

Common Assumption of Rationality in Games*

Adam Brandenburger
Stern School of Business
New York University
44 West Fourth Street
New York, NY 10012
abranden@stern.nyu.edu

Amanda Friedenberg
Olin School of Business
Washington University
1 Brookings Drive
St. Louis, MO 63130
friedenberg@olin.wustl.edu

First Version: January 2002
Current Version: December 2003

Abstract

A fundamental question in decision theory asks: What is rational behavior? There is a natural corresponding question in interactive decision theory: What behavior is consistent with *rationality and common assumption of rationality* (RCAR)? We seek an analysis of RCAR that is *invariant*—i.e., is unaffected by inessential transformations of the tree. In this case, we show that RCAR can be characterized by a solution concept we call a *self-admissible set*. Further, we characterize self-admissible sets for various well-known games, including Centipede and the Finitely Repeated Prisoner’s Dilemma. We also show that in the particular case of a perfect-information game satisfying the “single payoff condition,” a self-admissible set yields a Nash outcome.

1 Introduction

In decision theory, a natural question arises: “What is rational behavior?” Corresponding questions arise in interactive decision theory. Suppose players are rational and assume others to be rational—

*We are indebted to Johannes Hörner and Gus Stuart for many conversations relevant to this paper and to Pierpaolo Battigalli, Martin Osborne, and Marciano Siniscalchi for important comments. Example 2 is due to Ken Cortis. Drew Fudenberg pointed us to an example similar to Example E1. Geir Asheim, Chris Avery, Oliver Board, Giacomo Bonanno, Konrad Grabiszewski, Elon Kohlberg, Jeroen Swinkels, and participants of the 14th Italian Meeting on Game Theory and Applications (Ischia, July 2001), the 2001 Summer Workshop of the Stanford Institute for Theoretical Economics (August 2001), the 56th European Meeting of the Econometric Society (August 2001), the North American Summer Meetings of the Econometric Society (June 2002), and seminars at Harvard, Northwestern, Penn State, Pittsburgh, Rochester, Stanford, Washington, and Yale Universities provided important input. Brandenburger gratefully acknowledges support from Harvard Business School and the Stern School of Business. Friedenberg thanks the CMS-EMS at Northwestern University, the Department of Economics at Yale University, and the Olin School of Business for their support. carg-12-22-03

where does this lead? Now, what if players also assume this? And so on. Proceeding along these lines, it is natural to ask: What behavior is consistent with *rationality and common assumption of rationality (RCAR)*?¹

Bernheim [12, 1984] and Pearce [36, 1984] argued that RCAR leads to the set of *rationalizable* strategies. In a two-player game, this is equal to the set of strategies that survive iterated strong dominance.²

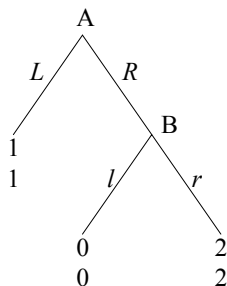


Figure 1

It might be tempting to conclude with rationalizability. But while rationalizability may seem natural in a simultaneous-move game, it lacks the same appeal in an extensive-form game. For instance, in the game in Figure 1, $\{L, R\} \times \{l, r\}$ is the set of rationalizable strategies. Yet, if Bob is rational, then surely, conditional upon his node being reached, he will play r . Thus, if Ann assumes Bob is rational, she should begin the game with a hypothesis that places probability 1 on Bob’s playing r . If so, Ann will rationally play R and not L .

Such considerations led Pearce to define *extensive-form rationalizability (EFR)*, which was later simplified by Battigalli [9, 1997]. Shimoji-Watson [42, 1998] showed EFR is equivalent to iterated deletion of conditionally dominated strategies—the natural extensive-form analog to iterated strong dominance on the matrix. In Figure 1, Bob’s strategy l is conditionally dominated by r , at his node. On the second round, Ann’s strategy L is then conditionally dominated by R . We get the desired outcome (R, r) .³

Battigalli-Siniscalchi [10, 2002] provide epistemic conditions for EFR. They demonstrate that EFR is characterized by an RCAR-like condition, provided the latter is posed in a “complete” type structure. In effect, the completeness condition requires the players to optimize with respect to ‘rational’ hypotheses—even at information sets that are precluded by the EFR strategies.⁴

¹We use the term “assumption” rather than the more customary “belief,” because it will fit better with our formal treatment later. For now, the reader can substitute “belief,” if preferred.

²Under Bernheim’s and Pearce’s original definition, the rationalizable set may be strictly contained in the iteratively not-strongly-dominated set. This is because of an independence assumption they make. Recent definitions (see, e.g., Osborne-Rubinstein [35, 1994]) allow for correlation; in this case, the two sets are equal.

³More generally, EFR gives the backward-induction outcome in perfect-information games satisfying certain conditions on payoffs. Battigalli [9, 1997] showed the result for games with “no relevant ties.” Brandenburger-Friedenberg [16, 2003] contains a proof for games satisfying the “single payoff condition” (Definition 8 below).

⁴More formally, a complete type structure is one that contains all possible types of the players. See Sections 8 and 9 below. Apart from completeness, the precise condition Battigalli-Siniscalchi [10, 2002] impose is “rationality and common strong belief of rationality.” See their paper for details.

Yet, EFR has one undesirable property. The solution may change when the game tree is changed in an inessential manner. The game in Figure 2 is the opening example transformed by “addition of decision nodes.” Here, the entire strategy set $\{L, R\} \times \{l, r\}$ is EFR. So we see that EFR may depend on irrelevant features of the tree. More generally, we maintain that decision theory—and so interactive decision theory—should be invariant to presentation effects. (In Section 2, we’ll return to the issue of invariance and say more about why we ask for this property.)

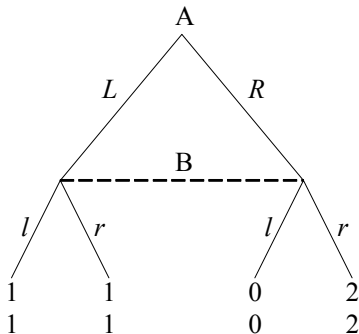


Figure 2

Given this, neither strategic-form rationalizability nor extensive-form rationalizability are fully satisfactory answers to the question of what RCAR in a game means. The goal of this paper is to find such an answer—that is, to find a solution concept satisfying the following two requirements:

- **RCAR** We seek an analysis of rationality and common assumption of rationality.
- **Invariance** We seek an analysis that is invariant to how the tree is presented.

The answer we’ll get is a new solution concept, which we call a *self-admissible set* (SAS). Here, we preview some properties of SAS’s:

- The iteratively admissible strategy profiles⁵ constitute an SAS—thus, every (finite) game has at least one SAS. Typically, though, a game has other SAS’s, too.
- An SAS is fully invariant—i.e., it is invariant to the Elmes-Reny [19, 1994] transformations and also the Kohlberg-Mertens [28, 1986] convex-combination transformation.
- In a game of perfect information satisfying the “single payoff condition”, each SAS yields a (possibly different) single outcome, which is that reached by a pure-strategy Nash equilibrium.

The rest of the paper is organized as follows. Section 2 discusses the invariance requirement. Sections 3 through 5 develop the interactive decision-theoretic (“epistemic”) model we use to formalize our requirements of RCAR and invariance. Section 6 defines SAS’s, and gives some properties and examples. Section 7 shows that this is indeed the solution concept characterized by our two

⁵That is, the profiles that survive iterated (simultaneous maximal) elimination of weakly dominated strategies.

requirements. Section 8 looks at SAS’s in perfect-information games. Section 9 adds an additional requirement on a solution concept, and concludes.

The reader who is more interested in the main results, and less in motivation, should focus on Sections 4 through 9.

2 Invariance

Figure 2 differed from Figure 1 by “addition of decision nodes.” In general, we see this transformation along with the two other transformations of Elmes-Reny [19, 1994]—namely, “coalesce” and “interchange” of moves—to be irrelevant to strategic decision-making.⁶ Up to the duplication of pure strategies, two games with perfect recall differ by a sequence of these transformations if and only if they have the same strategic form. As such, we require any characterization of RCAR—in fact, rationality alone—to depend only on the strategic form, up to the duplication of pure strategies.⁷

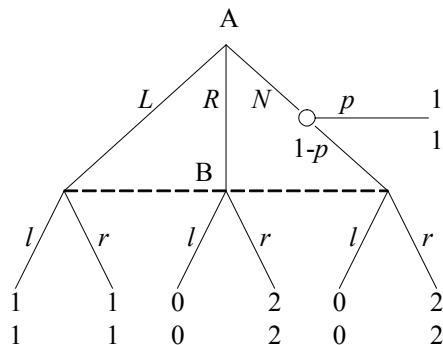


Figure 3

Kohlberg-Mertens [28, 1986] propose a fourth transformation: The addition or deletion of a pure strategy that yields payoffs for each player that are a convex combination of the payoffs from other pure strategies. To see the logic behind this transformation, look at Figure 3, which applies the transformation to the game in Figure 2.⁸ Ann now has an extra strategy, which is a $p:(1-p)$ mixture of L and R . But since Ann has already considered all such mixtures in forming her preferences (à la

⁶That is, we require that a player think about different—though equivalent—trees, when choosing a strategy. This is very similar to the situation in decision theory, where the decision maker is asked to consider all possible acts, not just the available ones.

⁷In a similar fashion, Kohlberg-Mertens [28, 1986] argued that an equilibrium of a game should depend only on the reduced strategic form. Kohlberg-Mertens framed their argument in terms of the Dalkey [18, 1953]-Thompson [45, 1952] transformations. We appeal to the Elmes-Reny transformations because, unlike the Dalkey-Thompson transformations, they preserve perfect recall.

Friedenberg-Hörner [23, 2002] provide an axiomatization of preferences, under which a solution is invariant to these transformations.

⁸Here Ann can move L , R , or N . If she plays N , then Nature gets to move and chooses the right-hand branch with probability $0 < p < 1$. The cases $p = 0$ and $p = 1$ (with Nature’s node deleted) are covered by “addition of decision nodes.”

Anscombe-Aumann [2, 1963], say), explicitly including this mixture shouldn't change the strategic situation. Thus, we require another form of invariance: If two trees differ only by the fourth transformation, the set of strategies consistent with RCAR should differ only by this transformation.

To summarize, we maintain that a player's decision making should not be misled by presentation effects. Therefore, RCAR should be invariant to the four transformations above—or, equivalently, RCAR should depend only on the reduced strategic form, i.e., the strategic form after deletion of rows (resp. columns) that are convex combinations of other pure strategies. In particular, in any tree that reduces to the matrix corresponding to Figure 2, Bob's rationality should imply that Bob will play r .

That said, we do not require EFR in every tree with the same reduced strategic form.⁹ A priori, there is no reason to suspect that this would correspond to a characterization of RCAR for a given game.¹⁰ While invariance is a desirable property, it should not be imposed on the solution itself. To borrow from Kohlberg-Mertens [28, 1986, Appendix E]:

“While ... invariance ... [is a] reasonable property against which a proposed solution concept may be checked, [it] cannot serve as a definition or an axiom.... Presumably, a correct definition or axiom system should involve only rationality criteria (like admissibility) about the game itself as opposed to criteria (like invariance) about the solution correspondence. In addition, the criteria should be phrased in purely decision theoretic terms—e.g., depend only on the best reply correspondence—instead of depending on the tree.”

So how to get invariance without simply imposing it? The key is the following result from decision theory: Fix a decision matrix M , where the decision maker chooses the row and Nature the column. A decision maker's strategy is conditionally undominated in each tree whose reduced strategic form is M if and only if the strategy is *admissible* (i.e., not weakly dominated) in M . (See Brandenburger-Keisler [17, 2000, Appendix A], who note that the idea for this result is in Kohlberg-Mertens [28, 1986, Section 2.7].¹¹)

This result tells us that we should seek an analysis of RCAR, in which rationality embodies an admissibility requirement. This way, the resulting solution concept should be invariant, just as desired.¹²

⁹Nor do we require iterated conditional dominance with respect to the “normal-form information sets” (Mailath-Samuelson-Swinkels [31, 1993]), as is permitted by the definition of conditional dominance (Shimoji-Watson [42, 1998]).

¹⁰Although Brandenburger-Friedenberg [16, 2003] show that this set does exist: The set of iteratively admissible outcomes is contained in the intersection (over every equivalent tree) of the EFR outcomes.

¹¹The “If” direction is straightforward. To see the “Only if” direction, let r be a row that is weakly dominated by a mixture of rows σ , and let c be the set of columns on which σ does strictly better than r . Now look at the following tree: The decision maker moves first and chooses between the single move $\{r, \sigma\}$ and any of the other rows. Nature then moves, in ignorance of this move. Finally, if the decision maker chose $\{r, \sigma\}$ and Nature chose one of the columns from c , there is a single information set h at which the decision maker gets to choose between r and σ . It follows that r is conditionally dominated (at h) by σ in this tree.

¹²As it will be—see Proposition 2 in Section 6.1. This is a more subtle point than it may at first appear. A solution

We note that there is good reason to adopt admissibility, apart from the fact that it gives us invariance. It is a very natural axiom in and of its own right, embodying the idea that players take every state into consideration. It is standard in decision theory, and seems equally compelling in interactive decision theory.

Let's repeat our two desiderata from the end of Section 1:

- **RCAR** We seek an analysis of rationality and common assumption of rationality.
- **Invariance** We seek an analysis that is invariant to the four transformations above.

We shall achieve RCAR by developing an epistemic model similar to other such models now standard in the literature. And we have argued that the way to satisfy invariance is to build the epistemic model from a decision theory that satisfies admissibility.

3 Epistemic Analysis

Return to the opening example (Figure 1), and ask: What does RCAR mean in this game? To answer, we have to say what each player hypothesizes about the other players' strategies, about their hypotheses, and so on. As usual, we do this by appending to the game an epistemic model, as in Figure 4.

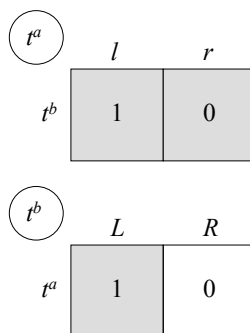


Figure 4

Here, there is one type of Ann, viz. t^a , and one type of Bob, viz. t^b . Type t^a assigns probability 1 to Bob's playing l and being type t^b . Type t^b assigns probability 1 to Ann's playing L and being type t^a . Given her type, Ann has a unique best response, L . Given his type, both l and r are optimal for Bob. (Bob's expected payoff of 1 is independent of his strategy.) The rational strategy-type pairs of each player are shaded in Figure 4.

Let the true state of the structure be (L, t^a, l, t^b) . We just saw that both players are rational at this state. Moreover, each player assigns probability 1 to this state; so, presumably, each player also assumes the other is rational. In fact, it seems we get RCAR at (L, t^a, l, t^b) .

concept may satisfy admissibility, yet fail invariance (e.g., proper equilibrium). We start with an underlying definition of rationality—not a solution concept!—and build in admissibility there. At this level, admissibility implies invariance (in fact, is equivalent to it). This invariance then carries through to our solution concept (SAS).

But this answer doesn't make conceptual sense, as pointed out in Section 1. The problem is that, when Bob is asked to move in the tree, he can no longer maintain his hypothesis that Ann plays L —he knows that she has played R ! Bob must also form a hypothesis about Ann, conditional on his node being reached. Indeed, at this node, Bob must assign probability 1 to Ann's playing R , which makes r not l rational for him.

In the matrix, the problem is that l is weakly dominated by r . Indeed, the game of Figure 1 gives a (simple) instance of the equivalence between admissibility in the matrix and rationality in the tree. (Refer back to Section 2 for a statement of the equivalence.) Both approaches require Bob to choose r not l .

Given our interest in invariance, we solve the problem in the matrix—anticipating that our answer will also be satisfactory in the tree.¹³ Specifically, to get admissibility of choices, we require that the players give every state positive probability, as in Figure 5 (where ε is a small positive number).

