The Implications of Finite-Order Reasoning

Adam Brandenburger†  Alexander Danieli‡  Amanda Friedenberg§

November 4, 2020

Abstract

The epistemic conditions of rationality and $m$th-order strong belief of rationality ($R^m$SBR, Battigalli and Siniscalchi, 2002) formalize the idea that players engage in contextualized forward-induction reasoning. This paper characterizes the behavior consistent with $R^m$SBR across all type structures. In particular, in a class of generic games, $R(m − 1)$SBR is characterized by a new solution concept we call an $m$-best response sequence ($m$-BRS). Such sequences are an iterative version of extensive-form best response sets (Battigalli and Friedenberg, 2012). The strategies that survive $m$ rounds of extensive-form rationalizability are consistent with an $m$-BRS, but there are $m$-BRS’s which are disjoint from the former set. As such, there is behavior that is consistent with $R(m − 1)$SBR but inconsistent with $m$ rounds of extensive-form rationalizability. We use our characterization to draw implications for the interpretation of experimental data. Specifically, we show that the implications are non-trivial in the 3-repeated Prisoner’s Dilemma and Centipede games.

1 Introduction

Suppose that each player is rational, each player thinks other players are rational, and so on ad infinitum. What are the implications for behavior? This fundamental question has organized the epistemic game theory literature. It has been asked in the context of both strategic-form and extensive-form games. It has been asked when rationality corresponds to ordinary subjective expected utility maximization, when rationality incorporates an admissibility requirement, and even when rationality departs from subjective expected utility maximization. Likewise, it has been asked when the phrase “thinks” corresponds to knowledge, to belief, and to many other modalities (e.g., assumption, initial belief, strong belief, etc).

This paper focuses on a bounded version of the basic question: What are the implications for behavior if each player is rational, each player thinks other players are rational, and so on up to $m$ levels—but

---

*We thank George Mailath and three referees for many helpful comments. In addition, we have benefited from conversations with Vince Crawford, Jessy Hsieh, Terri Kneeland, Xiaomin Li, Elliot Lipnowski, and Ernesto Rivera Mora. We also thank audiences at Bristol, Oxford, Penn State, Washington University in St. Louis, McGill, National University of Singapore, Florida State University, University of Chicago, the 2013 Asian Meeting of the Econometric Society, the 2013 Workshop on Logical Structure of Correlated Information Change, SAET, TARK, and the Game Theory Society World Congress for important input. Adam Brandenburger thanks the NYU Stern School of Business, NYU Shanghai, and J.P. Valles for financial support. Amanda Friedenberg and Alex Danieli thank NSF SES-1358008 for financial support. Amanda Friedenberg thanks the UCL Economics Department for its unbounded hospitality. Previous versions of this paper were circulated under the titles: “How Many Levels Do Players Reason: An Observational Challenge and Solution” and “Identification of Reasoning about Rationality.”

†Stern School of Business, Tandon School of Engineering, NYU Shanghai, New York University, adam.brandenburger@stern.nyu.edu, adambrandenburger.com
‡W.P. Carey School of Business, Arizona State University, Alexander.Danieli@asu.edu
§Department of Economics, University of Arizona, amanda@amandafriedenberg.org, amandafriedenberg.org
not \((m + 1)\) levels? We address this question in the context of extensive-form games, where players engage in so-called “contextualized forward-induction reasoning.” We go on to spell out implications for the interpretation of experimental data.

Consider the game in Figure 1.1, Battle of the Sexes (BoS) with an Outside Option.\(^1\) A natural first approach would involve some notion of iterated dominance—here, extensive-form rationalizability (EFR, Pearce, 1984). EFR gives the standard forward-induction prediction: Ann plays \(In-U\) and Bob plays \(L\). In particular, the strategy \(In-D\) is dominated, but \(In-U\) is not. So, under EFR, Bob believes Ann will play \(U\) conditional on BoS; therefore, he plays \(L\). Ann expects this and, so, plays \(In-U\). With this, the analyst may be tempted to conclude that, if Ann plays Out, it is not the case that Ann is rational and reasons (in the sense of forward induction) two levels about Bob’s rationality. However, we will see that this conclusion is incorrect.

To see this, we adopt the epistemic approach. In particular, we expand the description of the game to include the players’ hierarchies of beliefs about the play of the game. To understand why this is needed, note that we will be interested in which of Ann’s strategies are rational. The strategy \(Out\) is rational—i.e., a best response—if Ann believes that Bob will play \(R\); it is irrational if she believes that Bob will play \(L\). Thus, we cannot specify whether \(Out\) is rational or irrational without specifying Ann’s beliefs about Bob’s play of the game.

\[\begin{array}{c|cc}
   & Bob & \\
\hline
Ann & & \\
   \hspace{1cm} & 3 & \\
   & \text{Out} & \hspace{1cm} \text{In} \\
\hline
   U & 4,2 & 0,0 \\
   D & 0,0 & 2,4 \\
\end{array}\]

Figure 1.1: Battle of the Sexes with an Outside Option

In fact, we will need to specify hierarchies of conditional beliefs about the play of the game. To see this, consider the case where Ann believes that Bob plays \(R\). This strategy will be rational for Bob if, conditional on BoS being played, Bob believes that Ann plays \(D\); but this same strategy is irrational for Bob if, conditional on BoS being played, Bob believes Ann plays \(U\). Thus, to specify whether Ann is “rational and believes that Bob is rational,” we need to specify both (i) Ann’s belief about the strategy Bob plays and (ii) her belief about “Bob’s conditional belief about her own play” (where the conditioning is on BoS). Continuing along these lines, we need to specify Ann’s hierarchies of conditional beliefs about the play of the game.

To do this, we build on the framework and analyses in Battigalli and Siniscalchi (2002) and Battigalli and Friedenberg (2012). We describe the situation by what is called an epistemic game. This is the game itself (e.g., BoS with an Outside Option) plus an epistemic type structure. Each type in an epistemic type

\(^1\) Here, and in other examples, \(\ast\) indicates that the payoff is irrelevant.
structure is associated with a conditional probability system, which describes a player’s belief at every
information set—including beliefs at information sets that are, \textit{ex ante}, not expected to be reached. For
instance, in BoS with an Outside Option, if Bob believes Ann plays \textit{Out}, he does not expect his information
set to be reached. Nevertheless, a conditional probability system specifies his belief conditional on this
event. This framework allows us to specify the requirement that each player is rational, thinks the other
is rational, etc.

Our baseline epistemic condition is rationality and common strong belief of rationality (RCSBR). There
are two ingredients to this condition: (extensive-form) rationality and strong belief. Extensive-form ra-
tionality requires that a player chooses a strategy that, at each information set, maximizes her subjective
expected utility given her type’s conditional probability system. For strong belief note that, in an extensive
form, we cannot simply require that a player “believes” an event—we must specify where, in the game,
the player believes that event.\footnote{For this reason, we follow the modern literature and refrain from using the phrase “mutual belief of rationality” for the idea that a player thinks or reasons that the other player is rational.} Strong belief requires that a player begins the game with a belief that the
event is true and maintains that hypothesis so long as it is not contradicted by evidence. Returning to
BoS with an Outside Option, if Bob strongly believes Ann is rational and reaching the BoS subgame is
consistent with Ann’s rationality, then he assigns probability 1 to her rationality conditional upon the BoS
subgame. (If reaching the BoS subgame is inconsistent with Ann’s rationality, strong belief of rationality
only requires that Bob begin the game assigning probability 1 to Ann’s rationality.) Thus, under RCSBR,
a player rationalizes past behavior as long as possible.

Battigalli and Siniscalchi (2002) and Battigalli and Friedenberg (2012) both study the behavioral impli-
cations of RCSBR. However, they do so under importantly different assumptions. Battigalli and Siniscalchi
(2002) focus on the case of a complete type structure (Brandenburger, 2003). Loosely, this corresponds to
an assumption that players have all possible hierarchies of beliefs. (See Friedenberg, 2010.) By contrast,
Battigalli and Friedenberg (2012) depart from the complete type structure assumption. The idea is that
there may be a context to the strategic situation and that context may frame the hierarchies of beliefs that
the players consider possible. For instance, history or social/cultural norms may impact the players’ beliefs.
(See, the discussion in Brandenburger, Friedenberg and Keisler, 2008, Section 2.8.) Under this viewpoint,
a complete type structure represents a special “context-free” case, in which there are no restrictions on the
players’ beliefs.

The distinction between the complete (context-free) vs. incomplete (contextualized) cases has important
implications for RCSBR (and, in particular, for RCSBR behavior). Battigalli and Siniscalchi (2002) show
that, for a complete type structure, RCSBR behavior corresponds to EFR. So, in BoS with an Outside
Option, Ann plays \textit{In-U} and Bob plays \textit{L}. By contrast, Battigalli and Friedenberg (2012) characterize
RCSBR across all type structures. In particular, it is characterized by a solution concept they call an
extensive-form best response set (EFBRS), a natural extensive-form analog to Pearce’s (1984) (strategic-
form) best response solution concept. The EFR strategy set is one EFBRS, but there are others. For
instance, in BoS with an Outside Option, there is an EFBRS where Ann plays \textit{Out}. This can arise if,
for instance, history leads Ann to maintain the hypothesis (throughout the game) that Bob is “tough”
and so “goes for his best outcome” and plays \textit{R}. (This is formalized as an incomplete type structure.
See Example 3.1.) This would lead any rational Ann to play \textit{Out}. Moreover, this is consistent with
contextualized forward-induction reasoning. In particular, when Bob has such a reputation, the observation
of \textit{In} is inconsistent with Ann being rational. As such, Bob can both strongly believe Ann is rational and,
conditional on BoS, believe that playing $R$ would lead to his best outcome.

This paper studies the context-dependent behavioral implications of rationality and $m^{th}$-order strong belief of rationality, namely $R_m\text{SBR}$. (We take $R_0\text{SBR}$ to mean just “rationality.”) As in Battigalli and Friedenberg (2012), we seek to characterize the $R_m\text{SBR}$ behavior across all type structures. This characterization is of particular interest when the analyst does not know the relevant type structure. As such, to infer that a strategy is consistent vs. inconsistent with $R_m\text{SBR}$, the analyst needs the $R_m\text{SBR}$ prediction across all type structures.

There is a natural approach to characterize the $R_m\text{SBR}$ behavior across all type structures: “unwrap” the EFBRS solution concept, thereby converting it from a fixed-point definition into an iterative definition. This leads to a sequence of predictions, viz. $(Q^0, \ldots, Q^m)$. We call such a sequence an $m$-(extensive-form) best response sequence (or an $m$-BRS). For a given type structure, the sequence of behavior consistent with $R_0\text{SBR}$, $\ldots$, $R(m-1)\text{SBR}$ gives rise to an $m$-BRS. (See Proposition 6.1.) However, there are $m$-BRS’s that, in a certain sense, do not reflect $R(m-1)\text{SBR}$ behavior. (Examples 6.1-6.2 make this precise.) Nonetheless, the Theorem 6.1 establishes that, in a class of generic games, the $m$-BRS’s characterize $R_0\text{SBR}$, $\ldots$, $R(m-1)\text{SBR}$ behavior. (See Definition 6.4 for the meaning of “generic.”

There is behavior consistent with an $m$-BRS but inconsistent with $m$-EFR—i.e., the strategies that survive $m$ rounds of EFR. More precisely, we can have an $m$-BRS $(Q^0, \ldots, Q^m)$ with $Q^m$ disjoint from the set of $m$-EFR strategies. This is to be expected given that an EFBRS may be disjoint from EFR. Both scenarios can arise because strong belief is non-monotonic.\(^3\) (The discussion following Examples 4.1-4.2 elaborates on this point.) This said, the $m$-EFR strategies are consistent with one $m$-BRS.

What does this mean for identifying levels of forward-induction reasoning from experimental data? In BoS with an Outside Option, there is a 3-BRS in which Ann plays $Out$—something that is inconsistent with 3-EFR. However, this fact is not surprising, given that $Out$ is also consistent with an EFBRS. In fact, in BoS with an Outside Option, all undominated strategies are consistent with an EFBRS. But, there are other games where the distinction between $m$-BRSs and $m$-EFR is both informative and important for the interpretation of experimental data. We give two such examples.

First, we look at the 3-repeated Prisoner’s Dilemma. Round-for-round, EFR gives the expected theoretical prediction. Round 1 has players defect at all third-period histories, round 2 has players defect at all second-period histories, and round 3 has players defect initially. The $m$-BRS’s give the same path of play, but they allow for additional strategies. As such, the inference from observed behavior is more subtle. In particular, if the experimentalist uses the direct response method and observes a subject choose to cooperate in the second-period, then—irrespective of the behavior of the other subjects—she cannot conclude that the behavior is inconsistent with “rationality and strong belief of rationality.” If, instead, the experimentalist uses the strategy method and observes a subject choose to cooperate at every second-period history, then she can conclude that the behavior is inconsistent with “rationality and strong belief of rationality.”

Second, we turn to the Centipede game (Rosenthal, 1981), where EFR corresponds round-for-round to backward induction. Thus, one might conjecture that, if we observe a subject play $in$, the behavior indicates a bound on strategic reasoning and, the longer the subject plays $in$, the lower that bound. This intuition is incomplete, at least if we take “strategic reasoning” to reflect RCSBR. In particular, we will see that the intuition is correct for the first mover, but incorrect for the second mover.

The paper proceeds as follows. Sections 2-4 introduce the formalism and epistemic conditions. Section

\(^3\)Note, this is different from (strategic-form) rationalizability: Best response sets are contained in the rationalizable strategy set. This follows from the fact that belief is monotonic, i.e., if $E$ is believed and $E \subseteq F$ then $F$ is believed.
5 reviews the characterization of RCSBR. It is used as a benchmark result, to motivate our main result. Section 6 defines the \( m \)-BRS concept and shows the main result: A characterization of \( R(m-1) \)SBR sequences in terms of \( m \)-BRS’s. Section 7 uses the \( m \)-BRS concept to analyze the 3-repeated Prisoner’s Dilemma and Centipede. Section 8 concludes with a discussion. It highlights several issues that are important for applying the \( m \)-BRS concept, including which games are generic (8.a), termination of the \( m \)-BRS procedure (8.b), and when arrays can replace CPS’s (8.c). Proofs can be found in the appendices.

2 Extensive-Form Games

Write \( \Gamma \) for a finite two-player extensive-form game with perfect recall, with (potentially) simultaneous moves, and without moves by nature. The players are \( a \) (Ann) and \( b \) (Bob).\(^4\) Write \( c \) for an arbitrary player in \( \{a, b\} \) and \( -c \) for the player in \( \{a, b\}\setminus\{c\} \). The underlying game tree has a set of non-terminal nodes (or vertices) \( V \) and a set of terminal nodes \( Z \). Write \( \phi \in V \) for the initial node. As in Osborne and Rubinstein (1994), we often identify nodes with histories. Let \( H_c \) be the set of information sets of \( c \). The set of information sets is \( H = H_a \cup H_b \). Player \( c \)'s extensive-form payoff function is given by \( \Pi_c : Z \to \mathbb{R} \).

