



Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19103-6297
pier@ssc.upenn.edu
<http://www.econ.upenn.edu/pier>

PIER Working Paper 97-011

“How Proper is Sequential Equilibrium”

by

George J. Mailath, Larry Samuelson and Jeroen M. Swinkels

How Proper is Sequential Equilibrium?*

George J. Mailath
Department of Economics
3718 Locust Walk
University of Pennsylvania
Philadelphia, PA 19104-6297

Larry Samuelson
Department of Economics
1180 Observatory Drive
University of Wisconsin
Madison, WI 53706

Jeroen M. Swinkels
Kellogg Graduate School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2009

July 16, 1996

Abstract

A strategy profile of a normal form game is proper if and only if it is quasi-perfect in every extensive form (with that normal form). Thus, properness requires optimality along a sequence of supporting trembles, while sequentiality only requires optimality in the limit. A decision-theoretic implementation of sequential rationality, strategic independence respecting equilibrium (SIRE), is defined and compared to proper equilibrium, using lexicographic probability systems. Finally, we give tremble-based characterizations, which do not involve structural features of the game, of the rankings of strategies that underlie proper equilibrium and SIRE.

Keywords: Refinements, proper equilibrium, sequential equilibrium, perfect equilibrium, trembles, lexicographic probability systems, indifferences.

JEL classification numbers: C70, C72.

*This is a substantial revision of "Structural Indifference in Normal Form Games." We thank John Hillas, Phil Reny, an associate editor, and two referees for helpful comments. Financial support from the National Science Foundation is gratefully acknowledged. Mailath and Samuelson thank the Economics Department at the University of Bonn for its hospitality and acknowledge financial support from Deutsche Forschungsgemeinschaft (SFB 303). They also thank the Center for Rationality and Interactive Decision Theory and the Institute for Advanced Studies at the Hebrew University of Jerusalem for their hospitality and financial support during a revision of this paper.

How Proper is Sequential Equilibrium?

by George J. Mailath, Larry Samuelson, and Jeroen M. Swinkels

1. Introduction

Among the more surprising results in game theory is that a proper equilibrium in a normal form game induces a sequential equilibrium in every corresponding extensive form (van Damme [11] and Kohlberg and Mertens [4]). The converse however, can be false: a strategy profile can be sequential in every extensive form with a given normal form without being proper. Proper equilibrium is thus stronger than sequential-in-every-tree. What is the difference between the two, and how much stronger is proper equilibrium?

The paper addresses these questions by exploring the relationship between proper equilibrium and two other concepts: quasi-perfect in every tree and strategic independence respecting equilibrium (SIRE). *Quasi-perfect equilibrium*, which van Damme [11] defined and showed is implied by proper equilibrium, is an extensive-form concept closely related to trembling-hand perfect equilibrium.¹ *Strategic independence respecting equilibrium* is a normal-form concept that is, we argue, the appropriate decision-theoretic description of sequential rationality (a key element of sequentiality).²

We begin by showing that properness is equivalent to quasi-perfect-in-every-tree: A converging sequence of perturbed normal form strategy profiles supports its limit as a proper equilibrium if and only if the sequence of strategies induced in every corresponding extensive form supports its limit as a quasi-perfect equilibrium.³ We thus obtain an extensive-form characterization of the distinction between properness and sequentiality, since quasi-perfection requires players to act optimally against each term in the sequence of opponents' strategies, while sequential equilibrium requires optimality only against the limit of this sequence.

While the difference in the definitions of quasi-perfection and sequentiality is easy to understand, the differences in the *implications* of those definitions are more subtle.

¹It differs from extensive-form trembling-hand perfection in that, at each information set, the player choosing an action ignores the possibility of her own future mistakes.

²Both sequential equilibrium in the extensive form and SIRE (like most existing solution concepts) impose rather strong cross-player consistency conditions on beliefs. We will not try to argue that these conditions are natural.

³A key ingredient in this result is the requirement that the same sequence be used in all corresponding extensive forms. The result is false without the restriction that the same sequence is used in all corresponding extensive forms (see Hillas [3]).

In particular, sequential rationality is often described as the requirement that, at every information set, a player’s choice makes sense if that information set is reached, even if the player learns something unexpected about the play of her opponents (i.e., the information set is off the play path). In contrast, quasi-perfection seems to have little to do with the information a player learns about her opponents during play.⁴ Moreover, every extensive-form game used in the proof of the equivalence of proper and quasi-perfection has the property that each player receives no information about the play of her opponents.⁵ Since sequential rationality appears to depend in a central way on the notion of information, while quasi-perfection does not, our extensive-form characterization is an incomplete description of the relationship between properness and sequential-in-every-tree.

The remainder of the paper is concerned with a normal-form characterization of the relationship between properness and sequential-in-every-tree. The requirement that planned actions at an information set make sense when that information set is reached seems intrinsically an extensive-form notion. However, Mailath, Samuelson, and Swinkels [5] (henceforth MSS) argue that this is not so: statements about “when” an action will matter can be translated into statements about “if an action matters.” That is, sequential rationality can be rephrased as the requirement that a decision that only matters for a given subset of strategies by the opponents should be made as if a strategy from that subset had been chosen. In the normal form, the phrase “only matters for a given subset of strategies by the opponents” comes down to a particular pattern of payoff ties for the player making the decision. We call this pattern of ties a *strategic independence*. Like an information set in an extensive form, a strategic independence captures a situation in which a player can uncouple her decision into two parts, one of which is relevant if the information set is reached, and one if it is not.⁶

Requiring sequential rationality in the normal form (optimal play in the limit at all strategic independences) yields *strategic independence respecting equilibrium (SIRE)*. The normal-form characterization of the relationship between proper and sequential equilibria then involves two elements: the connection between sequential-in-every-tree

⁴Sequentiality is more sensitive than quasi-perfection to the information structure of extensive-form games. Consider the following two-player game of perfect information: Player I chooses “Out” or “In;” the game ends after O yielding payoffs 0 and 1 to players I and II, respectively; and player II has two possible actions $\{A, C\}$ after I , with (I, A) yielding payoffs 1 and 0 and (I, C) yielding payoffs -1 and -1 to players I and II, respectively. The only sequential equilibrium is (I, A) . However, in the simultaneous-move extensive-form game in which player II does not know if player I has chosen I when choosing between A and C when choosing between A and C , (O, C) is also a sequential equilibrium. In contrast, only (I, A) is quasi-perfect in either extensive form.

⁵This does not mean that the dynamic structure of the extensive form is irrelevant. The different extensive forms are used to force players to make choices between (and so rank) every pair of strategies.

⁶A strategic independence for player i is a subset of strategy profiles, $X_i \times X_{-i}$, and is *reached* if the opponents’ choices are in X_{-i} .

and SIRE, and the connection between SIRE and proper.

The connection between sequential-in-every-tree and SIRE is straightforward: In an extensive form, if an information set is not reached, not only is the decision maker indifferent about her choice at that information set, *but all other players are indifferent as well*. In contrast, a strategic independence in the normal form is defined solely in terms of the payoffs to the decision maker. A strategic independence for which the payoff ties also hold for the other players is called a *normal form information set*. Requiring sequential rationality only at normal form information sets rather than at all strategic independences yields *normal form sequential equilibrium*. MSS show that an assessment induces a sequential equilibrium in every extensive form with a given reduced normal form if and only if it constitutes a normal form sequential equilibrium in that reduced normal form. The difference between SIRE and sequential-in-every-tree is thus the difference between sequential rationality at all strategic independences and sequential rationality only at those strategic independencies satisfying the extra condition on other players' payoffs. From the point of view of the decision maker, payoff ties for the other players are irrelevant and so there is no decision-theoretic reason to distinguish between the different strategic independences. Hence, SIRE is the appropriate decision-theoretic formulation of sequential rationality.

The relationship between proper and SIRE is the normal-form analog of the relationship between quasi-perfection and sequentiality. In particular, properness is equivalent to optimal play along the sequence at all strategic independences, while SIRE requires optimality against the limit.

We next turn to the decision-theoretic foundations of SIRE and proper. Blume, Brandenburger, and Dekel [2] provide a decision-theoretic characterization of proper equilibrium in terms of *lexicographic probability systems* (LPSs). We provide a similar decision-theoretic characterization of SIRE.

In the Blume, Brandenburger, and Dekel [2] characterization, a proper equilibrium can be viewed as the result of players' ranking their strategies according to a lexicographic probability system (i.e., a hierarchy of beliefs) over the play of their opponents, where the hierarchy of beliefs reflects the opponents' payoffs in an intuitive manner. A player first consults her first-level belief about opponents' play and ranks her strategies according to their expected payoffs given this belief. If any indifferences occur, the player appeals to her second-level beliefs, breaking indifferences accordingly to their expected payoffs against this belief. Further indifferences are appealed to a third-level belief, and so on. This process continues until either all indifferences have been broken or sufficiently many beliefs have been encountered that their supports exhaust the opponents' strategy spaces.

Like proper, SIRE can be viewed as a ranking of strategies driven by a hierarchy of beliefs about opponents' play. However, in a SIRE a player appeals an indifference to the next level only if it is a *structural indifference*, that is, if the player's indifference is

caused by ties in the payoff matrix (as opposed to an indifference created by a fortuitous belief about opponents' play).⁷ The decision-theoretic difference between SIRE and properness thus hinges on how they treat cases in which players are indifferent between strategies.