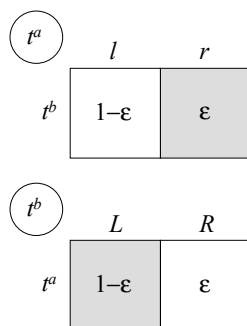


Figure 5

Note that now (r, t^b) is rational for Bob, but (l, t^b) is not. Since Ann assigns positive probability to (l, t^b) —in fact, probability close to 1—we should no longer say that Ann assumes Bob is rational. The next step, then, is to reverse Ann's probabilities, to get Figure 6:

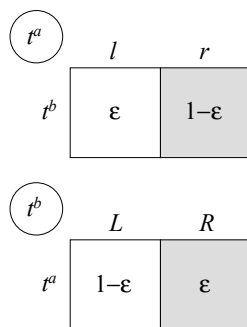


Figure 6

¹³In the matrix, we will introduce first infinitesimals, and then lexicographic probability systems (LPS's) in Section 4. In the tree, the solution is to introduce conditional probability systems (CPS's). (See Rényi [37, 1955] and Battigalli and Siniscalchi [10, 2002].) LPS's and CPS's are designed precisely for the strategic and extensive forms, respectively. LPS's involve 'full consideration of possibilities' (Gus Stuart suggested the term), which fits with choosing a strategy at the outset of the game, without knowing the other players' strategies. CPS's involve conditioning on observations along the game tree.

But now (R, t^a) —not (L, t^a) —is rational for Ann, so we should reverse Bob’s probabilities:

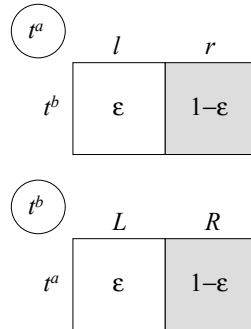


Figure 7

When the true state in Figure 7 is (R, t^a, r, t^b) , both players are rational, and each player assigns probability close to 1 to the rationality of the other. We also have the intuitively correct outcome (R, r) . But can we say that each player assumes the other is rational? The difficulty is that we have required the players to give every state positive probability—here, some states are irrational.

Our resolution will be to think of ε as an infinitesimal. Then, it is natural to say that Ann assumes Bob is rational: she considers the state in which Bob is irrational to be infinitely less likely than the state in which he is rational. Likewise, we can say that Bob assumes Ann is rational. In fact, we can conclude (by a simple induction) that RCAR holds, as desired.¹⁴

Summing up, the first part of our formulation is to require that the players give every state positive probability. This way, we guarantee admissibility. The second part is to allow the players to give some states infinitesimal probability. This way, we can talk about assumption of rationality—and higher-order assumption of rationality—even when there are irrational states.

4 Lexicographic Probability Systems

We formalize the idea that players give some states infinitesimal probability via the concept of a *lexicographic probability system* (LPS), due to Blume-Brandenburger-Dekel [13, 1991a]. Instead of a single probability measure, an LPS gives a player a sequence of probability measures on the state space, with the property that each state receives positive probability under one—and only one—measure. If Ann is rational, she will choose the strategy s over the strategy t when the sequence of expected payoffs associated with s is lexicographically greater than the sequence associated with t . Thus, she views states given positive probability under her primary measure as infinitely more likely than states given positive probability under her secondary measure, which she views as infinitely more likely than states given positive probability under her tertiary measure, and so on. But however unlikely, no state is considered entirely impossible.

We now give a formal definition. Fix a finite set Ω .¹⁵ Let $\mathcal{M}(\Omega)$ be the set of all probability

¹⁴Of course, all this is only informal. The formal treatment of RCAR is in the next two sections.

¹⁵Appendix D treats the infinite case.

measures on Ω , and let $\mathcal{M}^{<\mathbb{N}}(\Omega)$ denote the set of all finite sequences from $\mathcal{M}(\Omega)$. Thus, a typical element of $\mathcal{M}^{<\mathbb{N}}(\Omega)$ is a sequence $(\mu_0, \dots, \mu_{n-1})$, where n is an integer and $\mu_i \in \mathcal{M}(\Omega)$ for all i .

Definition 1 *The set $\mathcal{L}(\Omega)$ of **lexicographic probability systems (LPS's)** is the set of all $(\mu_0, \dots, \mu_{n-1}) \in \mathcal{M}^{<\mathbb{N}}(\Omega)$ such that:*

- (a) $(\mu_0, \dots, \mu_{n-1})$ has **full support**, that is, $\Omega = \bigcup_{i < n} \text{Supp } \mu_i$;
- (b) $(\mu_0, \dots, \mu_{n-1})$ is **mutually singular**, that is, $\mu_i(\omega) > 0$ implies $\mu_j(\omega) = 0$ for $j \neq i$.¹⁶

Given an LPS $(\mu_0, \dots, \mu_{n-1}) \in \mathcal{L}(\Omega)$, the measure μ_0 is the primary hypothesis about the true state, the measure μ_1 is the secondary hypothesis, and so on until the n th hypothesis μ_{n-1} . Conditions (a) and (b) can both be seen as expressions of the idea of ‘full consideration of possibilities’ mentioned in the previous section. Condition (a) is clear: it says that no state is ruled out. Condition (b) is more subtle. We can say that Ann moves on to her secondary hypothesis μ_1 once she has ‘ruled out’ in her mind the states given positive probability under μ_0 . But if $\mu_0(\omega) > 0$ and $\mu_1(\omega) > 0$, then Ann wants to consider ω under μ_1 but has already ruled ω out, which seems contradictory. (Likewise for higher-order hypotheses.¹⁷)

For the application to games, we will need definitions of *rationality* and *assumption* when players have LPS's. Rationality means lexicographically maximizing expected utility, as noted above. (See also Definition 4 in the next section.) We gave the idea for defining assumption at the end of Section 3: Say Ann assumes an event $E \subseteq \Omega$ if she considers the states in E to be infinitely more likely than the states not in E . Formally:

Definition 2 (*Brandenburger-Keisler [17, 2000]*) *An event E is **assumed under** $(\mu_0, \dots, \mu_{n-1})$ (at level j) if there exists $j < n$ such that:*

- (a) $\mu_i(E) = 1$ for all $i \leq j$,
- (b) $\mu_i(E) = 0$ for all $j < i < n$.

Example 1 (*Brandenburger-Keisler [17, 2000]*) *The state space is $\Omega = \{a, b, c, d\}$, and Ann has the LPS (μ_0, μ_1, μ_2) , where $\mu_0(a) = 1, \mu_1(b) = \frac{1}{3}, \mu_1(c) = \frac{2}{3}$, and $\mu_2(d) = 1$. Then Ann assumes each of the events $\{a\}$, $\{a, b, c\}$, and $\{a, b, c, d\}$. Note that in each case, any state not in the event is considered infinitely less likely than any state in the event. In this sense, we can say that Ann is prepared to bet on the event vs. its complement.¹⁸ The intuitive, everyday meaning of “assumption”*

¹⁶Blume, Brandenburger, and Dekel [13, 1991a] provide an axiomatic derivation of LPS's in terms of preferences over acts. N.B. They use the term LPS even if our Condition (b) does not hold. When (b) does hold, they refer to a “lexicographic conditional probability system.” In this paper, we drop the qualifier “conditional,” to avoid any confusion with (extensive-form) CPS's. Refer back to Footnote 13.

¹⁷We emphasize that the argument for mutual singularity of the measures applies when the entire state space is being considered. We certainly do not require it otherwise—e.g., when taking marginals, as in Definition 4 below.

¹⁸Brandenburger-Keisler [17, 2000, Appendix B] give a preference-based definition of assumption. They show it is equivalent to Definition 2 here.

seems similar—hence our use of the term. Another reason for using “assumption” (rather than “belief,” say) is that it is non-monotonic. Ann assumes $\{a\}$, but she does not assume $\{a, b\} \supseteq \{a\}$. This, too, seems to fit with everyday usage. We can say Ann assumes it will be sunny, without also meaning she assumes it will be sunny or cloudy. Ann considers bad weather (clouds or rain) much less likely than good weather. She is betting on sun. She recognizes that bad weather is possible even if unlikely, but if she brings in one possibility (clouds) she also wants to bring in the other (rain).

5 Interactive Structures

This section develops an epistemic formalism, based on LPS’s, that allows us to talk about rationality of the players in a game, what players assume about the rationality of other players, etc.

Fix a two-player finite strategic-form game $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$, where S^a, S^b are the strategy sets and π^a, π^b are the payoff functions of Ann and Bob, respectively.¹⁹ A *sequential probability structure* appends to the game a set of *types* for each player, where a type of Ann is associated with a finite sequence of measures on the strategies and types of Bob, and vice versa.

Definition 3 Fix nonempty finite sets S^a and S^b . An (S^a, S^b) -based (interactive) sequential probability structure is a structure

$$\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle,$$

where T^a, T^b are nonempty finite sets (called **type sets**) and

$$\begin{aligned} \lambda^a &: T^a \rightarrow \mathcal{M}^{<\mathbb{N}}(S^b \times T^b), \\ \lambda^b &: T^b \rightarrow \mathcal{M}^{<\mathbb{N}}(S^a \times T^a). \end{aligned}$$

Members of T^a, T^b are called **types**. Members of $S^a \times T^a \times S^b \times T^b$ are called **states (of the world)**.²⁰

The following definitions all have counterparts with a and b reversed.

Definition 4 A pair $(s^a, t^a) \in S^a \times T^a$ is **rational** if $\lambda^a(t^a) = (\mu_0, \dots, \mu_{n-1}) \in \mathcal{L}(S^b \times T^b)$ and for every $r^a \in S^a$,

$$\left(\sum_{s^b \in S^b} \pi^a(s^a, s^b) \text{marg}_{S^b} \mu_i(s^b) \right)_{i=0}^{n-1} \geq^L \left(\sum_{s^b \in S^b} \pi^a(r^a, s^b) \text{marg}_{S^b} \mu_i(s^b) \right)_{i=0}^{n-1},$$

where $\text{marg}_{S^b} \mu_i$ denotes the marginal on S^b of μ_i .²¹

¹⁹We focus on two players, but our analysis readily extends to n players.

²⁰Appendix D extends this definition to the case of infinite type sets T^a, T^b .

²¹If $x = (x_0, \dots, x_{n-1})$ and $y = (y_0, \dots, y_{n-1})$, then $x \geq^L y$ iff $y_j > x_j$ implies $x_k > y_k$ for some $k < j$.

Rationality is a property of a strategy-type pair: A strategy may be rational given certain hypotheses, and irrational given other hypotheses. In particular, a strategy-type pair (s^a, t^a) of a will be rational if $\lambda^a(t^a)$ is a LPS and s^a lexicographically maximizes a 's expected payoffs, calculated under the LPS $\lambda^a(t^a)$.²²

We now formalize our basic condition of *rationality and common assumption of rationality* (RCAR). For $E \subseteq S^b \times T^b$, let

$$A^a(E) = \{t^a \in T^a : E \text{ is assumed under } \lambda^a(t^a)\}.$$

Let R_1^a be the set of all rational pairs (s^a, t^a) , and define R_m^a inductively by

$$R_{m+1}^a = R_1^a \cap [S^a \times \bigcap_{i \leq m} A^a(R_i^b)].$$

Notice that for each m ,

$$R_{m+1}^a = R_m^a \cap [S^a \times A^a(R_m^b)].$$

Definition 5 *If $(s^a, t^a, s^b, t^b) \in R_{m+1}^a \times R_{m+1}^b$, say there is **rationality and m th-order assumption of rationality** at this state. If $(s^a, t^a, s^b, t^b) \in \bigcap_m R_m^a \times \bigcap_m R_m^b$, say there is **rationality and common assumption of rationality (RCAR)** at this state.*

6 Self-Admissible Sets

In the next section, we are going to show that RCAR is characterized by a new solution concept that we call a *self-admissible set* (SAS). Here, we give the definition of an SAS, and also some properties and examples of SAS's. The next definition and remark are standard:

Definition 6 *Fix $X \subseteq S^b$. A strategy $s^a \in S^a$ is **admissible with respect to X** if there is no $\sigma \in \mathcal{M}(S^a)$ such that*

$$\begin{aligned} \sum_{r^a \in S^a} \pi^a(r^a, s^b) \sigma(r^a) &\geq \pi^a(s^a, s^b) \quad \text{for every } s^b \in X, \\ \sum_{r^a \in S^a} \pi^a(r^a, s^b) \sigma(r^a) &> \pi^a(s^a, s^b) \quad \text{for some } s^b \in X. \end{aligned}$$

*If s^a is admissible with respect to S^b , we simply say s^a is **admissible**.*

²²Note that we allow a sequential probability structure (Definition 3) to include non-LPS measure sequences, although a rational strategy-type pair must have an LPS.

Refer back to Definition 1. We said that Conditions (a) and (b) there were expressions of the idea of 'full consideration of possibilities.' As such, we see them as part of what we mean by rationality in this paper. So the measure sequence associated with a rational strategy-type pair must satisfy these two conditions, but not if the strategy-type pair is irrational. Epistemic structures usually contain both rational and irrational strategy-type pairs—we allow the same here. See Brandenburger-Keisler [17, 2000] for further conceptual discussion of this point.

Remark 1 A strategy $s^a \in S^a$ is admissible with respect to $X \subseteq S^b$ if and only if there is a $\mu \in \mathcal{M}(S^b)$, with $\text{Supp } \mu = X$, such that for every $r^a \in S^a$,

$$\sum_{s^b \in S^b} \pi^a(s^a, s^b) \mu(s^b) \geq \sum_{s^b \in S^b} \pi^a(r^a, s^b) \mu(s^b).$$

Next, we need:

Definition 7 Fix some $s^a \in S^a$ and suppose there is $\varphi \in \mathcal{M}(S^a)$ such that for all $s^b \in S^b$,

$$\sum_{r^a \in S^a} \pi^a(r^a, s^b) \varphi(r^a) = \pi^a(s^a, s^b).$$

Then if $r^a \in \text{Supp } \varphi$, we say r^a **supports** s^a (via φ). We write $\text{su}(s^a)$ for the set of r^a that support s^a .

In words, this says that $\text{su}(s^a)$ consists of Ann's strategies that are part of some convex combination equivalent (for her) to s^a .

Definition 8 Fix $Q^a \subseteq S^a, Q^b \subseteq S^b$. The set $Q^a \times Q^b$ is a **self-admissible set (SAS)** if:

- (a) each $s^a \in Q^a$ is admissible;
- (b) each $s^a \in Q^a$ is admissible with respect to Q^b ; and
- (c) for any $s^a \in Q^a$, if r^a supports s^a (i.e. $r^a \in \text{su}(s^a)$) then $r^a \in Q^a$.

Likewise, for each $s^b \in Q^b$.

An SAS can be viewed as the weak-dominance analog to the concept (Pearce [36, 1984]) of a best-response set: Recall, this is a rectangular subset $Q^a \times Q^b$ of $S^a \times S^b$ such that for each $s^a \in Q^a$, there is a measure $\mu \in \mathcal{M}(S^b)$, with $\text{Supp } \mu \subseteq Q^b$, under which s^a is optimal among all strategies in S^a (and likewise for each $s^b \in Q^b$).²³ Our Condition (b) is similar—though it requires $s^a \in Q^a$ to be optimal under a measure with support equal to, not contained in, the set Q^b . Condition (a) is an extra requirement for us, since we again ask for a full-support measure. Condition (c) is also different: It is a maximality requirement, saying that if $s^a \in Q^a$ is a convex combination (for a) of other strategies, then those other strategies must lie in Q^a .²⁴

Of course, the real justification for Definition 7 is not the analogy to a best-response set, but the epistemic characterization in Section 7.

²³We've adapted Pearce's definition to the case of pure (not mixed) strategies.

²⁴If a Pearce best-response set fails Condition (c), then there is a larger best-response set that satisfies it. (Just take this larger set to be the set of iteratively not-strongly dominated profiles.) Adding Condition (c) to the definition would not affect which strategies are contained in some Pearce best-response set. By contrast, Condition (c) is essential to the definition of an SAS: A set $Q^a \times Q^b$ satisfying Conditions (a) and (b), but failing (c), may not be contained in an SAS. See Appendix B for an example.

6.1 Properties of SAS's

We begin by considering basic properties of an SAS. Proofs of the properties can be found in Appendix A. First, for finite games, SAS's exist:

Proposition 1 (Existence) *The set of profiles that survive iterated (maximal simultaneous) elimination of inadmissible strategies is an SAS.*²⁵

This result is obviously analogous to the fact that the set of profiles that survive iterated elimination of strongly dominated strategies is a Pearce best-response set. The set of iteratively strongly undominated strategies is the (component-by-component) union of all best-response sets. Interestingly, the analogous statement for SAS's is false: There are SAS's that are not contained in the iteratively admissible set. See Example 3 below.

Next is invariance. SAS's are defined on the strategic form. So, by the discussion in Section 2, we have to show invariance with respect to row or columns that are duplicates or convex combinations.

