Economic Notes by Monte dei Paschi di Siena, vol. 22, no. 3-1993, pp. 430-463

Rational Belief Equilibria

GIACOMO BONANNO*

Given an extensive game, from an assessment (σ, μ) (as defined by Kreps and Wilson, 1982) we obtain a belief for each player, defined as a map from the set of all nodes into the family of subsets of the set of terminal nodes. An assessment is defined to be a "Rational Belief Equilibrium" if, for each player, the associated belief satisfies three natural consistency properties. The two main results are that the notion of rational belief equilibrium strictly refines that of subgame-perfect equilibrium and that, in turn, sequential equilibria are a strict refinement of rational belief equilibria.

1. *Introduction*

Selten (1965, 1975) was the first one to point out one problem with the notion of Nash equilibrium in extensive games, namely the fact that it places no restrictions on choices at information sets that are not reached by the equilibrium path. The concept of subgame-perfect equilibrium (Selten 1965) constituted the first step in the general program of dealing with this problem. A stronger solution concept, widely used in the literature, is sequential equilibrium (Kreps and Wilson, 1982). The formal definition of sequential equilibrium is in terms of an assessment (σ, μ) where σ is a strategy profile and μ is a list of probability distributions, one for each information set. An assessment is a sequential equilibrium if it is sequentially rational and consistent. The substance of sequential rationality is that "the strategy of each player starting from each information set must be optimal starting from there according to some assessment over the nodes in the information set and the strategies of everyone else" (Kreps an Wilson, 1982, p. 871).

The notion of consistency places restrictions on out-of-equilibrium beliefs, by requiring μ to be the limit of a sequence of "Bayesian beliefs" obtained from a sequence of completely mixed strategies converging to the strategies under consideration. The consistency requirement is not without problems, both at the

* Department of Economics, University of California, Davis, Ca 95616-8578, U.S.A. The author is grateful to an anonymous referee for helpful and constructive comments.

practical and at the theoretical level. From a practical point of view, verifying consistency is a tedious process and in fact in many applications the focus is on sequential rationality, while often milder or no restrictions are imposed on beliefs. At the theoretical level, on the one hand - as Kreps and Ramey (1987, p. 1333) observe - "consistency itself does not encompass all the properties it was originally thought to", and, on the other hand - as Fudenberg and Tirole (1991b, p. 346) point out - "one would like to know more about what consistency implies for behavior".

In applications economists have often used a weaker notion of equilibrium, sometimes referred to as "perfect Bayesian equilibrium". However, there doesn't seem to be a well established definition of perfect Bayesian equilibrium. The weakest definition only requires sequential rationality together with Bayesian updating along the equilibrium path (see, for example, Rasmusen, 1989, p. 110), so that no restrictions at all are placed on out-of-equilibrium beliefs.

Recently Fudenberg and Tirole (1991a) have suggested a new definition of perfect Bayesian equilibrium for multi-period games of incomplete information with observed actions and studied the relationship between this notion and sequential equilibrium. They also suggested a way of extending their definition to general extensive games[1]. While the restriction they place on out-of-equilibrium beliefs has a very intuitive interpretation for the class of games they consider, namely that players should not signal what they do not know, the definition they propose seems to embody more than minimal "consistency" requirements on beliefs.

In this paper we put forward a definition of equilibrium - "rational belief equilibrium" - based entirely on properties of beliefs. We will show that (like perfect Bayesian equilibrium) our notion of equilibrium is stronger than subgame-perfect equilibrium but weaker than sequential equilibrium. Furthermore, our definition applies to general extensive games.

Our approach differs from the standard one and builds on the concepts introduced in Bonanno (1992a,b). The first of these two papers raises the question of what information the players receive during the play of an extensive game. Fix an extensive game and let Z be the set of terminal nodes. For every player i and for every node t, the information received by player i when the play of the game reaches node t is defined as a subset of Z, with the following interpretation. Suppose that, when node t is reached, player i receives information $\{z_1, z_3, z_7\}$. Then this means that player i is informed that the play of the game so far has been such that only terminal nodes z_1, z_3 or z_7 can be reached. In Bonanno (1992a) the main concern is with the notions of minimum

[1] Battigalli (1993) clarifies the relationship between "generally reasonable extended assessments" and sequential equilibria.

and maximum amount of information that can be conveyed to the players as well as a characterization of the notions of perfect information, perfect recall and simultaneity.

The second paper (Bonanno, 1992b) borrows one of the definitions of information suggested in the first paper, denoted by $K_i(t)$ (as explained above, for every player i and for every node t, $K_i(t)$ is a subset of Z, the set of terminal nodes), and introduces the notion of minimally rational profile of beliefs. A belief of player i is defined there as a function that associates with every node t an element of the set $K_i(t)$, denoted by $\beta_i(t)$. The interpretation is that if, say, $K_i(t)=\{z_1,z_3,z_7\}$ and $\beta_i(t)=z_3$ then player i knows (is informed) that the play of the game can only end either at node z_1 or z_3 or z_7 and believes that the outcome will actually be z_3. From a profile of beliefs one can extract a pure strategy profile in a natural way. Bonanno (1992b) only considers extensive games without chance moves and defines a profile of beliefs to be minimally rational if it satisfies four simple consistency properties. The main result of that paper is that if the profile of beliefs is minimally rational then the corresponding pure strategy profile is a subgame-perfect equilibrium.

In this paper we continue the analysis of Bonanno (1992b). First of all, we extend it to games with chance moves. Secondly, we change perspective: instead of starting from a profile of beliefs and extracting from it a strategy profile, we start from the notion of assessment introduced by Kreps and Wilson (1982) and extract from it a profile of beliefs as defined above. We then define an assessment to be a *rational belief equilibrium* if the associated profile of beliefs satisfies three of the four properties introduced in Bonanno (1992b). The main results of this paper are that (1) the notion of rational belief equilibrium refines that of subgame-perfect equilibrium, and (2) the notion of sequential equilibrium refines that of rational belief equilibrium.

In order to keep the exposition as semple as possible, we shall concentrate on pure beliefs (defined in section 2) and simple assessments (defined in section 3).

2. *Preliminary Definitions*

We begin by reviewing the notation and some of the definitions of Bonanno (1992a,b). Fix a finite extensive game. Let X be the set of *decision* nodes, Z the set of *terminal* nodes, and $T=X \cup Z$. [In general, we shall denote a decision node by x or y, a terminal node by z and a generic node - decision or terminal - by t]. For every $t \in T$, let $\theta(t) \subseteq Z$ be the set of terminal nodes that can be reached from t (for example, in the game of Figure 1, $\theta(x_3) = \{z_3,z_4,z_5,z_6\}$). Clearly, for every $z \in Z$, $\theta(z) = \{z\}$.

Recall that a choice c at information set $h=\{x_1,...x_m\}$ is a set of arcs $c=\{(x_1,y_1), (x_2,y_2),..., (x_m,y_m)\}$ where, for each $k=1,..., m$, node y_k is an immediate successor of node x_k. Define

$$\gamma(c) = \theta(y_1) \cup \theta(y_2) \cup ... \cup \theta(y_m),$$

that is, $\gamma(c)$ is the set of terminal nodes that can be reached from nodes in h by following the arcs that constitute choice c. For example, in the game of Figure 1, $\gamma(E) = \{z_1,z_3,z_4\}$.

We denote by x_0 the root of the tree and for every node $t \neq x_0$ we shall denote the immediate predecessor of t by p_t. Finally, for every node t and for every player i, $H_i(t)$ is the set of information sets *of player i* that satisfy the following property: $h \in H_i(t)$ if and only if there is a node $y \in h$ that is a successor of t.

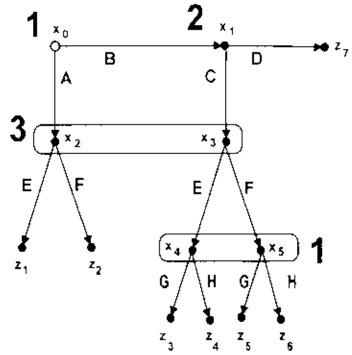

Figure 1

The information received by player i when the play of the game reaches node t is denoted by $K_i(t)$. The function $K: I \times T \to 2^Z$ (where I is the finite set of players and 2^Z denotes the set of subsets of Z) is defined as follows:[2]

(1) For every player i set $K_i(x_0) = Z$.

[2] For a more thorough discussion see Bonanno (1992a). One way of thinking about the proposed definition is as follows. At the root of the tree all players have the same information, namely Z. As the play of the game unfolds and new nodes are reached, an umpire gives (separately) to each player new information according to the following rules. If z is a terminal node, then every player is informed that the game ended at z. If node x belongs to information set h of player i, then player i is told that her information set h has been reached, but is not told which node in h was reached. If node x does not belong to player i and all the information sets of player i (if any) that are crossed by paths starting at x consist entirely of nodes that are successors of x, then player i is informed that node x has been reached (the justification for this rule is that, later on, at any of her information sets, player i will be able to deduce that the play of the game must have gone through node x; hence player i might as well be told at the time when x is reached). When the above condition is not satisfied, player i's information at x either doesn't change (that is, player i is not told anything new) or at most reflects the choice made by player i at the immediate predecessor of x, if that node belonged to player i.

(2) For every $z \in Z$ and for every player i, set $K_i(z)=\{z\}$.

(3) If x is a decision node that belongs to information set h *of player i*, set $K_i(x)= \bigcup_{y \in h} \theta(y)$, that is, $K_i(x)$ is the set of terminal nodes that are successors of nodes in h.

(4) If x is a decision node of a player different from player i and either $H_i(x)=\emptyset$ or, for every $h \in H_i(x)$, $\bigcup_{y \in h} \theta(y) \subseteq \theta(x)$ (that is, every node in h is a successor of x) then set $K_i(x)=\theta(x)$.

(5) If x is a decision node of a player different from player i and the condition given under (4) is not satisfied (that is, there exists an $h \in H_i(x)$ and a node $y \in h$ such that y is *not* a successor of x) and x is an immediate successor of decision node t of player i and c is the choice of player i that leads from t to x, then set $K_i(x) = \gamma(c)$, that is, $K_i(x)$ is the set of terminal nodes that can be reached from the information set containing node t by following choice c.

(6) Finally, if x is a decision node of a player different from player i and it does not satisfy conditions (4) and (5), then set $K_i(x)=K_i(p_x)$, that is, player i's information at node x is the same as it was at p_x, the immediate predecessor of x.

For example, in the game of Figure 1 we have:

By (1): $K_i(x_0) = Z = \{z_1,z_2,z_3,z_4,z_5,z_6,z_7\}$ for all i = 1,2,3.

By (2): $K_i(z_j) = \{z_j\}$ for all i = 1,2,3 and for all j = 1, ..., 7.

By (3): $K_2(x_1) = \theta\,(x_1) = \{z_3,z_4,z_5,z_6,z_7\}$.

By (4): $K_1(x_1) = \theta\,(x_1) = \{z_3,z_4,z_5,z_6,z_7\}$.

By (6): $K_3(x_1) = K_3(x_0) = Z$.

By (4): $K_1(x_3) = K_2\,(x_3) = \theta\,(x_3) = \{z_3,z_4,z_5,z_6\}$.

By (3): $K_3(x_3) = \theta\,(x_2) \cup \theta\,(x_3) = \{z_1,z_2,z_3,z_4,z_5,z_6\}$.

By (3): $K_1(x_4) = \theta\,(x_4) \cup \theta\,(x_5) = \{z_3,z_4,z_5,z_6\}$.

By (4): $K_2(x_4) = K_3(x_4) = \theta\,(x_4) = \{z_3,z_4\}$.

By (3): $K_1(x_5) = \theta(x_4) \cup \theta(x_5) = \{z_3,z_4,z_5,z_6\}$.

By (4): $K_2(x_5) = K_3(x_5) = \theta(x_5) = \{z_5,z_6\}$.

By (4): $K_1(x_2) = K_2(x_2) = \theta(x_2) = \{z_1,z_2\}$.

By (3): $K_3(x_2) = \theta(x_2) \cup \theta(x_3) = \{z_1,z_2,z_3,z_4,z_5,z_6\}$.