Our last result compares properness and sequentiality in terms of trembles, but without reference to information sets or strategic independences. We show that the structural ordering (underlying SIRE) ranks player i 's strategy r_i ahead of s_i if and only if r_i receives a higher payoff than s_i along *every* sequence of trembles that converges (in a sense made precise below) to the underlying hierarchy of beliefs about play. The lexicographic ordering (underlying properness) ranks r_i ahead of s_i if and only if r_i achieves a higher payoff along a converging sequence of trembles drawn from a particular subset of such sequences. Both proper and SIRE are characterized in terms of optimal behavior against sequences of trembles, but the set of sequences of trembles is different.⁸ The conditions for properness to rank a pair of strategies are weaker than the corresponding conditions for SIRE, and so properness accordingly imposes more restrictions on strategy choices than does SIRE.

The following section introduces notation. Section 3 presents the extensive-form characterization of proper equilibrium in terms of quasi-perfection, and so the characterization of the difference between properness and sequentiality in terms of optimality along a sequence of perturbations versus optimality in the limit. Section 4 introduces strategic independences and SIRE. Lexicographic belief systems and the strategy orderings underlying properness and SIRE are described in Section 5. Section 6 characterizes these strategy orderings in terms of sequences of perturbed strategies.

2. Preliminaries

We denote the set of players by N and player i 's (pure) strategy set by S_i , $i \in N$, with $|S_i| < \infty$. Typical strategies for player i are denoted r_i , s_i , and t_i . The set of strategy profiles is given by $S \equiv \prod_{i \in N} S_i$. A set of strategy profiles S and a payoff function $\pi : S \rightarrow \mathfrak{R}^N$ constitute the normal form game (S, π) . A subset of player i 's pure strategy space is denoted X_i . The set of probability distributions over a set $X_i \subseteq S_i$ is denoted $\Delta(X_i)$. A subscript $-i$ denotes $N \setminus \{i\}$ and a subscript $-I$ denotes $N \setminus I$. For any distribution $\beta \in \Delta(S_{-i})$, define expected payoffs given this distribution over

⁷Normal form sequential equilibrium appeals an indifference to the next level only if *every* player in the game is structurally indifferent between the outcomes involved (as opposed to indifference only on the part of the player making the decision).

⁸There is, as far we know, no corresponding tremble characterization of the ordering described in footnote 7 for normal form sequential equilibrium.

opponents' strategies as:

$$E\{\pi_i(s_i, \beta)\} = \sum_{s_{-i} \in S_{-i}} \pi_i(s_i, s_{-i})\beta(s_{-i}).$$

Definition 1. Strategies r_i and s_i agree for player i on $X_{-i} \subseteq S_{-i}$ if, $\forall s_{-i} \in X_{-i}$,

$$\pi_i(r_i, s_{-i}) = \pi_i(s_i, s_{-i}).$$

Strategies r_i and s_i agree for all players (or agree) on $X_{-i} \subseteq S_{-i}$ if, $\forall s_{-i} \in X_{-i}$,

$$\pi_j(r_i, s_{-i}) = \pi_j(s_i, s_{-i}), \quad \forall j \in N.$$

We use the term *agree for player i on X_{-i}* to emphasize that player i is indifferent between r_i and s_i for any fixed strategy profile, s_{-i} , of the other players in X_{-i} , but other players might not be. Note that these indifferences are due to the structure of payoff ties in the game, and do not depend on some particular mixed strategy profile of the other players. We will accordingly say that r_i and s_i are *structurally indifferent* on X_{-i} .

We say that the normal form (S, π) is a *pure strategy reduced normal form game (PRNF)* if there does not exist a pair $\{r_i, s_i\}$ with $r_i \neq s_i$ and r_i agreeing with s_i on S_{-i} . Any normal form is easily written as a PRNF by treating the set of pure strategies that agree on S_{-i} as a single strategy.

For any mixture on PRNF strategy profiles P , we also let P denote the mixture on normal-form strategies obtained by dividing the probability attached to each PRNF strategy equally among the corresponding normal-form strategies.⁹

A *probability sequence*, $\{\{P_i^n\}_{n=1}^\infty, i \in N\}$, is a collection of *independent* probability distributions such that each P_i^n is a completely mixed probability distribution on S_i . Given a probability sequence, we define P^n and P_{-i}^n in the obvious manner, i.e., $P^n(s) = \prod_{i \in N} P_i^n(s_i)$ and $P_{-i}^n(s) = \prod_{j \neq i} P_j^n(s_j)$. We say that the probability sequence $\{P^n\}$ is *conditionally convergent* if, for all subsets $X \subseteq S$, $\{P^n(\cdot | X)\}$ (and all its marginals) are convergent sequences, where $P^n(s | X) \equiv P^n(s)/P^n(X)$ for $s \in X$ and zero otherwise. We sometimes refer to converging probability sequences as “trembles” or “sequences of perturbed strategy profiles.”

Definition 2. (Myerson [9]) A strategy profile σ is **proper** if there is a probability sequence $\{P^n\}$ with $\lim P^n = \sigma$ and a sequence $\{\epsilon_n\}$, $\epsilon_n \rightarrow 0$, such that, for all i and all $r_i, s_i \in S_i$,

$$E\{\pi_i(s_i, P_{-i}^n)\} < E\{\pi_i(r_i, P_{-i}^n)\} \Rightarrow P_i^n(s_i) \leq \epsilon_n P_i^n(r_i). \quad (1)$$

⁹A PRNF strategy need not imply unique choices in an extensive form, because different actions at an extensive form information set may be consistent with the same PRNF strategy. This ambiguity is unimportant because the differing actions consistent with a given PRNF strategy profile do not affect any players' payoffs. Hence, any other completely-mixed division of the probability among corresponding normal form strategies could have been chosen.

By taking subsequences, we can always ensure that $\{P^n\}$ is conditionally convergent. If σ is a proper equilibrium and P^n is the probability sequence satisfying (1), then we say that σ is *supported* by $\{P^n\}$.

3. An Extensive-Form Characterization of Properness

This section characterizes the difference between the equilibrium concepts of sequential-in-every-tree and properness. Sequential-in-every-tree requires strategies at every information set (in every tree) to be best responses to the *limits* of sequences of opponents' strategies. Properness requires strategies at every information set to be best responses to *all of the terms* in the sequence of perturbed strategies.

The extensive-form equilibrium concept used to characterize proper equilibrium is that of a quasi-perfect equilibrium:

Definition 3. (van Damme [11]) *A conditionally convergent probability sequence $\{P^n\}$ with limit σ induces a quasi-perfect equilibrium in an extensive form game Γ if, for the corresponding sequence of completely mixed behavior strategies b^n and limit b^∞ , for each player i and information set h for that player, contingent on having reached h , b_i^∞ is a best reply to b_{-i}^n for all n .*

Quasi-perfection is closely related to sequentiality: sequentiality is obtained by replacing “ b_i^∞ is a best reply to b_{-i}^n for all n ” in the above definition with “ b_i^∞ is a best reply to $\lim_{n \rightarrow \infty} b_{-i}^n$.” That is, sequentiality requires best replies, at all information sets, to the limits of a perturbed sequence of opponents' strategies, while quasi-perfection requires best replies, at all information sets, to each element of the sequence of perturbed strategies.

Extensive-form trembling-hand perfection (hereafter, perfection) requires player i to play a best response at every information set to the perturbed strategies of her opponents *and* to perturbed versions of her own continuation strategies. Quasi-perfection, on the other hand, forces player i to ignore the perturbations in her own strategies. As a result, there is no inclusion relationship between perfection and quasi-perfection. The two standard examples illustrating this are in Figure 1. In the first extensive form, LL is quasi-perfect, but not perfect, while in the second, RL is perfect, but not quasi-perfect. A quasi-perfect equilibrium must be sequential, but the converse fails.

