Center for Research on Economic and Social Theory

CREST Working Paper

Lexiocographic Refinements of Nash Equilibrium

Lawrence E. Blume

April, 1986
Number 87--9

DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, Michigan 48109

LEXICOGRAPHIC REFINEMENTS OF NASH EQUILIBRIUM

April 1986
(revised August 1987)

Lawrence E. Blume
Department of Economics
University of Michigan
Ann Arbor, MI 48109

The author thanks Ehud Kalai for stimulating his interest in this topic, and Mark Bagnoli, Adam Brandenburger, Eddie Dekel, Ket Richter and Bill Zame for helpful discussion. Support from the National Science Foundation is gratefully acknowleged.

1. INTRODUCTION.

The fundamental idea of game theory is that each player in the game acts in his own best interests, given the actions of the other players. Nash equilibrium makes this idea precise by defining each player's best interests as the maximization of his expected utility, where the expectation is taken with respect to the (mixed) strategies played by the other players.

New equilibrium concepts and refinements of old equilibrium concepts should adhere to this fundamental idea of self-interested action. This is to say, they must be justified in decision-theoretic terms.[1] A notion of "self interest" must be defined by specifying the preferences of the players, and equilibrium must be defined with respect to these preferences. This paper characterizes preferences that justify perfect and proper equilibrium as the outcome of rational, self-interested behavior. They also provide a framework for discussing the consistency of various rationality hypotheses on individual choice with properties that we might look for in other refinements of Nash equilibrium.

Not surprisingly, the preferences which rationalize refinements of Nash equilibrium are in fact refinements of the preference orderings given by expected utility maximization. Each player has a preference ordering consistent with expected utility maximization, where the expectation is with respect to the joint distribution of all other players' pure strategies. When a particular Nash equilibrium is rejected — if it is not perfect, for example — the argument is that some player really does not prefer the strategy assigned to him in the equilibrium. Some other strategy, which by definition must be of equal expected utility, is in fact better. For example, consider the game:

		Player II: Left	Player II: Right
Player I	Top	2,2	1,1
	Bottom	1,1	1,1

In this game the strategy pair (*Bottom*, *Right*) is a Nash equilibrium which is not perfect. Given Player *II*'s equilibrium strategy, both *Bottom* and *Top* have the same expected utility for Player *I*. However, the possibility of trembles makes Player *I* actually prefer *Top* to *Bottom*. Consideration of trembles refines the preference order generated by expected utility maximization.

Refinements of Nash equilibrium focus on the theories players hold about the play of the game. Sequential, perfect and proper equilibrium refine Nash equilibrium by requiring "second order" theories about equilibrium play. In the game, (*Bottom*, *Right*) fails to be perfect because,

[1] This point of view is argued for by Kohlberg and Mertens [1985].

although *Bottom* maximizes utility for Player I given the theory, "my opponent will certainly play *Right*," it does not do well when the second order theory "perhaps, with tiny probability, my opponent will play *Left*" is considered. The strategy pair does not survive the imposition of second order theories which say, "perhaps there will be a mistake!"

The choice theory that rationalizes the preferences expressed in refinements such as perfect and proper equilibrium is lexicographic. Only after the first order theory about play of the game leaves a player indifferent between two alternatives is the second order theory applied. This is easily seen in the following perturbations of the game:

		Player II Left	Player II Right
Player I	Top	2,2	$1-\epsilon$,1
Player I	Bottom	1,$1-\epsilon$	1,1

If Player I really thought that the probability of Player II's playing *Left* was positive, he would prefer *Top* to *Bottom* for sufficiently small positive ϵ. But this is not what happens. The pair (*Bottom*, *Right*) is perfect in all the perturbed games with $\epsilon > 0$.

My purpose in this paper is threefold. First, I will characterize the choice theory that underlies perfect and proper equilibrium. This choice theory is a version of lexicographic expected utility. I call it expected utility with lexicographic beliefs because each component of the expected utility vector has the same utility indicator, but (perhaps) different beliefs. This choice theory is justified by a minimal weakening of the Archimedian postulates in an axiom system for subjective expected utility. It is consistent with postulates of rationality such as the sure thing principle. Second, I will define equilibrium play for these preferences, and characterize perfect and proper equilibrium in the context of this equilibrium theory. I will identify those properties of higher order beliefs such that preference maximization with respect to these beliefs gives perfect equilibria and proper equilibria. Finally, I will characterize those higher order beliefs that give rise precisely to the admissible equilibria — Nash equilibria in which no player plays a weakly dominated strategy.

The goal of this research is to examine the choice-theoretic foundations of Nash equilibrium refinements. Two recent papers have also examined these foundational questions. Brandenburger and Dekel [1986a] also investigate the connection between lexicographic choice and Nash equilibrium refinements. We differ more in technique than in substance. McLennan [1986] takes a different approach to choice-theoretic foundations by identifying best response correspondences on the space of consistent conditional systems whose fixed points are precisely the sequential equilibria. McLennan is concerned with the lexicographic nature of beliefs, but not with the other aspects of choice theory.

Lexicographic orderings have lexicographic representations in the real numbers, and numerical representations in some non-Archimedian ordered fields. Archimedian ordered fields (such as the real numbers) are familiar, but the numerical representations provided in the non-Archimedian fields are more convenient for calculations than are lexicographic representations in the reals. Furthermore, the non-Archimedian approach will be of more help in generalizing the lexicographic refinements of Nash equilibrium to infinite games. Accordingly, I will state all theorems in the paper using real-valued lexicographic representations, but I will use non-Archimedian representations in proving the theorems. The necessary algebra is sketched as it is needed in Appendices 1 and 2.

2. EXPECTED UTILITY WITH LEXICOGRAPHIC BELIEFS.

Imagine a decisionmaker confronting a choice situation. There are finite numbers of states of the world $\omega \in \Omega$ and outcomes $x \in X$. The decisionmaker chooses an act, which is a map from Ω to probability distributions on outcomes. Let $f(\omega)(x)$ denote the probability that act f assigns to outcome x in state ω.

This description of choice is a simple version of the Anscombe and Aumann [1963] framework, which is sufficiently rich for game theory.[2] If the decisionmaker is an expected utility maximizer, then there is a probability distribution ρ on Ω and a utility function u on outcomes such that act f is at least as good as act g ($f \succ g$) if and only if:

$$\sum_{\omega\in\Omega} \rho(\omega) \sum_{x\in X} u(x)f(\omega)(x) > \sum_{\omega\in\Omega} \rho(\omega) \sum_{x\in X} u(x)g(\omega)(x).$$

To reinforce the distinction between subjective uncertainty (on Ω) and objective uncertainty (on X) it is useful to define for each act f the random variable

$$U(f)(\omega) = \sum_{x\in X} u(x)f(\omega)(x).$$

Then $f \succeq g$ if and only if

$$E_\rho\{U(f)\} \geq E_\rho\{U(g)\}.$$

Lexicographic expected utility is the non-Archimedian version of this choice behavior. The decisionmaker is a lexicographic expected utility maximizer if there are K utility functions u_k

[2] Another version of this choice theory can be found in Brandenburger and Dekel [1986a,b].

on outcomes and K probability distributions ρ_k on Ω such that $f \succeq g$ if and only if the vector $(E_{\rho_k}\{U_k(f)\})_{k=1}^K$ equals or exceeds the vector $(E_{\rho_k}\{U_k(g)\})_{k=1}^K$ in the usual lexicographic ordering on $\mathbf{R}^{|\mathbf{A}|}$.[3] I will use the notation $\geq_L$ and $>_L$ to refer to the weak and strict lexicographic orders, respectively.

Expected utility with lexicographic beliefs is a special case of lexicographic expected utility, where the utility indicator u is independent of k, and only ρ varies with the index. In other words, $f \succeq g$ if and only if the vector $(E_{\rho_k}\{U(f)\})_{k=1}^K \geq_L (E_{\rho_k}\{U(g)\})_{k=1}^K$ in the usual lexicographic ordering on $\mathbf{R}^\mathbf{k}$.

The rationality of expected utility is best discussed in terms of the axioms on choice behavior which characterize those preferences with an expected utility representation. The following axioms characterize those preferences which can be represented by expected utility with lexicographic beliefs. As before, X will denote the finite set of outcomes and Ω is the finite set of states. The set of probability distributions on X is $P(X)$. An act is a function $f : \Omega \mapsto P(X)$. Let $\mathcal{W}$ denote the set of subsets of Ω. The set of acts is denoted by L. The set L is a mixture space. The act $\alpha f + (1-\alpha)g$ is the act that, in state ω, draws x from distribution $f(\omega)$ with probability α, and from distribution $g(\omega)$ with probability $1-\alpha$. I will assume that preferences are complete (Axiom 1.), and use the following definitions: $f \succ g$ iff $f \succeq g$ and not $g \succeq f$; $f \sim g$ if $f \succeq g$ and $g \succeq f$.

Axioms:

1. $\succeq$ is complete, transitive and reflexive on L.

2. (independence) For all $f, g, h \in L$, if $f \succeq g$ and $0 \geq \alpha \geq 1$ then $\alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h$, with strict preference if $f \succ g$ and $\alpha > 0$.

Axioms 1 and 2 allow for the definition of the orderings $\succeq_S$, as in Savage [1954].

Definition. $f \succeq_S g$ iff $f' \succeq g'$ for every pair of acts f' and g' such that $f' = f$ on S, $g' = g$ on S and $f' = g'$ on Ω/S.

Definition. An event S is **null** iff $f \sim_S g$ for all $f, g \in L$.

Proposition 1. The orderings $\succeq_S$ are complete, transitive and reflexive.

[3] In general, infinite K may be required, but due to the finiteness of Ω, finite K suffices for my purposes.

To say that $f \succeq_S g$ is to say that if the decisionmaker knew that S were going to occur, then, given this knowledge, he would find f at least as good as g. The proof of the Proposition can be found in Savage [1954] (pg. 23).

3. (Archimedian axiom) If $f \succeq_{\{\omega\}} g \succeq_{\{\omega\}} h$, then there is a γ, $0 \leq \gamma \leq 1$ s.t. $\gamma f + (1-\gamma)h \sim_{\{\omega\}} g$.

4. (non-triviality) There exists $\omega \in \Omega$, $f, g \in L$ s.t. $f \succ_{\{\omega\}} g$.

5. (Non-null state independence) For non-null states $\omega, \omega' \in \Omega$, if $f(x)(\omega) = f(x)(\omega')$ for all $x \in X$, if $g(x)(\omega) = g(x)(\omega')$ for all $x \in X$, and $f \succeq_{\{\omega\}} g$, then $f \succeq_{\{\omega'\}} g$.

Theorem 1. Let $\succeq$ on L satisfy Axioms 1–5. Then there exists a utility function $u : X \mapsto \mathbf{R}$ with $\min_x u(x) = 0$, $\max_x u(x) = 1$, and a finite sequence of probability measures $\{\rho_k\}_{k=1}^K$ on Ω such that $f \succeq g$ iff

$$\left(\sum_{\omega \in \Omega} \rho_n(\omega) \sum_{x \in X} u(x) f(x)(\omega)\right)_{k=1}^{K} \geq_L \left(\sum_{\omega \in \Omega} \rho_n(\omega) \sum_{x \in X} u(x) g(x)(\omega)\right)_{k=1}^{K}.$$

The utility function u is uniquely determined, as are first-order beliefs ρ_1.

Conversely, if a non-trivial preference order $\succeq$ on L has such a representation, then it satisfies Axioms 1–5.

I will call the representation of preferences (beliefs) presented in Theorem 1 a *factor representation.* This Theorem is proved in Appendix 1. I use techniques from non-standard analysis. Richter [1971] has shown that any reflexive, transitive and complete ordering can be represented by a utility function whose values are non-standard numbers. Using the techniques of Appendix 1 it is straightforward to show that if a preference order over lotteries satisfies all of the usual axioms except the continuity axiom, then there exists an expected utility representation where both probabilities and utilities are elements in a non-standard model of the real numbers. The particular representation theorem I am interested in investigates the implications of a partial relaxation of the continuity assumption. Utilities will be Archimedian but probabilities will not be.

The axioms are standard except for Axiom 3. If this statement were true for unconditional preference rather than just for conditional preference given $\{\omega\}$, then the conclusion of Theorem 1 would hold with $K = 1$, which is to say that the preference ordering would be representable by Archimedian expected utility.

Unlike the Archimedian analysis, the factor representation of beliefs in the non-Archimedian case is not unique. For example, in a two state world, an expected utility maximizer with lexicographic beliefs would make the same choices with lexicographic probability distribution $(1/2, 1/2)$, $(1, 0)$ as he would with lexicographic probability distribution $(1/2, 1/2)$, $(3/4, 1/4)$.

In applications to game theory, each player's strategy is a random variable that takes values in a finite set of pure strategies. In Nash equilibrium, each player believes that the strategies of the other players are independent random variables. The key decision-theoretic aspect of the independence of events A and B is that preferences over bets about the occurrence of event A are unaffected by knowledge of the occurrence or non-occurrence of event B. In the Archimedian world, this implies that beliefs are such that $\Pr(A \cap B) = \Pr(A)\Pr(B)$. The non-Archimedian version of the product rule for independent events is not a consequence of the independence of preferences over bets on A of knowledge about B. An adequate treatment of independence with non-Archimedian beliefs is somewhat technical and tedious. The interested reader can find a discussion of the issues, the correct definition of independence for my purposes and a characterization theorem in Appendix 2.

