Abstract

Iterated weak dominance, also called iterated admissibility (IA), has long been known as a powerful but conceptually puzzling solution concept. We give an epistemic foundation for IA. That is, we give conditions on the rationality of the players in the game, on what the players assume about one another’s rationality, etc., under which the players will choose IA strategies. Since IA gives the backward-induction outcome in perfect-information games and the forward-induction outcome in signalling games, our results also give a strategic-form—in fact, invariant—justification for these extensive-form solution concepts. A by-product of our analysis is an impossibility result on common (as opposed to finite-order) assumption of rationality.

1 Introduction

The concept of the iterated elimination of dominated strategies is basic in non-cooperative game theory. The underlying idea is that a rational player won’t play a dominated strategy. This reduces the original game matrix to a submatrix consisting of the undominated strategies. A player who is rational and also assumes that the other players are rational will then avoid any strategies that become dominated in the submatrix—and so on, until only the iteratively undominated strategies remain.

Can this justification for restricting the player’s choices to the iteratively undominated strategies be made formal? That is, can we give a mathematical treatment of the various ingredients—
rationality, assumption, etc.—that yields the restriction to the iteratively undominated strategies as a theorem?

For the case of strong dominance, such formal treatments do indeed exist. But, as we shall review below, giving a foundation for admissibility (i.e. weak dominance) has proved much harder. This is unfortunate because iterated admissibility (henceforth IA) can make sharp predictions in games. For example, it gives the backward-induction outcome in perfect-information games and also in the finitely repeated Prisoner’s Dilemma. It gives the forward-induction outcome in the original example of Kohlberg and Mertens (1986, Section 2.3) and other signalling games.

The goal of this paper is to give a formal epistemic justification for IA. As just indicated, we will, in doing so, also give an epistemic foundation for the use of backward- and forward-induction arguments in games.

Section 2 reviews the conceptual difficulties associated with IA. Sections 3 and 4 give an informal discussion of our resolution. Section 5 is a brief review of the relationship between IA and backward and forward induction. Sections 6 through 9 contain our formal treatment. (The main theorems are stated in Sections 8 and 9.) Section 10 discusses conceptual aspects of our result, and Section 11 reviews related work.

2 The Problem

Consider a finite strategic-form game between two players, Ann and Bob. A strategy of Ann is called admissible if there is no (possibly mixed) strategy of Ann that weakly dominates it. Equivalently, a strategy is admissible if it is optimal for Ann under some probability distribution that puts positive weight on each of Bob’s strategies.

Admissibility is an old criterion in decision theory, defended at length in Kohlberg and Mertens (1986, Section 2.7). It captures the idea that Ann takes all of Bob’s strategies into consideration; she rules none of them out. We take the requirement that a player choose only admissible strategies to be prima facie reasonable, though it can also be justified by an invariance condition. Specifically, we do our decision theory on a specific tree, but require that our analysis be unchanged if done on another tree that yields the same matrix. See Appendix A for details.

If it is reasonable to delete the inadmissible strategies from the game, then it also seems reasonable that any new strategies that become inadmissible in the resulting submatrix should likewise be deleted, and so on until no further deletion is possible, and we are left with the iteratively admissible (IA) strategies. The justification for this reduction should be that it follows from some sort of

---

1 Bernheim (1984) and Pearce (1984) revived interest in iterated strong dominance and their related concept of rationalizability, and gave an informal justification like that in the text here. Fully formal treatments were given by Brandenburger and Dekel (1987), Tan and Werlang (1988), and others.

2 We focus on two players in this paper. Our arguments readily extend to n players; see Section 10(iv).

3 As usual, a (mixed) strategy t of Ann weakly dominates a strategy s of Ann if t gives as high an expected payoff as s for each of Bob’s strategies, and a strictly higher expected payoff than s for at least one of Bob’s strategies. Strategy t strongly dominates s if t gives a strictly higher expected payoff than s for each of Bob’s strategies.
condition of rationality and common assumption of rationality,\footnote{Common assumption of rationality means that each player assumes that each player assumes ... that the players are rational, where the chain of assumptions is continued indefinitely. See Definition 5 in Section 7 for the precise formulation.} where rationality incorporates an admissibility requirement.

Is such a justification in fact possible? Various obstacles have been pointed out in the literature.

<table>
<thead>
<tr>
<th>Bob</th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>3,2</td>
<td>2,2</td>
</tr>
<tr>
<td>Ann</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1,1</td>
<td>0,0</td>
</tr>
<tr>
<td>B</td>
<td>0,0</td>
<td>1,1</td>
</tr>
</tbody>
</table>

Figure 1

i. Order Dependence Often noted is that the order of elimination can affect which strategies survive.\footnote{Here and throughout we use the terms “assumption” and “common assumption” rather than the more customary “belief” and “common belief.” The first terms appear to go better with our formal treatment, as we explain later.} Consider, for example, the game in Figure 1, which is from Kohlberg and Mertens (1986, p.1015). Both $M$ and $B$ are inadmissible (even strongly dominated). Simultaneous deletion of $M$ and $B$ yields the pair of profiles $\{(T, L), (T, R)\}$. But if $M$ is deleted and $B$ is kept, then $L$ must be deleted next, followed by $B$, and only the strategy profile $(T, R)$ survives. Similarly, if $B$ is deleted and $M$ is kept, only the profile $(T, L)$ survives.

The game in Figure 2, which is similar to Example 6 in Samuelson (1992), illustrates another kind of order dependence. The choices $T$ and $L$ are inadmissible, so simultaneous deletion for both players yields the profile $(B, R)$. But if we first delete $T$ for Ann, then $L$ becomes admissible in the resulting submatrix, and we end up with $\{(B, L), (B, R)\}$. Similarly, if we first delete $L$ for Bob, we end up with $\{(T, R), (B, R)\}$. It matters whether we do the deletion simultaneously for Ann and Bob, or first for one player and then for the other.

<table>
<thead>
<tr>
<th>Bob</th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>1,1</td>
<td>0,2</td>
</tr>
<tr>
<td>Ann</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,0</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Figure 2

There is also a more conceptual difficulty with IA. Mas-Colell, Whinston, and Green (1995, p.240) put it this way in their textbook:
The argument for deletion of a weakly dominated strategy for player $i$ is that he contemplates the possibility that every strategy combination of his rivals occurs with positive probability. However, this hypothesis clashes with the logic of iterated deletion, which assumes, precisely, that eliminated strategies are not expected to occur.

We can break this issue down into two parts, which we call the Non-Inclusion and Non-Exclusion problems.

ii. Non-Inclusion Suppose in the game of Figure 1 that we start by deleting $M$ (and not $B$). The issue here is whether we are then justified in deleting $L$. We did so on the basis that Bob would focus on the submatrix that excludes $M$. But doesn’t admissibility require Bob to take all of Ann’s strategies into consideration–i.e. to include all of them? If Bob does include $M$, then $L$ is no longer inadmissible, of course, and so presumably cannot be deleted.

The game in Figure 3, which is Example 8 in Samuelson (1992), makes clear that Non-Inclusion is a separate issue from Order Dependence. There is now a unique order of deletion–namely, the strategy $B$ followed by $R$. But the same conceptual puzzle arises: Is the deletion of $R$ on the second round is justified, given that it appears to depend on Bob’s failing to include the strategy $B$ of Ann?

![Figure 3](T/L 1, 1 1, 0

<table>
<thead>
<tr>
<th></th>
<th>$L$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>1, 1</td>
<td>1, 0</td>
</tr>
<tr>
<td>$B$</td>
<td>1, 0</td>
<td>0, 1</td>
</tr>
</tbody>
</table>

iii. Non-Exclusion There is another difficulty with the prediction $(T, L)$ for the game in Figure 3, which is, in a way, the opposite of the previous one. If we do conclude, the Non-Inclusion problem notwithstanding, that Bob will play $L$ and not $R$, then shouldn’t Ann exclude $R$? But if she does, she becomes indifferent between $T$ and $B$, and so perhaps she can safely play $B$ after all. In short, the elimination of $B$ is justified by the presence of $R$, but if $R$ is excluded this justification is lost.

We can go further. If Ann might indeed play $B$, then mustn’t we bring back the possibility that Bob might play $R$? But if both $L$ and $R$ are back under consideration, then Ann presumably will play $T$ and not $B$, and we seem to have gone round in a circle. This very nice point is due to Samuelson (1992).6

The Non-Exclusion problem also arises in the game of Figure 2. If Ann excludes $L$, she becomes indifferent between $T$ and $B$; if Bob excludes $T$, he becomes indifferent between $L$ and $R$. We again appear at risk of going around in a circle.

The Inclusion and Exclusion issues can be put together as follows. Take the game of Figure 3. We have one argument for Bob’s ruling out Ann’s strategy $B$, and another argument for Bob’s taking $B$ into account. Likewise, we have one argument for Ann’s taking Bob’s strategy $R$ into account, and another argument for Ann’s ruling out $R$. In short, we have two conflicting requirements. We want each player both to include and to exclude a given strategy of the other player.

---

6We come back to Samuelson (1992) in Section 11.
3 Resolution: Infinitesimal Probabilities

The Order Dependence problem is, in a sense, solved automatically in this paper. Our formal set-up, in which each player is rational, assumes that each other player is rational, etc., yields a specific order of elimination. This is simultaneous maximal deletion – i.e. the procedure in which all inadmissible strategies are deleted for all players on each round.

Arguably, this is in any case the natural procedure to consider. Certainly, simultaneous maximal deletion on the first round would seem to follow from simple rationality (without any higher-order conditions), provided this is defined to incorporate admissibility. In the game of Figure 1, for example, rationality alone must deem both M and B irrational, and so surely both choices must get discarded together. It seems hard to imagine a logic that would lead to discarding one choice but not the other. (We are then left with \{ (T, L), (T, R) \} as the set of iteratively admissible profiles.) Likewise, both T and L are irrational in the game of Figure 2 and so should be discarded together. (The iteratively admissible set is then \{ (B, R) \}.) It’s less immediately clear that simultaneous maximal deletion on subsequent rounds is implied by rationality together with higher-order conditions on rationality. But our formal set-up will deliver this result.

The bigger conceptual hurdle for us will be dealing with the conflicting requirements of Inclusion and Exclusion.

Our method for overcoming this difficulty will be to allow a player in a certain sense to both include and exclude a strategy at the same time. We do this by allowing Ann to consider some of Bob’s strategies as infinitely less likely to be played than others, but still possible. (Likewise for Bob, of course.) The strategies that get infinitesimal weight can be viewed as being both excluded (because they get only infinitesimal weight) and included (because they do not get zero weight). To see how this works, go back to the game of Figure 3 and suppose that Ann gives infinitesimal weight to R, while Bob gives infinitesimal weight to B, as in Figure 4.

![Figure 4](image)

7 Both choices are even strongly dominated, as already noted.
8 We have already noted that this game does exhibit the Non-Exclusion problem. Once we have arrived at the profile (B, R), we can certainly ask whether, at this point, we shouldn’t allow Ann to play T and Bob to play L. All we are saying so far is that deletions should be done simultaneously. Whether what we get here, the profile (B, R) is the ‘end of the story’ or leads to some kind of logical circle remains to be seen. We haven’t solved this conceptual problem yet, and indeed are coming to it next.
9 Of course, we don’t rule out that some other epistemic treatment could yield a different order of deletion.
10 Where the probabilities are unnormalized. We emphasize that for us the infinitesimal probabilities in Figure 4 aren’t ‘trembles.’ In the epistemic treatment to come, each player will make a definite choice of strategy. The probabilities will describe the uncertainty in the mind of the other player about what that choice is.

---

5
Recall the Non-Inclusion problem: If Bob considers it possible that Ann plays $B$, is the deletion of $R$ justified? Now it is. Bob does include $B$, since he gives it positive probability. But since he gives $B$ weight $\varepsilon$ only, his optimal choice is $L$ and not $R$. There was also the Non-Exclusion problem: If $R$ is deleted, is it right that Ann can’t play $B$? Now it is. Ann does exclude $R$ since she gives it infinitesimal probability. But since she does give $R$ weight $\varepsilon$, her optimal choice is $T$ and not $B$.

A similar construction works for the games of Figures 1 and 2. (In Figure 1, give each of $M$ and $B$ probability $\varepsilon$. In Figure 2, give each of $T$ and $L$ probability $\varepsilon$.) Of course, all this is only the idea for how a foundation for iterated admissibility can be built; it is far from a complete solution as yet. We don’t want simply to assume that the probabilities are arranged as in Figure 4, for example. That would be assuming the answer. We want to formalize rationality and common assumption of rationality in such a way that the probabilities must, under this condition, turn out this way. Indeed, the game in the next section will show that there is still another conceptual hurdle to be surmounted before we can hope to get the result we are looking for.

4 Resolution Contd.: Completeness

Consider the game in Figure 5 and associated (unnormalized) probabilities. Suppose that Ann assigns probability 1 to Bob’s playing $R$ and probability $\varepsilon$ to his playing $L$. Bob assigns probability 1 to Ann’s playing $T$, probability $\varepsilon$ to her playing $B$, and probability $\varepsilon^2$ to her playing $M$.

Suppose, moreover, that Ann plays $T$ and Bob plays $R$. Intuitively, both players’ choices are rational. Ann optimally chooses $T$ since she assigns probability 1 to Bob’s playing $R$. Bob optimally chooses $R$ since he considers it infinitely more likely than Ann will play $B$ than $M$. Does Ann assume that Bob is rational? This depends on what Ann thinks Bob thinks she will do. Let us suppose that Ann assumes Bob has the $(1, \varepsilon, \varepsilon)$ distribution that he actually has. Then, yes, we can say that Ann assumes that Bob is rational since she thinks it almost certain that Bob plays $R$, and we have already seen that this is optimal for Bob given his distribution. Similarly, if we suppose that Bob assumes Ann has the $(\varepsilon, 1)$ distribution she actually has, then we can say that Bob assumes that Ann is rational. Continuing in this fashion, if we suppose that the given distributions are actually commonly assumed, then the players’ rationality will also be commonly assumed.

