 0$  for  $i = 1, 2$ . Hence  $q_i \leq q^1 = B_i(q^0) = (1 - q^0)/2$  and  $S_i^1 = [0, q^1]$  for all  $i$ . But then  $q_i \geq q^2 = B_i(q^1) = (1 - q^1)/2$  and  $S_i^2 = [q^2, q^1]$  for all  $i$ . We obtain a sequence

$$q^0 \leq q^2 \leq \dots \leq q^{2k} \leq \dots \leq q_i \leq \dots \leq q^{2k+1} \leq \dots \leq q^1,$$

where  $q^{2k} = \sum_{l=1}^k 1/4^l = (1 - 1/4^k)/3$  and  $q^{2k+1} = (1 - q^{2k})/2$  for all  $k \geq 0$  such that  $S_i^k = [q^{k-1}, q^k]$  for  $k$  odd and  $S_i^k = [q^k, q^{k-1}]$  for  $k$  even. Clearly,  $\lim_{k \rightarrow \infty} q^k = 1/3$ , hence the only rationalizable strategy for firm  $i$  is  $q_i = 1/3$ . This is also the unique Nash equilibrium, which we define next. What are the rationalizable strategies when there are more than two firms?

We say that a strategy  $\sigma_i$  is *never a best response* for player  $i$  if it is not a best response to any  $\sigma_{-i} \in \Delta(S_{-i})$ . Recall that a strategy  $\sigma_i$  of player  $i$  is *strictly dominated* if there exists  $\sigma'_i \in \Delta(S_i)$  s.t.  $u_i(\sigma'_i, s_{-i}) > u_i(\sigma_i, s_{-i}), \forall s_{-i} \in S_{-i}$ .

**Theorem 1.** *In a finite game, a strategy is never a best response if and only if it is strictly dominated.*

*Proof.* Clearly, a strategy  $\sigma_i$  strictly dominated for player  $i$  by some  $\sigma'_i$  cannot be a best response for any belief  $\sigma_{-i} \in \Delta(S_{-i})$  as  $\sigma'_i$  yields a strictly higher payoff than  $\sigma_i$  against any such  $\sigma_{-i}$ .

We are left to show that a strategy which is never a best response must be strictly dominated. We prove that any strategy  $\tilde{\sigma}_i$  of player  $i$  which is not strictly dominated must be a best response for some beliefs. Define the set of “dominated payoffs” for  $i$  by

$$D = \{x \in \mathbb{R}^{S-i} \mid \exists \sigma_i \in \Delta(S_i), x \leq u_i(\sigma_i, \cdot)\}.$$

Clearly  $D$  is non-empty, closed and convex. Also,  $u_i(\tilde{\sigma}_i, \cdot)$  does not belong to the interior of  $D$  because it is not strictly dominated by any  $\sigma_i \in \Delta(S_i)$ . By the supporting hyperplane theorem, there exists  $\alpha \in \mathbb{R}^{S-i}$  different from the zero vector s.t.  $\alpha \cdot u_i(\tilde{\sigma}_i, \cdot) \geq \alpha \cdot x, \forall x \in D$ . In particular,  $\alpha \cdot u_i(\tilde{\sigma}_i, \cdot) \geq \alpha \cdot u_i(\sigma_i, \cdot), \forall \sigma_i \in \Delta(S_i)$ . Since  $D$  is not bounded from below, each component of  $\alpha$  needs to be non-negative. We can normalize  $\alpha$  so that its components sum to 1, in which case it can be interpreted as a belief in  $\Delta(S_{-i})$  with the property that  $u_i(\tilde{\sigma}_i, \alpha) \geq u_i(\sigma_i, \alpha), \forall \sigma_i \in \Delta(S_i)$ . Thus  $\tilde{\sigma}_i$  is a best response to  $\alpha$ .  $\square$

**Corollary 1.** *Correlated rationalizability and iterated strict dominance coincide.*

**Theorem 2.** *For every  $k \geq 0$ , each  $s_i \in S_i^k$  is a best response (within  $S_i$ ) to a belief in  $\Delta(S_{-i}^{k-1})$ .*

*Proof.* Fix  $s_i \in S_i^k$ . We know that  $s_i$  is a best response within  $S_i^{k-1}$  to some  $\sigma_{-i} \in \Delta(S_{-i}^{k-1})$ . If  $s_i$  was not a best response within  $S_i$  to  $\sigma_{-i}$ , let  $s'_i$  be such a best response. Since  $s_i$  is a best response within  $S_i^{k-1}$  to  $\sigma_{-i}$ , and  $s'_i$  is a strictly better response than  $s_i$  to  $\sigma_{-i}$ , we need  $s'_i \notin S_i^{k-1}$ . Then  $s'_i$  was deleted at some step of the iteration, say  $s'_i \in S_i^{l-1}$  but  $s'_i \notin S_i^l$  for some  $l \leq k-1$ . This contradicts the fact that  $s'_i$  is a best response in  $S_i^{l-1}$  to  $\sigma_{-i}$ , which belongs to  $\Delta(S_{-i}^{k-1}) \subseteq \Delta(S_{-i}^{l-1})$ .  $\square$

**Corollary 2.** *If the game is finite, then each  $s_i \in S_i^\infty$  is a best response (within  $S_i$ ) to a belief in  $\Delta(S_{-i}^\infty)$ .*

**Definition 5.** A set  $Z = Z_1 \times \dots \times Z_n$  with  $Z_i \subseteq S_i$  for  $i \in N$  is *closed under rational behavior* if, for all  $i$ , every strategy in  $Z_i$  is a best response to a belief in  $\Delta(Z_{-i})$ .

**Theorem 3.** *If the game is finite (or if  $S$  is a compact metric space and  $u$  is continuous), then  $S^\infty$  is the largest set closed under rational behavior.*

*Proof.* Clearly,  $S^\infty$  is closed under rational behavior by Corollary ???. Suppose that there exists  $Z_1 \times \dots \times Z_n \not\subseteq S^\infty$  that is closed under rational behavior. Consider the smallest  $k$  for which there is an  $i$  such that  $Z_i \not\subseteq S_i^k$ . It must be that  $k \geq 1$  and  $Z_{-i} \subset S_{-i}^{k-1}$ . By assumption, every element in  $Z_i$  is a best response to an element of  $\Delta(Z_{-i}) \subset \Delta(S_{-i}^{k-1})$ , contradicting  $Z_i \not\subseteq S_i^k$ .  $\square$

Rationalizability has strong epistemic foundations—it characterizes the strategic implications of common knowledge of rationality (see next section). As we will see later, it also has some evolutionary foundations. In any adaptive process the proportion of players who play a non-rationalizable strategy vanishes as the system evolves.

#### 4. COMMON KNOWLEDGE OF RATIONALITY AND RATIONALIZABILITY

We now formalize the idea of common knowledge and show that rationalizability captures the idea of common knowledge of rationality (and payoffs) precisely.<sup>3</sup> We first introduce the notion of an incomplete-information epistemic model.

**Definition 6.** (Information Structure) An information (or belief) structure is a list  $(\Omega, (I_i)_{i \in N}, (p_i)_{i \in N})$  where

- $\Omega$  is a finite state space;
- $I_i : \Omega \rightarrow 2^\Omega$  is a partition of  $\Omega$  for each  $i \in N$  such that  $I_i(\omega)$  is the set of states that  $i$  thinks are possible when the true state is  $\omega$ ; it assumed that  $\omega' \in I_i(\omega) \Leftrightarrow \omega \in I_i(\omega')$ ;
- $p_{i, I_i(\omega)}$  is a probability distribution on  $I_i(\omega)$  representing  $i$ 's belief at  $\omega$ .

The state  $\omega$  summarizes all the relevant facts about the world. Note that only one of the state is the true state of the world; all others are hypothetical states needed to encode players' beliefs. In state  $\omega$ , player  $i$  is informed that the state is in  $I_i(\omega)$  and gets no other information. Such an information structure arises if each player observes a state-dependent

<sup>3</sup>This section builds of notes by Muhamet Yildiz.

signal, where  $I_i(\omega)$  is the set of states for which player  $i$ 's signal is identical to the signal at state  $\omega$ . The next definition formalizes the idea that  $I_i$  summarizes all of the information of  $i$ .

**Definition 7.** For any event  $F \subseteq \Omega$ , player  $i$  *knows* at  $\omega$  that  $F$  obtains if  $I_i(\omega) \subseteq F$ . The event that  $i$  *knows*  $F$  is

$$K_i(F) = \{\omega | I_i(\omega) \subseteq F\}.$$

The event that *everyone knows*  $F$  is defined by

$$K(F) = \bigcap_{i \in N} K_i(F).$$

Let  $K^0(F) = F$  and  $K^{t+1}(F) = K(K^t(F))$  for  $t \geq 0$ . Set  $K^\infty(F) = \bigcap_{t \geq 0} K^t(F)$ .  $K^\infty(F)$  is the set of states where  $F$  is *common knowledge*.

Note that  $K(K^\infty(F)) = K^\infty(F)$ . This leads to an alternative definition of common knowledge. An event  $F'$  is *public* if  $F' = \bigcup_{\omega' \in F'} I_i(\omega')$  for all  $i$ , which is equivalent to  $K(F') = F'$  (and  $K^\infty(F') = F'$ ). Then an event  $F$  is common knowledge at  $\omega$  if and only if there exists a public event  $F'$  with  $\omega \in F' \subseteq F$ .

We have so far considered an abstract information structure for the players in  $N$ . Fix a game  $(N, S, u)$ . In order to give strategic meaning to the states, we also need to describe what players play at each state by introducing a strategy profile  $s : \Omega \rightarrow S$ .

**Definition 8.** A strategy profile  $s : \Omega \rightarrow S$  is adapted with respect to  $(\Omega, (I_i)_{i \in N}, (p_i)_{i \in N})$  if  $s_i(\omega) = s_i(\omega')$  whenever  $I_i(\omega) = I_i(\omega')$ .

Players must choose a constant action at all states in each information set since they cannot distinguish between states in the same information set.

**Definition 9.** An *epistemic model*  $(\Omega, (I_i)_{i \in N}, (p_i)_{i \in N}, s)$  consists of an information structure and an adapted strategy profile.

The ideas of rationality and common knowledge of rationality can be formalized as follows.

**Definition 10.** For any epistemic model  $(\Omega, (I_i)_{i \in N}, (p_i)_{i \in N}, s)$  and any  $\omega \in \Omega$ , a player  $i$  is said to be rational at  $\omega$  if

$$s_i(\omega) \in \arg \max_{s_i \in S_i} \sum_{\omega' \in I_i(\omega)} u_i(s_i, s_{-i}(\omega')) p_{i, I_i(\omega)}(\omega').$$

**Definition 11.** A strategy  $s_i \in S_i$  consistent with *common knowledge of rationality* if there exists a model  $(\Omega, (I_j)_{j \in N}, (p_j)_{j \in N}, s)$  and state  $\omega^* \in \Omega$  with  $s_i(\omega^*) = s_i$  at which it is common knowledge that all players are rational (i.e., the event  $R := \{\omega \in \Omega \mid \text{every player } i \in N \text{ is rational at } \omega\}$  is common knowledge at  $\omega^*$ ).

Given the alternative definition of common knowledge in terms of public events,  $s_i \in S_i$  consistent with common knowledge of rationality if there exists an epistemic model  $(\Omega', (I_j)_{j \in N}, (p_j)_{j \in N}, s)$  such that  $s_j(\omega)$  is a best response to  $s_{-j}$  at each  $\omega \in \Omega$  for every player  $j \in N$  (simply consider the restriction of the original model to  $\Omega' = K^\infty(R)$ ). The next result states that rationalizability is equivalent to common knowledge of rationality in the sense that  $S_i^\infty$  is the set of strategies that are consistent with common knowledge of rationality.

**Theorem 4.** *For any  $i \in N$  and  $s_i \in S_i$ , the strategy  $s_i$  is consistent with common knowledge of rationality if and only if  $s_i$  is rationalizable, i.e.,  $s_i \in S_i^\infty$ .*

*Proof.* ( $\Rightarrow$ ) First, take any  $s_i$  that is consistent with common knowledge of rationality. Then there exists a model  $(\Omega, (I_j)_{j \in N}, (p_j)_{j \in N}, s)$  with a state  $\omega^* \in \Omega$  such that  $s_i(\omega^*) = s_i$  and for each  $j$  and  $\omega$ ,

$$(4.1) \quad s_j(\omega) \in \arg \max_{s_j \in S_j} \sum_{\omega' \in I_j(\omega)} u_j(s_j, s_{-j}(\omega')) p_{j, I_j(\omega)}(\omega')$$

Define  $Z_j = s_j(\Omega)$ . Note that  $s_i \in Z_i$ . By Theorem ??, in order to show that  $s_i \in S_i^\infty$ , it suffices to show that  $Z$  is closed under rational behavior. Since for each  $z_j \in Z_j$ , there exists  $\omega \in \Omega$  such that  $z_j = s_j(\omega)$ , define belief  $\mu_{j, \omega}$  on  $Z_{-j}$  by setting

$$\mu_{j, \omega}(s_{-j}) = \sum_{\omega' \in I_j(\omega), s_{-j}(\omega') = s_{-j}} p_{j, I_j(\omega)}(\omega')$$

Then, by (??),

$$\begin{aligned} z_j &= s_j(\omega) \in \arg \max_{s_j \in S_j} \sum_{\omega' \in I_j(\omega)} u_j(s_j, s_{-j}(\omega')) p_{j, I_j(\omega)}(\omega') \\ &= \arg \max_{s_j \in S_j} \sum_{s_{-j} \in Z_{-j}} \mu_{j, \omega}(s_{-j}) u_j(s_j, s_{-j}), \end{aligned}$$

which shows that  $Z$  is closed under rational behavior.

( $\Leftarrow$ ) Conversely, since  $S^\infty$  is closed under rational behavior, for every  $s_i \in S_i^\infty$ , there exists a probability distribution  $\mu_{i,s_i}$  on  $S_{-i}^\infty$  against which  $s_i$  is a best response. Define the model  $(S^\infty, (I_i)_{i \in N}, (p_i)_{i \in N}, s)$  with

$$\begin{aligned} I_i(s) &= \{s_i\} \times S_{-i}^\infty \\ p_{i,s}(s') &= \mu_{i,s_i}(s'_{-i}) \\ s(s) &= s. \end{aligned}$$

In this model it is common knowledge that every player is rational. Indeed, for all  $s \in S^\infty$ ,

$$s_i(s) = s_i \in \arg \max_{s'_i \in S_i} \sum_{s_{-i} \in S_{-i}^\infty} u_i(s'_i, s_{-i}) \mu_{i,s}(s'_{-i}) = \arg \max_{s'_i \in S_i} \sum_{s' \in I_i(s)} u_i(s'_i, s_{-i}) p_{i,s}(s').$$

For every  $s_i \in S_i^\infty$ , there exists  $s = (s_i, s_{-i}) \in S^\infty$  such that  $s_i(s) = s_i$ , showing that  $s_i$  is consistent with common knowledge of rationality.  $\square$

## 5. NASH EQUILIBRIUM

Many games are not solvable by iterated strict dominance or rationalizability. The concept

	H	T		L	R		T	S
H	1, -1	-1, 1	L	1, 1	0, 0	T	3, 2	1, 1
T	-1, 1	1, -1	R	0, 0	1, 1	S	0, 0	2, 3

FIGURE 1. Matching Pennies, Coordination Game, Battle of the Sexes

of Nash (1950) equilibrium has more bite in some situations. The idea of Nash equilibrium was implicit in the particular examples of Cournot (1838) and Bertrand (1883) at an informal level.

**Definition 12.** A mixed-strategy profile  $\sigma^*$  is a Nash equilibrium if for each  $i \in N$

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(s_i, \sigma_{-i}^*), \forall s_i \in S_i.$$

Note that if a player uses a nondegenerate mixed strategy in a Nash equilibrium (one that places positive probability weight on more than one pure strategy) then he must be indifferent between all pure strategies in the support. Of course, the fact that there is no profitable deviation in pure strategies implies that there is no profitable deviation in mixed strategies either.

**Example 6** (Matching Pennies). This simple game shows that there may sometimes not be any equilibria in pure strategies. We will establish that equilibria in mixed strategies exist

	H	T
H	1, -1	-1, 1
T	-1, 1	1, -1

for any finite game.

**Example 7** (Partially Mixed Nash Equilibria). In these  $3 \times 3$  examples, we see that mixed strategy Nash equilibria may only put positive probability on some actions. The first matrix

	F	C	B
F	0, 5	2, 3	2, 3
C	2, 3	0, 5	3, 2
B	5, 0	3, 2	2, 3

represents a tennis service game, where player 1 chooses whether to serve to player 2's forehand, center or backhand side; player 2 similarly chooses which side to favor for the return. The game has a unique mixed strategy equilibrium, which puts positive probability only on strategies  $C$  and  $B$  for either player. Note first that choosing  $C$  with probability  $\epsilon$  and  $B$  with probability  $1 - \epsilon$  (for small  $\epsilon > 0$ ) strictly dominates  $F$  for player 1. If player 1 never chooses  $F$ , then  $C$  strictly dominates  $F$  for player 2. In the resulting  $2 \times 2$  game, there is a unique equilibrium, in which both players place probability  $1/4$  on  $C$  and  $3/4$  on  $B$ .

	H	T	C
H	1, -1	-1, 1	-1, -1
T	-1, 1	1, -1	-1, -1
C	-1, -1	-1, -1	3, 3

The second game is matching pennies with a third option: players may choose heads or tails as before, or they may cooperate. Cooperation produces the best outcome, but it is only worth it if both players choose it. The game has a total of 3 equilibria: a single

pure strategy equilibrium  $(C, C)$ , where players cooperate and ignore the matching pennies game; a partially mixed equilibrium  $((1/2, 1/2, 0), (1/2, 1/2, 0))$  where players play the matching pennies game and ignore the option of cooperating; and a totally mixed equilibrium  $((2/5, 2/5, 1/5), (2/5, 2/5, 1/5))$ .

To show that these are the only equilibria, we can proceed as follows: first, if player 1 is mixing between  $H$ ,  $T$  and  $C$ , he must be indifferent among all three actions, which implies that player 2 is also mixing between  $H$ ,  $T$  and  $C$ ; then we can calculate the equilibrium probabilities for the totally mixed equilibrium. If 1 is mixing between  $H$  and  $T$  (but not  $C$ ) then 2 must be mixing between  $H$  and  $T$  for this to be optimal, and 2 will never want to play  $C$  since 1 never does. This leads to the partially mixed equilibrium. If 1 mixes between  $H$  and  $C$  (but not  $T$ ), then 2 may only play  $T$  and  $C$ , but then 1 will never want to play  $H$ , a contradiction; so there are no equilibria of this form (the case where 1 mixes between  $T$  and  $C$  is analogous). Finally we check that the only pure equilibrium is  $(C, C)$ .

**Example 8** (Stag Hunt). This example shows the difficulty of predicting the outcome in games with multiple equilibria. In the stag hunt game, each player can choose to hunt hare by himself or hunt stag with the other player. Stag offers a higher payoff, but only if the players team up. The game has two pure strategy Nash equilibria,  $(S, S)$  and  $(H, H)$ . How

	S	H
S	9, 9	0, 8
H	8, 0	7, 7

should the hunters play? We may expect  $(S, S)$  to be played because it is Pareto dominant, that is, it is better for both players to coordinate on hunting stag. However, if one player expects the other to hunt hare, he is much better off hunting hare himself; and the potential downside of choosing stag is bigger than the upside. Thus, hare is the safer choice. In the language of Harsanyi and Selten (1988),  $H$  is the risk-dominant action: formally, if each player expects the other to play either action with probability  $1/2$ , then  $H$  has a higher expected payoff (7.5) than  $S$  (4.5). In fact, for a player to choose stag, he should expect the other player to play stag with probability at least  $7/8$ . Note that this coordination problem may persist even if players can communicate: regardless of what  $i$  intends to do, he would prefer  $j$  to play stag, so attempts to convince  $j$  to play stag may be cheap talk.

Nash equilibria are “consistent” predictions of how the game will be played—if all players expect that a specific Nash equilibrium will arise then no player has incentives to play differently. Each player must have a correct “conjecture” about the strategies of his opponents and play a best response to his conjecture.

Formally, Aumann and Brandenburger (1995) provide a framework that can be used to examine the epistemic foundations of Nash equilibrium. The primitive of their model is an *interactive belief system* in which there is a possible set of types for each player; each type has associated to it a payoff for every action profile, a choice of which action to play, and a belief about the types of the other players. Aumann and Brandenburger show that in a 2-player game, if the game being played (i.e., both payoff functions), the rationality of the players, and their conjectures are all *mutually known*, then the conjectures constitute a (mixed strategy) Nash equilibrium. Thus *common knowledge* plays no role in the 2-player case. However, for games with more than 2 players, we need to assume additionally that players have a *common prior* and that conjectures are *commonly known*. This ensures that any two players have identical and separable (i.e., independent) conjectures about other players, consistent with a (common) mixed strategy profile.

It is easy to show that every Nash equilibrium is rationalizable (e.g., by applying Theorem ?? to the strategies played with positive probability). The converse is not true. For example, in the battle of the sexes ( $S, T$ ) is not a Nash equilibrium, but both  $S$  and  $T$  are rationalizable for either player. Of course, these strategies correspond to some Nash equilibria, but one can easily construct a game in which some rationalizable strategies do not correspond to any Nash equilibrium.

So far, we have motivated our solution concepts by presuming that players make predictions about their opponents’ play by *introspection* and *deduction*, using knowledge of their opponents’ payoffs, knowledge that the opponents are rational, knowledge about this knowledge. . . . Alternatively, we may assume that players *extrapolate* from past observations of play in “similar” games, with either current opponents or “similar” ones. They form expectations about future play based on past observations and adjust their actions to maximize their current payoffs with respect to these expectations.

The idea of using adjustment processes to model *learning* originates with Cournot (1838). He considered the game in Example ??, and suggested that players take turns setting their

outputs, each player choosing a best response to the opponent's last-period action. Alternatively, we can assume simultaneous belief updating, best responding to sample average play, populations of players being anonymously matched, etc. In the latter context, mixed strategies can also be interpreted as the proportion of players playing various strategies. If the process *converges* to a particular steady state, then the steady state is a Nash equilibrium.

While convergence occurs in Example ??, this is not always the case. How sensitive is the convergence to the initial state? If convergence obtains for all initial strategy profiles sufficiently close to the steady state, we say that the steady state is asymptotically stable. See Figure ?? (FT, pp. 24-26). The Shapley (1964) cycling example from Figure ?? is also interesting.

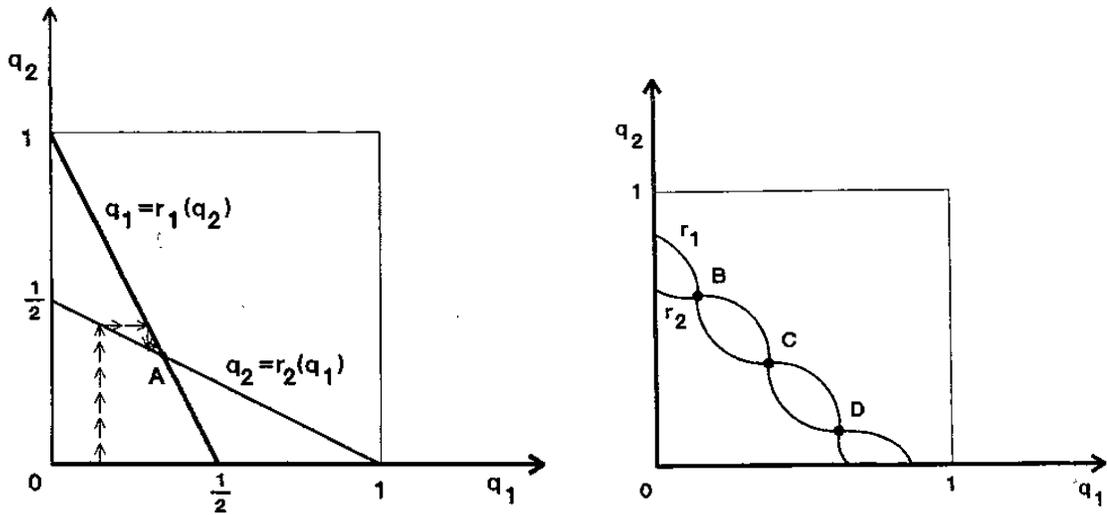


FIGURE 2

Courtesy of The MIT Press. Used with permission.

FIGURE 3

	L	M	R
U	0, 0	4, 5	5, 4
M	5, 4	0, 0	4, 5
D	4, 5	5, 4	0, 0

However, adjustment processes are *myopic* and do not offer a compelling description of behavior. Such processes definitely do not provide good predictions for behavior in the

actual *repeated game*, if players care about play in future periods and realize that their current actions can affect opponents' future play.

## 6. EXISTENCE AND CONTINUITY OF NASH EQUILIBRIA

We can show that a Nash equilibrium exists under broad regularity conditions on strategy spaces and payoff functions.<sup>4</sup> Some continuity and compactness assumptions are indispensable because they are usually needed for the existence of solutions to (single agent) optimization problems. Convexity is usually required for fixed-point theorems, such as Kakutani's.<sup>5</sup> Nash used Kakutani's fixed point theorem to show the existence of mixed strategy equilibria in finite games. We provide a generalization of his existence result. We start with some mathematical background.

**6.1. Topology Prerequisites.** Consider two topological vector spaces  $X$  and  $Y$ . A *correspondence*  $F : X \rightrightarrows Y$  is a set valued function taking elements  $x \in X$  into subsets  $F(x) \subseteq Y$ . The *graph* of  $F$  is defined by  $G(F) = \{(x, y) \mid y \in F(x)\}$ . A point  $x \in X$  is a *fixed point* of  $F$  if  $x \in F(x)$ . A correspondence  $F$  is *non-empty/closed-valued/convex-valued* if  $F(x)$  is non-empty/closed/convex for all  $x \in X$ .

The main continuity notion for correspondences we rely on is the following. A correspondence  $F$  has *closed graph* if  $G(F)$  is a closed subset of  $X \times Y$ . If  $X$  and  $Y$  are first-countable spaces (such as metric spaces), then  $F$  has closed graph if and only if for any sequence  $(x_m, y_m)_{m \geq 0}$  with  $y_m \in F(x_m)$  for all  $m \geq 0$ , which converges to a pair  $(x, y)$ , we have  $y \in F(x)$ . Note that correspondences with closed graph are closed-valued. The converse is false.

A related continuity concept is defined as follows. A correspondence  $F$  is *upper hemicontinuous* at  $x \in X$  if for every open neighborhood  $V_Y$  of  $F(x)$ , there exists a neighborhood  $V_X$  of  $x$  such that  $x' \in V_X \Rightarrow F(x') \subset V_Y$ . In general, closed graph and upper hemicontinuity may have different implications. For instance, the constant correspondence  $F : [0, 1] \rightrightarrows [0, 1]$  defined by  $F(x) = (0, 1)$  is upper hemicontinuous, but does not have a closed graph. However, the two concepts coincide for closed-valued correspondences in most spaces of interest.

<sup>4</sup>This presentation builds on lecture notes by Muhamet Yildiz.

<sup>5</sup>However, there are algebraic fixed point theorems that do not require convexity. We rely on such a result due to Tarski later in the course.

**Theorem 5** (Closed Graph Theorem). *A correspondence  $F : X \rightrightarrows Y$  with compact Hausdorff range  $Y$  is closed if and only if it is upper hemicontinuous and closed-valued.*

Another continuity property is *lower hemicontinuity*, which for compact metric spaces requires that for any sequence  $(x_m) \rightarrow x$  and for any  $y \in F(x)$ , there exists a sequence  $(y_m)$  with  $y_m \in F(x_m)$  for each  $m$  such that  $y_m \rightarrow y$ . In general, solution concepts in game theory are upper hemicontinuous but not lower hemicontinuous, a property inherited from optimization problems.

The maximum theorem states that in single agent optimization problems the optimal solution correspondence is upper hemicontinuous in parameters when the objective function and the domain of optimization vary continuously in all relevant parameters.

**Theorem 6** (Berge's Maximum Theorem). *Suppose that  $f : X \times Y \rightarrow \mathbb{R}$  is a continuous function, where  $X$  and  $Y$  are metric spaces and  $Y$  is compact.*

(1) *The function  $M : X \rightarrow \mathbb{R}$ , defined by*

$$M(x) = \max_{y \in Y} f(x, y),$$

*is continuous.*

(2) *The correspondence  $F : X \rightrightarrows Y$ ,*

$$F(x) = \arg \max_{y \in Y} f(x, y)$$

*is nonempty valued and has a closed graph.*

We lastly state the fixed point result.

**Theorem 7** (Kakutani's Fixed-Point Theorem). *Let  $X$  be a non-empty, compact, and convex subset of a Euclidean space and let the correspondence  $F : X \rightrightarrows X$  have closed graph and non-empty convex values. Then the set of fixed points of  $F$  is non-empty and compact.*

In game theoretic applications of Kakutani's theorem,  $X$  is usually the strategy space, assumed to be compact and convex when we include mixed strategies.<sup>6</sup>  $F$  is typically the best response correspondence, which is non-empty valued and has a closed graph by the

---

<sup>6</sup>We will see other applications of Kakutani's fixed point theorem and its extension to infinite dimensional spaces when we discuss my work on bargaining in dynamic markets.

Maximum Theorem. In that case, we can ensure that  $F$  is convex-valued by assuming that the payoff functions are quasi-concave.