Let \( S_c \) be the set of strategies for player \( c \) and let \( S = S_a \times S_b \). Assume the game is non-trivial, in the sense that \( |S_a|, |S_b| \geq 2 \). There is a mapping \( \zeta : S \to Z \) so that \( \zeta(s_a, s_b) \) is the terminal node reached by \( (s_a, s_b) \) reaches \( h \in H \) if the path from \( \phi \) to \( \zeta(s_c, s_{-c}) \) passes through some node in \( h \). Write \( S(h) \) for the set of strategy profiles that reach \( h \) and write \( S_c(h) = \text{proj}_{S_c}S(h) \). If a strategy \( s_c \in S_c(h) \), then we say that \( s_c \) allows \( h \in H \).

Player \( c \)'s strategic-form payoff function is given by \( \pi_c = \Pi_c \cdot \zeta \). We extend \( \pi_c \) to \( S_c \times \mathcal{P}(S_{-c}) \) in the usual way, i.e., \( \pi_c(s_c, \mu) = \sum_{s_{-c}} \pi_c(s_c, s_{-c})\mu(s_{-c}) \). Say \( s_c \in X_c \) is optimal under \( \mu \in \mathcal{P}(S_{-c}) \) given \( X_c \) if \( \pi_c(s_c, \mu) \geq \pi_c(r_c, \mu) \) or each \( r_c \in X_c \).

3 Type Structures

This section uses type structures to implicitly model the players’ hierarchies of beliefs about the play of the game. In defining such structures, we use the language of conditional probability systems. To understand the importance of doing so, refer back to BoS with an Outside Option. Suppose Bob begins the game assigning probability 1 to Ann playing \( \text{Out} \). Then, conditional upon observing Ann play \( \text{In} \), Bob can no longer hold that belief. So, to specify which strategy is rational for Bob, we must specify his belief conditional upon Ann playing \( \text{In} \).

**Conditional Probability Systems** Fix a compact metric space \( \Omega \). Write \( \mathcal{P}(\Omega) \) for the set of Borel probability measures on \( \Omega \). Endow \( \mathcal{P}(\Omega) \) with the topology of weak convergence so that it is again a compact metric space. Call \( (\Omega, \mathcal{E}) \) a conditional probability space if \( \mathcal{E} \subseteq 2^\Omega \setminus\{\emptyset\} \) is a finite set, where each \( E \in \mathcal{E} \) is Borel. The collection \( \mathcal{E} \) is a set of conditioning events.

**Definition 3.1.** An array on \( (\Omega, \mathcal{E}) \) is some \( p = (p(\cdot|E) : E \in \mathcal{E}) \) so that, for each \( E \in \mathcal{E}, \) \( p(\cdot|E) \in \mathcal{P}(\Omega) \) with \( p(E|E) = 1 \).

\(^4\)The analysis extends to three or more players, provided two assumptions hold. First, a player may have a correlated belief about other players. Second, players engage in so-called “correlated rationalization,” e.g., if they observe irrational behavior by one player they are willing to assume that other players may also be irrational.
Definition 3.2. A conditional probability system (CPS) on \((\Omega, \mathcal{E})\) is an array \(p = (p(\cdot | E) : E \in \mathcal{E})\) that satisfies the following criterion: If \(E, F \in \mathcal{E}\) with \(G \subseteq F \subseteq E\), then \(p(G | E) = p(G | F)p(F | E)\).

An array \(p\) specifies a belief, viz. \(p(\cdot | E)\), for each conditioning event \(E\). We refer to the beliefs \(p(\cdot | E)\) as conditional beliefs. If the array is a CPS, then the conditional beliefs must satisfy the rules of conditional probability when possible. Write \(\mathcal{A}(\Omega, \mathcal{E})\) for the set of arrays on \((\Omega, \mathcal{E})\) and write \(\mathcal{C}(\Omega, \mathcal{E})\) for the set of CPS’s on \((\Omega, \mathcal{E})\). Note that \(\mathcal{C}(\Omega, \mathcal{E}) \subseteq \mathcal{A}(\Omega, \mathcal{E}) \subseteq [\mathcal{P}(\Omega)]^\mathcal{E}\). Endow \([\mathcal{P}(\Omega)]^\mathcal{E}\) with the product topology and \(\mathcal{C}(\Omega, \mathcal{E})\) with the relative topology, so that \(\mathcal{C}(\Omega, \mathcal{E})\) is a compact metric space.

Type Structures In our analysis, player \(c\)’s set of conditioning events will correspond to
\[
\mathcal{E}_c = \{S_c(h) : h \in H_c \cup \{\phi\}\}.
\]
So, Ann has a conditioning event that corresponds to the beginning of the game, namely, \(S_b(\phi) = S_b\). She also has conditioning events \(S_b(h)\) corresponding to each information set \(h \in H_a\) at which she moves.

Definition 3.3. A \(\Gamma\)-based type structure is some \(T = (\Gamma; T_a, T_b; \beta_a, \beta_b)\) where
1. \(T_c\) is a compact metric type space for player \(c\), and
2. \(\beta_c : T_c \to \mathcal{C}(S_{-c} \times T_{-c}, \mathcal{E}_c \otimes T_{-c})\) is a continuous belief map for player \(c\).

So, each type of Ann, viz. \(t_a\), is associated with a CPS \(\beta_a(t_a)\) on \((S_b \times T_b, \mathcal{E}_a \otimes T_b)\). And, similarly for Bob. When \(\Gamma\) is a simultaneous-move game, the set of CPS’s is the set of probability measures and so \(\beta_c : T_c \to \mathcal{P}(S_{-c} \times T_{-c})\).

For any given game \(\Gamma\), there are infinitely many \(\Gamma\)-based type structures. Write \(T(\Gamma)\) for the family of \(\Gamma\)-based type structures. Battigalli and Siniscalchi (1999) construct a canonical type structure, which induces all possible hierarchies of conditional beliefs. Their type structure \(T^* = (\Gamma; T^*_a, T^*_b; \beta^*_a, \beta^*_b)\) has the property that it is type-complete (Brandenburger, 2003), i.e., for each CPS \(p_c \in \mathcal{C}(S_{-c} \times T^*_{-c}, \mathcal{E}_c \otimes T^*_{-c})\), there is a type \(t_c\) with \(\beta_c(t_c) = p_c\). Other type structures model an assumption that some event is (what is called) common full belief. (See Appendix A in Battigalli and Friedenberg, 2009 for a formal treatment.) The following example informally illustrates such an assumption.

Example 3.1. Consider BoS with an Outside Option. Suppose it is commonly understood that “Bob is tough” and, so, whenever a BoS game is played, he attempts to go for his best option and play \(R\). In particular:

Tough-1: at both the start of the game and conditional on BoS, Ann believes that Bob plays \(R\),

Tough-2: at both the start of the game and conditional on BoS, Bob believes “Tough-1,”

Tough-3: at both the start of the game and conditional on BoS, Ann believes “Tough-2,”

etc. This is a restriction on the hierarchies of beliefs that the players consider possible. There are no additional restrictions on the players’ beliefs.

We can capture this restriction on beliefs by a type structure \(T = (\Gamma; T_a, T_b; \beta_a, \beta_b)\) that satisfies the following properties:

- Each \(\beta_a(t_a)(\cdot | S_b \times T_b)\) assigns probability 1 to \(\{R\} \times T_b\).
For each CPS \( p_a \) with \( p_a(\{R\} \times T_b|S_b \times T_b) = 1 \), there is a type \( t_a \) with \( \beta_a(t_a) = p_a \).

For each CPS \( p_b \), there is a type \( t_b \) with \( \beta_b(t_b) = p_b \).

The first requirement says that, at the start of the game, each type of Ann assigns probability 1 to “Bob plays \( R \).” By the conditioning requirement, this implies that, conditional on BoS, each type continues to assign probability 1 to “Bob plays \( R \).” The second requirement says that, for each CPS that satisfies the first requirement, there is a type of Ann that holds that belief. Likewise, the third requirement says that for each CPS of Bob, there is a type of Bob that holds that belief. The second and third requirements capture the idea that there are no additional restrictions on the players’ beliefs. The fact that such a type structure exists follows from Battigalli and Friedenberg (2009).

**Epistemic Game** For a given game \( \Gamma \), write \( T(\Gamma) \) for the family of \( \Gamma \)-based type structures. Since \( \Gamma \) is non-trivial, there is an uncountable number of elements in \( T(\Gamma) \). An (extensive-form) epistemic game is some pair \((\Gamma, T)\) with \( T \in T(\Gamma) \). The epistemic game is the exogenous description of the strategic situation.

In what follows, we fix an extensive-form game \( \Gamma \). So, each epistemic game can be identified with a type structure in \( T(\Gamma) \). As such, we often conflate ‘type structure’ with ‘epistemic game.’ No confusion should result.

## 4 Epistemic Conditions

Fix an epistemic game \((\Gamma, T)\). It induces a set of states \( S_a \times T_a \times S_b \times T_b \). A state 

**Definition 4.1.** Say \( s_c \) is a sequential best response under \( p_c \in A(S_{-c}, \mathcal{E}_c) \) if, for each \( h \in H_c \) with \( s_c \in S_c(h) \), \( s_c \) is optimal under \( p_c(\cdot|S_{-c}(h)) \) given \( S_c(h) \).

Write \( \mathbb{BR}[p_c] \) for the set of strategies \( s_c \) that are a sequential best response under \( p_c \in A(S_{-c}, \mathcal{E}_c) \). So, \( s_c \in \mathbb{BR}[p_c] \) if and only if \( s_c \) is optimal under each of the conditional beliefs \( p_c(\cdot|S_{-c}(h)) \), provided \( h \) is an information set allowed by \( s_c \).

Each \( \beta_c(t_c) \) induces a CPS in \( \mathcal{C}(S_{-c}, \mathcal{E}_c) \) via marginalization. The marginal CPS \( \text{marg}_{S_{-c}} \beta_c(t_{-c}) \) is a CPS \( p_c \in \mathcal{C}(S_{-c}, \mathcal{E}_c) \) with \( p_c(\cdot|S_{-c}(h)) = \text{marg}_{S_{-c}} \beta_c(t_{-c})(\cdot|S_{-c}(h) \times T_{-c}) \) for each \( S_{-c}(h) \in \mathcal{E}_c \).

**Definition 4.2.** Say \((s_c, t_c)\) is rational if \( s_c \in \mathbb{BR}[\text{marg}_{S_{-c}} \beta_c(t_{c})] \). So, \((s_c, t_c)\) is rational if \( s_c \) is a sequential best response under the marginal CPS \( \text{marg}_{S_{-c}} \beta_c(t_{c}) \).

**Definition 4.3** (Battigalli and Siniscalchi, 2002). Say a CPS \( p \in \mathcal{C}(\Omega, \mathcal{E}) \) strongly believes an event \( F \) if, for each conditioning event \( E \in \mathcal{E} \), \( E \cap F \neq \emptyset \) implies \( p(F|E) = 1 \).

**Definition 4.4.** A type \( t_c \) strongly believes an event \( E_{-c} \subseteq S_{-c} \times T_{-c} \) if \( \beta_c(t_c) \) strongly believes \( E_{-c} \).

Strong belief asks that a type maintain a hypothesis so long as it is not contradicted by observed play. Thus, it requires that a type rationalize past behavior when possible. In this sense, it captures forward-induction reasoning (relative to the type structure \( T \)). See Battigalli and Siniscalchi (2002) and Battigalli and Friedenberg (2012) for a more complete discussion.
Set $R_e^0(T) = S_e \times T_e$. Let $R_e^1(T)$ be the set of rational strategy-type pairs $(s_e, t_e)$. Inductively define sets $R_a^m(T)$ and $R_b^m(T)$ by

$$R_e^{m+1}(T) = R_e^m(T) \cap [S_e \times \{ t_e : t_e \text{ strongly believes } R_e^m(T) \}] .$$

And, set $R_e^\infty(T) = \bigcap_{m \geq 0} R_e^m(T)$. Write $R^m(T) = R_a^m(T) \times R_b^m(T)$ and $R^\infty(T) = R_a^\infty(T) \times R_b^\infty(T)$.

**Definition 4.5.** The set $R^{m+1}(T)$ is the set of strategy-type pairs (in $(\Gamma, T)$) at which there is rationality and $m^{th}$-order strong belief of rationality (RmSBR). The set $R^\infty(T)$ is the set of strategy-type pairs (in $(\Gamma, T)$) at which there is rationality and common strong belief of rationality (RCSBR).

Note, the set of rational strategy-type pairs depends on the epistemic game $(\Gamma, T)$. In this paper, we fix $\Gamma$ but not necessarily the associated type structure. As such, we write $R^1(T)$ to indicate the set of states at which there is rationality in the epistemic game associated with $T$. And, analogously, for $R^m(T)$.

Observe that $\text{proj}_{S_a \times S_b} R^{m+1}(T)$ is the set of RmSBR predictions for the type structure $T$. A natural conjecture is that the set of RmSBR predictions is the set of strategies that survive $(m+1)$ rounds of extensive-form rationalizability (EFR, Pearce, 1984). EFR sequentially eliminates strategies that are not sequential best responses. Battigalli and Siniscalchi (2002, Proposition 6) show that, when the type structure $T^*$ is type-complete, the set of RmSBR predictions is the set of strategies that survive $(m+1)$ rounds of EFR. However, this need not be the case for a type-incomplete structure $T$. In that case, the set of RmSBR predictions may be disjoint from the set of strategies that survive $(m+1)$ rounds of EFR. The following examples illustrate this point.

**Example 4.1.** Consider BoS with an Outside Option. There is no CPS so that $\text{In-D}$ is a sequential best response. However, for each strategy $s_a \in \{ \text{Out}, \text{In-U} \}$ (resp. $s_b \in \{ L, R \}$), there is some CPS under which $s_a$ (resp. $s_b$) is a best response. Thus, one round of EFR gives

$$\text{EFR}^1_a \times \text{EFR}^1_b = \{ \text{Out}, \text{In-U} \} \times \{ L, R \} .$$

Now observe that a CPS on $S_a$ that strongly believes $\text{EFR}^1_a$ must assign probability 1 to $\text{In-U}$ conditional on BoS; $L$ is the unique sequential best response under any such CPS. As such,

$$\text{EFR}^2_a \times \text{EFR}^2_b = \{ \text{Out}, \text{In-U} \} \times \{ L \} .$$

So, $\text{In-U}$ is the unique sequential best response under any CPS on $S_b$ that strongly believes $\text{EFR}^2_b$. Thus,

$$\text{EFR}^3_a \times \text{EFR}^3_b = \{ \text{In-U} \} \times \{ L \} .$$

As such, there is one EFR strategy profile, $(\text{In-U}, L)$.