It would be consistent with common usage to characterize a strategy profile σ of the PRNF (S, π) as inducing a quasi-perfect equilibrium in an extensive form (with that PRNF) if there exists a sequence of completely mixed behavior strategy profiles whose limit is equivalent to σ and is a quasi-perfect equilibrium. In contrast, we have defined a *sequence* $\{P^n\}$ as inducing a quasi-perfect equilibrium in an extensive form if $\{P^n\}$ yields a sequence of completely mixed behavior strategies that converges to a limit that,

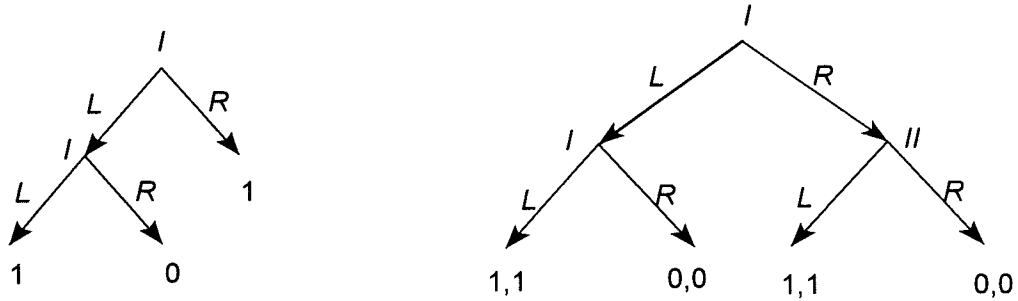


Figure 1: There is no inclusion relationship between perfection and quasi-perfection.

together with the sequence, satisfies the conditions for quasi-perfection. Thus, when $\{P^n\}$ induces a quasi-perfect equilibrium in every extensive form with a given PRNF, the strategy sequences supporting the quasi-perfect equilibrium in the different extensive forms are derived from the *same* sequence of completely mixed PRNF strategies.¹⁰

While a proper equilibrium need not induce a perfect equilibrium in every extensive form (LL in the first extensive form in Figure 1, for example), a proper equilibrium does induce a quasi-perfect equilibrium in every extensive form (van Damme [11]). Our first result is that this property characterizes proper equilibria:¹¹

Proposition 1. *A conditionally convergent probability sequence $\{P^n\}$ on S induces a quasi-perfect equilibrium in every extensive form with PRNF (S, π) if and only if the limit σ of $\{P^n\}$ is a proper equilibrium supported by $\{P^n\}$ in (S, π) .*

The proof of this proposition begins with a straightforward reformulation of proper equilibrium (whose proof is omitted):

¹⁰Figures 11-13 in MSS describe a simple example of a pair of extensive forms with the same normal form with the property that a strategy profile can be supported as a sequential equilibrium in each extensive form, but only by using different trembles. Hillas [3] contains an example of a strategy profile that is not proper, and yet can be supported by (necessarily different) trembles as a quasi-perfect in every tree.

¹¹In a previous version, we stated this proposition with the additional assumption of transference of decision maker indifference (see footnote 16). We are grateful to John Hillas for pointing out that our proof did not require this assumption. Also, see Hillas [3] for a different proof of this theorem.

Lemma 1. *The strategy profile σ is a proper equilibrium if and only if there exists a conditionally convergent probability sequence $\{P^n\}$, with $\lim P^n = \sigma$ and a sequence $\epsilon_n \rightarrow 0$ such that, for all i and all $r_i, s_i \in X_i$,*

$$E\{\pi_i(s_i, P_{-i}^n)\} < E\{\pi_i(r_i, P_{-i}^n)\} \Rightarrow P_i^n(s_i | \{r_i, s_i\}) \leq \epsilon_n. \quad (2)$$

Proof of Proposition 1: [ONLY IF] Suppose $\{P^n\}$ induces a quasi-perfect equilibrium in every tree. Since $\{P^n\}$ is completely mixed, it can be viewed as a completely mixed behavior strategy profile. Fix a pair of strategies s_i and r_i for player i such that

$$E\{\pi_i(s_i, P_{-i}^n)\} < E\{\pi_i(r_i, P_{-i}^n)\}. \quad (3)$$

The trivial extensive form representation of the PRNF, interpreted as a simultaneous move game, has one information set for each player, with $|S_j|$ choices for player j . Let Γ denote the extensive form obtained from this trivial extensive form by the following application of the ‘‘coalesce’’ transformation: Replace the single information set of player i with two sequential choices, the first information set has $|S_i| - 1$ actions, corresponding to $\{r_i, s_i\}$ and the strategies in $S_i \setminus \{r_i, s_i\}$, with the action $\{r_i, s_i\}$ leading to a second information set, h , with two actions, r_i and s_i . In Γ , if player i wishes to play either r_i or s_i , i must first select $\{r_i, s_i\}$, and then choose between r_i and s_i . Since $\{P^n\}$ induces a quasi-perfect equilibrium in Γ , (3) implies $P_i^n(s_i | \{s_i, r_i\}) \rightarrow 0$. Because S_i is finite, a sequence $\{\epsilon_n\}$ can be found such that (2) is satisfied, and so σ is proper.

[IF] This is van Damme [11, Theorem 1]. □

Thus, the distinction between properness and sequential-in-every-tree is that properness requires optimality along the sequence, while the latter only requires optimality in the limit. One of the attractive aspects of Proposition 1 is that it gives a characterization of properness that involves optimal play against a sequence of perturbed strategies, rather than condition (1), which only requires almost-optimal play. This characterization does, however, require considering different extensive forms (with the same PRNF). The key step in the proof of Proposition 1 involves finding, for every pair of strategies for each player i , an extensive form in which player i chooses from just these two strategies while all other players still have all of their strategies available. For any pair of strategies, there are extensive forms with such an information set, but, in general, there is no extensive form that captures all of these information sets for all players (see Mailath, Samuelson, and Swinkels [6, Section III] for a canonical example).¹²

¹²The distinction between properness and sequential-in-every-tree is reminiscent of that between trembling hand perfection and sequentiality in a given tree. Moreover, generically (in extensive form payoffs) the latter two coincide. This suggests that, in some sense, generically, proper and sequential-in-every-tree coincide. However, since we must deal simultaneously with several extensive forms at the same time, the description of the appropriate genericity requirement is a subtle issue. Section 10 of Mailath, Samuelson, and Swinkels [7] discusses this issue in some detail.

4. Sequential Rationality

4.1. Strategic Independence Respecting Equilibrium

The idea behind sequential equilibrium, optimality at every information set, is often phrased in terms of restrictions on the behavior of a player *when he is asked to make a decision*. This suggests that an important feature of the structure in an extensive form game is that a player need not make a decision until required to by the realized play of the game. In MSS, we argued that this is incorrect: it is not that the choice of an action at an information set need not be made until that information set is reached that is important, but rather that the choice of action, *whenever taken*, “matters” (i.e., affects the outcome of the game) only if that information set is reached. Hence, whether a player makes a decision at an information set or makes an ex ante contingency plan, the player’s action for that information set can be made as if the information set in question has been or will be reached. Sequential equilibrium requires that this choice be a best response to some belief about opponents’ play given that the information set in question is reached.

The normal form structure that captures situations in which a player can make a decision as if he knew that his opponents had chosen from a particular subset of their strategies was called a *strategic independence* in MSS:

Definition 4. *The set $X \subseteq S$ is strategically independent for player i if*

1. $X = X_i \times X_{-i}$, and
2. $\forall r_i, s_i \in X_i, \exists t_i \in X_i$ such that t_i and r_i agree for player i on X_{-i} and t_i and s_i agree for player i on $S_{-i} \setminus X_{-i}$.

Suppose X is a strategic independence for player i . If s_i and $t_i \in S_i$ agree for player i on $X_{-i} \subseteq S_{-i}$, then we say that s_i and t_i are X_{-i} -equivalent. If s_i and t_i agree for player i on $S_{-i} \setminus X_{-i}$, then we say that s_i and t_i are $S_{-i} \setminus X_{-i}$ -equivalent. Then when player i is evaluating strategies in X_i , we can think of him as independently choosing an X_{-i} -equivalence class of strategies in X_i and an $S_{-i} \setminus X_{-i}$ -equivalence class; these together determine a strategy in X_i . The optimality of the choice of an X_{-i} -equivalence class is a function only of beliefs over X_{-i} .¹³ Similarly, the optimality of the choice of an $S_{-i} \setminus X_{-i}$ -equivalence class is a function only of beliefs over $S_{-i} \setminus X_{-i}$. We can thus think of player i ’s choice in X_i as one of choosing behavior that is relevant if opponents choose from X_{-i} and independently choosing behavior that is relevant if opponents choose from $S_{-i} \setminus X_{-i}$.

¹³This follows from the fact that, in any two X_{-i} -equivalence classes, there are strategies s_i and s'_i which agree on $S_{-i} \setminus X_{-i}$, so that an X_{-i} -equivalence class can be chosen without worrying about $S_{-i} \setminus X_{-i}$.

The decision-theoretic analog of requiring best replies at all information sets is requiring best replies at all strategic independences:

Definition 5. *The limit of a conditionally convergent sequence $\{P^n\}$ is a **strategic independence respecting equilibrium (SIRE)** if for all i and any strategic independence X for player i , $\lim P_i^n(\cdot | X_i)$ is a best reply from among the elements of X_i to $\lim P_{-i}^n(\cdot | X_{-i})$, i.e., for all $r_i, s_i \in X_i$,*

$$E\{\pi_i(s_i, \lim P_{-i}^n(\cdot | X_{-i}))\} < E\{\pi_i(r_i, \lim P_{-i}^n(\cdot | X_{-i}))\} \Rightarrow \lim P_i^n(s_i | X_i) = 0. \quad (4)$$

As usual, we say that the strategic independence respecting equilibrium is *supported* by the sequence $\{P^n\}$. Some examples are discussed in Section 5.3.