Proposition 2 (Invariance) *Let \overline{G} be a strategic form that differs from G by the addition of a strategy that is a convex combination (for both a and b) of strategies in G . Then any SAS of G (resp. \overline{G}) is an SAS of \overline{G} (resp. G), up to the addition (resp. deletion) of this strategy.*²⁶

Next is an extensive-form property:

Proposition 3 (Projection) *Fix a game tree Γ , with strategic form G , and a proper subtree Δ of Γ . Any SAS of G that allows Δ induces an SAS on the strategic form of Δ .*

The general idea of Projection is that a 'good' solution concept should have the property that if a strategy profile lies in a solution for the whole tree, then it must also lie in a solution for any part of the tree that it reaches.²⁷ Proposition 3 says that SAS's satisfy this condition. In Section 7, we show that SAS's are characterized by RCAR (rationality and common assumption of rationality). So, Proposition 3 can also be viewed as saying: If a profile is to be consistent with RCAR on the whole tree, then it must be consistent with RCAR on any reached part of the tree.

Appendix B discusses other properties of SAS's.

6.2 Examples of SAS's

In each of our opening examples (Figures 1-3), there is a unique SAS, namely $\{R\} \times \{r\}$, as desired. Here, we find the SAS's in three examples: a Bertrand-style pricing game, the Battle of the Sexes with an outside option, and the Finitely Repeated Prisoner's Dilemma. (Other examples of SAS's are in Section 8.)

²⁵If a different order of elimination is used, the result may not be an SAS. It is also true that an SAS may not correspond to any order of elimination. See Appendix B for examples.

²⁶Precise statements of this and the following property can be found in Appendix A.

²⁷Kohlberg-Mertens [28, 1986] introduced (essentially) this property.

Example 2 (A Bertrand Pricing Game) Figure 8 depicts a game between two Bertrand competitors, when there are four available prices: “low” (L), which is equal to marginal cost; “low-medium” (m); “high-medium” (M); and “high” (H).

	H	M	m	L
H	3 3	4 0	3 0	0 0
M	0 4	2 2	3 0	0 0
m	0 3	0 3	1 1	0 0
L	0 0	0 0	0 0	0 0

Figure 8

If both players charge the same price, they receive the same payoff, which decreases in going from (H, H) to (L, L) . Yet each player has an incentive to charge just beneath the other, so long as that yields a price above marginal cost. An undercut competitor loses its market share.

Each of the four strategies survives iterated strong dominance. There are two Nash equilibria: (L, L) and (m, m) . But there is a unique SAS: $\{m\} \times \{m\}$. To see this, first note that both L and H are inadmissible and so can't be part of an SAS, and then check that there is no subset of $\{m, M\}$ with respect to which M is admissible. Notice, the SAS concept gives an intuitive answer in this game: The players don't price at marginal cost, since that is inadmissible. Yet, competition prevents them from sustaining a price higher than the first ‘price point’ above marginal cost.

Example 3 (Battle of the Sexes with an Outside Option)

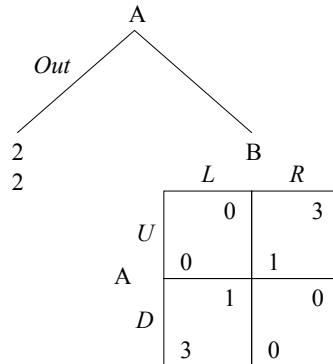


Figure 9

The reader can check that there are three SAS's in this game: $\{Out\} \times \{R\}$, $\{Out\} \times \{L, R\}$, and $\{D\} \times \{L\}$. The iteratively admissible set is $\{D\} \times \{L\}$, which is one of the SAS's, as per Proposition 1. (But there are other SAS's, too, as we noted earlier can happen.) Also, the set $\{D\} \times \{L\}$ is an SAS of the right-hand subtree, as per Proposition 3.

Example 4 (Finitely Repeated Prisoner’s Dilemma)

		B	
		C	D
A	C	c	d
	D	e	0
	D	d	0

$$d > c > 0 > e$$

Figure 10

Consider the Prisoner’s Dilemma played T times (for some integer T), and fix an SAS of the T -fold game. We show that any strategy profile in the SAS yields the Defect-Defect path throughout.²⁸

The proof is by induction on the number of rounds. For $T = 1$, the result is immediate from the fact that any strategy in an SAS is admissible (a fortiori, not strongly dominated). Now assume the result for T , and fix an SAS $Q^a \times Q^b$ of the $(T + 1)$ -fold game. Suppose $s^a \in Q^a$ involves Ann’s playing C on the first round. Then for any $s^b \in Q^b$, Ann gets a first-round payoff of c if s^b involves Bob’s playing C on the first round, and e if s^b involves Bob’s playing D on the first round. These are also Ann’s total payoffs from the game when (s^a, s^b) is played, since the Projection property of SAS’s (Proposition 3 above) and the induction hypothesis together imply that the profile (s^a, s^b) yields the Defect-Defect path on rounds $2, \dots, T + 1$. Suppose instead that Ann chooses the “Defect always” strategy. Then she gets a first-round payoff of d if s^b involves Bob’s playing C on the first round, and 0 if s^b involves Bob’s playing D on the first round. On subsequent rounds Ann gets at least 0 . But then the “Defect always” strategy does strictly better than s^a against every $s^b \in Q^b$, contradicting the definition of an SAS.²⁹

7 RCAR Yields an SAS

We now state and prove the main result of the paper—that rationality and common assumption of rationality (RCAR), as formulated in Section 5, yields an SAS.

²⁸Of course, we are not prescribing play in this game. Rather, there is a clear intuition that rationality and common assumption of rationality (RCAR) should yield the Defect-Defect path. We are showing that the SAS concept captures this intuition. If players cooperate, that simply means RCAR does not hold—a reasonable scenario.

Stuart [44, 1997] is a closely related analysis of the Finitely Repeated Prisoner’s Dilemma, also yielding the Defect-Defect path via a ‘forward-looking’ proof. (He doesn’t start with admissibility, as we do, but instead imposes a mutual absolute continuity condition on the type structure.)

²⁹Pure Nash equilibrium satisfies the Projection property. As such, the above argument also provides a proof of the well-known fact that any such equilibrium of the Finitely Repeated Prisoner’s Dilemma yields the Defect-Defect path.

Theorem 1 Fix a finite strategic-form game $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$.

- (i) Fix a sequential probability structure $\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$. If there is RCAR at the state (s^a, t^a, s^b, t^b) , then the strategy profile (s^a, s^b) is an element of an SAS of G .
- (ii) Fix a strategy profile (s^a, s^b) in an SAS of G . Then there is a sequential probability structure $\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$ and a state (s^a, t^a, s^b, t^b) at which there is RCAR.³⁰

The proof is quite short, though we do relegate a number of technical lemmas to the appendices.

Proof. For part (i), start with a state $(s^a, t^a, s^b, t^b) \in \bigcap_m R_m^a \times \bigcap_m R_m^b$. Since the sets R_m^a, R_m^b are decreasing and S^a, S^b, T^a, T^b are finite, there must be an M such that $R_m^a = R_M^a$ and $R_m^b = R_M^b$ for all $m \geq M$. We show that $\text{proj}_{S^a} R_M^a \times \text{proj}_{S^b} R_M^b$ is an SAS. Take $(s^a, t^a) \in R_M^a$. Then certainly $(s^a, t^a) \in R_1^a$. So, writing $\lambda^a(t^a) = (\mu_0, \dots, \mu_{n-1})$, and using the definition of rationality (Definition 4 in Section 5), we can form a (nested) convex combination of the measures $\text{marg}_{S^b} \mu_i$ to get a full-support measure on S^b under which s^a is optimal. (For the construction of the convex combination, see Proposition 1 in Blume-Brandenburger-Dekel [14, 1991b].) Thus s^a is admissible with respect to S^b (Condition (a) of an SAS). Next, note that $R_{M+1}^a = R_M^a$ implies $t^a \in A^a(R_M^b)$. Using the definition of assumption (Definition 2), there is then a $j < n$ such that $\mu_i(R_M^b) = 1$ for all $i \leq j$ and $\mu_i(R_M^b) = 0$ for all $i > j$. It follows that $\bigcup_{i \leq j} \text{Supp } \text{marg}_{S^b} \mu_i = \text{proj}_{S^b} R_M^b$ (Lemma C1 in Appendix C). Then, as above, we can form a (nested) convex combination of the measures $\text{marg}_{S^b} \mu_i$, for $i \leq j$, to get a measure on S^b , with support $\text{proj}_{S^b} R_M^b$, under which s^a is optimal. Thus s^a is admissible with respect to $\text{proj}_{S^b} R_M^b$ (Condition (b) of an SAS). Finally, suppose $r^a \in \text{su}(s^a)$. Then, for any t^a , we get $(s^a, t^a) \in R_1^a$ implies $(r^a, t^a) \in R_1^a$. (See Lemma A1 in Appendix A.) From this it is immediate that for all m , we have $(s^a, t^a) \in R_m^a$ only if $(r^a, t^a) \in R_m^a$. This establishes Condition (c) of an SAS.

For part (ii) of the theorem, start with an SAS $Q^a \times Q^b$. By assumption, for each $s^a \in Q^a$, there are measures ν and ξ on S^b , with $\text{Supp } \nu = S^b$ and $\text{Supp } \xi = Q^b$, under which s^a is optimal. We can choose ν so that r^a is optimal under ν if and only if $r^a \in \text{su}(s^a)$. (See Lemma C6 in Appendix C.) Note also that $\text{su}(s^a) \subseteq Q^a$, using Condition (c) of an SAS. Now define type sets $T^a = Q^a \cup \{t_*^a\}$ and $T^b = Q^b \cup \{t_*^b\}$, where t_*^a and t_*^b are arbitrary labels. For $t^a = s^a \in Q^a$, the associated $\lambda^a(t^a) \in \mathcal{M}^{<\mathbb{N}}(S^b \times T^b)$ will be a two-level LPS (μ_0, μ_1) where $\text{marg}_{S^b} \mu_0 = \xi$ and $\text{marg}_{S^b} \mu_1 = \nu$. Let $\lambda^a(t_*^a)$ be an arbitrary point in $\mathcal{M}^{<\mathbb{N}}(S^b \times T^b) \setminus \mathcal{L}(S^b \times T^b)$. Define the map λ^b similarly.

Refer now to Figure 11: Take some $s^b \in Q^b$. By construction, the point (s^b, s^b) on the diagonal is rational. Points (r^b, s^b) are rational if and only if $r^b \in \text{su}(s^b)$. In particular, the set R_1^b includes the diagonal and is contained in the rectangle $Q^b \times Q^b$. For this reason, we can take the measures μ_0 and μ_1 so that:

$$\begin{aligned} \text{Supp } \mu_0 &= R_1^b, & \text{marg}_{S^b} \mu_0 &= \xi, \\ \text{Supp } \mu_1 &= (S^b \times T^b) \setminus R_1^b, & \text{marg}_{S^b} \mu_1 &= \nu. \end{aligned}$$

³⁰Appendix D extends Theorem 1 to the case of infinite type spaces T^a, T^b .

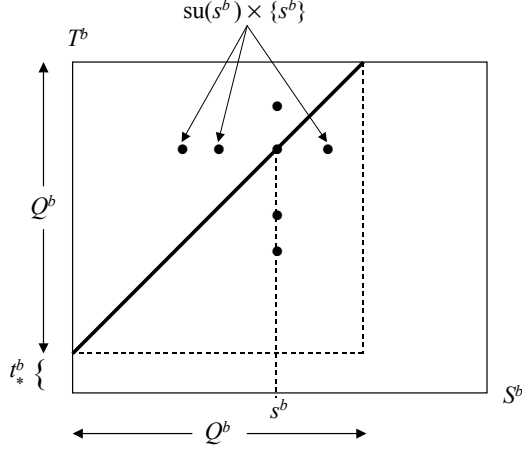


Figure 11

With this, each type $t^a \in Q^a$ assumes b is rational. That is, we have $A^a(R_1^b) = Q^a$. A corresponding argument established $A^b(R_1^a) = Q^b$. So we get that $R_2^b = R_1^b \cap [S^b \times A^b(R_1^a)] = R_1^b$, and symmetrically $R_2^a = R_1^a$. Thus $R_m^a = R_1^a$ and $R_m^b = R_1^b$ for all m , by induction. We have now shown that if $(s^a, s^b) \in Q^a \times Q^b$, then $(s^a, s^a, s^b, s^b) \in \bigcap_m R_m^a \times \bigcap_m R_m^b$, as required. ■

Some comments on Theorem 1. First, the proof immediately yields the following: For a given game and sequential probability structure, the set $\text{proj}_{S^a} \bigcap_m R_m^a \times \text{proj}_{S^b} \bigcap_m R_m^b$ is an SAS. And, conversely, for a given game and SAS, we can construct a sequential probability structure so that $\text{proj}_{S^a} \bigcap_m R_m^a \times \text{proj}_{S^b} \bigcap_m R_m^b$ is this SAS.

Next, notice that Condition (c) of an SAS (the maximality requirement) plays an essential role. Without it, part (ii) of Theorem 1 is false. Appendix C provides an example.³¹

Finally, a natural question is whether we really need the full force of RCAR in Theorem 1(i). We are concerned with finite games. So, for a given game G , shouldn't it be enough to have rationality and finite-order assumption of rationality, to a sufficiently high order? We certainly know the following from the proof of Theorem 1(i) and Appendix D: Fix a game G and an associated sequential probability structure. Then there is indeed a number M such that for any $m \geq M$, $\text{proj}_{S^a} R_m^a \times \text{proj}_{S^b} R_m^b$ is $\text{proj}_{S^a} R_M^a \times \text{proj}_{S^b} R_M^b$. But is M independent of the sequential probability structure? In Appendix C, we provide an example that shows the answer is no. Thus, we can have a game where, for arbitrarily large m , the prediction of rationality and m th-order assumption of rationality is different from the prediction of rationality and $(m + 1)$ th-order assumption of rationality. But the example still leaves open the question: Is there a number M such that, for any sequential probability structure and any $m \geq M$, $\text{proj}_{S^a} R_m^a \times \text{proj}_{S^b} R_m^b$ is some SAS?

³¹The same example answers the following question: Let (σ^a, σ^b) be a perfect equilibrium of a game G . Then $\text{Supp } \sigma^a \times \text{Supp } \sigma^b$ certainly satisfies Conditions (a) and (b) of an SAS. Is it an SAS? The answer is that it may not be. (Similarly, if (σ^a, σ^b) is a proper equilibrium, then $\text{Supp } \sigma^a \times \text{Supp } \sigma^b$ may not be a SAS. An example is available upon request.)

8 Perfect-Information Games

In this section, we consider the particular class of perfect-information (PI) games. We begin with the well-known Centipede game (Rosenthal [38, 1981]), and then give a general result.

Example 5 (Centipede) Consider n -legged Centipede, as in Figure 12. We show that RCAR implies that Ann plays Out at the first node.³²

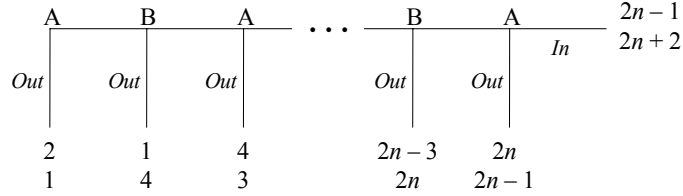


Figure 12

By Theorem 1, we have to show that if $Q^a \times Q^b$ is an SAS for Centipede, and $(s^a, s^b) \in Q^a \times Q^b$, then s^a is Ann's strategy of playing Out at the first node. Suppose, to the contrary, that there is an $(s^a, s^b) \in Q^a \times Q^b$ such that s^a involves Ann's playing across at the first node. In particular, pick the profile (s^a, s^b) that yields the longest path of play (before Ann or Bob plays Out). Let h be the node on this path at which Out is played. Suppose Bob moves at h . (A similar argument applies if Ann moves at h .) Then, by Condition (b) of an SAS, Ann's strategy s^a must be optimal under a measure that assigns: (i) probability 1 to Bob's playing Out at node h or earlier, and (ii) positive probability to Bob's playing across until node h and Out at h . Now consider the strategy r^a for Ann that plays across until node $h-1$ and plays Out at $h-1$. Then r^a does strictly better than s^a against any such measure—a contradiction.