REMARK 1. It is clear that if h is an information set of player i, and x and y are two nodes in h, then $K_i(x)=K_i(y)$. Thus it makes sense to write $K_i(h)$ for player i's information at her information set h.

The following properties are proved in Bonanno (1992a).

PROPERTY 1. For every node t and for every player i, $\theta(t) \subseteq K_i(t)$.

PROPERTY 2. For a game with *perfect recall* the following is true: if node t is a successor of node x, then, for every player i, $K_i(t) \subseteq K_i(x)$. That is, at every node each player knows at least as much as she knew before that node was reached.

PROPERTY 3. If x is the root of a subgame, then, for every player i, $K_i(x)=\theta(x)$.

For notational simplicity, we shall follow Kreps and Wilson (1982) and assume that Nature moves at most at the root of the tree. Given an extensive game, we define the associated *set of events*, denoted by $\mathcal{E}$, as follows:

(1) if x_0 (the root) is a decision node of a personal player, then th ame has no chance moves and we set $\mathcal{E}=\{Z\}$;

(2) if Nature moves at x_0, let $c_1, c_2,..., c_m$ be Nature's choices at x_0 and let $(p(c_1),...,p(c_m))$ be the corresponding probability distributio   We shall assume that each chance move has a strictly positive probability (thus, $p(c_j) > 0$, for all $j=1, ..., m$, and $\sum_{j=1}^{m} p(c_j) = 1$).

Define $\mathcal{E} = \{\gamma(c_1), \gamma(c_2),..., \gamma(c_m)\}$. An element $E \in \mathcal{E}$ is called an event, and if $E = \gamma(c_j)$, then the probability of E, denoted by Pr(E), is defined as $Pr(E)=p(c_j)$.

For example, in the game of Figure 2 the set of events is $\mathcal{E} = \{\gamma(c_1)=\{z_1,z_2,z_3,z_4\},\ \gamma(c_2)=\{z_5,z_6,z_7\},\ \gamma(c_3)=\{z_8,z_9,z_{10},z_{11},z_{12},z_{13}\}\}$ with respective probabilities (1-p-q), p and q.

The following properties are an immediate consequence of uniqueness of plays in extensive games:

(1) if $E,E' \in \mathcal{E}$ and $E \neq E'$ then $E \cap E' = \emptyset$;

(2) for every node $t \neq x_0$ there is a unique $E \in \mathcal{E}$ such that $\theta(t) \subseteq E$; we shall denote this unique event associated with t by E(t).

Finally, let

$$\pi: T \rightarrow [0,1]$$

be defined as follows: if the game has no chance moves, that is, if $\mathcal{E} = \{Z\}$, then $\pi(t)=1$ for all $t \in T$; otherwise set $\pi(x_0)=1$ and for $t \neq x_0$ set $\pi(t) = Pr(E(t))$.

For example, in the game of Figure 2, for j=1,2,3,4 and i=1,3,7 $\pi(z_j) = \pi(x_i) = 1-p-q$, for j=5,6,7 and i=4,8 $\pi(z_j)=\pi(x_i)=p$, for j=8,9,...,13 and i=2,5,6,9,10, $\pi(z_j)=\pi(x_i)=q$.

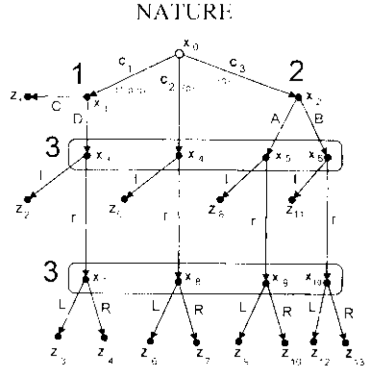

Figure 2

We now introduce the notion of belief.

DEFINITION. A *(pure) belief* of player i is a function

$$\beta_i: T \to 2^Z$$

satisfying the following properties:

(1) for every node t: $\emptyset \neq \beta_i(t) \subseteq K_i(t)$,

(2) if x and y belong to information set h of player i [so that $K_i(x) = K_i(y)$], then $\beta_i(x)=\beta_i(y)$,

(3) for every node t: if $z, z' \in \beta_i(t)$, z comes after choice c at information set h (of a personal player), z' comes after choice d at h, then c=d.

Condition (1) in the above definition says that what a player believes must be consistent with what she knows, and condition (2) says that a player cannot have different beliefs at two nodes that belong to one of her information sets, since her information is the same at both nodes. Thus it makes sense to write $\beta_i(h)$ for player i's belief at her information set h. Condition (3) is what makes a belief "pure": it says that player i believes that at every information set the relevant player will choose one action with probability 1.

It is easy to see that for games without chance moves, for every node t, $\beta_i(t)$ is a singleton. Thus for this class of games a belief of player i can be

Suppose not, that is, suppose there is a player i, a node t, and two terminal nodes z and z' such that $z,z' \in \beta_i(t)$ and $z \neq z'$ (recall that, by (1), $\beta_i(t) \neq \emptyset$). Let x be the node at which the path from x_0 to z and the path from x_0 to z' diverge. Since there are no chance moves, x is a decision node of a personal player. Let h be the information set to which x belongs. Let c be the choice at h that precedes z and c' the choice at h that precedes z'. Then it must be $c \neq c'$, violating condition (3).

defined more simply as a function

$$\beta_i : T \to Z$$

satisfying the following properties:

(i) $\beta_i(t) \in K_i(t) \quad \forall t \in T$,
(ii) if x and y belong to information set h of player i, then $\beta_i(x)=\beta_i(y)$.

When we consider a game without chance moves (or a proper subgame of a game with chance moves) we will make use of this simpler way of writing a belief of player i.

We shall employ the following notation: if h is an information set and c a choice at h, for every $x \in h$ we denote by $S(x|c)$ the immediate successor of x following choice c. Furthermore, if $\hat{h}$ is a subset of h, we denote by $\Sigma(\hat{h}|c)$ the set of immediate successors of nodes in $\hat{h}$ following choice c, that is,

$$\Sigma(\hat{h}|c) = \left\{ y \middle| y = S(x|c) \text{ for some } x \in \hat{h} \right\}.$$

From now on we shall restrict attention to games with perfect recall.

DEFINITION. We say that player i's belief β_i is *minimally rational* if it satisfies the following properties (which will be discussed immediately below):

(1) **[Contraction Consistency]** If y is a successor of x [so that, by property 2, $K_i(x) \supseteq K_i(y)$] and $\beta_i(x) \cap K_i(y) \neq \emptyset$, then

$$\beta_i(y) = \beta_i(x) \cap K_i(y).$$

(2) **[Tree Consistency]** Let h be an information set *of player i* and let $\hat{h} \subseteq h$ be the subset of h consisting of the predecessors of $\beta_i(h)$. Then for every choice c at h,

(a) $$\beta_i(y) \cap \theta(y) \neq \emptyset, \quad \forall y \in \Sigma(\hat{h}|c),$$

(b) $$\bigcup_{y \in \Sigma(\hat{h}|c)} \beta_i(y) = \bigcup_{y \in \Sigma(\hat{h}|c)} [\beta_i(y) \cap \theta(y)].$$

(3) **[Individual Rationality]** Let h be an information set of player i. Let $\hat{h}$ be the subset of h consisting of the predecessors of $\beta_i(h)$. Then for every choice c at h,

$$\sum_{z\in\beta_i(h)} U_i(z)\pi(z) \geq \sum_{z\in \bigcup_{x\in h}\beta_i(S(x\,c))} U_i(z)\pi(z)$$

where $U_i: Z \rightarrow \mathcal{R}$ is player i's payoff function ($\mathcal{R}$ denotes the set of real numbers).

Property (1) says that, as the information of a player evolves and becomes more refined, the player will not change his beliefs unless he has to, that is, unless his previous belief is inconsistent with the new information. This is a contraction consistency property which is implied, for example, by Bayesian updating.

Intuitively, property (2) can be interpreted as requiring that a player's beliefs about his opponent's previous moves be independent of his own choices. To see this, consider the example of Figure 3. There we have that $K_2(h)=\{z_1,z_2,z_3,z_4,z_5,z_6\}$ where $h=\{x_1,x_2\}$ is the first information set of player 2, and $K_2(g)=\{z_1,z_2,z_5,z_6\}$ where $g=\{x_3,x_4\}$ is the second information set of player 2.

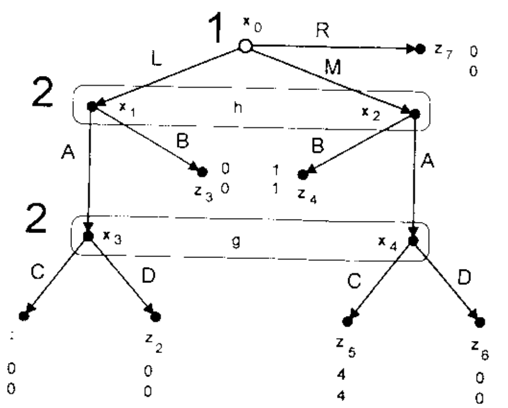

Figure 3

Suppose $\beta_2(h) = z_4$ and $\beta_2(g) = z_1$. This belief of player 2 is inconsistent because believing in z_4 at h means believing that node x_2 was reached. Given this belief, if player 2 takes action A, so that the play of the game proceeds to information set g, then node x_4 must be reached, and from x_4 terminal node z_1 cannot be reached. It is easy to see that part (a) of Tree Consistency is violated, since $\hat{h} = \{x_2\}$, $\Sigma(\hat{h}|A) = \{x_4\}$ and $\beta_2(x_4) = z_1 \notin \theta(x_4) = \{z_5, z_6\}$. In this example, part (a) of Tree Consistency requires that if $\beta_2(h) = z_4$ then either $\beta_2(g) = z_5$ or $\beta_2(g) = z_6$. Note that in games without chance moves, where $\beta_i(t)$

is a singleton for every i and for every t, part (b) of the definition of Tree Consistency is redundant, since it is implied by (a).

To see the role of part (b) in games with chance moves, consider the game of Figure 2 and the following belief of player 3: $\beta_3(x_0) = \beta_3(x_1) = \beta_3(x_2) = \{z_1,z_5,z_8\}$, $\beta_3(\{x_3,x_4,x_5,x_6\}) = \{z_5,z_8\}$, $\beta_3(\{x_7,x_8,x_9,x_{10}\}) = \{z_3,z_6,z_9\}$. Let $h = \{x_3,x_4,x_5,x_6\})$. Then the predecessors of $\beta_3(h)$ are x_4 and x_5, thus $\hat{h} = \{x_4,x_5\}$. The successors of $\hat{h}$ following choice r are x_8 and x_9. Part (a) of Tree Consistency is satisfied, since $\beta_3(x_8) \cap \theta(x_8) = \{z_6\}$ and $\beta_3(x_9) \cap \theta(x_9) = \{z_9\}$, but part (b) is not satisfied, since $\beta_3(x_8) \cup \beta_3(x_9) = \{z_3,z_6,z_9\}$, which is a proper superset of $\{z_6, z_9\}$.

To understand property (3), consider first the case of a game without chance moves. Let $z^*=\beta_i(h)$ and let x^* be the unique node in h which is on the path from the root to z^*. Then since player i believes in z^* at his information set h, it means that he believes that node x^* was reached. Property (3) requires that for every immediate successor y of x^*,

$$U_i(z^*) \geq U_i(\beta_i(y)).$$

Suppose instead that there were an immediate successor y of x^* such that $U_i(z^*) < U_i(\beta_i(y))$. Then believing in z^* (at h) is irrational for player i because, instead of making the choice required by z^*, he can - according to his beliefs and by making another choice - move the play to node y from where, again according to his beliefs, the game will evolve to outcome $\beta_i(y)$ that he prefers to z^*. When the game has chance moves, the interpretation of the inequality defining Individual Rationality is the same: the LHS represents player i's expected utility at h if he takes the choice implied by $\beta_i(h)$, while the RHS represents his expected utility if he takes choice c[1].

DEFINITION. A *profile of beliefs* is an n-tuple $\beta = (\beta_1,...,\beta_n)$ where β_i is a belief of player i, for each i=1,... n. We say that β is *minimally rational* if every β_i is minimally rational.