Finally, we need the notion of marginal distribution. If $\rho_1, \ldots, \rho_A$ is a joint lexicographic probability distribution of the two random variables $\tilde{x}$ and $\tilde{y}$, then the marginal lexicographic probability distribution of the random variable $\tilde{x}$ is given by

$$\rho_k(\tilde{x} = a) = \sum_b \rho_k(\tilde{x} = a, \tilde{y} = b).$$

It is easy to see that this definition of marginal distributions has all the required properties.

3. LEXICOGRAPHIC CHOICE: AN EXAMPLE.

An example will help to distinguish the difference between expected utility with lexicographic beliefs, Archimedian expected utility, and other lexicographic concepts that arise in game theory. Consider the decision tree in figure 1. There is one decisionmaker, playing against nature. Nature chooses A, B or C. The decisionmaker then chooses R or L. If nature has chosen A or B and the decisionmaker has chosen L, the decisionmaker gets to choose again from among r and l. Utility payoffs are given at the terminal nodes.

The decision problem has three states of nature: A, B and C. The decisionmaker can choose from among four actions: R, Lr, Lm and Ll. Let us suppose that, initially, the decisionmaker is an Archimedian expected utility maximizer, and that his beliefs assign $Pr\{C\} = 1$. Then all actions have expected utility 1, and so $R \sim Lr \sim Lm \sim Ll$.

Suppose now that the decisionmaker has lexicographic beliefs, and his second order beliefs assign $Pr\{A\} = 1/3$, $Pr\{B\} = 2/3$, $Pr\{C\} = 0$. Since first order beliefs leave him indifferent

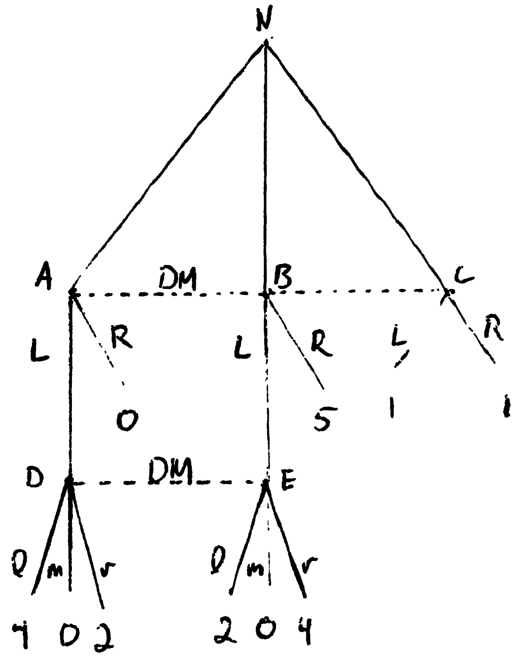

fig 1

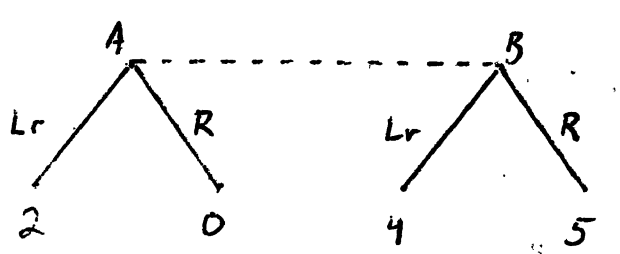

fig 1a

among all his choices, he resolves the tie with his second order beliefs. Second order beliefs assign $E\{U(R)\} = 10/3$, $E\{U(Lr)\} = 5/3$, $E\{U(Ll)\} = 8/3$ and $E\{U(Lm)\} = 0$, so $R \succ Lr \succ Ll \succ Lm$.

Another notion of Lexicographic choice underlies sequential equilibrium. The choice model of Kreps and Wilson [1982] has the decisionmaker use his second order beliefs only in the event that he should find himself in the second information set. Put differently, second order beliefs can only be used to break the tie between Lr and Ll. This theory of choice, without further refinement, can lead to choice behavior which cannot be rationalized by expected utility theory (Archimedian or not). Consider the strategy Lr. The choice of local strategy L at the decisionmaker's first information set and r at the second information set is rationalized only by beliefs which assign probability 1 to state C. Beliefs at the second information set are really conditional beliefs on the event $\{A, B\}$, and must have $Pr\{B\} > 2/3$. Given these beliefs it must be the case that $r \simeq l \sim m$.

Although Lr is rationalized by sequential rationality with lexicographically consistent beliefs, it runs afoul of the sure thing principle. Since $r \succ m$, the event $\{A, B\}$ is not null. Given the choice r at the second information set, the choice problem at the first information set becomes as described in Figure 1a. Since $Pr\{B\} \geq 2Pr\{A\}$, it must be the case that, given the event $\{A, B\}$, $R \succ L$. Given the event $\{C\}$, $R \sim L$. Since $\{A, B\}$ is not null, we should conclude from the sure thing principle that $R \succ L$.

The point of this example is to illustrate that the choice criteria embodied in the sequential equilibrium concept are insufficient by themselves to guarantee consistency with conventional theories of choice under uncertainty.[4]

4. LEXICOGRAPHIC CHOICE IN NORMAL FORM GAMES.

In this section I will define an equilibrium concept for normal form games when players preferences are representable by expected utility with lexicographic beliefs. I will discuss the relationship of this equilibrium concept to Nash equilibrium with Archimedian expected utility.

An N-person game in normal form is a $2N$-tuple $G = < I_n, u_n >_{n=1}^{N}$. Player n has the set I_n of pure strategies available to him. I will use the symbol I_n both for the set of pure strategies available to player n and its cardinality. Let $I = \prod_{n=1}^{N} I_n$. This is the set of possible pure strategy vectors. I will also use the symbol I to refer to the cardinality of this set. Let I_{-n} denote the set of

[4] In this paper I want to discuss only normal form games. I will discuss sequential equilibrium, other extensive form equilibrium concepts and lexicographic choice in another paper.

pure strategy combinations for all traders other than n. Let S_n denote the set of mixed strategies available to player n; this is the set

$$S_n = \{(p(1), \dots, p(I_n)) \ : \ p(i) \in \mathbf{R}_+, \ \sum_{i=1}^{I_n} p(i) = 1\}.$$

Let $S = \prod_{n=1}^{N} S_n$, and let S_{-n} denote the set of probability distributions on I_{-n}.

If $s \in S$ is a mixed strategy vector, then s_n is player n's mixed strategy and s_{-n} is the marginal distribution of plays of the players other than n. The same notation will apply to the set I and any pure strategy vector $i \in I$. Player n's utility function is a map $u_n : I \mapsto \mathbf{R}$.

The games and strategies I consider are just those that arise in normal form game theory with Archimedian preferences. Now, however, I take up beliefs. Nash equilibrium in an n-player game requires that each player have a theory of how the game is to be played, that each player choose a strategy which maximizes his utility given his theory, and that each player's theory be consistent with other players' theories. Any n-tuple of strategies which meets these three criteria is a Nash equilibrium strategy n-tuple. Lexicographic beliefs allows for a richer specification of beliefs than is available with Archimedian expected utility, and so they must be formally examined.

Let $K \geq I$ be an arbitrarily chosen integer. Player n has beliefs about the actions of other players. These (non-Archimedian) beliefs can be represented as elements of the set $R_n = \prod_{k=1}^{K} S_{-n}$. Let $r_n = (r_n^1, \dots, r_n^K)$ be an element of R_n. Probability distribution r_n^1 represents player n's first order beliefs, r_n^2 represents his second order beliefs, and so forth. Were player n to play strategy $s_n \in S_n$, his lexicographic expected utility vector would be:

$$U_n(s_n; r_n) = \Big(\sum_{i \in I} u_n(i_n, i_{-n}) s_n(i_n) r_n^1(i_{-n}), \dots, \sum_{i \in I} u_n(i_n, i_{-n}) s_n(i_n) r_n^K(i_{-n})\Big).$$

It is this vector which must be maximized with respect to the lexicographic ordering $\geq_L$ on $\mathbf{R}^{\mathbf{K}}$.

Let $R = \prod_{n=1}^{N} R_n$. An element of R is a vector of joint beliefs.

Definition. A strategy vector-belief vector pair $(s, r) \in S \times R$ is a *lexicographic Nash equilibrium* if for each player n,

$$\Big(\sum_{i \in I} u_n(i_n, i_{-n}) s_n(i_n) r_n^1(i_{-n}), \dots, \sum_{i \in I} u_n(i_n, i_{-n}) s_n(i_n) r_n^K(i_{-n})\Big) \geq_L$$
$$\Big(\sum_{i \in I} u_n(i_n, i_{-n}) s'_n(i_n) r_n^1(i_{-n}), \dots, \sum_{i \in I} u_n(i_n, i_{-n}) s'_n(i_n) r_n^K(i_{-n})\Big)$$

for all $s'_n \in S_n$, and

$$r_n^1 = s_{-n}.$$

The first condition can be more conveniently written as $U_n(s_n; r_n) \geq_L U_n(s'_n, r_n)$.

The first requirement of the definition is the maximization hypothesis — that each player is acting in his own self-interest given the expected actions of others. The second requirement is consistency of each player's theory with actual play of the game.

The purpose of lexicographic Nash equilibrium is to refine Nash equilibrium by imposing constraints on higher order beliefs. To see how this works, notice that, by constraining first order and higher order beliefs to be identical, any Nash equilibrium play is also realizable as a lexicographic Nash equilibrium. By imposing other constraints on higher order beliefs, we can, as we shall see, pick out various refinements of Nash equilibrium.

Theorem 2. *i*) If s is a Nash equilibrium of a game G, then there exists a $r \in R$ such that (s, r) is a Lexicographic Nash equilibrium.

ii) If (s, r) is a lexicographic Nash equilibrium of a game G, then s is a Nash equilibrium of G.

Proof. *i*) Obvious from the definition. Set $r_n^k = s_{-n}$ for all n and k.

ii) If (s, r) is a lexicographic Nash equilibrium, then

$$u_n(i_n, i_{-n})s_n(i_n)r_n^1(i_{-n}) \geq u_n(i_n, i_{-n})s'(i_n)r_n^1(i_{-n})$$

for all $s' \in S_n$. Since $r_n^1 = s_{-n}$, s_n is a Nash best response to s_{-n}. ∎

Again, this Theorem clarifies the purpose for introducing lexicographic beliefs. The set of Nash equilibria strategy vectors and the set of lexicographic Nash equilibria strategy vectors for a game G are identical. Further constraints on the nature of lexicographic beliefs, beyond what is already required for the existence of equilibrium, will serve to refine the Nash equilibrium concept.

5. PERFECTION, PROPERNESS AND LEXICOGRAPHIC NASH EQUILIBRIA.

In this section I identify those constraints on higher order belief which characterize perfect and proper equilibria among the class of lexicographic Nash equilibria of a normal form game G.[5]

[5] For another discussion of the lexicographic characterization for perfect and proper equilibrium, see Brandenburger and Dekel [1986a].

The idea behind perfect equilibrium is that mistakes are possible, although extremely improbable, and that, consequently, no strategy combination can be regarded as impossible. Formally, a Nash equilibrium $s \in S$ is perfect if and only if there exists a sequence $\{s^j\}_{j=1}^{\infty}$ of completely mixed strategies converging to s such that, for player n, s_n is a best response to s_{-n}^j for all j. This idea has a natural expression in terms of lexicographic beliefs. Each player has higher-order beliefs which assign positive probability to every strategy combination. An important aspect of perfect equilibrium is that each player's strategy choice is required to be robust to the same test sequence of completely mixed strategies. This idea can be expressed lexicographically in the following condition:

Definition. A belief vector $r \in R$ is **independent** if each player believes the play of all other players j and k to be independent random variables. It is **shared** if there is a lexicographic distribution s on I such that r_i is the marginal distribution under s of the actions of all players other than i.

The definition of shared beliefs requires that player i's first order beliefs be the the marginal distribution of s^1 on I_{-i}, that his second order beliefs be the marginal distribution of s^2 on I_{-i}, and so forth. The shared beliefs condition says that all players share the same theories about how the game is to be played. Independence is the appropriate generalization of Archimedian independence to the case of lexicographic probabilities. Notions of independence are discussed in some detail in Appendix 2.

Theorem 3. A strategy vector $s \in S$ is a perfect Nash equilibrium of G if and only if there exists a belief vector $r \in R$ with shared, independent beliefs such that (s, r) is a lexicographic Nash equilibrium and such that, for all n and all i_{-n}, there exists a k such that $r_n^k(i_{-n}) > 0$.