As a second manifestation of the limit of optima vs. optima of limits distinction, we note:

Lemma 2. *A strategy profile σ is proper if and only if there is a conditionally convergent probability sequence $\{P^n\}$, with $\lim P^n = \sigma$, and a sequence $\{\epsilon_n\}$, $\epsilon_n \rightarrow 0$, such that, for all i , all strategic independences X for i , and all $r_i, s_i \in X_i$,*

$$E\{\pi_i(s_i, P_{-i}^n(\cdot | X_{-i}))\} < E\{\pi_i(r_i, P_{-i}^n(\cdot | X_{-i}))\} \Rightarrow P_i^n(s_i | X_i) \leq \epsilon_n. \quad (5)$$

Proof. [ONLY IF]: Let σ be a proper equilibrium and let $\{P^n\}$ and $\{\epsilon_n\}$ be sequences satisfying (1). Suppose $E\{\pi_i(s_i, P_{-i}^n(\cdot | X_{-i}))\} < E\{\pi_i(r_i, P_{-i}^n(\cdot | X_{-i}))\}$. Since X is a strategic independence, there exists $t_i \in X_i$ that agrees with s_i on $S_{-i} \setminus X_{-i}$ and agrees with r_i on X_{-i} . Then $E\{\pi_i(s_i, P_{-i}^n)\} < E\{\pi_i(t_i, P_{-i}^n)\}$ and hence properness implies $P_i^n(s_i) \leq \epsilon_n P_i^n(t_i)$. Since $s_i, t_i \in X_i$, we then have $P_i^n(s_i | X_i) \leq \epsilon_n P_i^n(t_i | X_i) \leq \epsilon_n$, giving the result.

[IF]: Since $X \equiv \{r_i, s_i\} \times S_{-i}$ is a strategic independence for player i , and on that strategic independence, $E\{\pi_i(r_i, P_{-i}^n(\cdot | X_{-i}))\} = E\{\pi_i(r_i, P_{-i}^n)\}$ and $E\{\pi_i(s_i, P_{-i}^n(\cdot | X_{-i}))\} = E\{\pi_i(s_i, P_{-i}^n)\}$, the result follows from Lemma 1. \square

Condition (5) implies that if there exists an n^* such that the antecedent of (5) holds for all $n > n^*$, then $\lim P_i^n(s_i | X_i) = 0$.¹⁴ The definition of SIRE is thus the limit of the characterization in Lemma 2: proper equilibrium is the limit of a sequence of strategies, each element of which satisfies an optimality property, while SIRE requires optimality only with respect to the limiting strategies induced by such a sequence.¹⁵ This also implies, of course, that a proper equilibrium is a SIRE, and so SIRE exist.

¹⁴Note that, since $\{P^n\}$ is conditionally convergent, if (5) holds for all n , either $E\{\pi_i(s_i, P_{-i}^n(\cdot | X_{-i}))\} \leq E\{\pi_i(r_i, P_{-i}^n(\cdot | X_{-i}))\}$ holds for all n sufficiently large, or the reverse inequality holds for all n sufficiently large.

¹⁵Since the statement and proof of Lemma 2 are still valid if strategic independence is replaced by normal form information set, a similar distinction holds between properness and normal form sequential equilibrium.

4.2. Sequential-in-every-tree

If the definition of a strategic independence is strengthened by requiring agree for all players, rather than just agree for player i , then the resulting normal form structure is called a *normal form information set*. This structure is the focus of MSS, where it is shown that a set of strategy profiles X of a PRNF (S, π) is a normal form information set for player i if and only if there exists an extensive form game without nature with PRNF (S, π) with an information set h for player i such that the set of PRNF strategies that make h reachable is precisely X (Theorem 1).

The definition of *normal form sequential equilibrium* is obtained by replacing “strategic independence” with “normal form information set” in the definition of SIRE. Normal form sequential equilibrium is precisely sequential-in-every-tree, in the sense that the limit of a conditionally convergent sequence $\{P^n\}$ in a PRNF is a normal form sequential equilibrium if and only if $\{P^n\}$ induces a sequential equilibrium in every extensive form with that PRNF (MSS, Theorems 7 and 8). While the lexicographic belief systems formulation of SIRE has a clear decision-theoretic character, the corresponding formulation of normal form sequential equilibrium does not (see footnote 20). Since any normal form information set for a player is also a strategic independence for that player, any SIRE is a normal form sequential equilibrium.¹⁶

5. Decision-Theoretic Foundations of Properness and SIRE

In this section, we begin by briefly reviewing Blume, Brandenburger, and Dekel [2]’s notion of lexicographic probability systems (LPSs). We then recall their characterization of properness in terms of an ordering induced by an LPS.¹⁷ Finally, we show how that order can be modified to yield a characterization of SIRE.

5.1. Lexicographic Probability Systems

Consider a finite state space Ω . In a game theoretic context, the appropriate choice for Ω is S_i , S_{-i} , or S . For example, the state space when describing player i ’s beliefs

¹⁶Since not every strategic independence is a normal form information set, the converse can fail. Following Marx and Swinkels [8], say that a game satisfies *transference of decision-maker indifference* if, for all i and any pair of strategy profiles (s_i, s_{-i}) and (t_i, s_{-i}) , $\pi_i(s_i, s_{-i}) = \pi_i(t_i, s_{-i})$ implies $\pi_j(s_i, s_{-i}) = \pi_j(t_i, s_{-i})$ for all $j \in N$. Every strategic independence for a player is also a normal form information set for that player if and only if (S, π) satisfies transference of decision-maker indifference. Moreover, if the game satisfies transference of decision-maker indifference, strategic independence respecting equilibrium and normal form sequential equilibrium coincide.

¹⁷Blume, Brandenburger, and Dekel [1] provide an axiomatic characterization of decision making that yields a subjective expected utility theory based on LPSs. Myerson’s [10] notion of a *conditional probability system* (which is equivalent to the notion of a lexicographic conditional probability system, see footnote 21) is an alternative description of beliefs about out-of-equilibrium play.

about opponents' play is the space of strategy choices for the other players, S_{-i} . Where convenient, we will define the concepts for an arbitrary state space Ω .

Definition 6. A **Lexicographic Probability System (LPS)** on Ω is a K -tuple $\rho = (\rho^0, \dots, \rho^{K-1})$, for some integer K , of probability distributions on Ω .

Blume, Brandenburger, and Dekel [2, page 82] interpret an LPS ρ as follows: “The first component of the LPS can be thought of as representing the player’s primary theory of how the game will be played, the second component the player’s secondary theory, and so on.” In a Nash equilibrium, players’ primary theories about the play of the game will be correct and so ρ_i^0 also describes player i ’s behavior.

Each player i has an LPS ρ_{-i} describing his or her beliefs about S_{-i} . A collection of lexicographic probability systems, one for each player, is denoted $(\rho_{-1}, \dots, \rho_{-N})$. We (like Blume, Brandenburger, and Dekel [2]) impose three restrictions on the lexicographic probability systems held by players:

1. **Common Prior Assumption:** There exists an LPS ρ on S such that for all i , ρ_{-i} is the marginal on S_{-i} of ρ .¹⁸
2. **Strong Independence:** There exists a sequence of vectors $\{r(n)\}$, with $r(n) \equiv (r^1(n), \dots, r^{K-1}(n)) \in (0, 1)^{K-1}$ and $r(n) \rightarrow 0$, such that the probability distribution $r(n) \square \rho$ is a product distribution for all n , where $r \square \rho \equiv (1 - r^1)\rho^0 + r^1[(1 - r^2)\rho^1 + r^2[(1 - r^3)\rho^2 + r^3[\dots + r^{K-2}[(1 - r^{K-1})\rho^{K-2} + r^{K-1}\rho^{K-1}]]]]$ for $r \in (0, 1)^{K-1}$.
3. **Full Support:** For all i and $s_{-i} \in S_{-i}$, there exists κ such that $\rho_{-i}^\kappa(s_{-i}) > 0$.

The first condition is the usual requirement that different players have the same beliefs about the behavior of other players. We let ρ_i denote the lexicographic belief system that players other than i hold on S_i . The second condition ensures that player i believes that the other players are independently choosing strategies.¹⁹ The third condition ensures that a player can evaluate the relative likelihood of any two strategy profiles chosen by the other players.

Every LPS induces a “more likely than” ordering:

¹⁸The marginal of an LPS $(\rho^0, \dots, \rho^{K-1})$ is the LPS whose κ^{th} probability distribution is the marginal of ρ^κ , $\kappa = 0, \dots, K - 1$.

¹⁹Blume, Brandenburger, and Dekel [2, Proposition 1] show that as $n \rightarrow \infty$, and hence $r(n) \rightarrow 0$, the sequence of probability distributions $\{r(n) \square \rho\}$ “captures” the hierarchy of beliefs described by the LPS, in the sense that strategies are ranked the same by the LPS ρ and expected payoffs under the sequence of probability distributions $\{r(n) \square \rho\}_{n=1}^\infty$ (see Proposition 4 below).

Definition 7. Given an LPS ρ on Ω and $\omega, \omega' \in \Omega$, write $\omega \geq_\rho \omega'$ if

$$\min\{\kappa : \rho^\kappa(\omega) > 0\} \leq \min\{\kappa : \rho^\kappa(\omega') > 0\}.$$

The order \geq_ρ captures the ranking on states induced by the order in which these states appear in the levels of the belief systems. As usual, we have $\omega >_\rho \omega'$ if $\omega \geq_\rho \omega'$ holds but $\omega' \geq_\rho \omega$ does not; and $\omega =_\rho \omega'$ if both $\omega \geq_\rho \omega'$ and $\omega' \geq_\rho \omega$ hold. The order $>_\rho$ is a complete and transitive order on Ω . Loosely, if $\omega >_\rho \omega'$, then ω is “infinitely more likely” than ω' under ρ .