3. *Rational Belief Equilibrium*

Fix an extensive game with perfect recall. As in Kreps and Wilson (1982) an *assessment* is defined as a pair (σ,μ), where $\sigma = (\sigma_1,..., \sigma_n)$ is a profile of behaviour strategies and μ is a function (called a "system of beliefs" by Kreps and Wilson)

$$\mu: T \rightarrow [0,1]$$

[1] It will be shown in the proof of proposition 2 (see Appendix) that this interpretation is indeed correct.

satisfying the property that, for every information set h, $\sum_{x \in h} \mu(x) = 1$. We shall call an assessment (σ, μ) *simple* if σ is a pure-strategy profile and μ satisfies the following properties:

(i) if x and x' belong to the same information set h, x comes after choice c at information set g (of a personal player), x' comes after choice c' at g and $c \neq c'$, then either $\mu(x) = 0$ or $\mu(x') = 0$ or both;

(ii) if x and x' belong to the same information set h and $\mu(x) = 0$ while $\mu(x') > 0$, then there exist two choices c and d of personal players (not necessarily the same player, hence not necessarily at the same information set) with $c \neq d$ such that x comes after c and x' comes after d;

(iii) let h be an information set and define $supp(\mu|h) = \{t \in h | \mu(t) > 0\}$; then, for every $x \in supp(\mu|h)$

$$\mu(x) = \frac{\pi(x)}{\sum_{y \in supp(\mu|h)} \pi(y)}$$

Note that if the game has no chance moves, then property (i) implies that, for every node t, $\mu(t) = 0$ or $\mu(t) = 1$, so that properties (ii) and (iii) become redundant.

For example, in the game of Figure 4 (taken from Kreps and Wilson, 1982) every simple assessment must have $\mu(x_3) = 1/3$ and $\mu(x_4) = 2/3$.

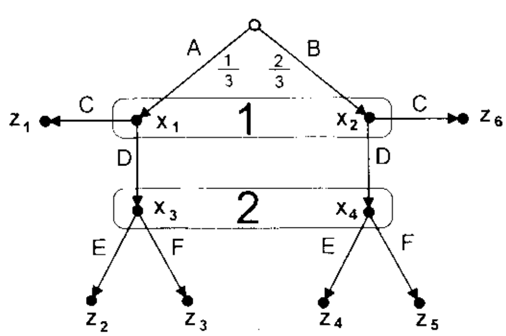

Figure 4

Given a pure-strategy profile σ, for every node $t \neq x_0$, let $\zeta(t|\sigma)$ be the unique terminal node reached from t by following σ [for every terminal node z we set by definition $\zeta(z|\sigma) = z$]. Clearly, $\zeta(t|\sigma) \in \theta(t)$. Finally, let $\Sigma(t)$ denote the set of immediate successors of node t, and recall that p_t denotes the immediate predecessor of node t.

We now show how to extract a profile of beliefs from a simple assessment.

DEFINITION [the function $\chi(\sigma,\mu)$]. Given a *simple* assessment (σ,μ) we can associate with it a profile of beliefs $\beta=\chi(\sigma,\mu)$ as follows[5]:

(1) If x_0(the root) is a decision node of a personal player (so that the game has no chance moves) set, for every player i,

$$\beta_i(x_0) = \{\zeta(x_0|\sigma)\},$$

otherwise, set

$$\beta_i(x_0) = \{z | z = \zeta(y|\sigma) \text{ for some } y \in \Sigma(x_0)\};$$

(2) If $x \neq x_0$ and $K_i(x)=\theta(x)$ [recall that, in particular, this is true if x is a terminal node], set

$$\beta_i(x) = \{\zeta(x|\sigma)\};$$

(3) If x is a decision node that belongs to information set h of player i, set

$$\beta_i(x) = \{z | z = \zeta(y|\sigma) \text{ for some } y \in \text{supp}(\mu|h)\};$$

(4) If x is a decision node that does *not* belong to player i and $\beta_i(p_x) \cap K_i(x) \neq \varnothing$ set

$$\beta_i(x) = \beta_i(p_x) \cap K_i(x).$$

(5) If x is a decision node that does *not* belong to player i and $K_i(x) \neq \theta(x)$ and $\beta_i(p_x) \cap K_i(x) = \varnothing$, then it must be $K_i(x) \neq K_i(p_x)$. It follows from the definition of the function $K_i(\bullet)$ that p_x belongs to an information set of player i, call it h, and $K_i(x) = \gamma(c)$, where c is the choice at h that leads from p_x to x[6]. Set

$$\beta_i(x) = \{z | z = \zeta(S(y|c)|\sigma) \text{ for some } y \in \text{supp}(\mu|h)\}.$$

[Recall that $S(y|c)$ denotes the immediate successor of node y following choice c.]

Example: consider the game of Figure 5 and the simple assessment (σ,μ)

[5] It is easy to check that β so constructed is indeed a profile of beliefs as defined in the previous section (property (i) of the definition of simple assessment is crucial in this respect).

[6] Since x is *not* a decision node of player i, case (3) of the definition of $K_i(\bullet)$ is ruled out. Since $K_i(x) \neq \theta(x)$, cases (1), (2) and (4) are ruled out. Finally, since $\beta_i(p_x) \cap K_i(x) = \varnothing$ implies that $K_i(x) \neq K_i(p_x)$, case (6) is also ruled out. Thus we are left with case (5).

given by $\sigma = ((B,M,P),(D,F,H))$ and $\mu(x_1) = p$, $\mu(x_2) = q$, $\mu(x_3) = 1\text{-}p\text{-}q$, $\mu(x_7) = \frac{p}{p+q}$, $\mu(x_8) = \frac{q}{p+q}$, $\mu(x_9) = 0$, $\mu(x_{10}) = 1$.

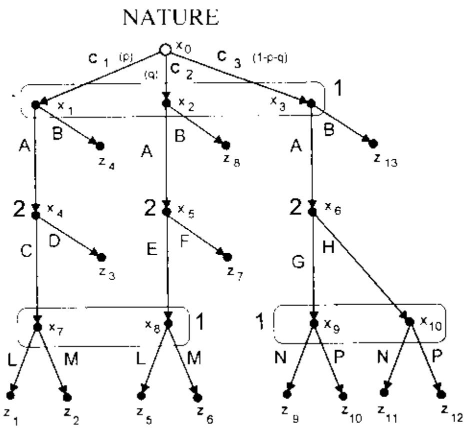

Figure 5

Then the associated profile of beliefs $\beta=\chi(\sigma,\mu)$ is as follows [Note that, for every node t, $K_2(t)=\theta(t)$]:

By (1): $\beta_1(x_0) = \beta_2(x_0) = \{z_4,z_8,z_{13}\}$;

By (2): $\beta_2(x_1) = \{z_4\}$, $\beta_2(x_2) =\{z_8\}$, $\beta_2(x_3)=\{z_{13}\}$, $\beta_2(x_4) = \{z_3\}$, $\beta_2(x_5) = \{z_7\}$, $\beta_2(x_6) =\{z_{12}\}$, $\beta_2(x_7) = \{z_2\}$, $\beta_2(x_8) = \{z_6\}$, $\beta_2(x_9) = \{z_{10}\}$, $\beta_2(x_{10}) = \{z_{12}\}$, $\beta_1(x_6) = \{z_{12}\}$;

By (3): $\beta_1(\{x_1,x_2,x_3\}) = \{z_4,z_8,z_{13}\}$, $\beta_1(\{x_7,x_8\}) = \{z_2,z_6\}$, $\beta_1(\{x_9,x_{10}\}) = z_{12}\}$;

By (5): $\beta_1(x_4)=\beta_1(x_5) =\{z_3,z_7,z_{12}\}$.

DEFINITION. A simple assessment (σ,μ) is a (pure-strategy) *rational belief equilibrium* if the associated profile of beliefs $\beta=\chi(\sigma,\mu)$ is minimally rational.

PROPOSITION 1. Let (σ,μ) be a simple assessment. If (σ,μ) is a rational belief equilibrium then σ is a (pure strategy) subgame-perfect equilibrium.

Proof. See the appendix.

The example given in Bonanno (1992b, Figure 1) can easily be adapted to show that the converse of proposition 1 is not true, that is, not every subgame-perfect equilibrium is (part of) a rational belief equilibrium.

PROPOSITION 2. Let (σ,μ) be a simple assessment. If (σ,μ) is a sequential equilibrium then it is a rational belief equilibrium.

Proof. See the appendix.

The example given in Bonanno (1992b, Figure 5) can easily be adapted to show that the converse of proposition 2 is not true, that is, not every rational belief equilibrium is a sequential equilibrium.

4. *Conclusion*

Given an extensive game, we associated with every node t and every player i a subset $K_i(t)$ of the set of terminal nodes, interpreted as player i's information when the play of the game reaches node t. A belief of player i was then defined as a map from the set of all nodes into the family of subsets of the set of terminal nodes satisfying two main properties: what a player believes must be consistent with what she knows, and a player's belief must be the same at any two nodes that belong to one of her information sets (since her information is the same at those two nodes). Three natural properties (Contraction Consistency, Tree Consistency and Individual Rationality) were used to define the notion of minimally rational belief. Having shown how to extract a belief for each player from a simple assessment (σ,μ), we defined a simple assessment to be a rational belief equilibrium if for each player the associated belief is minimally rational. The two main results of this paper are: (1) the notion of rational belief equilibrium refines that of subgame-perfect equilibrium and (2) the notion of sequential equilibrium is, in turn, a refinement of the notion of rational belief equilibrium.

APPENDIX

In this appendix we prove propositions 1 and 2. We shall begin with a few lemmas.

LEMMA 1. Fix an extensive game. Let β_i be a belief of player i. Then for every node t and for every $x \neq x_0$, if $\beta_i(t) \cap \theta(x) \neq \emptyset$ then $\beta_i(t) \cap \theta(x)$ is a singleton.

Proof. Suppose that $z,z' \in \beta_i(t) \cap \theta(x)$ with $z \neq z'$. Let y be the node at which the path from x_0 to z and the path from x_0 to z' diverge. Then y is either x itself or a successor of x. Hence y is a decision node of a personal player. Let c be the choice at y that precedes z and c' the choice at y that precedes z'. Then it must be $c \neq c'$, contradicting property (3) of the definition of $\beta_i(\bullet)$.

LEMMA 2. Let (σ,μ) be a simple assessment and let $\beta=\chi(\sigma,\mu)$ be the corresponding profile of beliefs. Then for every node $x \neq x_0$ and for every player i,

$$\beta_i(x) \cap \theta(x) \neq \emptyset \ \Rightarrow \ \beta_i(x) \cap \theta(x) = \{\zeta(x|\sigma)\}.$$

Proof. From the definition of $\chi(\bullet)$ we have that for every $x \neq x_0$ and for every player i, $\beta_i(x)=\{z|z = \zeta(t|\sigma)$ for some $t \in Y\}$ for some set of nodes Y, none of which is a successor of x. If $\beta_i(x) \cap \theta(x) \neq \emptyset$, then by lemma 1 and by the definition of $\chi(\bullet)$, $\beta_i(x) \cap \theta(x)=\{\zeta(t|\sigma)\}$ for some node $t \in Y$. Since $\zeta(t|\sigma) \in \theta(x)$, either t=x or t is a predecessor of x. Hence $\zeta(t|\sigma)=\zeta(x|\sigma)$.

PROOF OF PROPOSITION 1. Fix an extensive game G with perfect recall. Let the simple assessment (σ,μ) be a rational belief equilibrium. If the game has no chance moves then proposition 1 follows from proposition 1 in Bonanno (1992b). In fact, if $\beta=\chi(\sigma,\mu)$ then it easy to verify that β satisfies the property of Choice Consistency defined there. Suppose therefore that Nature moves at the root of the game. Fix an arbitrary subgame of G and let x^* be the root of the subgame. Two cases are possible: (1) $x^* \neq x_0$ and (2) $x^* = x_0$.