![Figure 5](image)

11This is also why we do not want simply to observe that $(T, L)$ is the unique perfect equilibrium of the game in Figure 3, and leave it at that. This wouldn’t solve our problem anyway, since, as is well known, a perfect (or even proper) equilibrium may involve the play of non-IA strategies. (The game in Figure 5 below provides an example.)
It seems that bringing in infinitesimal probabilities is not enough. The unique IA profile in this game is \((M, L)\). Yet we have a set-up with rationality (incorporating admissibility) and common assumption of rationality, where the players choose other strategies.\(^{12}\)

It is not hard to see what has gone wrong. We want Bob to consider \(M\) infinitely more likely than \(B\), rather than vice versa. He will then play \(L\) rather than \(R\), and Ann presumably will be led to play \(M\). There is even a good reason to hope that we can arrange the probabilities this way. The choice \(B\) is deleted on the first round of IA (followed by \(R\) and then \(T\)). Perhaps Bob should indeed consider \(B\) less likely than \(M\), which is not deleted at all. But we have to be careful, as earlier, not to assume the answer. We don’t want simply to require Bob to order Ann’s strategies according to their order of deletion. Let’s go back to Ann, however. Something we can say is that, regardless of what Ann thinks Bob will do, the choice \(B\) can never be rational for her precisely because it is inadmissible (even strongly dominated). By contrast, the choice \(M\) can be rational if Ann assigns sufficiently high probability to \(L\). This gives a solid reason for Bob to consider \(B\) less likely than \(M\)—that he can rationalize Ann’s playing the second strategy but not the first.

Note carefully that with the given \((\varepsilon, 1)\) distribution, both \(M\) and \(B\) are irrational for Ann, and so there is no clear reason for Bob to order these strategies one way or the other.\(^{13}\) The key is to allow Bob to vary Ann’s distribution, so to speak. This suggests the following general approach: formulate rationality and common assumption of rationality in a set-up that includes, in some sense, all possible distributions. This is exactly the concept of completeness that we shall formalize in Section 8.\(^{14}\) It is the remaining missing ingredient. We do then get that the strategies chosen by the players must be IA (Theorem 1 in Section 8).

Let’s recap the two ingredients of our approach:

- **Infinitesimal Probabilities** We want a player to be able, in a sense, to both include and exclude a given strategy of the other player. We do this by allowing one strategy to be considered infinitely less (or more) likely than another strategy. Figure 4 in Section 3 illustrated the idea, where we put the infinitesimal probabilities on the strategies that were eliminated.

- **Completeness** But we also had Figure 5 above, where we deliberately put the probabilities in the wrong order vis-à-vis the order of elimination. Our question was whether we could get the right order without assuming it. And it seemed that by having all possible distributions present—not just certain distributions as in Figure 5—the right order might indeed emerge.

We will begin the formal treatment of this in Section 6. But, even before that, we can see the conceptual meaning of the result we will get. The ‘philosophy’ underneath admissibility is, as noted earlier, that Bob should consider each of Ann’s strategies possible (likewise for Ann). But this is not enough to get IA, as the example of Figure 5 shows. To get IA, Bob must consider possible both every strategy that Ann might play and every distribution that Ann might hold (likewise for Ann). It is this broader “Everything is possible” philosophy that is needed.

---

\(^{12}\)True, we have only described the set-up loosely. But later on (in Section 7), we will show that what we have said holds up in a formal epistemic treatment.

\(^{13}\)In the spirit of properness, one can even argue that, with the given distribution, the choice \(B\) should be considered more likely than \(M\) (as it is) since it is the less costly error. Formally, note that the profile \((T, R)\) is a proper equilibrium.

\(^{14}\)Completeness is closely related to universality, as defined in Mertens and Zamir (1985) and elsewhere. We come back to this in Section 10(ii).
5 Backward and Forward Induction

This section, which is separate from the main argument, is further motivation for IA. We review how IA gives the outcome of interest in a number of well-known games.

Start with perfect-information games and the backward-induction algorithm (BI). The use of BI in such games is completely traditional, of course. Its logical basis also seems clear. Suppose Ann moves at a last node $n$. If she is rational, she will make the payoff-maximizing — i.e. BI — choice at $n$. Let Bob move at the node $m$ preceding $n$. If he is rational and assumes that Ann is rational, he will make the payoff-maximizing choice at $m$, given the BI choice at $n$; i.e. he will make the BI choice at $m$. And so on. The argument sounds obvious, but, on closer inspection, becomes much less obvious. What if the node $m$ is reached only after Ann makes an irrational move? What should Bob think then? Perhaps Bob will assume that Ann will make another irrational move, at her node $n$, in which case he may himself rationally make a non-BI choice at $m$. Arguing this way, it seems that the BI path might not result after all.\footnote{Various authors have challenged the use of BI. See e.g. Basu (1990), Ben Porath (1997), Bicchieri (1989, 1992), Bimmore (1987), Bonanno (1991), Reny (1992), and Rosenthal (1981).}

A number of recent papers offer resolutions of this difficulty and give formal justifications for BI; see, inter alia, Aumann (1995), Battigalli and Siniscalchi (2002), Feinberg (2001), Halpern (2001), Samet (1996), and Stalnaker (1998).\footnote{We return to Battigalli and Siniscalchi (2002) and Stalnaker (1998) in Section 11.} These analyses are conducted on the extensive-form representation of the game. This paper gives a strategic-form justification for BI. Indeed, IA in a perfect-information game gives the BI outcome, provided payoffs are dealt with appropriately.\footnote{Battigalli (1997), Brandenburger and Friedenberg (2003), and Marx and Swinkels (1997) give different proofs, under different conditions on the payoffs.} Thus, conditions on the players’ rationality and assumptions that yield IA, as we provide in this paper, are also conditions for BI.

The strategic-form analysis of BI that we give is in line with the classical approach in game theory. Following Selten (1975), it has become common to couch analyses on the extensive form. But Kohlberg and Mertens (1986) re-emphasized the strategic-form approach. They argued that a good solution concept should be defined on the strategic form and should also be invariant to the addition or deletion of strategies that are convex combinations of other strategies.\footnote{This way, the solution will be the same in any two trees that differ from each other by a sequence of the Dalkey (1953)-Thompson (1952) or Elmes-Reny (1994) transformations, augmented by the convex-combination transformation proposed by Kohlberg-Mertens (op. cit., p.1009).} It can be checked that IA does indeed satisfy Kohlberg-Mertens invariance, and so our analysis of BI is invariant in their sense. These same remarks apply to what we say next about the Prisoner’s Dilemma and forward induction.

Luce and Raiffa (1957, p.109) note that IA gives the Defect-Defect path in the finitely repeated Prisoner’s Dilemma. Thus, our conditions for IA are also conditions for the Nash-equilibrium outcome in this particular game.\footnote{Actually, it can be shown that each player has a unique IA strategy, namely “Defect always.” Stuart (1996) is a different epistemic analysis of the finitely repeated Prisoner’s Dilemma. He uses ordinary probabilities, unlike the current paper, and imposes a mutual absolute continuity condition.} (Of course, IA doesn’t always coincide with Nash equilibrium.)
Turn next to signalling games and the use of forward-induction reasoning (FI). Osborne and Rubinstein (1994, pp.110-111) observe that IA gives the FI outcome in the original example of Kohlberg and Mertens (1986, Section 2.3). Indeed, Figure 5 above is the strategic form of this game, and the unique IA profile there, namely $(M, L)$, is also the FI profile. Van Damme (1989) and Ben Porath and Dekel (1992) employ IA to generate the FI outcome in their striking Burn-a-Dollar game. IA also gives the desired outcome in the well-known Beer-Quiche game of Cho and Kreps (1987).

6 Lexicographic Probability Systems

We now begin the formal development. The first step is to develop a suitable one-person decision theory from which we can then build a multi-person ("interactive") formalism.

In our heuristic treatment in Sections 3 and 4, we brought infinitesimal probabilities into the analysis. The formal counterpart will be the following lexicographic decision theory. Ann (or Bob) possesses not one probability measure on the state space, but a sequence of measures, with the property that every state receives positive probability under one (and only one) measure. Ann will prefer strategy $s$ to strategy $t$ if the sequence of expected payoffs associated with $s$ is lexicographically greater than the sequence associated with $t$. Thus, Ann views states that receive positive probability under her first measure as infinitely more likely than states that get positive probability under her second measure, which she views as infinitely more likely than states that get positive probability under her third measure, and so on. Alternatively put, the first measure can be thought of as Ann’s primary ‘hypothesis’ about the true state. Ann is almost, but not quite, certain of her primary hypothesis. So, she also forms a secondary hypothesis, as represented by her second measure. And so on.

The Inclusion criterion from Section 3 will be satisfied: no state is ruled out completely since every state gets positive probability under some measure. Exclusion will also be satisfied since some states are considered infinitely less likely than others.

We now give a formal definition. By a Polish space we will mean a complete separable metric space. Given a Polish space $\Omega$, let $\mathcal{M}(\Omega)$ be the space of all Borel probability measures on $\Omega$ with the Prohorov metric. Then $\mathcal{M}(\Omega)$ is again Polish, with diameter $\leq 1$. Another standard fact: Given a Polish space $\Omega$ with metric $d$ and diameter $\leq 1$, 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 $\leq 1$. By the empty measure on $\Omega$ we will mean the function that assigns 0 to every Borel subset of $\Omega$. 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 $\leq 1$ (and the empty measure as an isolated point), and so $(\mathcal{M}'(\Omega))^\mathbb{N}$ is also Polish. Now define $\mathcal{M}^\omega(\Omega)$ to be the subspace of $(\mathcal{M}'(\Omega))^\mathbb{N}$ consisting of all infinite sequences

---

20 When IA is performed on the two-person strategic form, where one player is the Sender and the other is the Receiver, and the Sender’s payoffs are calculated as expected payoffs before Nature chooses the Sender’s type. The significance of doing IA on the two-person game—rather than on the three-person game obtained by treating the Sender’s two types as separate players—is that the Sender’s types are required to have the same probability distribution over the Receiver’s strategies. This is certainly a restriction, but it is a long way from assuming equilibrium. The IA analysis does differ from an equilibrium analysis. (Feinberg [2001] gives a similar analysis of Beer-Quiche to ours, albeit on the extensive not strategic form.)

21 In the finite case. For the general condition, see Definition 1 below. The next section describes what the relevant state space and states are in the game-theoretic context we are interested in.

22 Note that we fix a metric. (In Kechris [1995], a space is Polish if it is separable and completely metrizable.)
from $\mathcal{M}(\Omega)$ and all finite sequences from $\mathcal{M}(\Omega)$ followed by infinitely many empty measures. That is: $\sigma = (\mu_0, \ldots) \in \mathcal{M}^\infty(\Omega)$ if and only if for each $n$, if $\mu_n$ is the empty measure then $\mu_{n+1}$ is the empty measure. Then $\mathcal{M}^\infty(\Omega)$ with the induced metric is a closed subset of $(\mathcal{M}')(\Omega)^\infty$, and hence is a Polish space.

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

**Definition 1** The set $\mathcal{L}(\Omega)$ of **lexicographic probability systems (LPS’s)** is the set of all $\sigma = (\mu_0, \ldots) \in \mathcal{M}^\infty(\Omega)$ such that:

(a) $0 < l(\sigma) < \infty$;

(b) there are Borel sets $U_i$ in $\Omega$, for $i < l(\sigma)$, such that $\mu_i(U_i) = 1$ and $\mu_i(U_j) = 0$ for $i \neq j$;

(c) $\sigma$ has full support, that is, $\Omega = \bigcup_{i<l(\sigma)} \text{Supp } \mu_i$.

For a given $\sigma = (\mu_0, \ldots, \mu_{n-1}) \in \mathcal{L}(\Omega)$, the first measure $\mu_0$ is the player’s primary hypothesis about the true state, the measure $\mu_1$ is his secondary hypothesis, and so on until the $n$th hypothesis $\mu_{n-1}$.23 Condition (c) is the requirement, appropriately stated for an infinite space $\Omega$, that the player rules nothing out.

LPS’s on finite spaces were introduced by Blume, Brandenburger, and Dekel (1991a), where an axiomatic derivation is provided in terms of the player’s preferences over acts.24 Our Definition 1 here is designed to be a natural extension to infinite spaces of the BBD (1991a) definition.25 Infinite spaces will play an essential role later on.

**Example 1** We give a simple, finite example adapted from BBD (1991a). (Example 2 below involves infinite $\Omega$.) The state space $\Omega = \{\text{Heads}, \text{Tails}, \text{Edge}\}$, and Ann has the LPS $\sigma = (\mu_0, \mu_1)$, where the primary measure $\mu_0$ assigns probability $\frac{1}{2}$ each to Heads and Tails, and the secondary measure $\mu_1$ assigns probability 1 to Edge. Consider, now, two bets. Bet $x$ pays off $\$1$ if Heads comes up and nothing otherwise. Bet $y$ pays off $\$2$ if Tails or Edge comes up and nothing otherwise. If Ann lexicographically maximizes expected value, then she strictly prefers $x$ to $y$ whenever $v > 1$. This is true even if $v$ is arbitrarily close to 1, which is a way of saying that the state Edge is considered infinitely unlikely. But Ann strictly prefers $y$ to $x$ when $v \leq 1$. And this is true, in particular, when $v = 1$, which says that Edge is not ruled out completely.

Let us record some immediate properties of LPS’s.26

---

23 Please note the abuse of notation: Here and later we write an LPS $\sigma$ as a finite sequence by leaving out the empty measures.

24 They modify the axioms of the Anscombe and Aumann (1963) version of subjective expected utility theory.

A note on terminology: BBD (1991a) 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 will use the simpler term LPS throughout. See also Section 10(iii).

25 Halpern (2003) has suggested a further extension, where the measures are indexed by some initial segment of the ordinals. An axiomatic treatment of LPS’s on infinite $\Omega$ would certainly be of interest, but we do not attempt one here.