Intuitively, our analysis of Centipede starts at the beginning of the tree and works forwards—reaching a contradiction because if Bob is the player to end the game (playing Out at node h), then Ann rationally should have ended the game earlier.³³ This is quite different from the usual approach of starting at the end of the tree and working backwards, to construct the backward-induction strategies.³⁴ But it is closely related to the argument that a Nash equilibrium of Centipede must involve Ann's playing Out at the first node. (Indeed, if (σ^a, σ^b) is an equilibrium, then exactly the argument we just gave can be applied to $\text{Supp } \sigma^a \times \text{Supp } \sigma^b$.)

Below, we show that this is a general result: For any PI game (satisfying a certain restriction on payoffs), the strategy profiles in an SAS yield a Nash outcome.

Before giving the result, we note that forward-looking arguments arise with SAS's quite generally. Our analysis of the Finitely Repeated Prisoner's Dilemma (Example 4) was forward-looking. More fundamentally, the Projection property of SAS's (Proposition 3) is clearly a forward-looking idea,

³²As in Footnote 28, we are analyzing not prescribing play.

³³Aumann [6, 1998] is another forward-looking epistemic analysis of Centipede.

³⁴Also, note that we do not get the backward-induction (BI) strategies, only the BI outcome. (It is easy to construct a situation satisfying RCAR where Ann plays Out at the first node and Bob plays across at the second node.)

saying that SAS's induce SAS's along the path of play. The Projection property will be used below to prove our general result on PI games.³⁵

8.1 A General Result

We begin with some definitions. Given an extensive game Γ , let S^a (resp. S^b) be the set of strategies of a (resp. b). Let Z denote the set of terminal nodes in Γ , and let $\zeta : S^a \times S^b \rightarrow Z$ be the map from strategy profiles to terminal nodes. Extensive-form payoff functions are maps $\Pi^a : Z \rightarrow \mathbb{R}$ and $\Pi^b : Z \rightarrow \mathbb{R}$. The strategic form of Γ is then $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$, where $\pi^a = \Pi^a \circ \zeta$ and $\pi^b = \Pi^b \circ \zeta$.

An *outcome* is a payoff vector $\Pi(z) = (\Pi^a(z), \Pi^b(z))$. Two terminal nodes z and z' are *outcome equivalent* if $\Pi(z) = \Pi(z')$. We also say that two strategy profiles (s^a, s^b) and (r^a, r^b) are outcome equivalent if $\zeta(s^a, s^b)$ and $\zeta(r^a, r^b)$ are outcome equivalent.

Definition 9 Fix a finite extensive-form game Γ . Say Γ satisfies the **Single Payoff Condition (SPC)** if the following holds for every pair of terminal nodes z, z' in Z : If a (resp. b) moves at the last common predecessor of z and z' , then $\Pi^a(z) = \Pi^a(z')$ implies $\Pi^b(z) = \Pi^b(z')$ (resp. $\Pi^b(z) = \Pi^b(z')$ implies $\Pi^a(z) = \Pi^a(z')$).

In words: Ann is indifferent between two terminal nodes over which she is decisive, only if those two terminal nodes are outcome equivalent.³⁶

With this condition, we can show the following:

Theorem 2 Fix a finite extensive-form game Γ of perfect information satisfying SPC, and let G be the strategic form of Γ .

- (i) Fix a sequential probability structure $\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$. If there is RCAR at the state (s^a, t^a, s^b, t^b) , then the strategy profile (s^a, s^b) is outcome equivalent to a (pure) Nash equilibrium of G .
- (ii) Fix a (pure) admissible Nash equilibrium (s^a, s^b) of G . Then there is a sequential probability structure $\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$ and a state (s^a, t^a, s^b, t^b) at which there is RCAR.

³⁵Interestingly, forward reasoning played a large role in the early literature on PI games. Kalmar [27, 1928-9] proved the Minimax Theorem for Chess via a forward-looking argument. (Schwalbe and Walker [41, 2001] document that this result, often attributed to Zermelo, was actually established by Kalmar. Ewerhart [21, 2002] details the forward proof.) The von Neumann-Morgenstern [46, 1944] proof that a two-player zero-sum game is strictly determined is forward-looking. So is Kuhn's [29, 1950], [30, 1953] proof that every finite PI game has a pure-strategy Nash equilibrium.

³⁶A generic tree satisfies SPC. But non-generic trees can also satisfy SPC, and many games of interest are non-generic. (See, e.g., Example 6 below, and the discussions in Mertens [33, 1989] and Marx-Swinkels [32, 1997].)

SPC is a no-ties condition, similar to simply assuming no ties at all or to assuming "no relevant ties" (Battigalli [9, 1997])—though weaker than both these conditions. In a PI tree, it is also equivalent to the "transfer of decisionmaker indifference" condition (Marx-Swinkels [32, 1997]).

Theorem 2 is proved in Appendix E.³⁷ Note that part (i) of the theorem produces a strategy profile that is admissible (Condition (a) of an SAS) and that is outcome equivalent to a Nash equilibrium. Part (ii) starts with an admissible Nash profile. It is false if we begin only with an admissible profile that is outcome equivalent to a Nash equilibrium.³⁸

The conditions of Theorem 2(i) are quite different from those usually associated with Nash equilibrium. Aumann-Brandenburger [7, 1995, Preliminary Observation] observe that if each player is rational and places probability 1 on the actual strategy choices of the other players, then the strategy choices constitute a Nash equilibrium. By contrast, given a profile (s^a, s^b) in an SAS $Q^a \times Q^b$, Ann must assign positive probability to Bob's actual strategy s^b , and Bob must assign positive probability to s^a , but these probabilities can be arbitrarily small. Nevertheless, in the special case of a PI game satisfying SPC, a Nash outcome must result.³⁹

To see why, fix a strategy profile (s^a, s^b) in an SAS $Q^a \times Q^b$, and consider the last non-terminal node along the path induced by (s^a, s^b) . Call this node h , and suppose Ann moves at h . Since s^a is admissible, Ann's choice at h must certainly be optimal under some measure μ . Write this choice as $s^a(h)$. (This choice may not be unique, but by SPC, if Ann is indifferent between two choices, then so is Bob.⁴⁰ Thus $s^a(h)$ is 'effectively' unique.) Now turn to the preceding node $h - 1$, and suppose Bob moves here. Any of Ann's strategies in Q^a that allow $h - 1$ (and there is at least one, viz. s^a) must involve her playing $s^a(h)$, should h be reached. Since s^b is admissible with respect to Q^a , Bob's choice at $h - 1$ must be optimal under a measure that assigns probability 1 to Ann's choosing $s^a(h)$ at h ; write this choice of Bob's as $s^b(h - 1)$. SPC implies that $s^b(h - 1)$ is 'effectively' unique. Continuing in this way, we get that (s^a, s^b) is a profile with the property: The strategy s^a is optimal under a measure that assigns probability 1 to the outcome induced by (s^a, s^b) ; and vice versa for s^b .⁴¹ This implies that (s^a, s^b) yields a Nash outcome.

The formal proof (in Appendix E) proceeds by induction on the length of the tree, and uses the Projection property (Proposition 3).⁴²

As for part (ii) of Theorem 2, the key step we show in Appendix E is that in a PI game, if an admissible strategy is a convex combination of other strategies, then those strategies must all be copies of the first strategy. With this, we can show that an admissible Nash profile (together with any copies of the two strategies) constitutes an SAS, and we then appeal to Theorem 1(ii).

³⁷Appendix E also provides examples to show both directions are false without SPC.

We first reported this result in "When Does Common Assumption of Rationality Lead to Nash Equilibrium?" (unpublished, July 2001).

³⁸Example E3 in Appendix E.

³⁹Our Theorem 2 appears similar in spirit to a result in Ben Porath [11, 1997]. He formulates a concept called common certainty of rationality, and adds a 'grain of truth' assumption, to get a Nash outcome in PI games.

⁴⁰This isn't as straightforward to see when μ is non-degenerate, but it can still be argued.

⁴¹But note that off the path, the measure may well assign positive probability to choices that are inconsistent with s^b . (Cf. Footnote 34.)

⁴²Thus, as mentioned earlier, the proof is forward looking, just like our analyses of the Finitely Repeated Prisoner's Dilemma and Centipede. The intuition we gave is also forward looking, even though it may appear backward looking! The key is that we don't consider an arbitrary pre-terminal node, as with the backward-induction algorithm, but the pre-terminal node along the path of play. That is, we go forwards to find this node!

8.2 Backward-Induction Outcomes

A PI tree satisfying SPC clearly has a unique backward-induction (BI) outcome. We then get from Theorem 2:

Corollary 1 *Fix a finite extensive-form PI game Γ satisfying SPC, and let G be the strategic form of Γ . Fix also a sequential probability structure $\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$ for G . Consider the Nash equilibria of G that are each outcome equivalent to an admissible profile, and suppose that these equilibria are all outcome equivalent. Then, if there is RCAR at the state (s^a, t^a, s^b, t^b) , the strategy profile (s^a, s^b) yields the BI outcome of Γ .*

Centipede satisfies SPC and has a unique Nash outcome, and so fits Corollary 1. (In Example 5, we proved directly that RCAR yields the BI outcome in Centipede.) Here is another instance of Corollary 1.

Example 6 (Zero-sum Games) *Let Γ be a finite zero-sum game of perfect information.⁴³ Then Γ satisfies SPC and has a unique Nash outcome. So, by Corollary 1, RCAR implies the BI (also the Minimax) outcome of Γ .*

Of course, many PI games (satisfying SPC) have more than one admissible Nash outcome. When this is so, Theorem 2(ii) implies that RCAR need not yield the BI outcome.

Example 7 (A Coordination Game) *The game in Figure 13 satisfies SPC. One admissible Nash equilibrium is (Out, Down). So, by Theorem 2(ii), these strategies can be played under RCAR. Of course, the BI outcome is (3, 3) not (2, 2).*

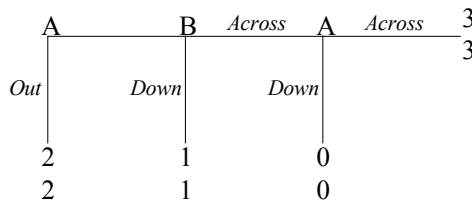


Figure 13

Figure 14 gives a specific set-up in which (Out, Down) is played and RCAR holds. There is one type of each player: $T^a = \{t^a\}$ and $T^b = \{t^b\}$. The LPS $\lambda^a(t^a)$ is the two-level LPS depicted in the top matrix (where brackets indicate the second-order measure). Thus, the type t^a assigns first-order probability one to (Down, t^b), and second-order probability one to (Across, t^b). The LPS $\lambda^b(t^b)$ is a three-level LPS as depicted in the bottom matrix. (The double brackets represent b 's third-order

⁴³Chess is often modelled as such a game (sometimes, though, as an infinite game).

hypothesis.) We have:

$$\begin{aligned} R_1^a &= \{(Out, t^a)\}, \\ R_1^b &= \{(Down, t^b)\}, \\ A^a(R_1^b) &= \{t^a\}, \\ A^b(R_1^a) &= \{t^b\}. \end{aligned}$$

Thus $(Out, t^a, Down, t^b) \in \bigcap_m R_m^a \times \bigcap_m R_m^b$, by induction.⁴⁴

t^a	<i>Down</i>	<i>Across</i>
t^b	1	[1]

t^b	<i>Out</i>	<i>Down</i>	<i>Across</i>
t^a	1	[1]	[[1]]

Figure 14

The key to this example is that Ann assumes Bob plays *Down*, and so has a unique optimal choice, *Out*. Both *Down* and *Across* are irrational for her. In fact, Bob considers it infinitely more likely that Ann play *Down* than *Across* (which is why he plays *Down* himself). Bob is quite free to assign the probabilities this way. To assume Ann is rational, it is enough that he considers *Out* infinitely more likely than both *Down* and *Across*, as he does.

There is a line of reasoning that says Bob should consider *Across* infinitely more likely than *Down*, not vice versa. If Ann foregoes the payoff of 2 she can get by playing *Out*, mustn't she be planning to play *Across* and get a payoff of 3 (rather than play *Down* and get 0)? Bob should assign his probabilities correspondingly, and play *Across* himself. The BI outcome results.

This would be right if Bob could indeed make Ann's playing *Across* rational. But in the present example, he can't. Since Ann assumes Bob plays *Down*, a choice of *Across* is irrational.⁴⁵ This said, it is certainly important to identify epistemic conditions that do yield the BI outcome in a PI game.⁴⁶ The example suggests what is needed—to add more types. Suppose we add a type of Ann that assumes Bob plays *Across*; call this type u^a . There would now be two rational strategy-type pairs of Ann, the pair (Out, t^a) as before and also the pair $(Across, u^a)$. Now, it is right that if Bob

⁴⁴The rational strategy-type pairs are shaded, as in the examples in Section 3.

⁴⁵The BI outcome of (3, 3) does look salient in this game, but that's only until we make the players' assumptions clear. In the example, the players happen to be trapped by a 'bad set of assumptions,' that prevents them reaching (3, 3). In any case, there are other games in which the BI outcome is unappealing, which gives all the more reason for an analysis that yields non-BI outcomes.

⁴⁶A number of recent papers give epistemic justifications for BI; see, inter alia, Asheim [3, 2001], Aumann [5, 1995], Battigalli-Siniscalchi [10, 2002], Brandenburger-Keisler [17, 2000], Feinberg [22, 2001], Halpern [25, 2001], Samet [39, 1996], and Stalnaker [43, 1998]. See also Section 9 below.

assumes Ann is rational, he must consider *Across* infinitely more likely than *Down* for her. As we said above, the BI outcome should then result.

The general counterpart to this idea is to use a so-called *complete* sequential probability structure—one that contains all possible types of the players. We turn to the completeness requirement below.

9 Three Desiderata

We began by asking for a solution concept that satisfies two requirements:

1. **RCAR** It is characterized by the condition of rationality and common assumption of rationality.
2. **Invariance** It is invariant to the four transformations described in Section 2.

This led to the concept of a self-admissible set (SAS). But while requirements 1 and 2 are basic, one may very well want to impose additional conditions. In the previous section, we saw one such possibility:

3. **Completeness** We seek to analyze 1 and 2 with respect to a complete epistemic model.

This is the condition that all possible types of the players be present in the model.⁴⁷ We saw that completeness is the extra condition that should yield the backward-induction outcome in perfect-information games. (Without completeness, we got admissible profiles yielding Nash outcomes.⁴⁸)

No doubt, there are other conditions one could impose.⁴⁹ But it turns out that asking for 1, 2, and 3 is already too much. Brandenburger-Keisler [17, 2001] formalize 1 and 2 as in this paper,⁵⁰ and show that if the epistemic model is complete, then, in fact, there can be no state at which there is RCAR.

Taken together, then, conditions 1, 2, and 3 are too stringent—we must give up on (at least) one. In this paper, we didn’t impose completeness, and showed that 1 and 2 are characterized by the SAS concept. Battigalli-Siniscalchi [10, 2002] essentially have 1 and 3. As we noted in the Introduction, they formalize rationality and common strong belief of rationality on the tree,⁵¹ and impose completeness, to characterize extensive-form rationalizability. They do not get invariance. Finally, Brandenburger-Keisler [17, 2000] give up on RCAR. They have 2 and 3, and show that

⁴⁷See Brandenburger-Keisler [17, 2000, Definition 6] for a precise statement.

⁴⁸We re-emphasize our view that these non-BI outcomes are sensible, even appealing, and should not be automatically rejected. A similar view is expressed by Mertens [33, 1989, pp.582-3], although he is coming from a different direction (equilibrium vs. epistemic analysis): “I had (and still have) some instinctive liking for the bruto Nash equilibrium, or a modest refinement like admissible equilibria, or normal-form perfect equilibria.”

⁴⁹For example, equilibrium-like conditions, stated as restrictions on the types.

⁵⁰With infinite type sets as in Appendix D.

⁵¹“Strong belief” is formalized via a conditional probability system on the tree, just as “assumption” is formalized via an LPS on the matrix. Refer back to Footnote 13.

rationality and m th-order assumption of rationality (for some number m) characterizes m rounds of elimination of inadmissible strategies. Table 1 summarizes.⁵²

Conditions	Solution Concept
1. Rationality & common assumption of rationality 2. Invariance 3. <i>NOT</i> Completeness	Self-admissible sets
1. Rationality & common strong belief of rationality 2. <i>NOT</i> Invariance 3. Completeness	Extensive-form rationalizability (Battigalli-Siniscalchi [10, 2002])
1. Rationality & m th-order assumption of rationality— <i>NOT</i> common assumption 2. Invariance 3. Completeness	m rounds of iterated admissibility (Brandenburger-Keisler [17, 2000])

Table 1

Is it really impossible to satisfy 1, 2, and 3 together, or is this an artefact of the formalism? Might there be another model in which all three conditions are consistent? We don't think so. We began with an invariance requirement on rationality, and saw that this was equivalent to admissibility. Our epistemic model was standard—excepting only the definition of rationality as incorporating admissibility. RCAR and completeness are both then defined (as they must be) with respect to the epistemic model. In sum, we see the formalism as natural, even inevitable.