Battigalli and Siniscalchi (2002) show that, if $T^*$ is type-complete, then EFR corresponds round-for-round to RmSBR in the associated epistemic game. That is, for each $m$, $\text{proj}_{S_a \times S_b} R^m(T^*) = \text{EFR}^m_a \times \text{EFR}^m_b$. So, the EFR predictions are the RmSBR predictions, when the type structure is type-complete. □

**Example 4.2.** Again, consider BoS with an Outside Option. Let $T$ be the type structure from Example 3.1, representing the case where it is commonly understood that “Bob is tough.” Now, for each $m \geq 1$, $\text{proj}_{S_a \times S_b} R^m(T) = \{ \text{Out} \} \times \{ L, R \}$. So, for each $m \geq 1$, there are types $t_a^m$ and $t_b^m$ so that $(\text{Out}, t_a^m, R, t_b^m) \in R^m(T)$.
Let us review the difference between Examples 4.1-4.2. Begin with Example 4.2. Observe that there is a type $t^*_b \in T_b$ that, at the initial node, assigns probability 1 to $\{Out\} \times T_a$ and, conditional upon Ann playing In, assigns probability 1 to $\{In-D\} \times T_a$. Certainly $(R, t^*_b)$ is rational. In addition, $t^*_b$ strongly believes the event that “Ann is rational in the epistemic game $(\Gamma, T)$,” viz. $R^1_a(T)$: At the initial node, the type assigns probability 1 to the event $R^*_a(T)$ and the event $R^1_a(T)$ is inconsistent with Ann playing In. Now turn to Example 4.1 and an associated complete type structure $T^*$. There is also a type $t^*_b \in T_b^*$ that, at the initial node, assigns probability 1 to $\{Out\} \times T^*_a$ and, conditional upon Ann playing In, assigns probability 1 to $\{In-D\} \times T^*_a$. However, this type does not strongly believe the event that “Ann is rational in the epistemic game $(\Gamma, T^*)$,” viz. $R^1_a(T^*)$, since

$$R^1_a(T^*) \cap (\{In-U\} \times T^*_a) \neq \emptyset$$

and, in particular, is contained in $\{In-U\} \times T^*_a$. The key is the following:

$$\{Out\} = \text{proj}_{S_a} R^1_a(T) \subseteq \text{proj}_{S_a} R^1_a(T^*) = \{Out, In-U\}.$$

As such, we can have a CPS that assigns positive probability to $\{In-D\} \times T_a$ conditional on BoS and strongly believes $R^1_a(T)$, but there is no CPS that assigns positive probability to $\{In-D\} \times T^*_a$ conditional on BoS and strongly believes $R^1_a(T^*)$. This is possible because of the non-monotonicity of strong belief.

## 5 The EFBRS Benchmark

We focus on the case where the analyst does not know the players’ type structure. Thus, we are interested in characterizing the RmSBR predictions across all type structures. We begin with the RCSBR benchmark, where the analogous characterization is known.

Fix a type structure $T$. The set of RCSBR predictions for $T$ is

$$\text{proj}_{S_a \times S_b} R^\infty_a(T) = \text{proj}_{S_a} R^\infty_a(T) \times \text{proj}_{S_b} R^\infty_b(T).$$

Fix a predicted strategy $s_a \in \text{proj}_{S_a} R^\infty_a(T)$ and a type $t_a$ so that $(s_a, t_a) \in R^\infty_a(T)$. Write $p_a$ for the marginal CPS of $t_a$, i.e. $p_a = \text{marg}_{S_b} \beta_a(t_a)$. Observe that $(s_a, p_a)$ must satisfy three properties: First, $s_a$ must be a sequential best response under $p_a$. (This follows from the fact that $(s_a, t_a)$ is rational.) Second, $p_a$ must strongly believe Bob’s RCSBR prediction for $T$, namely $\text{proj}_{S_a} R^\infty_b(T)$. (This follows from the fact that $t_a$ strongly believes $R^\infty_b(T)$.) Finally, if $r_a$ is also a sequential best response under $p_a$, $r_a$ must be contained in $\text{proj}_{S_a} R^\infty_a(T)$, i.e., $r_a$ must be one of Ann’s RCSBR predictions for $T$. (This follows from the following property: If $(r_a, t_a)$ is rational and $(s_a, t_a)$ satisfies RCSBR for $T$, then $(r_a, t_a)$ satisfies RCSBR for $T$.) This last property can be viewed as a maximality condition. These three properties motivate the definition of an EFBRS.

**Definition 5.1.** Call $Q_a \times Q_b \subseteq S_a \times S_b$ an extensive-form best response set (EFBRS) if, for each $s_c \in Q_c$, there exists some CPS $p_c \in C(S_{-c}, E_c)$ so that the following hold:

1. $s_c \in \text{BR}[p_c]$,
2. $p_c$ strongly believes $Q_{-c}$, and
(3) if \( r_c \in \mathbb{BR}[p_a] \), then \( r_c \in Q_c \).

An EFBRS is a subset of strategies \( Q_a \times Q_b \) that satisfies a certain fixed-point requirement: For each \( s_a \in Q_a \), there exists a CPS \( p_a \) defined only on \( S_b \) so that (i) \( s_a \) is a sequential best response under \( p_a \), (ii) \( p_a \) strongly believes Bob’s prediction \( Q_b \), and (iii) \( Q_a \) satisfies a requisite maximality property. These properties correspond to the properties derived from RCSBR—but are properties defined on the game \( \Gamma \) itself.

**Proposition 5.1** (Battigalli and Friedenberg, 2012).

(i) For each type structure \( T \), \( \text{proj}_S R_{\infty}(T) \) is an EFBRS.

(ii) Given an EFBRS \( Q_a \times Q_b \), there exists a type structure \( T \) so that \( \text{proj}_S R_{\infty}(T) = Q_a \times Q_b \).

**Corollary 5.1.** For each game \( \Gamma \),

\[
S_{\infty} := \bigcup_{Q_a \times Q_b \text{ is an EFBRS}} (Q_a \times Q_b) = \bigcup_{T \in T(\Gamma)} \text{proj}_S R_{\infty}(T).
\]

Corollary 5.1 says that the union over all EFBRSs corresponds to the union of the RCSBR predictions. So a strategy \( s_c \) is consistent with RCSBR if and only if \( s_c \in \text{proj}_S S_{\infty} \).

**Example 5.1.** Return to Figure 1.1. We will show that

\[
S_{\infty} = \left( \{\text{Out}\} \times \{L, R\} \right) \cup \left( \{\text{In-U}\} \times \{L\} \right).
\]

First, each EFBRS \( Q^a \times Q^b \subseteq S_{\infty} \). To see this, fix an EFBRS \( Q^a \times Q^b \). Since, at the start of the game, \( \text{In-D} \) is dominated, by Property (1), \( Q^a \subseteq \{\text{Out}, \text{In-U}\} \). So, if \( \text{In-U} \in Q_a \), then \( Q_b = \{L\} \). (This follows from Properties (1)-(2).) Thus, \( Q^a \times Q^b \subseteq S_{\infty} \). Next, \( \{\text{Out}\} \times \{L, R\} \), and \( \{\text{In-U}\} \times \{L\} \) are both EFBRSs.

Take \( Q_a \times Q_b = \{\text{Out}\} \times \{L, R\} \). \( \text{Out} \) is the unique sequential best response under a CPS that, at each information set, assigns probability 1 to \( \{R\} \subseteq Q_b \). \( L \) (resp. \( R \)) is the unique sequential best response under a CPS that initially assigns probability 1 to \( Q_a = \{\text{Out}\} \) and, conditional upon BoS, assigns probability 1 to \( \{\text{In-U}\} \) (resp. \( \{\text{In-D}\} \)). Analogously, for \( \{\text{In-U}\} \times \{L\} \).

Finally, note that \( S_{\infty} \) is not itself an EFBRS. Observe, it is not a product set. But, more importantly, \( R \) is not a best response given a CPS that strongly believes \( \text{proj}_{S_a} S_{\infty} = \{\text{Out}, \text{In-U}\} \).

\[ \square \]

6 The \( m \)-BRS

The EFBRS concept can be viewed as a collection of sets, each of which satisfy a certain fixed-point property: If \( s_a \) is contained in Ann’s solution \( Q_a \), then \( s_a \) is a sequential best response under a CPS that strongly believes Bob’s solution \( Q_b \). The EFBRS inherits this fixed-point property from RCSBR itself. If \( (s_a, t_a) \in R_{\infty}^a(T) \), then \( t_a \) strongly believes an event of the same order, namely \( R_{\infty}^b(T) \).

To obtain a finite-order analog, we need to depart from this fixed-point property—converting it into an iterative property. This is because \( Rm\text{SBR} \) is not a fixed-point concept: If \( (s_a, t_a) \in R_{\infty}^a(T) \setminus R_{\infty}^a(T) \), then \( t_a \) does not strongly believe the event \( R_{\infty}^b(T) \)—that is, \( t_a \) does not strongly believe the event of the same order. Instead, \( t_a \) strongly believes the lower-order events \( R_{\infty}^b(T), R_{\infty}^b(T), R_{\infty}^b(T) \). More generally, if \( (s_a, t_a) \in R_{\infty}^a(T) \setminus R_{\infty}^a(T) \), then \( t_a \) strongly believes the lower-order events \( R_{\infty}^b(T), \ldots, R_{\infty}^a(T) \). We will
build off this fact to go from the EFBRS concept to an iterative property. That property will apply to a
decreasing sequence of product sets.

Definition 6.1. Say \((Q^0, \ldots, Q^m)\) is a **decreasing sequence of product sets** if (i) \(Q^0 = S_a \times S_b\),
(ii) each \(Q^n = Q^n_a \times Q^n_b\) is a product set, and (iii) for each \(n = 0, \ldots, m - 1\), \(Q^{n+1} \subseteq Q^n\).

Definition 6.2. Say \(X = X_a \times X_b\) satisfies the *(extensive-form)* best response property relative to\n\((Q^0, \ldots, Q^m)\) if \((Q^0, \ldots, Q^m, X)\) is a decreasing sequence of product sets satisfying the following property:

For each \(s_c \in X_c\), there exists a CPS \(p_c \in C(S_c, E_c)\) so that

\[
\text{(BRP.1) } s_c \in \mathbb{BR}[p_c],
\]

\[
\text{(BRP.2) } p_c \text{ strongly believes } Q^0_{c}, \ldots, Q^m_{c}, \text{ and}
\]

\[
\text{(BRP.3) if } r_c \in \mathbb{BR}[p_c], \text{ then } r_c \in X_c.
\]

Definition 6.2 appears similar to Definition 5.1. The central difference arises in condition (BRP.2). Instead
of the CPS strongly believing \(X_{-c}\), the CPS strongly believes the lower-order sets \(Q^0_{-c}, \ldots, Q^m_{-c}\). (Note,
\(X_{-c} \subseteq Q^m_{-c} \subseteq Q^{m-1}_{-c} \subseteq \cdots \subseteq Q^0_{-c}\)).

Definition 6.3. Let \(m \geq 1\). Say \((Q^0, \ldots, Q^m)\) forms an **m-(extensive-form) best response sequence**
(m-BRS) if \(Q^1 \neq \emptyset\) and, for each \(n = 0, \ldots, m - 1\), \(Q^{n+1}\) satisfies the best response property relative to\n\((Q^0, \ldots, Q^n)\).

Remark 6.1. For each \(m \geq 2\), \((Q^0, \ldots, Q^m)\) is an m-BRS if and only if (i) \((Q^0, \ldots, Q^{m-1})\) is an
\((m - 1)\)-BRS and (ii) \(Q^m\) satisfies the best response property relative to \((Q^0, \ldots, Q^{m-1})\). \(\square\)

A 1-BRS is some \((Q^0, Q^1) = (S_a \times S_b, Q^1_a \times Q^1_b)\), where

\[
Q^1_c = \bigcup_{p_c \in E_c} \mathbb{BR}[p_c]
\]

for some non-empty \(E_c \subseteq \mathcal{C}(S_{-c}, E_c)\). An \((m + 1)\)-BRS is some \((Q^0, \ldots, Q^m, Q^{m+1})\), where \((Q^0, \ldots, Q^m)\) is
an m-BRS and \(Q^{m+1}\) satisfies the best response property relative to \((Q^0, \ldots, Q^m)\).\(^6\) Thus, it is an iterative
procedure that is a natural analog to the EFBRS. In fact:

Proposition 6.1. **For each** \(T\), the sequence \((\text{proj}_S R^0(T), \ldots, \text{proj}_S R^m(T))\) **forms an** m-BRS.

Say \(Q \subseteq S\) is **consistent with an** m-BRS if there exists some \((m - 1)\)-BRS, viz. \((Q^0, \ldots, Q^{m-1})\), so that \(Q\) satisfies the extensive-form best response property relative to \((Q^0, \ldots, Q^{m-1})\). By Proposition 6.1,

\[
\bigcup_{T \in \mathcal{T}(\Gamma)} \text{proj}_S R^m(T) \subseteq \bigcup_{Q \text{ is consistent with an } m\text{-BRS}} Q =: S^m.
\]

That is, the union over the m-BRS’s provides an upper bound on the behavior consistent with \(R(m-1)\)BRS
across all type structures. We denote the union over m-BRS’s as \(S^m\).

A natural analog to Corollary 5.1 is that

\[
S^m = \bigcup_{Q \text{ is consistent with an } m\text{-BRS}} Q = \bigcup_{T \in \mathcal{T}(\Gamma)} \text{proj}_S R^m(T), \quad (1)
\]

\(^6\)Note, if \((Q^0, \ldots, Q^{m+1})\) is an \((m + 1)\)-BRS and \(Q^m = Q_a^m \times Q_b^m\) with \(Q_a^m = \emptyset\), then \(Q_b^{m+1} = \emptyset\).
i.e., the union over the $m$-BRS’s is the set of strategies consistent with $R(m-1)$SBR in some type structure. In fact, there is a natural conjecture that would imply Equation (1): For each $m$-BRS $(Q^0, \ldots, Q^m)$, there exists a type structure $\mathcal{T}$ so that

$$Q^n \subseteq \text{proj}_S R^n(\mathcal{T}) \quad \text{for all } n = 1, \ldots, m. \quad (2)$$

However, this conjecture is incorrect. The next series of examples will illustrate the issues involved.

**Counterexamples** We begin by showing that Equation (2) cannot be strengthened to require equality.

**Example 6.1.** Consider the game in Figure 6.1. Let $(Q^0, Q^1, Q^2)$ be the decreasing sequence of product sets with

$$Q^1_a \times Q^1_b = S_a \times \{y_1q_1, y_1q_2, y_2\} \quad \text{and} \quad Q^2_a \times Q^2_b = \{x_2\} \times Q^1_b.$$

Note this is a 2-BRS.\(^7\) But, we show that there is no type structure $\mathcal{T}$ with $Q^1 = \text{proj}_S R^1(\mathcal{T})$ and $Q^2 = \text{proj}_S R^2(\mathcal{T})$.

Suppose otherwise. Then, there exists a type $t_a$ so that $(x_1z_1, t_a) \in R^1_a(\mathcal{T})$. Observe that, at each information set, $\beta_a(t_a)$ must assign probability 1 to $\{y_2\} \times T_b$. But, $y_2$ is a sequential best response under every CPS and, so, $\{y_2\} \times T_b \subseteq R^1_b(\mathcal{T})$. With this, $t_a$ strongly believes $R^1_b(\mathcal{T})$ and so $(x_1z_1, t_a) \in R^2_a(\mathcal{T})$. Thus, $Q^2_a \neq \text{proj}_S R^2_a(\mathcal{T})$. \(\square\)

Example 6.1 shows that we may have a 2-BRS $(Q^0, Q^1, Q^2)$ so that there is no type structure $\mathcal{T}$ with both $Q^1 = \text{proj}_S R^1(\mathcal{T})$ and $Q^2 = \text{proj}_S R^2(\mathcal{T})$. But, this is immaterial from the perspective of delivering Equation (1): There is some $\mathcal{T}$, with both $Q^1 \subseteq \text{proj}_S R^1(\mathcal{T})$ and $Q^2 \subseteq \text{proj}_S R^2(\mathcal{T})$. In fact, this conclusion holds more generally:

\(^7\)Let us point to three features of the example: First, $x_1z_1$ and $x_2$ are both sequential best responses under a CPS that assigns probability 1 to $y_2$; $x_1z_2$ is a unique sequential best response under a CPS that assigns probability 1 to $y_3$; and $x_2$ is a unique sequential best response under a CPS that assigns probability 1 to $\{y_1q_1, y_1q_2\}$. Second, $y_1q_1$ (resp. $y_1q_2$) and $y_2$ are the only strategies that are a sequential best response under a CPS that assigns probability 1 to $x_2$ at the initial information set and then assigns probability 1 to $x_1z_2$ (resp. $x_1z_1$) conditional upon observing $x_1$. Third, $y_2$ is a unique sequential best response under a CPS that assigns probability 1 to $x_1z_2$. 