Finally, given a lexicographic probability system, the number $k(r_i, s_i)$, identifies the first level in the beliefs ρ_{-i} at which strategies r_i and s_i receive different payoffs, with $k(r_i, s_i) = K$ if r_i and s_i agree at all levels.

These preliminaries in hand, we now turn to the respective orders characterizing proper equilibrium and SIRE.

5.2. Proper Equilibrium

The key to the LPS characterization of proper is the following order generated from the LPSs and payoffs:

Definition 8. Given a lexicographic probability system ρ , the lexicographic order \succeq_L on S_i is given by, for $r_i, s_i \in S_i$,

1. $r_i \succ_L s_i$ if $k(r_i, s_i) < K$ and, for $\kappa = k(r_i, s_i)$,

$$E\{\pi_i(r_i, \rho_{-i}^\kappa)\} > E\{\pi_i(s_i, \rho_{-i}^\kappa)\},$$

2. $r_i \sim_L s_i$ if $k(r_i, s_i) = K$, and

3. $r_i \succeq_L s_i$ if $r_i \succ_L s_i$ or $r_i \sim_L s_i$.

Note that \succeq_L is a complete and transitive order.

When comparing strategies r_i and s_i for player i , the lexicographic order seeks the first level in the beliefs ρ_{-i} at which r_i and s_i receive different payoffs, and ranks the higher payoff strategy (at this level) ahead of the other. Blume, Brandenburger, and Dekel [2] prove:

Proposition 2. The strategy profile σ is proper if and only if there is lexicographic probability system ρ with $\rho^0 = \sigma$ that satisfies the common prior assumption, strong independence, full support, and

$$r_i \succ_L s_i \Rightarrow r_i >_{\rho_i} s_i.$$

5.3. Strategic Independence Respecting Equilibrium

The LPS characterization of SIRE requires only a small modification to the lexicographic order used to characterize proper in Proposition 2.

Definition 9. Given a lexicographic probability system ρ , the **structural (partial) order** \succeq_S on S_i is given by $r_i \succeq_S s_i$ if

1. $r_i \succeq_L s_i$, and
2. for all $\kappa < k(r_i, s_i)$, r_i and s_i agree for player i on the support of ρ_{-i}^κ .

The structural order, like the lexicographic order, seeks the first level at which r_i and s_i receive different payoffs. However, the structural order then ranks the two strategies only if all preceding indifferences are structural, meaning that the indifferences are created by payoff ties in the normal form and hence would hold for any possible opponent strategy. As a result, \succeq_S in general will only be a partial order (unlike \succeq_L). In particular, if two strategies r_i and s_i have equal expected payoffs according to ρ_{-i}^0 but are not structurally indifferent on the support of ρ_{-i}^0 , then they are not comparable under \succeq_S . Conversely, if two strategies r_i and s_i are not comparable under \succeq_S , then for some k , the expected payoffs to r_i and s_i are equal under all ρ_{-i}^κ for $\kappa \leq k$, while r_i and s_i are indifferent but not structurally indifferent on the support of ρ_{-i}^k .

The structural order provides an intuitive characterization of SIRE:²⁰

Proposition 3. The strategy profile σ is a SIRE if and only if there is lexicographic probability system ρ with $\rho^0 = \sigma$ that satisfies the common prior assumption, strong independence, full support, and

$$r_i \succ_S s_i \Rightarrow r_i >_{\rho_i} s_i.$$

Proof. We first formulate SIRE in terms of LPSs. Let $\rho_i |_X$ denote the conditional distribution $\rho_i^k(\cdot | X_i)$, where $k = \min\{\kappa : \rho_i^\kappa(X_i) > 0\}$, and similarly for $\rho_{-i} |_X$. It

²⁰There is a similar characterization of normal form sequential equilibrium. The strategy profile σ is a normal form sequential equilibrium if and only if there is an LPS ρ with $\rho^0 = \sigma$ that satisfies the common prior assumption, strong independence, full support, and

$$r_i \succ_S^* s_i \Rightarrow r_i >_{\rho_i} s_i,$$

where $r_i \succeq_S^* s_i$ iff $r_i \succeq_L s_i$ and, for all $\kappa < k(r_i, s_i)$, r_i and s_i agree for all players on the support of ρ_{-i}^κ . The difficulty with this characterization is that the order \succeq_S^* requires player i to pay attention to the payoff structure of the other players, something that cannot be justified on purely decision-theoretic grounds. Moreover, there is no trembles-based characterization of \succeq_S^* analogous to the characterizations of \succeq_L and \succeq_S discussed in the next section.

is immediate from the definitions that a strategy profile σ is a SIRE if and only if there is a lexicographic probability system ρ with $\rho^0 = \sigma$ that satisfies the common prior assumption, strong independence, full support, and for every player i and every strategic independence for player i , X , $\rho_i | X$ is a best reply to $\rho_{-i} | X$ on X_i . We say that such an LPS is a SIRE, or that it supports σ as a SIRE.

(i) $(r_i \succ_S s_i \Rightarrow r_i \succ_{\rho_i} s_i) \Rightarrow \rho$ is SIRE.

The proof is by contradiction. Suppose X is a strategic independence for player i with $\rho_i | X$ not a best reply to $\rho_{-i} | X$. Let $k \equiv \min\{\kappa : \rho_i^\kappa(X_i) > 0\}$ and $k' \equiv \min\{\kappa : \rho_{-i}^\kappa(X_{-i}) > 0\}$. Then there exists $s_i, t_i \in X_i$, with $\rho_i^k(s_i) > 0$ and $\pi_i(s_i, \rho_{-i}^{k'} | X) < \pi_i(t_i, \rho_{-i}^{k'} | X)$. Since X is a strategic independence, there exists $r_i \in X_i$ agreeing for i with t_i on X_{-i} and agreeing for i with s_i on $S_{-i} \setminus X_{-i}$. So

$$\begin{aligned} E\{\pi_i(s_i, \rho_{-i}^{k'})\} &= \rho_{-i}^{k'}(X_{-i})\pi_i(s_i, \rho_{-i}^{k'} | X) + (1 - \rho_{-i}^{k'}(X_{-i}))\pi_i(s_i, \rho_{-i}^{k'} | S \setminus X) \\ &< \rho_{-i}^{k'}(X_{-i})\pi_i(r_i, \rho_{-i}^{k'} | X) + (1 - \rho_{-i}^{k'}(X_{-i}))\pi_i(r_i, \rho_{-i}^{k'} | S \setminus X) = E\{\pi_i(r_i, \rho_{-i}^{k'})\}, \end{aligned}$$

where $S \setminus X$ denotes the distribution conditional on $S_{-i} \setminus X_{-i}$ and the expressions are well defined when $\rho_{-i}^{k'}(S_{-i} \setminus X_{-i}) = 0$ (since $\rho_{-i}^{k'}(X_{-i}) = 1$ in that case). Now, r_i and s_i agree for i on $S_{-i} \setminus X_{-i}$ and so $r_i \succ_S s_i$. But then $r_i \succ_{\rho_i} s_i$, which is a contradiction (since $\rho_i^k(s_i) > 0$).

(ii) ρ is SIRE $\Rightarrow (r_i \succ_S s_i \Rightarrow r_i \succ_{\rho_i} s_i)$.

Fix r_i, s_i , and define X_{-i} to be the smallest subset of S_{-i} with the property that r_i and s_i agree for i on $S_{-i} \setminus X_{-i}$. Now, $r_i \succ_S s_i \Rightarrow \pi_i(r_i, \rho_{-i}^{k'}) > \pi_i(s_i, \rho_{-i}^{k'})$, where $k' \equiv \min\{\kappa : \rho_{-i}^\kappa(X_{-i}) > 0\}$. Since $\{r_i, s_i\} \times X_{-i}$ is a strategic independence, $\rho_i^k(s_i) = 0$, where $k = \min\{\kappa : \rho_i^\kappa(\{r_i, s_i\}) > 0\}$, and so $r_i \succ_{\rho_i} s_i$. \square

Thus, in a SIRE, strategies that rank higher in the payoff order \succeq_S also rank higher in the behavior order \succ_{ρ_i} . That is, players believe that player i is infinitely more likely to play r_i than s_i if player i 's preferences (as described by \succeq_S) rank r_i ahead of s_i .