Which condition, then, should be dropped? Each of them is restrictive. RCAR asks players to assume “rationality and m th-order assumption of rationality”—for all m . Invariance asks a player to consider all equivalent trees—or, since invariance is equivalent to admissibility, to consider all states within the model. Completeness asks players to consider all possible types that are implied by the model. (Arguably, the completeness condition is more stringent than invariance, since it may require ‘expanding’ the model. But our real point is simply that all three conditions are stringent.)

It seems to us that there is no ‘better’ condition to drop. One choice is to maintain RCAR and drop either invariance or completeness. This is saying that the context of the game matters—either in the form of its presentation (lack of invariance) or in the form of the particular hypotheses the players hold (incompleteness). The other choice is to let irrationality into the picture, by dropping RCAR. The right choice to make is probably different in different settings.

⁵²Both the Battigalli-Siniscalchi [10, 2002] and Brandenburger-Keisler [17, 2000] analyses yield the BI outcome in PI games satisfying SPC. (Battigalli-Siniscalchi assume a stronger condition, but SPC suffices.) This was the basis for our assertion in the previous section that completeness was the ‘missing ingredient’ in getting BI.

Appendix A: Properties of SAS's

Some notation: When $\mu \in \mathcal{M}(S^b)$, write $\pi^a(s^a, \mu)$ for $\sum_{s^b \in S^b} \pi^a(s^a, s^b) \mu(s^b)$; if $\varphi \in \mathcal{M}(S^a)$ write $\pi^a(\varphi, s^b)$ for $\sum_{s^a \in S^a} \pi^a(s^a, s^b) \varphi(s^a)$. (Likewise with a and b reversed.)

Lemma A1 *A strategy s^a is optimal under $\mu \in \mathcal{M}(S^b)$ if and only if all $r^a \in \text{su}(s^a)$ are optimal under μ .*

Proof. Suppose every r^a that supports s^a via φ is optimal under μ , i.e. $\pi^a(r^a, \mu) \geq \pi^a(u^a, \mu)$ for all $u^a \in S^a$. Then certainly $\sum_{r^a \in S^a} \pi^a(r^a, \mu) \varphi(r^a) \geq \pi^a(u^a, \mu)$. Notice that

$$\begin{aligned} \sum_{r^a \in S^a} \pi^a(r^a, \mu) \varphi(r^a) &= \sum_{s^b \in S^b} \sum_{r^a \in S^a} \pi^a(r^a, s^b) \varphi(r^a) \mu(s^b) \\ &= \sum_{s^b \in S^b} \pi^a(s^a, s^b) \mu(s^b) \end{aligned} \quad (\text{A1})$$

so that $\pi^a(s^a, \mu) \geq \pi^a(u^a, \mu)$, as required.

Next suppose that s^a is optimal under μ , i.e. $\pi^a(s^a, \mu) \geq \pi^a(u^a, \mu)$ for all $u^a \in S^a$. If $\pi^a(r^a, \mu) < \pi^a(s^a, \mu)$ for some r^a that supports s^a via φ , then equation (A1) implies that there exists $v^a \in \text{Supp } \varphi$ with $\pi^a(v^a, \mu) > \pi^a(s^a, \mu)$. This contradicts the optimality of s^a . ■

Lemma A2 *Fix $s^a \in S^a$ and $\varphi \in \mathcal{M}(S^a)$ with $\pi^a(\varphi, s^b) = \pi^a(s^a, s^b)$ for all $s^b \in S^b$. Let $X \subseteq S^a$ with $\text{Supp } \varphi \subseteq X$. Then s^b is admissible with respect to X if and only if it is admissible with respect to $X \cup \{s^a\}$.*

Proof. We can obviously assume that $s^a \notin X$. Now, if s^b is admissible with respect to X , there is a $\mu \in \mathcal{M}(S^a)$, with $\text{Supp } \mu = X$, such that s^b is optimal under μ . Define $\nu \in \mathcal{M}(S^a)$ by

$$\nu(r^a) = \begin{cases} \varepsilon & \text{if } r^a = s^a, \\ \mu(r^a) - \varepsilon\varphi(r^a) & \text{otherwise,} \end{cases}$$

where $\varepsilon > 0$ is chosen small enough that $\mu(r^a) - \varepsilon\varphi(r^a) > 0$. (Of course, this is possible since $\varphi(r^a) > 0$ implies $\mu(r^a) > 0$.) Then $\text{Supp } \nu = X \cup \{s^a\}$, and $\pi^b(\nu, r^b) = \pi^b(\mu, r^b)$ for all $r^b \in S^b$. Thus s^b is admissible with respect to $X \cup \{s^a\}$.

Conversely, if s^b is admissible with respect to $X \cup \{s^a\}$, there is a $\mu \in \mathcal{M}(S^a)$, with $\text{Supp } \mu = X \cup \{s^a\}$, such that s^b is optimal under μ . Define $\nu \in \mathcal{M}(S^a)$ by $\nu(r^a) = \mu(r^a) + \mu(s^a)\varphi(r^a)$ for $r^a \in S^a$. Then $\text{Supp } \nu = X$, and $\pi^b(\nu, r^b) = \pi^b(\mu, r^b)$ for all $r^b \in S^b$. Thus s^b is admissible with respect to X . ■

Next, given $X \subseteq S^a$ and $Y \subseteq S^b$, let $\pi^a|(X \times Y)$ denote the restriction of π^a to $X \times Y$ and $\pi^b|(X \times Y)$ the restriction of π^b to $X \times Y$. Let $S_0^a = S^a$ and define S_m^a inductively by

$$S_{m+1}^a = \{s^a \in S_m^a : s^a \text{ is admissible in the game } \langle S_m^a, S_m^b, \pi^a|(S_m^a \times S_m^b), \pi^b|(S_m^a \times S_m^b) \rangle\}.$$

Define S_m^b in parallel fashion.

Definition A1 A strategy $s^a \in S_m^a$ or $s^b \in S_m^b$ is called *m-admissible*. A strategy is *iteratively admissible (IA)* if it is *m-admissible* for all m .

For a finite game, there exists M such that $S_m^a = S_M^a$ and $S_m^b = S_M^b$ for all $m \geq M$.

Proof of Proposition 1. Fix $s^a \in S_M^a$. Certainly, $s^a \in S_1^a$, and so s^a is admissible with respect to S^b , establishing Condition (a). By definition, there is a $\mu \in \mathcal{M}(S^b)$, with $\text{Supp } \mu = S_M^b$, under which s^a is optimal among all strategies in S_M^a . Lemma D2 in Brandenburger-Keisler [17, 2000] then implies that s^a is admissible with respect to S_M^b (i.e., is optimal among all strategies in S^a), establishing Condition (b). For Condition (c), we show if $r^a \in \text{su}(s^a)$, then $r^a \in S_m^a$ for all m . The result is immediate for $m = 0$, so assume $r^a \in S_m^a$. Since $s^a \in S_{m+1}^a$, there is a $\mu \in \mathcal{M}(S^b)$, with $\text{Supp } \mu = S_m^b$, under which s^a is optimal (among all strategies in S^a). By Lemma A1 this is also true for r^a , so that $r^a \in S_{m+1}^a$, as required. The conditions for b are verified the same way. ■

We now give a formal statement and proof of Proposition 2. Consider games $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$ and $\bar{G} = \langle S^a \cup \{q^a\}, S^b, \bar{\pi}^a, \bar{\pi}^b \rangle$, where $q^a \notin S^a$, and:

- (a) $\bar{\pi}^a|S^a \times S^b = \pi^a$, $\bar{\pi}^b|S^a \times S^b = \pi^b$;
- (b) there is a $\varphi \in \mathcal{M}(S^a)$ such that $\bar{\pi}^a(q^a, s^b) = \pi^a(\varphi, s^b)$ and $\bar{\pi}^b(q^a, s^b) = \pi^b(\varphi, s^b)$ for each $s^b \in S^b$.

Proposition 2

- (i) Let $\bar{Q}^a \times \bar{Q}^b$ be an SAS of \bar{G} . Then $(\bar{Q}^a \setminus \{q^a\}) \times \bar{Q}^b$ is an SAS of G .
- (ii) Let $Q^a \times Q^b$ be an SAS of G . If q^a does not support any strategy in Q^a , then $Q^a \times Q^b$ is an SAS of \bar{G} . Otherwise, $(Q^a \cup \{q^a\}) \times Q^b$ is an SAS of \bar{G} .

Proof. For (i), it is immediate that each $s^a \in \bar{Q}^a \setminus \{q^a\}$ satisfies Conditions (a)-(c) of an SAS. Since each $s^b \in \bar{Q}^b$ is admissible with respect to $S^a \cup \{q^a\}$ (Condition (a) applied to \bar{G}), Lemma A2 implies that each $s^b \in \bar{Q}^b$ is admissible with respect to S^a . Likewise $s^b \in \bar{Q}^b$ is admissible with respect to \bar{Q}^a (Condition (b) applied to \bar{G}): It suffices to consider the case when $q^a \in \bar{Q}^a$. Then $\text{Supp } \varphi \subseteq \bar{Q}^a$ (Condition (c) applied to \bar{G}). It follows from Lemma A2 that s^b is admissible with respect to $\bar{Q}^a \setminus \{q^a\}$, establishing Condition (b) for b .

For Condition (c), suppose r^b supports $s^b \in \bar{Q}^b$, via $\rho \in \mathcal{M}(S^b)$, in the game G . We have to show that $r^b \in \bar{Q}^b$. This will follow from Condition (c) applied to \bar{G} , provided $\pi^b(q^a, \rho) = \pi^b(q^a, s^b)$. Notice that

$$\begin{aligned} \pi^b(q^a, \rho) &= \sum_{u^b \in S^b} \pi^b(s^a, u^b) \rho(u^b) \\ &= \sum_{u^b \in S^b} \sum_{s^a \in S^a} \pi^b(s^a, u^b) \varphi(s^a) \rho(u^b) \end{aligned}$$

$$\begin{aligned}
&= \sum_{s^a \in S^a} \sum_{u^b \in S^b} \pi^b(s^a, u^b) \rho(u^b) \varphi(s^a) \\
&= \sum_{s^a \in S^a} \pi^b(s^a, s^b) \varphi(s^a) \\
&= \pi^b(q^a, s^b)
\end{aligned}$$

as required.

For (ii), first suppose that q^a does not support any strategy in Q^a . Any $s^a \in Q^a$ is admissible with respect to S^b (resp. Q^b) among strategies in S^a . It follows from Lemma A1 that each $s^a \in Q^a$ is also admissible with respect to S^b (resp. Q^b) among strategies in $S^a \cup \{q^a\}$. This establishes Conditions (a) and (b) for $s^a \in Q^a$. Condition (c) is immediate for this case. Next, because each $s^b \in Q^b$ is admissible with respect to S^a , it is also admissible with respect to $S^a \cup \{q^a\}$ by Lemma A2. Condition (b) is immediate. Finally, notice that if r^b supports s^b in \overline{G} , then certainly it does in G , so Condition (c) follows.

Now suppose q^a supports some $s^a \in Q^a$, and write $\overline{Q^a} = Q^a \cup \{q^a\}$. Certainly each $r^a \in Q^a$ satisfies Conditions (a) and (b). Lemma A1 implies that q^a also satisfies (a) and (b), since s^a does. Condition (c) is clearly satisfied for any $r^a \in Q^a$, since $q^a \in \overline{Q^a}$. Condition (c) is also satisfied for q^a , since if some u^a supports q^a , then u^a also supports s^a , and Condition (c) applied to G then implies $u^a \in Q^a$. The verification of Conditions (a) and (c) for $s^b \in Q^b$ is as above, and (b) is satisfied since Lemma A2 gives that s^b is admissible with respect to $Q^a \cup \{q^a\}$. ■

Next is the formal statement and proof of Proposition 3. Fix a tree Γ , with strategic form G , and let Δ be a proper subtree of Γ , with strategic form D . Let S_Δ^a (resp. S_Δ^b) be the subset of Ann's (resp. Bob's) strategies that allow subtree Δ . Note that, up to duplication of strategies, we can identify S_Δ^a with the set of Ann's strategies on subtree Δ . (Likewise for Bob.) Since SAS's are invariant to the addition or deletion of duplicate strategies (Proposition 2), we can indeed make this identification.

Proposition 3 *Let $Q^a \times Q^b$ be an SAS of G , and suppose that $(Q^a \cap S_\Delta^a) \times (Q^b \cap S_\Delta^b) \neq \emptyset$. Then $(Q^a \cap S_\Delta^a) \times (Q^b \cap S_\Delta^b)$ is an SAS of D , up to the addition of strategies that are duplicates on Δ .*

Proof. Each $s^a \in Q^a \cap S_\Delta^a$ is optimal under some $\mu \in \mathcal{M}(S^b)$ with $\text{Supp } \mu = S^b$ (Condition (a)). We have $\mu(S_\Delta^b) > 0$, so $\mu(\cdot|S_\Delta^b)$ is well defined with $\text{Supp } \mu(\cdot|S_\Delta^b) = S_\Delta^b$. Suppose s^a is not optimal under $\mu^a(\cdot|S_\Delta^b)$, i.e., there is some $r^a \in S_\Delta^a$ with

$$\sum_{s^b \in S_\Delta^b} \pi^a(r^a, s^b) \mu(s^b|S_\Delta^b) > \sum_{s^b \in S_\Delta^b} \pi^a(s^a, s^b) \mu^a(s^b|S_\Delta^b).$$

Define a new strategy $q^a \in S_\Delta^a$ that agrees with r^a at each information set (of a) in Δ , and agrees with s^a elsewhere. Then

$$\sum_{s^b \in S^b} \pi^a(q^a, s^b) \mu(s^b) > \sum_{s^b \in S^b} \pi^a(s^a, s^b) \mu(s^b),$$

a contradiction. Thus s^a is admissible with respect to S_Δ^b .

The argument for Condition (b) is very similar. Each $s^a \in Q^a \cap S_\Delta^a$ is optimal under some $\nu \in \mathcal{M}(S^b)$ with $\text{Supp } \nu = Q^b$. Since $Q^b \cap S_\Delta^b \neq \emptyset$, we have $\nu(Q^b \cap S_\Delta^b) > 0$, so that $\nu(\cdot | Q^b \cap S_\Delta^b)$ is well defined with $\text{Supp } \nu(\cdot | Q^b \cap S_\Delta^b) = Q^b \cap S_\Delta^b$. Suppose there is some $r^a \in S_\Delta^a$ with

$$\sum_{s^b \in S_\Delta^b} \pi^a(r^a, s^b) \nu(s^b | Q^b \cap S_\Delta^b) > \sum_{s^b \in S_\Delta^b} \pi^a(s^a, s^b) \nu(s^b | Q^b \cap S_\Delta^b).$$

Define a new strategy $q^a \in S_\Delta^a$ that agrees with r^a at each information set (of a) in Δ , and agrees with s^a elsewhere. Then

$$\sum_{s^b \in S_\Delta^b} \pi^a(q^a, s^b) \nu(s^b) > \sum_{s^b \in S_\Delta^b} \pi^a(s^a, s^b) \nu(s^b),$$

a contradiction.

Finally, fix $s^a \in Q^a \cap S_\Delta^a$, and suppose there is $\varphi \in \mathcal{M}(S_\Delta^a)$ such that $\pi^a(\varphi, s^b) = \pi^a(s^a, s^b)$ for all $s^b \in S_\Delta^b$. Given each $r^a \in \text{Supp } \varphi$, let $f(r^a)$ be the strategy that agrees with r^a at each information set (of a) in Δ , and agrees with s^a elsewhere. Define a measure $\rho \in \mathcal{M}(S^a)$ that where $\rho(f(r^a)) = \varphi(r^a)$, for $r^a \in \text{Supp } \varphi$. Then $\pi^a(\rho, s^b) = \pi^a(s^a, s^b)$ for all $s^b \in S^b$. Thus Condition (c) applied to $Q^a \times Q^b$ implies that each $q^a \in \text{Supp } \rho$ is in Q^a . But each $r^a \in \text{Supp } \varphi$ agrees with $f(r^a) \in \text{Supp } \rho$ on subtree Δ , by construction. So, by our identification of strategies, we have that Condition (c) is satisfied. ■

Appendix B: Additional Properties of SAS's

This section discusses additional properties of SAS's. An SAS is a Pearce best-response set, i.e., any strategy in an SAS is rationalizable. (The converse is clearly false; see Figure 1 above.) A consistent pair (Börgers-Samuelson [15, 1992]) need not be an SAS; it is an SAS if it only contains admissible strategies.⁵³ Also, an SAS need not be a consistent pair. A modified consistent pair (Ewerhart [20, 1998]) is an SAS, but not every SAS is a modified consistent pair.⁵⁴ We can have an SAS that is not a (tight) CURB set (Basu-Weibull [8, 1991]), and a (tight) CURB set that is not an SAS. Likewise, an outcome may be allowed by a fully permissible set (Asheim-Dufwenberg [4, 2003]) yet not by an SAS, and conversely.