12
Proposition 6.2.

(i) For each 1-BRS $(Q^0, Q^1)$, there exists some $T$ so that $Q^1 = \text{proj}_S R^1(T)$.

(ii) For each 2-BRS $(Q^0, Q^1, Q^2)$, there exists some $T$ so that $Q^1 = \text{proj}_S R^1(T)$ and $Q^2 \subseteq \text{proj}_S R^2(T)$.

In light of Proposition 6.2, Equation (1) does indeed hold for $m = 1, 2$. However, we will next see that an analog of Proposition 6.2 does not hold for 3-BRS’s.

Example 6.2. Return to the game in Figure 6.1. Let $(Q^0, Q^1, Q^2)$, where $(Q^0, Q^1, Q^2)$ is the 2-BRS described in Example 6.1 and

$$Q^3_a \times Q^3_b = Q^3_a \times \{ y_1, y_2 \}.$$ 

We will show that there is no type structure $T$ so that $Q^3 \subseteq \text{proj}_S R^3(T)$ for each $n = 1, 2, 3$.

Suppose, contra hypothesis, that such a type structure $T$ exists. Since $Q^3 \subseteq \text{proj}_S R^3(T)$, there exists some $t_b$ with $(y_1, q_1, t_b) \in R^3_b(T)$. Then, $\beta_b(t_b)$ must assign positive probability to $\{ x_1, x_2 \} \times T_a$ conditional on $\{ x_1, x_2 \} \times T_a$. We will argue that $\{ x_1, x_2 \} \times T_a \cap R^2_a(T) \neq \emptyset$ but $\{ x_1, x_2 \} \times T_a \cap R^2_a(T) = \emptyset$, contradicting the fact that $t_b$ strongly believes $R^2_a(T)$.

First, observe that $(x_1, z_1) \in Q^3_a$ and so, by assumption, $(x_1, z_1) \in \text{proj}_S R^3_a(T)$. Thus, repeating the argument in Example 6.1 above, $(x_1, z_1) \in \text{proj}_S R^3_a(T)$. Second, observe that $x_1, z_2$ is only a sequential best response under a CPS that assigns positive probability to $\{ y_3 \} \times T_b$ at the initial information set. Since $y_3$ is dominated, no such CPS can strongly believe $R^1_b(T)$. Thus, $x_1, z_2 \notin \text{proj}_S R^3_a(T)$. \Box

Example 6.2 gives a 3-BRS so that, if $Q^1 \subseteq \text{proj}_S R^1(T)$, then there exists some strategy in $Q^3$ that is not contained in $\text{proj}_S R^3(T)$. The key is that there is a strategy in $Q^3_a$ that is a sequential best response under a CPS that strongly believes $Q^3_a$. But, that CPS cannot strongly believe $\text{proj}_S R^3_a(T)$; this arises because $Q^3_a$ is a strict subset of $\text{proj}_S R^3_a(T)$.

Let us review what led to the situation where $Q^3_a \subseteq \text{proj}_S R^3_a(T)$. The strategy $x_1, z_1$ is a sequential best response under a CPS $p_a$ on $S_b$. However, for any CPS $\hat{p}_a$ on $S_b \times T_b$ with $p_a = \text{marg}_{S_b} \hat{p}_a$, we have that $\hat{p}_a$ strongly believes that “Bob is rational.” With this in mind, we now restrict attention to a class of games that are generic; in such games, this phenomenon (essentially) cannot arise.

**Generic Games** Say two strategies $s_c$ and $r_c$ are equivalent if they induce the same plan of action, i.e., $\zeta(s_c, \cdot) = \zeta(r_c, \cdot)$. Write $[s_c]$ for the set of strategies that are equivalent to $s_c$, and observe that, since the game is non-trivial, each $[s_c] \subseteq S_c$. So, if $s_c$ and $r_c$ are equivalent, then $\pi_c(s_c, \cdot) = \pi_c(r_c, \cdot)$. It follows that $s_c \in \mathbb{B}[p_c]$ if and only if $[s_c] \subseteq \mathbb{B}[p_c].$

**Definition 6.4.** Call a game generic if the following property holds: There exists a CPS $p_c \in C(S_{-c}, E_c)$ so that $s_c \in \mathbb{B}[p_c]$ if and only if there exists a CPS $q_c \in C(S_{-c}, E_c)$ so that $[s_c] = \mathbb{B}[q_c]$.\footnote{In BoS with an Outside Option, O-L and O-R are two equivalent strategies. We have simply been writing Out; our notation formally describes an equivalence class of strategies.}

Thus, a game is generic if any sequential best response is a “unique” sequential best response under some—perhaps different—CPS. Here, “unique” is taken to mean “up to equivalent strategies.” Section 8.a discusses which games are generic.

By restricting attention to generic games, we solve the problem that arose in Examples 6.1-6.2. To see this, observe that $x_1, z_1$ is only a sequential best response under a CPS that assigns probability 1 to
In Figure 6.2a, the diagonal of $Q_1$ is the union over red boxes along the pictorial-diagonal of $Q_1^1 \times Q_1^1$ and the diagonal of $Q_2^2 \times Q_2^2$ is the union over blue boxes along the pictorial-diagonal of $Q_2^2 \times Q_2^2$. The off-diagonal of $Q_1^1 \times Q_1^1$ is the white area in $Q_1^1 \times Q_1^1$ (formally, $(Q_1^1 \times Q_1^1) \setminus \text{diag}_c^1$).

The idea will be to construct belief maps so that $R_1^1(T)$ is contained in the union of squares $(Q_1^1 \times Q_1^1) \cup (Q_2^2 \times Q_2^2)$ and $R_2^2(T)$ is contained in the square $(Q_2^2 \times Q_2^2)$. Moreover, the belief maps will separate 1-types and 2-types based on whether (or not) they strongly believe rationality. Specifically:

1. If $(s_c, s_c^1) \in Q_1^1 \times Q_1^1$, then $(s_c, s_c^1)$ is rational and does not strongly believe rationality.

2. If $(s_c, s_c^2) \in Q_2^2 \times Q_2^2$, then $(s_c, s_c^2) \in Q_1^1 \times Q_1^1$ is rational and strongly believes rationality.
Since \((s_c, s_c^{'}) \in R^1_c(T)\) implies \([s_c] \times \{s_c^{'}) \subseteq R^1_c(T)\), these properties of belief maps will give
\[
\text{diag}_c^2 \subseteq R^2_c(T) \subseteq (Q^2_c \times Q^2_c) \quad \text{and} \quad \text{diag}_c^1 \subseteq R^1_c(T) \setminus R^2_c(T) \subseteq (Q^1_c \times Q^1_c).
\]

We may well have \(\text{diag}_c^2 \subseteq R^2_c(T) \subseteq (Q^2_c \times Q^2_c)\). That is, pictorially, \(R^2_c(T)\) may well contain both the diagonal blue boxes and the (off-diagonal) green box. However, we will require that \(\text{diag}_c^1 = R^1_c(T) \setminus R^2_c(T) \subseteq (Q^1_c \times Q^1_c)\). With this, each off-diagonal point in \(Q^1_c \times Q^1_c\) will be irrational. The role of this requirement will become clear below.

First, we construct the beliefs associated with 2-types. By definition of a 2-BRS, for each \(s_c \in Q^2_c\), there is a CPS \(j_c(s_c^{'})\) on \((S_{-c}, E_c)\) so that \([s_c] \subseteq \mathbb{B}R[j_c(s_c^{'})] \subseteq Q^2_c\) and \(j_c(s_c^{'})\) strongly believes \(Q^1_{-c}\). Choose \(\beta_c(s_c^{'})\) so that \(\text{marg}_{S_{-c}} \beta_c(s_c^{'}) = j_c(s_c^{'})\). Moreover, if \(S_{-c}(h) \cap Q^1_{-c} \neq \emptyset\), we require that \(\beta_c(s_c^{'})\) on \([|S_{-c}(h) \times T_{-c}|\) be concentrated on the diagonal of \(Q^1_{-c} \times Q^1_{-c}\). (We can do this since, in that case, \(j_c(s_c^{'}) (Q^1_{-c} | S_{-c}(h)) = 1\).

Next construct the beliefs associated with 1-types. Since the game is generic, for each \(s_c \in Q^1_c\), there is a CPS \(j_c(s_c^{'})\) on \((S_{-c}, E_c)\) so that \([s_c] = \mathbb{B}R[j_c(s_c^{'})]\). For the purpose of illustrating the construction, suppose that \(Q^1_{-c}\) has at least two non-equivalent strategies.\(^9\) Figure 6.2b illustrates this case; note that the off-diagonal (illustrated by the orange lines) is non-empty. Moreover, the off-diagonal meets each strategy in \(Q^1_{-c}\). (Formally, for each \(s_{-c} \in Q^1_{-c}\), \(\{s_{-c}\} \times T_{-c} \cap (Q^1_{-c} \times Q^1_{-c}) \setminus \text{diag}_{-c} \neq \emptyset\).) We can then choose \(\beta_c(s_c^{'})\) so that (i) \(\text{marg}_{S_{-c}} \beta_c(s_c^{'}) = j_c(s_c^{'})\), (ii) for each \(h\), \(\beta_c(s_c^{'}) (S_{-c} \times Q^1_{-c} | S_{-c}(h) \times T_{-c}) = 1\), and (iii) for each \(h\), \(\beta_c(s_c^{'}) (\text{diag}^1_{-c} | S_{-c}(h) \times T_{-c}) = 0\). So, each \(\beta_c(s_c^{'})\) has beliefs that are concentrated on 1-strategy-type pairs, but off the diagonal.

\(^9\)The proof treats the case of \(Q^1_{-c} = [s_{-c}]\) differently. There, by genericity, we can choose \(j_c(s_c^{'})\) so that it does not strongly believe \(Q^1_{-c}\).
Observe that, under the construction,

$$R^1_c(T) = \bigcup_{i=1,2} \bigcup_{s^i_c \in Q^i_c} (\mathbb{BR}[j_c(s^i_c)] \times \{s^i_c\}) = \text{diag}^1 \cup \bigcup_{s^2_c \in Q^2_c} (\mathbb{BR}[j_c(s^2_c)] \times \{s^2_c\}).$$

Since the same holds for the other player \(-c\), the off-diagonal points of \(Q^1_{-c} \times Q^1_{-c}\) are irrational and the diagonal points of \(Q^1_{-c} \times Q^1_{-c}\) are rational. Thus, each 1-type \(s^1_c\) does not strongly believe \(R^1_{-c}(T)\) while each 2-type \(s^2_c\) strongly believes \(R^1_{-c}(T)\). As such,

$$R^2_c(T) = \bigcup_{s^2_c \in Q^2_c} (\mathbb{BR}[j_c(s^2_c)] \times \{s^2_c\}).$$

From this it follows that \(\text{proj}_{s_c} R^1_c(T) = Q^1_c\) and \(\text{proj}_{s_c} R^2_c(T) = Q^2_c\): By construction, \(Q^1_c \subseteq \text{proj}_{s_c} R^1_c(T)\) and \(Q^2_c \subseteq \text{proj}_{s_c} R^2_c(T)\). Moreover, each \(\mathbb{BR}[j_c(s^2_c)] \subseteq Q^2_c \subseteq Q^1_c\). So, \(\text{proj}_{s_c} R^1_c(T) \subseteq Q^1_c\) and \(\text{proj}_{s_c} R^2_c(T) \subseteq Q^2_c\).

### 7 Analyzing Games

Write \(\text{EFR}^m = \text{EFR}^m_a \times \text{EFR}^m_b\) for the set of strategies that survives \(m\) rounds of EFR. We have that \((S, \text{EFR}^1, \ldots, \text{EFR}^m)\) is an \(m\)-BRS. So, \(\text{EFR}^m\) will be contained in \(S^m\). However, the containment may be strict. In particular, there may be an \(m\)-BRS \((Q^0, Q^1, \ldots, Q^m)\) with \(\text{EFR}^m \cap Q^m = \emptyset\). This can already be seen from BoS with an Outside Option (Figure 1.1). There, for each \(m\), \((\text{Out}, R) \in S^m\); but \(R \notin \text{EFR}^3_b\) and \(\text{Out} \notin \text{EFR}^3_a\). This can arise because strong belief is non-monotonic. (See Examples 4.1-4.2.)

With Theorem 6.1 in mind, we turn to analyzing games via the \(m\)-BRS. We provide two examples: the 3-Repeated Prisoner’s Dilemma and the Centipede game (Rosenthal, 1981). Both have a somewhat subtle relationship to EFR. This, in turn, has implications for the interpretation of experimental data.

**Repeated Prisoner’s Dilemma**  Consider the Prisoner’s Dilemma in Figure 7.1, where \(\zeta > \kappa > \delta > \gamma\) and \(2\kappa > \zeta + \gamma\). For notational simplicity, we focus on the 3-repeated game. Corollary D.1 establishes that the game is generic. Thus, Theorem 6.1 applies and, so, we can focus on \(m\)-BRSs.

![Figure 7.1: Prisoner’s Dilemma](image)

Write \(V^t\) for the set of \(t\)-period nodes. Each non-terminal node is associated with a sequence of moves; for instance, \((\phi, (C, D))\) is the 2-period node that follows Ann playing \(C\) in the first period and Bob playing \(D\) in the first period. So, \(V^1 = \{\phi\}, V^2 = \{\phi\} \times \{(C, D)\}^2\), and \(V^3 = \{\phi\} \times \{(C, D)\}^4\). Since, at each node, all past moves are known, we can write a strategy of player \(c\) as a mapping \(s_c : V \rightarrow \{(C, D)\}\).
We begin with the EFR benchmark. Write $EFR_m^c = S_c$. Then, for each $m$ with $2 \geq m \geq 0$,

$$EFR_{c}^{m+1} = \{ s_c \in EFR_{c}^{m} : \text{ for each } v \in V_{3-m}^{c} \text{ with } s_c(v), s_c(v) = D \}.$$ 

(See Corollary D.1.) For each $m \geq 3$, $EFR_{c}^{m+1} = EFR_{c}^{m}$. So, if $s_c \in EFR_{c}^{1}\setminus EFR_{c}^{0}$, then $s_c$ plays $D$ at each 3-period node it allows, but it plays $C$ at some 2-period node it allows. Likewise, if $s_c \in EFR_{c}^{2}\setminus EFR_{c}^{1}$, then $s_c$ plays $D$ at each 2-period and 3-period node it allows, but it plays $C$ at the initial node $\phi$.

We know that, for each $m$, $EFR_{m} \subseteq S_{m}$. In fact, the two sets are equivalent for $m = 1$. However, this is not the case for $m \geq 2$. In particular:

**Proposition 7.1.**

(i) $S_{c}^{1} = EFR_{c}^{1}$.