In case (1), by property 3, $K_j(x^*) = \theta(x^*)$ for every player j. By property 2, since the game has perfect recall, for every player j and for every node t that belongs to the subgame, $K_j(t) \subseteq \theta(x^*)$. Thus for every player j and for every node t that belongs to the subgame, $\beta_j(t) \subseteq \theta(x^*)$. Hence $\beta_j(t) \cap \theta(x^*) = \beta_j(t)$. It follows from lemma 1 that $\beta_j(t)$ is a singleton. Hence we can apply proposition 1 in Bonanno (1992b) to the subgame and conclude that the restriction of σ to the subgame is a Nash equilibrium of the subgame.

Consider now case (2), namely the case where $x^* = x_0$. In order to complete the proof of proposition 1 we only need to show that σ is a Nash equilibrium of the entire game. Fix an arbitrary player i. Let σ_i' be a pure strategy of player i such that

(A.1) $$\sum_{y\in\Sigma(x_0)} U_i(\zeta(y|\sigma'))\pi(\zeta(y|\sigma')) \neq \sum_{y\in\Sigma(x_0)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma))$$

where $\sigma'=(\sigma_i',\sigma_{-i})$ [recall that $\Sigma(x_0)$ denotes the set of immediate successors of the root]. We want to show that

(A.2) $$\sum_{y\in\Sigma(x_0)} U_i(\zeta(y|\sigma'))\pi(\zeta(y|\sigma')) < \sum_{y\in\Sigma(x_0)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma))$$

Let $V \subseteq \Sigma(x_0)$ be defined as follows:

$$V = \left\{y \in \Sigma(x_0) \middle| \zeta(y|\sigma) \neq \zeta(y|\sigma')\right\}$$

By (A.1), $V\neq\varnothing$. Furthermore, it follows from (A.1) that

(A.3) $$\sum_{y\in V} U_i(\zeta(y|\sigma'))\pi(\zeta(y|\sigma')) \neq \sum_{y\in V} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma))$$

Fix an arbitrary $y_1 \in V$. Let h_1 be the information set of player i at which the path from y_1 to $\zeta(y_1|\sigma)$ and the path from y_1 $\zeta(y_1|\sigma')$ diverge(hence σ_i and σ_i' select different choices at h_1). Let

(A.4) $$V(h_1) = \{y\in \Sigma(x_0) \mid \text{the path from } y \text{ to } \zeta(y|\sigma) \text{ crosses } h_1\}$$

and

(A.5) $$V'(h_1) = \{y\in \Sigma(x_0) \mid \text{the path from } y \text{ to } \zeta(y|\sigma') \text{ crosses } h_1\}$$

Then it must be[1]

(A.6) $$V'(h_1) = V(h_1)$$

Clearly, $V(h_1)\neq\varnothing$, since $y_1 \in V(h_1)$. Furthermore, since σ_i and σ_i' select

[1] *Proof.* Clearly $V(h_1)\cap V'(h_1)\neq\varnothing$, since $y_1\in V(h_1)\cap V'(h_1)$. We first show that $V(h_1)\subseteq V'(h_1)$. Suppose not, that is, suppose there is a $y_0\in V(h_1)$ such that $y_0\notin V'(h_1)$. Then $y_0\neq y_1$; furthermore, the path from y_0 to $\zeta(y_0|\sigma)$ must cross another information set of player i, call it g, before it reaches h_1. It must also be true that σ_1 and σ_i' select different choices at g. As a consequence, the path from y_1 to $\zeta(y_1|\sigma)$ does not cross g before it crossess h_1 (if at all). Let x_1 be the node in h_1 that lies on the path from y_1 to $\zeta(y_1|\sigma)$ and x' be the node in h_1 that lies on the path from y_0 to $\zeta(y_0|\sigma)$. Then x' comes after a choice at g while x_1 does not, contradicting the assumption of perfect recall. The proof that $V'(h_1) \subseteq V(h_1)$ is similar.

different choices at h_1, it must be $V(h_1) \subseteq V$. Fix an arbitrary $y_2 \in V/V(h_1)$ and let h_2 be the information set of player i at which the path from y_2 to $\zeta(y_2|\sigma)$ and the path from y_2 to $\zeta(y_2|\sigma')$ diverge. Let

$$V(h_2)=\{y\in\Sigma(x_0)| \text{ the path from } y \text{ to } \zeta(y|\sigma) \text{ crosses } h_2\}.$$

Repeat this procedure until the set V has been partitioned into m non-empty subsets $V(h_1)$, $V(h_2)$, ..., $V(h_m)$. It follows from (A.3) that

$$\sum_{y\in V(h_1)} U_i(\zeta(y|\sigma'))\pi(\zeta(y|\sigma')) + \ldots + \sum_{y\in V(h_m)} U_i(\zeta(y|\sigma'))\pi(\zeta(y|\sigma')) \neq$$

$$\neq \sum_{y\in V(h_1)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma)) + \ldots + \sum_{y\in V(h_m)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma)) \tag{A.7}$$

We will now go through a number of steps to show that

$$\sum_{y\in V(h_1)} U_i(\zeta(y|\sigma'))\pi(\zeta(y|\sigma')) \leq \sum_{y\in V(h_1)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma)) \tag{A.8}$$

The same argument can then be repeated for every set $V(h_j)$, so that, for every $j=1,\ldots, m$,

$$\sum_{y\in V(h_j)} U_i(\zeta(y|\sigma'))\pi(\zeta(y|\sigma')) \leq \sum_{y\in V(h_j)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma)) \tag{A.9}$$

Hence, adding all the inequalities in (A.9), and taking into account (A.3) and (A.1), we obtain (A.2).

STEP 1. Recall that, by definition of $\chi(\bullet)$, $\beta_i(x_0) = \{z| z=\zeta(y|\sigma)$ for some $y \in \Sigma(x_0)\}$ and, by the definition of $K_i(\bullet)$, $K_i(h_1) = \bigcup_{t\in h_1} \theta(t)$. Thus, by Contraction Consistency

$$\beta_i(h_1) = \beta_i(x_0) \cap K_i(h_1) = \{z|z = \zeta(y|\sigma) \text{ for some } y \in V(h_1)\} \tag{A.10}$$

Hence,

$$\sum_{z\in\beta_i(h_1)} U_i(z)\pi(z) = \sum_{y\in V(h_1)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma)) \tag{A.11}$$

Let $\hat{h}_1$ be the subset of h_1 consisting of the predecessors of $\beta_i(h_1)$. Then

$$\hat{h}_1 = \{t \in h_1|t \text{ is a successor of some } y \in V(h_1)\} \tag{A. 12}$$

Let c' be the choice selected by σ_i' at h_1. By Individual Rationality and by (A.10) and (A.11),

$$\sum_{y\in V(h_1)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma)) \geq \sum_{z\in \bigcup_{x\in \hat{h}_1} \beta_i(S(x|c'))} U_i(z)\pi(z) \tag{A.13}$$

By part (a) of Tree Consistency, for every $x\in \hat{h}_1$, $\beta_i(S(x|c')) \cap \theta(S(x|c')) \neq \varnothing$. Thus, by lemma 2,

$$\beta_i(S(x|c')) \cap \theta(S(x|c')) = \{\zeta(S(x|c')|\sigma)\} \tag{A. 14}$$

By part (b) of Tree Consistency,

$$\bigcup_{x\in \hat{h}_1} \beta_i(S(x|c')) = \bigcup_{x\in \hat{h}_1} \left[\beta_i(S(x|c')) \cap \theta(S(x|c'))\right] \tag{A. 15}$$

Thus, using (A.13)-(A.15) we obtain

$$\sum_{y\in V(h_1)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma)) \geq \sum_{x\in \hat{h}_1} U_i(\zeta(S(x|c')|\sigma))\pi(\zeta(S(x|c')|\sigma)) \tag{A. 16}$$

Fix an arbitrary $x\in \hat{h}_1$. Then by (A. 12) there is a $y\in V(h_1)$ that is a predecessor of x. It follows that $S(x|c')$ is a successor of y. Hence $S(x|c')$ lies on the path from y to $\zeta(y|\sigma')$. If, for all $x\in \hat{h}_1$, $\zeta(S(x|c')|\sigma) = \zeta(S(x|c')|\sigma')$, that is, if σ_i and σ_i' do not differ at any information set of player i (if any) that comes after nodes in $\hat{h}_1$, then (A.8) is proved [recall (A.4)-(A.6)]. Otherwise proceed to step 2.

STEP 2. Recall the following notation: $\Sigma(\hat{h}_1|c') = \{t | t = S(x|c')$ for some $x\in \hat{h}_1\}$. Thus (A.16) ca be re-written as

$$\sum_{y\in V(h_1)} U_i(\zeta(y|\sigma))\pi(\zeta(y|\sigma)) \geq \sum_{t\in\Sigma(\hat{h}_1|c')} U_i(\zeta(t|\sigma))\ \pi(\zeta(t|\sigma)) \tag{A. 17}$$

Define $W\subseteq\Sigma(\hat{h}_1/c')$ as follows:

$$W = \{t \in \Sigma(\hat{h}_1|c') | \zeta(t|\sigma) \neq \zeta(t|\sigma')\}.$$

If $W=\varnothing$ then the proof of (A.8) is complete (cf. remark at the end of step 1). The same is true if

$$\sum_{t\in\Sigma(\hat{h}_1|c')} U_i(\zeta(t|\sigma))\pi(\zeta(t|\sigma)) \geq \sum_{t\in\Sigma(\hat{h}_1|c')} U_i(\zeta(t|\sigma'))\pi(\zeta(t|\sigma')).$$

Assume therefore that

$$\text{(A. 18)} \qquad \sum_{t\in\Sigma(\bar{h}_1|c')} U_i(\zeta(t|\sigma))\pi(\zeta(t|\sigma)) < \sum_{t\in\Sigma(\bar{h}_1|c')} U_i(\zeta(t|\sigma'))\pi(\zeta(t|\sigma')).$$

Then it follows from (A. 18) that

$$\text{(A. 19)} \qquad \sum_{t\in W} U_i(\zeta(t|\sigma))\pi(\zeta(t|\sigma)) < \sum_{t\in W} U_i(\zeta(t|\sigma'))\pi(\zeta(t|\sigma'))$$

Fix an arbitrary $t_1 \in W$ and let g_1 be the information set of player i at which the path from t_1 to $\zeta(t_1|\sigma)$ and the path from t_1 to $\zeta(t_1|\sigma')$ diverge. Let

$$W(g_1) = \{t \in W | \text{the path from } t \text{ to } \zeta(t|\sigma) \text{ crosses } g_1\}$$

and

$$W'(g_1) = \{t \in W | \text{the path from to } \zeta(t|\sigma') \text{ crosses } g_1\}$$

Clearly, $W'(g_1) = W(g_1)$ and $\varnothing \neq W(g_1) \subseteq W$. Fix an arbitrary $t_2 \in W/W(g_1)$ and let g_2 be the information set of player i at which the path from t_2 to $\zeta(t_2|\sigma)$ and the path from t_2 to $\zeta(t_2|\sigma')$ diverge. Repeat this procedure until W has been partitioned into m' non-empty subsets W (g1), W(g2), ..., W(gm'). It follows from (A. 19) that

$$\text{(A. 20)} \qquad \begin{aligned} &\sum_{t\in W(g_1)} U_i(\zeta(t|\sigma))\pi(z(t|\sigma)) + \ldots + \sum_{t\in W(g_{m'})} U_i(\zeta(t|\sigma))\pi(\zeta(t|\sigma)) < \\ &< \sum_{t\in W(h_1)} U_i(\zeta(t|\sigma'))\pi(z(t|\sigma')) + \ldots + \sum_{t\in W(h_{m'})} U_i(\zeta(t|\sigma'))\pi(\zeta(t|\sigma')) \end{aligned}$$

We will now show that

$$\text{(A. 21)} \qquad \sum_{t\in W(g_1)} U_i(\zeta(t|\sigma))\pi(z(t|\sigma)) \geq \sum_{t\in W(g_1)} U_i(\zeta(t|\sigma'))\pi(\zeta(t|\sigma'))$$