26 Proofs are in Appendix C.
Lemma 1 In Definition 1, the sets $U_i$ may be chosen so that they form a partition of $\Omega$ and each $U_i$ is contained in $\text{Supp}\mu_i$.

Lemma 2 In Definition 1, condition (c) may be replaced by:

(c') for each nonempty open set $U$ there is an $i$ such that $\mu_i(U) > 0$.

Lemma 3 The set $\mathcal{L}(\Omega)$ is Borel.

To prove our theorem on IA, we will need to formalize the concepts of rationality and assumption when players have LPS's. Rationality means lexicographically maximizing expected utility, as already noted. Definition 4 in the next section gives the formal definition in the game context. Next is assumption. Take a simple example: $\Omega = \{a, b, c, d\}$ and $\sigma = (\mu_0, \mu_1, \mu_2)$, where the primary measure $\mu_0$ is given by $\mu_0(a) = 1$, the secondary measure $\mu_1$ is given by $\mu_1(b) = \frac{1}{2}$ and $\mu_1(c) = \frac{1}{3}$, and the tertiary measure $\mu_2$ is given by $\mu_2(d) = 1$. Arguably, the only event $E \subseteq \Omega$ that Ann believes is $\Omega$ itself. She gives every state positive probability under some measure, and so rules no state out. This noted, we can still reasonably say that Ann assumes $E$ if she considers states not in $E$ to be much less likely—in fact, infinitely less likely—than states in $E$. Ann recognizes that $E$ may not happen. Nevertheless, she is prepared to ‘count on’ $E$ versus not-$E$. This seems to correspond quite well to the intuitive, everyday meaning of assumption.

With this definition, Ann assumes $\{a\}$ since she gives $b$, $c$, and $d$ infinitely less weight than $a$. Ann also assumes $\{a, b, c\}$ since $d$ gets infinitely less weight than $a$, $b$, or $c$. If you like, Ann starts with her initial assumption $\{a\}$, but if challenged to ‘think more,’ she brings in the states $b$ and $c$ since these are of the next order of likelihood. Finally, Ann also assumes $\{a, b, c, d\}$. She does not assume any other event. For example, it wouldn’t make sense to say that Ann assumes $\{a, b\}$, given that $c$ is of the same order of likelihood as $b$.

Our condition for $E$ to be assumed is obviously equivalent to requiring that one of the following holds: (i) $\mu_0(E) = 1$ and $\mu_1(E) = \mu_2(E) = 0$; or (ii) $\mu_0(E) = \mu_1(E) = 1$ and $\mu_2(E) = 0$; or (iii) $\mu_0(E) = \mu_1(E) = \mu_2(E) = 1$. This leads to our general definition:

Definition 2 A set $E$ is assumed under $\sigma$ (at level $j$) if $\sigma \in \mathcal{L}(\Omega)$ and there is a $j < l(\sigma)$ such that:

---

27 Formally, there are no (nonempty) Savage-null events. See Appendix B.
28 Appendix B gives a preference-based definition of assumption, making precise the idea of Ann’s betting on $E$. It is shown to be equivalent to the definition we are proposing here (Definition 2 below).
29 But we say Ann assumes $\{a, b, c\}$ and not just $\{b, c\}$. The latter wouldn’t be sensible since Ann clearly puts $a$ ahead of $b$ and $c$.
30 Assumption, then, is non-monotonic in that Ann may assume $E$, and not assume $F$, even though $F$ includes $E$. (Ann assumes $\{a\}$ but not $\{a, b\}$.) This fits with our using the term “assume” rather than “believe.” In everyday usage, if we say that Ann believes it will be sunny, then surely we would agree that, as a matter of logic, Ann must also believe that it will be sunny or cloudy. But we might say that Ann assumes it will be sunny, while not wishing to say that Ann also assumes it will be sunny or cloudy. The idea is precisely that Ann considers bad weather (clouds or rain) much less likely than good weather. She is counting on sun. She recognizes that bad weather is also possible, but if she brings in one possibility (clouds) she also wants to bring in the other (rain).

At the formal level, the monotonicity of belief is ensured if we reserve the term for events $E$ such that not-$E$ is Savage-null.
(a) $\mu_i(E) = 1$ for all $i \leq j$,
(b) $\mu_i(E) = 0$ for all $j < i < l(\sigma)$,
(c) $E \subseteq \bigcup_{i \leq j} \text{Supp} \mu_i$.

Conditions (a) and (b) are as just described. If $\Omega$ is finite, condition (b) implies condition (c). But this is not true if $\Omega$ is infinite. The next example illustrates the role that (c) then plays.\footnote{Appendix B gives a preference-based characterization of condition (c).}

**Example 2** The state space $\Omega = \{x, y, z\} \times [0, 1]$, and Ann has the LPS $\sigma = (\mu_0, \mu_1)$, where the primary measure $\mu_0$ is ‘uniform on’ $\{x\} \times [0, 1]$ and the secondary measure $\mu_1$ is ‘uniform on’ $\{y, z\} \times [0, 1]$. (Formally, the measures $\mu_0$ and $\mu_1$ are constructed in the obvious way starting from Lebesgue measure on $[0, 1]$.) Consider the event $E = (\{x\} \times [0, 1]) \cup \{(y, 0)\}$. We might be tempted to say that Ann assumes $E$, since $\mu_0(E) = 1$ and $\mu_1(E) = 0$, and so conditions (a) and (b) of Definition 2 are satisfied.

But without condition (c), assumption behaves badly under marginalization. To see this, let’s look at marginals on $\{x, y, z\}$, and ask whether Ann assumes $\text{proj}_{\{x, y, z\}} E$ under the LPS $(\text{ marg}_{\{x, y, z\}} \mu_0, \text{ marg}_{\{x, y, z\}} \mu_1)$. More precisely, do conditions (a) and (b) continue to hold? The answer is no, since $\text{ marg}_{\{x, y, z\}} \mu_0(\{x, y\}) = 1$ but $\text{ marg}_{\{x, y, z\}} \mu_1(\{x, y\}) = \frac{1}{2}$. Condition (c) rules out this effect. Indeed, in the current example we have $\text{Supp} \mu_0 = \{x\} \times [0, 1] \nsubseteq E$, so that (c) fails and we conclude that Ann does not assume $E$ according to Definition 2.

This example points to an interesting asymmetry caused by condition (c) of our definition of assumption. Start with the event $D = \text{Supp} \mu_0 = \{x\} \times [0, 1]$. Certainly, Ann assumes $D$ according to Definition 2. The event $E$ differs from $D$ by the addition of a state $(y, 0)$ that is ‘negligible.’ (It receives probability 0 under both $\mu_0$ and $\mu_1$.) We could also consider an event $F = D \setminus \{(x, 0)\}$ that differs from $D$ by the removal of a state $(x, 0)$ that is negligible in the same sense. Ann does not assume $E$, but she does assume $F$ since $\mu_0(F) = 1$, $\mu_1(F) = 0$, and $F \subseteq \text{Supp} \mu_0$. Evidently, addition and removal of negligible states behave differently.\footnote{This asymmetry makes conceptual sense. While building her primary measure $\mu_0$, Ann thinks about states that belong to its support $D$; she then assumes $D$. Now remove from $D$ a negligible state, such as $(x, 0)$. Ann has already thought about this state since it lies in $D$, and deemed it negligible. So it is natural to say that she [still] assumes $D$ minus this state. But Ann has not necessarily yet thought about a state like $(y, 0)$, which lies outside $D$. It isn’t obvious that we should require Ann to assume $D$ plus this state, and, indeed, we have seen that doing so would cause difficulties.}

### 7 Interactive Structures

This section develops a formalism, involving LPS’s, with which to talk about the rationality of the players in a game, what the players assume about one another’s rationality, etc.

Fix a two-player finite strategic-form game $(S^a, S^b, \pi^a, \pi^b)$, where $S^a, S^b$ are the strategy sets and $\pi^a, \pi^b$ are the payoff functions of Ann and Bob, respectively.\footnote{We make no attempt in this paper to treat infinite games, where various new issues would arise.} A probability structure adds to this structure a set of types for each player, where a type of Ann is associated with a probability measure on strategies and types of Bob, and vice versa. In the literature, this is by now a standard
epistemic model for treating uncertainty about strategy choices.\textsuperscript{34} The following definition changes the model to allow the players to have not single measures but sequences of measures, as is the case with LPS’s.

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

\[
(S^a, S^b, T^a, T^b, \lambda^a, \lambda^b)
\]

where \( T^a \) and \( T^b \) are nonempty Polish spaces, \( \lambda^a \) is a continuous mapping from \( T^a \) to \( \mathcal{M}^\infty(S^b \times T^b) \), and \( \lambda^b \) is a continuous mapping from \( T^b \) to \( \mathcal{M}^\infty(S^a \times T^a) \). Members of \( S^a \times T^a \times S^b \times T^b \) are called **types**. Members of \( S^a \times T^a \times S^b \times T^b \) are called **states** (of the world).

The definitions and results to come all have counterparts with \( a \) and \( b \) reversed.

**Definition 4** A pair \((s^a, t^a)\) \( \in S^a \times T^a \) is **rational** (with respect to the payoff function \( \pi^a \)) if \( \lambda^a(t^a) \in \mathcal{L}(S^b \times T^b) \) and for every \( r^a \in S^a \),

\[
\left( \int_{S^b \times T^b} \pi^a(s^a, s^b) d\mu_i(s^b, t^b) \right)_{i=0}^{n-1} \geq_L \left( \int_{S^b \times T^b} \pi^a(r^a, s^b) d\mu_i(s^b, t^b) \right)_{i=0}^{n-1},
\]

where we write \( \lambda^a(t^a) = (\mu_0, \ldots, \mu_{n-1}) \).

**Remark 1** Note that \((s^a, t^a)\) is rational if and only if 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 \( \mu_i \).

In words, rationality is a property of a strategy-type pair of a player. And a particular strategy-type pair \((s^a, t^a)\) of Ann is rational if \( s^a \) lexicographically maximizes her expected payoff, calculated under the LPS \( \lambda^a(t^a) \) associated with her type \( t^a \).

We now formalize common assumption of rationality. For \( E \subseteq S^b \times T^b \), let \( A^a(E) \) be the set of types of Ann that assume \( E \):

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

Define \( R^a_i \) to be the set of all rational pairs \((s^a, t^a)\), and define \( R^a_m \) inductively by

\[
R^a_{m+1} = R^a_1 \cap \bigcup_{i \leq m} A^a(R^b_i).
\]

\textsuperscript{34}There are also models that treat uncertainty about both strategies and payoff functions, but this is not our interest here. \textsuperscript{35}If \( x = (x_0, \ldots, x_{n-1}) \) and \( y = (y_0, \ldots, y_{n-1}) \), then \( x \geq_L y \) iff \( y_j > x_j \) implies \( x_k > y_k \) for some \( k < j \). \textsuperscript{36}Notice that, while rationality is defined in terms of LPS’s, we allow a sequential probability structure (Definition 3) to include sequences of measures that aren’t LPS’s. Each type \( t^a \) is associated via \( \lambda^a \) with an element of \( \mathcal{M}^\infty(S^b \times T^b) \), but not necessarily with an element of the subset \( \mathcal{L}(S^b \times T^b) \). We explain the reason for this in Section 10(i).
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. Say there is rationality and common assumption of rationality (RCAR) if there is rationality and \(m\)th-order assumption of rationality for all \(m\).

We can break this definition down into statements for Ann and Bob, respectively. We can say that rationality and \(m\)th-order assumption of rationality hold with respect to Ann if: (i) she is rational, and (ii) for each \(i < m\) she assumes that rationality and \(i\)th-order assumption of rationality hold with respect to Bob. We talk about rationality and \(m\)th-order assumption of rationality with respect to Bob in exactly the same way. Note that part (i) here is just the maximization condition, saying that the strategy \(s^a\) lexicographically maximizes Ann’s expected payoff, calculated under her LPS \(\lambda^a(t^a)\).

Lemma 4
(a) For each \(m\),

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

(b) For each \(m\) the set \(R_m^a\) is Borel in \(S^a \times T^a\).

We now give two examples of our epistemic formalism.

Example 3 Consider three-legged Centipede (Rosenthal [1981]) in Figure 6 and the associated sequential probability structure in Figure 7 below.

\[
\begin{array}{c|c|c}
\text{A} & \text{B} & \text{Across} \\
\hline
\text{Out} & \text{Down} & \text{Down} \\
\end{array}
\]

Figure 6

There are two types of each player: \(T^a = \{t^a, u^a\}\) and \(T^b = \{t^b, u^b\}\). The LPS \(\lambda^a(t^a)\) is the three-level LPS depicted in the top-left matrix (where square parentheses indicate the second-order measure and double square parentheses indicate the third-order measure). Thus, the type \(t^a\) assigns first-order probability one to (Across, \(u^b\)), second-order probability one to (Down, \(t^b\)), and equal third-order probability to the remaining two strategy-type pairs of Bob. (We explain in a moment why certain cells are shaded.) The LPS \(\lambda^a(u^a)\) is also a three-level LPS as depicted in the top-right matrix. Turning to Bob, the LPS’s \(\lambda^b(t^b)\) and \(\lambda^b(u^b)\) are both three-level LPS’s as depicted in the third and fourth matrices, respectively. We have:

\[
\begin{align*}
R_1^a &= \{(\text{Down}, t^a), (\text{Out}, u^a)\}, \\
R_2^a &= \{(\text{Down}, t^b), (\text{Across}, u^b)\}, \\
A^a(R_1^b) &= \{t^a, u^a\}, \\
A^b(R_1^a) &= \{t^b\}, \\
R_2^a &= R_1^b, \\
R_2^b &= \{(\text{Down}, t^b)\},
\end{align*}
\]

14
The elements of $R_a^3$ and $R_b^3$—i.e. the rational strategy-type pairs—are shaded. This should make it easier to verify the rest of the calculations.) Now let the true state be $(\text{Down}, t^a, \text{Down}, t^b)$. Then $(\text{Down}, t^a, \text{Down}, t^b) \in R_a^2 \times R_b^2$, but $(\text{Down}, t^a, \text{Down}, t^b) \notin R_a^3 \times R_b^3$. So, we have rationality and first-order assumption of rationality at this state, but not second-order assumption of rationality.