(ii) $S_{c}^{2}$ is the set of strategies $s_c \in EFR_{c}^{1}$ so that

(a) $D \in s_c(\{(\phi, (s_c(\phi), C)), (\phi, (s_c(\phi), D))\})$, and

(b) if $2\delta > \beta + \gamma$, then $s_c(\phi) = C$ implies $s_c(\phi, (C, C)) = D$.

(iii) $S_{c}^{3}$ is the set of strategies $s_c \in S_{c}^{2}$ so that $s_c(\phi) = D$.

(iv) For each $m \geq 4$, $EFR_{m}^{c} = S_{c}^{3}$ with $s_c(\phi, (D, D)) = D$.

So, $S_{c}^{2}$ requires that a strategy $s_c$ play $D$ at some second-period node that it allows; but, it does not require that $s_c$ play $D$ at each second-period node. This, in turn, implies that a strategy in $S_{c}^{2}$ initially plays $D$. With this, $S_{c}^{1}$ is the set of strategies $s_c \in EFR_{c}^{1}$ that satisfy (i) $s_c(\phi) = D$, and (ii) $s_c(\phi, (D, D)) = D$. But, note, $EFR_{c}^{1} \subseteq S_{c}^{1}$: It may well be that $s_a(\phi, (D, C))$ (resp. $s_b(\phi, (C, D))$) is $C$.

An implication of Proposition 7.1 is that, for each $m$, $S_{m}$ is outcome equivalent to $EFR_{m}$. More precisely, (i) for each $(s_a, s_b) \in S_{m}$, there exists some $(r_a, r_b) \in EFR_{m}$ so that $\zeta(s_a, s_b) = \zeta(r_a, r_b)$, and (ii) for each $(s_a, s_b) \in EFR_{m}$, $(s_a, s_b) \in S_{m}$. However, for each $m \geq 2$, $EFR_{m} \subseteq S_{m}$.

The distinction between $S_{m}$ and $EFR_{m}$ has important implications for the inferences that can be made from experimental data. To understand why, focus on the case where $\beta + \gamma \geq 2\delta$. Consider an experimental dataset obtained by the direct response method, where we observe Ann play $C$ in the second-period. Conventional wisdom would suggest that this indicates Ann cannot both be rational and “reason” that Bob is rational. Indeed, this is the conclusion that EFR would suggest. However, this behavior is indeed consistent with R1SBR. (In fact, if in the first-period, Ann and Bob played $(D, C)$, then this behavior is consistent with RmSBR for all $m$.) Suppose, instead, that the experimental dataset is obtained by the strategy method and we observe Ann play $C$ at each (relevant) second-period node (i.e., we observe $s_a(\phi, (s_a(\phi), C)) = s_a(\phi, (s_a(\phi), D)) = C$). In that case, we can conclude that Ann’s behavior is indeed inconsistent with R1SBR.

It is, of course, well understood that a benefit of the strategy method is that it provides additional data that may not be observed from the direct response method. (See, e.g., the opening example in Brandts and Charness, 2011.) Typically this occurs because there is a choice of Ann that can only be observed if Bob plays a particular prior action; in the direct response method, Bob may not play the action and so we may not be able to observe Ann’s choice. The situation here is different: If we only use the direct response method and we observe Ann choose $D$ in the final period, then we cannot contradict the hypothesis that
“Ann is rational and strongly believes Bob is rational” independent of what Bob plays. However, if we use the strategy method, we can observe Ann’s behavior at each second-period node and so there is potentially observed behavior that would contradict this hypothesis.\(^\text{10}\)

**Centipede Game** Figure 7.2 depicts the Centipede game. We order the non-terminal nodes as \(v = 1, 2, \ldots, |V|\), where \(|V| \geq 3\). (So, \(v = 1\) indicates the initial node and \(v = |V|\) indicates the last non-terminal node.) If the game ends after \(\text{out}_v\) is played and \(v\) is odd (resp. even), then the payoffs are \((x + (v - 1)y, x + (v - 2)y)\) (resp. \((x + (v - 3)y, x + vy)\)), where \(x, y > 0\).\(^\text{11}\) If the game ends after \(\text{in}_{|V|}\) is played and \(|V|\) is odd (resp. even), then the payoffs are \((x + (|V| - 2)y, x + (|V| + 1)y)\) (resp. \((x + |V|y, x + (|V| - 1)y)\)). Figure 7.2 depicts \(|V|\) odd.

Write \(\ell\) for the last player and \(-\ell\) for the second-to-last player. If \(|V|\) is odd then \(\ell\) is Ann and \(-\ell\) is Bob; if \(|V|\) is even then \(\ell\) is Bob and \(-\ell\) is Ann. Let \([\text{out}, v]\)\(_c\) be the set of strategies of player \(c\) that allow \(v\) and then play \(\text{out}_v\). Likewise, let \([\text{in}]_c\) be the set that contains the (unique) strategy of player \(c\) that specifies \(\text{in}_v\) at every node \(v\).

Let us point to four (related) features of the game: First, the player who moves at node \(v \leq |V| - 1\) strictly prefers \(\text{out}_{v+2}\) (resp. \(\text{in}_{|V|}\) if \(v = |V| - 1\)) to \(\text{out}_v\) and strictly prefers \(\text{out}_v\) to \(\text{out}_{v+1}\). Second, the player who moves at node \(|V|\), strictly prefers \(\text{out}_{|V|}\) to \(\text{in}_{|V|}\). Third, for each \(v \leq |V| - 1\) (resp. \(v = |V|\)), \([\text{out}, v]\)\(_c\) is the set of best responses under a CPS that strongly believes \([\text{out}, v + 1]\)\(_c\) (resp. \([\text{in}]\)\(_c\)). Fourth, the game is generic. (This follows from features two and three.) Thus, Theorem 6.1 applies and, so, we can focus on \(m\)-BRSs.\(^\text{12}\)

A useful benchmark will be \(m\) rounds of EFR. Observe that

\[
\text{EFR}^1_{\ell} \times \text{EFR}^1_{-\ell} = (S_{\ell}\setminus[\text{in}]_{\ell}) \times S_{-\ell} \quad \text{and} \quad \text{EFR}^2_{\ell} \times \text{EFR}^2_{-\ell} = \text{EFR}^1_{\ell} \times (S_{-\ell}\setminus[\text{in}]_{-\ell}).
\]

Moreover,

\[
\text{EFR}^m_{\ell} \times \text{EFR}^m_{-\ell} = \begin{cases} 
(\text{EFR}^{m-1}_{\ell}\setminus[\text{out}, |V| + 3 - m]\_\ell) \times \text{EFR}^{m-1}_{-\ell} & \text{if } m = 3, \ldots, |V| \text{ is odd}, \\
\text{EFR}^{m-1}_{\ell} \times (\text{EFR}^{m-1}_{\ell}\setminus[\text{out}, |V| + 3 - m]\_\ell) & \text{if } m = 4, \ldots, |V| \text{ is even.}
\end{cases}
\]

For all \(m \geq |V| + 1\), \(\text{EFR}^m_{\ell} \times \text{EFR}^m_{-\ell} = \text{EFR}^{|V|}_{\ell} \times \text{EFR}^{|V|}_{-\ell}\). Note, this also corresponds round-for-round with the backward-induction algorithm.

\(^\text{10}\)This, of course, presumes that the strategy method works, in that a strategy is optimal in the strategy-method game if and only if it is optimal in the direct-response game. Siniscalchi (2020) provides foundations for the strategy method.

\(^\text{11}\)The first component in the payoff vector is Ann’s payoffs.

\(^\text{12}\)Reny’s (1993) TOL game also satisfies these properties. Anything we say about Centipede also applies to that game.
Unlike EFR, the \( m \)-BRS procedure has very different implications for the first mover (Ann) and the second mover (Bob).

**Proposition 7.2.** In the Centipede game, the following hold for each finite \( m \geq 1 \):

(i) \( S^m_a = \text{EFR}^m_a \).

(ii) If \( |V| \) is odd, then \( S^m_b = S_b \). If \( |V| \) is even, then \( S^m_b = (S_b \setminus \text{in})_b \).

At first glance, part (i) of Proposition 7.2 may appear trivial: For each \( m \), \( \text{EFR}^m_a \times \text{EFR}^m_b \) is consistent with an \( m \)-BRS. Thus, \( \text{EFR}^m_a \subseteq S^m_a \). However, the key is showing that \( S^m_a \subseteq \text{EFR}^m_a \) and, as we have seen, this is not the case for the second mover Bob. (Appendix D.2 explains why this is the case.)

Proposition 7.2 points to a distinction between the first mover and second mover—one that is important for interpreting experimental data. Consider the case where \( |V| \) is odd, so that the first mover is the last mover. First, for each \( m \), \( S^m_a = \text{EFR}^m_a \). So, for instance, if we observe the first mover play \( \text{out}_{|V| - 2} \), then we can conclude that the first mover’s behavior is consistent with R3SBR but inconsistent with R4SBR. Second, for each \( m \), \( S^m_b = S_b \). So, in particular, any strategy that we observe the second mover play is consistent with RmSBR for each \( m \). (In fact, any strategy that we observe the second mover play is consistent with RCSBR.) This contradicts the conventional wisdom that, observing the second player choose in indicates that the second player exhibits some form of “bounded reasoning about rationality.”

To understand the difference between the first mover and second mover, note that—when \( |V| \) is odd—\( \text{out}_{1} \times S_b \) is an EFBRS. Thus, \( S_b \subseteq S^\infty_b \). But, for any non-empty EFBRS \( Q_a \times Q_b \), we have \( Q_a = \text{out}_{1} \). Thus, \( S^\infty_a = \text{out}_{1} \).

### 8 Discussion

#### 8.a Generic Games

Theorem 6.1 shows that, in generic games, \( m \)-BRS’s characterize the \( R(m-1) \) SBR sequences. This raises the question, which games are generic? We begin with NRT:

**Definition 8.1 (Battigalli, 1997).** A game satisfies no relevant ties (NRT) if \( \pi_c(s_c, s_{-c}) = \pi_c(r_c, s_{-c}) \) implies \( \zeta(s_c, s_{-c}) = \zeta(r_c, s_{-c}) \).

A game satisfies no relevant ties if, whenever player \( c \) is decisive over two distinct terminal nodes \( z \) and \( z^* \) (i.e., if there exists \( (s_c, s_{-c}) \) and \( (r_c, s_{-c}) \) with \( \zeta(s_c, s_{-c}) \neq \zeta(r_c, s_{-c}) \)), she is not indifferent between those terminal nodes.

We may have a game that satisfies NRT that is non-generic. (See Example E.1.) However, there is a subclass of NRT games that are generic—ones in which a strategy that is a best response under some CPS is a best response under a “degenerate CPS.” (See Definition E.2 and Proposition E.1.) Perfect-information games satisfy that condition and, so, a perfect-information game satisfying NRT is generic. (See Proposition E.2.)

There is a related condition that ensures genericity. Fix some \( X_{-c} \subseteq S_{-c} \) and some information set \( h \in H_c \) with \( s_c \in S_c(h) \). Say \( r_c \) supports \( s_c \) with respect to \( (X_{-c}, h) \) if there exists \( s \in \mathcal{P}(S_c(h)) \) with

1. \( \sigma(r_c) > 0 \), and
2. for all \( s_{-c} \in X_{-c} \cap S_{-c}(h) \), \( \sum_{s_{-c} \in S_c(h)} \pi_c(r_c, s_{-c})\sigma(r_c) = \pi_c(s_c, s_{-c}) \).

If \( s_c \notin S_c(h) \), then no \( r_c \) supports \( s_c \) with respect to \( (X_{-c}, h) \).

---

13Suppose not. Then there is some \( (s_a, s_b) \in Q_a \times Q_b \) where Ann plays \( i_{1} \). Consider the strategy profile that results in the maximum path of play of in. Specifically, it results in \( i_{1} \cdot \ldots \cdot i_{n} \) and the player, viz. \( v \), who moves at \( v + 1 \) plays \( \text{out}_{v+1} \). Let \( s_c \) be a strategy in \( Q_c \) that plays in up to and including \( v \). Any CPS that strongly believes \( Q_{-c} \) must, at \( v \), assign probability 1 to \( -c \) playing \( \text{out}_{v+1} \). Thus, \( s_c \) cannot be a sequential best response under any CPS that strongly believes \( Q_{-c} \).
Definition 8.2. A game satisfies no relevant convexities (NRC) if, for each $h \in H_c$, the following holds: If $s_c \in S_c(h)$ and $r_c$ supports $s_c$ with respect to some $(X_{-c}, h)$, then $\zeta(s_c, s_{-c}) = \zeta(r_c, s_{-c})$ for each $s_{-c} \in X_{-c} \cap S_{-c}(h)$.

Informally, a game satisfies NRC if strategies in the support of a mixture $\sigma \in P(S_c)$ induce the same path of play as $s_c$, whenever player $c$ is indifferent between $\sigma$ and $s_c$. Corollary E.1 (Appendix E.2) establishes that a game that satisfies NRC is generic.

8.b Termination of the $m$-BRS Procedure  Fix a decreasing sequence of strategies $(Q^0, Q^1, Q^2, \ldots)$, where each $(Q^0, \ldots, Q^m)$ forms an $m$-BRS. Since each $Q^{m+1} \subseteq Q^m$, $(Q^0, Q^1, Q^2, \ldots)$ defines an iterative elimination procedure. We refer to this as an $m$-BRS elimination procedure. Note, there may be many such elimination procedures, corresponding to distinct $(Q^0, Q^1, Q^2, \ldots)$ and $(Q^0, \hat{Q}^1, \hat{Q}^2, \ldots)$.

Because the strategy set is finite, this elimination procedure must terminate—i.e., there exists some $M$ so that, for each $m \geq M$, $Q^m = Q^M$. If the researcher knew at which $M$ this occurred, they could use that fact to determine that the elimination procedure has stopped.

At first glance, there may appear to be straightforward route to determine $M$. Typically, an elimination procedure stops shrinking at the first round where no strategy is eliminated for either player. However, this same principle does not apply to the $m$-BRS elimination procedure. We may have $Q^{m+1} \subseteq Q^m = Q^{m-1}$. To see this, refer to the simultaneous-move game given by Figure 8.1. For each $m$, there is an $m$-BRS with $(Q^0, \ldots, Q^m)$, so that (i) for each $n \leq m$, $Q^n = \{U, D\} \times \{L, R\}$, and (ii) $Q^{m+1} = \{U, D\} \times \{R\}$. Thus, the $(m+1)$-BRS procedure has no shrinkage up until round $m$, but a shrinkage at round $(m+1)$. Since $m$ can be any number, we can have arbitrarily long pauses before shrinkage. To understand why this can occur, note that we can have $R^{m+1}(T) \subseteq R^m(T)$ but $\text{proj}_S R^{m+1}(T) = \text{proj}_S R^m(T)$. When this happens, it may well be that $\text{proj}_S R^{m+2}(T) \subsetneq \text{proj}_S R^{m+1}(T)$.