Proper equilibrium is a refinement of perfect equilibrium. In proper equilibrium, trembles are constrained so that, loosely speaking, if strategy x is worse for a player than strategy y, then the sequence of trembles must assign probabilities to x and y such that the odds of x to y get arbitrarily large. Myerson [1978] motivates proper equilibria with explicit reference to beliefs about mistakes (as opposed to Selten [1975] who does not in his study of perfection). Formally, a Nash equilibrium $s \in S$ is proper if and only if there exists a sequence $\{s^j\}_{j=1}^{\infty}$ of completely mixed strategies converging to s and a sequence of positive real numbers $\{\epsilon^j\}_{j=1}^{\infty}$ with limit 0 such that if, for player n, $\sum_{i_{-n} \in I_{-n}} u_n(i_n, i_{-n}) s_{-n}^j > \sum_{i_{-n} \in I_{-n}} u_n(i'_n, i_{-n}) s_{-n}^j$, then $s_n^j(i_n) \geq \epsilon^j s_n^j(i'_n)$; and such that, for player n, s_n is a best response to s_{-n}^j for all j. The constraints on trembles in proper equilibrium also have a natural interpretation in terms of lexicographic beliefs. If strategy a is worse for player n

than strategy b in equilibrium, then, for all other players, player n's strategy a cannot have positive probability in any belief of order less than or equal to the order of belief where b first has positive probability. This says, roughly, that players must believe b to be infinitely more likely than a.

Theorem 4. A strategy vector $s \in S$ is a proper Nash equilibrium of G if and only if there exists an $r \in R$ with shared, independent beliefs such that (s, r) is a lexicographic Nash equilibrium, and such that if, for player n,

$$\sum_{i_{-n} \in I_{-n}} u_n(i_n, i_{-n}) r_{-n}(i_{-n}) >_L \sum_{i_{-n} \in I_{-n}} u_n(j_n, i_{-n}) r_{-n}(i_{-n}),$$

then for all $m \neq n$, $r_m^k(i_1, \ldots, i_{m-1}, i_{m+1}, \ldots, j_n, \ldots, i_N) > 0$ implies that there is an $l < k$ such that $r_m^l(i_1, \ldots, i_{m-1}, i_{m+1}, \ldots, i_n, \ldots, i_N) > 0$.

Proofs of these two theorems can be found in Appendix 3.

The source of many of the desirable properties of perfect and proper equilibrium can be found in the behavior of higher-order beliefs. For example, perfect equilibrium play does not involve the use of dominated strategies. The reason for this is, in Savage's [1954] language, that the beliefs of the players admit no null events. No combination of pure strategies is impossible, although their first order probabilities may be 0. Since no combination of plays can be completely neglected, dominated strategies can never be preference maximizing.

6. CORRELATION AND ADMISSIBLE EQUILIBRIA.

In this section I provide a lexicographic characterization of admissible equilibria — equilibria wherein no dominated strategies are played. This characterization involves relaxing the hypothesis of shared higher order beliefs and the hypothesis of independent beliefs.

Definition. A Nash equilibrium is **admissible** if no player plays a dominated strategy in equilibrium.

Theorem 1. A strategy vector $s \in S$ is an admissible Nash equilibrium if and only if there exists an $r \in R$ such that (s, r) is a lexicographic Nash equilibrium and such that, for all n and all i_{-n}, there exists a k such that $r_n^k(i_{-n}) > 0$.

The proof of Theorem 1 is in Appendix 3.

It is known that the set of perfect equilibria and admissible equilibria coincide in two-person games. The reason is clear in comparing Theorems 3 and 1. The assumptions of shared beliefs and independent beliefs impose no restrictions when there are only two players in the game. Each player is the only player who holds beliefs about the other, so the assumption of shared beliefs is trivially satisfied. The independent beliefs assumption also does not bind because each player holds beliefs about the play of only one other player. When in games containing three or more players, there can be admissible equilibria which are not perfect. In general n-person games, relaxation of both the shared beliefs and the independent beliefs assumptions is required to identify all the admissible equilibria, as the following two examples show.

The first example has an admissible equilibrium not supported by shared beliefs. The payoff matrix for this three-person game is

	l	r		l	r		l	r
T	$3,1,4$	$1,1,0$	T	$3,0,0$	$1,0,3$	T	$2,0,0$	$2,0,0$
B	$3,0,8$	$3,0,0$	B	$3,1,4$	$3,1,0$	B	$2,0,0$	$2,0,0$
	L			C			R	

Figure 2.

where player 1 chooses the matrix L,C or R; player 2 chooses the row T or B; and player 3 chooses the column l or r. The strategy triple (R,T,r) is a Nash equilibrium which is justified by higher order beliefs in which, given $\{A,B\}$, player 2 believes that L will be played with high probability, while player 3 believes that C will be played with high probability. It is easy to see that this equilibrium cannot be supported by any shared beliefs which assign positive probability to the strategy set $\{L,C\}$ of player 1.

The next example is a game with an admissible equilibrium supported only by dependent beliefs. This example comes from van Damme [1983].[6] The payoff matrix for this three-person game is

[6] See Figure 2.2.1 on page 31. Martin Osborne and Vijay Krishna called this example (and its required violation of independence) to my attention.

	l	r
T	1,1,1	1,0,1
B	1,1,1	0,0,1

L

	l	r
T	1,1,0	0,0,0
B	0,1,0	1,0,0

R

Figure 3.

where player 1 chooses a row T, B, player 2 chooses a column l, r and player 3 chooses a matrix L, R. The strategy combination (B, l, L) is an admissible equilibrium. Any vector of lexicographic beliefs which supports player 1's decision and which assigns positive probability to every combination of strategies for players 2 and 3 must treat individual defections by each player and simultaneous defections by both players as events of the same order. This stands in contradiction to the implication of independence that the joint probability of two independent second-order events must be at least third-order.

Dropping the independent and shared beliefs assumptions in the manner of Theorem 1 still gives lexicographic Nash equilibria because first order beliefs are shared and independent. Nonetheless, the dropping of these assumptions for higher order beliefs is far from the spirit of Nash equilibrium.

In a Nash equilibrium, the equilibrium strategies of all the players are common knowledge. One implication of this fact for Nash equilibrium is that (first order) beliefs are shared. Two kinds of arguments have traditionally been made for this assumption. Both are problematic. First, each player understands the game, and thus can compute the best responses of his opponents. Second, repeated play of the game may lead to equilibrium play as players observe and respond to the play of their opponents. Neither of these arguments seems to apply to higher order beliefs. Higher order beliefs may represent a theory of irrational play, and thus may not be susceptible to a priori analysis. Furthermore, higher order beliefs are never directly observable, unlike the first order beliefs which correspond to actual play of the game. If the assumption of shared higher order beliefs is to be defended, new arguments must be found.

Perhaps the most successful approach to the justification of equilibrium is provided by Aumann [1987]. Aumann shows that equilibrium is equivalent to the expression of Bayesian rationality on the part of the players, where the uncertainty is not just over nature's moves, but also over the play of other players in the game. The appropriate notion of equilibrium is not the Nash equilibrium, but rather the more general correlated equilibrium. His analysis is driven in part by the Common Prior Assumption, that all players share the same prior on the common state space of the players' decision problems. In Aumann's analysis, the Nash equilibria can be recovered from those information structures where the players' information partitions are independent.

One can imagine carrying out the same analysis in the domain of expected utility with non-Archimedian beliefs. Both formally and at the interpretative level, Aumann's story remains equally valid once the correlated equilibrium concept is modified to account for non-Archimedian beliefs just as the Nash equilibrium concept was modified in Section 4.

It is not my intention to redo Aumann's arguments in the non-Archimedian framework. Instead, I want to briefly discuss non-Archimedian correlated equilibrium and its relationship to lexicographic Nash equilibrium. The following definitions for correlated and subjective correlated equilibria are straightforward extensions to th non-Archimedian framework of Aumann's [1974, 1987] definitions and restatements of these definitions. I find it most convenient to work directly with correlated equilibrium distributions.

Definition. **A lexicographic subjective correlated equilibrium (l.s.c.e.)** is an $N-tuple$ of non-Archimedian probabilities $\left((\rho_n^k)_{k=1}^K\right)_{n=1}^N$ on I such that, for each player n,

$$\left(\sum_{i_{-n}} u_n(i_{-n}, i_n)\rho_n^k(i_{-n}, i_n)\right)_{k=1}^K \geq_L \left(\sum_{i_{-n}} u_n(i_{-n}, j_n)\rho_n^k(i_{-n}, i_n)\right)_{k=1}^K$$

for all $i_n, j_n \in I_n$.
A **correlated equilibrium (l.c.e)** is an l.s.c.e. in which, for all players m and n, $\rho_n^k = \rho_m^k$.

The first two observations are trivial. Every lexicographic Nash equilibrium is an l.s.c.e. If an l.s.c.e. is a lexicographic Nash equilibrium, then first order beliefs are shared and independent, and for each player n, strategy combinations i_{-n} of the other players are independent of his own action i_n. These observations follow immediately from the definitions. They are collected in the next Theorem, which also specifies the maps from l.s.c.e.'s to lexicographic Nash equilibria and back.

Theorem 2. 1. If (s, r) is a lexicographic Nash equilibrium, then there exists a lexicographic subjective correlated equilibrium (l.s.c.e.) $\left((\rho_n^k)_{k=1}^K\right)_{n=1}^N$ with $\rho_n^k(i_n, i_{-n}) = r_n^k(i_{-n})s_n(i_n)$.

2. If an l.s.c.e. is a lexicographic Nash equilibrium, then $\rho_n^1(i) = \rho^1(i_1)\cdots\rho^1(i_N)$, and $\rho_n^k(i) = \rho_n^k(i_{-n})\rho_n^k(i_n)$.

The question I want to raise concerns the role of l.c.e.'s. Is every lexicographic Nash equilibrium describable as a correlated equilibrium satisfying the Common Prior Assumption? I claim that the answer is "no". Neither of the examples in this section are l.c.e.'s. They can only be described as a lexicographic subjective correlated equilibria. Nonetheless, as Theorem 2 makes clear, the equilibrium is not totally subjective. First order priors agree. Players' priors differ only in higher orders. The relationship between l.c.e.'s and lexicographic Nash equilibria are detailed in the next Theorem, whose proof can be found in Appendix 3.

Theorem 3. If an l.c.e. $((\rho^k)_{k=1}^K$ is an lexicographic Nash equilibrium, then beliefs are shared and independent. If (s, r) is a lexicographic Nash equilibrium with shared and independent beliefs, then (s, r) describes an l.c.e. with $\rho^k(i) = r_n^k(i_{-n})s_n(i_n)$ for all n and $i \in I$.

The sets of perfect and proper equilibrium can be justified in the spirit of Aumann [1987] as expressions of Bayesian rationality where various common knowledge restrictions on beliefs are imposed, chief among them being the Common Prior Assumption. In light of Theorem 3, the significance of the first example is its demonstration that if the admissible Nash equilibria are to be justified along the lines laid out by Aumann, the non-Archimedian version of the Common Prior Assumption must be relaxed to a Common First Order Prior Assumption. Differences in higher order beliefs cannot be explained by information differences alone. Other "conceptual inconsistenc[ies] between players"[7] must be appealed to in order to justify the class of admissible equilibria. Exactly what is the subset of admissible equilibria that can be described as l.c.e.'s? Combining the criteria in Theorems 1 and 3, the answer is given by Theorem 3: the perfect equilibria.

[7] Aumann's words, [1987] pg 15.

Appendix 1: Proofs for Section 2.

In this Appendix I prove the theorems of Section 2. First I will prove a representation theorem where expected utility takes values in a non-Archimedian ordered field $\mathcal{R}$ that contains a proper subfield isomorphic to $\mathbf{R}$. Next I will use this representation to prove Theorem 1. Finally, I will discuss some of the issues surrounding the definition of independent events and independent random variables.

My notation will not distinguish between elements in $\mathbf{R}$, and elements of the subfield of $\mathcal{R}$ isomorphic to $\mathbf{R}$.

Theorem 1′. Let $\succeq$ on L satisfy Axioms 1–5. Then there exists a utility function $u : X \mapsto \mathbf{R}$ with $\min_x u(x) = 0$, $\max_x u(x) = 1$, a proper ordered-field extension $\mathcal{R}$ of the real numbers, and a $\mathcal{R}$-valued probability measure $\hat{\rho}$ on Ω such that $f \succeq g$ iff

$$\sum_{\omega\in\Omega} \hat{\rho}(\omega) \sum_{x\in X} u(x)f(x)(\omega) \geq \sum_{\omega\in\Omega} \hat{\rho}(\omega) \sum_{x\in X} u(x)g(x)(\omega).$$

Conversely, if a preference order $\succeq$ on L has such a representation, then it satisfies Axioms 1–5.

Proof of Theorem 1′.

First I will use the independence axiom (Axiom 2) to establish some monotonicity and substitution properties of preferences, and then the sure-thing principle.

Lemma 1. If $f \succeq g$ and $0 \leq \beta < \alpha \leq 1$, then $\alpha f + (1-\alpha)g \succeq \beta f + (1-\beta)g$ with strict preference if $f \succ g$.

Proof. From Axiom 2, $\alpha f + (1-\alpha)g \succeq g$, with strict preference if $f \succ g$. Let $\gamma = \beta/\alpha$. Applying Axiom 2 to the mixture of the lottery $\alpha f + (1-\alpha)g$ with probability γ and g with probability $1-\gamma$ gives $\alpha f + (1-\alpha)g \succ \beta f + (1-\beta g)$, with strict preference if $f \succ g$. ∎

Lemma 2. If $e \succeq f$ and $0 \leq \alpha \leq 1$, then $g \succeq h$ implies $\alpha e + (1-\alpha)g \succeq \alpha f + (1-\alpha)h$, with strict preference if $\alpha < 1$ and $g \succ h$.