Recall that a function  $f : X \rightarrow \mathbb{R}$  is *quasi-concave* when  $X$  is a convex subset of a real vector space if

$$f(tx + (1 - t)y) \geq \min(f(x), f(y)), \forall t \in [0, 1], x, y \in X.$$

In particular, quasi-concavity implies convex upper contour sets and convex arg max.

## 6.2. Existence of Nash Equilibrium.

**Theorem 8.** *Consider a game  $(N, S, u)$  such that  $S_i$  is a convex and compact subset of a Euclidean space and that  $u_i$  is continuous in  $s$  and quasi-concave in  $s_i$  for all  $i \in N$ . Then there exists a pure strategy Nash equilibrium.*

*Proof.* We construct a correspondence  $F : S \rightrightarrows S$  that satisfies the conditions of Kakutani's Fixed Point Theorem, whose fixed points constitute Nash equilibria.

Let  $F : S \rightrightarrows S$  be the *best response* correspondence defined by

$$F(s) = \{(s_1^*, \dots, s_n^*) \mid s_i^* \in B_i(s_{-i}), \forall i \in N\} = \prod_{i \in N} B_i(s_{-i}), \forall s \in S,$$

where  $B_i(s_{-i}) := \arg \max_{s'_i \in S_i} u_i(s'_i, s_{-i})$ .

Since  $S$  is compact and the utility functions are continuous, the Maximum Theorem implies that  $F$  is non-empty valued and has closed graph. Moreover, since  $u_i$  is quasi-concave in  $s_i$ , the set  $B_i(s_{-i})$  is convex for all  $i$  and  $s_{-i}$ . Then  $F$  is convex-valued. Therefore,  $F$  satisfies the conditions of Kakutani's fixed-point theorem and it must have a fixed point,

$$s^* \in F(s^*).$$

By definition,  $s_i^* \in B_i(s_{-i}^*)$  for all  $i \in N$ , thus  $s^*$  is a Nash equilibrium.  $\square$

For games with convex strategy sets and quasiconcave utility functions, Theorem ?? proves existence of a *pure* strategy Nash equilibrium. One can use this result to establish the existence of equilibrium in classical economic models, such as generalizations of the Cournot competition game introduced earlier. Theorem ?? also implies the existence of *mixed strategy* Nash equilibria in finite games.

**Corollary 3.** *Every finite game has a mixed strategy Nash equilibrium.*

*Proof.* Since  $S$  is finite, each  $\Delta(S_i)$  is isomorphic to a simplex in a Euclidean space, which is convex and compact. Player  $i$ 's expected utility  $u_i(\sigma) = \sum_s u_i(s) \sigma_1(s_1) \cdots \sigma_n(s_n)$  from a mixed strategy profile  $\sigma$  is continuous in  $\sigma$  and linear—hence also quasi-concave—in  $\sigma_i$ . Then the game  $(N, \Delta(S_1), \dots, \Delta(S_n), u)$  satisfies the assumptions of Theorem ???. Therefore, it admits a Nash equilibrium  $\sigma^* \in \Delta(S_1) \times \cdots \times \Delta(S_n)$ , which can be interpreted as a mixed Nash equilibrium in the original game.  $\square$

**6.3. Upperhemicontinuity of Nash Equilibrium.** The Maximum Theorem establishes that the best-response correspondence in optimization problems is upper hemicontinuous in parameters when the payoffs are continuous and the domain is compact. Hence the limits of optimal solutions is a solution to the optimization problem in the limit. One can then find a solution by considering approximate problems and taking the limit. There can be other solutions in the limit, so the best response correspondence is not lower hemicontinuous in general. Nash equilibrium (like many other solution concepts) inherits these properties of the best response correspondence.

Consider a compact metric space  $X$  of some payoff-relevant parameters. Fix a set  $N$  of players and set  $S$  of strategy profiles, where  $S$  is again a compact metric space (or a finite set). The payoff function  $u_i : S \times X \rightarrow \mathbb{R}$  of every  $i \in N$  is assumed to be continuous in both strategies and parameters. Write  $NE(x)$  and  $PNE(x)$  for the sets of all Nash equilibria and all pure Nash equilibria, respectively, of game  $(N, S, u(\cdot, x))$  in which it is common knowledge that the parameter value is  $x$ . Endow the space of mixed strategies with the weak topology.

**Theorem 9.** *Under the stated assumptions, the correspondences  $NE$  and  $PNE$  have closed graphs.*

*Proof.* We establish the result for  $PNE$ ; the proof for  $NE$  is similar. Consider any sequence  $(s^m, x^m) \rightarrow (s, x)$  with  $s^m \in PNE(x^m)$  for each  $m$ . Suppose that  $s \notin PNE(x)$ . Then

$$u_i(s'_i, s_{-i}, x) - u_i(s_i, s_{-i}, x) > 0$$

for some  $i \in N$ ,  $s'_i \in S_i$ . Then  $(s^m, x^m) \rightarrow (s, x)$  and the continuity of  $u_i$  imply that

$$u_i(s'_i, s^m_{-i}, x^m) - u_i(s^m_i, s^m_{-i}, x^m) > 0$$

for sufficiently large  $m$ . However,

$$u_i(s'_i, s^m_{-i}, x^m) > u_i(s^m_i, s^m_{-i}, x^m)$$

contradicts  $s^m \in PNE(x^m)$ . □

## 7. BAYESIAN GAMES

When some players are uncertain about the payoffs or *types* of others, the game is said to have *incomplete information*. Most often a player's type is simply defined by his payoff function. More generally, types may embody any private information that is relevant to players' decision making. This may include, in addition to the player's payoff function, his beliefs about other players' payoff functions, his beliefs about what other players believe his beliefs are, and so on. Modeling incomplete information about higher order beliefs is usually intractable and in most applications a player's uncertainty is assumed to be solely about his opponents' payoffs.<sup>7</sup>

A *Bayesian game* is a list  $\mathcal{B} = (N, S, \Theta, u, p)$  where

- $N = \{1, 2, \dots, n\}$  is a finite set of *players*
- $S_i$  is the set of *pure strategies* of player  $i$ ;  $S = S_1 \times \dots \times S_n$
- $\Theta_i$  is the set of *types* of player  $i$ ;  $\Theta = \Theta_1 \times \dots \times \Theta_n$
- $u_i : \Theta \times S \rightarrow \mathbb{R}$  is the *payoff function* of player  $i$ ;  $u = (u_1, \dots, u_n)$
- $p \in \Delta(\Theta)$  is a common prior (we can relax this assumption).

We often assume that  $\Theta$  is finite and the marginal  $p(\theta_i)$  is positive for each type  $\theta_i$ .

**Example 9** (First Price Auction with I.I.D. Private Values). One object is up for sale. Suppose that the value  $\theta_i$  of player  $i \in N$  for the object is uniformly distributed in  $\Theta_i = [0, 1]$  and that the values are independent across players. This means that if  $\tilde{\theta}_i \in [0, 1], \forall i$  then  $p(\theta_i \leq \tilde{\theta}_i, \forall i) = \prod_i \tilde{\theta}_i$ . Each player  $i$  submits a bid  $s_i \in S_i = [0, \infty)$ . The player with the

---

<sup>7</sup>See the slides for the general model and a rigorous treatment of higher order beliefs in Bayesian games.

highest bid wins the object and pays his bid. Ties are broken randomly. Hence the payoffs are given by

$$u_i(\theta, s) = \begin{cases} \frac{\theta_i - s_i}{|\{j \in N \mid s_i = s_j\}|} & \text{if } s_i \geq s_j, \forall j \in N \\ 0 & \text{otherwise.} \end{cases}$$

**Example 10** (An exchange game). Each player  $i = 1, 2$  receives a ticket on which there is a number in some finite set  $\Theta_i \subset [0, 1]$ . The number on a player's ticket represents the size of a prize he may receive. The two prizes are independently distributed, with the value on  $i$ 's ticket distributed according to  $F_i$ . Each player is asked independently and simultaneously whether he wants to exchange his prize for the other player's prize, hence  $S_i = \{\text{agree}, \text{disagree}\}$ . If both players agree then the prizes are exchanged; otherwise each player receives his own prize. Thus the payoff of player  $i$  is

$$u_i(\theta, s) = \begin{cases} \theta_{3-i} & \text{if } s_1 = s_2 = \text{agree} \\ \theta_i & \text{otherwise.} \end{cases}$$

In the *ex ante* representation  $G(\mathcal{B})$  of the Bayesian game  $\mathcal{B}$  player  $i$  has strategies  $(s_i(\theta_i))_{\theta_i \in \Theta_i} \in S_i^{\Theta_i}$ —that is, his strategies are functions from types to strategies in  $\mathcal{B}$ —and utility function  $U_i$  given by

$$U_i((s_i(\theta_i))_{\theta_i \in \Theta_i, i \in N}) = \mathbb{E}_p(u_i(\theta, s_1(\theta_1), \dots, s_n(\theta_n))).$$

The *interim* representation  $IG(\mathcal{B})$  of the Bayesian game  $\mathcal{B}$  has player set  $\cup_i \Theta_i$ . The strategy space of each player  $\theta_i$  is  $S_i$ . A strategy profile  $(s_{\theta_i})_{i \in N, \theta_i \in \Theta_i}$  yields utility

$$U_{\theta_i}((s_{\theta_i})_{i \in N, \theta_i \in \Theta_i}) = \mathbb{E}_p(u_i(\theta, s_{\theta_1}, \dots, s_{\theta_n}) | \theta_i)$$

for player  $\theta_i$ . For the conditional expectation to be well-defined we need  $p(\theta_i) > 0$ .

**Definition 13.** A Bayesian Nash equilibrium of  $\mathcal{B}$  is a Nash equilibrium of  $IG(\mathcal{B})$ .

**Proposition 1.** *Every Bayesian Nash equilibrium of  $\mathcal{B}$  is a Nash equilibrium of  $G(\mathcal{B})$ . If  $p(\theta_i) > 0$  for all  $\theta_i \in \Theta_i, i \in N$ , the converse also holds.*<sup>8</sup>

**Theorem 10.** *Suppose that*

- $N$  and  $\Theta$  are finite

<sup>8</sup>Strategies are mapped between the two games by  $s_i(\theta_i) \rightarrow s_{\theta_i}$ .

- each  $S_i$  is a compact and convex subset of a Euclidean space
- each  $u_i$  is continuous in  $s$  and concave in  $s_i$ .

Then  $\mathcal{B}$  has a pure strategy Bayesian Nash equilibrium.

*Proof.* We have to show that  $IG(\mathcal{B})$  has a pure Nash equilibrium. The latter follows from Theorem ???. We use the concavity of  $u_i$  in  $s_i$  to show that the corresponding  $U_{\theta_i}$  is quasi-concave in  $s_{\theta_i}$ . Quasi-concavity of  $u_i$  in  $s_i$  does not typically imply quasi-concavity of  $U_{\theta_i}$  in  $s_{\theta_i}$  because  $U_{\theta_i}$  is an integral of  $u_i$  over variables other than  $s_{\theta_i}$ .<sup>9</sup>  $\square$

**Example 11** (Study groups). Two students work on a joint project. Each student  $i$  can either exert effort ( $e_i = 1$ ) or shirk ( $e_i = 0$ ). The cost of effort is a fixed (and commonly known)  $c < 1$  for all students. The project is a success if at least one student puts in effort, but both fail otherwise. However, students vary in how much they care about the fate of the project. Concretely, each student  $i$  has a type  $\theta_i \sim U[0, 1]$ ; the types of both students are independently distributed and privately known. The payoff from success is  $\theta_i^2$ , so a student gets  $\theta_i^2 - c$  from working,  $\theta_i^2$  from shirking if  $j$  works, and 0 if both shirk.

This game has a unique Bayesian Nash equilibrium. In equilibrium, both players use a threshold strategy:  $i$  works if  $\theta_i \geq \theta^* = c^{\frac{1}{3}}$ , and shirks otherwise. For a proof, note that working is rational for  $i$  iff

$$\theta_i^2 - c \geq \theta_i^2 p_j \iff (1 - p_j)\theta_i^2 \geq c,$$

where  $p_j$  is  $i$ 's belief about the probability that  $j$  will work. (Crucially, since types are independent,  $p$  does not depend on  $\theta_i$ ). This implies that  $i$  must play a threshold strategy, with threshold  $\theta_i^* = \sqrt{\frac{c}{1-p_j}}$ . Since the same is true of player  $j$ , we have  $p_j = 1 - \theta_j^*$ , so  $\theta_i^* = \sqrt{\frac{c}{\theta_j^*}}$ . This is true for  $i = 1, j = 2$  and vice-versa, which implies the result.

Consider a family of Bayesian games  $\mathcal{B}^x$  parameterized by  $x \in X$ , with  $X$  compact, such that the payoff functions are continuous in  $x$ . If  $S, \Theta$  are finite, then the set of Bayesian Nash equilibria of  $\mathcal{B}^x$  is upper-hemicontinuous with respect to  $x$ . Indeed,  $BNE(\mathcal{B}^x) = NE(IG(\mathcal{B}^x))$ . Furthermore, we have upper-hemicontinuity with respect to beliefs.

**Theorem 11.** *Suppose that  $N, S, \Theta$  are finite. Let  $P \subset \Delta(\Theta)$  be such that for every  $p \in P$   $p(\theta_i) > 0, \forall \theta_i \in \Theta_i, i \in N$ . Then  $BNE(\mathcal{B}^p)$  is upper-hemicontinuous in  $p$  over  $P$ .*

<sup>9</sup>Sums of quasi-concave functions are not necessarily quasi-concave.

*Proof.* Since  $BNE(\mathcal{B}^p) = NE(IG(\mathcal{B}^p))$ , it is sufficient to note that

$$U_i((s_i(\theta_i))_{\theta_i \in \Theta_i, i \in N}) = \mathbb{E}_p(u_i(\theta, s_1(\theta_1), \dots, s_n(\theta_n)))$$

(as defined for  $G(\mathcal{B}^p)$ ) is continuous in  $p$ . □

## 8. AUCTIONS

In class we covered (see handouts)

- (1) the characterization of equilibria in first and second price auctions (pp. 14-20 in “Auction Theory” by Vijay Krishna)
- (2) the revenue equivalence theorem and optimal auctions (pp. 61-73 in “Auction Theory”)
- (3) all-pay and third-price auctions (pp. 31-34 in “Auction Theory”)
- (4) first-price auction with two asymmetric bidders (pp. 49-53 in “Auction Theory”)
- (5) bilateral trade and the inefficiency theorem of Myerson and Satterthwaite (1983).

## 9. EXTENSIVE FORM GAMES

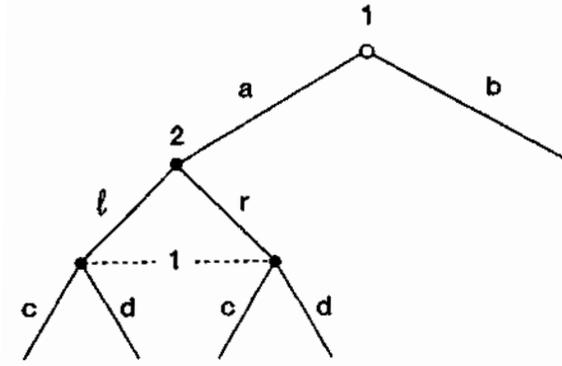
An *extensive form game* consists of

- a finite set of *players*  $N = \{1, 2, \dots, n\}$ ; nature is denoted as “player 0”
- the order of moves specified by a *tree*
- each player’s *payoffs* at the terminal nodes in the tree
- *information* partition
- the set of *actions* available at every information set and a description of how actions lead to progress in the tree
- moves by nature.

Go over the extensive form Figure ?? (FT, p. 86).

The *tree* is described by a *binary relationship*  $(X, >)$ , where  $x > y$  is interpreted as “node  $x$  precedes node  $y$ .” We assume that  $X$  is *finite*, there is an initial node  $\phi \in X$  ( $\phi > x$  for all  $x \neq \phi$ ),  $>$  is transitive ( $x > y, y > z \Rightarrow x > z$ ) and asymmetric ( $x > y \Rightarrow y \not> x$ ). Hence the tree has no cycles. We also require that each node  $x \neq \phi$  has exactly one immediate predecessor, i.e.,  $\exists x' > x$  such that  $x'' > x, x'' \neq x'$  implies  $x'' > x'$ . A node is *terminal* if it does not precede any other node; this means that the set of terminal nodes is

FIGURE 4



Courtesy of The MIT Press. Used with permission.

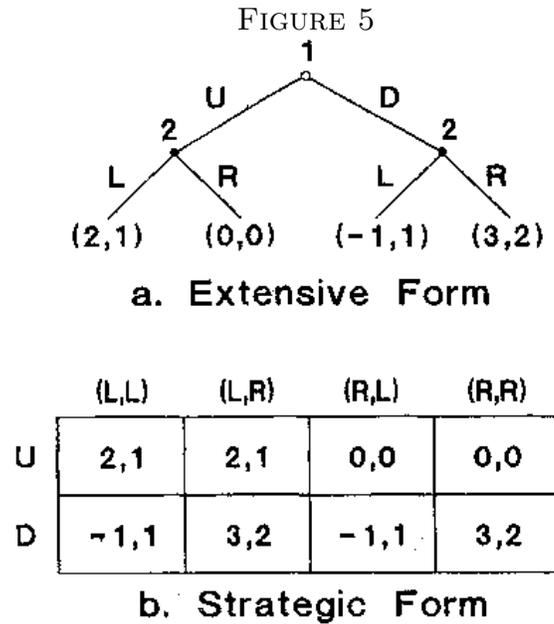
$Z = \{z \mid \nexists x, z > x\}$ . Each  $z \in Z$  completely determines a path of moves through the tree (recall the finiteness assumption), with associated payoff  $u_i(z)$  for player  $i$ .

An *information partition* is a partition of  $X \setminus Z$ . Node  $x$  belongs to the information set  $h(x)$ . For each information set  $h$ , there is a player  $i(h) \in N \cup \{0\}$ , who is to move at any node  $x \in h$ . The interpretation of the information partition is that  $i(h)$  knows that he is at some node of  $h$  but does not know which one. (We must assume the same player moves at all  $x \in h$ , otherwise players might disagree on whose turn it is.) We abuse notation writing  $i(x) = i(h(x))$ .

The set of available *actions* at  $x \in X \setminus Z$  for player  $i(x)$  is  $A(x)$ . We assume that  $A(x) = A(x') =: A(h), \forall x' \in h(x)$  (otherwise  $i(h)$  might play an infeasible action). A function  $l$  labels each node  $x \neq \phi$  with the last action taken to reach it. We require that each immediate successor of  $x$  is labeled with a different action in  $A(x)$ , and each such action is used for some successor of  $x$ . Finally, a move by nature at some node  $x$  corresponds to a probability distribution over  $A(x)$ .

Let  $H_i = \{h \mid i(h) = i\}$ . The set of *pure strategies* for player  $i$  is  $S_i = \prod_{h \in H_i} A(h)$ . As usual,  $S = \prod_{i \in N} S_i$ . A strategy is a complete *contingent plan* specifying an action to be taken at each information set (if reached). We can define *mixed strategies* as probability distributions over pure strategies,  $\sigma_i \in \Delta(S_i)$ . Any mixed strategy profile  $\sigma \in \prod_{i \in N} \Delta(S_i)$ , along with the distribution of the moves by nature and the labeling of nodes with actions, leads to a probability distribution  $O(\sigma) \in \Delta(Z)$ . We denote  $u_i(\sigma) = \mathbb{E}_{O(\sigma)}(u_i(z))$ . The *strategic form representation* of the extensive form game is the normal form game defined by  $(N, S, u)$ . A mixed strategy profile constitutes a *Nash equilibrium* of the extensive form

game if it induces a Nash equilibrium in its strategic form representation. See Figure ?? (FT, p. 85).



Courtesy of The MIT Press. Used with permission.

Two strategies  $\sigma_i, \sigma'_i \in S_i$  are *outcome equivalent* if  $O(\sigma_i, s_{-i}) = O(\sigma'_i, s_{-i}), \forall s_{-i}$ , that is, they lead to the same distribution over outcomes regardless of how the opponents play. See figure 3.9 in FT p. 86. The strategies  $(b, c)$  and  $(b, d)$  are equivalent in that example. Let  $S_i^R$  be a subset of  $S_i$  that contains exactly one strategy from each equivalence class. The *reduced normal form* game is given by  $(N, S^R, u)$ .

A *behavior strategy* specifies a distribution over actions for each information set. Formally, a behavior strategy  $b_i(h)$  for player  $i(h)$  at information set  $h$  is an element of  $\Delta(A(h))$ . We use the notation  $b_i(a|h)$  for the probability of action  $a$  at information set  $h$ . A behavior strategy  $b_i$  for  $i$  is an element of  $\prod_{h \in H_i} \Delta(A(h))$ . Note that behavior strategies assume *independent mixing* at each information set. A profile  $b$  of behavior strategies determines a distribution over  $Z$  in the obvious way. By definition, the behavior strategy  $b_i$  is outcome equivalent to the mixed strategy

$$(9.1) \quad \sigma_i(s_i) = \prod_{h \in H_i} b_i(s_i(h)|h),$$

where  $s_i(h)$  denotes the projection of  $s_i$  on  $A(h)$ .

To guarantee that every mixed strategy is equivalent to a behavior strategy (reinterpreted as a mixed strategy as above) we need to impose the additional requirement of *perfect recall*. Basically, perfect recall means that no player ever forgets any information he once had and all players know the actions they have chosen previously. Formally, perfect recall stipulates that if  $x'' \in h(x')$ ,  $x$  is a predecessor of  $x'$  and the same player  $i$  moves at both  $x$  and  $x'$  (and thus at  $x''$ ) then there is a node  $\hat{x}$  in the same information set as  $x$  (possibly  $x$  itself) such that  $\hat{x}$  is a predecessor of  $x''$  and the action taken at  $x$  along the path to  $x'$  is the same as the action taken at  $\hat{x}$  along the path to  $x''$ . Intuitively, the nodes  $x'$  and  $x''$  are distinguished by information  $i$  does not have, so he cannot have had it at  $h(x)$ ;  $x'$  and  $x''$  must be consistent with the same action at  $h(x)$  since  $i$  must remember his action there. Stated differently, every node in  $h \in H_i$  must be reached via the same sequence of  $i$ 's actions.

Discuss the absent-minded driver's paradox of Piccione and Rubinstein (1997).

Let  $R_i(h)$  be the set of pure strategies for player  $i$  that do not preclude reaching the information set  $h \in H_i$ , i.e.,  $R_i(h) = \{s_i | h \text{ is on the path of } (s_i, s_{-i}) \text{ for some } s_{-i}\}$ . If the game has perfect recall, a mixed strategy  $\sigma_i$  is equivalent to a behavior strategy  $b_i$  defined by

$$(9.2) \quad b_i(a|h) = \frac{\sum_{\{s_i \in R_i(h) | s_i(h)=a\}} \sigma_i(s_i)}{\sum_{s_i \in R_i(h)} \sigma_i(s_i)}$$

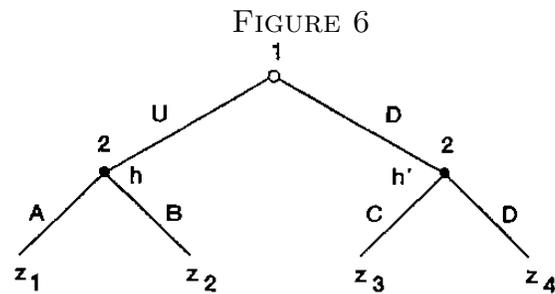
when the denominator is positive, and letting  $b_i(h)$  be any distribution when the above denominator is zero.

For some intuition, let  $h_1, \dots, h_{\bar{k}}$  be player  $i$ 's information sets preceding  $h$  in the tree. In a game of perfect recall, reaching any node in  $h$  requires  $i$  to take the same action  $a_k$  at each  $h_k$ . Then  $R_i(h)$  is simply the set of pure strategies  $s_i$  with  $s_i(h_k) = a_k$  for all  $k$ . Reaching any node  $x$  in  $h$  requires player  $i$  to choose  $a_1, \dots, a_{\bar{k}}$  and other players to also take specific actions. Conditional on getting to  $x$ , the distribution of continuation play at  $x$  is given by the relative probabilities of the actions available at  $h$  under the restriction of  $\sigma_i$  to the set of pure strategies  $\{s_i | s_i(h_k) = a_k, \forall k = \overline{1, \bar{k}}\}$  compatible with reaching  $h$ ,

$$b_i(a|h) = \frac{\sum_{\{s_i | s_i(h_k)=a_k, \forall k=\overline{1, \bar{k}} \ \& \ s_i(h)=a\}} \sigma_i(s_i)}{\sum_{\{s_i | s_i(h_k)=a_k, \forall k=\overline{1, \bar{k}}\}} \sigma_i(s_i)}.$$

Many different mixed strategies can generate the same behavior strategy. Consider the example from Figure ?? (FT, p. 88). Player 2 has four pure strategies,  $s_2 = (A, C), s'_2 =$

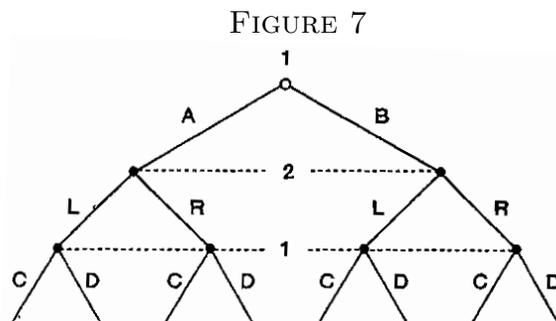
$(A, D), s_2'' = (B, C), s_2''' = (B, D)$ . Now consider two mixed strategies,  $\sigma_2 = (1/4, 1/4, 1/4, 1/4)$ , which assigns probability  $1/4$  to each pure strategy, and  $\sigma_2' = (1/2, 0, 0, 1/2)$ , which assigns probability  $1/2$  to each of  $s_2$  and  $s_2'''$ . Both of these mixed strategies generate the behavior strategy  $b_2$  with  $b_2(A|h) = b_2(B|h) = 1/2$  and  $b_2(C|h') = b_2(D|h') = 1/2$ . Moreover, for any strategy  $\sigma_1$  of player 1, all of  $\sigma_2, \sigma_2', b_2$  lead to the same probability distribution over terminal nodes. For example, the probability of reaching node  $z_1$  equals the probability of player 1 playing  $U$  times  $1/2$ .



Courtesy of The MIT Press. Used with permission

The relationship between mixed and behavior strategies is different in the game illustrated in Figure ?? (FT, p. 89), which is not a game of perfect recall (player 1 forgets what he did at the initial node; formally, there are two nodes in his second information set which obviously succeed the initial node, but are not reached by the same action at the initial node). Player 1 has four strategies in the strategic form,  $s_1 = (A, C), s_1' = (A, D), s_1'' = (B, C), s_1''' = (B, D)$ . Now consider the mixed strategy  $\sigma_1 = (1/2, 0, 0, 1/2)$ . This generates the behavior strategy  $b_1 = \{(1/2, 1/2), (1/2, 1/2)\}$ , where player 1 mixes 50 – 50 at each information set. But  $b_1$  is *not* equivalent to the  $\sigma_1$  that generated it. Indeed  $(\sigma_1, L)$  generates a probability  $1/2$  for the terminal node corresponding to  $(A, C, L)$  and a  $1/2$  probability for  $(B, L, D)$ . However, since behavior strategies describe independent randomizations at each information set,  $(b_1, L)$  assigns probability  $1/4$  to each of the four paths  $(A, C, L), (A, D, L), (B, C, L), (B, D, L)$ . Since both  $A$  vs.  $B$  and  $C$  vs.  $D$  are choices made by player 1, the strategy  $\sigma_1$  under which player 1 makes all his decisions at once allows choices at different information sets to be correlated. Behavior strategies cannot produce this correlation, because when it comes time to choose between  $C$  and  $D$ , player 1 has forgotten whether he chose  $A$  or  $B$ . By assumption, player 1 cannot distinguish between the nodes following  $A$  and  $B$ , and in line

with this assumption, behavior strategies cannot condition on the past choice between  $A$  and  $B$ .



Courtesy of The MIT Press. Used with permission.

**Theorem 12** (Kuhn 1953). *In extensive form games with perfect recall, mixed and behavior strategies are outcome equivalent under the formulae ??-??.*

Hereafter we restrict attention to games with perfect recall, and use mixed and behavior strategies interchangeably. Behavior strategies prove more convenient in many arguments and constructions. We drop the notation  $b$  for behavior strategies and instead use  $\sigma_i(a|h)$  to denote the probability with which player  $i$  chooses action  $a$  at information set  $h$ ...

## 10. BACKWARD INDUCTION AND SUBGAME PERFECTION

An extensive form game has *perfect information* if all information sets are *singletons*. *Backward induction* can be applied to any extensive form game of perfect information with finite  $X$  (which means that the number of “stages” and the number of actions feasible at any stage are finite). The idea of backward induction is formalized by Zermelo’s algorithm. Since the game is finite, it has a set of penultimate nodes, i.e., nodes whose immediate successors are (all) terminal nodes. Specify that the player who moves at each such node chooses the strategy leading to the terminal node with the highest payoff for him. In case of a tie, make an arbitrary selection. Next each player at nodes whose immediate successors are penultimate (or terminal) nodes chooses the action maximizing his payoff over the feasible successors, given that players at the penultimate nodes play as assumed. We can now roll back through the tree, specifying actions at each node. At the end of the process we have a pure strategy for each player. It is easy to check that the resulting strategies form a Nash equilibrium.