Now the relationship to Nash equilibrium: A profile $(s^a, s^b) \in S^a \times S^b$ may be a Nash equilibrium, and yet not be contained in any SAS. But if (s^a, s^b) is a strict Nash equilibrium, then $\{s^a\} \times \{s^b\}$ is an SAS. Similarly, if $(\sigma^a, \sigma^b) \in \mathcal{M}(S^a) \times \mathcal{M}(S^b)$ is a quasi-strict Nash equilibrium, then $\text{Supp } \sigma^a \times \text{Supp } \sigma^b$ is an SAS if and only if it only contains admissible strategies. If $(\sigma^a, \sigma^b) \in \mathcal{M}(S^a) \times \mathcal{M}(S^b)$ is a regular equilibrium (Harsanyi [26, 1973]), then $\text{Supp } \sigma^a \times \text{Supp } \sigma^b$ is an SAS.

If (σ^a, σ^b) is a perfect equilibrium, then $\text{Supp } \sigma^a \times \text{Supp } \sigma^b$ contains only admissible strategies, but still need not be an SAS. Consider the game in Figure B1. The mixed-strategy pair $(\text{Out}, \frac{1}{2}L:\frac{1}{2}R)$ is a perfect equilibrium, but $\{\text{Out}\} \times \{L, R\}$ is not an SAS. Conditions (a) and (b) are satisfied, but Condition (c) is violated. (Even a pure equilibrium in admissible strategies need not be contained in an SAS.)

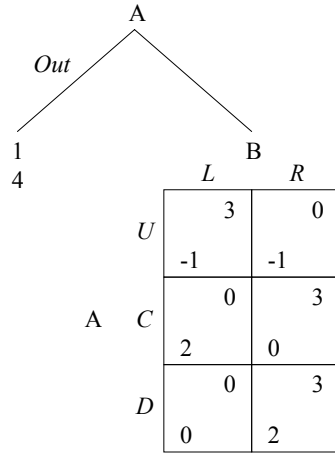


Figure B1

This example also shows why Condition (c) is needed in Theorem 1(ii). Specifically, we claim that *Out* cannot be played under RCAR. Indeed, if $(\text{Out}, t^a) \in R_1^a$, then $(C, t^a) \in R_1^a$ and $(D, t^a) \in R_1^a$. Now consider $(s^b, t^b) \in R_2^b$, so that $t^b \in A^b(R_1^a)$. Writing $\lambda^b(t^b) = (\mu_0, \dots, \mu_{n-1})$, there is then a j such that $\mu_i(R_1^a) = 1$ for all $i \leq j$ and $\mu_i(R_1^a) = 0$ for all $i > j$. But $\text{proj}_{S^a} R_1^a = \{\text{Out}, C, D\}$, so that we must have $s^b = R$. Now suppose $(s^a, t^a) \in R_3^a$. Then $t^a \in A^a(R_2^b)$, where $\text{proj}_{S^b} R_2^b = \{R\}$,

⁵³Samuelson [40, 1992] gives a mixed analog to consistent pairs as defined in Börgers-Samuelson [15, 1992].

⁵⁴We are adapting Ewerhart's definition to pure (not mixed) strategies.

from which $s^a = D$, not $s^a = Out$.

	L	M	R
U	1 0	2 1	0 1
C	1 1	0 0	2 0
D	-2 -1	-1 -1	-1 -1

Figure B2

Next, we give an example to show that while iterated simultaneous maximal deletion of inadmissible strategies yields an SAS (Proposition 1), a different order of deletion need not. In Figure B2, simultaneous maximal deletion gives $\{U\} \times \{M\}$, an SAS. But if we delete D first, then no further deletions are possible, and the remaining set $\{U, C\} \times \{L, M, R\}$ is not an SAS, as L is inadmissible in the overall game.

	L	M	R
U	1 1	0 1	0 0
D	0 0	1 1	0 1

Figure B3

Similarly, an SAS may not correspond to any order of deletion. In Figure B3, there exists a unique order of deletion leading to the set $\{U, L\}$. Yet, $\{D, M\}$ is an SAS.

	L	R
U	1 2	1 2
C	1 3	0 1
D	0 -1	2 -1

Figure B4

Kohlberg-Mertens [28, 1986, Sections 2.6 and 2.9] list various requirements of a solution concept. We've seen that SAS's satisfy several of these—Existence, Projection, Invariance, Admissibility (no strategy in the solution is weakly dominated). Another Kohlberg-Mertens requirement is Iterated Dominance: A solution for a game G should contain a solution for any game \overline{G} obtained from G by the deletion of an inadmissible strategy. SAS's fail this property. The set $\{U\} \times \{L, R\}$ is an SAS

of the game in Figure B4. But if D is deleted, then R becomes inadmissible, and the only SAS of the smaller game is $\{C\} \times \{L\}$.⁵⁵

We maintain that this failure to satisfy Iterated Dominance does not indicate a flaw in the SAS concept. Rather, the example points to the conceptual incompatibility of the Iterated Dominance and Admissibility requirements.⁵⁶ Return to the SAS $\{U\} \times \{L, R\}$ of the larger game. According to Iterated Dominance, this set should contain a solution of the smaller game (after deletion of D), because a solution for the original game should take ‘the whole situation’ into account. But the Admissibility requirement implies that ‘the whole situation’ has changed in going from the original game to the smaller game: The strategy R can be rational in the original game. But, for the game that results after D is deleted, it must be irrational.⁵⁷

⁵⁵Note that $\{C\} \times \{L\}$ is also an SAS of the larger game.

⁵⁶Abreu-Pearce [1, 1984] formulate a requirement similar to Iterated Dominance, and show a formal inconsistency with Admissibility.

⁵⁷Iterated admissibility—perhaps the most natural candidate to satisfy Iterated Dominance—also fails the requirement.

Appendix C: Proof of Theorem 1

Here we prove the two Lemmas referred to in the proof of Theorem 1 (Section 7).

Lemma C1 $\bigcup_{i \leq j} \text{Supp marg}_{S^b} \mu_i = \text{proj}_{S^b} R_M^b$.

Proof. Let $s^b \in \text{proj}_{S^b} R_M^b$, i.e. $(s^b, t^b) \in R_M^b$ for some t^b . Then $\mu_i(s^b, t^b) > 0$ for some $i \leq j$, from which $\text{marg}_{S^b} \mu_i(s^b) > 0$, i.e. $s^b \in \text{Supp marg}_{S^b} \mu_i$. For the other inclusion, suppose $s^b \notin \text{proj}_{S^b} R_M^b$, i.e. $\{s^b\} \times T^b$ is disjoint from R_M^b . But then for each $i \leq j$, we have $\mu_i(\{s^b\} \times T^b) = \text{marg}_{S^b} \mu_i(s^b) = 0$, i.e. $s^b \notin \text{Supp marg}_{S^b} \mu_i$. ■

Next, fix a polytope P in \mathbb{R}^K . Recall that a polytope is a polyhedral set (Grünbaum [24, 2002, Lemma 3.1.1]). Say the points $x^1, \dots, x^J \in P$ each support $x \in P$ if there exists $\lambda^1, \dots, \lambda^J$ each in $(0, 1]$ with $\sum_j \lambda^j = 1$ and $\sum_j \lambda^j x^j = x$. With slight abuse of notation, write $\text{su}(x)$ for the set of points that support $x \in P$.

Lemma C2 *Let H be a supporting hyperplane of P with $x \in H \cap P$. Then $\text{su}(x) \subseteq H \cap P$.*

The proof of Lemma C2 is the proof of Lemma A1, and so we omit it here.

Lemma C3 *Any open ball of x contained in P is also contained in $\text{su}(x)$.*

Proof. To see this, consider an open ball of x contained in P and fix a point x' in this ball. Look at the line going through both x' and x . Certainly, we can find a point x'' both on this line segment and in the ball, with $d(x', x) < d(x', x'')$ for the Euclidean metric d . Then there must exist $\lambda \in (0, 1)$ with $x = \lambda x' + (1 - \lambda)x''$. ■

Lemma C4 *If $x \in \text{int } P$ then $P \subseteq \text{su}(x)$.*

Proof. It suffices to show the following: If $x \in \text{int } P$ then every extreme point of P is contained in $\text{su}(x)$. If all extreme points of P are in $\text{su}(x)$ then $P \subseteq \text{su}(x)$, since any non-extreme point is itself a convex combination of the extreme points of P .

Let x' be an extreme point of P that is not contained in $\text{su}(x)$. Consider the line segment from x to x' , to be denoted by $L(x, x')$. Suppose there exists an open ball of x contained in P . This open ball must meet $L(x, x')$. Thus, some point in this open ball must be supported by x' . But Lemma C3 then implies that $x' \in \text{su}(x)$. As such, there can be no open ball of x contained in P . With this $x \notin \text{int } P$. ■

Lemma C5 *Suppose x^* is contained in a proper face of P . Then there exists a supporting hyperplane H where $x^* \in H \cap P$ and $\text{su}(x^*) = H \cap P$.*

Proof. Let H be a supporting hyperplane of P with $x^* \in H \cap P$. We can choose H so that there does not exist another supporting hyperplane with dimension strictly less than $\dim H$ and that contains x^* . We know from Lemma C2 that $\text{su}(x^*) \subseteq H \cap P$. It suffices to show $H \cap P \subseteq \text{su}(x^*)$.

If $x^* \in \text{int}(H \cap P)$ then $H \cap P \subseteq \text{su}(x^*)$, by Lemma C4. If $x^* \in \text{bd}(H \cap P)$, there exists a maximal proper face of $H \cap P$ that contains x^* (see Grünbaum [24, 2002, Lemma 2.6.3]). This face must itself be a face of P (see Grünbaum [24, 2002, Lemma 2.6.1]). By definition, the dimension of this face must be strictly less than $\dim H$, a contradiction. ■

Say H is a strictly positive supporting hyperplane if $H = \{x \in \mathbb{R}^K : \langle x, y \rangle = \alpha\}$, and $y \gg 0$.⁵⁸

Lemma C6 *Suppose $x^* \in P$ is contained in some strictly positive supporting hyperplane of P . Then there exists a strictly positive supporting hyperplane of P , viz. H , with $\text{su}(x^*) = H \cap P$.*

Proof. Let $H = \{x \in \mathbb{R}^K : \langle x, y \rangle = \alpha\}$ be a strictly positive supporting hyperplane that contains x^* . By Lemma C5, there exists a supporting hyperplane $H' = \{x \in \mathbb{R}^K : \langle x, y' \rangle = a'\}$ with $\text{su}(x^*) = H' \cap P$. Construct $y'' = y' + by$, where $b > 0$. Let $H'' = \{x \in \mathbb{R}^K : \langle x, y'' \rangle = a' + ba\}$. Notice, for $x \in H' \cap P$,

$$\begin{aligned} \langle x, y'' \rangle &= \sum_{k=1}^K x_k y'_k + b \sum_{k=1}^K x_k y_k \\ &= a' + ba, \end{aligned}$$

as $\text{su}(x^*) \subseteq H \cap P$. For $x \in P \setminus (H' \cap P)$,

$$\begin{aligned} \langle x, y'' \rangle &< a' + b \sum_{k=1}^K x_k y_k \\ &\leq a' + ba. \end{aligned}$$

So H'' is a supporting hyperplane with $\text{su}(x^*) = H'' \cap P$. Moreover, since b was chosen arbitrarily, we can choose b so that $y'' \gg 0$. ■

Next, we give the example showing that, for arbitrarily large m , the prediction of rationality and m th-order assumption of rationality is different from the prediction of rationality and $(m + 1)$ th-order assumption of rationality. Consider the game in Figure C1.

	X	Y
X	2	0
Y	0	1

Figure C1

⁵⁸Note, H is a supporting hyperplane if and only if it can be represented in the form $H = \{x \in \mathbb{R}^K : \langle x, y \rangle = \alpha\}$, $y \neq 0$, where $\langle x, y \rangle \leq \alpha$, for all $x \in P$, and $\langle x, y \rangle = \alpha$, for some $x \in P$. As such, it is without loss of generality to take this representation as the definition of a supporting hyperplane. With this, we say H is a strictly positive supporting hyperplane if, given the above definition, $y \gg 0$.

Proposition C1 For each $m \geq 1$, we can construct an interactive sequential probability structure

$$\langle S^a, S^b, T_m^a, T_m^b, \lambda_m^a, \lambda_m^b \rangle$$

with $\text{proj}_{S^a} R_m^a = \{X, Y\}$ and $\text{proj}_{S^a} R_{m+1}^a = \text{proj}_{S^a} \bigcap_i R_i^a = \{X\}$ (and similarly for b).

Proof. We give an induction construction. Let $\langle S^a, S^b, T_1^a, T_1^b, \lambda_1^a, \lambda_1^b \rangle$ be such that $T_1^a = \{X^a, Y_1^a\}$, $T_1^b = \{X^b, Y_1^b\}$, $\lambda_1^a(X^a) = (\mu_0, \mu_1, \mu_2, \mu_3)$ where

$$\begin{aligned} \mu_0(X, X^b) &= 1, & \mu_1(Y, Y_1^b) &= 1, \\ \mu_2(Y, X^b) &= 1, & \mu_3(X, Y_1^b) &= 1, \end{aligned}$$

and $\lambda_1^a(Y_1^a) = (\nu_0, \nu_1, \nu_2, \nu_3)$ where

$$\begin{aligned} \nu_0(Y, Y_1^b) &= 1, & \nu_1(Y, X^b) &= 1, \\ \nu_2(X, X^b) &= 1, & \nu_3(X, Y_1^b) &= 1. \end{aligned}$$

The map $\lambda_1^b : T_1^b \rightarrow \mathcal{M}^{<\mathbb{N}}(S^a \times T_1^a)$ is constructed symmetrically.

Now, assume the structure $\langle S^a, S^b, T_m^a, T_m^b, \lambda_m^a, \lambda_m^b \rangle$ is defined. We will define the structure $\langle S^a, S^b, T_{m+1}^a, T_{m+1}^b, \lambda_{m+1}^a, \lambda_{m+1}^b \rangle$. Set $T_{m+1}^a = T_m^a \cup \{Y_{m+1}^a\}$ and $T_{m+1}^b = T_m^b \cup \{Y_{m+1}^b\}$. For each $t^a \in T_m^a$, construct $\lambda_{m+1}^a(t^a) = (\mu'_0, \dots, \mu'_{2m+3})$ from $\lambda_m^a(t^a) = (\mu_0, \dots, \mu_{2m+1})$ as follows:

- i. $\mu'_0 = \mu_0$,
- ii. $\mu'_1(Y, Y_{m+1}^b) = 1$ and $\mu'_{2m+3}(X, Y_{m+1}^b) = 1$,
- iii. $\mu'_n = \mu_{n-1}$, for each $2 \leq n \leq 2m+2$.

To construct $\lambda_{m+1}^a(Y_{m+1}^a) = (\nu'_0, \dots, \nu'_{2m+3})$, consider $\lambda_m^a(Y_m^a) = (\nu_0, \dots, \nu_{2m+1})$ and set:

- i. $\nu'_0(Y, Y_{m+1}^b) = 1$,
- ii. $\nu'_1(Y, Y_m^b) = 1$,
- iii. $\nu'_2(X, X^b) = 1$,
- iv. $\nu'_3(Y, Y_{m-1}^b) = 1$,
- v. $\nu'_n = \nu_{n-1}$, for each $4 \leq n \leq 2m+2$,
- vi. $\nu'_{2m+3}(X, Y_{m+1}^b) = 1$.

The map $\lambda_{m+1}^b : T_{m+1}^b \rightarrow \mathcal{M}^{<\mathbb{N}}(S^a \times T_{m+1}^a)$ is defined symmetrically.