The unique IA profile in Centipede is $(\text{Out}, \text{Down})$. This induces the BI path, on which Ann ends the game at her first node. In our set-up, Ann plays across at her first node (and then Bob ends the game). But there is no contradiction here since we do not have RCAR (let alone completeness), and so the players need not play their IA strategies.

The next example formalizes our earlier intuitive discussion of the game in Figure 5.

**Example 4** Consider again the game in Figure 5, and the associated sequential probability structure depicted in Figure 8. Here there is just 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, and the LPS $\lambda^b(t^b)$ is the three-level LPS depicted in the bottom matrix.

**Figure 8**

The rational strategy-type pairs are again shaded, and we have:

- $R_1^a = \{(T, t^a)\}$,
- $R_1^b = \{(R, t^b)\}$,
- $A^a(R_1^a) = \{t^a\}$,
- $A^b(R_1^b) = \{t^b\}$,
- $R_2^a = R_1^a$,
- $R_2^b = R_1^b$.

It follows by induction that $R_m^a = R_1^a$ and $R_m^b = R_1^b$ for all $m$. Since $(T, t^a, R, t^b) \in R_1^a \times R_1^b$, there is actually RCAR at the state $(T, t^a, R, t^b)$. But the unique IA profile is $(M, L)$, as we noted in Section 4. Thus, we have now shown formally what we previously argued informally— that RCAR need not yield IA. We also suggested that the missing ingredient was a completeness condition. We formalize this next.

### 8 Main Result

We now state the completeness condition and then present our main result, giving epistemic conditions for IA.

**Definition 6** Fix nonempty finite sets $S^a, S^b$ and an associated sequential probability structure

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

The structure is **complete** if $\lambda^a$ and $\lambda^b$ are onto.

In words, a structure is complete if all possible types that can be defined in the structure are, in fact, present in the structure. That is, for every sequence of probability measures on $S^b \times T^b$ (the product of Bob’s strategy and type spaces), there is a type $t^a$ of Ann with that sequence, and similarly with Ann and Bob interchanged.
Proposition 1 For any nonempty sets \(S^a, S^b\) there is a complete \((S^a, S^b)\)-based sequential probability structure.

A note on cardinalities: We are assuming throughout that the strategy spaces \(S^a\) and \(S^b\) are finite. Even so, it is clear that in a complete structure the type spaces \(T^a\) and \(T^b\) must be uncountably infinite. This is why we defined LPS’s on infinite spaces (Definition 1) and not just finite spaces. It is also why we defined “assumption” for infinite spaces (Definition 2).\(^{37}\)

To state our main result, we need some standard definitions on admissibility.

Definition 7 A strategy \(s^a \in S^a\) is admissible if there is no \(\mu \in \mathcal{M}(S^a)\) such that
\[
\sum_{r \in S^a} \pi^a(r, s^b)\mu(r) \geq \pi^a(s^a, s^b) \quad \text{for every } s^b \in S^b,
\sum_{r \in S^a} \pi^a(r, s^b)\mu(r) > \pi^a(s^a, s^b) \quad \text{for some } s^b \in S^b.
\]

Given \(X \subseteq S^a\) and \(Y \subseteq S^b\), let \(\pi^a\)(\(X \times Y\)) denote the restriction of \(\pi^a\) to \(X \times Y\) and \(\pi^b\)(\(Y \times X\)) the restriction of \(\pi^b\) to \(Y \times X\). Now, let \(S^a_0 = S^a\) and define \(S^a_m\) inductively by
\[
S^a_{m+1} = \{s^a \in S^a_m : s^a \text{ is admissible in the game } \langle S^a_m, S^b_m, \pi^a\mid (S^a_m \times S^b_m), \pi^b\mid (S^b_m \times S^a_m) \rangle \}.
\]

Definition 8 A strategy \(s^a \in S^a_m\) is called \(m\)-admissible. A strategy is iteratively admissible (IA) if it is \(m\)-admissible for all \(m\).

Remark 2 Note that since \(S^a\) and \(S^b\) are finite, there is an \(M\) such that \(S^a_M = S^a\) and \(S^b_M = S^b\) for all \(m \geq M\).

The main result of the paper is:

Theorem 1 Fix a complete \((S^a, S^b)\)-based sequential probability structure
\[
\langle S^a, S^b, T^a, T^b, \lambda^a, \lambda^b \rangle
\]
and payoff functions \(\pi^a\) and \(\pi^b\).

(i) If there is rationality and \(m\)th-order assumption of rationality at the state \((s^a, t^a, s^b, t^b)\), then the strategies \(s^a\) and \(s^b\) are \((m+1)\)-admissible.

(ii) If \(s^a\) and \(s^b\) are \((m+1)\)-admissible strategies, then there is a state \((s^a, t^a, s^b, t^b)\) at which there is rationality and \(m\)th-order assumption of rationality.

Corollary 1

(i) If there is rationality and \(m\)th-order assumption of rationality at the state \((s^a, t^a, s^b, t^b)\), for some \(m \geq M - 1\), then the strategies \(s^a\) and \(s^b\) are IA.

(ii) If \(s^a\) and \(s^b\) are IA strategies, then for any \(m \geq M - 1\) there is a state \((s^a, t^a, s^b, t^b)\) at which there is rationality and \(m\)th-order assumption of rationality.

Appendix D gives the idea of the proof of Theorem 1 and then the formal proof.\(^{37}\)

\(^{37}\)Condition (c) of Definition 2 has bite in the infinite case, as we noted. We make essential use of it in the proof of Theorem 1 below.
9 An Impossibility Result

Theorem 1 actually talks about rationality and finite-order assumption of rationality, not common assumption of rationality. As Corollary 1 indicates, this is enough for us to get to the IA strategies: RCAR is not needed. Of course, this is quite intuitive. In a finite game, only finitely many rounds of elimination of inadmissible strategies are possible. In parallel fashion, an epistemic treatment of IA should need only exactly the sequence $\Omega$ considered. But Theorem 2 says we can’t have both RCAR and completeness. Also, without RCAR, completeness is possible. This is Proposition 1 in Section 8, and also the assumption of rationality. As Corollary 1 indicates, this is enough for us to get to the IA strategies: Theorem 1 actually talks about rationality and $\pi$-functions.

If we have a state $(s^a, t^b, s^a, t^b)$ at which there is RCAR, then $t^a$ must assume each of the decreasing sequence of events $R^a_1, R^a_2, \ldots$. To understand what this involves, take an abstract example: consider $\Omega = [0, 1]$ and the decreasing sequence of events $E_i = [0, \frac{1}{i+1}]$ for $i = 0, 1, \ldots$. For any finite sequence $E_0 \supseteq \cdots \supseteq E_{n-1}$, there is no difficulty in constructing an LPS $\sigma \in \mathcal{L}(\Omega)$ under which each $E_i$ is assumed. Just let the measure $\mu_0$ be uniform on $E_{n-1} = [0, \frac{1}{n}]$, the measure $\mu_1$ be uniform on $E_{n-2} \setminus E_{n-1} = (\frac{1}{n}, \frac{1}{n+1})$, \ldots, and the measure $\mu_{n-1}$ be uniform on $E_0 \setminus E_1 = (\frac{1}{2}, 1]$. But this construction also suggests that there can be no LPS $\sigma$ under which each member of the infinite sequence is assumed. This is correct. At some point in the construction we must ‘hit’ the first measure in the LPS, at which point there is no next (more likely) order of likelihood.

This example suggests that it may actually be impossible to have RCAR, depending on how exactly the sequence $R^a_1, R^a_2, \ldots$ behaves. It turns out that, subject to the following non-triviality condition, completeness provides just such a situation.

Definition 9 Say that player $a$ is indifferent if $\pi^n(r^a, s^b) = \pi^n(s^a, s^b)$ for all $r^a, s^a, s^b$.

Theorem 2 Fix a complete $(S^a, S^b)$-based sequential probability structure

\[ <S^a, S^b, T^a, T^b, \lambda^a, \lambda^b> \]

and payoff functions $\pi^a$ and $\pi^b$. If player $a$ is not indifferent, then there is no state at which there is RCAR. In fact,

\[ \bigcap_{m=1}^{\infty} R^n_m = \bigcap_{m=1}^{\infty} R^b_m = \emptyset. \]

Note well that without completeness, RCAR is possible, as Example 4 in Section 7 showed. Also, without RCAR, completeness is possible. This is Proposition 1 in Section 8, and also the context of Theorem 1 there. But Theorem 2 says we can’t have both RCAR and completeness.

---

38 In our informal treatment of Sections 2 through 4, we talked about common assumption of rationality. We could have guessed that finite-order assumption of rationality should be all that is needed, but it was simpler to postpone this point.

39 One motivation is that the appropriate value of $M$ in Corollary 1 does depend on the game $(S^a, S^b, \pi^a, \pi^b)$ in question. We can ask instead for an epistemic condition for IA that is independent of the particular game. The natural candidate, of course, is RCAR (together with completeness).

40 Formally, let $E_m$ be assumed at level $j$, so that $\mu_i(E_m) = 1$ for $i \leq j$ and $\mu_i(E_m) = 0$ for $i > j$. It is immediate that $E_{m+1}$ must then be assumed at a level $k \leq j$. If $k = j$, then the open interval $(\frac{1}{m+1}, \frac{1}{m})$ gets probability 0 under every measure in $\sigma$, contradicting Lemma 2. Thus $k < j$, and, continuing in this way, we will run out of levels.

Note that this argument would still hold if we allowed an LPS to have infinitely many measures—i.e. if we relaxed condition (a) in Definition 1. The issue is not that we have assumed a finite number of hypotheses, but that there is a primary hypothesis, secondary hypothesis, etc.

41 The sets $R^a_1, R^a_2, \ldots$ did not actually decrease, but remained equal. Under completeness, these sets decrease strictly and ‘sufficiently sharply’ to get the impossibility. See the proof of Theorem 2 in Appendix E.
What is this impossibility result telling us? Perhaps it is saying something about the ‘limits of analysis of games.’ True, this limit doesn’t prevent us from getting a foundation for IA. (For that, as we already emphasized, finite-order assumption of rationality is both adequate and natural.) But the impossibility seems worth noting in its own right and may also turn out to matter for some question beyond this paper.

Finally, a natural question is what does RCAR alone, without completeness, yield. (We know it doesn’t yield IA—recall Example 4.) Brandenburger and Friedenberg (2002) show that the answer is an object they call a “self-admissible set.” (The set of IA profiles in a game is one self-admissible set, but, often, there are several self-admissible sets.)

## 10 Discussion

This section discusses some technical and conceptual aspects of Theorem 1.

### i. Definition of an Interactive Structure

Let us go back to the question of why we allowed an interactive structure (Definition 3) to include sequences of measures that aren’t LPS’s. The technical answer is that while $\mathcal{M}^\infty(\Omega)$ is a complete separable metric space whenever $\Omega$ is, the natural metric on $\mathcal{L}(\Omega)$ is not complete.\(^\text{42}\) As a result, if we made $\lambda$ a map from $T^a$ to $\mathcal{L}(S^b \times T^b)$ rather than $\mathcal{M}^\infty(S^b \times T^b)$, our proof of the existence of a complete structure (Proposition 1) would break down.\(^\text{43}\) Conceptually, too, it makes sense to include non-lexicographic sequences in a complete structure. Take a sequence of LPS’s that approaches a non-lexicographic sequence $\sigma$. Since every member of the sequence will be present in a complete structure, it seems reasonable to require that $\sigma$ also be present. True, a strategy-type pair will be irrational if the associated measure sequence is non-lexicographic. (Refer back to Definition 4.) But irrational pairs have a natural place in an interactive structure.

### ii. Completeness and Universality

In the literature, the more common concept of a “space of all possible types” is the universal space (Armbruster and Böge [1979], Böge and Eisele [1979], Mertens and Zamir [1985], Brandenburger and Dekel [1993], Heifetz [1993], Battigalli and Siniscalchi [1999], et al.). Why do we define and use completeness instead?\(^\text{44}\) The reason is that it is exactly the completeness property that is used in our proofs, so it seems right to isolate and define this property (Definition 6). Also, we prove the existence of a complete sequential probability structure in a few lines—see the proof of Proposition 1 in Appendix C.\(^\text{45}\)

### iii. Definition of an LPS

We defined an LPS to be a finite sequence of probability measures. (This was condition (a) in Definition 1.) One could also imagine a definition involving infinite sequences. Our main reason for staying with finite sequences is that this is general enough for our purposes. We get the kind of result on IA (Theorem 1) that we want. But it would certainly be

---

\(^\text{42}\)Let $\Omega = \{x, y\}$, and consider a sequence $\sigma^m \in \mathcal{L}(\Omega)$, for $m = 1, 2, \ldots$, where each $\sigma^m = (\mu^m_0) \text{ and } \mu^m_0(x) = 1 - \frac{1}{m+1}$. Then $\sigma^m \rightarrow \sigma \in \mathcal{M}^\infty(\Omega)$, where $\sigma = (\mu_0)$ and $\mu_0(x) = 1$. Note that $\sigma \notin \mathcal{L}(\Omega)$ since it does not have full support. This shows that $\mathcal{L}(\Omega)$ is not closed in $\mathcal{M}^\infty(\Omega)$, so $\mathcal{L}(\Omega)$ with the induced metric fails to be complete.

\(^\text{43}\)We could replace $\lambda$ as given in Proposition 1 by a new function that agrees with $\lambda$ when $\lambda(t^m) \in \mathcal{L}(S^b \times T^b)$, and maps $t^m$ to some arbitrarily chosen point of $\mathcal{L}(S^b \times T^b)$ otherwise. But this new function wouldn’t be continuous, and we need continuity in the proof of Theorem 1.