![Figure 8.1](image_url)

Despite this, we can provide a bound on the elimination procedure $(S^0, S^1, S^2, \ldots)$, i.e., we can find some $M$ so that, for all $m \geq M$, $S^m = S^M$. To understand why, consider an $m$-BRS procedure $(Q^0, Q^1, Q^2, \ldots)$ with a pause at round $m$, i.e., $Q^{m+1} = Q^m$ but $Q^{m+2} \subseteq Q^{m+1}$. The key is that any eliminated strategy—i.e., any strategy in $Q^{m+1}\setminus Q^{m+2}$—must be contained in $S^{m+2}$. That is, there must exist some other $m$-BRS procedure $(\hat{Q}^0, \hat{Q}^1, \hat{Q}^2, \ldots)$ so that $Q^{m+1}\setminus Q^{m+2} \subseteq \hat{Q}^{m+2}$. This follows from the following:

Observation 8.1. If $(Q^0, \ldots, Q^m)$ is an $m$-BRS with $Q^m = Q^{m-1}$, then $Q^{m-1}$ is an EFBRS.

Fix some $(Q^0, Q^1, Q^2, \ldots)$ where (i) for each $m$, $(Q^0, \ldots, Q^m)$ is an $m$-BRS, and (ii) $Q^{n+1} = Q^n$. Then $Q^n$ is an EFBRS. We can define a new sequence $(\hat{Q}^0, \hat{Q}^1, \hat{Q}^2, \ldots)$ so that (i) for each $m \leq n$, $\hat{Q}^m = Q^m$, and (ii) for each $m > n$, $\hat{Q}^m = Q^n$. Then, for each $m$, $(\hat{Q}^0, \ldots, \hat{Q}^m)$ is an $m$-BRS. So, for each $m$, $Q^m \subseteq \hat{Q}^m \subseteq S^m$. From this, the following termination result follows:

---

14Importantly, an $m$-BRS is not an order of elimination of iterated conditional dominance (Chen and Micali, 2012).
Proposition 8.1. Set
\[ M = \begin{cases} 
2 \min\{|S_a|, |S_b|\} - 1 & \text{if } |S_a| \neq |S_b|, \\
2 \min\{|S_a|, |S_b|\} - 2 & \text{if } |S_a| = |S_b|. 
\end{cases} \]

Then, for all \( m \geq M \), \( S^m = S^\infty \).

Proposition 8.1 provides a bound \( M \) for the procedure \((S^0, S^1, S^2, \ldots)\). Thus, it suffices to compute all the \( M \)-BRSs, \((Q^0, \ldots, Q^M)\).

8.c Computing an \( m \)-BRS: NRC Games Games that satisfy NRC have a simple characterization of the \( m \)-BRS concept, one that allows for a “simpler” computation of the \( m \)-BRS concept.

Proposition 8.2. Suppose \( \Gamma \) satisfies NRC. Then \((Q^0, \ldots, Q^m)\) forms an \( m \)-BRS if and only if \( Q^1 \) is non-empty and, for each \( n = 1, \ldots, m \), the following holds: For each \( s_c \in Q^n_c \), there exists \( p_c \in \mathcal{A}(S_c, E_c) \) with \( s_c \in \text{BR}[p_c] \) and \( p_c \) strongly believes \( Q^0_c, \ldots, Q^n_{c-1} \).

Fix a decreasing sequence of product sets \((Q^0, \ldots, Q^{m-1}, Q^m)\) so that \((Q^0, \ldots, Q^{m-1})\) is an \((m-1)\)-BRS. We seek to determine whether \((Q^0, \ldots, Q^m)\) is also an \( m \)-BRS. If the game satisfies NRC, Proposition 8.2 provides two ways that simplify making that determination. First, we can replace CPS’s with arrays. Second, we can eliminate the maximality criterion. Appendix E.5 explains why each of these simplifications may fail absent NRC. It also shows that, when we reduce the definition in this way, repeated application of a simplex algorithm determines whether or not \((Q^0, \ldots, Q^m)\) is an \( m \)-BRS.

8.d Beyond Generic Games It would be desirable to have a procedure that determines the sets \( S^m \) in non-generic games. One possibility would be to amend the definition of an \( m \)-BRS. In light of Example 6.2 one might suggest the following: If \( s_a \in Q^1_a \setminus Q^2_a \), then there exists some CPS \( p_a \) that satisfies conditions (BRP.1)-(BRP.2)-(BRP.3) and does not strongly believe \( Q^1_b \). However, under that amendment, we lose an analog of Proposition 6.1: For a given \( \mathcal{T} \), \((\text{proj}_S R^n(\mathcal{T}), \ldots, \text{proj}_S R^m(\mathcal{T}))\) may fail the new property.

Example 6.2 illustrates that, for a given \( m \)-BRS \((Q^1, Q^2, \ldots, Q^m)\), there may be no \( \mathcal{T} \) so that \( Q^n \subseteq \text{proj}_S R^n(\mathcal{T}) \) for each \( n = 1, \ldots, m \). The example leaves open that there may be an alternate \( m \)-BRS \((\hat{Q}^1, \hat{Q}^2, \ldots, \hat{Q}^m)\) so that the following holds: (i) \( \hat{Q}^m = Q^m \), and (ii) there exists some type structure \( \mathcal{T} \) so
that \( \hat{Q}^n \subseteq \text{proj}_g R^n(T) \) for each \( n = 1, \ldots, m \). If correct, it would say that Equation (2) does hold for all games. We neither know this to be true nor have a counterexample.

8.e Finite-Order Type Structures Type structures induce hierarchies of conditional beliefs about the strategies played—i.e., \( m^{th} \)-order beliefs for all \( m \). This suggests that players can contemplate sentences of the form “I think that you think that I think . . .” As such, one might incorrectly hypothesize that our analysis requires that players have an unlimited ability to engage in interactive reasoning, despite the fact that they exhibit “bounded reasoning about rationality” (formalized as \( Rm \text{SBR} \), but not \( R(m + 1) \text{SBR} \)). However, there is no such requirement. The key observation is that hierarchies of beliefs beyond level \( m \) do not affect \( R(m - 1) \text{SBR} \). Formally, consider two types \( t_a \) and \( u_a \) with the same \( m^{th} \)-order beliefs about the strategies played. For any strategy \( s_a \), the strategy-type pair \( (s_a, t_a) \) is consistent with \( R(m - 1) \text{SBR} \) if and only if \( (s_a, u_a) \) is consistent with \( R(m - 1) \text{SBR} \). Thus, the higher-order beliefs become a consequence of our formalism and do not have implications for our characterization result. In particular, we could instead adapt the finite-order type structure approach proposed in Kets (2010) and Heifetz and Kets (2018)—amended for CPS’s—and apply \( R(m - 1) \text{SBR} \) in that framework. We would reach analogous conclusions. (Appendix A in Heifetz and Kets, 2018 makes a similar point, in a different context.)

8.f HK-Rationalizability Heifetz and Kets (2018) define a notion of rationalizability for finite-order settings. It is quite different from the analysis here. They focus on simultaneous-move games of incomplete information and use a finite-order type structure to model incomplete information. Strategic uncertainty is captured implicitly by their rationalizability concept, which we call HK-rationalizability. Importantly, their notion of rationalizability is different from the standard notion (to which ours reduces to in simultaneous-move games). If we apply their concept to a game of complete information, a strategy may survive two rounds of HK-rationalizability even though it does not survive two rounds of Bernheim (1984) and Pearce (1984) rationalizability. This arises for the same reason that a level-2 type in the cognitive hierarchy model assigns positive probability to a level-0 type and, so, may assign positive probability to a dominated strategy.\(^{15}\)

Appendix A Preliminaries

This appendix provides preliminary results, which are used in subsequent results.

Marginalization Property of Belief

Lemma A.1. If \( \beta_c(t_c) \) strongly believes \( S_{-c} \subseteq S_{-c} \times T_{-c} \), then marg \( S_{-c} \beta_c(t_c) \) strongly believes \( \text{proj}_{S_{-c}} E_{-c} \).

Proof. Suppose \( \beta_c(t_c) \) strongly believes \( S_{-c} \subseteq S_{-c} \times T_{-c} \). Fix some \( S_{-c}(h) \times T_{-c} \in E_{c} \otimes T_{-c} \). If \( \text{proj}_{S_{-c}} E_{-c} \cap S_{-c}(h) \neq \emptyset \), then there exists \( (s_{-c}, t_{-c}) \in E_{-c} \) so that \( s_{-c} \in S_{-c}(h) \). It follows that \( E_{-c} \cap (S_{-c}(h) \times T_{-c}) \neq \emptyset \) and so \( \beta_c(E_{-c}|S_{-c}(h) \times T_{-c}) = 1 \). Now note that

\[
\text{marg}_{S_{-c}} \beta_c(\text{proj}_{S_{-c}} E_{-c}|S_{-c}(h)) = \beta_c(\text{proj}_{S_{-c}} E_{-c} \times T_{-c}|S_{-c}(h) \times T_{-c}) \geq \beta_c(E_{-c}|S_{-c}(h) \times T_{-c}) = 1.
\]

\(^{15}\)This does not occur in the level-\( k \) model (Costa-Gomes, Crawford and Broseta, 2001) and, as such, a level-2 type in the level-\( k \) model does survive two rounds of rationalizability.

22
It follows that \( \text{marg}_{S_{-c}} \beta_c(\text{proj}_{S_{-c}} E_c|S_{-c}(h) \times T_{-c}) = 1 \), as desired. ■

**Image CPS’s:** Fix a CPS \( p_c \in \mathcal{C}(S_{-c}, E_c) \) and a measurable mapping \( \tau_{-c} : S_{-c} \rightarrow S_{-c} \times T_{-c} \). Define \( q_c \) as follows: For each conditioning event \( S_{-c}(h) \times T_{-c} \in E_c \otimes T_{-c} \) and each Borel \( E_{-c} \subseteq S_{-c} \times T_{-c} \), set

\[
q_c(E_{-c}|S_{-c}(h) \times T_{-c}) = p_c((\tau_{-c})^{-1}(E_{-c})|S_{-c}(h))
\]

We refer to \( q_c \) at the **image CPS of \( p_c \)** under \( \tau_{-c} \). So defined, \( q_c \) is indeed a CPS. See Battigalli, Friedenberg and Siniscalchi (2012, Part III, Chapter 4). Moreover, if \( \tau_{-c}(s_{-c}) \in \{s_{-c}\} \times T_{-c} \) for each \( s_{-c} \), then the image CPS of \( p_c \) under \( \tau_{-c} \), viz. \( q_c \), has marg \( \text{marg}_{S_{-c}} q_c = p_c \). As a consequence, for any given CPS \( p_c \in \mathcal{C}(S_{-c}, E_c) \), we can find some CPS \( q_c \in \mathcal{C}(S_{-c} \times T_{-c}, E_c \otimes T_{-c}) \) so that \( \text{marg}_{S_{-c}} q_c = p_c \).

**Structure of Games and Sequential Best Responses** By perfect recall, we have the following:

(i) For each \( h, h' \in H_c \), either \( S(h) \subseteq S(h') \), \( S(h') \subseteq S(h) \), or \( S(h) \cap S(h') = \emptyset \). (ii) For each \( h \in H_c \), \( S(h) = S_c(h) \times S_{-c}(h) \). The second of these implies the following:

**Lemma A.2.** Fix \( h, h' \in H_c \) so that \( S(h) \cap S(h') = \emptyset \). If \( S_{-c}(h) \cap S_{-c}(h') \neq \emptyset \), then \( S_c(h) \cap S_c(h') = \emptyset \).

**Proof.** Fix \( h, h' \in H_c \) so that \( S_c(h) \cap S_c(h') \neq \emptyset \) and \( S_{-c}(h) \cap S_{-c}(h') \neq \emptyset \). Then there exists \( s_c \in S_c(h) \cap S_c(h') \) and \( s_{-c} \in S_{-c}(h) \cap S_{-c}(h') \). It follows that \( (s_c, s_{-c}) \in S_c(h) \times S_{-c}(h) \) and \( (s_c, s_{-c}) \in S_c(h') \times S_{-c}(h') \).

By perfect recall, \( S(h) = S_c(h) \times S_{-c}(h) \) and \( S(h') = S_c(h') \times S_{-c}(h') \). Thus, \( S(h) \cap S(h') \neq \emptyset \).

**Lemma A.3.** Fix \( h^*, h^{**} \in H_c \) so that \( S(h^{**}) \subseteq S(h^*) \). Let \( \mu_c \in \mathcal{P}(S_{-c}) \) with \( \mu_c(S_{-c}(h^*)) = 1 \) and \( \mu_c(S_{-c}(h^{**})) > 0 \). If \( s_c \in S_c(h^{**}) \) is optimal under \( \mu_c \) given all strategies in \( S_c(h^*) \), then \( s_c \) is optimal under \( \mu_c(\cdot|S_{-c}(h^{**})) \) given all strategies in \( S_c(h^{**}) \).

**Proof.** Suppose that there exists some \( r_c \in S_c(h^{**}) \) so that

\[
\sum_{s_{-c}} [\pi_c(r_c, s_{-c}) - \pi_c(s_c, s_{-c})]\mu_c(s_{-c}|S_{-c}(h^{**})) > 0.
\]

Construct a strategy \( \tilde{r}_c \) so that

\[
\tilde{r}_c(h) = \begin{cases} 
  r_c(h) & \text{if } S(h) \subseteq S(h^{**}) \\
  s_c(h) & \text{otherwise.}
\end{cases}
\]

Fix some \( s_{-c} \in S_{-c}(h^{**}) \) and observe that \( (s_c, s_{-c}) \) and \( (r_c, s_{-c}) \) are both contained in \( S(h^{**}) = S_c(h^{**}) \times S_{-c}(h^{**}) \). (This follows from perfect recall.) Thus, \( (\tilde{r}_c, s_{-c}) \in S(h^{**}) \) and \( \tilde{r}_c \in S_c(h^{**}) \subseteq S_c(h^*) \).

We will show that

(i) \( \zeta(r_c, s_{-c}) = \zeta(\tilde{r}_c, s_{-c}) \) if \( s_{-c} \in S_{-c}(h^{**}) \), and

(ii) \( \zeta(s_c, s_{-c}) = \zeta(\tilde{r}_c, s_{-c}) \) if \( s_{-c} \in S_{-c}(h^*) \setminus S_{-c}(h^{**}) \).

From this, it follows that

\[
\sum_{s_{-c}} [\pi_c(\tilde{r}_c, s_{-c}) - \pi_c(s_c, s_{-c})]\mu_c(s_{-c}) > 0,
\]

contradicting the hypothesis that \( s_c \) is optimal under \( \mu_c \) given all strategies in \( S_c(h^*) \).
First, fix some \( s_{-c} \in S_{-c}(h^*) \) and note that, by perfect recall,
\[
(s_c, s_{-c}), (r_c, s_{-c}), (\tilde{r}_c, s_{-c}) \in S_c(h^*) \times S_{-c}(h^*) = S(h^*).
\]

Suppose, contra hypothesis, that \( \zeta(r_c, s_{-c}) \neq \zeta(\tilde{r}_c, s_{-c}) \). Then there exists some \( h \in H_c \) so that \((r_c, s_{-c}), (\tilde{r}_c, s_{-c}) \in S(h) = S_c(h) \times S_{-c}(h) \) but \( r_c(h) \neq \tilde{r}_c(h) = s_c(h) \). By construction, \( \neg S(h) \subseteq S(h^*) \). Since \( S(h^*) \cap S(h) \neq \emptyset \), it follows that \( S(h^*) \subseteq S(h) \). Thus we have established that \( s_c(h) = r_c(h) \) and \((s_c, s_{-c}), (r_c, s_{-c}) \in S(h^*) \); but this contradicts perfect recall.