Proof. This follows from repeated applications of Axiom 2. First, $\alpha e + (1-\alpha)g \succeq \alpha f + (1-\alpha)g$. Second, $\alpha f + (1-\alpha)g \succeq \alpha f + (1-\alpha)h$, with strict preference if $\alpha < 1$ and $g \succ h$. The conclusion follows from Axiom 1 (transitivity). ∎

If f and h are lotteries, define the lottery

$$f|_S h(\omega) = \begin{cases} f(\omega) & \text{if } \omega \in S, \\ h(\omega) & \text{if } \omega \notin S. \end{cases}$$

Definition. For $S \in \mathcal{W}$, write $f \succeq_S g$ if there exists a lottery h such that $f|_S h \succeq g|_S h$.

The sure-thing principle guarantees that these orderings are well-defined, complete, transitive and symmetric.

Lemma 3. If there exists a lottery h such that $f|_S h \succ g|_S h$, then for all lotteries k, $f|_S k \succ g|_S k$. If $f|_S h \sim g|_S h$, then for all lotteries k, $f|_S k \sim g|_S k$.

Proof. ($\succ$) Suppose not. Suppose instead that for lotteries f, g, h, k, $f|_S h \succ g|_S h$ and $g|_S k \succeq f|_S k$. Then by Lemma 2,

$$\begin{aligned} \frac{1}{2} f|_S h + \frac{1}{2} f|_S k &\succ \frac{1}{2} g|_S h + \frac{1}{2} f|_S k, \\ \frac{1}{2} g|_S k + \frac{1}{2} f|_S h &\succeq \frac{1}{2} f|_S k + \frac{1}{2} f|_S h. \end{aligned} \tag{1}$$

Each side of the two equations in (1) is a mixture. Computing them,

$$\begin{aligned} \frac{1}{2} g|_S h + \frac{1}{2} f|_S k &= (\frac{1}{2} g + \frac{1}{2} f)|_S(\frac{1}{2} h + \frac{1}{2} k) \\ &= \frac{1}{2} g|_S k + \frac{1}{2} f|_S h. \end{aligned}$$

From transitivity (Axiom 1), $\frac{1}{2} f|_S h + \frac{1}{2} f|_S k \succ \frac{1}{2} f|_S h + \frac{1}{2} f|_S k$, which contradicts reflexivity (Axiom 1).

For the case of indifference ($\sim$), replace $\succ$ with $\sim$ and $\succeq$ with $\succ$ in (1) to reach the same contradiction. ∎

Lemma 4. For all $S \in \mathcal{W}$, $\succeq_S$ is well-defined, complete, transitive and reflexive.

Proof. From Axiom 1, Lemma 3 and the definition. ∎

Lemma 5. Let $S \in \mathcal{W}$ and let $S_i \in \mathcal{W}$, $i = 1, \ldots, n$, be a partition of $\mathcal{W}$. If $f \succeq_{S_i} g$ for all i, then $f \succeq_S g$. If there exists an i such that $f \succ_{S_i} g$, then $f \succ_S g$.

Proof. By induction on n. For $n = 1$ the conclusion is trivial. For $n = 2$, define

$$f' = \begin{cases} f & \text{on } S_1, \\ g & \text{on } S_2, \\ g & \text{on } S^c, \end{cases} \qquad f'' = \begin{cases} f & \text{on } S_1, \\ f & \text{on } S_2, \\ g & \text{on } S^c. \end{cases}$$

Then $f \succeq_{S_1} g$ iff $f' \succeq g$. Also, $f \succeq_{S_2} g$ iff $f'' \succeq f'$. By transitivity, $f'' \succeq g$, so $f \succeq_S g$.

For arbitrary n, let $C = \cup_{i=1}^{n-1} S_i$. By the induction hypothesis, $f \succeq_{S_i} g$ for $i = 1, \ldots, n-1$ implies $f \succeq_C g$. Since C and S_n are a two-set partition of S, $f \succeq_S g$. ∎

Lemma 6. Given Axioms 1–3, there exists a utility function $u : X \times \Omega \mapsto \mathbf{R}$ such that for all $\omega \in \Omega$, $f \succeq_{\{\omega\}} g$ iff

$$U(f)(\omega) = \sum_{x \in X} u(x,\omega) f(x)(\omega) \geq \sum_{x \in X} u(x,\omega) g(x)(\omega) = U(g)(\omega).$$

The function u can be chosen such that, for all states ω for which the conclusion of Axiom 5 holds, $\max_x u(x,\omega) = 1$, $\min_x u(x,\omega) = 0$. If Axiom 4 holds as well, then for all $x \in X$, and all $\omega, \omega' \in \Omega$, $u(x,\omega) = u(x,\omega')$.

Proof. All of this follows from conventional arguments using Lemma 1 and the Archimedian assumption to construct the utility indicator for each state. ∎

Notice that if the vector $U(f) \in \mathbf{R}^{|\Omega|}$ is at least as big as $U(g)$ in the usual vector ordering, then $f \succeq g$. This follows from Lemma 5 since the sets $\{\omega\}_{\omega \in \Omega}$ forms a finite partition of Ω.

Let $v_1, \ldots, v_k$ and $w_1, \ldots, w_l$ be vectors in $\mathbf{R}^n$, and consider the following two propositions:

a) For some $x \geq \mathbf{0}$ in $\mathbf{R}^n$,

$$v_i \cdot x > 0 \quad \text{for all } i,$$
$$w_j \cdot x = 0 \quad \text{for all } j,$$

b) There exists $y \in \mathbf{R}^k_+$, $z \in \mathbf{R}^l$ such that

$$\sum_1^k y_i v_i + \sum_1^l z_j w_j \leq \mathbf{0}.$$

If $y = 0$, then $\sum_1^l z_j w_j \neq \mathbf{0}$.

Lemma 7. If *a*) is false, then *b*) is true.

Proof. I will prove that the failure of a) implies b), using a separation theorem. Equation system is satisfied if and only if

a') For some $x \geq \mathbf{0}$,

$$v_i \cdot x \geq 1 \quad \text{for all } i,$$
$$w_j \cdot x = 0 \quad \text{for all } j.$$

Any solution to a') also solves a). Any solution to a) can be multiplied by a positive scalar to give a solution to a').

Define the non-empty, closed convex sets

$$C = \{(v_1 \cdot x, \ldots, v_k \cdot x, w_1 \cdot x, \ldots, w_l \cdot x) \ : \ x \geq \mathbf{0}\},$$
$$D = \{(s_1, \ldots, s_k, 0, \ldots, 0) \ : \ s_1, \ldots, s_k \geq 1\}.$$

The set C is a cone. If a) has no solution, then a') has no solution and $C \cap D = \emptyset$. Thus C and D can be properly separated by a hyperplane through the origin (Rockafellar, Theorems 11.3 and 11.7). Thus there exists $p = (y, z) \in \mathbf{R}^{\mathbf{k+l}}$ such that

$$\inf_D p \cdot d \geq \sup_C p \cdot c = 0 \qquad (i),$$
$$\sup_D p \cdot d > \inf_C p \cdot c \qquad (ii).$$

(Proper separation means that both sets are not contained in the hyperplane.)

First, $\sum_1^k y_i v_i + \sum_1^l z_j w_j \leq \mathbf{0}$. This follows from $\sup_C p \cdot c = 0$ upon substituting the standard ordered basis in $\mathbf{R}^{\mathbf{n}}$ for x.

Second, $y_i \geq 0$. If not, $\inf_D p \cdot d = -\infty$, which contradicts (i).

Third, suppose that $y = \mathbf{0}$. Then $p \cdot d = 0$ for all $d \in D$, and $p \cdot c = \sum_1^l z_j w_j \cdot x$. Thus (ii) implies that $\sum_1^l z_j w_j \cdot x < 0$ for some $x \geq \mathbf{0}$. ∎

Lemma 8. Let l be any finite set of lotteries in L such that $f \succ g$ for some $f, g \in l$. Then there exists a probability measure ρ_l on Ω such that, if $f, g \in l$, $f \succeq g$ iff $\sum_{\omega \in \Omega} U(f)(\omega)\rho_l(\omega) \geq \sum_{\omega \in \Omega} U(g)(\omega)\rho_l(\omega)$.

Proof. Let v_i, $i = 1, \ldots, k$ be the vectors $v_i = U(f_i) - U(g_i)$ for all $f_i, g_i \in L'$ such that $f_i \succ g_i$. Let w_j, $j = 1, \ldots, l$ be the vectors $w_j = U(f_j) - U(g_j)$ for each pair $f_j, g_j \in l$ such that $f_j \sim g_j$. I need only prove that there is an x solving a) of Lemma 7. If so, since x is not $\mathbf{0}$

and non-negative, it can be normalized to a probability measure ρ_l, and by construction, it has the desired representation property.

To see that there is such a x, suppose not. Then system b) of Lemma 7 has a solution. Define $z_i' = |z|_i$, and $w_j' = \text{sgn } z_i w_i$. Then $(y, z') \geq \mathbf{0}$ but $(\neq \mathbf{0})$ and $\sum_1^k y_i v_i + \sum_1^l z_j' w_j' \leq \mathbf{0}$.. W.l.o.g. we can assume that $\sum_i y_i + \sum_j z_j' = 1$. Each v_i and w_j' is the difference of two vectors $U(f_j) - U(g_j)$, where $f_j \sim g_j$. Then for all $\omega \in \Omega$,

$$\sum_i y_i U(f_i)(\omega) + \sum_j z_j' U(f_j)(\omega) \leq \sum_i y_i U(g_i)(\omega) + \sum_j z_j' U(g_j)(\omega).$$

By linearity,

$$U(\sum_i y_i f_i + \sum_j z_j f_j)(\omega) \leq U(\sum_i y_i g_i + \sum_j z_j g_j)(\omega).$$

This implies, by Lemma 5, that $\sum_i y_i g_i + \sum_j z_j g_j \succeq \sum_i y_i f_i + \sum_j z_j f_j$. But $f_i \succ g_i$ and $f_j \sim g_j$, so by repeated application of Lemma 2, if any of the $y_i > 0$, $\sum_i y_i f_i + \sum_j z_j f_j \succ \sum_i y_i g_i + \sum_j z_j g_j$, a contradiction. If all the $y_i = 0$, then we can replace $\leq$ with $<$ in the preceding inequality. Thus, again by Lemma 5, $\sum_i y_i g_i + \sum_j z_j g_j \succ \sum_i y_i f_i + \sum_j z_j f_j$. By the same reasoning as before we get a contradiction, since repeated application of Lemma 2 yields $\sum_i y_i g_i + \sum_j z_j g_j \sim \sum_i y_i f_i + \sum_j z_j f_j$. ∎

Now I will construct a non-Archimedian ordered field representation of beliefs over Ω. Our field will be a reduced ultrapower of the reals $\mathbf{R}$.

The obvious thing to do at this point is to apply the compactness theorem from model theory. But in the next section I will use ultrafilter constructions so that I can explicitly relate non-Archimedian probabilities to sequences of Archimedian probabilities. For the sake of consistency, then, I will undertake an ultrafilter construction here.

Let $\mathcal{L}$ denote the collection of all finite sets of lotteries in L containing at least one strictly ordered pair. Let $L_f \in \mathcal{L}$ denote the set of all finite subsets containing f. For any finite set of acts $f_1, \ldots, f_n$, the intersection $\cap_{i=1}^n L_{f_i}$ is non-empty, since the set $\{f_1, \ldots, f_n\}$ is contained in each L_{f_i}. It is an easy exercise to show that there exists a free filter on $\mathcal{L}$ containing all the sets L_f for $f \in L$. It can be shown as a consequence of Zorn's lemma that there exists an ultrafilter $\mathcal{U}$ on $\mathcal{L}$ containing the filter, and hence all the sets L_f for $f \in L$. This ultrafilter must be free. Let $\mathcal{F}$ denote the set of all functions from $\mathcal{L}$ into $\mathbf{R}$. Two functions $r, s \in \mathcal{F}$ are equivalent if $\{l \in \mathcal{L} : r(l) = s(l)\}$ is a member of the ultrafilter $\mathcal{U}$. It is a standard exercise to show that this relationship is in fact an equivalence relationship, and that the set $\mathcal{R}$ of equivalence classes of functions is an ordered field, and an extension of the reals. Addition and multiplication of two elements are defined by pointwise addition and multiplication, respectively, on any two representatives of the two equivalence classes. The order relations are derived from applying the appropriate order relationship on $\mathbf{R}$ pointwise. Let r and s be two elements of $\mathcal{R}$, with representative functions r' and s'. Consider the complementary sets $\{l \in \mathcal{L} : r'(l) \geq s'(l)\}$ and $\{l \in \mathcal{L} : r'(l) < s'(l)\}$. One and only one of these sets is in $\mathcal{U}$ (since any ultrafilter contains either a set or its complement, but not both). If the first one is, then $r \geq s$; else $r < s$. The absolute value of $r \in \mathcal{R}$ is defined to be the equivalence class containing the function $|r'|$, where r' is a function in the equivalence class r.