\(^\text{44}\)The terminology is from Brandenburger (2003). The question of the existence of a space of all types with LPS’s is raised in Brandenburger (1996).

\(^\text{45}\)We have not tried to define and prove the existence of a universal sequential probability space.
worth exploring extensions of our definitions. We already mentioned Halpern (2003), which considers sequences of measures indexed by an initial segment of the ordinals.

Condition (b) of Definition 1 required that the measures in an LPS be disjoint. Blume, Brandenburger, and Dekel (1991a), where LPS’s were introduced, considers both the disjoint and non-disjoint cases. Both types of LPS are actually used here, too. Take a rational strategy-type pair \((s^a, t^a)\) of Ann. The type \(t^a\) is associated with an LPS \(\lambda^a(t^a) = (\mu_0, \ldots, \mu_{n-1})\) on \(S^b \times T^b\), where the measures \(\mu_i\) are disjoint by definition. But, as noted in Remark 1 in Section 7, the optimality of \(s^a\) under \(\lambda^a(t^a)\) really depends on the marginals on \(S^b\) of the \(\mu_i\). And these marginals need not be disjoint, of course.

Finally, the full-support requirement in Definition 1 (condition (c)) is obviously essential to getting the admissibility of rational choices.

iv. Extension to \(n > 2\) Players All of our results readily extend to games with more than two players. As always, though, we could choose to impose an independence condition in the \(n\)-player case. Call a strategy efficient if it is a best reply to some profile of completely mixed strategies for the other players. Efficiency and admissibility are the same in two-player games, of course. But it is well known that in games with three or more players, the efficient strategies can be a strict subset of the admissible strategies. Thus, if we wanted epistemic conditions for iterated efficiency instead of IA, we would have to make a suitable independence assumption in the epistemic set-up. We do not pursue this here.

11 The Literature

We already mentioned in Section 5 the literature on backward induction. Within this literature, this paper owes a particular debt to Stalnaker (1998) and Battigalli and Siniscalchi (2002). These papers took the key step of using large interactive structures—one containing many different types of each player. Battigalli-Siniscalchi, in particular, use a structure that is complete in a sense similar to ours (Definition 6) to give epistemic conditions for extensive-form rationalizability (EFR), a solution concept originally defined by Pearce (1984). EFR gives the backward-induction outcome in perfect-information games and also does forward induction. (Unlike IA, though, EFR is not invariant to the Dalkey-Thompson or Elmes-Reny transformations.)

---

\[\begin{array}{ccc}
L & C & R \\
T & 4, 0 & 4, 1 & 0, 1 \\
M & 0, 0 & 0, 1 & 4, 1 \\
B & 3, 0 & 2, 1 & 2, 1 \\
\end{array}\]

Here, the row \(B\) is admissible (even iteratively admissible). But if \(B\) is to be rational, then the associated LPS must have a primary measure that puts marginal probability of 1/2 on each of \(C\) and \(R\), and a secondary measure that puts positive marginal probability on \(R\). (For example, the secondary measure could put equal weight on \(L\) and \(R\).)

---

The modern, epistemic view is that the dependent (correlated) case is more basic.

Recall our discussion of invariance in Section 5.
Several papers have looked at the problem of giving a foundation for IA. We already mentioned Samuelson (1992). He gives a definition of common knowledge of rationality (where rationality incorporates an admissibility requirement) and shows that there are games where his condition is actually impossible. Indeed, Figure 3 in Section 2 of this paper is one of Samuelson’s examples, and we gave the essence of his argument there. Let us recap it, emphasizing the crucial role that knowledge—as opposed to assumption—plays in Samuelson’s result. Ann must play $T$ since $B$ is inadmissible. In Samuelson’s model, Bob knows this, and so must play $L$. But Ann then knows this and so must be indifferent between $T$ and $B$. But Bob then knows this, so perhaps he will play $R$ after all. But Ann then knows this and so won’t play $B$. We are led around in a circle, as before. The key is that Ann knows—not assumes—that Bob plays $L$. This is why she brings $B$ back in. If she merely assumes that Bob plays $L$, as we would have, then she doesn’t completely rule out his playing $R$, and so she sticks with $T$. There is then no impossibility. Samuelson (op.cit., p.312) writes: “It is the combination of admissibility and common knowledge that yields difficulties. It remains an open question which is the best candidate for deletion from the model.” Our paper gives an answer: keep admissibility, but drop knowledge in favor of assumption. One can then get a result on IA, as we have shown.

Other papers likewise keep admissibility and drop knowledge, but replace it with concepts different from assumption. Börgers (1994) uses “approximate common knowledge” (Monderer and Samet [1989] and Stinchcombe [1988]). Brandenburger (1991) uses lexicographic probabilities as the present paper does, but, unlike our definition of assumption of an event $E$, requires only that the player give $E$ probability 1 under his primary measure. Both Börgers (1994) and Brandenburger (1991) get that the players will choose so-called $S^\infty W$ strategies (Dekel and Fudenberg [1990])—i.e. strategies that survive one round of deletion of inadmissible strategies followed by iterated deletion of strongly dominated strategies. They do not get the IA strategies.

Ewerhart (2001) gets conditions for IA using provability (in the sense of mathematical logic). In his model, Ann assigns probability 0 to a choice of Bob if and only if it is not provable that it is possible that Bob plays that choice. In effect, Ann eliminates a choice of Bob unless it is provable that it should not be eliminated. The philosophy in Ewerhart’s paper seems almost opposite to ours. Ewerhart’s players are “aggressive” (his terminology) in eliminating strategies of the other player. Our players are, in a sense, cautious in eliminating strategies of the other player—since no strategy is, of course, ever entirely ruled out. Ewerhart shows that if the players follow his rule, and if they work in a self-referential system such as Peano Arithmetic, then they will choose IA strategies. Unlike in our paper, completeness does not appear to play a role in Ewerhart.

Gilli (1999) relates weak dominance to strong dominance. He characterizes IA as the set that survives iterated elimination of conditionally (strongly) dominated strategies (Shimoji and Watson [1998]) on the “strategic information sets” of the game. (Gilli also contains a good survey of applications of IA in the literature.)

Asheim (2001) provides epistemic conditions for “proper rationalizability” (Schuhmacher [1999]), a non-equilibrium analog to properness that is a refinement of $S^\infty W$. This yields the backward-induction outcome in perfect-information trees (as proper equilibrium does), but, in general, is different than IA. (It would allow the proper equilibrium $(T,R)$ in the game of Figure 5 from Section 4.) Asheim and Dufwenberg (2003) give epistemic conditions for a solution concept they call “fully permissible sets,” which is again a refinement of $S^\infty W$, and again different from IA.
Finally, Stahl (1995) uses lexicographic probabilities and supposes that Ann views Bob’s strategy $r$ as infinitely less likely than Bob’s strategy $s$ precisely if $r$ is eliminated on an earlier round of the IA procedure than is $s$. Our question was whether we could get the probabilities in the right order without assuming that order. We found that our completeness condition did exactly this.
Appendix A: Admissibility and Invariance

In this section, we give a result in decision theory that motivates the admissibility criterion. Specifically, we show that rationality in a decision tree together with an invariance requirement implies (and is implied by) admissibility on the matrix. The idea for this result comes from an (informal) argument in Kohlberg and Mertens (1986, Section 2.7).

First, we give a formal definition of a decision tree and of rationality in the tree. A decision tree is a finite two-player game in extensive form, where one player is the decision maker (D) and the other player is Nature (N), and we specify payoffs for D only. We allow imperfect information. Player D is assumed to have perfect recall.

Let \( \Omega \) be the set of strategies of N. Note that an information set \( i \) of D can be viewed as a subset of \( \Omega \), consisting of those strategies of N that allow \( i \). A strategy \( \pi \) of D is rational if, for each information set \( i \) of D allowed by \( \pi \), there is a probability measure \( p_i \) on \( \Omega \), with \( p_i(i) = 1 \), such that \( \pi \) maximizes D’s expected payoff under \( p_i \), among all strategies of D that allow \( i \).

Next, consider the (decision) matrix associated with a particular decision tree, and let the rows be D’s strategies and the columns be N’s strategies. Say matrix \( \Lambda \) reduces to matrix \( M \) if it differs from \( M \) by the addition of duplicate rows or columns, or rows that are convex combination of other rows.\(^50\) It is well known that two trees differ by a sequence of the Elmes-Reny (1994) transformations, augmented by the Kohlberg-Mertens (1986, p.1009) convex-combination transformation, if and only if their associated matrices are equal up to reduction.\(^51\)

We gave a definition of rationality for a fixed tree. We now want to consider only those strategies that are also rational in every equivalent tree. This is our decision-theoretic invariance condition. The following result characterizes those strategies.

**Theorem A1** A strategy is admissible in a decision matrix \( M \) if and only if it is rational in every decision tree that gives rise to a matrix that reduces to \( M \).\(^52\)

**Proof.** Fix a matrix \( M \) and an admissible row \( r \). Consider a tree \( T \) that gives rise to a matrix \( \Lambda \) that reduces to \( M \). We first show that \( r \) is admissible in \( \Lambda \). Indeed, since \( r \) is admissible in \( M \), there is a full-support measure \( q \) on the columns of \( M \), under which \( r \) is optimal among all rows in \( M \). But then \( r \) is also optimal under \( q \) among any duplicate rows or convex combinations of rows in \( M \). If \( \Lambda \) contains duplicate columns, we can clearly ‘distribute’ \( q \) over these duplicates.

Next, consider the tree \( T \), and set \( p_i = q(i \mid i) \) for each information set \( i \) of D. This is well defined since \( q(i) > 0 \) for each \( i \). Suppose \( r \) is irrational in \( T \). Then there is an information set \( i \) allowed by \( r \), and another strategy \( s \) that allows \( i \), such that \( s \) yields a higher expected payoff than \( r \) under \( p_i \). Define a strategy \( t \) that coincides with \( s \) at \( i \) and all succeeding information sets of D, and coincides with \( r \) elsewhere. Then \( t \) must yield a higher expected payoff than \( r \) under \( q \), contradicting the admissibility of \( r \) in \( \Lambda \).

For the converse direction, fix a matrix \( M \) and an inadmissible row \( r \). There is then another row (or mixture of rows) \( \sigma \) that weakly dominates \( r \). Let \( C \) denote the set of columns on which \( \sigma \) yields a strictly higher expected payoff than \( r \).

We now construct a tree \( T \) that gives rise to \( M \) (after reduction) and in which \( r \) is not rational. The tree is depicted in Figure A1 below. The initial node belongs to D, who chooses either: (i) one

---

\(^{50}\) Note that we consider convex combinations of rows only (and not columns). We adopt the Anscombe-Aumann (1963) viewpoint, under which the decision maker’s payoffs are really expected payoffs over objective lotteries. It is then natural to say that the decision maker can mix over objective lotteries. Nature does not mix.

\(^{51}\) The Elmes-Reny transformations, unlike the Dalkey (1953)-Thompson (1952) transformations, preserve perfect recall.

\(^{52}\) We are grateful to Pierpaolo Battigalli for pointing out a gap in an earlier version of the proof.
of the rows other than \( r \) or \( \sigma \); or (ii) a move labelled as \( \{ r, \sigma \} \). (In the case that \( \sigma \) is a mixture of rows, condition (i) means just a row different from \( r \).) Next, and without knowing \( D \)'s move, player \( N \) chooses one of the columns of \( M \). This ends the game, except on the paths where \( D \) chooses \( \{ r, \sigma \} \) and \( N \) then chooses a column from \( C \). Here, there is a final information set \( i \) of \( D \) at which he chooses between a move labelled \( r \) and a move labelled \( \sigma \). The payoffs come from \( M \) in the obvious fashion. (In the case that \( \sigma \) is a mixture of rows, a path that ends with \( D \)'s choosing \( \sigma \) gets a payoff equal to the expected payoff that \( \sigma \) yields against the relevant column.) It is clear that the tree just defined gives rise to \( M \) after reduction.

Note that, regardless of the measure \( p_i \) at \( i \), choosing \( \{ r, \sigma \} \) and then \( \sigma \) yields \( D \) a strictly higher expected payoff under \( p_i \) than does choosing \( \{ r, \sigma \} \) and then \( r \). This establishes that \( r \) is irrational, as required. ■
Appendix B: Axioms for Assumption

This section provides a characterization of conditions (a)-(c) in our definition of assumption (Definition 2 in Section 6) in terms of the player’s preferences over acts.

Give Ann two strategies s and t, where the payoffs to s and t differ only in states in some event E. Suppose that Ann takes the view that if she chooses s over t, then she would make the same choice even if the payoffs to the two strategies now differed in states outside E. If this is so, it seems reasonable to use the same terminology as before: Ann could be said to assume that she considers E almost sufficient for deciding what to do. (Only if she ranks s and t indifferently will she have to take account of payoffs outside E.) We now show that this definition is indeed the same as Definition 2 in the text.

Let Ω be a Polish space, as in Section 6, and let A be the space of all measurable functions from Ω to [0,1]. Here, a particular function x ∈ A is an act, where x(ω) is the payoff to the player of choosing the act x, if the true state is ω ∈ Ω. Fix an LPS σ = (µ_0, ..., µ_−1) ∈ L(Ω), and define a preference relation ≥^σ on A by:

\[
x ≥^σ y ⇔ \left( \int_Ω x(ω)dµ_i(ω) \right)_{i=0}^{n-1} ≥ L \left( \int_Ω y(ω)dµ_i(ω) \right)_{i=0}^{n-1}.
\]

Next, given acts x, z ∈ A and a Borel subset E of Ω, write (x_E, z_{Ω\backslash E}) for the element of A given by

\[
(x_E, z_{Ω\backslash E})(ω) = \begin{cases} x(ω) & \text{if } ω ∈ E, \\ z(ω) & \text{if } ω ∈ Ω \backslash E. \end{cases}
\]

Write x ≥^σ_E y (to be read as: x is weakly preferred to y, conditional on E) if there is a z ∈ A such that (x_E, z_{Ω\backslash E}) ≥^σ (y_E, z_{Ω\backslash E}). (It is readily checked that the definition does not depend on the choice of z.)