The same argument can then be repeated for every set j=1,..., m' to show that

$$\text{(A. 22)} \qquad \sum_{t\in W(g_j)} U_i(\zeta(t|\sigma))\pi(\zeta(t|\sigma)) \geq \sum_{t\in W(g_j)} U_i(\zeta(t|\sigma'))\pi(\zeta(t|\sigma'))$$

Hence adding up all the inequalities in (A. 22) we contradict (A. 20).
The argument parallels that of step 1. By perfect recall and Contraction Consistency,

$$\beta_i(g_1) = \{z | z = \zeta(t|\sigma) \text{ for some } t \in W(g_1)\}.$$

Hence

$$\text{(A. 23)} \qquad \sum_{z\in\beta_i(g_1)} U_i(z)\pi(z) = \sum_{t\in W(g_1)} U_i(\zeta(t|\sigma))\pi(\zeta(t|\sigma))$$

Let $\hat{g}_1$ be the subset of g_1 consisting of the predecessors of $\beta_1(g_1)$. Then

$$\text{(A. 24)} \qquad \hat{g}_1 = \{x \in g_1 | x \text{ is a successor of some } t \in W(g_1)\}$$

Let d' be the choice selected by σ_i' at g_1. By Individual Rationality and by (A.23) and (A.24),

$$\text{(A. 25)} \qquad \sum_{t\in W(g_1)} U_i(\zeta t|\sigma))\pi(\zeta(t|\sigma)) \geq \sum_{\substack{z\in\cup\beta_i(S(x|d'))\\ x\in\hat{g}_1}} U_i(z)\pi(z)$$

By part (a) of Tree Consistency, for every $x\in \hat{g}_1$, $\beta_1(S(x|d')) \cap \theta(S(x|d')) \neq \emptyset$. Thus, by lemma 2, for every $x\in \hat{g}_1$,

$$\text{(A. 26)} \qquad \beta_i(S(x|d')) \cap \theta(S(x|d')) = \{\zeta(S(x|d')|\sigma)\}$$

By part (b) of Tree Consistency,

$$\text{(A. 27)} \qquad \bigcup_{x\in\hat{g}_1} \beta_i(S(x|d')) = \bigcup_{x\in\hat{g}_1} [\beta_i(S(x|d')) \cap \theta(S(x|d'))]$$

Thus, using (A. 25)-(A. 27) we obtain

$$\text{(A. 28)} \qquad \sum_{t\in W(g_1)} U_i(\zeta(t|\sigma))\pi(\zeta(t|\sigma)) \geq \sum_{x\in\hat{g}_1} U_i(\zeta(S(x|d')|\sigma))\pi(\zeta(S(x|d')|\sigma))$$

Fix an arbitrary $x\in \hat{g}_1$. Then by (A. 24) there is a $t \in W(g_1)$ that is a predecessor of x. It follows that $S(x|d')$ is a succesor of t. Hence $S(x|d')$ lies on the path from t to $\zeta(t|\sigma')$. If, for all $x\in \hat{g}_1$, $\zeta(S(x|d')|\sigma) = \zeta(S(x|d')|\sigma')$, that is, if σ_i and σ_i' do not differ at any information set (if any) of player i that comes after nodes in $\hat{g}_1$, then (A.21) is proved. Otherwise we repeat the same argument until we reach an information set f_1 of player i where the condition that, for all $y\in \hat{f}_1$ [where $\hat{f}_1$ denotes the the subset of f_1 consisting of the predecessors of $\beta_1(f_1)$], $\zeta(S(y|e')|\sigma) = \zeta(S(y|e')|\sigma')$ - where e' is the choice selected by σ_i' at information set f_1 - is satisfied (either because σ_i and σ_i' do not differ at any information set of player i that comes after nodes in $\hat{f}_1$, or because there is no information set of player i that comes after nodes in $\hat{f}_1$).

We now turn to the proof of proposition 2. As before, we shall begin with a few lemmas.

LEMMA 3. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson. Then the following is true: if y belongs to information set u, d is the choice that leads from p_y to y [recall that p_y denotes the immediate prodecessor of y], and every node in u comes after choice d, then $\mu(y) \geq \mu(p_y)$.

Proof. Let h be the information set to which p_y belongs. Let $<\sigma^m>$ be the sequence of completely mixed strategies that converges to σ and from which the function μ is obtained (by applying Bayes' rule to σ^m and taking the limit). Then, for every m, Prob $\{u|\sigma^m\} \leq$ Prob $\{h|\sigma^m\}$ Prob $\{d|\sigma^m\}$. Also, Prob $\{y|\sigma^m\}$ = Prob $\{p_y|\sigma^m\}$ Prob $\{d|\sigma^m\}$. Thus

$$\text{Prob}\{y|u, \text{ given } \sigma^m\} = \frac{\text{Prob }\{y|\sigma^m\}}{\text{Prob }\{u|\sigma^m\}} \geq \frac{\text{Prob }\{p_y|\sigma^m\}\text{ Prob }\{d|\sigma^m\}}{\text{Prob }\{h|\sigma^m\}\text{ Prob }\{d|\sigma^m\}} = \frac{\text{Prob }\{p_y|\sigma^m\}}{\text{Prob }\{h|\sigma^m\}}$$

Since $\mu(y)$ is the limit of the LHS as $m\to\infty$ and $\mu(p_y)$ is the limit of the RHS as $m\to\infty$, the proof is complete.

COROLLARY 1. Fix a game with perfect recall. Let the simple assessment (σ,μ) be consistent in the sense of Kreps and Wilson and let $\beta=\chi(\sigma,\mu)$ be the corresponding profile of beliefs. Then the following is true for every player i: if h is an information set of player i, $x\in h$ is such that $\mu(x) > 0$ and y is an immediate successor of x, then $\zeta(y|\sigma) \in \beta_i(y)$.

Proof. If $K_i(y) = \theta(y)$ then $\beta_i(y) = \{\zeta(y|\sigma)\}$ by (2) of of the definition of $\chi(\bullet)$. If y belongs to information set u of player i and d is the choice that leads from x to y, then, by perfect recall, every node in u comes after choice d. Hence by lemma 3, $\mu(y) \geq \mu(x) > 0$. If follows from (3) of the definition of $\chi(\bullet)$ that $\zeta(y|\sigma) \in \beta_i(y)$. If $K_i(y) \neq \theta(y)$ and y is not a decision node of player i, then by (5) of the definition of information, $K_i(y)=\gamma(d)$ [where d is the choice at h that leads from x to y, that is, $y= S(x|d)$]. If $\beta_i(h) \cap \gamma(d) = \varnothing$, then $\zeta(y|\sigma) \in \beta_i(y)$ by (5) of the definition of $\chi(\bullet)$ [since, by assumption, $\mu(x) > 0$]. Finally, if $\beta_i(h) \cap \gamma(d) \neq \varnothing$, note that, by (2) of the definition of $\chi(\bullet)$, $\beta_i(h)=\{z| z=\zeta(t|\sigma)$ for some $t \in$ supp $(\mu|h)\}$. Thus σ_i selects choice d at h. Since $y = S(x|d)$, it follows that $\zeta(y|\sigma) = \zeta(x|\sigma)$. By (4) of the definition of $\chi(\bullet)$, $\beta_i(y) = \beta_i(h) \cap K_i(y) = \{z| z=\zeta(t|\sigma)$ for some $t \in$ supp $(\mu|h)\} \cap \gamma(d)$. Hence $\zeta(y|\sigma) \in \beta_i(y)$.

LEMMA 4. Fix an extensive game with perfect recall. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson. Let h and u be information sets of the same player and suppose that there exists a node $\tilde{y}\in u$ that is an immediate successor of a node $\tilde{x}\in h$ such that $\mu(\tilde{x}) > 0$. Then for every node $y\in u$ that is a successor (not necessarily an immediate successor) of a node $x\in h$ such that $\mu(x)=0$, $\mu(y)=0$.

Proof. Let $\tilde{y} \in u$ be an immediate successor of a node $\tilde{x} \in h$ such that $\mu(\tilde{x})>0$ and let $y \in u$ be a successor of a node $x \in h$ such that $\mu(x)=0$. By perfect recall, both $\tilde{y}$ and y come after the same choice at h, call it d. Let $<\sigma^m>$ be the sequence of completely mixed strategies that converges to σ and from which the function μ is obtained (applying Bayes' rule to σ^m and taking the limit). Then, for every m,

$$\text{Prob}\{\tilde{y}|\sigma^m\} = \text{Prob}\{\tilde{x}|\sigma^m\}\text{Prob}\{d|\sigma^m\}, \text{ and}$$

$$\text{Prob}\{y|\sigma^m\} \leq \text{Prob}\{x|\sigma^m\}\text{Prob}\{d|\sigma^m\}$$

[it is an equality if y is an immediate successor of x]. Thus

$$\text{(A.29)} \qquad \frac{\text{Prob}\ \{y|\sigma^m\}}{\text{Prob}\ \{\tilde{y}|\sigma^m\}} \leq \frac{\text{Prob}\ \{x|\sigma^m\}}{\text{Prob}\ \{\tilde{x}|\sigma^m\}}$$

Dividing numerator and denominator of the LHS by Prob $\{u|\sigma^m\}$ and numerator and denominator of the RHS by Prob $\{h|\sigma^m\}$ we obtain

$$\text{(A.30)}. \qquad \frac{\dfrac{\text{Prob}\ \{y|\sigma^m\}}{\text{Prob}\ \{u|\sigma^m\}}}{\dfrac{\text{Prob}\ \{\tilde{y}|\sigma^m\}}{\text{Prob}\ \{u|\sigma^m\}}} \leq \frac{\dfrac{\text{Prob}\ \{x|\sigma^m\}}{\text{Prob}\ \{h|\sigma^m\}}}{\dfrac{\text{Prob}\ \{\tilde{x}|\sigma^m\}}{\text{Prob}\ \{h|\sigma^m\}}}$$

Taking the limit as $m \to \infty$ and recalling that, by hypothesis, $\mu(\tilde{x}) > 0$, and, by lemma 3, $\mu(\tilde{y}) \geq \mu(\tilde{x})$, we obtain

$$\text{(A.31)} \qquad \frac{\mu(y)}{\mu(\tilde{y})} \leq \frac{\mu(x)}{\mu(\tilde{x})}$$

Since, by hypothesis, $\mu(x)=0$, it follows that $\mu(y)=0$.

COROLLARY 2. Fix an extensive game with perfect recall. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson. Let h and u be information sets of the same player and suppose that there exists a node $\tilde{y} \in u$ that is an immediate successor of a node $\tilde{x} \in h$ such that $\mu(\tilde{x}) > 0$. Let d be the choice at h that leads from $\tilde{x}$ to $\tilde{y}$ and let $\hat{h}$=supp.$(\mu|h)$. Then for every $y \in u$ such that $\mu(y)>0$, $\zeta(y|\sigma) = \zeta(t|\sigma)$ for some $t \in \Sigma(\hat{h}|d)$. [Recall that $\Sigma(\hat{h}|d)=\{t|\ t=S(x|d)$ for some $x \in \hat{h}\}$, where $S(x/d)$ denotes the immediate successor of x following choice d].