**Theorem 13** (Zermelo 1913; Kuhn 1953). *A finite game of perfect information has a pure-strategy Nash equilibrium.*

Moreover, the backward induction solution has the nice property that, if play starts at any intermediate node, each player's actions are again optimal if the play of the opponents is held fixed, which we call *subgame perfection*. This rules out non-credible threats in response to deviations from the prescribed play. More generally, subgame perfection extends the logic of backward induction to games with imperfect information. The idea is to replace the “smallest” subgame with one of its Nash equilibria and iterate this procedure on the reduced tree. In stages where the “smallest” subgame has multiple Nash equilibria, the procedure implicitly assumes that all players believe the same equilibrium will be played. To define subgame perfection formally we first need the definition of a subgame. Informally, a subgame is a portion of a game that can be analyzed as a game in its own right.

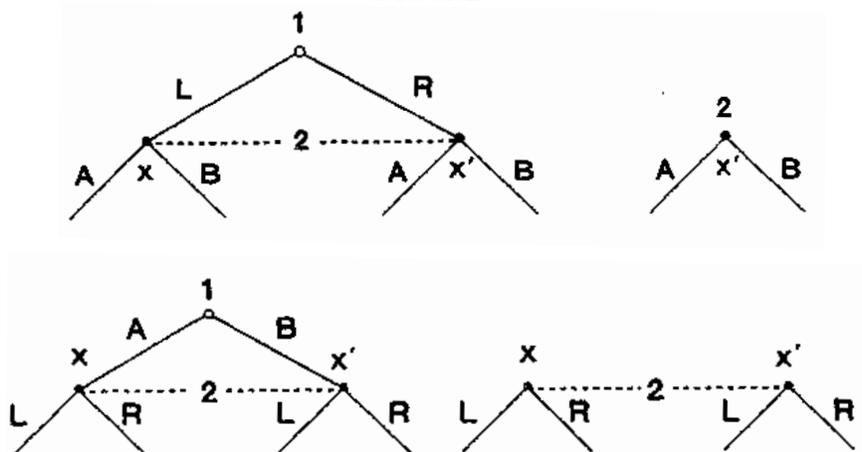
**Definition 14.** A *subgame*  $G$  of an extensive form game  $T$  consists of a single node  $x$  and *all* its successors in  $T$ , with the property that if  $x' \in G$  and  $x'' \in h(x')$  then  $x'' \in G$ . The information sets, actions and payoffs of the subgame are inherited from the original game. That is, two nodes are in the same information set in  $G$  if and only if they are in the same information set in  $T$ , and the payoff function on the subgame is just the restriction of the original payoff function to the terminal nodes of  $G$  (and likewise for the action sets and action labels).

The requirements that all the successors of  $x$  be in the subgame and that the subgame does not “chop up” any information set ensure that the subgame corresponds to a situation that could arise in the original game. At the top of Figure ?? (FT, p. 95), the game on the right is not a subgame of the game on the left, because on the right player 2 knows that player 1 has not played  $L$ , which he did not know in the original game.

Together, the requirements that the subgame begin with a single node  $x$  and respect information sets imply that in the original game  $x$  must be a singleton information set, i.e.  $h(x) = \{x\}$ . This ensures that the distribution over paths of play and payoffs in the subgame, conditional on the subgame being reached, are well defined. In the bottom of Figure ??, the “game” on the right has the problem that player 2's optimal choice may depend on the relative probabilities of nodes  $x$  and  $x'$ , but the specification of the game does

not provide these probabilities. In other words, the diagram on the right cannot be analyzed as a separate game.

FIGURE 8



Courtesy of The MIT Press. Used with permission.

Given any strategy profile  $\sigma$ , payoffs within a subgame  $G$  are well-defined—just start play at the initial node of  $G$ , follow the actions specified by  $\sigma$ , and take the payoffs of the resulting terminal node (or their expectations, if mixing is involved). So we can test whether strategies yield a Nash equilibrium when restricted to the subgame.

**Definition 15.** A behavior strategy profile  $\sigma$  of an extensive form game is a *subgame perfect equilibrium* if the restriction of  $\sigma$  to  $G$  is a Nash equilibrium of  $G$  for every subgame  $G$ .

Because any game is a subgame of itself, a subgame perfect equilibrium profile is necessarily a Nash equilibrium. If the only subgame is the whole game, the sets of Nash and subgame perfect equilibria coincide. If there are other subgames, some Nash equilibria may fail to be subgame perfect.

Subgame perfection coincides with backward induction in finite games of perfect information. Consider the penultimate nodes of the tree, where the last choices are made. Each of these nodes begins a trivial one-player subgame, and Nash equilibrium in these subgames requires that the player make a choice that maximizes his payoff. Thus any subgame perfect equilibrium must coincide with a backward induction solution at every penultimate node, and we can continue up the tree by induction.

## 11. IMPORTANT EXAMPLES OF EXTENSIVE FORM GAMES

## 11.1. Repeated games with perfect monitoring.

- *time*  $t = 0, 1, 2, \dots$  (usually infinite)
- *stage game* is a normal-form game  $G$
- $G$  is played every period  $t$
- players *observe* the realized actions at the end of each period
- *payoffs* are the sum of discounted payoffs in the stage game.

Repeated games are a widely studied class of dynamic games. A lot of recent research deals with the case of imperfect monitoring.

## 11.2. Multi-stage games with observable actions.

- *stages*  $t = 0, 1, 2, \dots$
- at stage  $t$ , after having *observed* a “non-terminal” history of play  $h = (a^0, \dots, a^{t-1})$ , each player  $i$  simultaneously chooses an *action*  $a_i^t \in A_i(h)$
- *payoffs* given by  $u(h)$  for each terminal history  $h$ .

Often it is natural to identify the “stages” of the game with time periods, but this is not always the case. A game of perfect information can be viewed as a multi-stage game in which every stage corresponds to some node. At every stage all but one player (the one moving at the node corresponding to that stage) have singleton action sets (“do nothing”; we can refer to these players as “inactive”). Repeated versions of perfect information extensive form games also lead to multi-stage games. Another important example is the Rubinstein (1982) alternating bargaining game, which we discuss later.

## 12. SINGLE (OR ONE-SHOT) DEVIATION PRINCIPLE

Consider a multi-stage game with observed actions. We show that in order to verify that a strategy profile  $\sigma$  constitutes a subgame perfect equilibrium, it suffices to check whether there are any histories  $h_t$  where some player  $i$  can gain by deviating from the action prescribed by  $\sigma_i$  at  $h_t$  and conforming to  $\sigma_i$  elsewhere. (The notation  $h_t$  denotes a history as of stage  $t$ .)

If  $\sigma$  is a strategy profile and  $h_t$  a history, write  $u_i(\sigma|h_t)$  for the (expected) payoff to player  $i$  that results if play starts at  $h_t$  and continues according to  $\sigma$  in each stage.

**Definition 16.** A (behavior) strategy  $\sigma_i$  is *unimprovable* given  $\sigma_{-i}$  if  $u_i(\sigma_i, \sigma_{-i} | h_t) \geq u_i(\sigma'_i, \sigma_{-i} | h_t)$  for every  $h_t$  and  $\sigma'_i$  with  $\sigma'_i(h'_t) = \sigma_i(h'_t)$  for all  $h'_t \neq h_t$ .

Hence a strategy  $\sigma_i$  is unimprovable if after every history, no strategy that differs from it only at that history can increase  $i$ 's expected payoff. Obviously, if  $\sigma$  is a subgame perfect equilibrium then  $\sigma_i$  is unimprovable given  $\sigma_{-i}$ . To establish the converse, we need an additional condition.

**Definition 17.** A game is *continuous at infinity* if for each player  $i$  the utility function  $u_i$  satisfies

$$\lim_{t \rightarrow \infty} \sup_{\{(h, \tilde{h}) | h_t = \tilde{h}_t\}} |u_i(h) - u_i(\tilde{h})| = 0.$$

Continuity at infinity requires that events in the distant future are relatively unimportant. It is satisfied if the overall payoffs are a discounted sum of per-period payoffs and the stage payoffs are uniformly bounded. It also holds in the (degenerate) case of finite-stage games. There exist versions for games with unobserved actions as well.

**Theorem 14.** Consider a (finite or infinite horizon) multi-stage game with observed actions<sup>10</sup> that is continuous at infinity. If  $\sigma_i$  is unimprovable given  $\sigma_{-i}$  for all  $i \in N$ , then  $\sigma$  constitutes an SPE.

*Proof.* Suppose that  $\sigma_i$  is unimprovable given  $\sigma_{-i}$ , but  $\sigma_i$  is not a best response to  $\sigma_{-i}$  following some history  $h_t$ . Let  $\sigma_i^1$  be a strictly better response and define

$$(12.1) \quad \varepsilon = u_i(\sigma_i^1, \sigma_{-i} | h_t) - u_i(\sigma_i, \sigma_{-i} | h_t) > 0.$$

Since the game is *continuous at infinity*, there exists  $t' > t$  and  $\sigma_i^2$  such that  $\sigma_i^2$  is identical to  $\sigma_i^1$  at all information sets up to (and including) stage  $t'$ ,  $\sigma_i^2$  coincides with  $\sigma_i$  across all longer histories and

$$(12.2) \quad |u_i(\sigma_i^2, \sigma_{-i} | h_t) - u_i(\sigma_i^1, \sigma_{-i} | h_t)| < \varepsilon/2.$$

In particular, ?? and ?? imply that

$$u_i(\sigma_i^2, \sigma_{-i} | h_t) > u_i(\sigma_i, \sigma_{-i} | h_t).$$

<sup>10</sup>We allow for the possibility that the action set be infinite at some stages.

Denote by  $\sigma_i^3$  the strategy obtained from  $\sigma_i^2$  by replacing the stage  $t'$  actions following any history  $h_{t'}$  with the corresponding actions under  $\sigma_i$ . Conditional on any history  $h_{t'}$ , the strategies  $\sigma_i$  and  $\sigma_i^3$  coincide, hence

$$(12.3) \quad u_i(\sigma_i^3, \sigma_{-i}|h_{t'}) = u_i(\sigma_i, \sigma_{-i}|h_{t'}).$$

As  $\sigma_i$  is *unimprovable* given  $\sigma_{-i}$ , and conditional on  $h_{t'}$  the subsequent play in strategies  $\sigma_i$  and  $\sigma_i^2$  differs only at stage  $t'$ , we must have

$$(12.4) \quad u_i(\sigma_i, \sigma_{-i}|h_{t'}) \geq u_i(\sigma_i^2, \sigma_{-i}|h_{t'}).$$

Then ?? and ?? lead to

$$u_i(\sigma_i^3, \sigma_{-i}|h_{t'}) \geq u_i(\sigma_i^2, \sigma_{-i}|h_{t'})$$

for all histories  $h_{t'}$  (consistent with  $h_t$ ). Since  $\sigma_i^2$  and  $\sigma_i^3$  coincide before reaching stage  $t'$ , we obtain

$$u_i(\sigma_i^3, \sigma_{-i}|h_t) \geq u_i(\sigma_i^2, \sigma_{-i}|h_t).$$

Similarly, we can construct  $\sigma_i^4, \dots, \sigma_i^{t'-t+3}$ . The final strategy  $\sigma_i^{t'-t+3}$  is identical to  $\sigma_i$  conditional on  $h_t$  and

$$u_i(\sigma_i, \sigma_{-i}|h_t) = u_i(\sigma_i^{t'-t+3}, \sigma_{-i}|h_t) \geq \dots \geq u_i(\sigma_i^3, \sigma_{-i}|h_t) \geq u_i(\sigma_i^2, \sigma_{-i}|h_t) > u_i(\sigma_i, \sigma_{-i}|h_t),$$

a contradiction. □

**12.1. Applications.** Apply the single deviation principle to repeated prisoners' dilemma to implement various equilibrium paths for high discount factors:

- (1)  $(C, C), (C, C), \dots$
- (2)  $(C, C), (C, C), (D, D), (C, C), (C, C), (D, D), \dots$
- (3)  $(C, D), (D, C), (C, D), (D, C) \dots$

In particular, note that cooperation is possible in repeated play.

	C	D
C	1, 1	-1, 2
D	2, -1	0, 0*

Also find the stationary equilibrium for the alternating bargaining game in which two players divide \$1. We will show that is the unique subgame perfect equilibrium.

## 13. ITERATED CONDITIONAL DOMINANCE

**Definition 18.** In a multi-stage game with observable actions, an action  $a_i$  is conditionally dominated at stage  $t$  given history  $h_t$  if, in the subgame starting at  $h_t$ , every strategy for player  $i$  that assigns positive probability to  $a_i$  is strictly dominated.

**Proposition 2.** *In any multi-stage game with observable actions, every subgame perfect equilibrium survives iterated elimination of conditionally dominated strategies.*

## 14. BARGAINING WITH ALTERNATING OFFERS

One important example of a multi-stage game with observed actions is the following bargaining game, analyzed by Rubinstein (1982).

The set of players is  $N = \{1, 2\}$ . For  $i = 1, 2$  we write  $j = 3 - i$ . The set of feasible utility pairs is

$$U = \{(u_1, u_2) \in [0, \infty)^2 \mid u_2 \leq g_2(u_1)\},$$

where  $g_2$  is some strictly decreasing, concave (and hence continuous) function with  $g_2(0) > 0$ .<sup>11</sup>

Time is discrete and infinite,  $t = 0, 1, \dots$ . Each player  $i$  discounts payoffs by  $\delta_i$ , so receiving  $u_i$  at time  $t$  is worth  $\delta_i^t u_i$ .

At every time  $t = 0, 1, \dots$ , player  $i(t)$  proposes an alternative  $u = (u_1, u_2) \in U$  to player  $j(t) = 3 - i(t)$ ; the bargaining protocol specifies that  $i(t) = 1$  for  $t$  even and  $i(t) = 2$  for  $t$  odd. If  $j(t)$  accepts the offer, then the game ends yielding a payoff vector  $(\delta_1^t u_1, \delta_2^t u_2)$ . Otherwise, the game proceeds to period  $t + 1$ . If agreement is never reached, each player receives a 0 payoff.

It is useful to define the function  $g_1 = g_2^{-1}$ . Notice that the graph of  $g_2$  (and  $g_1^{-1}$ ) coincides with the Pareto-frontier of  $U$ .

---

<sup>11</sup>The set of feasible utility outcomes  $U$  can be generated from a set of contracts or decisions  $X$  in a natural way. Define  $U = \{(v_1(x), v_2(x)) \mid x \in X\}$  for a pair of utility functions  $v_1$  and  $v_2$  over  $X$ . With additional assumptions on  $X, v_1, v_2$  we can ensure that the resulting  $U$  is compact and convex.

14.1. **Stationary subgame perfect equilibrium.** Let  $(m_1, m_2)$  be the unique solution to the following system of equations

$$\begin{aligned} m_1 &= \delta_1 g_1(m_2) \\ m_2 &= \delta_2 g_2(m_1). \end{aligned}$$

Note that  $(m_1, m_2)$  is the intersection of the graphs of the functions  $\delta_2 g_2$  and  $(\delta_1 g_1)^{-1}$ .

We are going to argue that the following “stationary” strategies constitute a subgame perfect equilibrium, and that any other subgame perfect equilibrium leads to the same outcome. In any period where player  $i$  has to make an offer to  $j$ , he offers  $u$  with  $u_j = m_j$  and  $j$  accepts only offers  $u$  with  $u_j \geq m_j$ . We can use the *single-deviation principle* to check that the constructed strategies form a subgame perfect equilibrium.

14.2. **Equilibrium uniqueness.** We can use *iterated conditional dominance* to rule out many actions and then prove that the stationary equilibrium is essentially the unique subgame perfect equilibrium.

**Theorem 15.** *The subgame perfect equilibrium is unique, except for the decision to accept or reject Pareto-inefficient offers.*

*Proof.* Player  $i$  cannot obtain a period  $t$  expected payoff greater than

$$M_i^0 = \delta_i \max_{u \in U} u_i = \delta_i g_i(0)$$

following a disagreement at date  $t$ . Hence rejecting an offer  $u$  with  $u_i > M_i^0$  is conditionally dominated by accepting such an offer for  $i$ . Once we eliminate these dominated actions,  $i$  accepts all offers  $u$  with  $u_i > M_i^0$  from  $j$ . Then making any offer  $u$  with  $u_i > M_i^0$  is dominated for  $j$  by an offer  $\bar{u} = \lambda u + (1 - \lambda)(M_i^0, g_j(M_i^0))$  for  $\lambda \in (0, 1)$ , since both offers will be accepted immediately and the latter is better for  $j$ . We remove all the strategies involving such offers.

Under the surviving strategies,  $j$  can always reject an offer from  $i$  and make a counteroffer next period that leaves him with slightly less than  $g_j(M_i^0)$ , which  $i$  accepts. Hence it is conditionally dominated for  $j$  to accept any offer that gives him less than

$$m_j^1 = \delta_j g_j(M_i^0).$$

After we eliminate the latter actions,  $i$  cannot expect to receive a continuation payoff greater than

$$M_i^1 = \max(\delta_i g_i(m_j^1), \delta_i^2 M_i^0) = \delta_i g_i(m_j^1)$$

in any future period following a disagreement. The second equality holds because  $\delta_i g_i(m_j^1) = \delta_i g_i(\delta_j g_j(M_i^0)) \geq \delta_i g_i(g_j(M_i^0)) = \delta_i M_i^0 \geq \delta_i^2 M_i^0$ .

We can recursively define the sequences

$$\begin{aligned} m_j^{k+1} &= \delta_j g_j(M_i^k) \\ M_i^{k+1} &= \delta_i g_i(m_j^{k+1}) \end{aligned}$$

for  $i = 1, 2$  and  $k \geq 1$ . Since both  $g_1$  and  $g_2$  are decreasing functions, we can easily show that the sequence  $(m_i^k)$  is increasing and  $(M_i^k)$  is decreasing. By arguments similar to those above, we can prove by induction on  $k$  that, in any strategy that survives iterated conditional dominance, player  $i = 1, 2$

- never accepts offers with  $u_i < m_i^k$
- always accepts offers with  $u_i > M_i^k$ , but making such offers is dominated for  $j$ .

One step in the inductive argument for the latter claim is that  $\max(\delta_i g_i(m_j^{k+1}), \delta_i^2 M_i^k) = \delta_i g_i(m_j^{k+1}) = M_i^{k+1}$ , which follows from  $\delta_i g_i(m_j^{k+1}) = \delta_i g_i(\delta_j g_j(M_i^k)) \geq \delta_i g_i(g_j(M_i^k)) = \delta_i M_i^k \geq \delta_i^2 M_i^k$ .

The sequences  $(m_i^k)$  and  $(M_i^k)$  are monotonic and bounded, so they need to converge. The limits satisfy

$$\begin{aligned} m_j^\infty &= \delta_j g_j(\delta_i g_i(m_j^\infty)) \\ M_i^\infty &= \delta_i g_i(m_j^\infty). \end{aligned}$$

It follows that  $(m_1^\infty, m_2^\infty)$  is the (unique) intersection point of the graphs of the functions  $\delta_2 g_2$  and  $(\delta_1 g_1)^{-1}$ . Moreover,  $M_i^\infty = \delta_i g_i(m_j^\infty) = m_i^\infty$ . Therefore, all strategies of  $i$  that survive iterated conditional dominance accept  $u$  with  $u_i > M_i^\infty = m_i^\infty$  and reject  $u$  with  $u_i < m_i^\infty = M_i^\infty$ .

This uniquely determines the reply to every offer that  $i$  makes that gives  $j$  an amount other than  $m_j^\infty$ . Now, at any history where  $i$  is the proposer, he has the option of making offers  $(u_i, g_j(u_i))$  for  $u_i$  arbitrarily close to (but less than)  $g_i(m_j^\infty)$ , which will be accepted by

$j$ . Hence  $i$ 's equilibrium payoff at such a history must be at least  $g_i(m_j^\infty)$ . On the other hand,  $i$  cannot get any more than  $g_i(m_j^\infty)$ . Indeed, any offer made by  $i$  specifying a payoff greater than  $g_i(m_j^\infty)$  for himself would leave  $j$  with less than  $m_j^\infty$ , and we have shown that such offers are rejected by  $j$ . Moreover,  $j$  never offers  $i$  more than  $M_i^\infty = \delta_i g_i(m_j^\infty) \leq g_i(m_j^\infty)$ . So  $i$ 's equilibrium payoff at any history where  $i$  is the proposer must be exactly  $g_i(m_j^\infty)$ , which can only be attained if  $i$  offers  $(g_i(m_j^\infty), m_j^\infty)$  and  $j$  accepts with probability 1.

This now uniquely pins down actions at every history, except those where agent  $j$  has just been given an offer  $(u_i, m_j^\infty)$  for some  $u_i < g_i(m_j^\infty)$ . In this case,  $j$  is indifferent between accepting and rejecting.  $\square$

**14.3. Properties of the subgame perfect equilibrium.** The subgame perfect equilibrium is efficient—agreement is obtained in the first period, without delay. The subgame perfect equilibrium payoffs are given by  $(g_1(m_2), m_2)$ , where  $(m_1, m_2)$  solve

$$\begin{aligned} m_1 &= \delta_1 g_1(m_2) \\ m_2 &= \delta_2 g_2(m_1). \end{aligned}$$

It can be easily shown that the payoff of player  $i$  is increasing in  $\delta_i$  and decreasing in  $\delta_j$ . For a fixed  $\delta_1 \in (0, 1)$ , the payoff of player 2 converges to 0 as  $\delta_2 \rightarrow 0$  and to  $\max_{u \in U} u_2$  as  $\delta_2 \rightarrow 1$ . If  $U$  is symmetric and  $\delta_1 = \delta_2$ , player 1 enjoys a first mover advantage because  $m_1 = m_2$  and  $g_1(m_2) > m_2$ .

## 15. NASH BARGAINING

Assume that  $U$  is such that  $g_2$  is decreasing, strictly concave and continuously differentiable (derivative exists and is continuous). The *Nash (1950) bargaining solution*  $u^*$  is defined by  $\{u^*\} = \arg \max_{u \in U} u_1 u_2 = \arg \max_{u \in U} u_1 g_2(u_1)$ . It is the outcome  $(u_1^*, g_2(u_1^*))$  uniquely pinned down by the first order condition  $g_2(u_1^*) + u_1^* g_2'(u_1^*) = 0$ . Indeed, since  $g_2$  is decreasing and strictly concave, the function  $f$ , given by  $f(x) = g_2(x) + x g_2'(x)$ , is strictly decreasing and continuous and changes sign on the relevant range.

**Theorem 16** (Binmore, Rubinstein and Wolinsky 1985). *Suppose that  $\delta_1 = \delta_2 =: \delta$  in the alternating bargaining model. Then the unique subgame perfect equilibrium payoffs converge to the Nash bargaining solution as  $\delta \rightarrow 1$ .*

*Proof.* <sup>12</sup> Recall that the subgame perfect equilibrium payoffs are given by  $(g_1(m_2), m_2)$  where  $(m_1, m_2)$  satisfies

$$\begin{aligned} m_1 &= \delta g_1(m_2) \\ m_2 &= \delta g_2(m_1). \end{aligned}$$

It follows that  $g_1(m_2) = m_1/\delta$ , hence  $m_2 = g_2(g_1(m_2)) = g_2(m_1/\delta)$ . We rewrite the equations as follows

$$\begin{aligned} g_2(m_1/\delta) &= m_2 \\ g_2(m_1) &= m_2/\delta. \end{aligned}$$

By the mean value theorem, there exists  $\xi \in (m_1, m_1/\delta)$  such that  $g_2(m_1/\delta) - g_2(m_1) = (m_1/\delta - m_1)g_2'(\xi)$ , hence  $(m_2 - m_2/\delta) = (m_1/\delta - m_1)g_2'(\xi)$  or, equivalently,  $m_2 + m_1g_2'(\xi) = 0$ . Substituting  $m_2 = \delta g_2(m_1)$  we obtain  $\delta g_2(m_1) + m_1g_2'(\xi) = 0$ .

Note that  $(g_1(m_2), m_2)$  converges to  $u^*$  as  $\delta \rightarrow 1$  if and only if  $(m_1, m_2)$  does. In order to show that  $(m_1, m_2)$  converges to  $u^*$  as  $\delta \rightarrow 1$ , it is sufficient to show that any limit point of  $(m_1, m_2)$  as  $\delta \rightarrow 1$  is  $u^*$ . Let  $(m_1^*, m_2^*)$  be such a limit point corresponding to a sequence  $(\delta_k)_{k \geq 0} \rightarrow 1$ . Recognizing that  $m_1, m_2, \xi$  are functions of  $\delta$ , we have

$$(15.1) \quad \delta_k g_2(m_1(\delta_k)) + m_1(\delta_k) g_2'(\xi(\delta_k)) = 0.$$

Since  $\xi(\delta_k) \in (m_1(\delta_k), m_1(\delta_k)/\delta_k)$  with  $m_1(\delta_k), m_1(\delta_k)/\delta_k \rightarrow m_1^*$  as  $k \rightarrow \infty$ , and  $g_2'$  is continuous by assumption, in the limit (??) becomes  $g_2(m_1^*) + m_1^* g_2'(m_1^*) = 0$ . Therefore,  $m_1^* = u_1^*$ .  $\square$

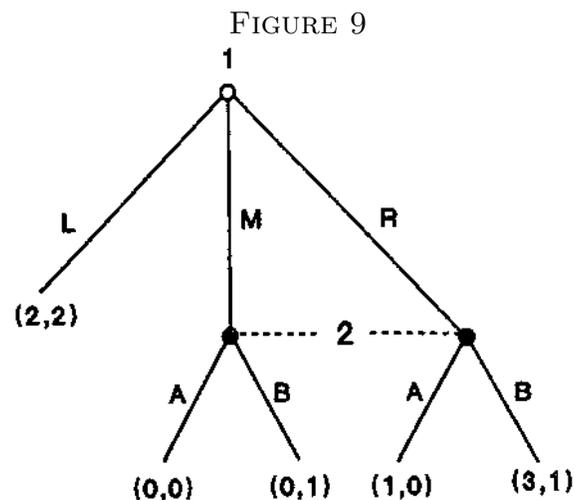
## 16. SEQUENTIAL EQUILIBRIUM

In multi-stage games with incomplete information, say where payoffs depend on initial moves by nature, the only subgame is the original game, even if players observe one another's actions at the end of each period. Thus the refinement of Nash equilibrium to subgame perfect equilibrium has no bite. Since players do not know each other's types, the continuation starting from a given period can be analyzed as a separate subgame only if we

<sup>12</sup>A simple graphical proof starts with the observation that  $m_1 g_2(m_1) = m_2 g_1(m_2)$ , hence the points  $(m_1, g_2(m_1))$  and  $(g_1(m_2), m_2)$  belong to the intersection of  $g_2$ 's graph with the same hyperbola, which approaches the hyperbola tangent to the boundary of  $U$  (at the Nash bargaining solution) as  $\delta \rightarrow 1$ .

have a specification of players' beliefs about which node they start at. The concept of sequential equilibrium provides a way to derive plausible beliefs at every information set. Based on the beliefs, one can test whether the continuation strategies form a Nash equilibrium.

The complications that incomplete information causes are evident in "signaling games," in which only one player has private information. The informed player moves first. The other player observes the informed player's action, but not her type, before choosing his own action. One example is Spence's (1974) model of the job market. In that model, a worker knows her productivity and must choose a level of education; a firm (or number of firms), observes the worker's education level, but not her productivity, and then decides what wage to offer her. In the spirit of subgame perfection, the optimal wage should depend on the firm's beliefs about the worker's productivity given the observed education. An equilibrium then needs to specify not only contingent actions, but also beliefs. At information sets that are reached with positive probability in equilibrium, beliefs should be derived using Bayes' rule. What about at information sets that are reached with probability zero? Some theoretical issues arise here.



Courtesy of The MIT Press. Used with permission.

Refer for more motivation to the example in Figure ?? (FT, p. 322). The strategy profile  $(L, A)$  is a Nash equilibrium, and it is subgame perfect, as player 2's information set does not initiate a subgame. However, it is not a very plausible equilibrium, since player 2 prefers playing  $B$  rather than  $A$  at his information set, regardless of whether player 1 has chosen

$M$  or  $R$ . So, a good equilibrium concept should rule out the solution  $(L, A)$  in this example and ensure that 2 always plays  $B$ .

For most definitions, we focus on extensive form games of perfect recall with finite sets of decision nodes. We use some of the notation introduced earlier.

To define *sequential equilibrium* (Kreps and Wilson 1982), we first define an *assessment* to be a pair  $(\sigma, \mu)$ , where  $\sigma$  is a (behavior) strategy profile and  $\mu$  is a *system of beliefs*. The latter component consists of a belief specification  $\mu(h)$  for each information set  $h$ ;  $\mu(h)$  is a probability distribution over the nodes in  $h$ . The definition of sequential equilibrium is based on the concepts of sequential rationality and consistency. *Sequential rationality* requires that conditional on every information set  $h$ , the strategy  $\sigma_{i(h)}$  be a best response to  $\sigma_{-i(h)}$  given the beliefs  $\mu(h)$ . Formally,

$$u_{i(h)}(\sigma_{i(h)}, \sigma_{-i(h)} | h, \mu(h)) \geq u_{i(h)}(\sigma'_{i(h)}, \sigma_{-i(h)} | h, \mu(h))$$

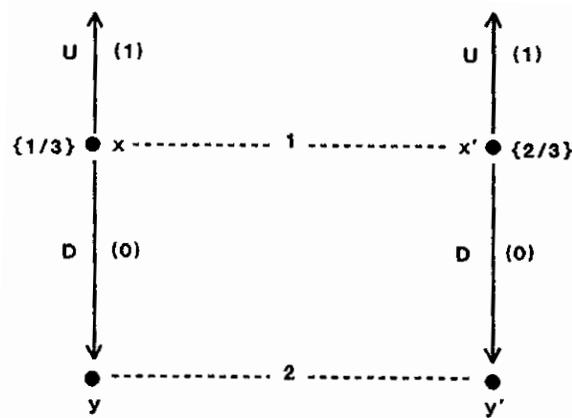
for all information sets  $h$  and alternative strategies  $\sigma'$ . Here, the conditional payoff  $u_i(\sigma | h, \mu(h))$  now denotes the payoff that results when play begins at a randomly selected node in the information set  $h$ , chosen according to the probability distribution  $\mu(h)$ , and subsequent play at each information set is as specified by the profile  $\sigma$ .