The ordered field $\mathcal{R}$ has a subfield isomorphic to $\mathbf{R}$, namely the field of equivalence classes containing constant functions. The field $\mathcal{R}$ is a proper extension of $\mathbf{R}$. Let $r(l) = \#l$. The function r assigns to each finite set of lotteries the number of lotteries in the set. This function is not equivalent to any constant function. For suppose it were equivalent to the function $k(l) = k$ for all $l \in \mathcal{L}$. First, since $r(l)$ is integer-valued, k must be an integer if it exists. Then the set K of all sets of lotteries of cardinality k must be in $\mathcal{U}$. Take any distinct $k+1$ sets L_f and intersect them. This set is a finite intersection of sets in $\mathcal{U}$, and thus is in $\mathcal{U}$. However each element of this set is a set of lotteries containing at least $k+1$ elements. This set is thus contained in K^c, and so K^c is also in $\mathcal{U}$, which is a contradiction. It is not hard to show that any proper extension of $\mathbf{R}$ contains infinite elements, elements greater in absolute value than any element in $\mathbf{R}$; and infinitessimal elements, elements less in absolute value than any element of $\mathbf{R}$. Finally, for every $s \in \mathcal{R}$ there is a unique r in the subfield isomorphic to $\mathbf{R}$ such that $|s - r|$ is infinitessimal. We then say that $s \approx r$, and write $r = {}^\circ s$.

Now for the construction. Define $\hat{\rho}(\omega)$ to be the equivalence class that contains the function assigning to each $l \in \mathcal{L}$ the number $\rho_l(\omega)$ constructed in Lemma 8. Since for all l, $\rho_l(\omega) \geq 0$, $\hat{\rho}(\omega) \geq 0$. Since $\sum_\omega \rho_l(\omega) = 1$ for all $l \in \mathcal{L}$, $\sum_\omega \hat{\rho}(\omega) = 1$. Thus $\hat{\rho}$ is an $\mathcal{R}$-valued probability distribution on Ω.

Finally, we need to show that $f \succeq g$ iff $\sum_\omega (U(f)(\omega) - U(g)(\omega))\hat{\rho}(\omega) \geq 0$. The set $A = L_f \cap L_g$ is in the ultrafilter $\mathcal{U}$. For any $l \in A$, both f and g are in l, and so $f \succeq g$ iff $\sum_\omega (U(f)(\omega) - U(g)(\omega))\rho_l(\omega) \geq 0$. This completes the proof of the sufficiency of the axioms for the representation. Necessity is straightforward to check, and therefore left to the reader. ∎

Proof of Theorem 1. Given Theorem 1′, it suffices to show that if $\hat{\rho}$ is an $\mathcal{R}$-valued probability distribution on Ω, then there exists $\rho_1, \ldots, \rho_K$, $K \leq |\Omega|$, such that

$$\sum_\omega U(f)(\omega)\hat{\rho}(\omega) \geq \sum_\omega U(g)(\omega)\hat{\rho}(\omega)$$

if and only if

$$\left(\sum_{\omega} U(f)(\omega)\rho_k(\omega)\right)_{k=1}^{K} \geq_L \left(\sum_{\omega} U(g)(\omega)\rho_k(\omega)\right)_{k=1}^{K};$$

and conversly. It is clear that we need only show that we can write

$$\hat{\rho}(\omega) = \alpha(\rho_1 + \delta_2\rho_2 + \delta_2\delta_3\rho_3 + \cdots + \delta_2\cdots\delta_K\rho_K),$$

where each ρ_k is $\mathbf{R}$-valued, $\alpha > 0$ and each δ_k is positive and infinitessimal. Notice that the ratio of any coefficient of a ρ_k to the coefficient of any predecessor ρ_l is infinitessimal.

Lemma 9. There exists a $\beta \in \mathcal{R}$, $\beta \approx 1$ and ρ', ν mapping Ω into $\mathcal{R}$ such that for all $\omega \in \Omega$,

$$\begin{aligned}
\beta\hat{\rho}(\omega) &= \rho'(\omega) + \nu(\omega), \\
{}^\circ\rho'(\omega) &= \rho'(\omega), \\
\sum_{\omega} \rho'(\omega) &= 1, \\
\nu(\omega) &\geq 0, \\
\nu(\omega) &\approx 0, \\
\nu(\omega') &= 0 \quad \text{for some } \omega'.
\end{aligned}$$

Proof. Let ω' be a minimizer of $\hat{\rho}(\omega) - {}^\circ\hat{\rho}(\omega)$, and let $\beta = \hat{\rho}(\omega')/{}^\circ\hat{\rho}(\omega')$. Then $\beta \approx 1$. Furthermore, since $\sum_\omega \hat{\rho}(\omega) = 1$, there must exist an ω' such that $\hat{\rho}(\omega') - {}^\circ\hat{\rho}(\omega') \leq 0$. Thus $\beta \geq 1$. Now computing,

$$\begin{aligned}
\beta\hat{\rho}(\omega) &= {}^\circ(\beta\hat{\rho}(\omega) + \beta\hat{\rho}(\omega) - {}^\circ(\beta\hat{\rho}(\omega)) \\
&\geq {}^\circ(\beta\hat{\rho}(\omega)) + \hat{\rho}(\omega) - {}^\circ\hat{\rho}(\omega) \\
&\geq {}^\circ(\beta\hat{\rho}(\omega)) + \hat{\rho}(\omega') - {}^\circ\hat{\rho}(\omega') \\
&\geq {}^\circ(\beta\hat{\rho}(\omega)).
\end{aligned}$$

Furthermore, equality will hold when $\omega = \omega'$. Taking $\rho'(\omega) = {}^\circ(\beta\hat{\rho}(\omega))$ and $\nu(\omega) = \beta\hat{\rho}(\omega - {}^\circ(\beta\hat{\rho}(\omega))$ proves the lemma. ∎

To complete the proof, write $\hat{\rho} = (1/\beta)(\rho' + \nu)$. Let $\rho_1 = \rho'$, and let $\delta_2 = \sum_\omega \nu(\omega)$. Define $\hat{\rho}_2 = (1/\delta_2)\nu$. Then

$$\hat{\rho} = \frac{1}{\beta}\left(\rho_1 + \delta_2\hat{\rho}_2\right).$$

The function $\hat{\rho}_2$ is a $\mathcal{R}$-valued probability distribution on Ω, and its support is of cardinality less than that of the support of $\hat{\rho}$. Repeat this argument again on $\hat{\rho}_2, \hat{\rho}_3$, etc. so long at the remainder term ν is not 0. Since at each iteration at least one more 0 is added to the remainder term, the process must stop in no more than K steps. This proves sufficiency of the axioms.

To prove necessity, construct the $\mathcal{R}$-valued representation from the lexicographic representation. Choose a positive infinitessimal ϵ, and let

$$\hat{\rho}(\omega) = \left(\sum_0^K \frac{1}{\epsilon^k}\right) \sum_0^K \epsilon^k \rho_k(\omega).$$

It is easy to see that expected utility computed with respect to $\hat{\rho}$ will represent the same ordering as that represented by the lexicographic ordering on the expected utility vectors. Then apply Theorem 1$'$. ∎

Appendix 2: Independence and Non-Archimedian Beliefs.

1. Weak Independence.

Let A and B be events in $\mathcal{W}$. The independence of these two events can be expressed in terms of properties of the preferences on acts. However, independence is more complicated with non-Archimedian beliefs than it is in the Archimedian domain.

Definition. A **bet on** A is an act f such that

$$f(\omega) = \begin{cases} \mu & \text{if } \omega \in A, \\ \nu & \text{if } \omega \in A^c, \end{cases}$$

where μ and ν are two probability distributions in $P(X)$.

Definition. The event A in $\mathcal{W}$ is **weakly independent** of event B in $\mathcal{W}$ if B is either certain, impossible, or, for any two bets f and g on A,

$$f \succeq_B g \quad \text{iff} \quad f \succeq_{B^c} g.$$

Another way to view the definition is as follows. Let:

$$\begin{aligned} \theta_A &= \sum_x u(x)f(x)(\omega) && \omega \in A, \\ \theta_{A^c} &= \sum_x u(x)f(x)(\omega) && \omega \in A^c, \\ \gamma_A &= \sum_x u(x)g(x)(\omega) && \omega \in A, \\ \gamma_{A^c} &= \sum_x u(x)g(x)(\omega) && \omega \in A^c. \end{aligned}$$

Let ρ denote a non-Archimedian representation ($\mathcal{R}$-valued representation) of beliefs. Then A and B are weakly independent iff B is certain, impossible, or

$$\theta_A\rho(A \cap B) + \theta_{A^c}\rho(A^c \cap B) \geq \gamma_A\rho(A \cap B) + \gamma_{A^c}\rho(A^c \cap B)$$

iff

$$\theta_A\rho(A \cap B^c) + \theta_{A^c}\rho(A^c \cap B^c) \geq \gamma_A\rho(A \cap B^c) + \gamma_{A^c}\rho(A^c \cap B^c),$$

for all $0 \leq \theta_A, \theta_{A^c}, \gamma_A, \gamma_{A^c} \leq 1$. This definition is easily seen to be equivalent to the usual multiplicative property $\rho(A \cap B) = \rho(A)\rho(B)$ when ρ is Archimedian.

The next Theorem characterizes weak independence in terms of the non-Archimedian representation ρ. If $\rho(B) > 0$, assume w.l.o.g. that $\rho(A^c \cap B) > 0$.

Theorem 4. A is weakly independent of B iff B is certain, impossible, or the interval in $\mathcal{R}$ whose endpoints are

$$\frac{\rho(A \cap B)}{\rho(A^c \cap B)} \quad \text{and} \quad \frac{\rho(A \cap B^c)}{\rho(A^c \cap B^c)}$$

contains no element of $\mathbf{R}$ unless it is degenerate.

Proof. The certain and impossible cases are statements from the definition. Suppose, then that B is neither certain nor impossible.

(If) Suppose the inequalities are not satisfied. In particular, suppose that the first inequality is true but the second is false. Since $\rho(A^c \cap B) > 0$, the truth of the first inequality implies that

$$(\theta_A - \gamma_A)\frac{\rho(A \cap B)}{\rho(A^c \cap B)} + \theta_{A^c} \geq \gamma_{A^c} - \theta_{A^c}.$$

If $\rho(A^c \cap B^c) = 0$ then one endpoint is finite while the other is infinite and so the interval must contain a standard number. If $\rho(A^c \cap B^c) > 0$, then, since the second inequality fails,

$$\gamma_{A^c} - \theta_{A^c} > (\theta_A - \gamma_A)\frac{\rho(A \cap B^c)}{\rho(A^c \cap B^c)}.$$

Putting these two inequalities together,

$$(\theta_A - \gamma_A)\frac{\rho(A \cap B)}{\rho(A^c \cap B)} \geq \gamma_{A^c} - \theta_{A^c} > (\theta_A - \gamma_A)\frac{\rho(A \cap B^c)}{\rho(A^c \cap B^c)}.$$

First, this implies that the standard number $(\theta_A - \gamma_A)$ is not 0. Dividing,

$$\frac{\rho(A \cap B)}{\rho(A^c \cap B)} \geq \frac{\gamma_{A^c} - \theta_{A^c}}{\theta_A - \gamma_A} > \frac{\rho(A \cap B^c)}{\rho(A^c \cap B^c)}.$$

The middle number is standard.

(Only if) Suppose that weak independence is satisfied, and set $\theta_{A^c} = \gamma_A = 0$. The weak independence condition becomes

$$\theta_A \rho(A \cap B) \geq \gamma_{A^c} \rho(A^c \cap B)$$

if and only if

$$\theta_A \rho(A \cap B^c) \geq \gamma_{A^c} \rho(A^c \cap B^c).$$

If $\rho(A^c \cap B^c)$ is 0 or infinitessimal, then $\rho(A^c \cap B)$ must be 0 infinitessimal or it is possible to choose standard parameters θ_A and γ_{A^c} to violate the condition. The converse is also true. If $\rho(A^c \cap B)$ is not infinitessimal, then $\rho(A^c \cap B^c)$ is finite and not infinitessimal, and so

$$\frac{\rho(A \cap B)}{\rho(A^c \cap B)} \geq \frac{\gamma_{A^c}}{\theta_A} \quad \text{iff} \quad \frac{\rho(A \cap B^c)}{\rho(A^c \cap B^c)} \geq \frac{\gamma_{A^c}}{\theta_A}$$

for all $0 < \gamma_{A^c}, \theta_A \leq 1$.

This implies that the interval cannot be non-degenerate and contain a standard number. ∎

The following Theorem reinterprets the necessary and sufficient condition of Theorem 4 in terms of a factor representation of lexicographic beliefs.