Second, fix some \( s_{-c} \in S_{-c}(h^*) \backslash S_{-c}(h^*) \) and suppose, contra hypothesis, that \( \zeta(s_c, s_{-c}) \neq \zeta(\tilde{r}_c, s_{-c}) \). Then there exists some \( h \in H_c \) with \((s_c, s_{-c}), (\tilde{r}_c, s_{-c}) \in S(h) = S_c(h) \times S_{-c}(h) \) and \( s_c(h) \neq \tilde{r}_c(h) = r_c(h) \). By construction, \( S(h) \subseteq S(h^*) \), contradicting the assumption that \( s_{-c} \in S_{-c}(h^*) \backslash S_{-c}(h^*) \).

**Appendix B  Proofs of Propositions 6.1-6.2**

**Proof of Proposition 6.1.** The proof is by induction on \( m \).

\( m = 1 \): If \( s_c \in \text{proj}_S R^1_c(T) \), then there exists some \( t_c \in T_c \) so that \((s_c, t_c) \in R^1_c(T) \). Take \( p_c = \text{marg}_S \beta_c(t_c) \). Note that \( s_c \in \mathbb{BR}[p_c] \). Moreover, if \( r_c \in \mathbb{BR}[p_c] \), then \((r_c, t_c) \in R^1_c(T) \) and so \( r_c \in \text{proj}_S R^1_c(T) \).

\( m \geq 2 \): Assume the claim holds for \( m \) and fix some \((\text{proj}_S R^0(T), \ldots, \text{proj}_S R^m(T), \text{proj}_S R^{m+1}(T)) \). Then, by the induction hypothesis, \((\text{proj}_S R^0(T), \ldots, \text{proj}_S R^m(T)) \) forms an \( m \)-BRS. Thus, it suffices to show that \( \text{proj}_S R^{m+1}(T) = \text{proj}_S R^{m+1}(T) \times \text{proj}_S R^{m+1}(T) \) satisfies the extensive-form best response property relative to \((\text{proj}_S R^0(T), \ldots, \text{proj}_S R^m(T)) \).

Fix some \( s_c \in \text{proj}_S R^m(T) \). There exists some \( t_c \in T_c \) so that \((s_c, t_c) \in R^m_c(T) \). Take \( p_c = \text{marg}_S \beta_c(t_c) \). Since \((s_c, t_c) \in R^m_c(T) \), \( s_c \in \mathbb{BR}[p_c] \). Moreover, \( \beta_c(t_c) \) strongly believes \( R^m_c(T) \). So applying Lemma A.1, \( \text{marg}_S \beta_c(t_c) \) strongly believes \( \text{proj}_S R^0_c(T), \ldots, \text{proj}_S R^m_c(T) \). Finally, if \( r_c \in \mathbb{BR}[p_c] \), then \((r_c, t_c) \in R^{m+1}_c(T) \) and so \( r_c \in \text{proj}_S R^{m+1}_c(T) \).

**Proof of Proposition 6.2(i).** Fix a 1-BRS \((Q^0, Q^1)\). Construct \( T \) as follows: Set \( T_c = Q^1_c \). For each \( s_c \in T_c = Q^1_c \), choose \( \beta_c(s_c) \) so that \( \text{marg}_S \beta_c(s_c) \) is a CPS \( p_c \) with \([s_c] \in \mathbb{BR}[p_c] \subseteq Q^1_c \). (The fact that such a CPS exists follows from the definition of a 1-BRS.) It follows that \( \text{proj}_{S_c} R^1_c(T) = Q^1_c \).

**Proof of Proposition 6.2(ii).** Fix a 2-BRS \((Q^0, Q^1, Q^2)\). For each \( s_c \in Q^1_c \), there exists some CPS \( j_c(s_c) \) so that \( s_c \in \mathbb{BR}[j_c(s_c)] \subseteq Q^1_c \). Moreover, if \( s_c \in Q^2_c \), we can take \( j_c(s_c) \) to strongly believe \( Q^1_c \) and so that \( \mathbb{BR}[j_c(s_c)] \subseteq Q^2_c \).

With this in mind, set \( T_c = Q^1_c \) and define \( \beta_c(s_c) \) so that \( \text{marg}_{S_{-c}} \beta_c(s_c) = j_c(s_c) \). Moreover, for each \( h \) with \( S_{-c}(h) \cap Q^1_c \neq \emptyset \) and each \( s_{-c} \in S_{-c}(h) \cap Q^1_{-c} \), set
\[
\beta_c(s_c)((s_{-c}, s_{-c})|S_{-c}(h) \times T_{-i}) = j_c(s_c)(s_{-c} | S_{-c}(h)).
\]

Then,
\[
R^1_c(T) = \bigcup_{s_c \in Q^1_c} (\mathbb{BR}[j_c(s_c)] \times \{s_c\}) \implies \text{proj}_{S_c} R^1_c(T) = Q^1_c.
\]
Moreover, if \( s_c \in Q^c \), type \( s_c \) strongly believes \( R^1_{-c}(T) \). So, \( Q^c \subseteq \text{proj}_{s_c}R^2(T) \). \( \blacksquare \)

**Appendix C  Proof of Theorem 6.1**

To show Theorem 6.1, it will be useful to introduce a strong justification property. With this in mind, refer to a set \( X_c \subseteq Q_c \) as an **effective singleton** if there exists some \( s_c \) so that \( X_c = [s_c] \). If \( X_c \subseteq Q_c \) is not effectively a singleton, then we simply say it is **non-singleton**.

**Definition C.1.** Fix an \( m \)-BRS \((Q^0, \ldots, Q^m)\). Say that the \( m \)-BRS satisfies the **strong justification property** if, for each player \( c \) and each \( n = 1, \ldots, m \), we can find a mappings \( j^n_c : Q^n_c \rightarrow C(S_{-c}, E_c) \) satisfying the following criteria:

(j.a) For each \( s_c \in Q^1_c \), \( \mathbb{B}[j^1_c(s_c)] = [s_c] \). Moreover, if \( Q^1_c \) if effectively a singleton, then \( j^1_c(s_c) \) does not strongly believe \( Q^1_c \).

(j.b) For each \( n = 2, \ldots, m \) and each \( s_c \in Q^n_c \), \( s_c \in \mathbb{B}[j^n_c(s_c)] \subseteq Q^n_c \) and \( j^n_c(s_c) \) strongly believes \( Q^{n-1}_c \).

Observe that, by definition of an \( m \)-BRS, we can always find mappings \( j^n_c : Q^n_c \rightarrow C(Q_{-c}, E_c) \) satisfying condition (j.a). But, condition (j.a) is stronger than that required by an \( m \)-BRS. If we find mappings \( j_c = (j^1_c, \ldots, j^m_c) \) satisfying these requirements, we say that \( j_c \) strongly justifies the \( m \)-BRS for player \( c \) or \( j_a \) and \( j_b \) strongly justify the \( m \)-BRS. Theorem 6.1 follows from the following two propositions.

**Proposition C.1.** Fix an \( m \)-BRS \((Q^0, \ldots, Q^m)\) satisfying the strong justification property. Then there exists a type structure \( T \) so that, for each \( n = 1, \ldots, m \), \( \text{proj}_{s}R^n(T) = Q^n \).

**Proposition C.2.** If the game is generic, then any \( m \)-BRS satisfies the strong justification property.

We now turn to proving these two results.

**C.1 Proof of Proposition C.1**

Throughout we fix an \( m \)-BRS \((Q^0, \ldots, Q^m)\) satisfying the strong justification property. Thus, for each player \( c \), there are mappings \( j_c = (j^1_c, \ldots, j^m_c) \) that strongly justify the \( m \)-BRS.

**Description of the Type Structure** For each player \( c \) and each \( n = 1, \ldots, m \), set \( U^n_c \equiv Q^n_c \) and write \( v^n_c : Q^n_c \rightarrow U^n_c \) for the identity map. The type set for player \( c \) will be \( T_c = \bigcup_{n=1}^m U^n_c \). We will refer to types in \( U^n_c \) as the \( n \)-types for player \( c \).

It will be convenient to specify the **diagonal** of \( Q^n_c \times U^n_c \). This will be given by

\[
\text{diag}^n_c = \bigcup_{s_c \in Q^n_c} ([s_c] \times v^n_c([s_c]))
\]

Observe that, if \( [s_c] = [r_c] \) then \( v^n_c([s_c]) = v^n_c([r_c]) \) and so \( [s_c] \times v^n_c([r_c]) \subseteq \text{diag}^n_c \). Moreover, if \( Q^n_c \) is non-singleton then, for each \( s_c \in Q^n_c \), there exists a type \( t_c \in U^n_c \) so that \( (s_c, t_c) \in (Q^n_c \times U^n_c) \) \( \text{diag}^n_c \).

For each \( n = 1, \ldots, m \), we define a mapping \( \tau^n_{-c} : S_{-c} \rightarrow S_{-c} \times T_{-c} \) with \( \tau^n_{-c}(s_{-c}) \in \{s_{-c}\} \times T_{-c} \). In addition, the mappings will satisfy the following: For \( n = 1 \), if \( Q^1_{-c} \) is non-singleton, then the range of \( \tau^1_{-c} \) is concentrated on \( S_{-c} \times U^1_{-c} \) but off of \( \text{diag}^1_{-c} \), i.e., each \( \tau^1_{-c}(s_{-c}) \in (S_{-c} \times U^1_{-c}) \setminus \text{diag}^1_{-c} \). For \( n = 2, \ldots, m \),
for each \( s_c \in Q^c \), \( \tau^n_c(s_c) \) is in the maximal diagonal \((\leq n-1)\) consistent with \( s_c \). Specifically, for a given \( s_c \in Q^c \), let \( \ell = \max\{k = 1, \ldots, n-1 : s_c \in Q^k \} \) and set \( \tau^n_c(s_c) = (s_c, v^n_{\ell}(s_c)) \).

The belief map is such that, for each \( v^n_c(s_c) \in U^n \), \( \beta(\psi^n(s_c)) \) is the image CPS of \( j^n_c(s_c) \) under \( \tau^n_c \). Observe that, for each \( s_c \in Q^c \), \( \max_{s_c \in Q^c} Q^{-1}_{\beta}(v^n_c(s_c)) = j^n_c(s_c) \).

**Analysis** It will be convenient to define sets of \( n \)-strategy-type pairs of the players. In particular, for each player \( c \) and each \( n = 1, \ldots, m \), set

\[
Q^n_c = \bigcup_{s_c \in Q^c} (\mathbb{R}[\tau^n_c(s_c)] \times \{v^n_c(s_c)\})
\]

By Conditions (j.a)-(j.b) of strong justification, \( \text{diag}^n_c \subseteq Q^n_c \).

**Lemma C.1.** For each \( n = 1, \ldots, m \), \( \text{proj}^n_c Q^n_c = Q^n_c \).

**Proof.** If \( s_c \in Q^c \), then \( s_c \in \mathbb{R}[\tau^n_c(s_c)] \) and so \( (s_c, v^n_c(s_c)) \in Q^n_c \). Fix some \( (s_c, v^n_c(r_c)) \in Q^n_c \). Then, \( r_c \in Q^c \) and \( s_c \in \mathbb{R}[\tau^n_c(r_c)] \). It follows that \( s_c \in \mathbb{R}[\tau^n_c(r_c)] \subseteq Q^n_c \), as required.

**Lemma C.2.** For each \( n = 1, \ldots, m \), \( R^n_c(T) \times R^n_c(T) = \bigcup_{k=1}^m (Q^k_c \times Q^k_c) \).

**Proof.** The case of \( n = 1 \) is immediate from the construction. Thus, we show \( n = 2, \ldots, m \). The proof is by induction on \( n \).

Fix some \( n = 2, \ldots, m \) and some \( k = n-1, \ldots, m \) and some \( (r_c, v^n_c(s_c)) \in \mathbb{R}[\tau^n_c(s_c)] \times \{v^n_c(s_c)\} \subseteq Q^n_c \). Since the claim holds for \( n = 1 \) it suffices to show the following:

(i) If \( k = n-1 \), then \( v^n_c(s_c) \) does not strongly believe \( R^{n-1}_c(T) \).

(ii) If \( k = n, \ldots, m \), then \( v^n_c(s_c) \) strongly believes \( R^{n-1}_c(T) \).

**n = 2:** Fix some \( k = 1, \ldots, m \) and some \( (r_c, v^n_c(s_c)) \in \mathbb{R}[\tau^n_c(s_c)] \times \{v^n_c(s_c)\} \subseteq Q^n_c \). We will show (i)-(ii) hold. To do so, we will make use of the following property: \( R^1_c(T) = \bigcup_{k=1}^m Q^{-1}_c \) and \( Q^{-1}_c \) is also defined as \( \text{proj}^c \bigcup_{k=1}^m Q^k_c = \text{proj}^c R^1_c(T) \) (Lemma C.1).

First, suppose that \( k = 1 \) and \( Q^1_c \) is an effective singleton. By Condition (j.a) of strong justification, \( j^n_1(s_c) \) does not strongly believe \( Q^1_c \), i.e., there exists some information set \( h \) with \( Q^1_c \cap S^{-1}_c(h) = \emptyset \) and \( j^n_1(s_c)(S^{-1}_c(Q^1_c \cap S^{-1}_c(h)) > 0 \). Since \( Q^1_c = \text{proj}^c R^1_c(T) \cap (S^{-1}_c(h) \times T^{-1}_c) \neq \emptyset \). Moreover, \( \beta_c(v^n_c(s_c))(Q^1_c \times T^{-1}_c) > 0 \) and, again using the fact that \( Q^1_c = \text{proj}^c R^1_c(T) \), \( (Q^1_c \times T^{-1}_c) \cap R^1_c(T) = 0 \). Thus, \( v^n_c(s_c) \) does not strongly believe \( R^1_c(T) \).

Next, suppose that \( k = 1 \) and \( Q^1_c \) is non-singleton. Observe that, in this case,

\[
\beta_c(v^n_c(s_c))(S^{-1}_c \times T^{-1}_c) = 1.
\]

By Condition (j.a) of strong justification, if \((s_c, t_c) \in (S^{-1}_c \times T^{-1}_c) \) then \( s_c \not\in \mathbb{R}[\tau^n_c(t_c)] \) and so \((s_c, t_c) \in R^1_c(T) \). Thus, \( v^n_1(s_c) \) does not strongly believe \( R^1_c(T) \).

Finally, suppose that \( k = 2, \ldots, m \). Fix a conditioning event \( S^{-1}_c(h) \times T^{-1}_c \) so that \( R^1_c(T) \cap (S^{-1}_c(h) \times T^{-1}_c) \neq \emptyset \). Since \( Q^1_c = \text{proj}^c Q^1_c = \text{proj}^c R^1_c(T) \), it follows that \( Q^1_c \cap S^{-1}_c(h) \neq \emptyset \). So, using the fact that \( j^n_c(s_c) \) strongly believes \( Q^1_c \), it follows that \( j^n_c(s_c)(Q^1_c \cap S^{-1}_c(h)) = 1 \). Now observe that, by
Fix a CPS to show the lemma, it will be useful to begin with a number of preliminary results. Now observe that, by construction,

\[ R_T((\bigcup_{l=1}^{k-1} \text{diag}_c^l)S_c(h) \times T_{-c}) = j_c(s_c)(Q_{-c}^l|S_c(h)) = 1. \]

Since \( \bigcup_{l=1}^{k-1} \text{diag}_c^l \subseteq \bigcup_{l=1}^{m} Q_{-c}^l \) and \( R_{-c}^1(T) = \bigcup_{l=1}^{m} Q_{-c}^l \) (the result shown for \( n = 1 \)), it follows that \( \beta_c(v_c^k(s_c))(R_{-c}^1(T)|S_c(h) \times T_{-c}) = 1 \), as desired.