Beliefs need to be *consistent* with strategies in the following sense. For any *totally mixed* strategy profile  $\tilde{\sigma}$ —that is, one where each action is played with positive probability at every information set—all information sets are reached with positive probability, and Bayes' rule leads to a unique system of beliefs  $\mu^{\tilde{\sigma}}$ . The assessment  $(\sigma, \mu)$  is consistent if there exists a sequence of totally mixed strategy profiles  $(\sigma^m)_{m \geq 0}$  converging to  $\sigma$  such that the associated beliefs  $\mu^{\sigma^m}$  converge to  $\mu$  as  $m \rightarrow \infty$ .

**Definition 19.** A *sequential equilibrium* is an assessment that is sequentially rational and consistent.

The definition of sequential equilibrium rules out the strange equilibrium in the earlier example (Figure ??). Since player 1 chooses  $L$  under the proposed equilibrium strategies, consistency does not pin down player 2's beliefs at his information set. However, sequential rationality requires that player 2 have some beliefs and best-respond to them, which ensures that  $A$  is not played.

FIGURE 10



Courtesy of The MIT Press. Used with permission.

Consistency imposes more restrictions on the beliefs than Bayes' rule alone. Consider the partial extensive form from Figure ?? (FT, p. 339). The information set  $h_1$  of player 1 consists of two nodes  $x, x'$ . Player 1 can take an action  $D$  leading to  $y, y'$  respectively. Player 2 cannot distinguish between  $y$  and  $y'$  at the information set  $h_2$ . If 1 never plays  $D$  in equilibrium, then Bayes' rule does not pin down beliefs at  $h_2$ . However, consistency implies that  $\mu_2(y|h_2) = \mu_1(x|h_1)$ . The idea is that since 1 cannot distinguish between  $x$  and  $x'$ , he is equally likely to play  $D$  at either node. Hence consistency ensures that players' beliefs "respect the information structure."

More generally, consistency imposes *common beliefs* following deviations from equilibrium behavior. There are criticisms of this requirement—why should different players have the same theory about something that was not supposed to happen? A counter-argument is that consistency matches the spirit of equilibrium analysis, which normally assumes that players agree in their beliefs about other players' strategies (and moreover these beliefs are correct).

## 17. PROPERTIES OF SEQUENTIAL EQUILIBRIUM

**Theorem 17.** *A sequential equilibrium exists for every finite extensive-form game.*

This is a consequence of the existence of perfect equilibria, which we prove later.

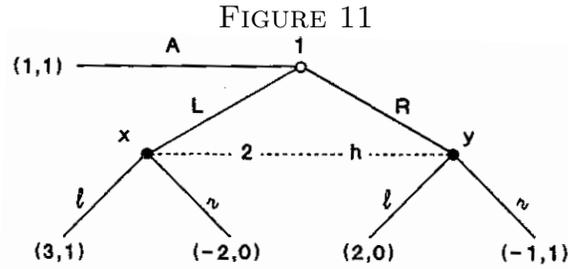
**Proposition 3.** *The sequential equilibrium correspondence has a closed graph with respect to payoffs.*

*Proof.* Let  $u^k \rightarrow u$  be a convergent sequence of payoff functions and  $(\sigma^k, \mu^k) \rightarrow (\sigma, \mu)$  be a convergent sequence of sequential equilibria of the games with corresponding payoffs  $u^k$ . We need to show that  $(\sigma, \mu)$  is a sequential equilibrium for the game with payoffs given by  $u$ . Sequential rationality of  $(\sigma, \mu)$  is straightforward because the expected payoffs conditional on reaching any information set are continuous in the payoff functions and beliefs.

We also have to check consistency of  $(\sigma, \mu)$ . As  $(\sigma^k, \mu^k)$  is a sequential equilibrium of the game with payoff function  $u^k$ , there exists a sequence of totally mixed strategies  $(\sigma^{m,k})_m \rightarrow \sigma^k$ , with corresponding induced beliefs given by  $(\mu^{m,k})_m \rightarrow \mu^k$ . For every  $k$ , we can find a sufficiently large  $m_k$  so that all components of  $\sigma^{m_k,k}$  and  $\mu^{m_k,k}$  are within  $1/k$  from the corresponding components of  $\sigma^k$  and  $\mu^k$ . Since  $\sigma^k \rightarrow \sigma, \mu^k \rightarrow \mu$ , it must be that  $\sigma^{m_k,k} \rightarrow \sigma, \mu^{m_k,k} \rightarrow \mu$ . Thus we have obtained a sequence of totally mixed strategies converging to  $\sigma$ , which induces beliefs converging to  $\mu$ .  $\square$

Kreps and Wilson show that in generic games (i.e., for a space of payoff functions such that the closure of its complement has measure zero, under any given tree structure), the set of outcome distributions that can be realized in some sequential equilibrium is finite. Nevertheless, it is not generally true that the set of sequential equilibria is finite, as there may be infinitely many belief specifications for off-path information sets that support some equilibrium strategies. It is not uncommon for the set of sequential equilibrium strategies to be infinite as well. Indeed, there may exist information sets off the equilibrium path that allow for consistent beliefs with the property that the corresponding players are indifferent between several actions. Many mixtures over the latter actions can be compatible with sequential rationality. See Figure ?? (FT, p. 359) for an example. That game has two sequential equilibrium outcomes,  $(L, l)$  and  $A$ . While there is a unique equilibrium leading to  $(L, l)$ , there are two one-parameter families of equilibria with outcome  $A$ . For  $A$  to be sequentially rational for player 1, it must be that 2 plays  $r$  with positive probability. In the first family of equilibria, player 2 chooses  $r$  with probability 1 and believes that  $\mu(x) < 1/2$ . In the second, player 2 chooses  $r$  with a probability in  $[2/5, 1]$  and believes that  $\mu(x) = 1/2$ .

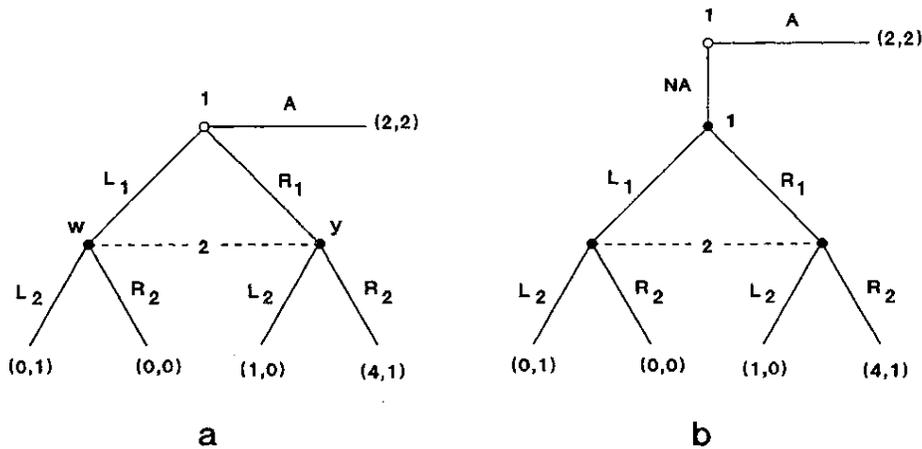
Kohlberg and Mertens (1986) criticized sequential equilibrium for allowing “strategically neutral” changes in the game tree to affect the equilibrium. Compare, for instance, the two games in Figure ?? (FT, p. 343). The game on the right is identical to the one on the left,



Courtesy of The MIT Press. Used with permission.

except that player 1's first move is split into two moves in a seemingly irrelevant way. Whereas  $(A, L_2)$  can be supported as a sequential equilibrium for the game on the left, the strategy  $A$  is not part of a sequential equilibrium for the one on the right. For the latter game, in the simultaneous-move subgame following  $NA$ , the only Nash equilibrium is  $(R_1, R_2)$ , as  $L_1$  is strictly dominated by  $R_1$  for player 1. Hence the unique sequential equilibrium strategies for the right-hand game are  $((NA, R_1), R_2)$ . This example demonstrates that eliminating a strictly dominated strategy affects the set of sequential equilibria.

FIGURE 12



Courtesy of The MIT Press. Used with permission.

Note that the sensitivity of sequential equilibrium to the addition of “irrelevant moves” is not a direct consequence of consistency, but is rather implied by sequential rationality. In the example above, the problem arises even for subgame perfect equilibria. Kohlberg and Mertens (1986) further develop these ideas in their concept of a stable equilibrium. However, their criticism that good mistakes should be much more likely than bad mistakes is not necessarily compelling. If we take seriously the idea that players make mistakes at each information set, then it is not clear that the two extensive forms in the above example

are equivalent. In the game on the right, if player 1 makes the mistake of not playing  $A$ , he is still able to ensure that  $R_1$  is more likely than  $L_1$ ; in the game on the left, he might take either action by mistake when intending to play  $A$ .

## 18. PERFECT BAYESIAN EQUILIBRIUM

Perfect Bayesian equilibrium was the original solution concept for extensive-form games with imperfect information, when subgame perfection does not have enough force. It incorporated the ideas of sequential rationality and Bayesian updating of beliefs. Nowadays sequential equilibrium (which was invented later) is the preferred way of expressing these ideas, but it's worthwhile to know about PBE since older papers refer to it.

The idea is similar to sequential equilibrium but with simpler requirements about how beliefs are updated. Fudenberg and Tirole (1991) have a paper that describes various formulations of PBE. The basic requirements are that strategies should be sequentially rational and that beliefs should be derived from Bayes' rule wherever applicable, with no constraints on beliefs at information sets reached with probability zero in equilibrium.

PBE is typically applied to multi-stage games with incomplete information in which nature assigns types to players and player actions are observed at the end of every stage. Here are some other restrictions that can be imposed for such games.

- If player types are drawn independently by nature, beliefs about different players should remain independent at each history.
- Updating should be "consistent": given a probability-zero history  $h^t$  at time  $t$ , at which strategies call for a positive-probability transition to history  $h^{t+1}$ , the beliefs at  $h^{t+1}$  should be given by updating beliefs at  $h^t$  via Bayes' rule.
- "Not signaling what you don't know": in multi-stage games with independent types, beliefs about player  $i$  at the beginning of period  $t + 1$  depend only on  $h^t$  and action by player  $i$  at time  $t$ , not also on other players' actions at time  $t$ .
- Two different players  $i, j$  should have the same belief about a third player  $k$  even at probability-zero histories.

All of these conditions are implied by consistency.

Anyhow, there does not seem to be a single clear definition of PBE in the literature. Different sets of conditions are imposed by different authors. This is one more reason why using sequential equilibrium is preferable.

### 19. PERFECT EQUILIBRIUM

Now consider the following game:

	L	R
U	1, 1	0, 0
D	0, 0	0, 0

Both  $(U, L)$  and  $(D, R)$  are sequential equilibria (sequential equilibrium coincides with Nash equilibrium in a strategic-form game). But  $(D, R)$  seems non-robust: if player 1 thinks that player 2 might make a mistake and play  $L$  with some small probability, he would rather deviate to  $U$ . This motivates the definition of (*trembling-hand*) *perfect equilibrium* (Selten, 1975) for strategic-form games. A profile  $\sigma$  is a PE if there is a sequence of “trembles”  $\sigma^m \rightarrow \sigma$ , where each  $\sigma^m$  is a totally mixed strategy, such that  $\sigma_i$  is a best reply to  $\sigma_{-i}^m$  for each  $m$  and all  $i \in N$ .

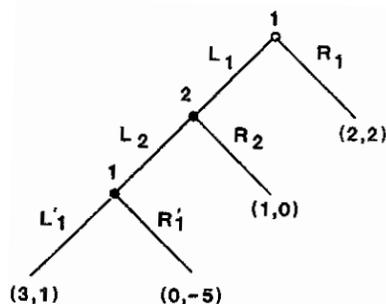
An equivalent approach is to define a strategy profile  $\sigma^\varepsilon$  to be an  $\varepsilon$ -*perfect equilibrium* if there exist  $\varepsilon(s_i) \in (0, \varepsilon)$  for all  $i$  and  $s_i \in S_i$  such that  $\sigma^\varepsilon$  is a Nash equilibrium of the game where players are restricted to play mixed strategies in which every pure strategy  $s_i$  has probability at least  $\varepsilon(s_i)$ . A PE is a profile that is a limit of some sequence of  $\varepsilon$ -perfect equilibria  $\sigma^\varepsilon$  as  $\varepsilon \rightarrow 0$ . (We will not show the equivalence here but it’s not too hard.)

**Theorem 18.** *Every finite strategic-form game has a perfect equilibrium.*

*Proof.* For any  $\varepsilon > 0$ , we can certainly find a Nash equilibrium of the modified game, where each player is restricted to play mixed strategies that place probability at least  $\varepsilon$  on every pure strategy. (Just apply the usual Nash existence theorem for compact strategy sets and quasiconcave payoffs.) By compactness, there is some subsequence of these strategy profiles as  $\varepsilon \rightarrow 0$  that converges, and the limit point is a perfect equilibrium by definition.  $\square$

We would like to extend this definition to extensive-form games. Consider the game in Figure ?? (FT, p. 353). They show an extensive-form game and its reduced normal form.

FIGURE 13



a. Extensive Form

	L <sub>2</sub>	R <sub>2</sub>
R <sub>1</sub>	2,2	2,2
L <sub>1</sub> ,L' <sub>1</sub>	3,1	1,0
L <sub>1</sub> ,R' <sub>1</sub>	0,-5	1,0

b. Strategic Form

Courtesy of The MIT Press. Used with permission.

There is a unique SPE  $(L_1L'_1, L_2)$ . But  $(R_1, R_2)$  is a PE of the reduced normal form. Thus perfection in the normal form does *not* imply subgame-perfection. The perfect equilibrium is sustained only by trembles such that, conditional on trembling to  $L_1$  at the first node, player 1 is also fairly likely to play  $R'_1$  rather than  $L'_1$  at his second node. This seems unreasonable— $R'_1$  is only explainable as a tremble. Perfect equilibrium as defined so far thus has the disadvantage of allowing correlation in trembles at different information sets.

The solution to this is to impose perfection in the *agent-normal form*. We treat the two different nodes of player 1 as being different players, thus requiring them to tremble independently. More formally, in the agent-normal form game, we have a player corresponding to every information set. Given a strategy profile in this game, the “player” corresponding to any information set  $h$  enjoys the same payoffs as player  $i(h)$  under the corresponding strategies in the extensive-form game. Thus, the game in Figure ?? turns into a three-player game. The only perfect equilibrium of this game is  $(L_1, L'_1, L_2)$ .

More generally, a *perfect equilibrium* for an extensive form game is defined to be a perfect equilibrium of the corresponding agent-normal form.

**Theorem 19.** *Every PE of a finite extensive-form game is a sequential equilibrium (for some appropriately chosen beliefs).*

*Proof.* Let  $\sigma$  be a PE of the extensive-form game. Then there exist totally mixed strategy profiles in the agent normal form game  $\sigma^m \rightarrow \sigma$  such that  $\sigma_h$  is a best reply to  $\sigma_{-h}^m$  for each  $m$  and all information sets  $h$ . For each  $\sigma^m$  we have a well-defined belief system  $\mu^m$  induced by Bayes' rule. Pick a subsequence for which these belief systems converge to some  $\mu$ . Then by definition  $(\sigma, \mu)$  is consistent. Sequential rationality follows exactly from the fact that  $\sigma_h$  is a best reply given  $\mu^m(\cdot|h)$  and  $\sigma_{-h}^m$  for each  $m$  along the subsequence, and hence also with respect to the limit beliefs  $\mu(\cdot|h)$  and strategies  $\sigma_{-h}$ . (More properly, perfection implies that there are no one-shot deviations that benefit any player; by an appropriate adaptation of the one-shot deviation principle we infer that  $\sigma$  is in fact sequentially rational at every information set.)  $\square$

The converse is not true—not every sequential equilibrium is perfect, as we already saw with the simple strategic-form example above. But for generic payoffs it is true (Kreps and Wilson, 1982).

The set of perfect equilibrium outcomes does not have a closed graph (unlike sequential or subgame-perfect equilibrium). Consider the following game:

	L	R
U	1, 1	0, 0
D	0, 0	$1/n, 1/n$

It has  $(D, R)$  as a perfect equilibrium for each  $n > 0$ , but in the limit where  $(D, R)$  has payoffs  $(0, 0)$  it is no longer a perfect equilibrium. We can think of this as an order-of-limits problem: as  $n \rightarrow \infty$  the trembles against which  $D$  and  $R$  remain best responses become smaller and smaller. Thus, whether or not  $(D, R)$  is a reasonable prediction in the limiting game depends on whether we think the approximation error in describing the payoffs is larger than the players' probability of trembling or vice versa.

## 20. PROPER EQUILIBRIUM

Myerson (1978) considered the notion that when a player trembles, he is still more likely to play better actions than worse ones. Specifically, a player's probability of playing the

second-best action is at most  $\varepsilon$  times the probability of the best action, the probability of the third-best action is at most  $\varepsilon$  times the probability of the second-best action, and so forth. Consider the game in Fig. 8.15 of FT (p. 357).  $(M, M)$  is a perfect equilibrium, but Myerson argues that it can be supported only using unreasonable trembles, where each player has to be likely to tremble to a very bad reply rather than an almost-best reply.

**Definition 20.** An  $\varepsilon$ -proper equilibrium is a totally mixed strategy profile  $\sigma^\varepsilon$  such that, if  $u_i(s_i, \sigma_{-i}^\varepsilon) < u_i(s'_i, \sigma_{-i}^\varepsilon)$ , then  $\sigma_i^\varepsilon(s_i) \leq \varepsilon \sigma_i^\varepsilon(s'_i)$ . A proper equilibrium is any limit of some  $\varepsilon$ -proper equilibria as  $\varepsilon \rightarrow 0$ .

**Theorem 20.** Every finite strategic-form game has a proper equilibrium.

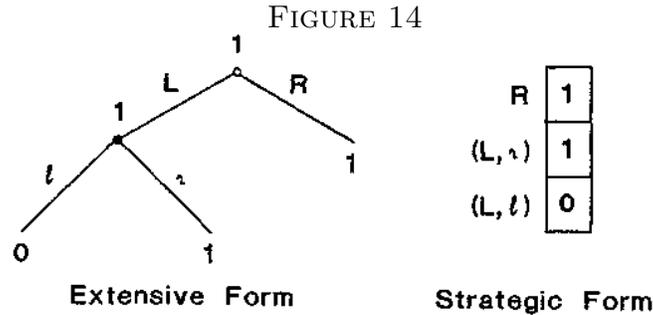
*Proof.* First prove existence of  $\varepsilon$ -proper equilibria, using the usual Kakutani argument applied to the “almost-best-reply” correspondences  $BR_i^\varepsilon$  rather than the usual best-reply correspondences. ( $BR_i^\varepsilon(\sigma_{-i})$  is the set of mixed strategies for player  $i$  in a suitable compact space of totally mixed strategies that satisfy the inequality in the definition of  $\varepsilon$ -proper equilibrium.) Then use compactness to show that there exists a sequence that converges as  $\varepsilon \rightarrow 0$ ; its limit is a proper equilibrium.  $\square$

Given an extensive-form game, a proper equilibrium of the corresponding normal form is automatically subgame-perfect; we don't need to go to the agent-normal form. We can show this by a backward-induction-type argument.

Kohlberg and Mertens (1986) showed that a proper equilibrium in a strategic-form game is sequential in every extensive-form game having the given normal form. However, it will not necessarily be a trembling-hand perfect equilibrium in (the agent-normal form of) every such game. See Figure ?? (FT, p. 358):  $(L, r)$  is proper (and so sequential) in the normal form but not perfect in the agent-normal form.

## 21. FORWARD INDUCTION

The preceding ideas are all attempts to capture some kind of *forward induction*: players should believe in the rationality of their opponents, even after observing a deviation; thus if you observe an out-of-equilibrium action being played, you should believe that your opponent expected you to respond in a way such that his action was reasonable, and this in turn is



Courtesy of The MIT Press. Used with permission.

informative about his type (or, in more general extensive forms, about how he plans to play in the future). Forward induction is not itself an equilibrium concept, since in equilibrium all players know that the specified strategies are to be exactly followed; rather, it is an attempt to describe reasoning by players who are not quite certain about what will be played. It also is not a single, rigorously defined concept, but rather a vague term for a form of reasoning about play.

Consider now the extensive-form game as follows: 1 can play  $O$ , leading to  $(2, 2)$ , or  $I$ , leading to the following battle-of-the-sexes game:

	T	W
T	0, 0	3, 1
W	1, 3	0, 0

There is an SPE in which player 1 first plays  $O$ ; conditional on playing  $I$ , they play the equilibrium  $(W, T)$ . But the following forward-induction argument suggests this equilibrium is unreasonable: if player 1 plays  $I$ , this suggests he is expecting to coordinate on  $(T, W)$  in the battle-of-the-sexes game, so player 2, anticipating this, will play  $W$ . If 1 can thus convince 2 to play  $W$  by playing  $I$  in the first stage, he can get the higher payoff  $(3, 1)$ .

This game can also be represented in (reduced) normal form.

	T	W
O	2, 2	2, 2
IT	0, 0	3, 1
IW	1, 3	0, 0

This representation of the game shows a connection between forward induction and strict dominance. We can rule out  $IW$  because it is dominated by  $O$ ; then the only perfect equilibrium of the remaining game is  $(IT, W)$  giving payoffs  $(3, 1)$ . However,  $(O, T)$  can be enforced as a perfect (in fact a proper) equilibrium in the original strategic-form game.

Kohlberg and Mertens (1986) argue that an equilibrium concept that is not robust to deletion of strictly dominated strategies is troubling. The above example, together with other cases of such non-robustness, leads them to define the notion of *stable equilibria*. It is a set-valued concept—not a property of an individual equilibrium but of sets of equilibria. Kohlberg and Mertens first argue that a solution concept should meet the following requirements:

- Iterated dominance: every stable set must contain a stable set of any game obtained by deleting a strictly dominated strategy.
- Admissibility: no mixed strategy appearing in a stable set assigns positive probability to a *weakly* dominated strategy.
- Invariance to extensive-form representation: they define an equivalence relation between extensive forms and require that any stable set in one game should be stable in any equivalent game.

Kohlberg and Mertens define strategic stability in a way such that these criteria are satisfied.

**Definition 21.** A closed set  $S$  of NE is *strategically stable* if it is minimal among sets with the following property: for every  $\eta > 0$ , there exists  $\varepsilon > 0$  such that all choices of  $\varepsilon(s_i) \in (0, \varepsilon)$ , the game where each player  $i$  is constrained to play every  $s_i$  with probability at least  $\varepsilon(s_i)$  has a Nash equilibrium which is within distance  $\eta$  of some equilibrium in  $S$ .

Any sequence of  $\varepsilon$ -perturbed games as  $\varepsilon \rightarrow 0$  should have equilibria corresponding to an equilibrium in  $S$ . Notice that we need the minimality property of  $S$  to give bite to this definition—otherwise, by upper hemi-continuity, we know that the set of all Nash equilibria would be strategically stable, and we get no refinement. The difference with trembling-hand perfection is that there should be convergence to one of the equilibria in  $S$  for *any* sequence of perturbations, not just some sequence of perturbations.

**Theorem 21.** *There exists a stable set that is contained in a connected component of the set of Nash equilibria. Generically, each component of the set of Nash equilibria leads to a*

single distribution over outcomes, so there exists a stable set that induces a unique outcome distribution. A stable set contains a stable set of any game obtained by eliminating a weakly dominated strategy and also of any game obtained by deleting a strategy that is not a best response to any of the opponents' strategy profiles in the set (“never a weak best reply” (NWBR)).

NWBR shows that the concept of stable equilibrium is robust to forward induction: knowing that player  $i$  will not use a particular strategy is consistent with the equilibrium theories from the stable set.

Every equilibrium in a stable set has to be a perfect equilibrium. This follows from the minimality condition—if an equilibrium is not a limiting equilibrium along some sequence of trembles, then there’s no need to include it in the stable set. But notice, these equilibria are only guaranteed to be perfect in the normal form, not in the agent-normal form (for a given extensive-form game). Actually, an example due to Gul proves that there exist stable sets that do not contain any sequential equilibrium.

Normal form games do not directly capture the reasoning of forward induction. Battigalli and Siniscalchi (2002) seek a general-purpose definition of forward induction by modeling the belief revision process explicitly in the context of extensive form games. They are interested in the epistemic conditions that lead to forward induction.

They propose an epistemic model, with each player having a set  $\Omega_i$  of states of the form  $\omega_i = (s_i, t_i)$ , where  $s_i$  represents player  $i$ 's *disposition to act* and  $t_i$  represents his *disposition to believe*.  $s_i$  specifies his action at each information set  $h$  of player  $i$ , and  $t_i$  specifies a belief  $g_{i,h} \in \Delta(\Omega_{-i})$  over states of the other players for each  $h$ . We say  $i$  is rational at state  $\omega$  if  $s_i$  is a best reply to his beliefs  $t_i$  at each information set. Let  $R$  be the set of states at which every player is rational. For any event  $E \subseteq \Omega$ , we can define the set  $B_{i,h}(E) = \{(s, t) \in \Omega \mid g_{i,h}(E) = 1\}$ , i.e. the set of states where  $i$  is sure that  $E$  has occurred (at information set  $h$ ). Finally  $SB_i(E) = \cap_{h \text{ reachable given } E} B_{i,h}(E)$  denotes the set of states at which  $i$  *strongly believes* in event  $E$ , meaning the set of states at which  $i$  would be sure of  $E$  as long as he’s reached an information set that is consistent with  $E$  occurring. Finally, Battigalli and Siniscalchi show that  $SB(R)$  identifies forward induction—that is, in the states of the world where everyone strongly believes that everyone is sequentially

rational, strategies must form a profile that is not ruled out by forward induction arguments of the sort discussed earlier.

Battigalli and Siniscalchi take this a level further by iterating the strong-beliefs operator—everyone strongly believes that everyone strongly believes that everyone is rational, and so forth—and this operator leads to backward induction in games of perfect information; without perfect information, it leads to iterated deletion of strategies that are never a best reply. This leads to a formalization of the idea of rationalizability in extensive-form games.

## 22. FORWARD INDUCTION IN SIGNALING GAMES

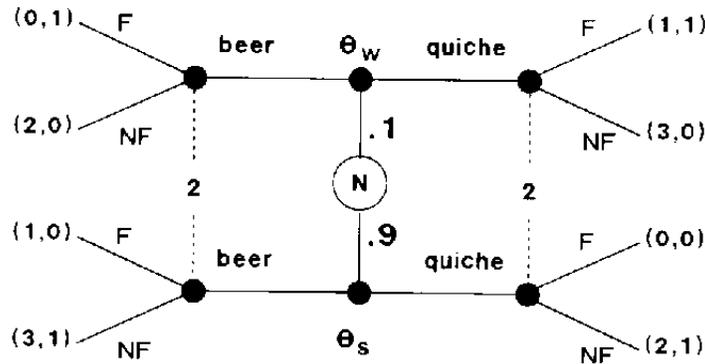
The NWBR property is a useful way to show that some components of equilibria are not stable. For instance, Cho and Kreps (1987) developed an equilibrium refinement for signaling games that is weaker than stability—the intuitive criterion—based on iterated applications of NWBR. Kohlberg and Mertens (1986) motivated their stability concept by mathematical properties they deemed desirable and robustness with respect to trembles a la Selten’s perfect equilibrium. By contrast, Cho and Kreps provided a behavioral foundation based on refining the set of plausible beliefs in the spirit of Kreps and Wilson’s sequential equilibrium.

In a *signaling game*, there are two players, a sender  $S$  and a receiver  $R$ . There is a set  $T$  of types for the sender; the realized type will be denoted by  $t$ .  $p(t)$  denotes the probability of type  $t$ . The sender privately observes his type  $t$ , then sends a message  $m \in M(t)$ . The receiver observes the message and chooses an action  $a \in A(m)$ . Finally both players receive payoffs  $u_S(t, m, a), u_R(t, m, a)$ ; thus the payoffs potentially depend on the true type, the message sent, and the action taken by the receiver. A signaling game is parameterized by the set  $T$ , the prior  $p$ , the feasible message and action correspondences  $M, A$ , and the payoff functions  $u_S, u_R$ . In such a game we will use  $T(m)$  to denote the set  $\{t \mid m \in M(t)\}$ .

Cho and Kreps’ intuitive criterion provides a behavioral explanation of one aspect of the NWBR property: robustness to replacing the equilibrium path by its expected payoff. This solution presumes that players are certain about play along the equilibrium path, but there is uncertainty about play off the path. If we begin with a stable set and then, using NWBR, delete a strategy in which type  $t$  plays an action  $m$ , the reduced game should have a stable component contained in the original component. This means that the surviving equilibria should assign probability zero to type  $t$  following message  $m$ .

Consider the beer-quiche signaling game from Figure ?? (FT, p. 450). In this example, player 1 is wimpy or surly, with respective probabilities 0.1 or 0.9. Player 2 is a bully who would like to fight the wimpy type but not the surly one. Player 1 orders breakfast and 2 decides whether to fight him after observing his breakfast choice. In the notation above,  $T = \{weak, surly\}$ ;  $M = M(t) = \{beer, quiche\}, \forall t \in T$ ;  $A(m) = \{fight, not\ fight\}, \forall m \in M$ .