Proof. Fix a $y \neq \tilde{y}$ such that $\mu(y) > 0$. By perfect recall, y comes after choice d. Hence by lemma 4, either $y \in \Sigma(\hat{h}|d)$ or y is a successor of a $t \in \Sigma(\hat{h}|d)$. Suppose the latter is the case and it is not true that $\zeta(y|\sigma) = \zeta(t|\sigma)$. Let $x \in h$ be the immediate predecessor of t and let $c_1, c_2, \ldots, c_r$ be the sequence of choices that leads from x to y (thus c_1=d and $r \geq 2$). Let $<\sigma^m>$ be the sequence of completely mixed strategies that converges to σ and from which the function μ is obtained (applying Bayes' rule to σ^m and taking the limit). Then, for at least one j=2, ..., r,

$$\lim_{m \to \infty} \text{Prob}\{c_j|\sigma^m\} = 0 \tag{A.32}$$

Now,

$$\text{Prob}\{y|\sigma^m\} = \text{Prob}\{x|\sigma^m\}\text{Prob}\{d|\sigma^m\}A(\sigma^m) \tag{A.33}$$

where $A(\sigma^m)=\text{Prob}\{c_2|\sigma^m\}\ \text{Prob}\{c_3|\sigma^m\} \ldots \text{Prob}\{c_r|\sigma^m\}$, and therefore, by (A.32)

$$\lim_{m \to \infty} A(\sigma^m) = 0 \tag{A.34}$$

Furthermore,

$$\text{Prob}\{\tilde{y}|\sigma^m\} = \text{Prob}\{\tilde{x}|\sigma^m\}\text{Prob}\{d|\sigma^m\} \tag{A.35}$$

Dividing (A.33) by (A.35) we obtain

$$\frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{\tilde{y}|\sigma^m\}} = \frac{\text{Prob}\{x|\sigma^m\}}{\text{Prob}\{\tilde{x}|\sigma^m\}} A(\sigma^m) \tag{A.36}$$

Dividing numerator and denominator of the LHS by Prob $\{u|\sigma^m\}$ and numerator and denominator of the RHS by Prob $\{h|\sigma^m\}$ and taking the limit as $m \to \infty$, we obtain [note that, by hypothesis, $\mu(\tilde{x})>0$, by lemma 4, $\mu(x)>0$, and, by lemma 3, $\mu(\tilde{y}) \geq \mu(\tilde{x})$]

$$\frac{\mu(y)}{\mu(\tilde{y})} = \frac{\mu(x)}{\mu(\tilde{x})} \lim_{m \to \infty} A(\sigma^m) \tag{A.37}$$

By (A.34) the RHS is equal to zero and therefore $\mu(y)=0$, yelding a contradiction.

LEMMA 5. Fix a game with perfect recall. Let the simple assessment (σ,μ) be consistent in the sense of Kreps and Wilson and let $\beta=\chi(\sigma,\mu)$ be the corresponding belief system. Then β satisfies the property of Tree Consistency.

Proof. Part (a) of Tree Consistency follows from corollary 1, since $\zeta(y|\sigma) \in \theta(y)$. It only remains to prove that if h is an information set of player i, $\hat{h}$ is the subset of h consisting of the predecessors of $\beta_i(h)$ [so that, by (3) of the definition of $\chi(\bullet)$, $\hat{h}=\text{supp}(\mu|h)$] and c is a choice at h then

$$\bigcup_{y\in\Sigma(\hat{h}|c)} \beta_i(y) = \bigcup_{y\in\Sigma(\hat{h}|c)} [\beta_i(y)\cap\theta(y)] \tag{A.38}$$

By corollary 1, for all $y\in\Sigma(\hat{h}|c)$, $\zeta(y|\sigma) \in \beta_i(y)$. By lemma 1, $\beta_i(y) \cap \theta(y)$ is a singleton. Thus, for all $y\in\Sigma(\hat{h}|c)$,

$$\beta_i(y)\cap\theta(y) = \{\zeta(y|\sigma)\} \tag{A.39}$$

Hence (A.38) is equivalent to

$$\bigcup_{y\in\Sigma(\hat{h}|c)} \beta_i(y) = \{z|z = \zeta(y|\sigma) \text{ for some } y \in \Sigma(\hat{h}|c)\} \tag{A.40}$$

Let $Y_1=\{y\in\Sigma(\hat{h}|c)| K_i(y)=\theta(y)\}$. Then, for every $y\in Y_1$, $\beta_i(y) \cap \theta(y) = \beta_i(y)$. Hence, by (A.39),

$$\begin{aligned}\bigcup_{y\in Y_1} \beta_i(y) &= \{z|z = \zeta(y|\sigma) \text{ for some } y \in Y_1\} \\ &\subseteq \{z|z = \zeta(y|\sigma) \text{ for some } y \in \Sigma(\hat{h}|c)\}\end{aligned} \tag{A.41}$$

Let $Y_2 = \{y\in\Sigma(\hat{h}|c)| K_i(y) \neq \theta(y)\}$ and y is not a decision node of player i} and $Y_3 = \{y\in\Sigma(\hat{h}|c)| K_i(y) \neq \theta(y)$ and y is a decision node of player i}. Obviously, $\Sigma(\hat{h}|c) = Y_1\cup Y_2\cup Y_3$. By (5) of the definition of information, for every $y\in Y_2$, $K_1(y)=\gamma(c)$. If $\beta_i(h) \cap \gamma(c) = \varnothing$, then, by (5) of the definition of $\chi(\bullet)$, for every $y\in Y_2$,

$$\beta_i(y) = \{z|z = \zeta(y|\sigma) \text{ for some } y \in \Sigma(\hat{h}|c)\} \tag{A.42a}$$

If $\beta_i(h) \cap \mu(c) = \varnothing$, then σ_i selects choice c at h and therefore for every node $x\in\hat{h}$, $\zeta(x|\sigma) = \zeta(S(x|c)|\sigma)$.
Thus $\beta_i(h) = \{z| z=\zeta(x|\sigma)$ for some $x\in\hat{h}\} = \{z| z=\zeta(x|\sigma)$ for some $y\in\Sigma(\hat{h}|c)\} \subseteq \gamma(c)$.
By (4) of the definition of $\chi(\bullet)$, for every $y\in Y_2$,

$$\beta_i(y) = \beta_i(h)\cap K_i(y) = \beta_i(h)\cap\gamma(c) = \{z|z = \zeta(y|\sigma) \text{ for some } y \in \Sigma(\hat{h}|c)\} \tag{A.42b}$$

Thus, by (A.42a) and (A.42b),

$$\bigcup_{y\in Y_2} \beta_i(y) \subseteq \{z|z = \zeta(y|\sigma \text{ for some } y \in \Sigma(\hat{h}|c)\} \tag{A.43}$$

It only remains to prove that

$$\bigcup_{y \in Y_3} \beta_i(y) \subseteq \{z | z = \zeta(y|\sigma) \text{ for some } y \in \Sigma(\hat{h}|c)\} \tag{A.44}$$

Fix an arbitrary $\tilde{y} \in Y_3$. Let u be the information set of player i to which $\tilde{y}$ belongs. By corollary 2, for every $y \in u$, if $\mu(y)>0$ then $\zeta(y|\sigma) = \zeta(t|\sigma)$ for some $t \in \Sigma(\hat{h}|c)$. It follows from(3) of the definition of $\chi(\bullet)$ that

$$\beta_i(\tilde{y}) \subseteq \{z | z = \zeta(y|\sigma) \text{ for some } y \in \Sigma(\hat{h}|c)\} \tag{A.45}$$

Since $\tilde{y}$ was chosen arbitrarily, (A.44) is proved. By (A.42a,b)-(A.44), the proof is complete.

LEMMA 6. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson. If y is a successor of x on the path from x to $\zeta(x|\sigma)$, y belongs to information set u and all the nodes in u are successors of x, then $\mu(y)=1$.

Proof. Let $c_1, c_2, \ldots, c_r$ be the choices that lead from x to y. Let $<\sigma^m>$ be the sequence of completely mixed strategies that converges to σ and from which the function μ is obtained (by applying Bayes' rule to σ^m and taking the limit as $m \to \infty$). Since y is reached from x by following σ,

$$\lim_{m \to \infty} \text{Prob}\{c_k|\sigma^m\} = 1, \quad \text{for every } k = 1, 2, \ldots, r. \tag{A.46}$$

Now, Prob $\{y|\sigma^m\}$ = Prob $\{x|\sigma^m\}$ Prob $\{c_1|\sigma^m\}$ Prob $\{c_2|\sigma^m\}$...Prob $\{c_r|\sigma^m\}$. Furthermore,

$$\text{Prob}\{y|u, \text{ given } \sigma^m\} = \frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{u|\sigma^m\}} = \frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{x|\sigma^m\}} \frac{\text{Prob}\{x|\sigma^m\}}{\text{Prob}\{u|\sigma^m\}}$$

Since all the nodes in u are successors of x, $\dfrac{\text{Prob}\{x|\sigma^m\}}{\text{Prob}\{u|\sigma^m\}} = 1$. Thus

$$\frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{u|\sigma^m\}} = \frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{x|\sigma^m\}} = \frac{\text{Prob}\{x|\sigma^m\}\text{Prob}\{c_1|\sigma^m\}\text{Prob}\{c_2|\sigma^m\}\cdots\text{Prob}\{c_r|\sigma^m\}}{\text{Prob}\{x|\sigma^m\}}$$

$$= \text{Prob}\{c_1|\sigma^m\}\text{Prob}\{c_2|\sigma^m\}\cdots\text{Prob}\{c_r|\sigma^m\}.$$

Since, by (A.46) the limit of the RHS as $m \to \infty$ is equal to 1 and $\mu(y)$ is the limit of the LHS as $m \to \infty$, it follows that $\mu(y)=1$.

LEMMA 7. Fix a game with perfect recall. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson. If x and y are decision nodes of player i and y is successor of x on the path from x to $\zeta(x|\sigma)$, then $\mu(y) \geq \mu(x)$.

Proof. Let h be the information set (of player i) to which x belongs, and u the information set (of player i) to which y belongs. Let $c_1, c_2, \ldots, c_r$ be the choices that lead from x to y. Let $<\sigma^m>$ be the sequence of completely mixed strategies that converges that converges to σ and from which the function μ is obtained (by applying Bayes' rule to σ^m and taking the limit as $m \to \infty$). Since y is reached from x by following σ,

$$\text{(A.47)} \qquad \lim_{m \to \infty} \text{Prob}\{c_k|\sigma^m\} = 1, \quad \text{for every } k = 1, 2, \ldots, r,$$

Now,

$$\text{Prob}\{y|\sigma^m\} = \text{Prob}\{x|\sigma^m\}\text{Prob}\{c_1|\sigma^m\}\text{Prob}\{c_2|\sigma^m\}\ldots\text{Prob}\{c_r|\sigma^m\}.$$

By perfect recall, all the nodes in u come after choice c_1 at h. Thus

$$\text{Prob}\{u|\sigma^m\} \leq \text{Prob}\{h|\sigma^m\} \ \text{Prob}\{c_1|\sigma^m\}.$$

Therefore,

$$\text{Prob}\{y|u, \text{ given } \sigma^m\} = \frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{u|\sigma^m\}} \geq \frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{h|\sigma^m\}\text{Prob}\{c_1|\sigma^m\}}$$

$$= \frac{\text{Prob}\{x|\sigma^m\}\text{Prob}\{c_1|\sigma^m\}\text{Prob}\{c_2|\sigma^m\}\ldots\text{Prob}\{c_r|\sigma^m\}}{\text{Prob}\{h|\sigma^m\}\text{Prob}\{c_1|\sigma^m\}}$$

$$= \frac{\text{Prob}\{x|\sigma^m\}}{\text{Prob}\{h|\sigma^m\}}[\text{Prob}\{c_2|\sigma^m\}\ldots\text{Prob}\{c_r|\sigma^m\}]$$

The limit as $m \to \infty$ of the first term on the RHS is equal to $\mu(x)$, while, by (A.47), the limit of the second term on the RHS is equal to 1. Since $\mu(y)$ is the limit of the LHS $m \to \infty$, the proof is complete.

LEMMA 8. Fix an extensive game with perfect recall. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson. Let h and u be information sets of the same player and suppose there exist nodes $\tilde{x} \in h$ and $\tilde{y} \in u$ such that: (i) $\mu(\tilde{x})>0$, (ii) $\tilde{y}$ is a successor of $\tilde{x}$ and (iii) $\zeta(\tilde{x}|\sigma)=\zeta(\tilde{y}|\sigma)$. Then for every $y \in u$ with $\mu(y)>0$, $\zeta(y|\sigma)=\zeta(x|\sigma)$ for some $x \in \text{supp}(\mu|h)$.