**Definition B1** A set E is assumed under ≥^σ if:

(i) there are x, y ∈ A such that x ≥^σ_E y,

(ii) for all x, y ∈ A, we have x ≥^σ_E y implies x ≥^σ y.

Condition (ii) says that in the comparison between any two acts x and y, the payoffs in E are determining for strict preference. Condition (i) is a non-triviality requirement, stating that E is not Savage-null. (Recall that E would be Savage-null if x ∼^σ_E y for all x, y.)

**Proposition B1** A set E is assumed under ≥^σ if and only if conditions (a) and (b) of Definition 2 hold, that is, there is a j < n such that:

(a) µ_i(E) = 1 for all i ≤ j,

(b) µ_i(E) = 0 for all i < j < n.

**Proof.** Suppose first that there is a j < n such that conditions (a) and (b) are satisfied. Set x(ω) = 1 and y(ω) = z(ω) = 0 for every ω. The act (x_E, z_{Ω\backslash E}) is evaluated as (1, ..., 1, 0, ..., 0) where the last 1 corresponds to µ_j, and (y_E, z_{Ω\backslash E}) is evaluated as (0, ..., 0). Thus x ≥^σ_E y, establishing condition (i) above.
Next, note that \( x \triangleright_{E} y \) implies
\[
\left( \int_{E} x d\mu_{0}, \ldots, \int_{E} x d\mu_{j}, \int_{\Omega \setminus E} x d\mu_{j+1}, \ldots, \int_{\Omega \setminus E} x d\mu_{n-1} \right) > L
\]
\[
= \left( \int_{E} y d\mu_{0}, \ldots, \int_{E} y d\mu_{j}, \int_{\Omega \setminus E} y d\mu_{j+1}, \ldots, \int_{\Omega \setminus E} y d\mu_{n-1} \right).
\]
Therefore
\[
\left( \int_{E} x d\mu_{0}, \ldots, \int_{E} x d\mu_{j}, \int_{\Omega \setminus E} x d\mu_{j+1}, \ldots, \int_{\Omega \setminus E} x d\mu_{n-1} \right) > L
\]
\[
= \left( \int_{E} y d\mu_{0}, \ldots, \int_{E} y d\mu_{j}, \int_{\Omega \setminus E} y d\mu_{j+1}, \ldots, \int_{\Omega \setminus E} y d\mu_{n-1} \right).
\]
Thus \( x \triangleright_{\sigma} y \), establishing condition (ii) above.

Next, suppose that either of conditions (a) and (b) (or both) fails to hold. There are three cases
to consider. Case 1: \( \mu_{i}(E) = 0 \) for all \( i \). Case 2: \( \mu_{i}(E) = 0 \) and \( \mu_{h}(E) = 1 \) where \( h > i \). Case 3: \( 0 < \mu_{i}(E) < 1 \) for some \( i \). We take each case in turn.

Case 1: Here, condition (i) above is clearly violated.

Case 2: Let
\[
x(\omega) = \begin{cases} 
1 & \text{if } \omega \in E \cap U_{h}, \\
0 & \text{otherwise}; 
\end{cases}
\]
\[
y(\omega) = \begin{cases} 
1 & \text{if } \omega \in U_{i} \setminus E, \\
0 & \text{otherwise}; 
\end{cases}
\]
\[
z(\omega) = 0 \quad \text{for every } \omega.
\]
The act \((x_{E}, z_{\Omega \setminus E})\) is evaluated as \((0, \ldots, 0, 1, 0, \ldots, 0)\) where the 1 corresponds to \( \mu_{h} \). (Here, we use \( \mu_{k}(U_{h}) = 0 \) for all \( k \neq h \).) The act \((y_{E}, z_{\Omega \setminus E})\) is evaluated as \((0, \ldots, 0)\), and so \( x \triangleright_{E} y \). But \( x \) is evaluated the same as \((x_{E}, z_{\Omega \setminus E})\), while \( y \) is evaluated as \((0, \ldots, 0, 1, 0, \ldots, 0)\) where the 1 corresponds to \( \mu_{i} \). Since \( h > i \), this establishes that \( x \prec_{\sigma} y \), violating condition (ii) above.

Case 3: Let
\[
x(\omega) = \begin{cases} 
\mu_{i}(U_{i} \setminus E) & \text{if } \omega \in E \cap U_{i}, \\
0 & \text{otherwise}; 
\end{cases}
\]
\[
y(\omega) = \begin{cases} 
1 & \text{if } \omega \in U_{i} \setminus E, \\
0 & \text{otherwise}; 
\end{cases}
\]
\[
z(\omega) = 0 \quad \text{for every } \omega.
\]
The act \((x_{E}, z_{\Omega \setminus E})\) is evaluated as
\[
(0, \ldots, 0, \mu_{i}(U_{i} \setminus E), \mu_{i}(E \cap U_{i}), 0, \ldots, 0),
\]
where the non-zero entry corresponds to \( \mu_{i} \). This entry is indeed non-zero, since the assumption
that \( 0 < \mu_{i}(E) < 1 \) implies that \( \mu_{i}(U_{i} \setminus E) > 0 \) and \( \mu_{i}(E \cap U_{i}) > 0 \). The act \((y_{E}, z_{\Omega \setminus E})\) is evaluated as \((0, \ldots, 0)\), and so \( x \triangleright_{E} y \). The act \( x \) is evaluated the same as \((x_{E}, z_{\Omega \setminus E})\), while \( y \) is evaluated as
\[
(0, \ldots, 0, \mu_{i}(U_{i} \setminus E), 0, \ldots, 0),
\]
26
where the non-zero entry corresponds to \( \mu_i \). Note that this entry is indeed non-zero. Now from
\[ \mu_i(E \cap U_i) < 1 \]
it follows that \( x \prec y \), violating condition (ii) above. 

We now relate the remaining condition (c) of Definition 2 to a preference-based notion, which says that ‘all of \( E \) should matter.’

**Definition B2** A set \( E \) is **whole under** \( \succeq^\sigma \) if for each \( e \in E \) and open neighborhood \( U \) of \( e \), there are \( x, y \in A \) such that \( x \succ^\sigma_{E \cap U} y \).

**Proposition B2** Suppose that \( E \) is assumed under \( \succeq^\sigma \), so that there is a \( j < n \) such that:

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

Then \( E \) is whole under \( \succeq^\sigma \) if and only if condition (c) of Definition 2 holds, that is:

(c) \( E \subseteq \bigcup_{i \leq j} \text{Supp} \mu_i \).

**Proof.** Suppose first that condition (c) holds, and consider some \( e \in E \) and an open neighborhood \( U \) of \( e \). Set

\[
\begin{align*}
x(\omega) &= \begin{cases} 1 & \text{if } \omega \in E \cap U, \\ 0 & \text{otherwise}; \end{cases} \\
y(\omega) &= z(\omega) = 0 \text{ for every } \omega.
\end{align*}
\]

Choose an \( i \leq j \) such that \( U \cap \text{Supp} \mu_i \neq \emptyset \). Then \( \mu_i(U) > 0 \), from which \( \mu_i(E \cap U) > 0 \). Thus \( x \succ^\sigma_{E \cap U} y \), as required.

Next, suppose that condition (c) fails. Then there is an \( e \in E \) such that \( U = \Omega \setminus \bigcup_{i \leq j} \text{Supp} \mu_i \) is an open neighborhood of \( e \). But \( \mu_i(E \cap U) = 0 \) for all \( i \). (If \( i \leq j \) then \( \mu_i(U) = 0 \); if \( i > j \) then \( \mu_i(E) = 0 \).) 

27
Appendix C: Proofs for Sections 6 and 7

**Proof of Lemma 1.** We give a proof that works even in the case \( l(\sigma) = \infty \). Let \( U_i \), for \( i < l(\sigma) \), satisfy (b). For \( 0 < i \) let \( V_i = U_i \setminus \bigcup_{j<i} U_j \). Now let \( V_0 = \Omega \setminus \bigcup_{i>0} V_i \). Then \( V_i \), for \( i < l(\sigma) \), partitions \( \Omega \) and satisfies (b).

Next, let \( W = \bigcup_j (V_j \setminus \text{Supp} \mu_j) \). We have \( \mu_i(W) = 0 \) for all \( i \). Let \( W_0 = W \cap \text{Supp} \mu_0 \) and

\[
W_i = [W \setminus (W_0 \cup \cdots \cup W_{i-1})] \cap \text{Supp} \mu_i
\]

for \( i > 0 \). Using \( \bigcup_j \text{Supp} \mu_j = \Omega \), we see that the \( W_i \) partition \( W \). Also \( \mu_i(W_j) = 0 \) for all \( i \) and \( j \), since \( W_j \subseteq W \). Finally, set \( X_i = (V_i \cap \text{Supp} \mu_i) \cap W_i \). Then the \( X_i \) partition \( \Omega \) and have the required properties: \( \mu_i(X_i) = 1 \), \( \mu_i(X_j) = 0 \) for \( j \neq i \), and \( X_i \subseteq \text{Supp} \mu_i \). ■

**Proof of Lemma 2.** Suppose (c) fails, so that \( U = \Omega \setminus \bigcup_{i<l(\sigma)} \text{Supp} \mu_i \) is nonempty. Then \( U \) is open (using \( l(\sigma) < \infty \)) and \( \mu_i(U) = 0 \) for all \( i \), and hence (\( e' \)) fails.

For the converse, suppose (c) holds. If (\( e' \)) fails, there is a nonempty open \( U \) with \( \mu_i(U) = 0 \) for all \( i \). By (c), \( U \cap \text{Supp} \mu_i \neq \emptyset \) for some \( i \). Then \( (\Omega \setminus U) \cap \text{Supp} \mu_i \) is closed and strictly contained in \( \text{Supp} \mu_i \), so that \( \mu_i((\Omega \setminus U) \cap \text{Supp} \mu_i) < 1 \), from which \( \mu_i(U) > 0 \), a contradiction. Therefore (\( e' \)) holds as required. ■

**Proof of Lemma 3.** In this and the next proof, Borel without qualification means Borel in \( (\mathcal{M}'(\Omega))^N \). We make repeated use of the following facts:

1. There is a countable open basis \( E_1, E_2, \ldots \) for \( \Omega \).
2. For each Borel set \( B \) in \( \Omega \) and \( r \in [0,1] \), the set of \( \mu \) such that \( \mu(B) > r \) is Borel in \( \mathcal{M}'(\Omega) \).
3. For each Borel set \( Y \) in \( \mathcal{M}'(\Omega) \) and each \( k \), the set of \( \sigma \) such that \( \mu_k \in Y \) is Borel.

It is enough to show that for each \( n \in N \), the set \( \mathcal{L}^n(\Omega) \) of all \( \sigma \in \mathcal{L}(\Omega) \) of length \( n \) is Borel, because \( \mathcal{L} = \bigcup_n \mathcal{L}^n \). Fix \( n \). It follows from (3) that \( \mathcal{M}^\infty(\Omega) \) is Borel, and also the set \( X^n_1 \) of \( \sigma \in \mathcal{M}^\infty(\Omega) \) of length \( n \) is Borel. Using Lemma 2 and (1)-(3), the set \( X^n_1 \) of \( \sigma \in X^n_1 \) such that \( \sigma \) has full support is Borel. The set \( \mathcal{L}^n(\Omega) \) is equal to the set of all \( \sigma \in X^n_2 \) satisfying condition (b) of Definition 1. Let us write \( \mu \perp \nu \) if there is a Borel set \( U \in \Omega \) such that \( \mu(U) = 1 \) and \( \nu(U) = 0 \). It is easy to see that (b) holds for an element \( \sigma \in X^n_2 \) if and only if \( \mu_i \perp \mu_j \) for all \( i < j < l(\sigma) \). To complete the proof it suffices to prove that for each \( i < j \), the set of \( \sigma \) such that \( \mu_i \perp \mu_j \) is Borel. Note that \( \mu_i \perp \mu_j \) if and only if for each \( m \), there is an open set \( V \) such that \( \mu_i(V) = 1 \) and \( \mu_j(V) < \frac{1}{m} \). By (1), this in turn holds if and only if for each \( m \) there exists \( k \) such that \( \mu_i(E_k) > 1 - \frac{1}{m} \) and \( \mu_j(E_k) < \frac{1}{m} \). By (2) and (3), the set of \( \sigma \) such that \( \mu_i(E_k) > 1 - \frac{1}{m} \) is Borel, and the set of \( \sigma \) such that \( \mu_j(E_k) < \frac{1}{m} \) is Borel. The set of \( \sigma \) such that \( \mu_i \perp \mu_j \) is a countable Boolean combination of these sets, and hence is Borel as required. ■

**Proof of Lemma 4.** Part (a) is immediate.

For part (b), first note that since \( \lambda^a \) is continuous, Lemma 3 implies that for each \( n \) the set \( \lambda^a(\mathcal{L}^n(S^b \times T^b)) \) is Borel in \( T^a \). From Remark 1, for each \( s^a \in S^a \) there is a finite Boolean combination \( C \) of linear equations in \( n \cdot | S^b | | S^b | | T^b | \) variables such that whenever \( \lambda^a(t^a) = (\mu_0, \ldots, \mu_{n-1}) \in \mathcal{L}^n(S^b \times T^b) \), the pair \( (s^a, t^a) \) is rational if and only if \( C \) holds for \( \{ \text{margin of } \mu_i(s^b) : i < n, s^b \in S^b \} \). Since \( S^a \) and \( S^b \) are finite, this shows that \( R^a_{ik} \) is Borel in \( S^a \times T^a \).