Fix the structure $\langle S^a, S^b, T_m^a, T_m^b, \lambda_m^a, \lambda_m^b \rangle$. We will show that, for each $i \geq 1$,

$$\begin{aligned} R_i^a &= \{(Y, Y_k^a) : k \geq i\} \cup \{(X, X^a)\}, \\ R_i^b &= \{(Y, Y_k^b) : k \geq i\} \cup \{(X, X^b)\}. \end{aligned}$$

It is immediately verified that this is true for $i = 1$. Assume it is true for some $i \geq 2$. Begin with the type $Y_i^a \in T_m^a$, with $\lambda_m^a(Y_i^a) = (\mu_0, \dots, \mu_{2m+1})$. By construction, there exists n with $\mu_n(Y, Y_{i-1}^b) = 1$ and $\mu_{n+1}(X, X^b) = 1$. By the induction hypothesis, $(Y, Y_{i-1}^b) \notin R_i^b$ and $(X, X^b) \in R_i^b$, so Y_i^a does not assume R_i^b . Next consider $Y_k^a \in T_m^a$, where $k \geq i + 1$ and $\lambda_m^a(Y_k^a) = (\mu_0, \dots, \mu_{2m+1})$. By construction, $\mu_n(R_i^b) = 1$ for all $0 \leq n \leq m - k + 2$ and $\mu_n(R_i^b) = 0$ for all $2m + 1 \geq n > m - k + 2$, and so Y_k^a assumes R_i^b . Finally, $\lambda_m^a(X^a) = (\mu_0, \dots, \mu_{2m+1})$ with $\mu_n(R_i^b) = 1$ for all $0 \leq n \leq 1 + m - i$ and $\mu_n(R_i^b) = 0$ otherwise. ■

Appendix D: Infinite Type Spaces

Here we extend our epistemic model to infinite type spaces, so we can state and prove an infinite analog to Theorem 1 (which was stated for finite type sets). Definitions D1-D3 below are taken directly from Brandenburger-Keisler [17, 2000], which should be consulted for a full discussion. First, the definition of an LPS for infinite spaces.

By a Polish space we will mean a complete separable metric space. Given a Polish space Ω , let $\mathcal{M}(\Omega)$ be the space of all Borel probability measures on Ω with the Prohorov metric. Then $\mathcal{M}(\Omega)$ is again Polish, with diameter ≤ 1 . Another standard fact: Given a Polish space Ω with metric d and diameter ≤ 1 , let $\Omega^{\mathbb{N}}$ be the (countable) product space with the metric $d^{\mathbb{N}}(x, y) = \sum_{n \in \mathbb{N}} 2^{-(n+1)} d(x_n, y_n)$. Then $\Omega^{\mathbb{N}}$ is Polish, with diameter ≤ 1 . By the *empty measure* on Ω we will mean the function that assigns 0 to every Borel subset of Ω . Let $\mathcal{M}'(\Omega)$ be $\mathcal{M}(\Omega)$ with the empty measure added as a new point at distance 1 from every point of $\mathcal{M}(\Omega)$. Then $\mathcal{M}'(\Omega)$ is Polish with diameter ≤ 1 (and the empty measure as an isolated point), and so $(\mathcal{M}'(\Omega))^{\mathbb{N}}$ is also Polish. Now define $\mathcal{M}^{\infty}(\Omega)$ to be the subspace of $(\mathcal{M}'(\Omega))^{\mathbb{N}}$ consisting of all infinite sequences from $\mathcal{M}(\Omega)$ and all finite sequences from $\mathcal{M}(\Omega)$ followed by infinitely many empty measures. That is: $\sigma = (\mu_0, \dots) \in \mathcal{M}^{\infty}(\Omega)$ if and only if for each n , if μ_n is the empty measure then μ_{n+1} is the empty measure. Then $\mathcal{M}^{\infty}(\Omega)$ with the induced metric is a closed subset of $(\mathcal{M}'(\Omega))^{\mathbb{N}}$, and hence is a Polish space.

The length $l(\sigma)$ of a sequence $\sigma = (\mu_0, \dots) \in \mathcal{M}^{\infty}(\Omega)$ is defined as the least n such that μ_n is the empty measure, and is ∞ if there is no such n . (We allow the sequence σ of length 0 where $\mu_n = 0$ for all n .)

Definition D1 *The set $\mathcal{L}(\Omega)$ of **lexicographic probability systems (LPS's)** is the set of all $\sigma = (\mu_0, \dots) \in \mathcal{M}^{\infty}(\Omega)$ such that:*

- (a) $0 < l(\sigma) < \infty$;
- (b) $\Omega = \bigcup_{i < l(\sigma)} \text{Supp } \mu_i$;
- (c) *there are Borel sets U_i , for $i < l(\sigma)$, such that $\mu_i(U_i) = 1$ and $\mu_i(U_j) = 0$ for $i \neq j$.*

Definition D2 *An event E is assumed under σ (at level j) if $\sigma \in \mathcal{L}(\Omega)$ and there is a $j < l(\sigma)$ such that:*

- (a) $\mu_i(E) = 1$ for all $i \leq j$,
- (b) $\mu_i(E) = 0$ for all $j < i < l(\sigma)$,
- (c) $E \subseteq \bigcup_{i \leq j} \text{Supp } \mu_i$.

(If Ω is finite, Condition (b) implies Condition (c), but not if Ω is infinite. Again, see Brandenburger-Keisler [17, 2000] for further discussion.)

Definition D3 Fix nonempty finite sets S^a and S^b . An (S^a, S^b) -based (interactive) sequential probability structure is a structure

$$\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$$

where T^a and T^b are nonempty Polish spaces, λ^a is a continuous mapping from T^a to $\mathcal{M}^\infty(S^b \times T^b)$, and λ^b is a continuous mapping from T^b to $\mathcal{M}^\infty(S^a \times T^a)$.

With these definitions, the sets R_m^a and R_m^b are defined exactly as in Section 5, and just as before, if $(s^a, t^a, s^b, t^b) \in \bigcap_m R_m^a \times \bigcap_m R_m^b$, we say there is *rationality and common assumption of rationality (RCAR)* at this state.

We can now state and prove the infinite analog to Theorem 1(i). (Of course, Theorem 1(ii) constructs (finite) type spaces, so no extension to this part of the theorem is needed.)

Theorem D1 Fix a finite strategic-form game $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$. Fix a sequential probability structure $\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$. If there is RCAR at the state (s^a, t^a, s^b, t^b) , then the strategy profile (s^a, s^b) is an element of an SAS of G .⁵⁹

Proof. We first show that if $\bigcap_m R_m^a \neq \emptyset$, then there is an M such that $\text{proj}_{S^a} R_m^a = \text{proj}_{S^a} R_M^a \neq \emptyset$ for all $m \geq M$. Certainly $\text{proj}_{S^a} \bigcap_m R_m^a \subseteq \bigcap_m \text{proj}_{S^a} R_m^a$. But $\text{proj}_{S^a} \bigcap_m R_m^a \neq \emptyset$ since $\bigcap_m R_m^a \neq \emptyset$, and so $\bigcap_m \text{proj}_{S^a} R_m^a \neq \emptyset$. Since S^a is finite, the sets $\text{proj}_{S^a} R_m^a$ must eventually stop shrinking (and are nonempty). Of course, the same argument shows that there is also an integer (without loss of generality, call it M) such that $\text{proj}_{S^b} R_m^b = \text{proj}_{S^b} R_M^b \neq \emptyset$ for all $m \geq M$.

We next show that $\text{proj}_{S^a} R_M^a \times \text{proj}_{S^b} R_M^b$ is an SAS. The arguments for Conditions (a) and (c) are identical to those in the proof of Theorem 1. For (b), take $s^a \in \text{proj}_{S^a} R_M^a$. Since $\text{proj}_{S^a} R_{M+1}^a = \text{proj}_{S^a} R_M^a$, there is a t^a such that $(s^a, t^a) \in R_{M+1}^a$. Thus $t^a \in A^a(R_M^b)$. Write $\lambda^a(t^a) = (\mu_0, \dots, \mu_{n-1})$. Mirroring the argument for (b) in the proof of Theorem 1, it will suffice to show that $\bigcup_{i \leq j} \text{Supp marg}_{S^b} \mu_i = \text{proj}_{S^b} R_M^b$ (which, in the finite case, was Lemma C1). The proof that the left-hand side is contained in the right-hand side is the same as in the proof of Lemma C1. For the reverse inclusion, let $s^b \in \text{proj}_{S^b} R_M^b$, i.e. $(s^b, t^b) \in R_M^b$ for some t^b . Using Condition (c) of Definition D2, we get $(s^b, t^b) \in \text{Supp } \mu_i$ for some $i \leq j$. Thus $0 < \mu_i(\{s^b\} \times T^b) = \text{marg}_{S^b} \mu_i(s^b)$, i.e. $s^b \in \text{Supp marg}_{S^b} \mu_i$.⁶⁰ ■

⁵⁹The result concerns the case in which there is a state at which there is RCAR. For a given sequential probability structure, there may or may not be such a state (just as in the finite case). Indeed, Brandenburger-Keisler [17, 2000] show that there is no such state in a complete structure. (Refer back to the discussion in Section 9.)

⁶⁰This argument is also given in the proof of Lemma D5 in Brandenburger-Keisler [17, 2000].

Appendix E: Proof for Perfect-Information Games

The formal set-up is as at the beginning of Section 8.1 in the text. The first result is standard (see, e.g., Ben Porath [11, 1997]).

Lemma E1 *A finite PI game satisfies*

$$\min_{s^b \in S^b} \max_{s^a \in S^a} \pi^a(s^a, s^b) = \max_{s^a \in S^a} \min_{s^b \in S^b} \pi^a(s^a, s^b).$$

We now turn to the proof of Theorem 2(i). Using Theorem 1(i), it will suffice to establish:

Proposition E1 *Fix a PI game Γ satisfying SPC. Let G be the strategic form of Γ , and let $Q^a \times Q^b$ be an SAS of G . Then there is a (pure) Nash equilibrium of G such that each profile $(s^a, s^b) \in Q^a \times Q^b$ is outcome equivalent to this equilibrium.*

Proof. By induction on the length of the tree.

Consider a tree of length 1, and assume that a moves at the initial node. (We then define π^a and π^b on S^a alone.) If $s^a \in Q^a$, then $\pi^a(s^a) \geq \pi^a(r^a)$ for all $r^a \in S^a$. Thus (s^a) is an admissible Nash equilibrium. If $r^a \in Q^a$, then we must have $\pi^a(r^a) = \pi^a(s^a)$. Since a moves at the last common predecessor of $\zeta(r^a)$ and $\zeta(s^a)$, SPC implies $\pi^b(r^a) = \pi^b(s^a)$, establishing that (r^a) is outcome equivalent to (s^a) .

Now suppose the result is true for any tree of length l or less, and consider a tree of length $l + 1$. We can again assume that a moves at the initial node. Denote the subtrees that follow a 's initial move by $\Delta_1, \dots, \Delta_K$. Also, let S_k^a be the subset of a 's strategies that reach subtree Δ_k . (Up to duplication, this subset can be identified with a 's strategies for Δ_K .) Let S_k^b be the set of b 's strategies for Δ_k . Thus $S^b = \times_{k=1}^K S_k^b$.⁶¹

By the Projection property (Proposition 3), for each k such that $Q^a \cap S_k^a \neq \emptyset$, we have that $(Q^a \cap S_k^a) \times \text{proj}_{S_k^b} Q^b$ is an SAS of (the strategic form of) Δ_k . By the induction hypothesis, for each such k , each profile in $(Q^a \cap S_k^a) \times \text{proj}_{S_k^b} Q^b$ is outcome equivalent.

Suppose there are strategies $s^a, r^a \in Q^a$ that reach two distinct subtrees, say Δ_1 and Δ_2 respectively. By Condition (b) of an SAS, there is a $\mu \in \mathcal{M}(S^b)$, with $\text{Supp } \mu = Q^b$, such that s^a is optimal under μ . Since s^a reaches Δ_1 and r^a reaches Δ_2 , this implies

$$\sum_{s_1^b \in \text{proj}_{S_1^b} Q^b} \text{marg}_{S_1^b} \mu(s_1^b) \pi^a(s^a, s_1^b) \geq \sum_{s_2^b \in \text{proj}_{S_2^b} Q^b} \text{marg}_{S_2^b} \mu(s_2^b) \pi^a(r^a, s_2^b).$$

(We use π^a for the induced payoff functions on $S_1^a \times S_1^b$ and $S_2^a \times S_2^b$; no confusion should result.) The induction hypothesis gives that $\pi^a(s^a, s_1^b)$ is constant for all $s_1^b \in \text{proj}_{S_1^b} Q^b$, and, likewise, that $\pi^a(r^a, s_2^b)$ is constant for all $s_2^b \in \text{proj}_{S_2^b} Q^b$. Thus we have $\pi^a(s^a, s_1^b) \geq \pi^a(r^a, s_2^b)$ for all such s_1^b and s_2^b . But, symmetrically, we can apply Condition (b) of an SAS to r^a , to get the opposite inequality. Thus $\pi^a(s^a, s_1^b) = \pi^a(r^a, s_2^b)$ for all $s_1^b \in \text{proj}_{S_1^b} Q^b$ and $s_2^b \in \text{proj}_{S_2^b} Q^b$. Also, since a moves at the last common predecessor of $\zeta(s^a, s_1^b)$ and $\zeta(r^a, s_2^b)$, SPC implies that $\pi^b(s_1^b, s^a) = \pi^b(s_2^b, r^a)$.

⁶¹Do notice, this approach differs slightly from that in Appendix A.

We have now shown that every profile in $Q^a \times Q^b$ is outcome equivalent.

The final step is to construct a Nash equilibrium to which (s^a, s_1^b) is outcome equivalent. By the induction hypothesis, there is a Nash equilibrium of Δ_1 to which (s^a, s_1^b) is outcome equivalent. Write it as (q_1^a, q_1^b) . We show that for each $k = 2, \dots, K$, there is a strategy $q_k^b \in S_k^b$ such that $\pi^a(q_1^a, q_1^b) \geq \pi^a(r^a, q_k^b)$ for all $r^a \in S_k^a$. The profile $(q_1^a, (q_1^b, \dots, q_K^b))$ will then be the desired equilibrium.

Since s^a is optimal under μ , as above, we have for each $k = 2, \dots, K$,

$$\pi^a(q_1^a, q_1^b) = \pi^a(s^a, s_1^b) \geq \sum_{s_k^b \in \text{proj}_{S_k^b} Q^b} \text{marg}_{S_k^b} \mu(s_k^b) \pi^a(r^a, s_k^b)$$

for all r^a that reach Δ_k . Letting $(\bar{r}_k^a, \bar{r}_k^b) \in \arg \max_{S_k^a} \min_{S_k^b} \pi^a(\cdot, \cdot)$, we have in particular

$$\pi^a(q_1^a, q_1^b) \geq \sum_{s_k^b \in \text{proj}_{S_k^b} Q^b} \text{marg}_{S_k^b} \mu(s_k^b) \pi^a(\bar{r}_k^a, s_k^b).$$

But $\pi^a(\bar{r}_k^a, r_k^b) \geq \pi^a(\bar{r}_k^a, \bar{r}_k^b)$ for any $r_k^b \in S_k^b$, by definition. So

$$\begin{aligned} \pi^a(q_1^a, q_1^b) &\geq \sum_{s_k^b \in \text{proj}_{S_k^b} Q^b} \text{marg}_{S_k^b} \mu(s_k^b) \pi^a(\bar{r}_k^a, \bar{r}_k^b) \\ &= \pi^a(\bar{r}_k^a, \bar{r}_k^b). \end{aligned}$$

Letting $(\underline{r}_k^a, \underline{r}_k^b) \in \arg \min_{S_k^b} \max_{S_k^a} \pi^a(\cdot, \cdot)$, and using Lemma E1, we get $\pi^a(q_1^a, q_1^b) \geq \pi^a(\underline{r}_k^a, \underline{r}_k^b)$. But $\pi^a(\underline{r}_k^a, \underline{r}_k^b) \geq \pi^a(r_k^a, \underline{r}_k^b)$ for any $r_k^a \in S_k^a$, by definition. So $\pi^a(q_1^a, q_1^b) \geq \pi^a(r_k^a, \underline{r}_k^b)$. Setting $q_k^b = \underline{r}_k^b$ gives the desired profile. ■

We now turn to the proof of Theorem 2(ii).