FIGURE 15



Courtesy of The MIT Press. Used with permission.

Player 1 gets a utility of 1 from having his favorite breakfast—beer if surly, quiche if weak—but a disutility of 2 from fighting. When player 1 is weak, player 2’s utility is 1 if he fights and 0 otherwise; when 1 is surly, the payoffs to the two actions are reversed.

One can show that all equilibria involve pooling. The key idea in this game is to compare  $\sigma_2(F|beer)$  and  $\sigma_2(F|quiche)$ . The breakfast leading to a smaller probability of fighting must be selected with probability 1 in equilibrium by the type of player 1 who likes it. . . We find that there are two classes of sequential equilibria, corresponding to two distinct outcomes. In one set of sequential equilibria, both types of player 1 drink beer, while in the other both types of player 1 eat quiche. In both cases, player 2 must fight with probability at least 1/2 when observing the out-of-equilibrium breakfast since otherwise one of the two types of player 1 would want to deviate to the other breakfast. Note that either type of equilibrium can be supported with any belief for player 2 placing a probability weight of at least 1/2 on player 1 being wimpy following the out-of-equilibrium breakfast (hence there is an infinity of sequential equilibrium assessments).

Cho and Kreps argued that the equilibrium in which both types choose quiche is unreasonable for the following reason. It does not make any sense for the weak type to deviate

to ordering beer, no matter how he thinks that the receiver will react, because he is already getting payoff 3 from quiche, whereas he cannot get more than 2 from switching to beer. On the other hand, the surly type can benefit if he thinks that the receiver will react by not fighting. Thus, conditional on seeing beer ordered, the receiver should conclude that the sender is surly and so should not want to fight. Clearly, the class of equilibria in which both types choose quiche violates NWBR.

On the other hand, this argument does not rule out the equilibrium in which both types drink beer. In this case, in equilibrium the surly type is getting 3, whereas he gets at most 2 from deviating no matter how the receiver reacts; hence he cannot want to deviate. The weak type, on the other hand, is getting 2, and he can get 3 by switching to quiche if he thinks this will induce the receiver not to fight him. Thus only the weak type would deviate, so the sender's belief (that the receiver is weak if he orders quiche) is reasonable.

Now consider modifying the game by adding an extra option for the receiver: paying a million dollars to the sender. Now the preceding argument doesn't rule out the quiche equilibrium—either type of sender might deviate to beer if he thinks this will induce the receiver to pay him a million dollars. Hence, in order to apply the same argument, we need to make the additional assumption that the sender cannot expect the receiver to play a bad strategy. Then we can restrict attention to the game without the million-dollar option, and the argument goes through.

Cho and Kreps formalized this line of reasoning in the *intuitive criterion*, as follows. For any set of types  $T' \subseteq T$  and any message  $m$ , write

$$BR(T', m) = \cup_{\mu \mid \mu(T')=1} BR(\mu, m)$$

—the set of strategies that  $R$  could rationally play if he observes  $m$  and is sure that the sender's type is in  $T'$ . Now with this notation established, consider any sequential equilibrium, and let  $u_S^*(t)$  be the equilibrium payoff to a sender of type  $t$ . Define

$$\tilde{T}(m) = \{t \mid u_S^*(t) > \max_{a \in BR(T(m), m)} u_S(t, m, a)\}.$$

This is the set of types that do better in equilibrium than they could possibly do by sending  $m$ , no matter how  $R$  reacts, as long as  $R$  is playing a best response to some belief. We then say that the proposed equilibrium fails the intuitive criterion if there exist a type  $t'$  and a

message  $m \in M(t')$  such that

$$u_S^*(t') < \min_{a \in BR(T(m) \setminus \tilde{T}(m), m)} u_S(t', m, a).$$

In words, the equilibrium fails the intuitive criterion if some type  $t'$  of the sender is getting less than any payoff he could possibly get by playing  $m$ , assuming he could thereby convince the sender that he could not possibly be in  $\tilde{T}(m)$ .

In the beer-quiche example, the all-quiche equilibrium fails this criterion: let  $t' = \textit{surly}$  and  $m = \textit{beer}$ ; check that  $\tilde{T}(m) = \{\textit{weak}\}$ .

We can apply this procedure repeatedly, giving the *iterated intuitive criterion*. Given a proposed equilibrium, we can use the intuitive criterion as above to rule out pairs  $(t, m)$ —types  $t$  that cannot conceivably want to send message  $m$ . We can then rule out some actions of the receiver, by requiring that the receiver should best respond to a belief about the types that have not yet been eliminated given the message. Under the surviving actions, we can possibly rule out more pairs  $(t, m)$ , and so forth.

This idea has been further developed by Banks and Sobel (1987). They say that type  $t'$  is *infinitely more likely* to choose the out-of-equilibrium message  $m$  than type  $t$  under the following condition: the set of possible best-replies by the receiver (possibly mixed) that make  $t'$  strictly prefer to deviate to  $m$  is a strict superset of the possible best-replies that make  $t$  weakly prefer to deviate. If this holds, then conditional on observing  $m$ , the receiver should put probability 0 on type  $t$ . The analogue of the Intuitive Criterion under this elimination procedure is known as D1. If we allow  $t'$  to vary across different best replies by the sender, requiring only that every best response that weakly induced  $t$  to deviate would also strictly induce *some*  $t'$  to deviate, then this gives criterion D2. We can also apply either of these restrictions on beliefs to eliminate possible actions by the receiver, and proceed iteratively. Iterating D2 leads to the equilibrium refinement criterion known as *universal divinity*.

The motivating application is Spence's job-market signaling model. With just two types of job applicant, the intuitive criterion selects the equilibrium where the low type gets the lowest level of education and the high type gets just enough education to deter the low type from imitating her. With more types, the intuitive criterion no longer accomplishes this. D1 does manage to uniquely select the socially optimal separating equilibrium by having each type get just enough education to deter the next-lower type from imitating her.

## 23. THE SPENCE SIGNALING MODEL

We next consider Spence's (1973) signaling model of the job market.<sup>13</sup> An employer faces a worker of unknown ability  $\theta$ . The ability of the worker is known to the worker though, and is either  $\theta = H$  or  $\theta = L$ , where  $H > L > 0$ . Interpret these numbers as the money value of what the worker would produce working in the firm.

The worker would like to transmit the knowledge of her ability to the firm; the problem is how to do so in a credible way. Think of education as just such a device.

**23.1. The Game.** Specifically, suppose that a worker can choose to acquire  $e$  units of education, where  $e$  is any nonnegative number. Of course, the worker will have to study hard to obtain her education, and this creates disutility (studying for exams, doing homework, etc.). Assume that a worker of true ability  $\theta$  expends  $e/\theta$  in disutility. The point is, then, that  $H$ -types can acquire education easier than  $L$ -types.

The game proceeds as follows:

1. Nature moves and chooses a worker type,  $H$  or  $L$ . The type is revealed to the worker but not to the employer.
2. The worker then chooses  $e$  units of education. This is perfectly observed by the employer.
3. The employer observes  $e$  and forms an estimate of  $\theta$ . He then pays the worker a salary equal to this estimate, which is just the conditional expectation of  $\theta$  given  $e$ , written  $\mathbb{E}(\theta|e)$ .
4. The  $H$ -worker's payoff is  $\mathbb{E}(\theta|e) - e/H$ , and the  $L$ -worker's payoff is  $\mathbb{E}(\theta|e) - e/L$ .

The game is set up so simply that the employer's expected payoff is zero. Essentially, we assume that the worker's choice of education is visible to the world at large so that perfect competition must push her wage to  $\mathbb{E}(\theta|e)$ , the conditional expectation of  $\theta$  given  $e$ .

*Very Important.* Note well that  $\mathbb{E}(\theta|e)$  is not just a given but it is an endogenous object derived from strategies. How it is computed will depend on worker strategies and how they translate in beliefs.

**23.2. Single Crossing.** Suppose that a worker of type  $\theta$  uses a probability distribution  $\mu_\theta$  over different education levels. First observe that if  $e$  is a possible choice of the high worker

---

<sup>13</sup>This exposition is based on notes by Debraj Ray (<http://www.econ.nyu.edu/user/debraj/Courses/05UGGameLSE/Handouts/05uggl10.pdf>).

and  $e'$  a possible choice of the low worker, then it must be that  $e \geq e'$ . This follows from the following important *single-crossing* argument.

The  $H$ -type *could* have chosen  $e'$  instead of  $e$ , so

$$(23.1) \quad \mathbb{E}(\theta|e) - \frac{e}{H} \geq \mathbb{E}(\theta|e') - \frac{e'}{H},$$

while the  $L$ -type *could* have chosen  $e$  instead of  $e'$ , so

$$(23.2) \quad \mathbb{E}(\theta|e') - \frac{e'}{L} \geq \mathbb{E}(\theta|e) - \frac{e}{L}.$$

Adding both sides in (??) and (??), we see that

$$(e - e') \left( \frac{1}{L} - \frac{1}{H} \right) \geq 0.$$

Because  $1/L > 1/H$ , it follows that  $e \geq e'$ .

Essentially, if the low type *weakly* prefers a higher education to a lower one, the high type would *strictly* prefer it. So a high type can never take strictly less education than a low type in equilibrium.

This sort of result typically follows from the assumption that being a high type reduces not just the *total* cost from taking an action but also the *marginal* cost of that action; in this case, of acquiring one more unit of education. As long as this feature is present, we could replace the cost function  $e/\theta$  by any cost function and the same analysis goes through.

**23.3. Equilibrium.** Now that we know that the high type will not invest any less than the low type, we are ready to describe the equilibria of this model. There are three kinds of equilibria here; the concepts are general and apply in many other situations.

1. *Separating Equilibrium.* Each type takes a different action, and so the equilibrium action reveals the type perfectly. It is obvious that in this case,  $L$  must choose  $e = 0$ , for there is nothing to be gained in making a positive effort choice.

What about  $H$ ? Note: she cannot play a mixed strategy because each of her actions fully reveals her type, so she might as well choose the least costly of those actions. So she chooses a single action: call it  $e^*$ , and obtains a wage equal to  $H$ . Now these are the crucial incentive constraints; we must have

$$(23.3) \quad H - \frac{e^*}{L} \leq L,$$

otherwise the low person will try to imitate the high type, and

$$(23.4) \quad H - \frac{e^*}{H} \geq L,$$

otherwise the high person will try to imitate the low type.

Look at the smallest value of  $e^*$  that just about satisfies (??); call it  $e_1$ . And look at the largest value of  $e^*$  that just about satisfies (??); call it  $e_2$ . Clearly,  $e_1 < e_2$ , so the two restrictions above are compatible.

Any outcome in which the low type chooses 0 and the high type chooses some  $e^* \in [e_1, e_2]$  is supportable as a separating equilibrium. To show this we must also specify the beliefs of the employer. There is a lot of leeway in doing this. Here is one set of beliefs that works: the employer believes that any  $e < e^*$  (if observed) comes from the low type, while any  $e > e^*$  (if observed) comes from the high type. These beliefs are consistent because sequential equilibrium in this model imposes no restrictions on off-the-equilibrium beliefs. Given these beliefs and equations (??) and (??), we can check that no type has incentives to deviate.

2. *Pooling Equilibrium.* There is also a family of pooling equilibria in which only one signal is received in equilibrium. It is sent by both types, so the employer learns nothing new about the types. So if it sees that signal — call it  $e^*$  — it simply pays out the expected value calculated using the prior beliefs:  $pH + (1 - p)L$ .

Of course, for this to be an equilibrium two conditions are needed. First, we need to specify employer beliefs off the equilibrium path. Again, a wide variety of such beliefs are compatible; here is one: the employer believes that any action  $e \neq e^*$  is taken by the low type. [It does not have to be this drastic.<sup>14</sup>] Given these beliefs, the employer will “reward” any signal not equal to  $e^*$  with a payment of  $L$ . So for the types not to deviate, it must be that

$$pH + (1 - p)L - \frac{e^*}{\theta} \geq L,$$

but the binding constraint is clearly for  $\theta = L$ , so rewrite as

$$pH + (1 - p)L - \frac{e^*}{L} \geq L.$$

---

<sup>14</sup>For instance, the employer might believe that any action  $e < e^*$  is taken by the low type, while any action  $e > e^*$  is taken by types in proportion to their likelihood:  $p : 1 - p$ .

This places an upper bound on how big  $e^*$  can be in any pooling equilibrium. Any  $e^*$  between 0 and this bound will do.

3. *Hybrid Equilibria.* There is also a class of “hybrid equilibria” in which one or both types randomize. For instance, here is one in which the low type chooses 0 while the high type randomizes between 0 (with probability  $q$ ) and some  $e$  with probability  $1 - q$ . If the employer sees  $e$  he knows the type is high. If he sees 0 the posterior probability of the high type there is — by Bayes’ Rule — equal to

$$\frac{qp}{qp + (1 - p)},$$

and so the employer must pay out a wage of precisely

$$\frac{qp}{qp + (1 - p)}H + \frac{1 - p}{qp + (1 - p)}L.$$

But the high type must be *indifferent* between the announcement of 0 and that of  $e$ , because he willingly randomizes. It follows that

$$\frac{qp}{qp + (1 - p)}H + \frac{1 - p}{qp + (1 - p)}L = H - \frac{e}{H}.$$

To complete the argument we need to specify beliefs everywhere else. This is easy as we’ve seen more than once (just believe that all other  $e$ -choices come from low types). We therefore have a hybrid equilibrium that is “semi-separating”.

In the Spence model all three types of equilibria coexist. Part of the reason for this is that beliefs can be so freely assigned off the equilibrium path, thereby turning lots of outcomes into equilibria. What we turn to next is a way of narrowing down these beliefs. To be sure, to get there we have to go further than just sequential equilibrium.

**23.4. The Intuitive Criterion.** Consider a sequential equilibrium and a non-equilibrium announcement (such as a nonequilibrium choice of education in the example above). What is the other recipient of such a signal (the employer in the example above) to believe when she sees that signal?

Sequential equilibrium imposes little or no restrictions on such beliefs in signalling models. [We have seen, of course, that in other situations — such as those involving moves by Nature — that it does impose several restrictions, but not in the signalling games that we have been

studying.] The purpose of the Intuitive Criterion is to try and narrow beliefs further. In this way we eliminate some equilibria and in so doing sharpen the predictive power of the model.

Consider some non-equilibrium signal  $e$ . Consider some type of a player, and suppose even if she were to be treated in the best possible way following the emission of the signal  $e$ , she *still* would prefer to stick to her equilibrium action. Then we will say that signal  $e$  is *equilibrium-dominated* for the type in question. She would never want to emit that signal, except purely by error. Not strategically.

The Intuitive Criterion (IC) may now be stated.

*If, under some ongoing equilibrium, a non-equilibrium signal is received which is equilibrium-dominated for some types but not others, then beliefs cannot place positive probability weight on the former set of types.*

Notice that IC places no restrictions on beliefs over the types that are *not* equilibrium dominated, and in addition it also places no restrictions if *every* type is equilibrium-dominated. For then the deviation signal is surely an error, and once that possibility is admitted, all bets about who is emitting that signal are off.

The idea behind IC is the following “speech” that a sender (of signals) might make to a recipient:

Look, I am sending you this signal which is equilibrium-dominated for types  $A$ ,  $B$  or  $C$ . But it is not so for types  $D$  and  $E$ . Therefore you *cannot* believe that I am types  $A$ ,  $B$  or  $C$ .

Let us apply this idea to the Spence model.

**Proposition 4.** *In the Spence Signalling model, a single equilibrium outcome survives the IC, and it is the separating equilibrium in which  $L$  plays 0 while  $H$  plays  $e_1$ , where  $e_1$  solves (??) with equality.*

*Proof.* First we rule out all equilibria in which types  $H$  and  $L$  play the same value of  $e$  with positive probability. [This deals with all the pooling and all the hybrid equilibria.]

At such an  $e$ , the payoff to each type  $\theta$  is

$$\lambda H + (1 - \lambda)L - \frac{e}{\theta},$$

where  $\lambda$  represents the employer's posterior belief after seeing  $e$ . Now, there always exists an  $e' > e$  such that

$$\lambda H + (1 - \lambda)L - \frac{e}{L} = H - \frac{e'}{L} < H - \frac{e'}{H}$$

If we choose  $e''$  very close to  $e'$  but slightly bigger than it, it will be equilibrium-dominated for the low type —

$$\lambda H + (1 - \lambda)L - \frac{e}{L} > H - \frac{e''}{L},$$

while it is not equilibrium-dominated for the high type:

$$\lambda H + (1 - \lambda)L - \frac{e}{H} < H - \frac{e''}{H}.$$

But now the equilibrium is broken by having the high type deviate to  $e''$ . By IC, the employer must believe that the type there is high for sure and so must pay out  $H$ . But then the high type benefits from this deviation relative to playing  $e$ .

Next, consider all separating equilibria in which  $L$  plays 0 while  $H$  plays some  $e > e_1$ . Then a value of  $e'$  which is still bigger than  $e_1$  but smaller than  $e$  can easily be seen to be equilibrium-dominated for the low type but not for the high type. So such values of  $e'$  must be rewarded with a payment of  $H$ , by IC. But then the high type will indeed deviate, breaking the equilibrium.

This proves that the only equilibrium that can survive the IC is the one in which the low type plays 0 and the high type chooses  $e_1$ .  $\square$

The heart of the intuitive criterion is an argument which is more general: it is called *forward induction*. The basic idea is that an off-equilibrium signal can be due to one of two things: an error, or strategic play. If at all strategic play can be suspected, the error theory must play second fiddle: that's what a forward induction argument would have us believe.

## 24. FORWARD INDUCTION AND ITERATED WEAK DOMINANCE

In the same way that iterated strict dominance and rationalizability can be used to narrow down the set of predictions without pinning down strategies perfectly, the concept of iterated weak dominance (IWD) can be used to capture some of the force of forward and backward induction without assuming that players coordinate on a certain equilibrium. Since the idea of forward induction is that players interpret a deviation as a signal of future play, forward

induction is more compatible with a situation of considerable strategic uncertainty—a non-equilibrium model—rather than a theory in which players are certain about the opponents’ strategies.

In games with perfect information iterated weak dominance implies backward induction. Indeed, any suboptimal strategy at a penultimate node is weakly dominated, then we can iterate this observation.

IWD also captures part of the forward induction notion implicit in stability, since stable components contain stable sets of games obtained by removing a weakly dominated action. For instance, applying IWD to the motivating example of Kohlberg and Mertens we obtain

	T	W
O	2, 2	2, 2
IT	0, 0	3, 1
IW	1, 3	0, 0

the unique outcome  $(IT, W)$  predicted by stability.

Similarly, we can solve the beer-quiche game using IWD. Consider the ex ante game in which the types of player 1 are treated as two distinct information sets for the same player. Player 1’s strategy (beer if wimp, quiche if surly) is strictly dominated by a strategy under which with probability .9 both types of player 1 eat quiche and with probability .1 both drink beer. Indeed, for any strategy of player 2, the latter strategy involves the same total probability that player 1 is fought by player 2 as the former, but the latter leads to player 1’s favorite breakfast with higher probability. Once we eliminate (beer if wimp, quiche if surly), only the strategies (beer if wimp, beer if surly) and (quiche if wimp, beer if surly) generate a breakfast of beer for player 1. Then the decision of whether player 2 should fight after observing a breakfast of beer makes a difference only in the event that player 1 uses one of these two strategies. The best response to either strategy is not fighting because it implies a probability of at least .9 of confronting the surly type. This means that any strategy for 2 that involves fighting after observing beer is weakly dominated in the strategic form by one with no fighting after beer. Then the surly type should choose beer in any surviving equilibrium, which generates his highest possible payoff of 3—he has his preferred breakfast and is not challenged by player 2.

Ben-Porath and Dekel (1992) consider the following striking example in which the mere option of “burning money” selects a player’s favorite equilibrium in the following battle of the sexes game. The outcome  $(U, L)$  is preferred by player 1 to any other outcome, and is

	L	R
U	5, 1	0, 0
D	0, 0	1, 5

a strict Nash equilibrium. Suppose we extend the game to include a signaling stage, where player 1 has the possibility of burning, say, 2 units of utility before the game begins. Hence player 1 first chooses between the game above and the following game. Burning and then

	L	R
U	3, 1	-2, 0
D	-2, 0	-1, 5

playing  $D$  is strongly dominated for player 1 (by not burning and playing  $D$ ) hence if player 2 observes 1 burning, then 2 can conclude that 1 will play  $U$ . Therefore player 1 can guarantee herself a payoff of 3 by burning and playing  $U$ , since 2 (having concluded that 1 will play  $U$  after burning) will play  $L$ . Formally, any strategy in which 2 plays  $R$  after burning is weakly dominated by playing  $L$  after burning (the two strategies lead to the same outcome in the event that player 1 does not burn, hence the weak domination). Now, even if player 1 does not burn, player 2 should conclude that 1 will play  $U$ . This is because, by playing  $D$ , player 1 can receive a payoff of at most 1, while the preceding argument demonstrated that player 1 can guarantee 3 (by burning). That is, among the surviving strategies, player 1’s strategy of playing  $D$  after not burning is strictly dominated by burning and playing  $U$ . Hence, if 2 observes that 1 does not burn then 2 will play  $L$ —playing  $R$  after 1 does not burn is weakly dominated among the surviving strategies by playing  $L$ —leading to player 1’s preferred outcome which involves no burning and  $(U, L)$ . Thus player 1 can ensure that his most preferred equilibrium is played even *without* burning. Ben-Porath and Dekel show that in any game where a player has a unique best outcome that is a strict Nash equilibrium and can signal with a sufficiently fine grid of burning stakes, she will attain her most preferred outcome under IWD.

## 25. REPEATED GAMES

We now move on to consider another important topic: repeated games. Let  $G = (N, A, u)$  be a normal-form stage game. At time  $t = 0, 1, \dots$ , the players simultaneously play game  $G$ . At each period, the players can all observe play in each previous period; the history is denoted  $h^t = (a^0, \dots, a^{t-1})$ . Payoffs in the repeated game  $RG(\delta)$  are given by  $U_i = (1 - \delta) \sum_{t=0}^{\infty} \delta^t u_i(a^t)$ . The  $(1 - \delta)$  factor normalizes the sum so that payoffs in the repeated game are on the same scale as in the stage game. We assume players follow behavior strategies (by Kuhn's theorem), so a strategy  $\sigma_i$  for player  $i$  is given by a choice of  $\sigma_i(h^t) \in \Delta(A_i)$  for each history  $h^t$ . Given such strategies, we can define continuation payoffs after any history  $h^t$ :  $U_i(\sigma|h^t)$ .

If  $\alpha^*$  is a Nash equilibrium of the static game, then playing  $\alpha^*$  at every history is a subgame-perfect equilibrium of the repeated game. Conversely: for any finite game  $G$  and any  $\varepsilon > 0$ , there exists  $\bar{\delta}$  with the property that, for any  $\delta < \bar{\delta}$ , any SPE of the repeated game  $RG(\delta)$  has the property that, at every history, play is within  $\varepsilon$  of a static NE (in the strategy space). However, interesting results generally occur when players have high discount factors, not low discount factors.

The main results for repeated games are “Folk Theorems”: for high enough  $\delta$ , every feasible and individually rational payoff vector in the stage game can be attained in an equilibrium of the repeated game. There are several versions of such a theorem, which is why we use the plural. For now, we look at repeated games with perfect monitoring (the class of games defined above), where the appropriate equilibrium concept is SPE. We can check if a strategy profile is an SPE by using the one-shot deviation principle. Conditional on a history  $h^t$ ,  $i$ 's payoff from playing  $a$  and then following  $\sigma$  in the continuation is given by the value function

$$(25.1) \quad V_i(a) = (1 - \delta)u_i(a) + \delta U_i(\sigma|h^t, a).$$

This gives us an easy way to check whether or not a player wants to deviate from a proposed strategy, given other player's strategies.  $\sigma$  is an SPE if and only if, for every history  $h^t$ ,  $\sigma|h^t$  is a NE of the induced game  $G(h^t, \sigma)$  whose payoffs are given by (??).

To state a folk theorem, we need to explain the terms “individually rational” and “feasible.” The *minmax payoff* of player  $i$  is the worst payoff his opponents can hold him down

to if he knows their strategies:

$$\underline{v}_i = \min_{\alpha_{-i} \in \prod_{j \neq i} \Delta(A_j)} \left[ \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}) \right].$$

We will let  $m^i$ , a *minmax profile for  $i$* , denote a profile of strategies  $(a_i, \alpha_{-i})$  that solves this minimization and maximization problem. Note that we require *independent* mixing by  $i$ 's opponents. It is important to consider mixed, rather than just pure, strategies for  $i$ 's opponents. For instance, in the matching pennies game the minmax when only pure strategies are allowed for the opponent is 1, while the actual minmax, involving mixed strategies, is 0.

In any SPE—in fact, any Nash equilibrium— $i$ 's payoff is at least his minmax payoff, since he can always get at least this much by just best-responding to his opponents' (possibly independently mixed) actions in each period separately. This motivates us to say that a payoff vector  $v$  (i.e. an element of  $\mathbb{R}^N$ , specifying a payoff for each player) is *individually rational* if  $v_i \geq \underline{v}_i$  for each  $i$ , and it is *strictly individually rational* if the inequality is strict for each  $i$ .

The set of *feasible payoffs* (properly, feasible payoff vectors) is the convex hull of the set  $\{u(a) \mid a \in A\}$ . Again note that this can include payoffs that are not obtainable in the stage game using mixed strategies, because some such payoffs may require correlation among players to achieve. Under the *common discount factor* assumption, the normalized payoffs along any path of play in the repeated game are certainly in the feasible set.

Also, in studying repeated games we usually assume the availability of a *public randomization device* that produces a publicly observed signal  $\omega^t \in [0, 1]$ , uniformly distributed and independent across periods, so that players can condition their actions on the signal. Properly, we should include the signals (or at least the current period's signal) in the specification of the history, but it is conventional not to write it out explicitly. The public randomization device is a convenient way to convexify the set of possible equilibrium payoff vectors: for example, given equilibrium payoff vectors  $v$  and  $v'$ , any convex combination of them can be realized by playing the equilibrium with payoffs  $v$  conditional on some realizations of the device and  $v'$  otherwise. (Fudenberg and Maskin (1991) showed that one can actually do

this without the public randomization device for sufficiently high  $\delta$ , while *preserving incentives*, by appropriate choice of which periods to play each action profile involved in any given convex combination.)

An easy folk theorem is that of Friedman (1971):

**Theorem 22.** *If  $e$  is the payoff vector of some Nash equilibrium of  $G$ , and  $v$  is a feasible payoff vector with  $v_i > e_i$  for each  $i$ , then for all sufficiently high  $\delta$ , there exists an SPE with payoffs  $v$ .*

*Proof.* Just specify that the players play whichever action profile gives payoffs  $v$  (using the public randomization device to correlate their actions if necessary), and revert to the static Nash permanently if anyone has ever deviated. When  $\delta$  is high enough, the threat of reverting to Nash is severe enough to deter anyone from deviating.  $\square$

So, in particular, if there is a Nash equilibrium that gives everyone their minmax payoff (for example, in the prisoner's dilemma), then every strictly individually rational and feasible payoff vector is obtainable in SPE.

However, it would be nice to have a full, or nearly full, characterization of the set of possible equilibrium payoff vectors (for large  $\delta$ ). In many repeated games, the Friedman folk theorem is not strong enough for this. A more general folk theorem would say that every individually rational, feasible payoff is achievable in SPE under general conditions. This is harder to show, because in order for one player to be punished by minmax if he deviates, others need to be willing to punish him. Thus, for example, if all players have equal payoff functions, then it may not be possible to punish a player for deviating, because the punisher hurts himself as well as the deviator.

For this reason, the standard folk theorem (due to Fudenberg and Maskin, 1986) requires a full-dimensionality condition.

**Theorem 23.** *Suppose the set of feasible payoffs has full dimension  $n$ . For any feasible and strictly individually rational payoff vector  $v$ , there exists  $\underline{\delta}$  such that whenever  $\delta > \underline{\delta}$ , there exists an SPE of  $RG(\delta)$  with payoffs  $v$ .*

Actually we don't quite need the full-dimensionality condition—all we need, conceptually, is that there are no two players who have the same payoff functions; more precisely, no

player's payoff function can be a positive affine transformation of any other's (Abreu, Dutta, and Smith, 1994). But the proof is easier under the stronger assumption.

*Proof.* We will first give the construction assuming that  $i$ 's minmax action profile  $m^i$  is pure. Consider the action profile  $a$  for which  $u(a) = v$ . Choose  $v'$  in the interior of the feasible, individually rational set with  $v'_i < v_i$  for each  $i$ . Let  $w^i$  denote  $v'$  with  $\varepsilon$  added to each player's payoff except for player  $i$ ; with  $\varepsilon$  low enough, this will again be a feasible payoff vector.

Strategies are now specified as follows.

- Phase  $I$ : play  $a$ , as long as there are no deviations. If  $i$  deviates, switch to  $II_i$ .
- Phase  $II_i$ : play  $m^i$ . If player  $j$  deviates, switch to  $II_j$ . (If several players deviate simultaneously, we may arbitrarily choose  $j$  among them; this makes little difference, since verification of the equilibrium will only require checking single deviations.) Note that if  $m^i$  is a pure strategy profile it is clear what we mean by  $j$  deviating. If it requires mixing it is not so clear; this will be discussed in the second part of the proof. Phase  $II_i$  lasts for  $T$  periods, where  $T$  is a number, *independent of*  $\delta$ , to be determined, and if there are no deviations during this time, play switches to  $III_i$ .
- Phase  $III_i$ : play the action profile leading to payoffs  $w^i$  forever. If  $j$  deviates, go to  $II_j$ . (This is the "reward" phase that gives players  $-i$  incentives to punish in phase  $II_i$ .)