To show that \( R^a_{ik} \) is Borel in \( S^a \times T^a \) for each \( m > 1 \), it now suffices to prove that for each Borel set \( E \) in \( S^b \times T^b \), the set \( A^a(E) \) is Borel in \( T^a \). For this it suffices to show that for each Polish space
\( \Omega \) and Borel set \( E \) in \( \Omega \), the set of \( \sigma \) such that \( E \) is assumed under \( \sigma \) is Borel. Fix \( n \) and \( j < n \). The sets of \( \mu \) such that \( \mu(E) = 1 \) and such that \( \mu(E) = 0 \) are Borel in \( \mathcal{M}'(\Omega) \). Therefore the set of \( \sigma = (\mu_0, \ldots, \mu_{n-1}) \in \mathcal{L}^n(\Omega) \) such that conditions (a) and (b) in Definition 2 hold is Borel. Let \( \{d_0, d_1, \ldots\} \) be a countable dense subset of \( E \). For each \( k \) and \( \mu \in \mathcal{M}(\Omega) \), we have \( d_k \in \text{Supp} \mu \) if and only if \( \mu(B) > 0 \) for every open ball \( B \) with center \( d_k \) and rational radius. Therefore the set of \( \mu \) such that \( d_k \in \text{Supp} \mu \) is Borel in \( \mathcal{M}(\Omega) \). We have \( E \subseteq \bigcup_{i \geq j} \text{Supp} \mu_i \) if and only if \( d_k \in \bigcup_{i \geq j} \text{Supp} \mu_i \) for all \( k \in \mathbb{N} \). Therefore, the set of \( \sigma \in \mathcal{L}^n(\Omega) \) satisfying condition (c) in Definition 2 is Borel. Thus the set of \( \sigma \in \mathcal{L}(\Omega) \) such that \( E \) is assumed under \( \mu \) is Borel, as required.

**Proof of Proposition 1.** Let \( \mathcal{N} \) be the Baire space, i.e. the metric space \( \mathbb{N}^\mathbb{N} \) with the product metric, where \( \mathbb{N} \) has the discrete metric. Then \( \mathcal{N} \) is Polish and there is a continuous surjection from \( \mathcal{N} \) to any nonempty Polish space (Kechris [1995, p.13 and Theorem 7.9]). Set \( T^a = T^b = \mathcal{N} \). The space \( \mathcal{M}^\infty(S^b \times T^b) \) is a Polish space, so there is a continuous surjection from \( T^a \) to \( \mathcal{M}^\infty(S^b \times T^b) \), and similarly with \( a \) and \( b \) reversed. ■
Appendix D: Proof of Theorem 1

This section proves our main result—Theorem 1 in Section 8 of the text. We begin by giving the idea of the proof. The formal proof follows.

First, the idea of the proof. Start with a rational pair \((s^a, t^a)\). (Formally: \((s^a, t^a) \in R_a^1\).) Writing \(\lambda^a(t^a) = (\mu_0, \ldots, \mu_{n-1})\) and using Remark 1 in Section 7, 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. This establishes that \(s^a\) is admissible. (Formally: \(s^a \in S^a_1\).

Next, fix any admissible strategy \(r^a\). There is then a full-support measure \(\nu_0\) on \(S^b\) under which \(r^a\) is optimal. Construct a full-support measure \(\mu_0\) on \(S^b \times T^b\) with \(\text{marg}_{S^b} \mu_0 = \nu_0\). Using completeness, we can find a type \(u^a\) such that \(\lambda^a(u^a)\) is the one-level LPS \((\mu_0)\). The pair \((r^a, u^a)\) is rational by construction. So, we have now shown that \(\text{proj}_{S^a} R^b_1 = S^a_1\). Of course, a parallel argument establishes that \(\text{proj}_{S^b} R^a_1 = S^b_1\).

Next, take a rational pair \((s^a, t^a)\) where \(t^a\) assumes \(b\) is rational. (Formally: \((s^a, t^a) \in R^b_2\).) Again writing \(\lambda^a(t^a) = (\mu_0, \ldots, \mu_{n-1})\), and using Definition 2, there is then a \(j < n\) such that \(\mu_i(R^b_1) = 1\) for all \(i \leq j\) and \(\mu_i(R^b_1) = 0\) for all \(i > j\). It follows that

\[\bigcup_{i \leq j} \text{Supp}(\text{marg}_{S^b} \mu_i) = \text{proj}_{S^b} R^b_1.\]

(Condition \((c)\) of Definition 2 is also used here.) This step is illustrated in Figure D1. Using \(\text{proj}_{S^b} R^b_1 = S^b_1\), we can then 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 \(S^b_1\), under which \(s^a\) is optimal. This establishes that \(s^a \in S^a_2\). That is, the strategy \(s^a\) is 2-admissible.

![Figure D1](image)

Next, fix any 2-admissible strategy \(q^a\). Thus, there is a measure \(\nu_0\) with support \(S^b_1\) under which \(q^a\) is optimal among strategies in \(S^a_1\). We can show that \(q^a\) is, in fact, optimal among all strategies in \(S^a\). Also, since \(q^a\) is 1-admissible, there is a full-support measure \(\nu_1\) on \(S^b\) under which \(q^a\) is optimal among all strategies in \(S^a\). Certainly then, the strategy \(q^a\) lexicographically maximizes expected payoff (among all strategies in \(S^a\)) under the sequence of measures \((\nu_0, \nu_1)\). We now construct
measures $\mu_0$ and $\mu_1$ on $S^b \times T^b$ such that

\begin{align*}
\mu_0(R_1^b) &= 1, \\
\mu_1((S^b \times T^b) \setminus R_1^b) &= 1, \\
R_1^b &\subseteq \text{Supp } \mu_0, \\
S^b \times T^b &= \text{Supp } \mu_0 \cup \text{Supp } \mu_1, \\
\text{marg}_{S^b} \mu_0 &= \nu_0, \\
\text{marg}_{S^b} \mu_1 &= \nu_1.
\end{align*}

This step is illustrated in Figure D2 (and again uses $\text{proj}_{S^b} R_1^b = S_1^b$). Using completeness, we can find a type $v^a$ such that $\lambda^a(v^a)$ is the two-level LPS $(\mu_0, \mu_1)$. By construction, the pair $(q^a, v^a)$ is rational and $v^a$ assumes $R_1^b$. We have now shown that $\text{proj}_{S^a} R_2^a = S_2^a$ (and likewise with $a$ and $b$ interchanged). The formal proof proceeds in this way to establish that $\text{proj}_{S^a} R_m^a = S_m^a$ for all $m$, which establishes Theorem 1.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure_d2.png}
\caption{Figure D2}
\end{figure}

We now give the formal proof. The first result is standard.

**Lemma D1** A strategy $s^a$ is $m$-admissible if and only if there is a $\mu \in \mathcal{M}(S^b)$, with $\text{Supp } \mu = S_{m-1}^b$, such that for every $r^a \in S_{m-1}^a$,

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

(D1)

**Lemma D2** If a strategy $s^a$ is $m$-admissible, then there is a $\mu \in \mathcal{M}(S^b)$, with $\text{Supp } \mu = S_{m-1}^b$, such that (D1) holds for every $r^a \in S^a$.

**Proof.** Fix a $\mu$ such that (D1) holds for every $r^a \in S_{m-1}^a$. Suppose (D1) fails for some $r^a \in S_{m-1}^a \setminus S_{m-1}^a$. We have $r^a \in S_{l+1}^a \setminus S_{l+1}^a$ for some $l < m - 1$. Pick some $\nu \in \mathcal{M}(S^b)$, with $\text{Supp } \nu = S_{l}^b$, and define a sequence of measures $\mu^n \in \mathcal{M}(S^b)$, for each $n \in \mathbb{N}$, by $\mu^n = (1 - \frac{1}{n}) \mu + \frac{1}{n} \nu$. Note
that $\mathrm{Supp} \mu^n = S^b_1$ for each $n$. Using $r^a \not\in S^a_{l+1}$, and Lemma D1 applied to the $(l+1)$-admissible strategies, it follows that for each $n$ there is a $q^n \in S^a_l$ s.t.
\[
\sum_{s^b \in S^b} \pi^n(q^a, s^b)\mu^n(s^b) > \sum_{s^b \in S^b} \pi^n(r^a, s^b)\mu^n(s^b).
\] (D2)

We can assume that $q^a \in S^a_{l+1}$. (Choose $q^a \in S^a_l$ to maximize the left-hand side of (D2) among all strategies in $S^a_l$.) Also, since $S^a_{l+1}$ is finite, there is a $q^a \in S^a_{l+1}$ s.t. (D2) holds for infinitely many $n$. Letting $n \to \infty$ yields
\[
\sum_{s^b \in S^b} \pi^n(q^a, s^b)\mu(s^b) \geq \sum_{s^b \in S^b} \pi^n(s^a, s^b)\mu(s^b).
\] (D3)

From (D3) and the negation of (D1) for $r^a$ we get
\[
\sum_{s^b \in S^b} \pi^n(q^a, s^b)\mu(s^b) > \sum_{s^b \in S^b} \pi^n(s^a, s^b)\mu(s^b).
\]
This is a contradiction if $q^a \in S^a_{m-1}$. If not, we have $q^a \in S^a_k \setminus S^a_{k+1}$ for some $l < k < m - 1$. Now repeat with $q^a$ the argument made with $r^a$. This way, a contradiction must arise in at most $(m - 1 - l)$ steps. ■

The next lemma will guarantee that we will always have enough room to build the measures we need in the subsequent steps in the proof. For $t^a, u^a \in T^a$, let us write $t^a \approx u^a$ if for each $i$ the component measures $(\lambda^a(t^a))_i$ and $(\lambda^a(u^a))_i$ have the same marginals on $S^b$ and are mutually absolutely continuous (have the same null sets).

Lemma D3

(i) If $\lambda^a(t^a)$ has full support, then there are continuum many $u^a$ such that $u^a \approx t^a$.

(ii) For each set $E \subseteq S^b \times T^b$, the set $A^a(E)$ is closed under the relation $\approx$.

(iii) If $t^a \approx u^a$ then for each $m$ and $s^a \in S^a$, we have $(s^a, t^a) \in R^a_m$ if and only if $(s^a, u^a) \in R^a_m$.

Proof. (i) Full support implies that $\mu_i = (\lambda^a(t^a))_i$ has infinite support for some $i$, which implies that there are continuum many different measures $\nu_i$ which have the same null sets and marginal on $S^b$ as $\mu_i$. By completeness, the sequence of measures obtained by replacing $\mu_i$ by $\nu_i$ is equal to $\lambda^a(u^a)$ for some $u^a$, and $u^a \approx t^a$. Part (ii) and part (iii) for $m = 1$ are immediate, and (iii) for $m > 1$ is then proved by an easy induction. ■

It will simplify the presentation from now on to set $R^a_0 = S^a \times T^a$.

Lemma D4 $\proj_{S^a} R^a_m = \proj_{S^a}(R^a_m \setminus R^a_{m+1})$ for each $m \geq 0$.

Proof. We argue by induction on $m$. The result is true $m = 0$, as can be seen by choosing $t^a$ such that $\lambda^a(t^a) \not\in \mathcal{L}(S^b \times T^b)$ and then noting that $S^a \times \{t^a\}$ is disjoint from $R^a_0$. Let $m \geq 1$ and assume the result for $m - 1$. Let $(s^a, t^a) \in R^a_m$ and $\lambda^a(t^a) = \sigma = (\mu_0, \ldots, \mu_{m-1})$. Then $t^a \in A^a(R^a_k)$ for each $k < m$. We must find a $u^a$ such that $(s^a, u^a) \in R^a_m \setminus R^a_{m+1}$.

By the inductive hypothesis and the fact that $S^b$ is finite, there is a finite set $U \subseteq R^b_{m-1} \setminus R^b_m$ such that $\proj_{S^b} U = \proj_{S^b} R^b_{m-1}$. When $m > 1$ we have $U \subseteq R^b_1$, so $U$ has the property that

32
\(\lambda^b(t^b)\) has full support for each \((s^b, t^b) \in U\). We can also take \(U\) to have this property when \(m = 1\).

Therefore by Lemma D3, the set \(U\) can be chosen so that \(\mu_i(U) = 0\) for all \(i\).

We will get a point \((s^b, u^b) \in R^a_m \setminus R^a_{m+1}\) by adding a measure to the beginning of the finite sequence \(r\). Since \(U\) is finite, \(\text{proj}_{S^b} U = \text{proj}_{S^b} R^a_{m-1}\), and \(\mu_0(R^a_{m-1}) = 1\), there is a probability measure \(\nu\) such that \(\nu(U) = 1\) and \(\text{marg}_{S^b} \nu = \text{marg}_{S^b} \mu_0\). Let \(\tau\) be the finite sequence \((\nu, \mu_0, \ldots, \mu_{m-1})\). By completeness there exists \(u^a \in T^a\) such that \(\lambda^a(u^a) = \tau\). Since \(r \in \mathcal{L}(S^b \times T^b)\) and \(\mu_1(U) = 0\) for each \(i\), we see that \(\tau \in \mathcal{L}(S^b \times T^b)\). Since \(\nu\) has the same marginal on \(S^b\) as \(\mu_0\) and \((s^a, u^a) \in R^a_i\), we have \((s^a, u^a) \in R^a_i\). Since \(U \subseteq R^a_{m-1}\) and \(u^a \in A^a(R^a_k)\) for each \(k < m\), it follows that \(u^a \in A^a(R^a_k)\) for each \(k < m\). Then by Lemma 4 in the text we have \((s^a, u^a) \in R^a_m\). However, since \(U\) is disjoint from \(R^b_m\) we have \(\nu(R^b_m) = 0\), so \(u^a \notin A^a(R^b_m)\) and hence \((s^a, u^a) \notin R^a_{m+1}\). This completes the induction.

\[\text{Lemma D4}}\]

Lemma D4 shows that the sets \(R^a_{m}\) keep shrinking forever.

It will be useful to single out a subset of \(R^a_{m}\) whose members have a particularly simple structure.

\[\text{Definition D1} \]

For \(m > 0\), let \(\hat{R}^a_m\) be the set of all \((s^a, t^a) \in R^a_m\) such that \(\lambda^a(t^a)\) has length \(m\) and for each \(j < m\), the set \(\hat{R}^b_j\) is assumed under \(\lambda^a(t^a)\) at level \(m - 1 - j\). We also put \(\hat{R}^a_0 = R^a_0\).