Theorem 5. Let $\rho_1, \ldots, \rho_K$ be a factor representation of beliefs. Then A is weakly independent of B iff B is certain, impossible, or (assuming, w.l.o.g. that $\rho_1(A^c \cap B) > 0$),

$$\frac{\rho_1(A \cap B)}{\rho_1(A^c \cap B)} = \frac{\rho_1(A \cap B^c)}{\rho_1(A^c \cap B^c)} \equiv r$$

and, if i is the first index for which one of the inequalities

$$\frac{\rho_i(A \cap B)}{\rho_i(A^c \cap B)} = \frac{\rho_i(A \cap B^c)}{\rho_i(A^c \cap B^c)} = r$$

fails, then

$$\operatorname{sgn}\left(\frac{\rho_i(A \cap B)}{\rho_i(A^c \cap B)} - r\right) = \operatorname{sgn}\left(\frac{\rho_i(A \cap B^c)}{\rho_i(A^c \cap B^c)} - r\right).$$

Proof. It suffices to consider the case where B is neither certain nor impossible. Given the factor representation $(\rho_1, \ldots, \rho_k)$, any $\mathcal{R}$-valued probability distribution $(1/\sum_{k=1}^K \epsilon_k)\sum_{k=1}^K \epsilon_k \rho_k$, where $\epsilon_1 = 1$ and $\epsilon_{k+1}/\epsilon_k$ is infinitessimal,represents beliefs. Accordingly, the conditions of Theorem 4 must be satisfied for all values of the ϵ_k satisfying the aforementioned properties. The interval whose endpoints are

$$\frac{\sum_{k=1}^K \epsilon_k \rho_k(A \cap B)}{\sum_{k=1}^K \epsilon_k \rho_k(A^c \cap B)} \quad \text{and} \quad \frac{\sum_{k=1}^K \epsilon_k \rho_k(A \cap B^c)}{\sum_{k=1}^K \epsilon_k \rho_k(A^c \cap B^c)}$$

must contain no standard number unless the two endpoints are equal. To simplify calculations define

$$r_k = \frac{\rho_k(A \cap B)}{\rho_k(A^c \cap B)}, \quad s_k = \frac{\rho_k(A \cap B^c)}{\rho_k(A^c \cap B^c)}.$$

Then the endpoints can be written as

$$e = \frac{\sum_{k=1}^K \epsilon_k r_k \rho_k(A^c \cap B)}{\sum_{k=1}^K \epsilon_k \rho_k(A^c \cap B)}, \quad \text{and} \quad f = \frac{\sum_{k=1}^K \epsilon_k s_k \rho_k(A^c \cap B^c)}{\sum_{k=1}^K \epsilon_k \rho_k(A^c \cap B^c)},$$

respectively.

The conditions of Theorem 4 are true iff $e \approx f$. Note that ${}^\circ e = r_1$, and ${}^\circ f = s_1$ so $r_1 = s_1$. Now look at the remainders $e - r_1$ and $f - s_1$. A necessary and sufficient condition for the condition of Theorem 4 to hold is that $r_1 = s_1$ and $\operatorname{sgn}(e - r_1) = \operatorname{sgn}(f - s_1)$. Computing,

$$\operatorname{sgn}(e - r_1) = \operatorname{sgn}\left(\sum_{k=2}^K \epsilon_k (r_k - r_1)\rho_k(A^c \cap B)\right)$$

$$\operatorname{sgn}(f - s_1) = \operatorname{sgn}\left(\sum_{k=2}^K \epsilon_k (s_k - s_1)\rho_k(A^c \cap B^c)\right).$$

The equality of the signs of the differences is easily seen to be equivalent to the stated conditions in the Theorem by noting that each term in each sum is infinitely larger than its successors. ∎

In game theory, independence is a property that applies to n-tuples of strategies — to random variables. The definition of weak independence is easily defined for random variables.

I will consider the case of two random variables; finite numbers of random variables can be accounted for by induction. Denote the two random variables by $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$, which can take on I and J different values, respectively. Then $\Omega = \{1,\ldots,I\} \times \{1,\ldots,J\}$. Let i denote the event $\{\tilde{\mathbf{a}} = i\}$, and j the event $\{\tilde{\mathbf{b}} = j\}$. Similarly for i' and j'. Let G denote the set of mixtures of bets on $\tilde{\mathbf{a}}$. An the outcome of an act in G depends only on the value of $\tilde{\mathbf{a}}$, and not on $\tilde{\mathbf{b}}$.

Definition. The random variable $\tilde{\mathbf{a}}$ is **weakly independent** of the random variable $\tilde{\mathbf{b}}$ iff the preference orderings $\succeq_j$ are independent of j on G.

This assumption says that knowledge of the value of $\tilde{\mathbf{b}}$ does not affect preferences over bets (mixtures of bets) on the value of $\tilde{\mathbf{a}}$. It is easy to check that the random variables $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$ are weakly independent if and only if the events $\{\tilde{\mathbf{a}} = i\}$ and $\{\tilde{\mathbf{b}} = j\}$ are weakly independent for all i and j.

When the representation of beliefs ρ is Archimedian, weak independence is equivalent to the condition $\rho(A \cap B) = \rho(A)\rho(B)$. When the representation is non-Archimedian, weak independence implies that this multiplicative condition must hold approximately — that the difference between the right and left hand sides must be infinitessimal.[8] It is also easy to see that if there exists a non-Archimedian representation of beliefs for which the multiplicative condition holds exactly, then the event A is weakly independent of B. However, it is not always true that there exists a representation of beliefs which will satisfy the multiplicative condition for weakly independent events. Consider the following example. A gambler is to bet on the flip of two coins. Three outcomes are conceivable for each coin: The coin can come up heads H, the coin can come up tails T, or the coin can land on edge E. The state space Ω consists of the nine possible combinations of H, T and E for each of the two coins. Define the following acts, where payoffs are in terms of utility:

$$f_{xy} = \begin{cases} 1 & \text{if } \omega = xy, \\ 0 & \text{if otherwise,} \end{cases} g_r = \begin{cases} r & \text{if } \omega = HH, \\ 0 & \text{if otherwise,} \end{cases} h_s = \begin{cases} s & \text{if } \omega = HE, \\ 0 & \text{if otherwise,} \end{cases}.$$

Let $\mathbf{0}$ denote the act that pays off 0 in each state. Suppose the gambler's preferences include the

[8] Non-standard probability theory makes use of a concept called S-independence, which requires that $\rho(A \cap B) - \rho(A)\rho(B)$ is infinitessimal. Weak independence is stronger than S-independence. The concept of independence which I define below is identical with the notion of *-independence from non-standard probability.

following specifications:

$$f_{HH} \sim f_{HT} \sim f_{TH} \sim f_{TT},$$
$$f_{HE} \sim f_{TE} \sim f_{EE} \sim f_{ET} \sim f_{EH},$$
$$g_r \succ h_s \quad \text{for all } 0 < r, s \geq 1,$$
$$h_s \succ \mathbf{0} \quad \text{for all } 0 < s < 1.$$

These preferences are consistent with lexicographic expected utility maximization, and it is easy to see that any non-Archimedean representation of beliefs must look as follows:

$$\begin{array}{cccc} & H & T & E \\ H & \frac{1-5\epsilon}{4} & \frac{1-5\epsilon}{4} & \epsilon \\ T & \frac{1-5\epsilon}{4} & \frac{1-5\epsilon}{4} & \epsilon \\ E & \epsilon & \epsilon & \epsilon \end{array}$$

where the rows index the flip of coin 1, the columns index the flip of coin 2, and ϵ is a positive infinitessimal. Simple computations now show that the events {coin 1=H} and {coin 2=H} are weakly independent:

$$1 > \frac{\rho\{HH\}}{\rho\{H^cH\}} = \frac{1-5\epsilon}{1-\epsilon} > \frac{1-\epsilon}{1+7\epsilon} = \frac{\rho\{HH^c\}}{\rho\{H^cH^c\}} \approx 1$$

in every representation. The multiplicative condition implies that these ratios are equal, so in fact the multiplicative condition is violated in every representation. From direct calculation it is clear that

$$\rho\{HH\} = \frac{1-5\epsilon}{4} > \frac{1-6\epsilon+9\epsilon^2}{4} = \rho\{H\}\rho\{H\}.$$

In the previous example the events {coin 1=H} and {coin 2=H} are weakly independent, but the random variables **coin 1** and **coin 2** are not. The following example, essentially due to Roger Myerson[9] shows that the independence of conditional preferences (weak independence of random variables) is not enough to guarantee a multiplicative representation. There are two random variables, each with possible values A, B and C. A non-Archimedean representation of beliefs is given:

$$\begin{array}{cccc} & A & B & C \\ A & \frac{1}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} & \frac{\epsilon}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} & \frac{\epsilon^2}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} \\ B & \frac{2\epsilon}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} & \frac{\epsilon^2}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} & \frac{\epsilon^3}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} \\ C & \frac{\epsilon^2}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} & \frac{\epsilon^3}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} & \frac{\epsilon^4}{1+3\epsilon+3\epsilon^2+2\epsilon^3+\epsilon^4} \end{array}$$

9 Personal communication.

where, for example, cell BA contains the joint probability that random variable 1 takes on the value B and random variable 2 takes on the value A. It is easy to see that this particular representation does not satisfy the product rule for independent random variables. Furthermore, no multiplicative representation can generate the same preferences as does this probability representation. To see this last fact, observe that the following four acts are indifferent to the act that pays off 0 in each state:

$$\begin{pmatrix} 0 & 2 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

where the number in each cell corresponds to a utility payoff in the corresponding state. One can verify that no product representation will make these four acts indifferent to the act paying off 0 in each state.

A further perversity of weak independence is that the weak independence relationship is not symmetric. Consider the following example (matrix entries are again non-Archimedian probabilities):

$$\begin{array}{cccc} & A & B & C \\ A & \dfrac{1}{3+2\epsilon+2\epsilon^2} & \dfrac{\epsilon}{3+2\epsilon+2\epsilon^2} & \dfrac{\epsilon^2}{3+2\epsilon+2\epsilon^2} \\ B & \dfrac{2}{3+2\epsilon+2\epsilon^2} & \dfrac{\epsilon}{3+2\epsilon+2\epsilon^2} & \dfrac{\epsilon^2}{3+2\epsilon+2\epsilon^2} \end{array}$$

Observe that preferences given top and preferences given bottom are identical, but that preferences given left are different from preferences given center and preferences given right. Random variable $\tilde{\mathbf{b}}$ (columns) is weakly independent of random variable $\tilde{\mathbf{a}}$ (rows) but not vice versa.

2. Independence.

The weakly independent strategies that arise in the lexicographic characterization of perfect and proper equilibrium all satisfy the multiplicative condition. In this section I characterize this stronger condition on representations in terms of the decisionmaker's preference order. In this section I will consider the case of two random variables. The generalization to any finite number of random variables is straightforward but tedious.

The state space will again be the cartesian product $I \times J$, where $I = \{1, \dots, I\}$ is the range of the random variable $\tilde{\mathbf{a}}$ and $J = \{1, \dots, J\}$ is the range of the random variable $\tilde{\mathbf{b}}$. Let Δ^I be the unit simplex in $\mathbf{R}^{\mathbf{I}}_{+}$, and Δ^J is the unit simplex in $\mathbf{R}^{\mathbf{J}}_{+}$. I seek conditions under which there exists a representation of beliefs for which $\rho(\tilde{\mathbf{a}} = i, \tilde{\mathbf{b}} = j) = \rho(\tilde{\mathbf{a}} = i)\rho(\tilde{\mathbf{b}} = j)$. Beliefs for which this is true will be said to satisfy the **product rule.**

Associated with each act is a utility vector which specifies the utility payoff in each state of nature. The preference ordering on acts induces a preference ordering on utility vectors. Two acts with identical utility vectors are indifferent — in fact, conditionally indifferent in every state. The preference ordering on utility vectors inherits all of the important properties of the preference ordering on acts. One way of viewing the representation question posed by the product rule is to note that it asks for a bilinear representation of preferences over the vectors of utilities. The Independence Axiom (Axiom 3) is an invariance property. It requires that preferences be invariant to objective mixtures of acts. The existence of a bilinear representation requires additional invariance properties with respect to a different kind of mixture.

Definition. A **conditional act** on J is a function $f : J \mapsto P(X)$.

Definition. A **subjective mixture** over I of the act $f : I \times J \mapsto P(X)$ is an act $g : I \times J \mapsto P(X)$ such that each conditional act $g(i, \cdot)$ on J is a mixture of the conditional acts $f(k, \cdot)$, $k \in I$.

Let $[f]$ denote the $I \times J$ matrix whose (i,j)'th element is the conditional utility payoff of the act given that state (i,j) occurs. The following Lemma can be verified immediatgely from the definitions:

Lemma 10. An act g is a **subjective mixture over** I of f if and only if there is a Markov matrix A such that $[g] = A[f]$.

In this case $g(i, \cdot)$ is just the mixture whose coefficients are described by the i'th row of the matrix A. I will denote this mixture as $A \otimes f$.

Subjective mixing is linear with respect to objective mixing. Thus the Independence Axiom applies. If $A \otimes f \succeq A \otimes g$, then for all $0 \leq \lambda \leq 1$, $A \otimes \big(\lambda f + (1-\lambda)h\big) \succeq A \otimes \big(\lambda g + (1-\lambda)h\big)$, with strict preference if $f \succ g$ and $\lambda > 0$.

A special class of mixtures is important for characterizing independence. Recall that a bet on J is an act f such that $[f]_{ij} = [f]_{kj}$ for all $i, k \in I$. The characterization of independence is concerned with those mixtures whose range is the set of bets on J. To always generate a bet on J, the subjective mixture must mix the acts $f(i, \cdot)$ in the same proportions for each i. This requires that the rows of A be identical. The subjective mixture $A \otimes f$ is a bet on J for all f if and only if A has rank 1. I will refer to the mixture as $\mathbf{p} \otimes f$, where $\mathbf{p} \in \Delta^I$ is the row vector of A.

The condition which characterizes independence is an invariance property for subjective mixtures onto bets:

Definition. The random variables $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$ are **independent** iff for all f_1, f_2 and g_1, g_2 in L with $f_k \succeq g_k$, there exists an $\mathbf{p} \in \Delta^I$ such that for all $0 \leq \lambda \leq 1$ there is an event $\{j\}$ such that $\mathbf{p} \otimes \lambda f_1 + (1-\lambda) f_2 \succeq_{\{j\}} \mathbf{p} \otimes \lambda g_1 + (1-\lambda) g_2$. If $f_k \succ g_k$ for some k and $\lambda_k > 0$, then substitute strict preferences for the relationship between the subjective mixtures.