Propositions 2 and 3 show that the decision-theoretic difference between SIRE and proper consists precisely of the difference between \succeq_S and \succeq_L .²¹ Suppose that player i finds himself indifferent between strategies r_i and s_i , given that the opponents' play is described by ρ_{-i} . SIRE appeals to the next level in the belief hierarchy ρ_{-i} in order to rank r_i and s_i if and only if r_i and s_i are structurally indifferent. Proper equilibrium

²¹One difference between the orders \succeq_S and \succeq_L is reflected in the ability, when using \succeq_S , to work with an LPS whose various levels have disjoint supports. A lexicographic probability system is a lexicographic *conditional* probability system if the probability distributions $(\rho^0, \dots, \rho^{K-1})$ have pairwise disjoint supports. An LPS ρ is a SIRE if and only if the lexicographic conditional probability system $\hat{\rho}$ defined by $\hat{\rho}_{-i}^\kappa(\cdot) = \rho_{-i}^\kappa(\cdot | S_{-i} \setminus \cup_{\ell < \kappa} \text{supp}(\rho_{-i}^\ell))$ for all κ is a SIRE. A restriction to lexicographic conditional probability systems is thus without loss of generality for the order \succeq_S and SIRE. The same is not true for \succeq_L and proper equilibria (Blume, Brandenburger, and Dekel [2, p. 89]).

		2		
		L	C	R
1	A	1,2	2,1	2,1
	B	2,2	2,0	2,0
	C	0,0	4,1	0,3
	D	0,0	1,3	3,1

Figure 2: $(.5C + .5D, .5C + .5R)$ is a SIRE but is not proper

		2				
		.25 ⁰	.75 ⁰	.75 ¹	.25 ¹	1 ²
		α	β	γ	δ	θ
1	.25 ⁰ A	7,7	5,5	2,2	0,0	11, -1
	.25 ⁰ B	4,4	6,6	2,2	0,0	9, -2
	.25 ⁰ C	4,4	6,6	1,1	3,3	10, -3
	.25 ⁰ D	7,7	5,5	1,1	3,3	8, -4

Figure 3: A SIRE that is not proper. Superscripts indicate levels of the LPS.

always appeals to the next level. As a result, we have $r_i \succeq_S s_i \Rightarrow r_i \succeq_L s_i$. Proper equilibrium thus imposes more stringent requirements than does SIRE.

Since \succeq_L appeals any indifferences between strategies at a given level of beliefs to a higher level, while \succeq_S appeals only indifferences that are structural, these orders, and hence SIRE and proper equilibrium, will coincide if all payoff ties arise out of structural indifferences. In particular, if ρ is an LPS supporting ρ^0 as a SIRE and the support of ρ^k is a singleton for each k , then ρ supports ρ^0 as a proper equilibrium.

Figure 2 shows that a SIRE need not be proper. Since B weakly dominates $\frac{1}{2}C + \frac{1}{2}D$, the profile $(\frac{1}{2}C + \frac{1}{2}D, \frac{1}{2}C + \frac{1}{2}R)$ cannot be a proper equilibrium. However, normal form information sets and strategic independences include $\{C, D\} \times \{C, R\}$ for both players. Letting $\rho_1^0(C) = \rho_1^0(D) = \rho_2^0(C) = \rho_2^0(R) = 0.5$, $\rho_1^1(B) = \rho_2^1(L) = 1$, and $\rho_1^2(A) = 1$ then gives best replies on all strategic independences and hence ensures that $(\frac{1}{2}C + \frac{1}{2}D, \frac{1}{2}C + \frac{1}{2}R)$ is a SIRE. In particular, ρ_2^0 makes player 1 indifferent between B , C , and D , but the indifference is not structural and hence is not appealed to ρ_2^1 .

The SIRE in Figure 2 is not even a normal form perfect equilibrium, since it attaches probability to the dominated strategy $\frac{1}{2}C + \frac{1}{2}D$.²² However, Figure 2 is special in that the payoffs to A , B , C , and D all fortuitously equal 2 when player 2 plays $\frac{1}{2}C + \frac{1}{2}R$. That is, the game is not robust to tie-preserving perturbations of payoffs.

An example in which SIRE and proper do not coincide that cannot be destroyed by tie-preserving payoff perturbations is given in Figure 3.²³ Let ρ be as indicated, so that $\rho_1^0(A) = \rho_1^0(B) = \rho_1^0(C) = \rho_1^0(D) = .25$, $\rho_2^0(\alpha) = .25$, $\rho_2^0(\beta) = 0.75$, $\rho_2^1(\gamma) = .75$, $\rho_2^1(\delta) = 0.25$, and $\rho_2^2(\theta) = 1$. Then the lexicographic probability system ρ is a SIRE. However, there exists no lexicographic probability system that will support this outcome as a proper equilibrium. To verify this, consider the specification of ρ_2^1 that would be required for properness. Since player 1 is indifferent between strategies A and D against ρ_2^0 and A earns a higher payoff than D against ρ_2^1 , ρ_2^1 must attach a probability to δ higher than .25 (otherwise A surely earns a higher payoff than D , precluding indifference). But player 1 is also indifferent between strategies B and C , requiring ρ_2^1 to attach a probability to γ higher than .75, a contradiction.

6. Tremble-Based Characterizations of \succeq_L and \succeq_S

The characterizations of properness and sequentiality in Sections 3 and 4 are in terms of perturbed strategy profiles and the structural features of games (information sets in the extensive form and structural indifferences in the normal form). In contrast, the characterizations in Sections 5.2 and 5.3 are not in terms of structural features of games, but do use LPSs. In this section, we provide tremble-based characterizations that do not use the structural features of games. The original definition of properness is, of course, one such characterization of properness. The substantive result here is the characterization of \succeq_S in terms of trembles. This is then compared with a tremble-based characterization of \succeq_L .

First, note the following equivalence between lexicographic probability systems and probability sequences: Given an LPS satisfying common prior, strong independence, and full support, $\{P^n\} = \{r(n) \square \rho\}$ is a conditionally convergent probability sequence (where $r(n)$ is the sequence of vectors from strong independence). Moreover, $P_i^n = r(n) \square \rho_i$, so that, for $s_i \in S_i$ and any $k \leq \min\{\kappa : \rho^\kappa(s_i) > 0\}$, $\rho_i^k(s_i) = \lim P_i^n(s_i) / \prod_{\kappa=1}^k r^\kappa(n)$

²²It is easy to show that a SIRE cannot attach positive weight to a *pure* strategy that is weakly dominated by another pure strategy, though it can attach positive probability to a pure strategy dominated by a mixed strategy.

²³This example is generic in the sense that any perturbations in payoffs that preserve the ties in player 1's payoffs yields a nearby SIRE that is not a proper equilibrium. The ties in player 2's payoffs appear for simplicity; they are not important to the example. There exist nearby specifications of player 2's payoffs that feature no ties and again yield a SIRE that is not proper. This example is then only nongeneric if one considers *all* perturbations in payoffs, including those that disrupt ties.

(where we avoid division by zero by defining $\prod_{\kappa=1}^0 r^\kappa(n) \equiv 1$). In particular, $\rho^0 = \lim P^n$. Thus, $\rho_i |_X$ is a best reply to $\rho_{-i} |_X$ on X_i if and only if the distribution $\lim P_i^n(\cdot | X_i)$ is a best reply to the distribution $\lim P_{-i}^n(\cdot | X_{-i})$ on X_i . Conversely, given a sequence of completely mixed behavior strategy profiles and so, trivially, a sequence of completely mixed PRNF strategy profiles, $\{P^n\}$, there exists an LPS ρ such that a subsequence $\{P^m\}$ of $\{P^n\}$ can be written as $P^m = r(m) \square \rho$ for a sequence of vectors $\{r(m)\} \subseteq (0, 1)^{K-1}$ with $r(m) \rightarrow 0$ (see Blume, Brandenburger, and Dekel [2, Proposition 2]).

Definition 10. The LPS ρ and the probability sequence $\{P^n\}$ are **tail equivalent** if there exists n^* , such that for $n > n^*$, $P^n = r(n) \square \rho$ for some $r(n) \in (0, 1)^{K-1}$ with $r(n) \rightarrow 0$.

The LPS ρ and the probability sequence P^n are **limit equivalent** if for all $i \in N$, $s_{-i}, t_{-i} \in S_{-i}$,

$$s_{-i} >_{\rho_{-i}} t_{-i} \iff \lim_{n \rightarrow \infty} \frac{P_{-i}^n(t_{-i})}{P_{-i}^n(s_{-i})} = 0$$

and

$$P_{-i}^n |_{A^\kappa} \rightarrow \rho_{-i}^\kappa |_{A^\kappa}, \quad \forall \kappa,$$

where $A^0 = \text{supp}(\rho_{-i}^0)$ and $A^\kappa = \text{supp}(\rho_{-i}^\kappa) \setminus \cup_{\ell < \kappa} A^\ell$.