Proof. Fix a $y \neq \tilde{y}$ such that $(\mu(y) > 0$. By perfect recall, there is a node $x \in h$ that is a predecessor of y. Furthermore, there is a choice d at h that precedes all the nodes in u. We first show that $\mu(x) > 0$. Let $<\sigma^m>$ be the sequence of completely mixed strategies that converges to σ and from which the function μ is obtained (applying Bayes' rule to σ^m and taking the limit). Let $c_1, c_2, \ldots, c_r$ be the choices that lead from $\tilde{x}$ to $\tilde{y}$ (thus c_1=d and $r \geq 1$). Then, since $\zeta(\tilde{y}|\sigma)=\zeta(\tilde{x}|\sigma)$,

$$\lim_{m\to\infty} \text{Prob}\{c_k|\sigma^m\} = 1, \quad \text{for every } k = 1,\ldots,r, \tag{A.48}$$

For every m,

$$\text{Prob}\{\tilde{y}|\sigma^m\} = \text{Prob}\{\tilde{x}|\sigma^m\}\text{Prob}\{d|\sigma^m\}\text{Prob}\{c_2|\sigma^m\}\ldots\text{Prob}\{c_r|\sigma^m\} \tag{A.49}$$

and

$$\text{Prob}\{y|\sigma^m\} \leq \text{Prob}\{x|\sigma^m\}\text{Prob}\{d|\sigma^m\}$$

[it is equality if y is an immediate successor of x and a strict inequality otherwise]. Thus,

$$\frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{\tilde{y}|\sigma^m\}} \leq \frac{\text{Prob}\{x|\sigma^m\}}{\text{Prob}\{\tilde{x}|\sigma^m\}}\text{Prob}\{c_2|\sigma^m\}\ldots\ \text{Prob}\{c_r|\sigma^m\}, \tag{A.50}$$

Dividing numerator and denominator of the LHS by Prob $\{u|\sigma^m\}$ and numerator and denominator of the RHS by Prob $\{h|\sigma^m\}$ and taking the limit as $m\to\infty$, we obtain [using (A.48) and noting that, by hypothesis, $\mu(\tilde{x})>0$ and, by lemma 7, $\mu(\tilde{y}) \geq (\tilde{x})$],

$$\frac{\mu(y)}{\mu(\tilde{y})} \leq \frac{\mu(x)}{\mu(\tilde{x})} \tag{A.51}$$

Hence $\mu(y)>0$ implies $\mu(x)>0$.
Now suppose that $\zeta(x|\sigma)=\zeta(y|\sigma)$. Let $c_1', c_2', \ldots, c_s'$ be the choices that lead from x to y (thus c_1'=d; furthermore, $s>1$, since by hypothesis (iii), σ_i selects choice d at h). Then for at least one $j=2,\ldots,s$,

$$\lim_{m\to\infty} \text{Prob}\{c_j'|\sigma^m\} = 0 \tag{A.52}$$

Now,

(A.53) $$\text{Prob}\{y|\sigma^m\} = \text{Prob}\{x|\sigma^m\}\text{Prob}\{d|\sigma^m\}A(\sigma^m)$$

where $A(\sigma^m)=\text{Prob}\{c_2'|\sigma^m\}...\text{Prob}\{c_s'|\sigma^m\}$, and therefore, by (A.52),

(A.54) $$\lim_{m\to\infty} A(\sigma^m) = 0$$

Dividing (A.53) by (A.49) we obtain

(A.55) $$\frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{\tilde{y}|\sigma^m\}} = \frac{\text{Prob}\{x|\sigma^m\}A(\sigma^m)}{\text{Prob}\{\tilde{x}|\sigma^m\}B(\sigma^m)}$$

where $B(\sigma^m)=\text{Prob}\{c_2|\sigma^m\}...\text{Prob}\{c_r|\sigma^m\}$ and, by (A.48), $\lim_{m\to\infty} B(\sigma^m)=1$. Dividing numerator and denominator of the LHS by $\text{Prob}\{u|\sigma^m\}$ and numerator and denominator of the RHS by $\text{Prob}\{h|\sigma^m\}$ and taking the limit as $m\to\infty$, we obtain, using (A. 48) [recall that, $\mu(x) > 0$, $\mu(\tilde{x}) > 0$ and $\mu(\tilde{y}) \geq \mu(\tilde{x})$],

$$\frac{\mu(y)}{\mu(\tilde{y})} = \frac{\mu(x)}{\mu(\tilde{x})} \lim_{m\to\infty} A(\sigma^m).$$

By (A.52) the RHS is equal to zero and therefore $\mu(y)=0$, yielding a contradiction.

LEMMA 9. Fix a game with chance moves. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson. Let $\tilde{y}$ belong to information set u of player i and suppose there exists a $\tilde{t}\in\Sigma(x_0)$ such that $\zeta(\tilde{t}/\sigma) = \zeta(\tilde{y}/\sigma)$. Then for every $y\in u$ such that $\mu(y)>0$, $\zeta(y/\sigma)=\zeta(t/\sigma)$ for some $t\in\Sigma(x_0)$.

Proof. Let $p>0$ be the probability of the choice of Nature that leads from x_0 to t and let $d_1,...,d_s$ be the choices that lead from $\tilde{t}$ to $\tilde{y}$ [note that $s=0$, that is, $\tilde{t} = \tilde{y}$ is a possibility]. Let $<\sigma^m>$ be the sequence of completely mixed strategies that converges to σ and from which the function μ is obtained (applying Bayes' rule to σ^m and taking the limit). Then, since $\zeta(\tilde{t}/\sigma)=\zeta(\tilde{y}/\sigma)$, for every $k=1,...,s$

(A.56) $$\lim_{m\to\infty} \text{Prob}\{d_k|\sigma^m) = 1$$

Furthermore, for every m,

(A.57) $$\text{Prob}\{\tilde{y}|\sigma^m\} = p\ \text{Prob}\{d_1|\sigma^m\}...\text{Prob}\{d_s|\sigma^m\}$$

Fix an arbitrary $y \in u$ such that $y \neq \tilde{y}$ and $\mu(y)>0$. If $y \in \Sigma(x_0)$ there is nothing to prove. Suppose therefore that y is a successor of $t \in \Sigma(x_0)$ and $\zeta(y|\sigma) \neq \zeta(t|\sigma)$. Let $q>0$ be the probability of the choice of Nature that leads from x_0 to t. Let $c_1,...c_r$ be the choices that lead from t to y (thus $r \geq 1$). Then for at least one $j=1,...,r$,

(A.58) $$\lim_{m \to \infty} \text{Prob}\{c_j|\sigma^m\} = 0$$

Now,

(A.59) $$\text{Prob}\{y|\sigma^m\} = q \ \ \text{Prob}\{c_1|\sigma^m\} \ldots \text{Prob}\{c_r|\sigma^m\}$$

Furthermore,

(A.60) $$\text{Prob}\{u|\sigma^m\} \geq \text{Prob}\{\tilde{y}|\sigma^m\}$$

Thus, using (A.57), (A.59) and (A.60)

(A.61) $$\frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{u|\sigma^m\}} \leq \frac{\text{Prob}\{y|\sigma^m\}}{\text{Prob}\{\tilde{y}|\sigma^m\}} = \frac{q}{p} \frac{\text{Prob}\{c_1|\sigma^m\} \ldots \text{Prob}\{c_r|\sigma^m\}}{\text{Prob}\{d_1|\sigma^m\} \ldots \text{Prob}\{d_s|\sigma^m\}}$$

Since the limit of the LHS as $m \to \infty$ is $\mu(y)$, and by (A.56) and (A.58) the limit of the RHS as $m \to \infty$ is zero, it follows that $\mu(y)=0$, yelding a contradiction.

LEMMA 10. Fix a game with perfect recall. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson. Let x_1 belong to information set h of player i, x_2 be the immediate successor of x_1 following choice c at h and x_3 be the immediate successor of x_2. Let $\hat{h} = \text{supp}(\mu/h)$. Suppose x_3 belongs to information set u of player i. Suppose also that $\mu(x_1)>0$ (i.e. $x_1 \in \hat{h}$) and $\zeta(x_2|\sigma) = \zeta(x_3|\sigma)$. Then for every $t \in u$ such that $\mu(t)>0$, $\zeta(t|\sigma) = \zeta(y|\sigma)$ for some $y \in \Sigma(\hat{h}|c)$.

Proof. We first show that $\mu(x_3) \geq \mu(x_1)$. Let d be the choice leading from x_2 to x_3. Let $<\sigma^m>$ be the sequence of completely mixed strategies that converges to σ and from which the function μ is obtained (applying Bayes' rule to σ^m and taking the limit). Since, by hypothesis, $\zeta(x_2|\sigma) = \zeta(x_3|\sigma)$, it must be

(A.62) $$\lim_{m \to \infty} \text{Prob}\{d|\sigma^m\} = 1$$

For every m,

(A.63) $$\text{Prob}\{x_3|\sigma^m\} = \text{Prob}\{x_1|\sigma^m\}\text{Prob}\{c|\sigma^m\}\text{Prob}\{d|\sigma^m\}$$

By perfect recall, every node in u comes after choice c at h, hence

(A.64) $$\text{Prob}\{u|\sigma^m\} \leq \text{Prob}\{h|\sigma^m\}\text{Prob}\{c|\sigma^m\}$$

Thus

(A.65) $$\frac{\text{Prob}\{x_3|\sigma^m\}}{\text{Prob}\{u|\sigma^m\}} \geq \frac{\text{Prob}\{x_1|\sigma^m\}}{\text{Prob}\{h|\sigma^m\}}\text{Prob}\{d|\sigma^m\}$$

Taking the limit as $m\to\infty$ we obtain, using (A.62), $\mu(x_3) \geq \mu(x_1)$.

Now, fix an arbitrary $t\in u$ such that $t\neq x_3$ and $\mu(t)>0$. Let $x\in h$ be the predecessor of t and let $y=S(x|c)$ [note that t=y is a possibility]. Next we show that $x\in \hat{h}$, i.e. $\mu(x) > 0$. For every m,

(A.66) $$\text{Prob}\{t|\sigma^m\} \leq \text{Prob}\{x|\sigma^m\}\text{Prob}\{c|\sigma^m\}$$

[it is an equality if t=y]. Dividing (A.66) by (A.63) we obtain

(A.67) $$\frac{\text{Prob}\{t|\sigma^m\}}{\text{Prob}\{x_3|\sigma^m\}} \leq \frac{\text{Prob}\{x|\sigma^m\}}{\text{Prob}\{x_1|\sigma^m\}}\text{Prob}\{d|\sigma^m\}$$

Dividing numerator and denominator of the LHS by Prob $\{\mu|\sigma^m\}$ and numerator and denominator of the RHS by Prob $\{h|\sigma^m\}$ and taking the limit as $m\to\infty$ we obtain, using (A.62) [recall that, by hypothesis, t, $x_3\in u$, $x,x_1\in h$, $\mu(x_1)>0$ and, by the above argument $\mu(x_3) \geq \mu(x_1)$]:

(A.68) $$\frac{\mu(t)}{\mu(x_3)} \leq \frac{\mu(x)}{\mu(x_1)}$$

Thus $\mu(t) > 0$ implies $\mu(x) > 0$.

It follows that either $t\in\Sigma(\hat{h}|c)$ or t is a successor of a node $y\in\Sigma(\hat{h}|c)$. In the first case there is nothing left to prove. Consider therefore the latter case. Suppose it is not true that $\zeta(y|\sigma) = \zeta(t|\sigma)$. Let $c_1,....,c_r$ be the choices that lead from y to t (hence $r\geq 1$). Then at least one $j=1,...,r$,

(A.69) $$\lim_{m\to\infty} \text{Prob}\{c_j|\sigma^m) = 0$$

Now,

(A.70) $$\text{Prob}\{t|\sigma^m\} = \text{Prob}\{x|\sigma^m\}\text{Prob}\{c|\sigma^m\}A(\sigma^m)$$

where $A(\sigma^m)=\text{Prob}\{c_1|\sigma^m\}\ldots\text{Prob}\{c_r|\sigma^m\}$ and, therefore, by (A.69),

$$\lim_{m\to\infty} A(\sigma^m) = 0 \tag{A.71}$$

Dividing (A.70) by (A.63) we obtain

$$\frac{\text{Prob}\{t|\sigma^m\}}{\text{Prob}\{x_3|\sigma^m\}} \geq \frac{\text{Prob}\{x|\sigma^m\}}{\text{Prob}\{x_1|\sigma^m\}} \frac{A(\sigma^m)}{\text{Prob}\{d|\sigma^m\}} \tag{A.72}$$

Dividing numerator and denominator of the LHS by Prob $\{u|\sigma^m\}$ and numerator and denominator of the RHS by Prob $\{h|\sigma^m\}$ and taking the limit as $m\to\infty$, we obtain [recall that $\mu(x_1)>0$ and $\mu(x_3)>0$]

$$\frac{\mu(t)}{\mu(x_3)} = \frac{\mu(x)}{\mu(x_1)} \frac{\lim_{m\to\infty} A(\sigma^m)}{\lim_{m\to\infty} \text{Prob}\ \{d|\sigma^m\}} \tag{A.73}$$

By (A.62) and (A.70) the RHS is equal to zero. Hence $\mu(t)=0$, contradicting our supposition.