Independence is a symmetric relationship. The proof of the representation theorem will make clear that the roles of I and J in the definition are interchangeable. It can also be seen that this condition implies weak independence, since this condition is equivalent to the existence of beliefs satisfying the product rule, and any such beliefs are weakly independent.

Theorem 6. The preference order $\succeq$ has an expected utility representation with beliefs ρ such that, for all i and j, $\rho(\tilde{\mathbf{a}} = i, \tilde{\mathbf{b}} = j) = \rho(\tilde{\mathbf{a}} = i)\rho(\tilde{\mathbf{b}} = j)$ if and only if I and J are independent.

Proof. By examining the usual ultrapower construction, it is clear that non-Archimedian beliefs satisfying the product rule can be found if and only if for every finite set of lotteries there exists an Archimedian distribution satisfying the product rule and representing preferences restricted to that set. Thus I need to show that the independence condition is equivalent to the existence for any finite set of binary comparisons of acts of a representation satisfying the product rule.

Let $f_1, \ldots, f_N$ and $g_1, \ldots, g_N$ be a finite collection of acts such that $f_n \succ g_n$ for $1 \leq n \leq M$ and $f_n \sim g_n$ for $M < n \leq N$. A representation of beliefs satisfying the product rule is given by any solution $\mathbf{p} \in \Delta^I$, $\mathbf{q} \in \Delta^J$ of the following system of inequalities:

$$\begin{aligned} \mathbf{p}^T([f_n] - [g_n])\mathbf{q} &> 0, & 1 \leq n \leq M, \\ \mathbf{p}^T([f_n] - [g_n])\mathbf{q} &\geq 0, & M < n \leq N, \\ \mathbf{p}^T([g_n] - [f_n])\mathbf{q} &\geq 0, & M < n \leq N. \end{aligned} \tag{1}$$

Let $\Delta^J_+ = \{\mathbf{q} \in \mathbf{R}^J_+ : \sum_j \mathbf{q}_j \geq 1\}$. System (1) has a solution if and only if there is an $\mathbf{p} \in \Delta^I$ and $\mathbf{q} \in \Delta^J_+$ such that

$$\begin{aligned} \mathbf{p}^T([f_n] - [g_n])\mathbf{q} &\geq 1, & 1 \leq n \leq M, \\ \mathbf{p}^T([f_n] - [g_n])\mathbf{q} &\geq 0, & M < n \leq N, \\ \mathbf{p}^T([g_n] - [f_n])\mathbf{q} &\geq 0, & M < n \leq N. \end{aligned} \tag{2}$$

Lemma 11. One and only 1 of the following alternatives holds:

i) Equation system (2) has a solution.

ii) For all $\mathbf{p} \in \Delta^I$ there is a $\lambda \in \Delta^N$ such that $\sup\{\sum_n \lambda_n \mathbf{p}^T([f_n] - [g_n])\mathbf{q} \ : \mathbf{q} \in \Delta^J_+\} \leq 0$. If $\lambda_n = 0$ for all $n \leq M$, then the supremum is negative.

Proof. Both alternatives cannot simultaneously be true. Suppose that (2) has no solution. I will show that *ii)* is true. Let

$$D = \{\mathbf{d} \in \mathbf{R}^\mathbf{N}_+ \ : \ \mathbf{d}_n \geq 1 \text{ for } i \leq M\}, \qquad \text{and}$$
$$C = \{\Big(\mathbf{p}^T([f_1] - [g_1])\mathbf{q}, \dots, \mathbf{p}^T([f_N] - [g_N])\mathbf{q}\Big) \ : \ \mathbf{q} \in \Delta^J_+\}.$$

The sets C and D are polyhedral convex sets. They are disjoint if (2) has no solution. Thus there is a $\lambda \in \mathbf{R}^\mathbf{N}$ such that $\inf_D \lambda \cdot \mathbf{d} > \sup_C \lambda \cdot \mathbf{c}$ (Rockafellar TK). Rescale any such λ to be an element of Δ^N. Each $\lambda_n \geq 0$, or else $\inf_D \lambda \cdot \mathbf{d} = -\infty$. It must be true that $\sup_C \lambda \cdot \mathbf{c} \leq 0$ because $\alpha\mathbf{c} \in C$ for all $\alpha \geq 1$ and $\mathbf{c} \in C$, so $\lambda \cdot \mathbf{c} > 0$ for some $\mathbf{c} \in C$ implies $\sup_C \lambda \cdot \mathbf{c} = +\infty$. Finally, if $\lambda_n = 0$ for $n \leq M$, then $\inf_D \lambda \cdot \mathbf{d} = 0$. In this case, $\sup_C \lambda \cdot \mathbf{c} < 0$. ∎

Choose $\mathbf{p} \in \Delta^I$. For act $\mathbf{p} \otimes f$, $[\mathbf{p} \otimes f]_{ij} = \mathbf{p}^T \cdot [f]_j$, where $[f]_j$ is the j'th column vector of $[f]$. From Lemma 11, if (2) has no solution, then for each $\mathbf{p} \in \Delta^I$ there exists $\lambda \in \Delta^N$ such that $[\mathbf{p} \otimes \sum_n \lambda_n f_n] \cdot \mathbf{q} \leq [\mathbf{p} \otimes \sum_n \lambda_n g_n] \cdot \mathbf{q}$ for all $\mathbf{q} \in \Delta^J$. Since all the rows within each matrix are equal, conclude that

$$[\mathbf{p} \otimes \sum_n \lambda_n f_n]_{ij} \leq [\mathbf{p} \otimes \sum_n \lambda_n g_n]_{ij} \tag{3}$$

for all i and j. Thus, for all j,

$$\mathbf{p} \otimes \sum_n \lambda_n g_n \succeq_j \mathbf{p} \otimes \sum_n \lambda_n f_n. \tag{4}$$

Since the acts being compared are mixtures of bets on j, and thus constant across i given j, (4) implies (3). If $\lambda_1 = \cdots = \lambda_M = 0$, then we have strict inequality in (3) and consequently strict preference in (4).

Suppose w.l.o.g. that $\lambda_k > 0$. Then let

$$f' = \frac{1}{1 - \lambda_k} \sum_{n \neq k} \lambda_n f_n, \qquad g' = \frac{1}{1 - \lambda_k} \sum_{n \neq k} \lambda_n g_n.$$

Then $f' \succeq g'$, and $\mathbf{p} \otimes \lambda_k g_k + (1 - \lambda_k) g' \succeq_j \mathbf{p} \otimes \lambda_k g_k + (1 - \lambda_k) g'$ for all j. Furthermore, if any $\lambda_n > 0$ for $n \leq M$, then either $f_k \succ g_k$ or $f' \succ g'$. This contradicts independence.

Conversely, suppose independence holds. Then in the same manner the Theorem of the alternative (Lemma 11) shows that for no finite set of binary choices can (4) hold for every $\mathbf{p}$. Thus every finites set of binary choices of lotteries admits an Archimedian product representation of beliefs and so a non-Archimedian representation of beliefs satisfying the product rule exists for $\succeq$ on L. ∎

Appendix 3: Proofs for Section 5.

In this Appendix I prove Theorems 3, 4, 1and 3. I will use an ultrapower construction to prove Theorems 3 and 4. The technique is very powerful, and can provide lexicographic characterizations and sequential characterizations for other refinements, such as strictly perfect and weakly proper equilibrium.

Let N denote the set of positive integers. The set of all co-finite subsets of N is a filter, called the *Fréchet filter*. Let $\mathcal{U}$ be an ultrafilter containing the Fréchet filter. This ultrafilter is free. It contains no finite subsets of N. Let $\mathcal{F}$ denote the set of all functions from N to $\mathbf{R}$. As in Appendix 1, two functions are equivalent if they agree on a set in the ultrafilter $\mathcal{U}$. The set $\mathcal{R} = \mathcal{F}/\sim$ is a proper ordered-field extension of $\mathbf{R}$. See Appendix 1 for more details.

Proof of Theorem 3. (Sufficiency) Let t denote the lexicographic probability distribution on joint strategies which generates the common, strictly independent beliefs.

Since beliefs are strictly independent and common, we can choose positive elements $\epsilon^1, \ldots, \epsilon^K$ in $\mathcal{R}$ with the following two properties: First, $\epsilon^1 = 1$, and $\epsilon^k/\epsilon^{k-1}$ is infinitessimal for all k. Let $\alpha = \sum_{k=1}^K \epsilon^k$, and let $\hat{t} = \alpha^{-1}\sum_{k=1}^K \epsilon^k t^k$. Then r_n, the beliefs of player n, are represented by the marginal distribution of $\hat{t}$ on the set I_{-n}. Second, $\hat{t}(i_1, \ldots, i_n) = \hat{t}(i_1)\cdots\hat{t}(i_n)$. Choose sequences $\{\epsilon_j^k\}_{j=1}^\infty$ from the equivalence classes ϵ^k, and define α_j as above. Since ϵ^k is infinitessimal for $k > 1$, $\lim_j \epsilon_j^k = 0$. The sequence $\{\alpha_j^{-1}({}^*t^1 + \epsilon_j^k {}^*t^2 + \cdots + \epsilon_j^K {}^*t^K)\}_{j=1}^\infty$ is a member of the equivalence class $\hat{t}$.

First, there exists a set $A \in \mathcal{U}$ such that ϵ_j^k is positive for all k and $j \in A$. Second,

$$\sum_{i \in I} {}^*u_n(i_n, i_{-n}) s_n(i_n)\hat{t}(i_{-n}) \geq \sum_{i \in I} {}^*u_n(i_n, i_{-n}) s'(i_n)\hat{t}(i_{-n})$$

for all $s' \in S_n$. This means that the set B of all j such that

$$\sum_{i \in I} u_n(i_n, i_{-n}) s_n(i_n) t_j(i_{-n}) \geq \sum_{i \in I} u_n(i_n, i_{-n}) s'(i_n) t_j(i_{-n})$$

for all $s' \in S_n$ is contained in $\mathcal{U}$. There exists a set C in $\mathcal{U}$ such that, for all $j \in C$, each t_j is a product distribution of mixed strategy vectors $s_1^j, \ldots s_n^j$ (by strict independence). The set $D = A \cap B \cap C$ is contained in $\mathcal{U}$. The sequence $\{s_1^j, \ldots s_n^j\}_{j \in D}$ has all of the desired properties. Each s_n is optimal against all of the $s_1^j, \ldots s_n^j$. Each s_n^j assigns positive probability to all $i_n \in I_n$. To see this, note that i_n is assigned positive probability by some t^k, and for all k and $j \in A$,

$c_j^k > 0$, so the particular t^k gets counted in the weighted average t_j. Having constructed the test sequence, this proves sufficiency.

(Necessity). Let $\{s^j\}_{j=1}^{\infty}$ denote a suitable test sequence, and let t_j denote the joint distribution on I constructed from the belief vector s^j. Define $\hat{t}(i)$ to be the equivalence class containing the function $t_j(i)$. Clearly the probability distribution $\hat{t}$ has the strategies of different players being independent. Since $t_j(i) > 0$ for all j, $\hat{t}(i) > {}^*0$. Thus the factor representation of each r_n will assign positive probability to each i_{-n}. (For more details on constructing the factor representation, see Appendix 1). Since, for all j, s_n is best given beliefs $r_n^j(i_{-n}) = t_j(i_{-n})$, s_n is best against $\hat{r}_n(i_{-n}) = \hat{t}(i_{-n})$. This completes the proof of necessity. ∎

Proof of Theorem 4. The proof of sufficiency is the same as that for Theorem 3, except that one more requirement must be checked. Choose $\delta > 0$ in $\mathcal{R}$ such that $\epsilon^k/\epsilon^{k-1} < \delta$ for all k. The requirement on beliefs in the Theorem implies that if

$$\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})r_{-n}(i_{-n}) >_L \sum_{i_{-n}\in I_{-n}} u_n(i'_n, i_{-n})r_{-n}(i_{-n}),$$

then $\hat{t}(i'_n)/\hat{t}(i_n) \leq \delta$. In terms of the $\mathcal{R}$-valued representation,

$$\sum_{i_{-n}\in I_{-n}} {}^*u_n(i_n, i_{-n})\hat{t}(i_{-n}) > \sum_{i_{-n}\in I_{-n}} {}^*u_n(j_n, i_{-n})\hat{t}(i_{-n}).$$

Choose a sequence $\{\delta_j\}_{j=1}^{\infty}$ in the equivalence class δ. Since δ is infinitessimal, $\lim_{j\to\infty}\delta_j = 0$. Since $\delta > 0$, there exists a set $E \in \mathcal{U}$ such that $\delta_j > 0$ for all $j \in E$. If the hypothesis of the lexicographic condition is true, then for some set $F \in \mathcal{U}$, $j \in F$ implies

$$\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})t_j(i_{-n}) > \sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})t_j(i_{-n}).$$

If the conclusion of the lexicographic condition is true, then there exists a set $G \in \mathcal{U}$ such that for all $j \in G$, $t_j(i'_n) \leq \delta_j t_j(i_n)$. Thus on the set $D' = D \cap E \cap F \cap G$, the additional requirement of proper equilibrium is satisfied. Since $D' \in \mathcal{U}$, D' is infinite. As before, a suitable test sequence can be constructed from $\{t_j\}_{j\in D'}$.