Note that \(\hat{R}^a_m \subseteq R^a_m\) for each \(m\), and that the family of sets \(\hat{R}^a_m\), for \(m > 0\), is pairwise disjoint. The next lemma establishes Theorem 1.

\[\text{Lemma D5} \]

\(\text{proj}_{S^b} R^a_m = \text{proj}_{S^b} \hat{R}^a_m = S^a_m\) for all \(m\).

\[\text{Proof.}\]

We argue by induction on \(k\). The result is trivial for \(k = 0\). Assume the result for all \(k \leq m\). We give the proof for \(k = m + 1\). Let \((s^a, t^a) \in R^a_{m+1}\). We show that \(s^a \in S^a_{m+1}\). Write \(\lambda^a(t^a) = (\mu_0, \ldots, \mu_{m-1})\).

We need a separate argument for the case that \(k = 1\), so that \((s^a, t^a) \in R^a_1\). Since \(\lambda^a(t^a)\) has full support, for each \(s^b\) there exists \(i\) such that \(\text{marg}_{S^b} \mu_i(s^b) > 0\). It then follows from Proposition 1 in Blume, Brandenburger, and Dekel (1991b) that we can form a nested convex combination of the measures \(\text{marg}_{S^b} \mu_i\), for \(0 \leq i < n\), to get a measure \(\mu \in \mathcal{M}(S^b)\), with \(\text{Supp} \mu = S^b\), such that equation (D1) holds for every \(r^a \in S^a\). This shows that \(s^a \in S^a_m\).

We now take up the case that \(k > 1\). In that case \((s^a, t^a) \in R^a_m\) and \(t^a \in A^a(R^a_m)\) at some level \(j < n\). By the inductive hypothesis we have \(s^a \in S^a_m\). Hence it will be enough to show that we can form a nested convex combination of the measures \(\text{marg}_{S^b} \mu_i\), for \(0 \leq i \leq j\), to get a measure \(\mu \in \mathcal{M}(S^b)\), with \(\text{Supp} \mu = S^a_m\), such that equation (D1) holds for every \(r^a \in S^a\) (a fortiori, for every \(r^a \in S^a_m\)). This will be possible if

\[S^b_m = \bigcup_{i \leq j} \text{Supp}(\text{marg}_{S^b} \mu_i). \quad \text{(D4)}\]

Let \(s^b \in S^b_m\). By the inductive hypothesis we have \((s^b, t^b) \in R^b_m\) for some \(t^b\). Then \((s^b, t^b) \in \text{Supp} \mu_i\) for some \(i \leq j\). Therefore \(0 < \mu_i(s^b) = \mu_i(t^b) = \text{marg}_{S^b} \mu_i(s^b)\), and hence \(s^b \in \text{Supp}(\text{marg}_{S^b} \mu_i)\).

Thus the left side of (D4) is contained in the right side.

Now suppose \(s^b \notin S^b_m\). By the inductive hypothesis, the set \(\{s^b\} \times T^b\) is disjoint from \(R^b_m\). But for each \(i \leq j\) we have \(\mu_i(R^b_m) = 1\), so \(\mu_i(\{s^b\} \times T^b) = \text{marg}_{S^b} \mu_i(s^b) = 0\) and hence \(s^b \notin \text{Supp}(\text{marg}_{S^b} \mu_i)\). This proves (D4).
We have shown that \( \text{proj}_{S^a} R^a_{m+1} \subseteq S^a_{m+1} \). Since \( \hat{R}^a_m \subseteq R^a_m \), our remaining task is to let \( s^a \in S^a_{m+1} \) and find a \( t^a \) such that \( (s^a, t^a) \in \hat{R}^a_{m+1} \). We use Lemma D2 to infer that for each \( k \leq m \), there is a \( \mu_k \in \mathcal{M}(S^b) \), with \( \text{Supp} \mu_k = S^b_k \), such that for every \( r^a \in S^a \),

\[
\sum_{s^b \in S^b} \pi^a(s^a, s^b) \mu_k(s^b) \geq \sum_{s^b \in S^b} \pi^a(t^a, s^b) \mu_k(s^b).
\]

Certainly, then, the strategy \( s^a \) lexicographically maximizes expected payoff (among all strategies in \( S^a \)) under the sequence of measures \((\mu_0, \ldots, \mu_m)\). By completeness, it suffices to find a sequence \( \sigma = (\nu_0, \ldots, \nu_m) \in \mathcal{L}(S^b \times T^b) \) such that for each \( k \leq m \):

(i) \( \text{marg}_{S^b} \nu_k = \mu_k \);

(ii) \( R^b_k \) is assumed under \( \sigma \) at level \( m - k \).

Let \( s^b \in S^b \) and let \( h \) be the greatest \( k \leq m \) such that \( s^b \in S^b_k \). Let \( X = \{s^b\} \times T^b \). For each \( k \leq h \) we have \( s^b \in S^b_k = \text{Supp} \mu_k \) and thus \( \mu_k(s^b) > 0 \). By rescaling and combining the measures over different \( s^b \), we see that it is enough to find \((\xi_0, \ldots, \xi_h) \in \mathcal{L}(X) \) such that:

(iii) \( \xi_0(X \cap R^b_k) = 1 \);

(iv) \( \xi_k(X \cap (R^b_{h-k} \setminus R^b_{h-k+1})) = 1 \) for each \( 1 \leq k \leq h \);

(v) \( X \cap R^b_{h-j} \subseteq \bigcup_{k=0}^j \text{Supp} \xi_k \) for each \( 0 \leq j \leq h \).

By the inductive hypothesis and Lemma D4, for each \( 1 \leq k \leq h \) we have \( s^b \in \text{proj}_{S^b} R^b_{h-k} = \text{proj}_{S^b} (R^b_{h-k} \setminus R^b_{h-k+1}) \). Therefore for each \( 1 \leq k \leq h \) the set \( X_k = X \cap (R^b_{h-k} \setminus R^b_{h-k+1}) \) is nonempty, and the set \( X_0 = X \cap R^b_h \) is also nonempty. By separability, each of the sets \( X_k \) has a countable dense subset \( Y_k \). By assigning a positive weight to each point in \( Y_k \) we can build a probability measure \( \xi_k \) on \( X \) such that \( \xi_k(Y_k) = 1 \) and \( \text{Supp} \xi_k \) is the closure of \( Y_k \). Then the sequence \((\xi_0, \ldots, \xi_h) \) belongs to \( \mathcal{L}(X) \) and has the required properties (iii)-(v). This completes the induction. \( \blacksquare \)

**Definition D2** If \((s^a, t^a, s^b, t^b) \in \hat{R}^a_{m+1} \times \hat{R}^b_{m+1} \), say there is simple rationality and \( m \)-th order assumption of rationality at this state.

The following is immediate from Lemma D5.

**Corollary D1** Fix a complete \((S^a, S^b)\)-based sequential probability structure

\[
(S^a, S^b, T^a, T^b, \lambda^a, \lambda^b)
\]

and payoff functions \( \pi^a \) and \( \pi^b \). If \( s^a \) and \( s^b \) are \((m+1)\)-admissible strategies, then there is a state \((s^a, t^a, s^b, t^b) \) at which there is simple rationality and \( m \)-th order assumption of rationality.
Appendix E: Proof of Theorem 2

Lemma E1  Suppose \( E \subseteq F \subseteq \Omega, \sigma \in \mathcal{L}(\Omega), \) and both \( E \) and \( F \) are assumed under \( \sigma \) at the same level \( j \). Then \( F \subseteq \text{cl}(E) \) (where \( \text{cl}(\cdot) \) denotes closure).

**Proof.** Let \( X = \bigcup_{i \leq j} \text{Supp} \mu_i \). Then \( F \subseteq X \). But since \( \mu_i(E) = 1 \) for each \( i \leq j \), we have \( X \subseteq \text{cl}(E) \). \( \blacksquare \)

Lemma E2

(i) If player \( a \) is not indifferent, then \( R_0^a \setminus \text{cl}(R_1^a) \) is uncountable.

(ii) If neither player is indifferent, then both \( R_0^a \setminus \text{cl}(R_1^a) \) and \( R_0^b \setminus \text{cl}(R_1^b) \) are uncountable.

**Proof.** (i) We have that \( \pi^a(r^a, s^b) < \pi^a(s^a, s^b) \) for some \( r^a, s^a, s^b \). Let \( t^a \) be such that \( \lambda^a(t^a) = (\mu_0, \mu_1) \) has length 2 and full support, and \( \mu_0(\{s^b\} \times T^b) = 1 \). Let \( U \) be the set of all \( u^a \) such that for some \( s^a \in S^a \),

\[
\sum_{s^b \in S^b} \pi^a(r^a, s^b) \text{mark}_{S^b}(\lambda^a(u^a))_0(s^b) < \sum_{s^b \in S^b} \pi^a(s^a, s^b) \text{mark}_{S^b}(\lambda^a(u^a))_0(s^b).
\]

Then \( \{r^a\} \times U \) is an open set with no rational points. For any \( u^a \approx t^a \), we have \( u^a \in U \), and hence \( (r^a, u^a) \notin \text{cl}(R_1^a) \). By Lemma D3, there are uncountably many \( u^a \) such that \( u^a \approx t^a \), so \( R_0^a \setminus \text{cl}(R_1^a) \) is uncountable. (ii) is immediate from (i). \( \blacksquare \)

Lemma E3  Suppose that \( m > 0 \) and \( R_{m-1}^b \setminus \text{cl}(R_m^b) \) is uncountable. Then \( R_m^b \setminus \text{cl}(R_{m+1}^b) \) is uncountable.

**Proof.** The proof is similar to the proof of Lemma D4. By Lemma D5, there is a point \( (s^a, u^a) \in R_m^a \). Let \( \lambda^a(t^a) = \sigma = (\mu_0, \ldots, \mu_{m-1}) \). We will get uncountably many points \( (s^a, u^a) \in R_m^a \setminus \text{cl}(R_{m+1}^a) \) by adding one more measure to the beginning of the finite sequence \( \sigma \) and using Lemma D3. As in Lemma D4, there is a finite set \( U \subseteq R_{m-1}^a \setminus R_m^a \) such that \( \text{proj}_{S^b} U = \text{proj}_{S^b} R_{m-1}^b \) and \( \mu_i(U) = 0 \) for all \( i < m \). Since \( R_{m-1}^b \setminus \text{cl}(R_m^b) \) is uncountable, there is a point \( (s^b, t^b) \in R_{m-1}^b \setminus \text{cl}(R_m^b) \) such that \( \mu_i(s^b, t^b) = 0 \) for all \( i < m \). Therefore we may also take \( U \) to contain such a point \( (s^b, t^b) \). Let \( \nu \) be a probability measure such that \( \nu(U) = 1, \text{mark}_{S^b} \nu = \text{mark}_{S^b} \mu_0 \), and \( \nu(s^b, t^b) = \text{mark}_{S^b} \mu_0(s^b) \). Since \( R_{m-1}^b \) is assumed under \( \sigma \) at level 0, we have \( (s^b, t^b) \in \text{Supp} \mu_0 \), and thus \( \mu_0(\{s^b\} \times T^b) = \text{mark}_{S^b} \mu_0(s^b) > 0 \). Therefore \( \nu(s^b, t^b) > 0 \).

Let \( \tau \) be the finite sequence \( (\nu, \mu_0, \ldots, \mu_{m-1}) \). By completeness there exists \( v^a \in T^a \) such that \( \lambda^a(v^a) = \tau \). As in Lemma D4, we have \( (s^a, u^a) \in R_m^a \). Suppose \( u^a \approx v^a \). By Lemma D3, we have \( (s^a, u^a) \in R_m^a \). However, since \( (s^b, t^b) \notin \text{cl}(R_m^b) \), the measure \( \nu \) has an open neighborhood in which \( R_m^b \) receives probability < 1. (An example of such a neighborhood is the set \( \{v' : \nu'(V) > \nu(s^b, t^b)/2\} \), where \( V \) is an open neighborhood of \( (s^b, t^b) \) which is disjoint from \( R_m^b \).) Thus \( (s^a, u^a) \) has an open neighborhood which is disjoint from \( R_{m+1}^b \). By Lemma D3, there are uncountably many \( u^a \approx v^a \), and therefore \( R_m^a \setminus \text{cl}(R_{m+1}^a) \) is uncountable. \( \blacksquare \)

**Proof of Theorem 2.** By Lemma E2, the set \( R_0^a \setminus \text{cl}(R_1^a) \) is uncountable. Suppose that \( (s^b, t^b) \in \bigcap_{m} R_m^b \). Then, for each \( m \), we have that \( R_m^b \) is assumed under \( \lambda^b(t^b) \) at some level \( j(m) \). Moreover, the sequence \( j(m) \) is non-increasing. Using Lemma E3, we see by induction that for each \( m \), the set \( R_{2m}^b \setminus \text{cl}(R_{2m+1}^b) \) is uncountable and the set \( R_{2m+1}^b \setminus \text{cl}(R_{2m+2}^b) \) is uncountable. Then by
Lemma E1, for each $m$ we have $j(2m + 1) < j(2m)$. But this contradicts the fact that $\lambda^b(t^b)$ has finite length. ■

**Corollary E1** Suppose that neither player is indifferent. If $(s^a, t^a) \in R^a_m$ then $\lambda^a(t^a)$ has length at least $m$. Moreover, $(s^a, t^a) \in R^a_m$ if and only if $(s^a, t^a) \in R^a_m$ and $\lambda^a(t^a)$ has length $m$.

**Proof.** In this case, the set $R^b_m \setminus \text{cl}(R^b_{m+1})$ is uncountable for all $m$. Thus if $(s^a, t^a) \in R^a_m$ and $\lambda^a(t^a)$ has length $n$, then for each $0 < k \leq m$ the set $R^b_{m-k}$ is assumed under $\lambda^a(t^a)$ at some level $j$ where $j \geq k - 1$. ■
References