LEMMA 11. Let (σ,μ) be an assessment that is consistent in the sense of Kreps and Wilson and let $\beta=\chi(\sigma,\mu)$ be the corresponding profile of beliefs. Then β satisfies the property of Contraction Consistency.

Proof. We want to prove that if y is a successor of x and $\beta_i(x)\cap K_i(y)\neq\emptyset$ then $\beta_i(y)=\beta_i(x)\cap K_i(y)$. It will be sufficient to prove this for the case where y is an *immediate* successor of x. If y is not a decision node of player i, then it follows from (4) of the definition of $\chi(\bullet)$. Assume, therefore, that y belongs to information set u of player i. Then, by (3) of the definition of $\chi(\bullet)$, $\beta_i(y)=\{z|\ z=\zeta(t|\ \sigma)$ for some $\in$ supp $(\mu|\ u)\}$.

CASE 1: (y belongs to information set u of player i and) x belongs to an information set of player i, call it h. Then, by (3) of the definition of $\chi(\bullet)$, $\beta_i(x)=\{z|\ z=\zeta(w|\ \sigma)$ for some $w\in$ supp $(\mu|\ h)\}$. Since $\beta_i(x)\cap K_i(y)\neq\emptyset$, there must be a node $v\in u$ such that $\zeta(v|\ \sigma) = \zeta(w|\ \sigma)$ for some node $w\in$ supp $(\mu|\ h)$ of which v is a successor. Then by lemma 8, $\beta_i(y)=\beta_i(x)\cap K_i(y)$.

CASE 2: (y belongs to information set u of player i and) x is not a decision node of player i and $K_i(x)=\theta(x)$ [note that if $x=x_0$, then we are either in case 1 or in case 2, since $K_i(x_0)=\theta(x_0)$]. Then by definition of $K_i(\bullet)$ it must be true that $\bigcup_{t\in u}\theta(t)\subseteq\theta(x)$, that is, all the nodes in u are successors of x. Consider first the case where either the game has no chance moves (in which case we do not rule out the possibility that $x=x_0$) or the game has chance moves and $x\neq x_0$. Then by (2) of the definition of $\chi(\bullet)$, $\beta_i(x)=\zeta(x|\ \sigma)$. Since $\beta_i(x)\cap K_i(y)\neq\emptyset$, there must be a node $v \in u$ that is a successor of x on the path from x to $\zeta(x|\sigma)$, so that $\zeta(v|\sigma) = \zeta(x|\sigma)$. Then, by lemma 6, $\mu(v)=1$, and by (3) of the definition of $\chi(\bullet)$, $\beta_i(y)=\{\zeta(v|\sigma)\} = \beta_i(x) \cap K_i(y)$. Consider now the case where the game has chance moves and $x=x_0$. Then by (1) of the definition of $\chi(\bullet)$, $\beta_i(x_0)=\{z|\, z=\zeta(x|\sigma)$ for some $x \in \Sigma(x_0)\}$; by (3) of the definition of $\chi(\bullet)$, $\beta_i(y)=\{z|\, z=\zeta(t|\sigma)$ for some $t \in u$ with $\mu(t) > 0\}$ and by (2) of the definition of information $K_i(y) = \bigcup_{t \in u} \theta(t)$. Since, by lemma 9, for every node $t \in u$ with $\mu(t)>0$, $\zeta(t|\sigma)=\zeta(x|\sigma)$ for some $x \in \Sigma(x_0)$, it follows that $\beta_i(y)=\beta_i(x_0) \cap K_i(y)$.

CASE 3: (y belongs to information set u of player i and) x is not a decision node of player i and $K_i(x) \neq \theta(x)$ and $K_i(x) \neq K_i(p_x)$. Then it follows from the definition of $K_i(\bullet)$ that p_x is a decision node of player i. Let h be the information set to which p_x belongs. Let $\hat{h}=\text{supp}(\mu|h)$. Then, by (3) of the definition of $\chi(\bullet)$, $\beta_i(p_x)=\{z|\, z=\zeta(w|\sigma)$ for some $w \in \hat{h}\}$. Let c be the choice that leads from p_x to x. Then, by (5) of the definition of $\chi(\bullet)$, $\beta_i(x)=\{z|\, z=\zeta(y|\sigma)$ for some $y \in \Sigma(\hat{h}|c)\}$. Since $\beta_i(x) \cap K_i(y) \neq \varnothing$, there must be a node $s \in u$ that lies on the path from some $y \in \Sigma\}(\hat{h}|c)$ to $\zeta(y|\sigma)$. Hence, $\zeta(s|\sigma)=\zeta(y|\sigma)$. By lemma 10, $\beta_i(y)=\beta_i(x) \cap K_i(y)$.

CASE 4: (y belongs to information set u of player i and) x is not a decision node of player i and $K_i(x) \neq \theta(x)$ and $K_i(x) = K_i(p_x)$. Then by (4) of the definition of $\chi(\bullet)$, $\beta_i(x)=\beta_i(p_x) \cap K_i(x)$. Consider the path from x_0 to x. If $\beta_i(w)=\beta_i(x_0)$, for every node w on this path, and the game has no chance moves, then the result follows from lemma 6 [by (1) of the definition of $\chi(\bullet)$, $\beta_i(x_0)=\{\zeta(x_0|\sigma)\}$; since $\beta_i(x_0) \cap K_i(y) \neq \varnothing$, there exists a node $t \in u$ that lies on the path from x_0 to $\zeta(x_0|\sigma)$, hence $\zeta(t|\sigma) = \zeta(x_0|\sigma)$; by lemma 6, $\mu(t)=1$; by (3) of the definition of $\chi(\bullet)$, $\beta_i(y)=\{\zeta(t|\sigma)\}$]. If $\beta_i(w)=\beta_i(x_0)$, for every node on the path from x_0 to x and the game has chance moves, then the result follows from lemma 9. If $\beta_i(x) \neq \beta_i(x_0)$, let w be the node such that $\beta_i(w)=\beta_i(x)$ and $\beta_i(p_w) \neq \beta_i(x)$. Then either w is a decision node of player i, in which case the result follows from the argument of case 1, or w is not a decision node of player i, in which case if $K_i(w)=\theta(w)$, the argument of case 2 applies, while if $K_i(w) \neq \theta(w)$ then the argument of case 3 applies, since it must be $K_i(w) \neq K_i(p_w)$.

PROOF OF PROPOSITION 2. Let the simple assessment (σ,μ) be a sequential equilibrium and let $\beta=\chi(\sigma,\mu)$ be the corresponding profile of beliefs. In virtue of lemmas 5 and 11, it only remains to show that β satisfies the property of Individual Rationality. We need to show that if h is an information set of player i, $\hat{h}$ is the subset of h consisting of the predecessors of $\beta_i(h)$ - so that, by (3) of the definition of $\chi(\bullet)$, $\hat{h} = \text{supp}(\mu|h)$ - then for every choice c at h,

$$\sum_{z \in \beta_i(h)} U_i(z)\pi(z) \geq \sum_{z \in \bigcup_{x \in \hat{h}} \beta_i(\Sigma(x|c))} U_i(z)\pi(z) \qquad \text{(A.74)}$$

By (3) of the definition of $\chi(\bullet)$

(A.75) $$\sum_{z\in\beta_i(h)} U_i(z)\pi(z) = \sum_{x\in supp(\mu|h)} U_i(\zeta(x|\sigma))\pi(\zeta(x|\sigma))$$

By Contraction Consistency and by lemma 2 [cf. the argument that led to (A.14) and (A.15)], for every choice c at h,

(A.76) $$\sum_{z\in \bigcup_{x\in h} \beta_i(S(x|c))} U_i(z)\pi(z) = \sum_{x\in supp(\mu|h)} U_i(\zeta(S(x|c)|\sigma))\pi(\zeta(S(x|c)|\sigma))$$

Thus, using (A.75) and (A.76), (A.74) is equivalent to

(A.77) $$\sum_{x\in supp(\mu|h)} U_i(\zeta(x|\sigma))\pi(\zeta(x|\sigma)) \geq \sum_{x\in supp(\mu|h)} U_i(\zeta(S(x|c)|\sigma))\pi(\zeta(S(x|c)|\sigma))$$

First of all we show that for every $x\in h$ and for every choice c at h,

(A.78) $$\pi(\zeta(x|\sigma)) = \pi(\zeta(S(x|c)|\sigma))$$

If the game has no chance moves, $\pi(z)=1$ for all z and therefore (A.78) is true. Suppose the game has chance moves. Fix an arbitrary $x\in h$ and an arbitrary choice c at h. Let E be the unique event to which $\zeta(x|\sigma)$ belongs and E' be the unique event to which $\zeta(S(x|c)|\sigma)$ belongs. If $E\neq E'$, then the path from x_0 to $\zeta(x|\sigma)$ and the path from x_0 to $\zeta(S(x|c)|\sigma)$ follow different arcs at x_0 and reach the same node $x\in h$, contradicting uniqueness of plays in extensive games. Thus E=E' and (A. 77) follows from the definition of $\pi(\bullet)$.

Now, note that, by definition of $\pi(\bullet)$,

(A79) $$\pi(\zeta(x|\sigma)) = \pi(x)$$

Thus, using (A.78) and (A.79), we can re-write (A.77) as follows:

(A.80) $$\sum_{x\in supp(\mu|h)} U_i(\zeta(x|\sigma))\pi(x) \geq \sum_{x\in supp(\mu|h)} U_i(\zeta(S(x|c)|\sigma))\pi(x)$$

Dividing both sides of (A.80) by $\sum_{y\in supp(\mu|h)} \pi(y)$ we obtain, using the fact that, by definition of simple assessment, $\mu(x) = \dfrac{\pi(x)}{\sum_{y\in supp(\mu|h)} \pi(y)}$,

(A.81) $$\sum_{x\in supp(\mu|h)} U_i(\zeta(x|\sigma))\mu(x) \geq \sum_{x\in supp(\mu|h)} U_i(\zeta(S(x|c)|\sigma))\mu(x)$$

which is implied by sequential rationality.

REFERENCES

P. BATTIGALLI (1993), "Strategic independence and perfect Bayesian equilibria", Rapporto Intervento No. 93-011, Dipartimento di Economia e Produzione, Politecnico di Milano, Milano (Italy).

G. BONANNO (1992a), "Players' information in extensive games", *Mathematical Social Sciences*, 24, 35-48.

G. BONANNO (1992b), "Rational beliefs in extensive games", *Theory and Decision*, 33, 153-176.

D. FUDENBERG - J. TIROLE (1991a), "Perfect Bayesian Equilibrium and Sequential Equilibrium", *Journal of Economic Theory*, 53, 236-260.

D. FUDENBERG - J. TIROLE (1991b), *Game theory*, M.I.T. Press, Cambridge, Mass.

D. KREPS - G. RAMEY (1987), "Structural consistency, consistency, and sequential rationality", *Econometrica*, 55 (6), 1331-1348.

D. KREPS - R. WILSON (1982), "Sequential Equilibria", *Econometrica*, 50 (4), 863-894.

E. RASMUSEN (1989), *Games and information*, Basil Blackwell, Oxford.

R. SELTEN (1965), "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfraegetragheit", *Zeitschrift für die gesamte Staatswissenschaft*, 12, 301-324 and 667-689.

R. SELTEN (1975), "Re-examination of the perfectness concept for equilibrium points in extensive games", *International Journal of Game Theory*, 4, 25-55.