The proof of necessity is also similar to that of Theorem 3. Starting from the test sequence, construct $\hat{t}$, as before. Properness implies

$$\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})t_j(i_{-n}) > \sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})t_j(i_{-n}) \tag{2}$$

implies that $t_j(i'_n) \leq \epsilon_j t_j(i_n)$. Suppose now that

$$\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})r_{-n}(i_{-n}) >_L \sum_{i_{-n}\in I_{-n}} u_n(i'_n, i_{-n})r_{-n}(i_{-n}),$$

so

$$\sum_{i_{-n}\in I_{-n}} {}^*u_n(i_n, i_{-n})\hat{t}(i_{-n}) > \sum_{i_{-n}\in I_{-n}} {}^*u_n(i'_n, i_{-n})\hat{t}(i_{-n}).$$

Then for some $A \in \mathcal{U}$, $j \in A$ implies that equation 2 is satisfied. Then for all $j \in A$, $t_j(i'_n) \leq \epsilon_j t_j(i_n)$. Thus $\hat{t}(i'_n) \leq \dot{\epsilon}\hat{t}(i_n)$. Constructing the factor representation using the algorithm of Appendix 1, we see that the first k for which $t^k(i_n) > 0$ is strictly smaller than the first l for which $t^l(i'_n) > 0$. This completes the proof of necessity. ∎

Proof of Theorem 1. Sufficiency is clearly true, since every possible event has positive probability. We prove necessity using the following easy separation theorem for polyhedral convex sets, whose proof is left to the reader.

Lemma 12. Let C be a closed polyhedral convex set in $\mathbf{R}^l$, and suppose that $C \cap \mathbf{R}^l_{++} = \{0\}$. Then there exists a vector $p \in \mathbf{R}^l_{++}$ such that $p \cdot c \leq 0$ for all $c \in C$.

Suppose that s is an admissible Nash equilibrium. The Theorem will be proved if we can find second-order beliefs r^2_n for each player n that assign positive probability to every pure strategy combination, and such that the expected utility of s_n computed with respect to beliefs r^2_n is at least as high as that of any other possible strategy $s'_n \in S_n$. Suppose that there are l elements in I_{-n}, identified as $e_1, \ldots, e_l$. Let

$$C'_n = \{x \in \mathbf{R}^l \ : \ x_k = \sum_{i\in I_n} u(i, e_k)s'_n(i) \text{ for some } s'_n \in S_n\}.$$

In other words, $x \in C'_n$ if there is a strategy vector s'_n for player n which, for all k, pays off x_k in the event that the other players choose joint strategy combination e_k. Let

$$x^*_n = (\sum_{i\in I_n} u(i, e_1)s_n(i), \ldots, \sum_{i\in I_n} u(i, e_l)s_n(i))$$

be the vector in C'_n associated with the equilibrium strategy vector s_n. Let $C_n = \{x^*_n\} - C'_n$. If the equilibrium s is admissible, $C_n \cap \mathbf{R}^l_{++} = \mathbf{0}$. Furthermore, it is easy to see that C_n is closed, polyhedral and convex. The separation result, Lemma 12, guarantees the existence of a $p \in \mathbf{R}^l_{++}$ such that $x \cdot p \leq 0$ for all $x \in C_n$. The vector p, normalized to be a probability distribution on I_{-n}, is an r^2_n with the required property. ∎

Proof of Theorem 3. The second assertion in the Theorem is an immediate consequence of the definitions. Here I prove the first statement — that if an l.c.e. is a lexicographic Nash equilibrium, then beliefs are independent and shared. Again the proof is most easily completed using $\mathcal{R}$-valued representations of beliefs. The convenience is due to the fact that conditional probabilities are defined just as in the Archimedian case: $\rho^*(A|B) = \rho^*(A \cap B)/\rho^*(B)$.

I begin with the definition of an l.c.e. using the non-Archimedian representation: A correlated strategy is now a $\mathcal{R}$-valued probability distribution on I. (A subjective correlated strategy would be an N-tuple of these objects, one for each player.) An l.c.e. is a correlated strategy σ^* satisfying for each player n and all $i_n, j_n \in I_n$,

$$\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})\sigma^*(i_n, i_{-n}) \geq \sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})\sigma^*(i_n, i_{-n}).$$

Rewriting in terms of conditional probabilities,

$$\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})\sigma^*(i_{-n}|i_n)\sigma^*(i_n) \geq \sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})\sigma^*(i_{-n}|i_n)\sigma^*(i_n).$$

Dividing by $\sigma^*(i_n)$ and multiplying by $\sigma^\circ(i_n)$ gives, for all $i_n, j_n \in I_n$,

$$\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})\sigma^*(i_{-n}|i_n)\sigma^\circ(i_n) \geq \sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})\sigma^*(i_{-n}|i_n)\sigma^\circ(i_n).$$

Since this l.c.e. is in fact a lexicographic Nash equilibrium, it follows that $\sigma^*(i_{-n}|i_n) = \sigma^*(i_{-n})$ for all $i_n \in I_n$ and $i_{-n} \in I_{-n}$. Substituting,

$$\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})\sigma^*(i_{-n})\sigma^\circ(i_n) \geq \sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})\sigma^*(i_{-n})\sigma^\circ(i_n).$$

Summing over $i_n \in I_n$,

$$\begin{aligned}\sum_{i_n\in I_n}\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})\sigma^*(i_{-n})\sigma^\circ(i_n) &\geq \sum_{i_n\in I_n}\sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})\sigma^*(i_{-n})\sigma^\circ(i_n)\\ &= \sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})\sigma^*(i_{-n})\sigma^\circ(i_n),\end{aligned}$$

for all $j_n \in I_n$. Thus for all n and any $\mathbf{R}$-valued probability distribution τ° on I_n,

$$\sum_{i_n\in I_n}\sum_{i_{-n}\in I_{-n}} u_n(i_n, i_{-n})\sigma^*(i_{-n})\sigma^\circ(i_n) \geq \sum_{j_n\in I_n}\sum_{i_{-n}\in I_{-n}} u_n(j_n, i_{-n})\sigma^*(i_{-n})\tau^\circ(j_n).$$

This is the definition (in $\mathcal{R}$-valued probabilities) of a lexicographic Nash equilibrium with shared beliefs.

Finally, since the l.c.e. is not correlated,

$$\sigma^*(i) = \sigma^*(i_{-1}|i_1)\sigma^*(i_1) = \sigma^*(i_{-1,2})\sigma^*(i_2)\sigma^*(i_1)\cdots = \sigma^*(i_N)\sigma^*(i_{N-1})\cdots\sigma^*(i_1).$$

In other words, beliefs are independent. ∎

REFERENCES

Anscombe, F.J. and R.J. Aumann, A definition of subjective probability, *Annals of Mathematical Statistics* 34 (1963), 199–205.

Brandenburger, A. and E. Dekel, On an axiomatic approach to refinements of Nash equilibrium, unpublished manuscript, University of California, Berkeley, 1986a.

Subjective expected utility, admissibility, and posterior rationality, unpublished manuscript, Cambridge University, 1986b.

Chipman, J.S., The foundations of utility, *Econometrica* 28 (1960), 193–224.

Fishburn, P., *The foundations of expected utility*, 1982, Reidel.

Kohlberg, E. and J.-F. Mertens, On the strategic stability of equilibrium, *Econometrica* 54 (1986), 1003-1037.

Kreps, D.M. and R. Wilson, Sequential equilibria, *Econometrica* 50 (1982), 863–894.

McLennan, A., Consistent conditional systems in non-cooperative game theory, unpublished manuscript MSRI 05217-86, Mathematical Sciences Research Institute, Berkeley, 1986.

Myerson, R., Refinements of the Nash equilibrium concept, *International Journal of Game Theory* 7 (1978), 73–80.

Richter, M.K., Rational choice, in Chipman, Hurwicz, Richter and Sonnenschein, eds., *Preferences, Utility and Demand*, 1971, Harcourt Brace Jovanovich, New York.

Rockafellar, R. Tyrrell, *Convex analysis*, 1970, Princeton University Press, Princeton N.J.

Savage, L.J., *Foundations of statistics*,1954, Wiley, New York.

Selten, R., Reexamination of the perfectness concept for equilibrium points in sequential games, *International Journal of Game Theory* 4 (1975), 25–55.

Recent CREST Working Papers

87–1: Jeffrey K. MacKie-Mason, "Nonlinear Taxation of Risky Assets and Investment, With Application to Mining" September, 1984.

87–2: Jeffrey K. MacKie-Mason, "Sequential Decision Problems and Asymmetric Information" September, 1985.

87–3: Michelle J. White, "Contract Breach and Contract Discharge due to Impossibility: A Unified Theory" July 29, 1987.

87–4: Ted Bergstrom, "Systems of Benevolent Utility Interdependence" May 19, 1987.

87–5: Ted Bergstrom, "A Fresh Look at the Rotten Kid Theorem—and Other Household Mysteries" November, 1986.

87–6: Michelle J. White, "The Corporate Bankruptcy Decision" July, 1987.

87–7: Michelle J. White, "Location Choice and Commuting Behavior in Cities with Decentralized Employment" July, 1987.

87–8: Lawrence E. Blume and David Easley, "Implementation of Walrasian Expectations Equilibria" December, 1985.

87–9: Lawrence E. Blume, "Lexiocographic Refinements of Nash Equilibrium" April, 1986.

87–10: David Lam, "Lorenz Curves, Inequality, and Social Welfare Under Changing Population Composition" June 16, 1987.

87–11: Mark Bagnoli and Naveen Khanna, "Equilibrium with Debt and Equity Financing of New Projects: Why More Equity Financing Occurs When Stock Prices are High" June, 1987.

87–12: Mark Bagnoli and Barton L. Lipman, "Provision of Public Goods: Fully Implementing the Core through Private Contributions" March, 1987.

87–13: Mark Bagnoli and Barton L. Lipman, "Successful Takeovers without Exclusion" August, 1987.

87–14: Mark Bagnoli and Michael McKee, "Controlling the Game: Political Sponsors and Bureaus" May, 1987.

87–15: Mark Bagnoli and Michael McKee, "Can the Private Provision of Public Goods be Efficient?—Some Experimental Evidence" March, 1987.

87–16: Mark Bagnoli, "Non-Market Clearing Prices in a Dynamic Oligopoly with Incomplete Information" January, 1986.

87–17: John Laitner, "Bequests, Gifts, and Social Security" February 28, 1986.

87–18: John Laitner, "Dynamic Determinacy and the Existence of Sunspot Equilibria" May 12, 1986.

87–19: David Lam, "Does a Uniform Age Distribution Minimize Lifetime Wages?" August 12, 1987.

87–20: David Lam, "Assortative Mating with Household Public Goods" April, 1987.

87–21: Jeffrey A. Miron and Stephen P. Zeldes, "Production, Sales, and the Change in Inventories: An Identity that Doesn't Add Up" June 1987.

87–22: Jeffrey A. Miron and Stephen P. Zeldes, "Seasonality, Cost Shocks, and the Production Smoothing Model of Inventories" December, 1986.

87–23: Hal R. Varian, "Differences of Opinion in Financial Markets" March, 1985.

87–24: Roger H. Gordon and Hal R. Varian, "Taxation of Asset Income in the Presence of a World Securities Market" August, 1986.

87–25: Hal R. Varian, "Measuring the Deadweight Costs of DUP and Rent Seeking Activities" November, 1982.

87–26: Hal R. Varian, "Price Discrimination" January, 1987.

87–27: Roger H. Gordon and Hal R. Varian, "Intergenerational Risk Sharing" October, 1985.

87–28: Hal R. Varian, "Three Papers on Revealed Preference" August, 1987.

87–29: Hal R. Varian, "Optimal Tariffs and Financial Assets" April, 1987.

87–30: Jonathan Cave and Stephen W. Salant, "Cartels That Vote: Agricultural Marketing Boards and Induced Voting Behavior" August, 1987.

87–31: Stephen W. Salant and Donald H. Negri, "Pastures of Plenty: When is the Standard Analysis of Common Property Extraction Under Free Access Incorrect?" July 10, 1987.

87–32: Stephen W. Salant, "When is Inducing Self-Selection Sub- optimal for a Monopolist?" February, 1987.

87–33: Stephen W. Salant, "Treble Damage Awards in Private Lawsuits for Price-Fixing" August, 1987.

87–34: Stephen W. Salant and Roy Danchick, "Air Force Academy Attrition: A New Perspective on the College Dropout Problem" August, 1987.

87–35: Stephen W. Salant and Eban Goodstein, "Committee Voting Under Alternative Procedures and Preferences: An Experimental Analysis" April 20, 1987.

87–36: Robert B. Barsky and Jeffrey A. Miron, "The Seasonal Cycle and the Business Cycle" June, 1987.

87–37: Robert B. Barsky, N. Gregory Mankiw, Jeffrey A. Miron and David N. Weil, "The Worldwide Change in the Behavior of Interest Rates and Prices in 1914" July, 1987.

87–38: Jeffrey K. MacKie-Mason, "Taxes, Information and Corporate Financing Choices" April 1986.