Limit equivalence is less demanding than tail equivalence. For example, the sequence $\{\frac{1}{2} + \frac{1}{n}, \frac{1}{2} - \frac{1}{n}\}_{n=0}^\infty$ is limit, but not tail, equivalent to the constant sequence $\{\frac{1}{2}, \frac{1}{2}\}$. More generally, if the probability sequence $\{P^n\}$ is tail equivalent to ρ , then it is limit equivalent to ρ and there exists n^* , such that for $n > n^*$,

$$P_{-i}^n |_{A^\kappa} = \rho_{-i}^\kappa |_{A^\kappa}, \quad \forall \kappa.$$

Given an LPS with pairwise disjoint supports (i.e., a lexicographic *conditional* probability system, see footnote 21), the payoffs to any two strategies, r_i and s_i , must be ranked the same way by P_{-i}^n for all sufficiently large n and for every probability sequence that is tail equivalent to ρ . Hence, if r_i receives a higher payoff than s_i along *some* probability sequence that is tail equivalent to ρ , then r_i receives a higher payoff than s_i along *every* probability sequence that is tail equivalent to ρ . This is also true when the LPS does not have pairwise disjoint supports (Proposition 4 below). The same does not hold for limit equivalence; it is easy to find two probability sequences that are limit equivalent to the same LPS but rank strategies differently.²⁴

²⁴In Figure 3, the sequences $\{\frac{1}{4} - \frac{1}{n}, \frac{3}{4} - \frac{1}{n}, \frac{3}{4n} - \frac{1}{n^2}, \frac{1}{4n}, \frac{1}{n^2}\}$ and $\{\frac{1}{4} - \frac{1}{n}, \frac{3}{4} - \frac{1}{n}, \frac{3}{4n}, \frac{1}{4n} - \frac{1}{n^2}, \frac{1}{n^2}\}$ are both limit equivalent to ρ_2 , but the first gives a higher payoff to C than to A (for large n) while the second gives a higher payoff to A than to C .

The lexicographic order induced by a lexicographic probability system ρ ranks r_i ahead of s_i if and only if r_i receives a higher payoff than s_i along probability sequences that are tail equivalent to ρ :

Proposition 4. *Suppose ρ is an LPS satisfying the common prior, strong independence, and full support assumptions. Then, for all i , $r_i \succ_L s_i$ if and only if for every probability sequence $\{P^n\}$ tail equivalent to ρ , there exists n^* such that for all $n > n^*$,*

$$E\{\pi_i(r_i, P_{-i}^n)\} > E\{\pi_i(s_i, P_{-i}^n)\}.$$

Proof: The proof of [2, Proposition 1] applies here with the modification that (using their notation) n^* is chosen so that $n > n^*$ implies $r^k(n) < r^* \forall n > n^*$, where r^* solves $(1 - r^*)B + r^*W > 0$. \square

The structural order can be similarly characterized, but the relevant probability sequences now consist of all limit equivalent sequences. The appendix proves the following:

Proposition 5. *Suppose ρ is an LPS satisfying the common prior, strong independence, and full support assumptions. Then for all i , $r_i \succ_S s_i$ if and only if for every probability sequence $\{P^n\}$ limit equivalent to ρ , there exists n^* such that for all $n > n^*$,*

$$E\{\pi_i(r_i, P_{-i}^n)\} > E\{\pi_i(s_i, P_{-i}^n)\}.$$

Propositions 4 and 5 describe the difference between tail and limit equivalence, and correspondingly between the lexicographic and structural orders, without reference to structural features of the game. Fix an LPS ρ and suppose that every tail-equivalent probability sequence ranks r_i ahead of s_i . Then $r_i \succ_L s_i$. However, $r_i \succ_S s_i$ may fail to hold, as there may be limit-equivalent probability sequences that either fail to rank r_i and s_i or disagree in their ranking.²⁵ The structural order thus requires more stringent conditions than the lexicographic order to rank strategies and the structural order can decline to rank strategies that are ranked under the lexicographic order.

We can illustrate this difference by returning to Figure 2. Consider the specification $\rho_2^0(C) = \rho_2^0(R) = .5$ and $\rho_2^1(L) = 1$. How should player 1's strategies be ranked? Against ρ_2^0 , player 1 is indifferent between A and B . Properness then demands that the decision between A and B be appealed to ρ_2^1 , which suffices to rank B ahead of A . The set $\{A, B\} \times \{L\}$ is a strategic independence for player 1, and SIRE also ranks B ahead of A . It is obvious that for any $\{P_2^n\}$ converging to ρ_2 , B receives a higher payoff than A along every term of the sequence.

²⁵A probability sequence fails to rank r_i and s_i if each strategy earns a higher payoff than the other for infinitely many terms of the sequence.

Player 1 is also indifferent between strategies B , C and D given ρ_2^0 . Properness again appeals to ρ_2^1 , ranking B ahead of C and D . A similar ranking is given along every term of any sequence that is tail equivalent to ρ . The key here is that tail equivalence preserves any indifferences that appear at any level of beliefs in ρ . In particular, a tail equivalent probability sequence is simply a collection of convex combinations of the distributions ρ_2^n , $n = 1, \dots, k-1$, with the weight on ρ_2^k becoming arbitrarily high relative to ρ_2^{k+1} . Using ρ to evaluate strategies according to the order \succ_L is then equivalent to looking at sequences of tail-equivalent strategies.

Because player 1 has no information set or strategic independence that includes B and either C or D , and in which player 2's strategy set is $\{L\}$, the order \succ_S and hence SIRE does not rank B , C and D . It is easy to find sequences that are limit (but not tail) equivalent to ρ in which either C or D gets a higher payoff than B along every element of the sequence. More generally, when must SIRE rank B ahead of C ? A necessary condition must be that along every converging sequence of perturbed strategies, B does strictly better than C , since otherwise we can find limits in which C does as least as well as B and hence the SIRE need not rank B ahead of C . The proof of Proposition 5 involves showing that this condition is sufficient as well, by showing that if B fares better than C along every limit-equivalent sequence, then all indifferences between B and C must be structural indifferences, causing \succ_S to rank B ahead of C .

7. Conclusion

In this paper, we provide three results on the relationship between properness and sequential-in-every-tree. First, we show that properness is equivalent to quasi-perfect-in-every-tree. Since quasi-perfection is optimality along the sequence of perturbed strategies, while sequentiality is optimality in the limit, the distinction between properness and sequential-in-every-tree can be similarly phrased. Second, we provide a lexicographic probability system characterization of SIRE, the normal form implementation of sequential rationality. This characterization uses the structural order on a player's strategy space, denoted \succeq_S . Blume, Brandenburger, and Dekel [2] have a similar characterization of properness, based on the lexicographic order, \succ_L . The distinction between \succ_S and \succ_L describes the difference in the decision theories that underlie sequential rationality and properness. Third, we give tremble-based characterizations of the orders \succ_S and \succ_L that do not involve structural features of the game, such as information sets or strategic independences.

Appendix: Proof of Proposition 5

The result is trivial if r_i and s_i agree for i on S_{-i} . So suppose not. Let k be the largest

index satisfying: r_i and s_i do not agree for i on $\text{supp}(\rho_{-i}^\kappa)$ for $\kappa < k$. Let $A^0 = \text{supp}(\rho_{-i}^0)$, $A^\kappa = \text{supp}(\rho_{-i}^\kappa) \setminus \cup_{\ell < \kappa} A^\ell$, for all κ . Define $\Delta\pi_i(s_{-i}) \equiv \pi_i(r_i, s_{-i}) - \pi_i(s_i, s_{-i})$.

(\Rightarrow) Suppose $\{P^n\}$ is limit equivalent to ρ and $r_i \succ_S s_i$. Then r_i has a strictly higher expected payoff than s_i under ρ_{-i}^k . Let $C = \sum \rho_{-i}^k(s_{-i}) \Delta\pi_i(s_{-i})$ and $B = \max |\Delta\pi_i(s_{-i})|$. Note that $A^k \neq \emptyset$, so that $\epsilon(n) = P_{-i}^n(A^k)/\rho_{-i}^k(A^k)$ is well defined. Since P^n is completely mixed, $\epsilon(n) \neq 0$. Choose n' so that for $n > n'$ and for all $s_i \notin \cup_{\kappa \leq k} \text{supp}(\rho_{-i}^\kappa)$, $P_{-i}^n(s_{-i}) < C\epsilon(n)/(3B |S_{-i}|)$. Since P is limit equivalent to ρ , there is an n'' such that for $n > n''$,

$$\left| \sum_{s_{-i} \in A^k} \rho_{-i}^k(s_{-i}) \Delta\pi_i(s_{-i}) - (\epsilon(n))^{-1} \sum_{s_{-i} \in A^k} P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) \right| < C/3.$$

Set $n^* = \max\{n', n''\}$. Since r_i and s_i agree for i on $\text{supp}(\rho_{-i}^\kappa)$ for $\kappa < k$,

$$\begin{aligned} \sum P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) &= \sum_{s_{-i} \notin \cup_{\kappa < k} \text{supp}(\rho_{-i}^\kappa)} P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) \\ &= \sum_{s_{-i} \in A^k} P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) + \sum_{s_{-i} \notin \cup_{\kappa \leq k} \text{supp}(\rho_{-i}^\kappa)} P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) \\ &> \sum_{s_{-i} \in A^k} P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) - C\epsilon(n)/3 \\ &> \epsilon(n)(C - C/3 - C/3) = \epsilon(n)C/3 > 0. \end{aligned}$$

(\Leftarrow) Suppose for all $\{P^n\}$ limit equivalent to ρ and for n sufficiently large,

$$\sum P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) > 0. \quad (6)$$

We suppose $r_i \succ_S s_i$ does not hold and derive a contradiction. If $r_i \succ_S s_i$ does not hold, then:

$$\begin{aligned} \sum \rho_{-i}^\kappa(s_{-i}) \Delta\pi_i(s_{-i}) &= 0 \quad \text{for all} \\ \sum \rho_{-i}^k(s_{-i}) \Delta\pi_i(s_{-i}) &\leq 0. \end{aligned} \quad (7)$$

The definition of k and (6) implies:

$$\sum_{s_{-i} \in A^k} P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) + \sum_{s_{-i} \notin \cup_{\kappa \leq k} \text{supp}(\rho_{-i}^\kappa)} P_{-i}^n(s_{-i}) \Delta\pi_i(s_{-i}) > 0.$$

Dividing by $P_{-i}^n(A^k)$ and taking limits yields

$$\sum_{s_{-i} \in A^k} \rho_{-i}^k(s_{-i}) \Delta\pi_i(s_{-i}) \geq 0.$$

Combining with (7) yields

$$\sum_{s_{-i} \in A^k} \rho_{-i}^k(s_{-i}) \Delta \pi_i(s_{-i}) = 0. \quad (8)$$

We now argue that there exists a probability sequence Q that is limit equivalent to ρ but reverses the inequality in (6), which is a contradiction.