We check that there are no incentives to deviate, using the one-shot deviation principle for each of the three phases: calculate the payoff to  $i$  from complying and possible deviations in each phase. Phases  $II_i$  and  $II_j$  ( $j \neq i$ ) need to be considered separately, as do  $III_i$  and  $III_j$ .

- Phase  $I$ : deviating gives at most  $(1 - \delta)M + \delta(1 - \delta^T)v_i + \delta^{T+1}v'_i$ , where  $M$  is some upper bound on all of  $i$ 's feasible payoffs, and complying gives  $v_i$ . Whatever  $T$  we have chosen, it is clear that as long as  $\delta$  is sufficiently close to 1, complying produces a higher payoff than deviating, since  $v'_i < v_i$ .
- Phase  $II_i$ : Suppose there are  $T' \leq T$  remaining periods in this phase. Then complying gives  $i$  a payoff of  $(1 - \delta^{T'})v_i + \delta^{T'}v'_i$ , whereas since  $i$  is being minmaxed, deviating can't help in the current period and leads to  $T$  more periods of punishment, for a

total payoff of at most  $(1 - \delta^{T+1})\underline{v}_i + \delta^{T+1}v'_i$ . Thus deviating is always worse than complying.

- Phase  $II_j$ : With  $T'$  remaining periods,  $i$  gets  $(1 - \delta^{T'})u_i(m^j) + \delta^{T'}(v'_i + \varepsilon)$  from complying and at most  $(1 - \delta)M + (\delta - \delta^{T+1})\underline{v}_i + \delta^{T+1}v'_i$  from deviating. When  $\delta$  is large enough, complying is preferred.
- Phase  $III_i$ : This is the one case that affects the choice of  $T$ . Complying gives  $v'_i$  in every period, while deviating gives at most  $(1 - \delta)M + \delta(1 - \delta^T)\underline{v}_i + \delta^{T+1}v'_i$ . Rearranging, the comparison is between  $(\delta + \delta^2 + \dots + \delta^T)(v'_i - \underline{v}_i)$  and  $M - v'_i$ . For any  $\underline{\delta} \in (0, 1)$ , there exists  $T$  such that the desired inequality holds for all  $\delta > \underline{\delta}$ .
- Phase  $III_j$ : Complying gives  $v'_i + \varepsilon$  forever, whereas deviating leads to a switch to phase  $II_i$  and so gives at most  $(1 - \delta)M + \delta(1 - \delta^T)\underline{v}_i + \delta^{T+1}v'_i$ . Again, for sufficiently large  $\delta$ , complying is preferred.

Now we need to deal with the part where minmax strategies are mixed. For this we need to change the repeated-game strategies so that, during phase  $II_j$ , player  $i$  is indifferent among all the possible sequences of  $T$  realizations of his prescribed mixed action. We accomplish this by choosing a different reward  $\varepsilon$  for each such sequence, so as to balance out their different short-term payoffs. We're not going to talk about this in detail; see the Fudenberg and Maskin paper for this.  $\square$

## 26. REPEATED GAMES WITH FIXED $\delta < 1$

The folk theorem shows that many payoffs are possible in SPE. But the construction of strategies in the proof is fairly complicated, since we need to have punishments and then rewards for punishers to induce them not to deviate. In general, an equilibrium may be supported by an elaborate hierarchy of punishments, and punishments of deviations from the prescribed punishments, and so on. Also, the folk theorem is concerned with limits as  $\delta \rightarrow 1$ , whereas we may be interested in the set of equilibria for a particular value of  $\delta < 1$ .

We will now approach the question of identifying equilibrium payoffs for a given  $\delta < 1$ . In repeated games with perfect information, it turns out that an insight of Abreu (1988) will simplify the analysis greatly: equilibrium strategies can be enforced by using a worst possible punishment for any deviator. First we need to show that there is a well-defined worst possible punishment.

**Theorem 24.** *Suppose each player's action set in the stage game is a compact subset of a Euclidean space and payoffs are continuous in actions, and some pure-strategy SPE of the repeated game exists. Then, among all pure-strategy SPEs, there is one that is worst for player  $i$ .*

That is, the infimum of player  $i$ 's payoffs, across all pure-strategy SPEs, is attained.

*Proof.* We prove this for every player  $i$  simultaneously.

An *equilibrium play path* is an infinite sequence of action profiles, one for each period, that is attained in some pure-strategy SPE. Fix a sequence of such play paths  $a^{i,k}$ ,  $k = 0, 1, 2, \dots$  such that  $U_i(a^{i,k})$  converges to the specified infimum  $y(i)$ , as  $k \rightarrow \infty$ . We want to define a limit of the play paths, in such a way that the limiting path is again achieved in some SPE, with payoff  $y(i)$  to player  $i$ . The constructed equilibria rely on each other for punishments off the equilibrium path.

Each play path is an element of the strategy space  $\prod_{t \geq 0} A$ , where  $A$  is the action space of the stage game. Endow this space with the product topology. Convergence in the product topology is defined componentwise—that is,  $a^{i,k} \rightarrow a^{i,\infty}$  if and only if  $a_t^{i,k} \rightarrow a_t^{i,\infty}$  for each  $t$ . Because the space of paths is sequentially compact,<sup>15</sup> by passing to a subsequence if necessary, we can ensure that the  $a^{i,k}$  have a limiting play path  $a^{i,\infty}$ . It is easy to check that the resulting payoff to player  $i$  is  $y(i)$ .

Now we just have to check that this limiting play path  $a^{i,\infty}$  is supportable as an SPE by some strategy profile. We construct the following profile. Play starts in regime  $i$ . A deviation by player  $j$  from the current regime leads to regime  $j$ . In each regime  $i$ , all players play according to  $a^{i,\infty}$ .

<sup>15</sup>By a diagonalisation argument, a countable product of sequentially compact spaces is sequentially compact. Note that while the set of play paths in Abreu's setting is sequentially compact, the space of pure strategy profiles is not. This space is an uncountable product of Euclidean sets. For instance, second-period strategies depend on the first period action profile and are represented by the set  $A^A$ . Even when  $A$  is an closed interval,  $A^A$  is not sequentially compact. (Note, however, that by Tychonoff's theorem  $A^A$  is a compact set.) For a proof, consider the set  $[0, 1]^{[0,1]}$  with the product topology. We can think of each point in the set as a function  $f : [0, 1] \rightarrow [0, 1]$ . Convergence in the product topology reduces to point-wise convergence for such functions. Let  $f_n(x)$  denote the  $n$ th digit in the binary expansion of  $x$ . The sequence  $(f_n)$  does not admit a convergent subsequence. Indeed, for any subsequence  $(f_{n_k})_{k \geq 0}$  consider an  $x \in [0, 1]$  whose binary expansion has the  $n_k$  entry equal to 0 for  $k$  even and 1 for  $k$  odd. Then  $(f_{n_k}(x))_{k \geq 0}$  is not a convergent sequence of real numbers, and hence  $(f_{n_k})_{k \geq 0}$  does not converge in the product topology.

Now we need to check that the  $|N|$  strategy profiles constructed this way are indeed SPEs. Consider a deviation by player  $j$  from stage  $\tau$  of regime  $i$  to an action  $\widehat{a}_j$ . His payoff from deviating is

$$(1 - \delta)u_j(\widehat{a}_j, a_{-j}^{i,\infty}(\tau)) + \delta y(j).$$

We want to show that this is at most the continuation payoff from complying,

$$(1 - \delta) \sum_{t=0}^{\infty} \delta^t u_j(a^{i,\infty}(\tau + t)).$$

But we know that for each  $k$ , there is some SPE whose equilibrium play path is  $a^{i,k}$ ; in SPE,  $j$  is disincentivized from deviating, and we also know that by deviating his value in future periods is at least  $y(j)$  (by definition of  $y(j)$ ). So for each  $k$  we have

$$(1 - \delta) \sum_{t=0}^{\infty} \delta^t u_j(a^{i,k}(\tau + t)) \geq u_j(\widehat{a}_j, a_{-j}^{i,k}(\tau)) + \delta y(j).$$

By taking limits at  $k \rightarrow \infty$ , we see that there is no incentive to deviate in the strategy profile supporting  $a^{i,\infty}$ , either.

This shows there are never incentives for a one-shot deviation. So by the one-shot deviation principle, we do have an SPE giving  $i$  his infimum of SPE payoffs, for each player  $i$ .  $\square$

Abreu refers to an SPE that gives  $i$  his worst possible payoff as an *optimal penal code*.

The above theorem applies when there exists a pure-strategy SPE. If the stage game is finite, there frequently will not be any pure-strategy SPE. In this case, there will be mixed-strategy SPE, and we would like to again prove that an optimal (mixed-strategy) penal code exists. A different method is required; we invoke a theorem of Fudenberg and Levine (1983).

**Theorem 25.** *Consider an infinite-horizon repeated game with a finite stage game. The set of strategy profiles is simply the countable product  $\prod_{h_t} \prod_i \Delta(A_i)$ , taken over all possible finite histories  $h_t$  and players  $i$ . Put the product topology on this space. Then the set of SPE profiles and payoffs are nonempty and compact.*

The set of SPEs is nonempty because it includes strategies that play the same static Nash equilibrium following any history. Since the stage game is finite,  $\prod_{h_t} \prod_i \Delta(A_i)$  is a countable product of sequentially compact spaces, so it is sequentially compact (see also footnote ??). One can easily show that payoffs vary continuously in the strategy profile for the repeated

game (with the product topology). Indeed, for any sequence  $\sigma^n \rightarrow \sigma$  and any fixed  $t$ , the distribution over date  $t$  histories/actions induced by  $\sigma_n$  converges to the one induced by  $\sigma$  as  $n \rightarrow \infty$ . Then the expected payoffs under  $\sigma_n$  converge to those under  $\sigma$  as  $n \rightarrow \infty$ . This immediately implies that the set of SPEs is closed.<sup>16</sup> Since  $\prod_{h_t} \prod_i \Delta(A_i)$  is compact by Tychonoff's theorem and closed subsets of compact sets are compact, it follows that the set of SPE strategies is compact. As payoffs are continuous in strategies, the set of SPE payoffs is also compact.<sup>17</sup> In particular, for every player  $i$  there exists an SPE that minimizes  $i$ 's payoff, that is, an optimal penal code for  $i$ .

The following result holds in either of the settings where we proved the existence of an optimal penal code—either for pure strategies when the stage game has continuous action spaces (and some SPE exists) or for mixed strategies when the stage game is finite.

**Theorem 26.** (*Abreu, 1988*) *Any distribution over play paths achievable by an SPE can be generated by an SPE enforced by optimal penal codes off the equilibrium path, i.e. when  $i$  is the first to deviate, continuation play follows the optimal penal code for  $i$ .*

For mixed-strategy equilibria, “off the path” means “at histories that occur with probability zero.”

*Proof.* Let  $\hat{\sigma}$  be the given SPE. Form a new strategy profile  $s$  by leaving play on the equilibrium path as proposed by  $\hat{\sigma}$ , and replacing play off the equilibrium path by the optimal penal code for  $i$  when  $i$  is the first deviator (or one of the first deviators, if there is more than one). By the one-shot deviation principle and the fact that off-path play follows an SPE, we need only check that  $i$  does not want to deviate when play so far is on the equilibrium path—but this is immediate, because  $i$  is punished with  $y(i)$  in the continuation if he deviates, whereas in the original profile  $\hat{\sigma}$  he would get at least  $y(i)$  in the continuation (by definition of  $y(i)$ ) and we know this was already low enough to deter deviation (because  $\hat{\sigma}$  was an SPE).  $\square$

Abreu (1986) looks at symmetric games and considers *strongly symmetric* equilibria—equilibria in which all players behave identically at every history, including asymmetric

<sup>16</sup>Since countable product of metric spaces is metrizable, the product topology on  $\prod_{h_t} \prod_i \Delta(A_i)$  is metrizable, so closed sets can be defined in terms of convergent sequences.

<sup>17</sup>These conclusions extend to multistage games with observable actions that have a finite set of actions at every stage and are continuous at infinity. See Theorem 4.4 in FT, relying on approximate equilibria of truncated games.

histories. This is a simple setting because everyone gets the same payoff, so there is one such equilibrium that is worst for everyone. One can similarly show that there is an equilibrium that is best for everyone. Abreu considers a stage game that is a general version of a Cournot oligopoly. The action spaces are given by  $[0, \infty)$  (however, there exists an  $M$  such that taking an action higher than  $M$  is never rational). He assumes that payoffs are continuous and bounded from above as well as (a) the payoff at a symmetric action profile where all payers choose action  $a$  are quasi-concave and decrease without bound as  $a \rightarrow \infty$  and (b) the maximum payoff a player can achieve by responding to a profile in which all of his opponents play the same action  $a$  is decreasing in  $a$ .

**Theorem 27.** *Let  $e^*$  and  $e_*$  denote the highest and lowest payoff per player in a pure-strategy strongly symmetric equilibrium.*

- *The payoff  $e_*$  can be attained in an equilibrium with strongly symmetric strategies of the following form: “Begin in phase A, where players choose an action  $a_*$  that satisfies*

$$(1 - \delta)u(a_*, \dots, a_*) + \delta e^* = e_*.$$

*If there are no deviations, switch to an equilibrium with payoff  $e^*$  (phase B). Otherwise, continue in phase A.”*

- *Phase B: the payoff  $e^*$  can be attained with strategies that play a constant action  $a^*$  as long as there are no deviations and switch to the worst strongly symmetric equilibrium (phase A) if there are any deviations.*

For a proof of the first part of the statement, fix some strongly symmetric equilibrium  $\hat{\sigma}$  with payoff  $e_*$  and first period action  $a$ . Since the continuation payoffs under  $\hat{\sigma}$  cannot be more than  $e^*$ , the first period payoffs  $u(a, \dots, a)$  must be at least  $(-\delta e^* + e_*)/(1 - \delta)$ . Thus, under condition (a) there is an  $a_* \geq a$  with  $u(a_*, \dots, a_*) = (-\delta e^* + e_*)/(1 - \delta)$ . Let  $\sigma_*$  denote the strategies constructed in phase A. By definition, the strategies  $\sigma_*$  are subgame perfect in phase B. In phase A, condition (b) and  $a_* \geq a$  imply that the short-run gain to deviating is no more than that in the first period of  $\hat{\sigma}$ . Since the punishment for deviating in phase A is the worst possible punishment, the fact that no player prefers to deviate in the first period of  $\hat{\sigma}$  implies that no player prefers to deviate in phase A of  $\sigma_*$ .

The good equilibrium can be sustained by punishments that last only one period due to assumption (a), which ensures that punishments can be made arbitrarily bad. This is an important simplifying assumption. Then describing the set of strongly symmetric equilibrium payoffs is simple—there are just two numbers,  $a_*$  and  $a^*$ , and we just have to write the incentive constraints relating the two, which makes computing these extremal equilibria fairly easy. For either of the extremal equilibria, a first-period deviation leads to one period of punishment with the profile  $(a_*, \dots, a_*)$  and playing  $(a^*, \dots, a^*)$  thereafter. Abreu shows that this simple “stick and carrot” structure implies that  $a_*$  is the highest action and  $a^*$  is the lowest (recall that payoffs are decreasing in actions) among the pairs  $(a_*, a^*)$  for which the corresponding incentive constraints bind,

$$\begin{aligned} \max_{a_i \in A_i} u_i(a_i, a_{*-i}) - u_i(a_*, \dots, a_*) &= \delta(u_i(a^*, \dots, a^*) - u_i(a_*, \dots, a_*)) \\ \max_{a_i \in A_i} u_i(a_i, a_{-i}^*) - u_i(a^*, \dots, a^*) &= \delta(u_i(a^*, \dots, a^*) - u_i(a_*, \dots, a_*)). \end{aligned}$$

Typically the best outcome is better (and the worst punishment is worse) than the static Nash equilibrium.

## 27. IMPERFECT PUBLIC MONITORING

Next we describe the paradigm of repeated games with imperfect public monitoring: players only see a signal of other players’ past actions, rather than observing the actions fully. We spell out the general model while simultaneously giving a classic motivating example, the collusion model of Green and Porter (1984).

More specifically, each period there is a publicly observed signal  $y$ , which follows some probability distribution conditional on the action profile  $a$ ; write  $\pi_y(a)$  for the probability of  $y$  given  $a$ . Each player  $i$ ’s payoff is  $r_i(a_i, y)$ , something that depends only on his own action and the signal. His expected payoff from a strategy profile is then  $u_i(a) = \sum_{y \in Y} r_i(a_i, y) \pi_y(a)$ . In the Green-Porter model, each player is a firm in a cartel that sets a production quantity. Quantities are only privately observed. There is also a market price, which is publicly observed and depends stochastically on the players’ quantity choices (thus there is an unobserved demand shock each period). Each firm’s payoff is the product of the market price and its quantity, as usual. So the firms are trying to collude by keeping quantities low and prices high, but in any given period prices may be low, and each firm doesn’t know if prices

are low because of a demand shock or because some other firm deviated and produced a high quantity. In particular, Green and Porter assume that the support of the price signal  $y$  does not depend on the action profile played, which ensures that a low price may occur even when no firm has deviated.

Green and Porter did not try to solve for all equilibria of their model. Instead they simply discussed the idea of threshold equilibria: everyone plays the collusive action profile  $\hat{a}$  for a while; if the price  $y$  is ever observed to be below some threshold  $\hat{y}$ , revert to static Nash for some number of periods  $T$ , and then return to the collusion phase. (Note: this is not pushing the limits of what is feasible, since, for example, Abreu's work implies that there can be worse punishments possible than just reverting to static Nash.) In general, the optimal choice of  $T$  will be finite, since the punishment phase can be triggered accidentally in equilibrium and it is not optimal to end up stuck there forever.

Define  $\lambda(a) = P(y > \hat{y}|a)$ , the probability of seeing a high price when action profile  $a$  is played. Equilibrium values are then given by

$$\hat{v}_i = (1 - \delta)u_i(\hat{a}) + \delta\lambda(\hat{a})\hat{v}_i + \delta(1 - \lambda(\hat{a}))\delta^T\hat{v}_i$$

(after normalizing the static Nash payoffs to 0). This lets us calculate  $\hat{v}$  for any proposed  $\hat{a}$  and  $T$ ,

$$\hat{v} = \frac{(1 - \delta)u_i(\hat{a})}{1 - \delta\lambda(\hat{a}) - \delta^{T+1}(1 - \lambda(\hat{a}))}.$$

These strategies form an equilibrium only if no player wants to deviate in the collusive phase:

$$u_i(a'_i, \hat{a}_{-i}) - u_i(\hat{a}) \leq \frac{\delta(1 - \delta^T)(\lambda(\hat{a}) - \lambda(a'_i, \hat{a}_{-i}))\hat{v}_i}{1 - \delta} = \frac{\delta(1 - \delta^T)(\lambda(\hat{a}) - \lambda(a'_i, \hat{a}_{-i}))u_i(\hat{a})}{1 - \delta\lambda(\hat{a}) - \delta^{T+1}(1 - \lambda(\hat{a}))}$$

for all possible deviations  $a'_i$ . This compares the short-term incentives to deviate, the relative probability that deviation will trigger a reversion to static Nash, and the severity of the punishment.

It is possible to sustain payoffs at least slightly above static Nash with trigger strategies for high  $\delta$ . One can check that the incentive constraints hold for  $T = \infty$  and  $\hat{a}$  just below the static Nash actions, with  $\delta$  and  $\lambda$  close to 1 (and low  $\hat{y}$ ) and some bounds on the derivative of  $\lambda$ . As already remarked, Green and Porter did not identify the best possible equilibria.

To describe how one would find better equilibria, we need a general theory of repeated games with imperfect public monitoring. Accordingly, we return to the general setting; the

notation is as laid out at the beginning of this section. We will present the theory of these games as developed by Abreu, Pearce, and Stacchetti (1990) (hereafter referred to as APS).

For convenience we will assume that the action spaces  $A_i$  and the space  $Y$  of possible signals are finite. Recall that we write  $\pi_y(a)$  for the probability distribution over  $y$  given action profile  $a$ . It is clear how to generalize this to the distribution  $\pi_y(\alpha)$  where  $\alpha$  is a mixed action profile.

If there were just one period, players would just be playing the normal-form game with action sets  $A_i$  and payoffs  $u_i(a) = \sum_{y \in Y} \pi_y(a) r_i(a_i, y)$ . With repetition, this is no longer the case since play can be conditioned on the history—but may not be able to be conditioned exactly on past actions of opponents, as in the earlier, perfect-monitoring setting, because players do not see their opponents' actions.

Notice that the perfect monitoring setting can be embedded into this framework, by simply letting  $Y = A$  be the space of action profiles, and  $y$  be the action profile actually played with probability 1. We can also embed “noisy” repeated games with perfect monitoring, where each agent tries to play a particular action  $a_i$  in each period but ends up playing any other action  $a'_i$  with some small probability  $\varepsilon$ ; each player can only observe the action profile actually played, rather than the actions that the opponents “tried” to play.

In a repeated game with imperfect public monitoring, at any time  $t$ , player  $i$ 's information is given by his *private history*

$$h_i^t = (y^0, \dots, y^{t-1}; a_i^0, \dots, a_i^{t-1}).$$

That is, he knows the history of public signals and his own actions (but not others' actions). He can condition his action in the present period on this information. The *public history*  $h^t = (y^0, \dots, y^{t-1})$  is commonly known.

In their original paper, APS restrict attention to pure strategies, which is a nontrivial restriction.

A strategy  $\sigma_i$  for player  $i$  is a *public strategy* if  $\sigma_i(h_i^t)$  depends only on the history of public signals  $y^0, \dots, y^{t-1}$ .

**Lemma 1.** *Every pure strategy is equivalent to a public strategy.*

*Proof.* Let  $\sigma_i$  be a pure strategy. Define a public strategy  $\sigma'_i$  on length- $t$  histories by induction:  $\sigma'_i(y^0, \dots, y^{t-1}) = \sigma_i(y^0, \dots, y^{t-1}; a_i^0, \dots, a_i^{t-1})$  where  $a_i^s = \sigma'_i(y^0, \dots, y^{s-1})$  for each  $s < t$ . That is, at each period,  $i$  plays the actions specified by  $\sigma_i$  for the given public signals and the history of private actions that  $i$  was supposed to play. It is straightforward to check that  $\sigma'_i$  is equivalent to  $\sigma_i$ , since they differ only at “off-path” histories reachable only by deviations of player  $i$ .  $\square$

This shows that if attention is restricted to pure strategies, it is no further loss to restrict in turn to public strategies. However, instead of doing this, we will follow the exposition of Fudenberg, Levine, and Maskin (1994) and restrict attention to public (but potentially mixed) strategies.

**Lemma 2.** *If  $i$ 's opponents use public strategies, then  $i$  has a best response in public strategies.*

*Proof.* At every date  $i$  knows what the other players will play, since their actions depend only on the public history; hence  $i$  can just play a best response to their anticipated future play, which does not depend on  $i$ 's private history of past actions.  $\square$

This allows us to define a *perfect public equilibrium* (PPE): a profile  $\sigma = (\sigma_i)$  of public strategies such that, at every public history  $h^t = (y^0, \dots, y^{t-1})$ , the strategies  $\sigma_i|_{h^t}$  form a Nash equilibrium of the continuation game.

(This is the straightforward adaptation of the concept of subgame-perfect equilibrium to our setting. Notice that we cannot simply use subgame-perfect equilibrium because it has no bite in general—there are no subgames.)

The set of PPE's is stationary—they are the same at every history. This is why we look at PPE. Sequential equilibrium does not share this stationarity property, because a player may want to condition his play in one period on the realization of his mixing in a previous period. Such correlation across periods can be self-sustaining in equilibrium: if  $i$  and  $j$  both mixed at a previous period, then the signal in that period gives  $i$  information about the realization of  $j$ 's mixing, which means it is informative about what  $j$  will do in the current period, and therefore affects  $i$ 's current best response. On the other hand, some third player  $k$  may be unable to infer what  $j$  will do in the current period, since he does not know what

$i$  played in the earlier period. Consequently, different players can have different information at time  $t$  about what will be played at time  $t$ , and stationarity is destroyed. We stick to public equilibria in order to avoid this difficulty.

Importantly, the one-shot deviation principle applies to our setting. That is, a set of public strategies constitutes a PPE if and only if there is no beneficial one-shot deviation for any player.

Let  $w : Y \rightarrow \mathbb{R}^n$  be a function. We interpret  $w_i(y)$  as the continuation payoff player  $i$  expects after signal  $y$  is realized.

**Definition 22.** A pair consisting of a (mixed) action profile  $\alpha$  and payoff vector  $v \in \mathbb{R}^n$  is *enforceable* with respect to  $W \subseteq \mathbb{R}^n$  if there exists  $w : Y \rightarrow W$  such that

$$v_i = (1 - \delta)u_i(\alpha) + \delta \sum_{y \in Y} \pi_y(\alpha)w_i(y)$$

and

$$v_i \geq (1 - \delta)u_i(a'_i, \alpha_{-i}) + \delta \sum_{y \in Y} \pi_y(a'_i, \alpha_{-i})w_i(y)$$

for all  $i$  and all  $a'_i \in A_i$ .

The idea of enforceability is that it is incentive-compatible for each player to play according to  $\alpha$  in the present period if continuation payoffs are given by  $w$ , and the resulting (expected) payoffs starting from the present period are given by  $v$ .

Let  $B(W)$  be the set of all  $v$  that are enforceable with respect to  $W$  for some action profile  $\alpha$ . This is the set of payoffs *generated* by  $W$ .

**Theorem 28.** *Let  $E$  be the set of payoff vectors that are achieved by some PPE. Then  $E = B(E)$ .*

*Proof.* For any  $v \in E$  generated by some equilibrium strategy profile  $\sigma$ , let  $\alpha_i = \sigma_i(\emptyset)$  and  $w_i(y)$  be the expected continuation payoff of player  $i$  in subsequent periods given that  $y$  is the realized signal. Since play in subsequent periods again forms a PPE,  $w(y) \in E$  for each signal realization  $y$ . Then  $(\alpha, v)$  is enforced by  $w$  on  $E$ —this is exactly the statement that  $v$  represents overall expected payoffs and players do not have incentives to deviate from  $\alpha$  in the first period. So  $v \in B(E)$ .

Conversely, if  $v \in B(E)$ , let  $(\alpha, v)$  be enforced by  $w$  on  $E$ . Consider the strategies defined as follows: play  $\alpha$  in the first period, and whatever signal  $y$  is observed, play in subsequent periods follows a PPE with payoffs  $w(y)$ . These strategies form a PPE, by the one-shot deviation principle: enforcement means that there is no incentive to deviate in the first period, and the fact that continuation play is given by a PPE ensures that there is no incentive to deviate in any subsequent period. Finally it is straightforward from the definition of enforcement that the payoffs are in fact given by  $v$ . Thus  $v \in E$ .  $\square$

**Definition 23.**  $W \subseteq \mathbb{R}^n$  is *self-generating* if  $W \subseteq B(W)$ .

We have shown that  $E$  is self-generating. The next result shows that  $E$  is the largest bounded self-generating set.

**Theorem 29.** *If  $W$  is a bounded, self-generating set, then  $W \subseteq E$ .*

*Proof.* Let  $v \in W$ . We want to construct a PPE with payoffs given by  $v$ . We construct the strategies iteratively, simultaneously specifying, for each public history  $h^t = (y^0, \dots, y^{t-1})$ , the strategies to be played and the continuation values that each player should expect from subsequent play for each realization of the signal.

The base case,  $t = 0$ , has players receiving continuation payoffs given by  $v$ . Now suppose we have specified play for periods  $0, \dots, t - 1$ , and promised continuation payoffs for each history of signals  $y^0, \dots, y^{t-1}$ .

Suppose the history of public signals so far is  $y^0, \dots, y^{t-1}$  and promised continuation payoffs are given by  $v' \in W$ . Because  $W$  is self-generating, there is some action profile  $\alpha$  and some  $w : Y \rightarrow W$  such that  $(\alpha, v')$  is enforced by  $w$ . Specify that the players play  $\alpha$  at this history, and whatever signal  $y$  is observed, their continuation payoffs starting from the next period should be  $w(y)$ .

The expected payoffs following any public history match the target continuation payoffs; in particular, the constructed strategies generate expected payoffs of  $v$  at time 0. This follows from the adding-up identities, since they ensure that the continuation payoff following any public history  $h^t$  equals the expected total payoffs across the following  $s$  periods plus  $\delta^s$  times the promised continuation payoff from period  $t+s$  onward, and the latter converges to zero as  $s \rightarrow \infty$ . Here the assumption that  $W$  is bounded is essential; otherwise we could run a Ponzi

scheme with promised continuation payoffs. Finally, these strategies form a PPE—this is easily checked using the one-shot deviation principle. Enforcement means exactly that there are no incentives to deviate at any history.  $\square$

In addition to obtaining this characterization of the set of PPE payoffs, Abreu, Pearce, and Stacchetti also prove a monotonicity property with respect to the discount factor. Let  $E(\delta)$  be the set of PPE payoffs when the discount factor is  $\delta$ . Suppose that  $E(\delta)$  is convex: this can be achieved either by incorporating public randomization into the model, or by having a sufficiently rich space of public signals (however, in our version of the model  $Y$  is finite). Then if  $\delta_1 < \delta_2$  we have  $E(\delta_1) \subseteq B(E(\delta_1), \delta_2)$ , and therefore, by the previous theorem,  $E(\delta_1) \subseteq E(\delta_2)$ . This is shown by the following approach: given  $v \in E(\delta_1) = B(E(\delta_1), \delta_1)$ , find  $\alpha$  and  $w$  that enforce  $v$  when the discount factor is  $\delta_1$ ; then we can enforce  $(\alpha, v)$  for discount factor  $\delta_2$  with continuation payoffs given by a suitable convex combination of  $w$  and (the constant function)  $v$ . The operator  $B$  has the following properties.