Remark E1 *If s^a is admissible, each $r^a \in \text{su}(s^a)$ is admissible.*

This follows immediately from the definition of admissibility and Lemma A1.

Lemma E2 *Fix a PI game. If s^a is admissible, then $\pi^a(s^a, s^b) = \pi^a(r^a, s^b)$ for each $r^a \in \text{su}(s^a)$ and $s^b \in S^b$.*

Proof. Let s^a be admissible. Fix also $\varphi \in \mathcal{M}(S^a)$ with $\pi^a(\varphi, s^b) = \pi^a(s^a, s^b)$ for all $s^b \in S^b$. Suppose, contra hypothesis, there exists $r^a, q^a \in \text{Supp} \varphi$ with $\pi^a(r^a, r^b) > \pi^a(q^a, r^b)$ for some $r^b \in S^b$. Denote by h the last common predecessor of the terminal nodes $\zeta(r^a, r^b)$ and $\zeta(q^a, r^b)$. Of course, there may be many choices of profiles (r^a, r^b) and (q^a, r^b) . If so, choose a profile so that there does not exist another whose last common predecessor (strictly) follows h .

Notice that there exists a strategy v^a (perhaps not in $\text{Supp} \varphi$) that agrees with r^a at the node h and the nodes that follow h , but otherwise agrees with q^a . Then v^a and q^a yield a the same payoff against any $s^b \in S^b$ that precludes h . It follows from the admissibility of q^a that there exists a q^b that allows h with $\pi^a(q^a, q^b) > \pi^a(r^a, q^b)$.

For the remainder of the proof, we will take $\pi^a(r^a, r^b) \neq \pi^a(r^a, q^b)$. If this is not the case, then $\pi^a(q^a, q^b) > \pi^a(r^a, q^b) = \pi^a(r^a, r^b) > \pi^a(q^a, r^b)$, and a corresponding argument can then be made with these pairs.

Construct strategies \hat{r}^b and \hat{q}^b , each of which allows the last common predecessor of $\zeta(r^a, r^b)$ and $\zeta(r^a, q^b)$. The two strategies agree in all but one respect: At and beyond the last common predecessor of $\zeta(r^a, r^b)$ and $\zeta(r^a, q^b)$, the strategy \hat{r}^b agrees with r^b and the strategy \hat{q}^b agrees with q^b . Notice that for each $u^a \in \text{Supp } \varphi$ that allows this last common predecessor,⁶² we have $\pi^a(u^a, \hat{r}^b) \neq \pi^a(u^a, \hat{q}^b)$ if and only if $\pi^a(u^a, s^b) = \pi^a(r^a, s^b)$ for each s^b that allows h . The ‘‘If’’ direction is immediate. To establish the ‘‘Only if’’ direction, suppose there exists $u^a \in \text{Supp } \varphi$ that allows this last common predecessor, but $\pi^a(u^a, s^b) \neq \pi^a(r^a, s^b)$ for some s^b that allows h . Then the last common predecessor of $\zeta(u^a, s^b)$ and $\zeta(r^a, s^b)$ strictly follows h , contradicting our choice of profiles.

We can then write

$$\begin{aligned}\pi^a(s^a, \hat{r}^b) &= \alpha \pi^a(r^a, \hat{r}^b) + (1 - \alpha)c, \\ \pi^a(s^a, \hat{q}^b) &= \alpha \pi^a(r^a, \hat{q}^b) + (1 - \alpha)c,\end{aligned}$$

for $\alpha \in (0, 1)$ and some number c . (Recall that $q^a \in \text{Supp } \varphi$ implies $\alpha < 1$.) Thus $\pi^a(s^a, \hat{r}^b) \neq \pi^a(s^a, \hat{q}^b)$. But we could have initially specified φ so that $s^a \in \text{Supp } \varphi$. It follows that $\pi^a(s^a, \hat{r}^b) = \pi^a(r^a, \hat{r}^b)$ and $\pi^a(s^a, \hat{q}^b) = \pi^a(r^a, \hat{q}^b)$. This implies

$$\begin{aligned}(1 - \alpha) \pi^a(s^a, \hat{r}^b) &= (1 - \alpha)c \\ &= (1 - \alpha) \pi^a(s^a, \hat{q}^b),\end{aligned}$$

which can occur only if $\alpha = 1$, a contradiction. ■

To prove Theorem 2(ii), it is enough to show Proposition E2 below and then appeal to Theorem 1(ii).

Proposition E2 *Fix a finite perfect information game Γ satisfying SPC. Let G be the strategic form of Γ , and let (s^a, s^b) be an admissible Nash equilibrium of G . Then $\text{su}(s^a) \times \text{su}(s^b)$ is an SAS of G .*

Proof. Each $r^a \in \text{su}(s^a)$ is admissible, since s^a is admissible. So Condition (a) of an SAS is satisfied.

Next to Condition (b). Since (s^a, s^b) is a Nash equilibrium, we have $\pi^a(s^a, s^b) \geq \pi^a(q^a, s^b)$ for all $q^a \in S^a$. It suffices to show that, for each $r^b \in \text{su}(s^b)$, we must also have $\pi^a(s^a, r^b) \geq \pi^a(q^a, r^b)$ for all $q^a \in S^a$. If so, the strategy s^a is optimal under any measure $\mu \in \mathcal{M}(S^b)$ with $\text{Supp } \mu = \text{su}(s^b)$, and so s^a is admissible with respect to $\text{su}(s^b)$.

Fix $r^b \in \text{su}(s^b)$. By Lemma E2, we have $\pi^b(q^a, r^b) = \pi^b(q^a, s^b)$ for any $q^a \in S^a$. If the profile (q^a, r^b) reaches the same terminal node as (q^a, s^b) , certainly $\pi^a(q^a, r^b) = \pi^a(q^a, s^b)$. If

⁶²That is, the last common predecessor of $\zeta(r^a, r^b)$ and $\zeta(r^a, q^b)$.

not, then the last common predecessor of $\zeta(q^a, r^b)$ and $\zeta(q^a, s^b)$ belongs to b . SPC then implies $\pi^a(q^a, r^b) = \pi^a(q^a, s^b)$. This establishes the above claim.

Finally, Condition (c) is immediate from Lemma E2. ■

Example E1 Here we show that Theorem 2(i) is false in the absence of SPC. Consider the game in Figure E1 (which fails SPC) and the associated sequential probability structure is given in Figure E2.

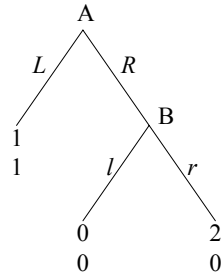


Figure E1

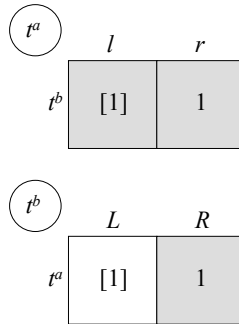


Figure E2

As in the text, the rational strategy-type pairs are shaded. We have:

$$\begin{aligned} R_1^a &= \{(R, t^a)\}, \\ R_1^b &= \{(l, t^b), (r, t^b)\}, \\ A^a(R_1^b) &= \{t^a\}, \\ A^b(R_1^a) &= \{t^b\}. \end{aligned}$$

Thus $(R, t^a, l, t^b) \in \bigcap_m R_m^a \times \bigcap_m R_m^b$, by induction. But (R, l) does not yield a Nash outcome.

Notice that Ann assumes that Bob plays r , while in fact he plays l . Bob is indifferent between the two terminal nodes following these choices. Under SPC, Ann would be indifferent, too, and her incorrect assumption wouldn't matter for her own payoff. But here, without SPC, Ann is not indifferent, Her 'prediction' of her own payoff is, in fact, then wrong. As such, a non-Nash outcome can result under RCAR. We can say that the failure of SPC effectively converts a game of perfect information into one of imperfect information.

Example E2 Here we show that Theorem 2(ii) is false in the absence of SPC. Consider the game in Figure E3 (which fails SPC).

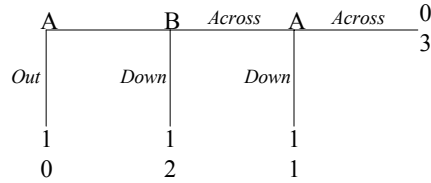


Figure E3

The profile $(Out, Across)$ is an admissible Nash equilibrium. But the unique SAS is $\{Out, Down\} \times \{Down\}$. For Ann, the strategy *Across* is inadmissible, and *Out* and *Down* are copies, so must both be present in an SAS. But for Bob, only *Down* is admissible with respect to $\{Out, Down\}$.

Example E3 Here we show that Theorem 2(ii) cannot be improved to read: Fix a (pure) admissible profile (s^a, s^b) of G that is outcome equivalent to a Nash equilibrium of G . Then there is a sequential probability structure $\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle$ and a state (s^a, t^a, s^b, t^b) at which there is RCAR.

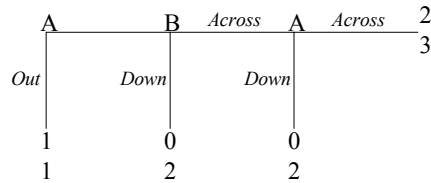


Figure E4

The game in Figure E4 satisfies SPC, and $(Out, Across)$ is an admissible profile that is outcome equivalent to the Nash equilibrium $(Out, Down)$. But the unique SAS is $\{Across\} \times \{Across\}$. For Bob, only *Across* is admissible. For Ann, only *Across* is admissible with respect to $\{Across\}$.

References

- [1] Abreu, D. and D. Pearce. 1984. "On the Inconsistency of Certain Axioms on Solution Concepts for Noncooperative Games." *Journal of Economic Theory*. 34:169-174.
- [2] Anscombe, F. and R. Aumann. 1963. "A Definition of Subjective Probability." *Annals Math. Stat.* 34:199-205.
- [3] Asheim, G. 2001. "Proper Rationalizability in Lexicographic Beliefs." *International Journal of Game Theory*. 30:453-478.
- [4] Asheim, G. and M. Dufwenberg. 2003. "Admissibility and Common Belief." *Games and Economic Behavior*. 42:208-234.
- [5] Aumann, R. 1995. "Backward Induction and Common Knowledge of Rationality." *Games and Economic Behavior*. 8:6-19.
- [6] Aumann, R. 1998. "On the Centipede Game." *Games and Economic Behavior*. 23:97-105.
- [7] Aumann, R. and A. Brandenburger. 1995. "Epistemic Conditions for Nash Equilibrium." *Econometrica*. 63:1161-1180.
- [8] Basu K. and J. Weibull. 1991. "Strategy Subsets Closed Under Rational Behavior." *Economics Letters*. 36:141-146.
- [9] Battigalli, P. 1997. "On Rationalizability in Extensive Games." *Journal of Economic Theory*. 74:40-61.
- [10] Battigalli, P. and M. Siniscalchi. 2002. "Strong Belief and Forward-Induction Reasoning." *Journal of Economic Theory*. 106:356-391.
- [11] Ben Porath, E. 1997. "Rationality, Nash Equilibrium and Backwards Induction in Perfect-Information Games." *Review of Economic Studies*. 64:23-46.
- [12] Bernheim, D. 1984. "Rationalizable Strategic Behavior." *Econometrica*. 52:1007-1028.
- [13] Blume, L., A. Brandenburger, and E. Dekel. 1991a. "Lexicographic Probabilities and Choice Under Uncertainty." *Econometrica*. 59:61-79.
- [14] Blume, L., A. Brandenburger, and E. Dekel. 1991b. "Lexicographic Probabilities and Equilibrium Refinements." *Econometrica*. 59:81-98.
- [15] Börgers, T. and L. Samuelson. 1992. "Cautious Utility Maximization and Iterated Weak Dominance." *International Journal of Game Theory*. 21:13-25.
- [16] Brandenburger, A. and A. Friedenberg. 2003. "The Relationship Between Rationality on the Matrix and the Tree." <http://www.olin.wustl.edu/faculty/Friedenberg/>

- [17] Brandenburger, A. and H.J. Keisler. 2000. "Epistemic Conditions for Iterated Admissibility." Available at www.stern.nyu.edu/~abranden.
- [18] Dalkey, N. 1953. "Equivalence of Information Patterns and Essentially Determinate Games." In H. Kuhn and A. Tucker (eds.), *Contributions to the Theory of Games*, Vol. 2. Princeton: Princeton University Press, 217-244.
- [19] Elmes, S. and P. Reny. 1994. "On the Strategic Equivalence of Extensive Form Games." *Journal of Economic Theory*. 62:1-23.
- [20] Ewerhart, C. 1998. "Rationality and the Definition of Consistent Pairs." *International Journal of Game Theory*. 27:49-59.
- [21] Ewerhart, C. 2002. "Backward Induction and the Game-Theoretic Analysis of Chess." *Games and Economic Behavior*. 39:206-214.
- [22] Feinberg, Y. 2001. "Subjective Reasoning in Games: Parts 1 and 2." Unpublished, Stanford University.
- [23] Friedenberg, A. and J. Hörner. 2002. "When are Two Extensive Games Equivalent? A Preference Based Characterization of the Transformations." Unpublished, Washington University.
- [24] Grünbaum, B. 2002. *Convex Polytopes*. 2nd edition. New York: Springer.
- [25] Halpern, J. 2001. "Substantive Rationality and Backward Induction." *Games and Economic Behavior*. 37:425-435.
- [26] Harsanyi, J. 1973. "Oddness of Number of Equilibrium Points: A New Proof." *International Journal of Game Theory*. 2:235-250.
- [27] Kalmar, L. 1928-9. "Über Eine Schlussweise aus dem Endlichen ins Unendliche." *Acta Universitatis Szegediensis/Sectio Scientiarum Mathematicarum*. 3:121-130.
- [28] Kohlberg, E. and J.F. Mertens. 1986. "On the Strategic Stability of Equilibria." *Econometrica*. 54:1003-1038.
- [29] Kuhn, H. 1950. "Extensive Games." *Proceedings of the National Academy of Sciences*. 36:570-576.
- [30] Kuhn, H. 1953. "Extensive Games and the Problem of Information." In H. Kuhn and A. Tucker (eds.), *Contributions to the Theory of Games*, Vol. 2. Princeton: Princeton University Press, 193-216.
- [31] Mailath, G., L. Samuelson, and J. Swinkels. 1993. "Extensive Form Reasoning in Normal Form Games." *Econometrica*. 61:273-302.
- [32] Marx, L. and J. Swinkels. 1997. "Order Independence for Iterated Weak Dominance." *Games and Economic Behavior*. 18:219-245.

- [33] Mertens, J-F. 1989. "Stable Equilibria—A Reformulation." *Mathematics of Operations Research*. 14:575-625.
- [34] Mertens, J.F. and S. Zamir. 1985. "Formulation of Bayesian Analysis for Games with Incomplete Information." *International Journal of Game Theory*. 14:1-29.
- [35] Osborne, M., and A. Rubinstein. 1994. *A Course in Game Theory*. Cambridge: MIT Press.
- [36] Pearce, D. 1984. "Rationalizable Strategic Behavior and the Problem of Perfection." *Econometrica*. 52:1029-1050.
- [37] Rényi, A. 1955. "On a New Axiomatic Theory of Probability." *Acta Mathematica Academiae Scientiarum Hungaricae*. 6:285-335.
- [38] Rosenthal, R. 1981. "Games of perfect-information, Predatory Pricing, and the Chain-store Paradox." *Journal of Economic Theory*. 25:92-100.
- [39] Samet, D. 1996. "Hypothetical Knowledge and Games with Perfect Information." *Games and Economic Behavior*. 17:230-251.
- [40] Samuelson, L. 1992, "Dominated Strategies and Common Knowledge," *Games and Economic Behavior*. 4:284-313.
- [41] Schwalbe U. and P. Walker. 2001. "Zermelo and the Early History of Game Theory." *Games and Economic Behavior*. 34:123-137.
- [42] Shimoji, M. and J. Watson. 1998. "Conditional Dominance, Rationalizability and Game Forms." *Journal of Economic Theory*. 83:161-195.
- [43] Stalnaker, R. 1998. "Belief Revision in Games: Forward and Backward Induction." *Mathematical Social Sciences*. 36: 31-56.
- [44] Stuart, H. 1997. "Common Belief of Rationality in the Finitely Repeated Prisoners' Dilemma." *Games and Economic Behavior*. 19:133-143.
- [45] Thompson, F. 1952. "Equivalence of Games in Extensive Form." Research Memorandum RM-759, The RAND Corporation.
- [46] Von Neumann, J. and O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.