Define $k_j(s_j) \equiv \min\{\kappa : \rho_j^\kappa(s_j) > 0\}$ and $k_{-i}(s_{-i}) \equiv \min\{\kappa : \rho_{-i}^\kappa(s_{-i}) > 0\}$. Note that $k = k_{-i}(s_{-i})$, $\forall s_{-i} \in A^k$. Since ρ is strongly independent, there exists $r(n) \in (0, 1)^{K-1}$, $r(n) \rightarrow 0$ as $n \rightarrow \infty$, such that $r(n) \square \rho = \prod_i (r(n) \square \rho)_i = \prod_i (r(n) \square \rho_i)$. Fixing $s_{-i} \in S_{-i}$, and letting $k_j = k_j(s_j)$, $k_{-i} = k_{-i}(s_{-i})$, we have

$$\begin{aligned} & r^1(n) \cdots r^{k_{-i}}(n) \left((1 - r^{k_{-i}+1}(n)) \rho_{-i}^{k_{-i}}(s_{-i}) + r^{k_{-i}+1}[\dots] \right) \\ &= \prod_{j \neq i} r^1(n) \cdots r^{k_j}(n) \left((1 - r^{k_j+1}(n)) \rho_j^{k_j}(s_j) + r^{k_j+1}[\dots] \right). \end{aligned}$$

(We follow the convention that $r^1(n) \cdots r^0(n) = 1$.) Dividing by $r^1(n) \cdots r^{k_{-i}}(n)$ and taking n to infinity shows that

$$\rho_{-i}^{k_{-i}}(s_{-i}) = \alpha(k^{-i}, k_{-i}) \times \prod_{j \neq i} \rho_j^{k_j}(s_j), \quad (9)$$

where k^{-i} is the vector $(k_j)_{j \neq i}$ and

$$\alpha(k^{-i}, k_{-i}) = \lim_{n \rightarrow \infty} \left[\frac{\prod_{j \neq i} r^1(n) \cdots r^{k_j}(n)}{r^1(n) \cdots r^{k_{-i}}(n)} \right] \neq 0.$$

Let $A_j^k \equiv \{s_j \in S_j : (s_j, s_{-\{i,j\}}) \in A^k \text{ for some } s_{-\{i,j\}}\}$. Define $u(s_{-i}) \equiv \alpha(k^{-i}(s_{-i}), k_{-i}(s_{-i})) \times \Delta \pi_i(s_{-i})$ and consider the function $\Phi : \prod_{j \neq i} \mathfrak{R}^{A_j^k} \rightarrow \mathfrak{R}$, given by, for $p_{-i} = (p_j)_{j \neq i}$, $p_j \in \mathfrak{R}^{A_j^k}$:

$$\Phi(p_{-i}) = \sum_{s_{-i} \in A^k} \left(\prod_{j \neq i} p_j(s_j) \right) u(s_{-i}).$$

From (8) and (9), $\Phi(p_{-i}^*) = 0$, where $p_j^*(s_j) = \rho_j^{k_j(s_j)}(s_j)$. Since $u(s_{-i}) \neq 0$ for at least one $s_{-i} \in A^k$, Φ is not identically zero on any neighborhood of p^* . Fix $\delta > 0$ and p^0 such that $|p^* - p^0| < \delta$ and $\Phi(p^0) \neq 0$. Define $p^j \equiv (p_1^*, \dots, p_j^*, p_{j+1}^0, \dots, p_N^0)$, for $j = 1, \dots, N$, so that $p^N = p^*$. Let j' be the first index j such that $\Phi(p^j) = 0$. Since $\Phi((1 - \lambda)p^{j'-1} + \lambda p^j)$ is affine in λ and $\Phi(p^{j'-1}) \neq 0$, there is a $\lambda \in \mathfrak{R}$ such that $p^\delta = (1 - \lambda)p^{j'-1} + \lambda p^j$ satisfies $\Phi(p^\delta) < 0$ and $|p^* - p^\delta| < \delta$. Since $p_j^*(s_j) \neq 0$, for small δ , $p_j^\delta(s_j) \neq 0$ for all $s_j \in A_j^k$ and j .

We now define a pseudo-LPS for player j , $\sigma_j(\cdot; \delta)$, as follows: $\sigma_j^{k_j(s_j)}(s_j; \delta) \equiv p_j^\delta(s_j)$, and $\sigma_j^\kappa(s_j; \delta) \equiv \rho_j^\kappa(s_j)$ for $\kappa \neq k_j(s_j)$. The pseudo-LPS fails to be a true LPS only because it may be that $\sum \sigma_j^\kappa(s_j; \delta) \neq 1$ for some κ .

Consider, for fixed δ , the probability sequence $\{P_j^n(\cdot; \delta)\}$ given by $P_j^n(\cdot; \delta) \equiv q_j(n; \delta) \cdot (r(n) \square \sigma_j(\cdot; \delta))$, where $q_j(n; \delta) \equiv \left(\sum_{s_j} (r(n) \square \sigma_j(s_j; \delta)) \right)^{-1} \in \mathfrak{R}_+$ is a scaling factor ($q_j(n; \delta) \rightarrow 1$ as $\delta \rightarrow 0$, and $q_j(n; \delta) = 1$ if $\sigma_j(\cdot; \delta)$ is a true LPS). We now argue that for n sufficiently large,

$$\sum P_{-i}^n(s_{-i}; \delta) \Delta \pi_i(s_{-i}) < 0. \quad (10)$$

This is clearly equivalent to:

$$\sum (r^1(n) \cdots r^k(n))^{-1} \prod_{j \neq i} (r(n) \square \sigma_j(s_j; \delta)) \Delta \pi_i(s_{-i}) < 0.$$

Using the fact that $k_j(s_j) = \min\{\kappa : \rho_j^\kappa(s_j) > 0\} = \min\{\kappa : \sigma_j^\kappa(s_j; \delta) > 0\}$, the left hand side converges to

$$\sum_{s_{-i} \in A^k} \left(\prod_{j \neq i} \sigma_j^{k_j(s_j)}(s_j; \delta) \right) \alpha(k^{-i}(s_{-i}), k_{-i}(s_{-i})) \Delta \pi_i(s_{-i}) = \Phi(p^\delta) < 0$$

as $n \rightarrow \infty$, and so (10) holds for n sufficiently large.

Let $Q_j^m = P_j^{n(m)}(\cdot; 1/m)$, where $\{n(m)\}_{m \uparrow \infty}$ is an increasing sequence with the property that (10) holds when $\delta = 1/m$ and $n = n(m)$. It is immediate that $\{Q_j^m\}$ is limit equivalent to ρ and reverses the inequality in (6). This is the desired contradiction and so $r_i \succ_S s_i$. \square

References

- [1] L. E. Blume, A. Brandenburger, and E. Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica*, 59:61–79, 1991.
- [2] L. E. Blume, A. Brandenburger, and E. Dekel. Lexicographic probabilities and equilibrium refinements. *Econometrica*, 59:81–98, 1991.
- [3] J. Hillas. On the relation between perfect equilibria in extensive form games and proper equilibria in normal form games. Mimeo, State University of New York at Stony Brook, 1996.
- [4] E. Kohlberg and J.-F. Mertens. On the strategic stability of equilibria. *Econometrica*, 54:1003–1037, 1986.

- [5] G. J. Mailath, L. Samuelson, and J. M. Swinkels. Extensive form reasoning in normal form games. *Econometrica*, 61:273–302, 1993.
- [6] G. J. Mailath, L. Samuelson, and J. M. Swinkels. Normal form structures in extensive form games. *Journal of Economic Theory*, 64:325–371, 1994.
- [7] G. J. Mailath, L. Samuelson, and J. M. Swinkels. How proper is sequential equilibrium? CARESS Working Paper #94-14R, University of Pennsylvania, January 1995.
- [8] L. M. Marx and J. M. Swinkels. Order independence for iterated weak dominance. *Games and Economic Behavior*, forthcoming.
- [9] R. B. Myerson. Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, 7:73–80, 1978.
- [10] R. B. Myerson. Multistage games with communication. *Econometrica*, 54:323–358, 1986.
- [11] E. van Damme. A relation between perfect equilibria in extensive form games and proper equilibria in normal form games. *International Journal of Game Theory*, 13:1–13, 1984.