- If  $W$  is compact, so is  $B(W)$ . This is shown by a straightforward topological argument.
- $B$  is monotone: if  $W \subseteq W'$  then  $B(W) \subseteq B(W')$ .
- If  $W$  is nonempty, so is  $B(W)$ . To show this, just let  $\alpha$  be a Nash equilibrium of the stage game,  $w : Y \rightarrow W$  a constant function, and  $v$  the resulting payoffs.

Now let  $V$  be the set of all feasible payoffs, which is certainly compact. Consider the sequence of iterates  $B^0(V), B^1(V), \dots$ , where  $B^0(V) = V$  and  $B^k(V) = B(B^{k-1}(V))$ . By induction, these sets are compact and form a decreasing sequence. Hence, their intersection  $B^\infty(V)$  is non-empty and compact. Since  $E \subseteq V$  and  $E = B(E)$ , the set  $E$  is contained in each term of the sequence, so  $E \subseteq B^\infty(V)$ .

**Theorem 30.**  $E = B^\infty(V)$ .

*Proof.* We are left to prove that  $B^\infty(V) \subseteq E$ . It suffices to show that  $B^\infty(V)$  is self-generating. Suppose  $v \in B^\infty(V)$ . Then there exists  $(\alpha^k, w^k)_{k \geq 1}$  such that  $(\alpha_k, v)$  is enforced by some  $w^k : Y \rightarrow B^{k-1}(V)$ . By compactness, this sequence has a convergent subsequence. Let  $(\alpha^\infty, w^\infty)$  denote the limit of such a subsequence. It must be that  $w^\infty(y) \in B^\infty(V)$  since  $w^\infty(y)$  is a limit point of the closed set  $B^k(V)$  for all  $k$ . By continuity,  $(\alpha^\infty, v)$  is enforced by  $w^\infty : Y \rightarrow B^\infty(V)$ , so  $v \in B(B^\infty(V))$ .  $\square$

This result characterizes the set of PPE payoffs: if we start with the set of all feasible payoffs and apply the operator  $B$  repeatedly, then the resulting sequence of sets converges to the set of equilibrium payoffs.

**Corollary 4.** *The set of PPE payoff vectors is nonempty and compact.*

(Nonemptiness is immediate because, for example, the infinite repetition of any static NE is a PPE.)

In their setting with finite action spaces and continuous signals with a common support, APS also show a “bang-bang” property of perfect public equilibria. We say that  $w : Y \rightarrow W$  has the bang-bang property if  $w(y)$  is an extreme point of  $W$  for each  $y$ . They show that if  $(\alpha, v)$  is enforceable on a compact  $W$ , it is in fact enforceable on the set  $\text{ext}(W)$  of extreme points of  $W$ . Consequently, every vector in  $E$  can be achieved as the vector of payoffs from a PPE such that the vector of continuation payoffs at every history lies in  $\text{ext}(E)$ .

## 28. THE FOLK THEOREM FOR IMPERFECT PUBLIC MONITORING

Fudenberg, Levine, and Maskin (1994) (hereafter FLM) prove a folk theorem for repeated games with imperfect public monitoring. They identify conditions on the stage game—particularly on how informative public signals need to be about actions—under which it is possible to construct convex sets with a smoothly curved boundary, approximating the set of feasible, individually rational payoffs arbitrarily closely that are *self-generating* for sufficiently high discount factors. This implies that a folk theorem obtains.

The proof is fairly complicated. We will briefly discuss the technical difficulties involved. First, there has to be statistical identifiability of each player’s actions. If player  $i$ ’s deviation to  $\alpha'_i$  generates exactly the same distribution over signals as some mixed action  $\alpha_i$  he is supposed to play (given opponents’ play  $\alpha_{-i}$ ), but gives him a higher payoff on average, then clearly there is no way to enforce the action profile  $\alpha$  in equilibrium. To avoid this problem, FLM assume an *individual full-rank* condition: given  $\alpha_{-i}$ , the different signal distributions generated by varying  $i$ ’s pure actions  $a_i$  are linearly independent.

They need to further assume a *pairwise full rank* condition: deviations by player  $i$  are statistically distinguishable from deviations by player  $j$ . Intuitively this is necessary because, if the signal suggests that someone has deviated, the players need to know who to punish.

(Radner, Myerson, and Maskin (1986) give an example of a game that violates this condition and where the folk theorem does not hold. There are two workers who put in effort to increase the probability that a project succeeds; they both get 1 if it succeeds and 0 otherwise. The outcome of the project does not statistically distinguish between shirking by player 1 and shirking by player 2. So if the project fails, both players have to be punished by giving them lower continuation payoffs than if it succeeds. Because it fails some of the time even if both players are working, this means that equilibrium payoffs are bounded away from efficiency, even as  $\delta \rightarrow 1$ .)

The statement of the pairwise full rank condition is as follows: given the action profile  $\alpha$ , if we form one matrix whose rows represent the signal distributions from  $(a_i, \alpha_{-i})$  as  $a_i$  varies over  $A_i$ , and another matrix whose rows represent the signal distributions from  $(a_j, \alpha_{-j})$  as  $a_j$  varies over  $A_j$ , and stack these two matrices, the combined matrix has rank  $|A_i| + |A_j| - 1$ . (This is effectively “full rank”—it is not possible to have literal full rank  $|A_i| + |A_j|$ , since the signal distribution generated by  $\alpha$  is a linear combination of the rows of the first matrix and is also a linear combination of the rows of the second matrix.)

When this condition is satisfied, it is possible to use continuation payoffs to transfer utility between the two players  $i, j$  in any desired ratio, depending on the signal, so as to incentivize  $i$  and  $j$  to play according to the desired action profile.

Having imposed appropriate formulations of these conditions, FLM show that the  $W$  they construct is *locally self-generating*: for every  $v \in W$ , there is an open neighborhood  $U$  and a  $\underline{\delta} < 1$  such that  $U \cap W \subseteq B(W)$  when  $\delta > \underline{\delta}$ . This definition allows  $\underline{\delta}$  to vary with  $v$ . For  $W$  compact and convex, they show that local self-generation implies self-generation for all sufficiently high  $\delta$ .

The intuition behind their approach to proving local self-generation is best grasped with a picture. Suppose we want to achieve some payoff vector  $v$  on the boundary of  $W$ . The full-rank conditions ensure we can enforce it using some continuation payoffs that lie below the tangent hyperplane to  $W$  at  $v$ , by “transferring” continuation utility between players as described above. As  $\delta \rightarrow 1$ , the continuation payoffs sufficient to enforce  $v$  contract toward  $v$ , and the smoothness condition on the boundary of  $W$  ensures that they will eventually lie inside  $W$ . Thus  $(\alpha, v)$  is enforced on  $W$ .

[PICTURE—See p. 1013 of Fudenberg, Levine, Maskin (1994)]

Some extra work is needed to take care of the points  $v$  where the tangent hyperplane is a coordinate hyperplane (i.e. one player's payoff is constant on this hyperplane).

An argument along these lines shows that every vector on the boundary of  $W$  is achievable using continuation payoffs in  $W$ , when  $\delta$  is high enough. Using public randomization among boundary points, we can then achieve any payoff vector  $v$  in the interior of  $W$  as well. It follows that  $W$  is self-generating (for high  $\delta$ ).

## 29. CHANGING THE INFORMATION STRUCTURE WITH TIME PERIOD

Follow FT pp. 197-200. Suppose players have a discount rate  $r$  and can only update their actions at times  $t, 2t, \dots$ . Then the effective discount rate is  $\delta = e^{-rt}$ . Hence the limit  $\delta \rightarrow 1$  has two interpretations: either players become patient ( $r \rightarrow 0$ ) or periods are short ( $t \rightarrow 0$ ). In games where actions are observable, as well as games with imperfect public monitoring in which the amount of information revealed does not change with  $t$ , the variables  $r$  and  $t$  enter symmetrically in  $\delta$  and the limit set of PPE payoffs as  $\delta \rightarrow 1$  can be interpreted both as the outcome when players are patient and when periods are short. However, if monitoring is imperfect and the quality of signals deteriorates as  $t \rightarrow 0$ , then the short period interpretation is lost.

Abreu, Milgrom, and Pearce (1991) point out that the two limits  $r \rightarrow 0$  and  $t \rightarrow 0$  may lead to distinct predictions. They focus on partnership games where the expected payoffs in the stage game induce the structure of prisoners' dilemma. Players do not directly observe each other's level of effort. Instead, the total level of effort is imperfectly reflected in a public signal, interpreted as the number of "successes." The signal has a Poisson distribution with an intensity parameter  $\lambda$  if both players cooperate and  $\mu$  if one of them defects. Assume that  $\lambda > \mu$ , so that signals indeed represent "good news."<sup>18</sup> For small  $t$ , the probability of observing more than one success is of order  $t^2$ . As in FT, we simplify the analysis by approximating the signal structure with a setting in which there are either 0 or 1 successes observed, with probabilities  $e^{-\theta t}$  and  $1 - e^{-\theta t}$ , respectively, for  $\theta \in \{\lambda, \mu\}$ .

Let  $c$  denote the common payoff when both players cooperate, and  $c + g$  the payoff a player obtains from defecting when the other cooperates; payoffs when both players defect are normalized to 0. The static Nash equilibrium (*defect, defect*) generates the minmax

<sup>18</sup>AMP also analyze the case of "bad news," where signals indicate failures.

payoff for both players. Hence the worst equilibrium for either player in the repeated game delivers zero payoffs.

Restrict attention to pure strategy strongly symmetric equilibria. Let  $v^*$  denote the payoff in an optimal equilibrium within this class. It can be easily seen that such an equilibrium must specify cooperation in the first period. Suppose that a public randomization device is available, so that continuation play when the number of successes observed is  $i = 0, 1$  can be described by playing the worst equilibrium (minmax, static Nash), with 0 payoffs, with probability  $\alpha(i)$  and the best equilibrium, which yields common continuation payoffs  $v^*$ , with probability  $1 - \alpha(i)$ .

The relevant PPE constraints are

$$\begin{aligned} v^* &= (1 - e^{-rt})c + e^{-rt} (e^{-\lambda t}(1 - \alpha(0)) + (1 - e^{-\lambda t})(1 - \alpha(1))) v^* \\ &\geq (1 - e^{-rt})(c + g) + e^{-rt} (e^{-\mu t}(1 - \alpha(0)) + (1 - e^{-\mu t})(1 - \alpha(1))) v^* \end{aligned}$$

Solving for  $v^*$  in the first equation, we obtain

$$v^* = \frac{(1 - e^{-rt})c}{1 - e^{-rt}(1 - \alpha(1) - e^{-\lambda t}(\alpha(0) - \alpha(1)))}.$$

The incentive constraint becomes

$$(1 - e^{-rt})g \leq e^{-rt}(e^{-\mu t} - e^{-\lambda t})(\alpha(0) - \alpha(1))v^*,$$

which simplifies to

$$g \leq \frac{ce^{-rt}(e^{-\mu t} - e^{-\lambda t})(\alpha(0) - \alpha(1))}{1 - e^{-rt}(1 - \alpha(1) - e^{-\lambda t}(\alpha(0) - \alpha(1)))}.$$

Note that  $v^*$  is decreasing in  $\alpha(1)$  and the incentive constraint is also relaxed by decreasing  $\alpha(1)$ . Hence an optimal symmetric equilibrium in pure strategies specifies  $\alpha(1) = 0$ . Intuitively, an optimal equilibrium should not involve any punishment if a success occurs.

Setting  $\alpha(1) = 0$ , the constraints become

$$\begin{aligned} v^* &= \frac{(1 - e^{-rt})c}{1 - e^{-rt}(1 - e^{-\lambda t}\alpha(0))} \\ g/c &\leq \frac{e^{-rt}(e^{-\mu t} - e^{-\lambda t})\alpha(0)}{1 - e^{-rt}(1 - e^{-\lambda t}\alpha(0))} \end{aligned}$$

It is possible to satisfy the inequality for  $\alpha(0) \leq 1$  only if

$$(29.1) \quad g/c \leq \frac{e^{-rt}(e^{-\mu t} - e^{-\lambda t})}{1 - e^{-rt}(1 - e^{-\lambda t})}$$

Note that

$$\frac{e^{-rt}(e^{-\mu t} - e^{-\lambda t})}{1 - e^{-rt}(1 - e^{-\lambda t})} \leq e^{(\lambda-\mu)t} - 1.$$

The RHS of the inequality above converges to 0 as  $t \rightarrow 0$ . Hence an equilibrium with payoffs above static Nash does not exist for small  $t$ . The term  $e^{(\lambda-\mu)t}$  can be interpreted as the likelihood ratio for no success. As  $t \rightarrow 0$  this ratio converges to 1. Since we are almost certain that no success occurs even when both players exert effort, the information provided by the public signal is too poor for there to be an equilibrium that improves on the static Nash outcome.

Taking the limit  $r \rightarrow 0$  in ??, we obtain

$$(29.2) \quad g/c \leq e^{(\lambda-\mu)t} - 1.$$

Hence an equilibrium with the desired properties exists for small  $r$  and certain values of  $t$ . For the “optimal” (minimum)  $\alpha(0)$ , we find that when ?? holds,

$$\lim_{r \rightarrow 0} v^* = c - \frac{g}{e^{(\lambda-\mu)t} - 1},$$

which is greater than  $\lim_{t \rightarrow 0} v^* = 0$  when ?? is satisfied with strict inequality.

### 30. REPUTATION

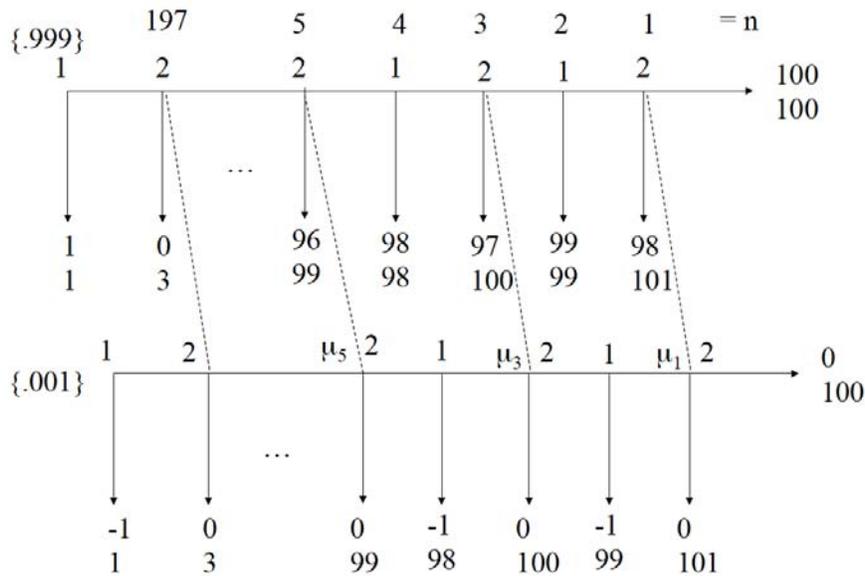
Repeated games provide a useful setting for studying the concept of reputation building. The earliest repeated-game models of reputation were by the Gang of Four (Kreps, Milgrom, Roberts, and Wilson); in various combinations they wrote three papers that were simultaneously published in JET 1982.

The motivating example was the “chain-store paradox.” In the chain-store game, there are two players, an entering firm and an incumbent monopolist. The entrant (player 1) can enter or stay out; if it enters, the incumbent (player 2) can fight or not. If the entrant stays out, payoffs are  $(0, a)$  where  $a > 1$ . If the entrant enters and the incumbent does not fight, the payoffs are  $(b, 0)$  where  $b \in (0, 1)$ . If they do fight, payoffs are  $(b - 1, -1)$ . There is a unique SPE, in which the entrant enters and the incumbent does not fight.

In reality, incumbent firms seem to fight when a rival enters, and thereby deter other potential rivals. Why would they do this? In a one-shot game, it is irrational for the incumbent to fight the entrant. As pointed out by Selten, even if the game is repeated finitely

many times, the unique SPE still has the property that there is entry and accommodation in every period, by backward induction.

The Kreps-Wilson explanation for entry deterrence is as follows: with some small positive probability, the monopolist does not have the payoffs described above, but rather is obsessed with fighting and has payoffs such that it always chooses to fight. Then, when there are a large number of periods, they show that there is no entry for most of the game, with entry occurring only in the last few periods.



Their analysis is tedious, so we will instead illustrate the concepts with a simpler example due to Muhamet Yildiz: the centipede game from Figure ???. Initially each player has \$1. Player 1 can end the game (giving payoffs (1, 1)), or he can give up \$1 for player 2 to get \$2. Player 2 can then end the game (giving (0, 3)), or can give up \$1 for player 1 to get \$2. Player 1 can then end the game (with payoffs (2, 2)), or can give up \$1 for player 2 to get \$2. And so forth—until the payoffs reach (100, 100), at which point the game automatically ends. We will refer to continuing the game as “playing across” and ending as “playing down,” due to the shape of the centipede diagram.

There is a unique SPE in this game, in which both players play down at every opportunity. But believing in SPE requires us to hold very strong assumptions about the players’ higher-order knowledge of each other’s rationality.

Suppose instead that player 1 has two types. With probability 0.999, he is a “normal” type and his payoffs are as above. With probability 0.001, he is a “crazy” type who always

gets utility  $-1$  if he ends the game and  $0$  if player 2 ends the game. (Player 2's payoffs are the same regardless of 1's type.) The crazy type of player 1 thus always wants to continue the game. Player 2 never observes player 1's type.

What happens in equilibrium? Initially player 1 has a low probability of being the crazy type. If the normal player 1 plays down at some information set, and the crazy player 1 across, then after 1 plays across, player 2 must infer that 1 is crazy. But if player 1 is crazy then he will continue the game until the end; knowing this, player 2 also wants to play across in order to accumulate money. Anticipating this, the normal type of player 1 in turn also wants to play across in order to get a high payoff.

With this intuition laid out, we analyze the game formally and describe the sequential equilibria. Number the periods, starting from the end, with 1 being player 2's last information set, 2 being player 1's previous information set, ..., 198 being 1's first information set. The crazy player 1 always plays across.

Player 2 always plays across with positive probability at every period  $n > 1$ . (Proof: if not, then the normal player 1 must play down at period  $n + 1$ . Then, conditional on reaching  $n$ , player 2 knows that 1 is crazy with probability 1, hence he would rather go across and continue the game to the end.)

Hence there is positive probability of going across at every period, so the beliefs are uniquely determined from the equilibrium strategies by Bayes' rule.

Next we see that the normal player 1 plays across with positive probability at every  $n > 2$ . Proof: if not, then again, at  $n - 1$  player 2 is sure that he is facing a crazy type and therefore wants to go across. Given this strategy by player 2, then, the normal 1 also has incentives to go across at  $n$  so that he can go down at  $n - 2$ , contradicting the assumption that 1 only goes down at  $n$ .

Next, if 2 goes across with probability 1 at  $n$ , then 1 goes across with probability 1 at  $n + 1$ , and this in turn implies that 2 goes across with probability 1 at  $n + 2$ . This is also seen by the same argument as in the previous paragraph. Therefore there is some cutoff  $n^* \geq 3$  such that both players play across with probability 1 at  $n > n^*$ , and there is mixing for  $2 < n \leq n^*$ . (We know that both the normal 1 and 2 play down with probability 1 at  $n = 1, 2$ .)

Let  $p_n$  be the probability of the normal type of player 1 going down at  $n$ , if  $n$  is even. Let  $\mu_n$  be the probability player 2 assigns to the crazy type at node  $n$ .

At each odd node  $n$ ,  $2 < n \leq n^*$ , player 2 is to be indifferent between going across and down. The payoff to going down is some  $x$ . The payoff to going across is  $(1 - \mu_n)p_{n-1}(x - 1) + [1 - (1 - \mu_n)p_{n-1}](x + 1)$ , using the fact that player 2 is again indifferent (or strictly prefers going down) two nodes later. Hence we get  $(1 - \mu_n)p_{n-1} = 1/2$ : player 2 expects player 1 to play down with probability  $1/2$ . But  $\mu_{n-2} = \mu_n / (\mu_n + (1 - \mu_n)(1 - p_{n-1}))$  by Bayes' rule; this simplifies to  $\mu_{n-2} = \mu_n / (1 - (1 - \mu_n)p_{n-1}) = 2\mu_n$ . We already know that  $\mu_1 = 1$  since the normal player 1 goes down with certainty at node 2. Therefore  $\mu_3 = 1/2$ ,  $\mu_5 = 1/4$ , and so forth; and in particular  $n^* \leq 20$ , since otherwise  $\mu_{21} = 1/1024 < 0.001$ , but clearly the posterior probability of the crazy type at any node cannot be lower than the prior. This shows that for all but the last 20 periods, both players are going across with probability 1 in equilibrium.

(One can in fact continue to solve for the complete description of the sequential equilibrium: now that we know player 2's posterior at each period, we can compute player 1's mixing probabilities from Bayes' rule, and we can also compute player 2's mixing probabilities given that 1 must be indifferent whenever he mixes.)

This model illustrates the way that the concept of reputation is generally modeled in repeated games. A player develops a reputation for playing a certain *action*; in equilibrium it is rational for him to continue with that action in order to maintain the reputation, even though it would not be rational in a one-shot setting. Unraveling is prevented by having a small probability of a type that is committed to that action.

The papers by the Gang of Four consider repeated interactions between the same players, with one-sided incomplete information. Inspired by this work, Fudenberg and Levine (1989) consider a model in which a long-run player faces a series of short-run players, and where there are many possible "crazy" types of the long-run player, each with small positive probability. They show that if the long-run player is sufficiently patient, he will get close to his Stackelberg payoff in any *Nash* equilibrium of the repeated game.

The model is as follows. There are two players, playing the finite normal-form game  $(N, A, u)$  (with  $N = \{1, 2\}$ ) in each period. Player 1 is a long-run player. Player 2 is a short-run player (which we can think of as a series of players who play for one period each,

or one very impatient player). Incentive for short-run players are simple: they best respond to the long-run player's anticipated action in each stage.

Define

$$u_1^* = \max_{a_1 \in A_1} \min_{\sigma_2 \in BR_2(a_1)} u_1(a_1, \sigma_2).$$

This is player 1's *Stackelberg payoff*; the action  $a_1^*$  that achieves this maximum is the *Stackelberg action*. Fudenberg and Levine (1989) only considers pure action selections by player 1. The analysis is extended to mixed actions in a follow-up paper published in 1992.

A strategy for player 1 consists of a function  $\sigma_1^t : H^{t-1} \rightarrow \Delta(A_1)$  for each  $t \geq 0$ . A strategy for the player 2 who plays at time  $t$  consists of a function  $\sigma_2^t : H^{t-1} \rightarrow \Delta(A_2)$ . With the usual discounted payoff formulation, the game described constitutes the *unperturbed game*. Fudenberg, Kreps, and Maskin (1988) prove a version of the folk theorem for this game. Let  $B_2$  denote the set of mixed strategy best responses in the stage game for the short run players to mixed strategies of the long run player. Set

$$\underline{u}_1 = \min_{\sigma_2 \in B_2} \max_{a_1 \in A_1} u_1(a_1, \sigma_2).$$

Fudenberg, Kreps, and Maskin show that any payoff above  $\underline{u}_1$  can be sustained in SPE for high enough  $\delta$ . The main reputation result of Fudenberg and Levine shows that if there is a rich space of crazy types of player 1, each with positive probability, this folk theorem is completely overturned—player 1 obtains a payoff of at least  $u_1^*$  in any Nash (not necessarily subgame perfect) equilibrium for high  $\delta$ . Note that for the standard Cournot duopoly with two firms and linear demand,  $\underline{u}_1$  and  $u_1^*$  correspond to the follower and leader payoffs, respectively.

Accordingly, we consider the perturbed game, where there is a countable state space  $\Omega$ . Player 1's payoff depends on the state  $\omega \in \Omega$ ; thus write  $u_1(a_1, a_2, \omega)$ . Player 2's payoff does not depend on  $\omega$ . There is some prior distribution  $\mu$  on  $\Omega$ , and the true state is known only to player 1. When the state is  $\omega_0 \in \Omega$ , player 1's payoffs are given by the original  $u_1$ ; we call this the “rational” type of player 1.

Suppose that for every  $a_1 \in A_1$ , there is a state  $\omega(a_1)$  for which playing  $a_1$  at every history is a strictly dominant strategy in the *repeated* game.<sup>19</sup> Thus, at state  $\omega(a_1)$ , player

<sup>19</sup>Assuming it is strictly dominant in the stage game is not enough. For instance, defection is a dominant strategy in prisoners' dilemma, but always defecting is not a best response against tit-for-tat in the repeated game.

1 is guaranteed to play  $a_1$  at every history. Write  $\omega^* = \omega(a_1^*)$ . We assume also that  $\mu^* = \mu(\omega^*) > 0$ . That is, with positive probability, player 1 is a type who is guaranteed to play  $a_1^*$  in every period.

Any strategy profile generates a joint probability distribution  $\pi$  over play paths and states,  $\pi \in \Delta((A_1 \times A_2)^\infty \times \Omega)$ . Let  $h^*$  be the event (in this path-state space) that  $a_1^t = a_1^*$  for all  $t$ . Let  $\pi_t^* = \pi(a_1^t = a_1^* | h^{t-1})$ , the probability of seeing  $a_1^*$  at period  $t$  given the previous history; this is a random variable (defined on path-state space) whose value is a function of  $h^{t-1}$ . For any number  $\bar{\pi} \in (0, 1)$ , let  $n(\pi_t^* \leq \bar{\pi})$  denote the number of periods  $t$  such that  $\pi_t^* \leq \bar{\pi}$ . This is again a random variable, whose value may be infinite.

The next result provides the main ingredient of the analysis. Conditional on observing  $a_1^*$  every period, it is guaranteed that there are at most  $\ln \mu^* / \ln \bar{\pi}$  periods in which  $a_1^*$  is expected with probability below  $\bar{\pi}$  conditional on the history. In other words, player 2 can be surprised by seeing  $a_1^*$  only a finite number of times.

**Lemma 3.** *Let  $\sigma$  be a strategy profile such that  $\pi(h^* | \omega^*) = 1$ . Then*

$$\pi \left( n(\pi_t^* \leq \bar{\pi}) \leq \frac{\ln \mu^*}{\ln \bar{\pi}} \mid h^* \right) = 1.$$

Given that  $\pi(h^* | \omega^*) = 1$ , if the true state is  $\omega^*$ , then player 1 will always play  $a_1^*$ . Every time the probability of seeing  $a_1^*$  next period is less than  $\bar{\pi}$ , if  $a_1^*$  is in fact played, the posterior probability of  $\omega^*$  must increase by a factor of at least  $1/\bar{\pi}$ . The posterior probability starts out at  $\mu^*$  and can never exceed 1, so it can increase no more than  $\ln \mu^* / \ln \bar{\pi}$  times.

Formally, consider any finite history  $h^t$  at which  $a_1^*$  has been played every period, and such that  $\pi(h^t) > 0$ . Write  $h^{t,1}$  ( $h^{t,2}$ ) for the event where  $h^{t-1}$  is observed and then at period  $t$  player 1 (2) plays as in  $h^t$ . We have that

$$\begin{aligned} \pi(\omega^* | h^t) &= \frac{\pi(h^t \ \& \ \omega^* | h^{t-1})}{\pi(h^t | h^{t-1})} = \frac{\pi(\omega^* | h^{t-1}) \pi(h^t | \omega^*, h^{t-1})}{\pi(h^t | h^{t-1})} \\ &= \frac{\pi(\omega^* | h^{t-1}) \pi(h^{t,1} | \omega^*, h^{t-1}) \pi(h^{t,2} | \omega^*, h^{t-1})}{\pi(h^{t,1} | h^{t-1}) \pi(h^{t,2} | h^{t-1})} \\ &= \frac{\pi(\omega^* | h^{t-1}) \pi(h^{t,2} | \omega^*, h^{t-1})}{\pi(h^{t,1} | h^{t-1}) \pi(h^{t,2} | h^{t-1})} \\ &= \frac{\pi(\omega^* | h^{t-1})}{\pi_t^*}. \end{aligned}$$

Here the first line of equalities uses Bayes' rule, the second holds because 1 and 2 mix independently at period  $t$ , the third holds because if  $\omega^*$  occurs then  $a_1^t = a_1^*$ , and the fourth holds because player 2's behavior conditional on the history  $h^{t-1}$  cannot depend on 1's type.

Repeatedly expanding, we have

$$\pi(\omega^*|h^t) = \frac{\pi(\omega^*|h^{t-1})}{\pi_t^*} = \dots = \frac{\pi(\omega^*|h^0)}{\pi_t^* \pi_{t-1}^* \dots \pi_0^*} = \frac{\mu^*}{\pi_t^* \pi_{t-1}^* \dots \pi_0^*}.$$

Since  $\pi(\omega^*|h^t) \leq 1$ , at most  $\ln \mu^* / \ln \bar{\pi}$  terms in the denominator of the last expression can be less than or equal to  $\bar{\pi}$ . Therefore,  $n(\pi_t^* \leq \bar{\pi}) \leq \ln \mu^* / \ln \bar{\pi}$  with probability 1.

Now we get to the main theorem. Let  $u_m = \min_{\sigma_2} u_1(a_1^*, \sigma_2, \omega_0)$  denote the worst possible stage payoff for player 1 when he takes action  $a_1^*$ . Note that the payoff  $u_1^*$  is a ‘‘lower Stackelberg payoff.’’ There is also an ‘‘upper Stackelberg payoff’’ in the stage game,

$$\bar{u}_1 = \max_{a_1} \max_{\sigma_2 \in BR_2(a_1)} u_1(a_1, a_2).$$

Let  $u_M = \max_a u_1(a, \omega_0)$  denote the highest payoff for the rational type of player 1 in the stage game. Denote by  $\underline{v}_1(\delta, \mu, \omega_0)$  and  $\bar{v}_1(\delta, \mu, \omega_0)$  the infimum and supremum, respectively, of rational player 1's payoffs in the repeated game across all Nash equilibria in which player 1 uses a pure strategy, for given discount factor  $\delta$  and prior  $\mu$ .

**Theorem 31.** *For any value  $\mu^*$ , there exists a number  $\kappa(\mu^*)$  with the following property: for all  $\delta$  and all  $(\mu, \Omega)$  with  $\mu(\omega^*) = \mu^*$ , we have*

$$\underline{v}_1(\delta, \mu, \omega_0) \geq \delta^{\kappa(\mu^*)} u_1^* + (1 - \delta^{\kappa(\mu^*)}) u_m.$$

Moreover, there exists  $\kappa$  such that for all  $\delta$ , we have

$$\bar{v}_1(\delta, \mu, \omega_0) \leq \delta^\kappa \bar{u}_1 + (1 - \delta^\kappa) u_M.$$

As  $\delta \rightarrow 1$ , the payoff bounds converge to  $u_1^*$  and  $\bar{u}_1$ , which are generically identical.

*Proof.* First, note that there exists a  $\bar{\pi} < 1$  such that, in every play path of every Nash equilibrium, at every stage  $t$  where  $\pi_t^* > \bar{\pi}$ , player 2 plays a best response to  $a_1^*$ . This follows from the fact that the pure strategy best response correspondence has a closed graph and the assumption that action spaces are finite.

Thus, by the lemma, we have a number  $\kappa(\mu^*)$  of periods such that  $\pi(n(\pi^* \leq \bar{\pi}) > \kappa(\mu^*) \mid h^*) = 0$ . Now, whatever player 2's equilibrium strategy is, if the rational player

1 deviates to simply playing  $a_1^*$  every period, there are at most  $\kappa(\mu^*)$  periods in which player 2 will not play a best response to  $a_1^*$ —since player 2 is playing a best response to player 1's expected play in each period. Thus the rational player 1 gets a stage payoff of at least  $u_m$  in each of these periods, and least  $u_1^*$  in all the other periods. This immediately gives that player 1's payoff from deviating is at least  $\delta^{\kappa(\mu^*)}u_1^* + (1 - \delta^{\kappa(\mu^*)})u_m$ . Since we have a Nash equilibrium, player 1's payoff in equilibrium is at least his payoff from deviating.

An argument similar to the one above establishes the second bound. The idea is to obtain a version of Lemma ?? with  $\omega^*$  replaced by  $\omega_0$ . Players may be surprised by the behavior of type  $\omega_0$  only a finite number of times. In all other periods, they must play a best response to the expected play of  $\omega_0$ .  $\square$

Fudenberg and Levine (1992) extend the result to mixed strategy Nash equilibria. In generic games, the lower and upper Stackelberg payoffs coincide and we get a unique equilibrium payoff for the rational player 1 in the limit as  $\delta \rightarrow 1$ .

### 31. REPUTATION AND BARGAINING

Abreu and Gul (2000) consider reputation in the context of bargaining. Two players need to divide \$1. Every player can be either rational or a crazy type who always demands a fixed share of the dollar. Each player wants to develop a reputation for being irrational in order to get his opponent to concede to his demand.

The bargaining protocol is very general. Every player is allowed to make offers at a discrete set of dates. The analysis focuses on the continuous time limit in which each player gets the opportunity to make offers in every time interval. It turns out that the details of the bargaining protocol do not affect the limit outcomes.

Abreu and Gul show that whenever either player  $i$  has revealed himself to be rational by doing anything other than demanding  $\alpha_i$ , there will be almost immediate agreement:  $j$  can get himself a share close to  $\alpha_j$  by continuing to use his reputation, leading  $i$  to concede quickly in equilibrium. This is similar to the Fudenberg-Levine reputation result, but it turns out to be complicated to prove. So what happens in equilibrium if both players are rational? They play a war of attrition—each player pretends to be irrational but has some probability of conceding at each period (by revealing rationality), and as soon as one concedes the ensuing payoffs are those given by the reputation story. These concession probabilities

must make each player indifferent between conceding and not; from this we can show that the probabilities are stationary, up to some finite time, and if both players have not conceded by that time they must be irrational (and so will never concede).

The setting is as follows. There are two players  $i = 1, 2$ . Player  $i$  has discount rate  $r_i$ . If an agreement  $(x_1, x_2)$  is reached at time  $t$ , the payoffs (if the players are rational) are  $(x_1e^{-r_1t}, x_2e^{-r_2t})$ . Each player  $i$ , in addition to his rational type, has an irrational type, whose behavior is fixed: this type always demands  $\alpha_i$ , and always accepts offers that give him at least  $\alpha_i$  and rejects lower offers. We assume  $\alpha_1 + \alpha_2 > 1$ . The prior probability that player  $i$  is irrational is  $z_i$ .

We consider bargaining protocols that are a generalization of the Rubinstein alternating-offers protocol. A protocol is given by a function  $g : [0, \infty) \rightarrow \{0, 1, 2, 3\}$ . If  $g(t) = 0$ , then nothing happens at time  $t$ . If  $g(t) = 1$  then player 1 makes an offer, and 2 immediately decides whether to accept or reject. If  $g(t) = 2$  then the same happens with players 1 and 2 reversed. If  $g(t) = 3$  then both players simultaneously offer. If their offers are incompatible (the amount player 1 demands plus the amount player 2 demands exceeds 1) then both offers are rejected and the game continues; otherwise each player gets what he demands and the remaining surplus is split equally.

The protocol is discrete, meaning that for every  $t$ ,  $g^{-1}(\{1, 2, 3\}) \cap [0, t)$  is finite. A sequence of such protocols  $(g_n)$  converges to the continuous limit if, for all  $\varepsilon > 0$ , there exists  $n^*$  such that for all  $n > n^*$ , and for all  $t$ ,  $\{1, 2\} \subseteq g_n([t, t + \varepsilon])$ . For example, this is satisfied if  $g_n$  is the Rubinstein alternating protocol with time increments of  $1/n$  between offers. As Abreu and Gul show, each  $g_n$  induces a game with a unique equilibrium outcome, and these equilibria converge to the unique equilibrium outcome of the continuous-time limit game.

The continuous-time limit game is a war of attrition. Each player initially demands  $\alpha_i$ . At any time, each player can concede or not. Thus, rational player  $i$ 's strategy is a probability distribution over times  $t \in [0, \infty]$  at which to concede (given that  $j$  has not already conceded);  $t = \infty$  corresponds to never conceding. If player  $i$  concedes at time  $t$ , the payoffs are  $(1 - \alpha_j)e^{-r_it}$  for  $i$  and  $\alpha_j e^{-r_jt}$  for  $j$ . With probability  $z_i$ , player  $i$  is the irrational type who never concedes. (If there is no concession, both players get payoff 0.)

Bargaining follows a Coasian dynamics once one of the players is revealed to be rational. The Coase conjecture asserts that when the time between offers is sufficiently small, bargaining between a seller with known valuation  $v$  and a buyer who may have one of many reservation values, all greater than  $v$ , results in almost immediate agreement at the lowest buyer valuation. Myerson's (1991) text (pp. 399-404) offers a different perspective on this result by recasting it in a reputational setting. The low valuation buyer is replaced by an irrational type who demands some constant amount and accepts no less than this amount. In an alternating offer bargaining game, he shows that as the time between offers goes to zero, agreement is reached without delay at the constant share demanded by the irrational type. Similarly, in the Coase conjecture there is immediate agreement at the lowest buyer valuation. Both results are independent of the ex ante probability of the low type and the players' relative discount factors so long as they are both close to 1, as implied by the assumption that offers are frequent. Thus, Myerson observes that the influence of asymmetric information overwhelms the effect of impatience in determining the division of surplus. Abreu and Gul extend Myerson's result as follows.

**Lemma 4.** *For any  $\varepsilon > 0$ , if  $n$  is sufficiently high, then after any history in  $g_n$  where  $i$  has revealed rationality and  $j$  has not, in equilibrium play of the continuation game,  $i$  obtains at most  $1 - \alpha_j + \varepsilon$  and  $j$  obtains at least  $\alpha_j - \varepsilon$ .*

*Proof.* Consider the equilibrium continuation play starting from some history at which  $i$  has revealed rationality and  $j$  has not as of time  $t$ . It is sufficient to show that player  $j$ 's payoff if he continues to act irrationally converges to  $\alpha_j$  as  $n \rightarrow \infty$ . Let  $\hat{t}$  be any time increment such that, with positive probability (in this continuation), the game still has not ended at time  $t + \hat{t}$ . We will first show that there is an upper bound on  $\hat{t}$ .

Let  $\pi$  be the probability that  $j$  does not reveal rationality under the equilibrium strategies in the interval  $[t, t + \hat{t})$ . Then  $i$ 's expected continuation payoff as of time  $t$  satisfies  $v_i \leq 1 - \pi + \pi e^{-r_i \hat{t}}$ . We also have  $v_i \geq (1 - \alpha_j)z_j^t$  where  $z_j^t$  is the posterior that  $j$  is irrational as of time  $t$ , since  $i$  could get this much by immediately conceding. Then

$$1 - \pi + \pi e^{-r_i \hat{t}} \geq (1 - \alpha_j)z_j^t \geq (1 - \alpha_j)z_j.$$

It must be that  $\pi$  is bounded above by some  $\bar{\pi} < 1$  for large enough  $\hat{t}$ .

Now we apply the reasoning from Fudenberg and Levine (1989). Assume  $\hat{t}$  is large enough that  $j$  always has a chance to offer in any interval of length  $\hat{t}$ . Each time an interval of length  $\hat{t}$  goes by without  $j$  conceding, the posterior probability that  $j$  is irrational increases by a factor of at least  $1/\bar{\pi} > 1$ . The number of such increases that can occur is bounded above (by  $\ln(z_i)/\ln(\bar{\pi})$ ). Thus there is an upper bound on the amount of time the game can continue, as claimed.

The argument above shows that for every  $n$ , if player  $j$  continues to behave irrationally, player  $i$  concedes in finite time  $\bar{t}(n)$ . This time depends on  $n$  since we chose  $\hat{t}$  such that  $j$  has an opportunity to make an offer. We next show that  $\bar{t}(n)$  converges to zero as  $n$  increases.

Consider the last  $\epsilon$  units of time before player  $i$  would concede with certainty if  $j$  sticks to his demand. Without loss of generality, assume  $r_j = 1$  and  $r_i = r$ . Since with probability at least  $z_j$  player  $j$  is irrational, with positive probability player  $i$  is using some strategy that does not end the game for at least  $\epsilon$  longer. Fix  $\beta \in (0, 1)$ . The expected payoff from such strategy is at most

$$(1 - \zeta)x + \zeta y$$

where

- $x$  is  $i$ 's expected payoff if  $j$  agrees to an offer worse than  $\alpha_j$  by time  $\beta\epsilon$ ;
- $y$  is  $i$ 's payoff if  $j$  does not agree to such an offer by time  $\beta\epsilon$ ;
- $\zeta$  is the probability  $i$  assigns to the latter event.

For  $i$  to have incentives to wait out  $\epsilon$  more time rather than accept the offer  $\alpha_j$ , it must be that

$$(31.1) \quad 1 - \alpha_j \leq (1 - \zeta)x + \zeta y.$$

If  $j$  agrees to a payoff less than  $\alpha_j$ , then  $i$  will find out he is rational. But then  $j$  knows that if he holds out for  $\epsilon$  longer, he will get  $\alpha_j$  and, hence,  $j$ 's payoff is at least  $e^{-\epsilon}\alpha_j$ . Therefore  $x \leq 1 - e^{-\epsilon}\alpha_j$ . Similarly, if  $j$  does not agree to an offer by  $\beta\epsilon$ , then the best that  $i$  can do after that time is  $1 - e^{-(1-\beta)\epsilon}\alpha_j$ . So  $y \leq e^{-\beta r\epsilon}(1 - e^{-(1-\beta)\epsilon}\alpha_j)$ . Note that  $y < 1 - \alpha_j$  reduces to

$$\alpha_j < \frac{1 - e^{-\beta r\epsilon}}{1 - e^{-(\beta r + 1 - \beta)\epsilon}}.$$

In the limit  $\varepsilon \rightarrow 0$  the latter inequality becomes  $\alpha_j < \beta r / (\beta r + 1 - \beta)$ , which is satisfied for  $\beta > \alpha_j / (\alpha_j + r(1 - \alpha_j))$ .

Plugging our bounds into inequality (??) we get

$$1 - \alpha_j \leq 1 - e^{-\varepsilon} \alpha_j + \zeta (e^{-\beta r \varepsilon} (1 - e^{-(1-\beta)\varepsilon} \alpha_j) - (1 - e^{-\varepsilon} \alpha_j)).$$

Note that for small  $\varepsilon$ , we have  $e^{-\beta r \varepsilon} (1 - e^{-(1-\beta)\varepsilon} \alpha_j) < 1 - \alpha_j < 1 - e^{-\varepsilon} \alpha_j$ , and hence the coefficient of  $\zeta$  above is negative. Reorganizing the terms to get a bound for  $\zeta$  and taking the limit  $\varepsilon \rightarrow 0$  we obtain that

$$(31.2) \quad \zeta \leq \frac{\alpha_j}{\beta(\alpha_j + (1 - \alpha_j)r)} + k(\varepsilon)$$

where  $\lim_{\varepsilon \rightarrow 0} k(\varepsilon) = 0$ .

Fix  $\beta \in \left( \frac{\alpha_j}{\alpha_j + (1 - \alpha_j)r}, 1 \right)$  and let  $\bar{\varepsilon} > 0$  be small enough so that the right-hand side of the inequality (??) is strictly less than some  $\delta < 1$  whenever  $\varepsilon < \bar{\varepsilon}$ .

Suppose we have a history at time  $t$  where  $i$  has revealed rationality,  $j$  has not, and the latest possible end of the game (if  $j$  continues not conceding) is  $t + \varepsilon$ . If  $j$  does not reveal rationality by time  $t + \beta\varepsilon$ , the posterior probability that  $j$  is irrational must increase by at least a factor  $1/\delta > 1$ . If  $j$  does not reveal rationality by the time  $(1 - \beta)^2\varepsilon$  before the end of the game, the posterior probability of irrationality must increase by another factor of  $1/\delta$ , and so forth. There can only be some number  $k$  of such increments before the posterior belief exceeds 1. Note that our argument for  $i$ 's incentives assumes that  $i$  is able to make offers sufficiently close to each of the cutoffs in the argument above.

As  $n \rightarrow \infty$ , because the offers in the games  $g_n$  become increasingly frequent, the corresponding upper bounds on  $\varepsilon$  go to 0. Thus, once  $i$  has revealed rationality, the maximum amount of time that it can take before the game ends if  $j$  continues to act irrationally goes to 0 as  $n \rightarrow \infty$ . This means that by acting irrationally,  $j$  can guarantee himself a payoff arbitrarily close to  $\alpha_j$  for  $n$  sufficiently high.  $\square$

This leads (with a little further technical work) to the result that the continuous-game equilibrium is the limit of the discrete-game equilibria. When one agent is known to be rational and there is a positive probability that her opponent is irrational, delay is not possible. This means that either the player  $i$  known to be rational gives in to the irrational demand of the other player  $j$  or player  $j$  also reveals himself to be rational. The latter

outcome occurs only when  $j$  receives a payoff no less than  $\alpha_j$  when he reveals himself to be rational. Otherwise,  $j$  prefers to pretend to be irrational and be conceded to by  $i$  without delay.

With this conclusion in place, a war of attrition emerges: at any time  $t$ , player  $i$  can continue pretending to be irrational and wait for  $j$  to make the offer  $1 - \alpha_j$  or some offer that reveals  $j$ 's rationality. In either situation,  $i$  obtains a payoff no less than  $1 - \alpha_j$  since  $\alpha_i > 1 - \alpha_j$  by assumption and  $i$  obtains a payoff of at least  $\alpha_i$  once  $j$  reveals himself to be rational. Thus, player  $i$ 's equilibrium payoff before either player reveals rationality is at least  $1 - \alpha_j$ . Whenever  $j$  decides to reveal his rationality,  $i$  receives a payoff of at least  $\alpha_i$ , which means that *in equilibrium*  $i$  and  $j$  get exactly  $\alpha_i$  and  $1 - \alpha_i$ , respectively. But this is precisely the set-up of a war of attrition where  $i$ 's winning payoff is  $\alpha_i$  and her losing payoff is  $1 - \alpha_j$ .

It remains to analyze the continuous-time war of attrition. This is a well-known game, but with the twist that there are irrational types. In equilibrium, let  $F_i$  denote the cdf of times when  $i$  concedes—unconditional on  $i$ 's type; thus  $\lim_{t \rightarrow \infty} F_i(t) \leq 1 - z_i$  because the irrational player never concedes.

What are the rational player  $i$ 's payoffs from holding out until time  $t$ , then conceding? We get

$$u_i^i = \alpha_i \int_0^{t^-} e^{-r_i y} dF_j(y) + \frac{1}{2}(\alpha_i + 1 - \alpha_j)(F_j(t) - F_j(t^-)) + (1 - \alpha_j)(1 - F_j(t))e^{-r_i t}$$

(these terms correspond to  $i$  winning, both players conceding at the same time, and  $j$  winning, respectively). Assuming that  $F_1$  and  $F_2$  have a common support and are continuous, it must be that each  $u_t^i$  is constant and differentiable over the support. Taking derivatives with respect to  $t$  we obtain that  $dF_j/(1 - F_j) = r_i(1 - \alpha_j)/(\alpha_1 + \alpha_2 - 1) := \lambda_j$ . The concession rate of player  $j$  makes the rational type of player  $i$  indifferent about conceding everywhere on his support. Let  $(\hat{F}_1, \hat{F}_2)$  be the unique strategy profile satisfying  $T^0 = \min\{(-\log z_1)/\lambda_1, (-\log z_2)/\lambda_2\}$ ,  $c_i = z_i e^{\lambda_i T^0}$ , and  $\hat{F}_i(t) = 1 - c_i e^{-\lambda_i t}$ . The next result proves that these functions characterize the unique equilibrium. The proof resembles the argument showing equilibrium uniqueness in the first-price auction.

**Proposition 5.** *The unique sequential equilibrium is  $(\hat{F}_1, \hat{F}_2)$ .*

*Proof.* Let  $\bar{\sigma} = (F_1, F_2)$  define a sequential equilibrium. We will argue that  $\bar{\sigma}$  must have the form specified (i.e., uniqueness) and that these strategies do indeed define an equilibrium (existence).

Let  $u_s^i$  denote the expected utility of a rational player  $i$  who concedes at time  $s$ . Define  $A^i \equiv \{t | u_t^i = \max_s u_s^i\}$ . Since  $\bar{\sigma}$  is an equilibrium,  $A^i \neq \emptyset$  for  $i = 1, 2$ . Also, let  $\tau^i = \inf\{t \geq 0 | F_i(t) = \lim_{t' \rightarrow \infty} F_i(t')\}$ , where  $\inf \emptyset \equiv \infty$ .

- (a)  $\tau^1 = \tau^2$ . A rational player will not delay conceding once she knows that her opponent will never concede.
- (b) If  $F_i$  jumps at  $t \in \mathbb{R}$ , then  $F_j$  does not jump at  $t$ . If  $F_i$  had a jump at  $t$ , then player  $j$  receives a strictly higher utility by conceding an instant after  $t$  than by conceding exactly at  $t$ .
- (c) If  $F_i$  is continuous at  $t$ , then  $u_s^j$  is continuous at  $s = t$  for  $j \neq i$ . This follows immediately from the definition of  $u_s^i$ .
- (d) There is no interval  $(t', t'')$  such that  $0 \leq t' < t'' \leq \tau^1$  where both  $F_1$  and  $F_2$  are constant on the interval  $(t', t'')$ . Assume the contrary and without loss of generality, let  $t^* \leq \tau^1$  be the supremum of  $t''$  for which  $(t', t'')$  satisfies the above properties. Fix  $t \in (t', t^*)$  and note that for  $\varepsilon$  small there exists  $\delta > 0$  such that  $u_t^i - \delta \geq u_s^i$  for all  $s \in (t^* - \varepsilon, t^*)$  for  $i = 1, 2$ . By (b) and (c) there exists  $i$  such that  $u_s^i$  is continuous at  $s = t^*$ . Hence, for some  $\eta > 0$ ,  $u_s^i < u_t^i$  for all  $s \in (t^*, t^* + \eta)$  for this player  $i$ . Since  $F_i$  is optimal, we conclude that  $F_i$  is constant on the interval  $(t', t^* + \eta)$ . The optimality of  $F_j$  then implies that  $F_j$  is constant on  $(t', t^* + \eta)$ . Hence, both functions are constant on the latter interval. This contradicts the definition of  $t^*$ .

As we noted above if  $F_i$  is constant on some interval  $(t', t'')$ , then the optimality of  $F_j$  implies that  $F_j$  is constant on  $(t', t'')$ ; consequently, (d) implies (e):

- (e) If  $t' < t'' < \tau^1$ , then  $F_i(t'') > F_i(t')$  for  $i = 1, 2$ .
- (f)  $F_i$  is continuous at  $t > 0$ . Indeed, if  $F_i$  has a jump at  $t$  then  $F_j$  is constant on the interval  $(t - \varepsilon, t)$  for  $j \neq i$ . This contradicts (e).

From (e) it follows that  $A^i$  is dense in  $[0, \tau^i]$  for  $i = 1, 2$ . From (c) and (f) it follows that  $u_s^i$  is continuous on  $(0, \tau^1]$  and hence  $u_s^i = \text{constant}$  for all  $s \in (0, \tau^1]$ . Consequently

$A^i = (0, \tau^1]$ . Hence,  $u_t^i$  is differentiable as a function of  $t$  and  $du_t^i/dt = 0 \forall t \in (0, \tau^1)$ . Now

$$u_t^i = \int_{x=0}^t \alpha_i e^{-r_i x} dF_j(x) + (1 - \alpha_j) e^{-r_i t} (1 - F_j(t)).$$

The differentiability of  $F_j$  follows from the differentiability of  $u_t^i$  on  $(0, \tau^1)$ . Differentiating in the equation above leads to

$$0 = \alpha_i e^{-r_i t} f_j(t) - (1 - \alpha_j) r_i e^{-r_i t} (1 - F_j(t)) - (1 - \alpha_j) e^{-r_i t} f_j(t)$$

where  $f_j(t) = dF_j(t)/dt$ . This in turn implies  $F_j(t) = 1 - c_j e^{-\lambda_j t}$  where  $c_j$  is yet to be determined. At  $\tau^1 = \tau^2$  optimality for player  $i$  implies that  $F_i(\tau^i) = 1 - z_i$ . If  $F_j(0) > 0$  then  $F_i(0) = 0$  by (b), so  $c_i = 1$ . Let  $T^i$  solve  $1 - e^{-\lambda_i T^i} = 1 - z_i$ . Then  $\tau^1 = \tau^2 = T^0 \equiv \min\{T^1, T^2\}$  and  $c_i, c_j$  are determined by the requirement  $1 - c_i e^{-\lambda_i T^0} = 1 - z_i$ . So  $F_i = \hat{F}_i$  for  $i = 1, 2$ . If  $j$ 's strategy is  $\hat{F}_j$ , then  $u_t^i$  is constant on  $(0, \tau^1]$  and  $u_s^i < u_{T^0}^i \forall s > \tau^1$ . Hence any mixed strategy on this support, and, in particular,  $\hat{F}_i$  is optimal for player  $i$ . Hence  $(\hat{F}_1, \hat{F}_2)$  is indeed an equilibrium.  $\square$

#### Properties of the equilibrium

- At most one player concedes at time 0. If both conceded at time 0 with positive probability, then either player would prefer to wait and concede later; the loss from waiting is negligible while the gain from winning is discrete.
- There is no interval of time in which neither player concedes, but such that concessions do happen later with positive probability. There is also no interval during which only one player concedes with positive probability. Neither player's concession time distribution has a mass point on any positive time.
- After 0, each player concedes with a constant hazard rate. Moreover,  $i$  has to concede at a rate that makes  $j$  indifferent to conceding everywhere on his support. Writing down  $j$ 's local indifference condition, we see that it uniquely determines  $i$ 's instant hazard rate of concession:  $\lambda_i = r_j(1 - \alpha_j)/(\alpha_1 + \alpha_2 - 1)$ .
- Both players stop conceding at the same time, at which point they are both known to be irrational. This is because if player  $i$  continued to concede after  $j$  could no longer concede, then player  $i$  would prefer to deviate by conceding earlier. So they stop conceding at the same time, and no agreement can occur after that time. If  $i$

still had positive probability of being rational at that point, then  $i$  would prefer to continue conceding rather than not reaching an agreement.

The constant-hazard-rate finding tells us that  $F_i$  must have the form  $F_i(t) = 1 - c_i e^{-\lambda_i t}$  for some constant  $c_i$ . The constants  $c_1, c_2$  can be computed from the fact that both players become known to be irrational at the same time ( $F_1^{-1}(1 - z_1) = F_2^{-1}(1 - z_2)$ ) and that only one player can concede with positive probability at time 0 (so either  $c_1$  or  $c_2$  is 1).

If player  $i$  concedes with positive probability at time 0, then the expected equilibrium payoff of the rational type of player  $i$  must be  $1 - \alpha_j$ . Moreover,  $j$ 's indifference for conceding at any positive time implies that his ex ante expected payoff is  $F_i(0)\alpha_j + (1 - F_i(0))(1 - \alpha_i)$ . Note that  $T^i = -\log z_i/\lambda_i$  measures player  $i$ 's weakness, since  $T^i > T^j$  means that player  $i$  concedes at time 0. A more patient player has a stronger bargaining position. Indeed, a lower  $r_i$  implies a lower concession rate  $\lambda_j = r_i(1 - \alpha_i)/(\alpha_1 + \alpha_2 - 1)$  for player  $j$ . This, in turn, leads to an increase in  $T_j$  and the probability  $1 - c_j$  of  $j$  conceding at time 0. Hence, if player  $j$  concedes at time 0 ( $c_j < 1$ ), then a small decrease in  $r_i$  increases the probability of  $j$  conceding immediately. If player  $i$  concedes at time 0 ( $c_i < 1$ ), then a small decrease in  $r_i$  would lead to a reduction in the probability with which  $i$  concedes. In either case, a decrease in  $r_i$  makes player  $i$  better off and player  $j$  worse off. An increase in  $z_i$  has the same effect.

Finally, note that the equilibrium exhibits delay and hence inefficiency. Consider the symmetric case  $r_1 = r_2, \alpha_1 = \alpha_2 = \alpha, z_1 = z_2 = z$ . Then, in equilibrium,  $F_1(0) = F_2(0) = 0$ . The expected payoff of a rational player 1 is  $1 - \alpha$  since conceding at zero is in the support of his optimal concession times. The payoff of an irrational player cannot exceed  $1 - \alpha$  since it is bounded above by the payoff of a rational player who concedes at time  $T^0$ . Thus the expected payoff of either player is less than  $1 - \alpha$  and the total utility loss is in excess of  $2\alpha - 1$ , which may be substantial. The inefficiency is a consequence of delay to reaching agreement rather than not reaching agreement at all; the ex ante probability of disagreement is just  $z^2$ .

Abreu and Gul generalize the analysis to multiple irrational types. They show that existence and uniqueness of the equilibrium in the continuous time game extend to the more general model. In equilibrium players may mix between several irrational types and need not choose the type with the most extreme demand. The multiple types need to be mimicked with appropriate weights. The resulting posterior probabilities modulate the relative

strengths of the types such that all types mimicked with positive probability obtain the same equilibrium payoff.

The “strength” of a player depends upon the posterior probability of the type she mimics, and the latter probability decreases with the probability with which that type is mimicked. The payoffs to a type being conceded to with positive probability at time zero are strictly increasing in “strength.” Multiple equilibrium distributions over types being conceded to are in conflict with the requirement that types mimicked with positive probability must have equal payoffs that are not smaller than the payoffs of types that are not mimicked.

When the probability of irrational types is vanishingly small, one could think of the reputational bargaining model as a perturbation of (a more general version of) Rubinstein’s bargaining model. Abreu and Gul show that the complete information model is robust to such perturbations if type spaces are sufficiently rich. More precisely, if for any division of surplus there is an irrational type that makes a demand close to this division of surplus, then the inefficiency (in terms of delay) disappears as the probability of irrational types vanishes. In the limit, in both models, rational players choose to be virtually compatible and share surplus in proportion to impatience, i.e., player  $i$  mimics types close to  $r_i/(r_1 + r_2)$  with limit probability 1. This result confirms an earlier finding by Kambe (1994) who considers the model in which players are initially unrestricted in their demands and could gain commitment to the chosen posture later in the game.

MIT OpenCourseWare  
<https://ocw.mit.edu>

14.126 Game Theory  
Spring 2016

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.