**n ≥ 3**: Let \( n = 3, \ldots, m \) and suppose the result was shown for \( n - 1 \). Fix some \( k = n - 1, \ldots, m \) and some \( (r_c, v_c^k(s_c)) \in \mathbb{B}[j_c^k(s_c)] \times \{v_c^k(s_c)\} \subseteq Q_{-c}^k \). We show (i)-(ii).

First, suppose that \( k = n - 1 \). Fix \((s_{-c}, t_{-c})\) with \( \beta_c(v_c^k(s_c))((s_{-c}, t_{-c})|S_{-c} \times T_{-c}) > 0 \) and note that, by construction, \( t_{-c} = v_c^{k-1}(s_c) \). By the induction hypothesis (part (i)), \( v_c^{k-1}(s_c) \) does not strongly believe \( R_{-c}^{n-2}(T) \). Thus, \( v_c^k(s_c) \) does not strongly believe \( R_{-c}^{n-1}(T) \).

Second, suppose that \( k = n, \ldots, m \). Fix a conditioning event \( S_{-c}(h) \times T_{-c} \) so that \( R_{-c}^{n-1}(T) \cap (S_{-c}(h) \times T_{-c}) \neq \emptyset \). By the induction hypothesis and Lemma C.1

\[
\text{proj}_{S_{-c}} R_{-c}^{n-1}(T) = \text{proj}_{S_{-c}} \bigcup_{k=n-1}^{m} Q_{-c}^k = Q_{-c}^{n-1}
\]

and so \( Q_{-c}^{n-1} \cap S_{-c}(h) \neq \emptyset \). Since \( j_c^k(s_c) \) strongly believes \( Q_{-c}^{n-1} \), it follows that \( j_c^k(s_c)(Q_{-c}^{n-1}|S_{-c}(h)) = 1 \).

Now observe that, by construction,

\[
\beta_c(v_c^k(s_c))(\bigcup_{l=1}^{k-1} \text{diag}_c^l|S_c(h) \times T_{-c}) = j_c^k(s_c)(Q_{-c}^{n-1}|S_c(h)) = 1.
\]

Since \( \bigcup_{l=1}^{k-1} \text{diag}_c^l \subseteq \bigcup_{l=n-1}^{m} Q_{-c}^l \) and, by the induction hypothesis, \( R_{-c}^{n-1}(T) = \bigcup_{l=n-1}^{m} Q_{-c}^l \), it follows that \( \beta_c(v_c^k(s_c))(R_{-c}^{n-1}(T)|S_c(h) \times T_{-c}) = 1 \), as desired. \( \square \)

**Proof of Proposition C.1.** Immediate from Lemmata C.1-C.2. \( \square \)

**C.2 Proof of Proposition C.2**

Say a strategy \( s_c \) is **justifiable** if there exists some CPS \( p_c \) so that \( s_c \in \mathbb{B}[p_c] \). Proposition C.2 follows from the following lemma.

**Lemma C.3.** Suppose that the game is generic and let \([s_{-c}^*] \subseteq S_{-c} \). If \( s_{-c}^* \) is justifiable, then there exists some CPS \( p_c \) so that \([s_{-c}^*] = \mathbb{B}[p_c] \) and \( p_c \) does not strongly believe \([s_{-c}^*] \).

To show the lemma, it will be useful to begin with a number of preliminary results.

**Lemma C.4.** Fix a CPS \( p_c \in \mathcal{C}(S_{-c}, E_c) \) so that \([s_{-c}] = \mathbb{B}[p_c] \) and some \( r_c \notin [s_c] \). There exists some \( h \in H_c \cup \{\phi\} \) so that \( s_c, r_c \in S_c(h) \) and \( s_c(h) \neq r_c(h) \). Moreover, for any such \( h \),

\[
\sum_{s_{-c} \in S_{-c}(h)} [\pi_c(s_c, s_{-c}) - \pi_c(r_c, s_{-c})] p_c(s_{-c}|S_{-c}(h)) > 0.
\]
**Proof.** Fix \([s_c] \subseteq \mathbb{R}[p_c]\) and \(r_c \not\in [s_c]\). Then, for all \(h \in H_c \cup \{\phi\}\) with \(s_c, r_c \in S_c(h)\)

\[
\sum_{s_c \in S_c(h)} \left[ \pi_c(s_c, s_c) - \pi_c(r_c, s_c) \right] p_c(s_c|S_c(h)) \geq 0. \tag{3}
\]

Since \(r_c \not\in [s_c]\), there exists some \(h^* \in H_c\) so that \(s_c, r_c \in S_c(h^*)\) and \(s_c(h^*) \neq r_c(h^*)\). We will suppose that Equation (3) holds with equality at \(h = h^*\) and construct a new strategy \(r_c^*\) with \(r_c^* \not\in [s_c]\) and \(r_c^* \in \mathbb{R}[p_c]\). This will establish the result.

Construct the strategy \(r_c^*\) as follows: First, for each information set \(h\) with either \(S(h) \cap S(h^*) = \emptyset\) or \(S(h^*) \subseteq S(h)\), set \(r_c^*(h) = s_c(h)\). Second, for each information set \(h\) with \(S(h) \subseteq S(h^*)\) and \(p_c(S_c(h)|S_c(h^*)) > 0\), set \(r_c^*(h) = r_c(h)\).

Finally, for all remaining information sets, choose \(r_c^*\) to satisfy the following condition: If \(r_c^* \in S_c(h)\), then \(r_c^*\) solves

\[
\max_{s_c(h)} \sum_{s_c \in S_c(h)} \pi_c(\cdot, s_c) p_c(s_c|S_c(h)). \tag{4}
\]

The fact that we can choose \(r_c^*\) in this way follows from Lemma A.3.\(^\text{10}\)

Observe that \(r_c^* \not\in [s_c]\). Also observe that \(r_c^*\) is optimal under \(p_c(\cdot|S_c(h^*))\) given \(S_c(h^*)\). To see this, fix some \(s_c \in \text{Supp} p_c(\cdot|S_c(h^*))\). Since \((s_c, s_c), (r_c, s_c) \in S_c(h^*) \times S_c(h^*) = S(h^*)\), it follows from the construction that \((r_c^*, s_c) \in S(h^*)\). Thus, \(r_c^* \in S_c(h^*)\). Moreover, by construction, if \(s_c \in \text{Supp} p_c(\cdot|S_c(h^*))\) then \(\zeta(r_c^*, s_c) = r_c^*\). So, since \(r_c\) is optimal under \(p_c(\cdot|S_c(h^*))\) given \(S_c(h^*)\), it follows that \(r_c^*\) is also optimal under \(p_c(\cdot|S_c(h^*))\) given \(S_c(h^*)\).

We will show that \(r_c^* \in \mathbb{R}[p_c]\). Specifically, fix an information set \(h \in H_c \setminus \{h^*\}\) with \(r_c^* \in S_c(h)\). We will show that \(r_c^*\) is optimal under \(p_c(\cdot|S_c(h))\) given \(S_c(h)\).

First, suppose that \(S(h^*) \cap S(h) = \emptyset\). Fix some \(p_c(s_c|S_c(h)) > 0\). By construction, \(\zeta(r_c^*, s_c) = \zeta(s_c, s_c)\). Since \(s_c\) is optimal under \(p_c(\cdot|S_c(h))\) given \(S_c(h)\), it follows that \(r_c^*\) is also optimal under \(p_c(\cdot|S_c(h))\) given \(S_c(h)\).

Second, suppose that \(h \neq h^*, S(h) \subseteq S(h^*)\), and \(p_c(S_c(h)|S_c(h^*)) > 0\). Since \(r_c^*\) is optimal under \(p_c(\cdot|S_c(h^*))\) given \(S_c(h^*)\), Definition 3.2 and Lemma A.3 give that \(r_c^*\) is optimal under \(p_c(\cdot|S_c(h))\) given \(S_c(h)\).

Third, suppose that \(h \neq h^*, S(h) \subseteq S(h^*)\), and \(p_c(S_c(h)|S_c(h^*)) = 0\). In that case, by assumption, \(r_c^*\) is optimal under \(p_c(\cdot|S_c(h))\) given \(S_c(h)\).

Finally, suppose that \(S(h^*) \subseteq S(h)\). Fix some \(p_c(s_c|S_c(h)) > 0\). If \(s_c \not\in S_c(h^*)\), then \(\zeta(r_c^*, s_c) = \zeta(s_c, s_c)\). (This is by construction.) If \(s_c \in S_c(h^*)\), then \(\zeta(r_c^*, s_c) = \zeta(r_c, s_c)\). Observe that \(S_c(h^*) \subseteq S_c(h)\); so, by Definition 3.2, \(p_c(s_c|S_c(h)) > 0\) implies \(p_c(s_c|S_c(h^*)) > 0\). By construction, for any \(s_c\) with \(p_c(s_c|S_c(h^*)) > 0\), \(\zeta(r_c^*, s_c) = \zeta(r_c, s_c)\).

Let \(\alpha = p_c(S_c(h) \setminus S_c(h^*))|S_c(h)\). If \(\alpha > 0\), let \(\mu_c\) be \(p_c(\cdot|S_c(h))\) conditional on \(S_c(h) \setminus S_c(h^*)\).

\(^{10}\)Specifically: Let \(\tilde{H}^0_c\) be the set of all \(h \in H_c\) with \(S(h) \subseteq S(h^*)\), \(p_c(S_c(h)|S_c(h^*)) = 0\), and \(r_c \in S_c(h)\). Choose some \(h^1 \in \tilde{H}^0_c\) and note that \(r_c^* \in S_c(h^1)\). Choose \(r_c^*\) to solve Equation (4) for \(h = h^1\) and set \(r_c^*(h) = r_c^*(h^1)\). Then define \(\tilde{H}^k_c\) to be the set \(h \in \tilde{H}^0_c\) so that \(r_c^* \in S_c(h)\) and, if \(S_c(h) \subseteq S_c(h^k)\), then \(p_c(S_c(h)|S_c(h^k)) = 0\). Proceed inductively, until some \(\tilde{H}^k_c = \emptyset\) has been constructed. Then, “fill in” \(r_c^*(h)\) arbitrarily at all information sets \(h\) for which it has not been defined. (Note, \(r_c^*\) precludes those information sets.)
If $\alpha = 0$, let $\mu_c$ be the zero measure. Then

$$\sum_{s_c \in S_{c-}(h)} \left[ \pi_c(s_c, s_{c-}) - \pi_c(r^*_c, s_{c-}) \right] p_c(s_{c-}|S_{c-}(h)) =$$

$$\alpha \sum_{s_c \in S_{c-}(h)} \left[ \pi_c(s_c, s_{c-}) - \pi_c(r^*_c, s_{c-}) \right] \mu(s_{c-}) + (1 - \alpha) \sum_{s_c \in S_{c-}(h)} \left[ \pi_c(s_c, s_{c-}) - \pi_c(r^*_c, s_{c-}) \right] p_c(s_{c-}|S_{c-}(h^*)) .$$

Note that

$$\alpha \sum_{s_c \in S_{c-}(h)} \left[ \pi_c(s_c, s_{c-}) - \pi_c(r^*_c, s_{c-}) \right] \mu(s_{c-}) = 0$$

since $\mu_c(s_{c-}) > 0$ implies $\zeta(s_c, s_{c-}) = \zeta(r^*_c, s_{c-})$. Also note that

$$(1 - \alpha) \sum_{s_c \in S_{c-}(h)} \left[ \pi_c(s_c, s_{c-}) - \pi_c(r^*_c, s_{c-}) \right] p_c(s_{c-}|S_{c-}(h^*)) = 0 ,$$

since both $s_c$ and $r^*_c$ are optimal under $p_c(\cdot|S_{c-}(h^*))$. Thus,

$$\sum_{s_c \in S_{c-}(h)} \left[ \pi_c(s_c, s_{c-}) - \pi_c(r^*_c, s_{c-}) \right] p_c(s_{c-}|S_{c-}(h)) = 0 .$$

Now, it follows from the fact that $s_c \in S_c(h^*) \subseteq S_c(h)$ is optimal under $p_c(\cdot|S_{c-}(h))$ given $S_c(h)$ that $r_c$ is also optimal under $p_c(\cdot|S_{c-}(h))$. $\blacksquare$

**Lemma C.5.** Fix some $h^* \in H_c \cup \{\emptyset\}$ so that $s^*_c \in S_c(h^*)$, $s^*_{c-} \notin S_{c-}(h^*)$ and, for all $h \in H_c \cup \{\emptyset\}$ with $S(h^*) \subseteq S(h)$, $s^*_{c-} \in S_{c-}(h)$. Then $\zeta(s^*_c, s^*_{c-}) = \zeta(r_c, s^*_{c-})$ implies $r_c \in S_c(h^*)$.

**Proof.** We show the contrapositive. Suppose that $r_c \notin S_c(h^*)$. There exists some $(s^*_c, r_{c-}) \in S(h^*)$ so that $(r_c, r_{c-}) \notin S(h^*)$. Let $v$ be the last common predecessor of $\zeta(s^*_c, r_{c-})$ and $\zeta(r_c, r_{c-})$. Note that there exists some $h \in H_c$ so that $v \in h$ and $s^*_c \neq r_{c-}$. Observe that $S(h) \cap S(h^*) \neq \emptyset$. As such, either $S(h) \subseteq S(h^*)$ or $S(h^*) \subseteq S(h)$. Since $r_c \in S_c(h)$ but $r_c \notin S_c(h^*)$, it follows that $S(h^*) \subseteq S(h)$. By construction, $s^*_{c-} \in S_{c-}(h)$. Thus, $\zeta(s^*_c, s^*_{c-}) \neq \zeta(r_c, s^*_{c-})$. $\blacksquare$

**Proof of Lemma C.3.** Since the game is generic and $s^*_c$ is justifiable, there exists some CPS $p_c$ so that $[s^*_c] = \mathbb{E}[p_c]$. If $p_c$ does not strongly believe $[s^*_{c-}]$, then we are done. So throughout we suppose otherwise. We will show that we can tilt $p_c$ to construct a new CPS that satisfies the desired properties. We divide the argument into two cases.

**Case A.** Suppose that, for each $h \in H_c$ with $s^*_c \in S_c(h)$, $s^*_{c-} \in S_{c-}(h)$. So, for each $h \in H_c$ with $s^*_c \in S_c(h)$, $p_c(s^*_{c-}|S_{c-}(h)) = 1$. Lemma C.4 then implies that $\pi_c(s^*_c, s^*_{c-}) > \pi_c(s_c, s^*_{c-})$ for all $s_c \in S_c \setminus [s^*_c]$.

Since $S_{c-}\setminus [s^*_{c-}] \neq \emptyset$, we can choose $r^*_{c-} \in S_{c-}\setminus [s^*_{c-}]$. For each $\varepsilon \in (0, 1)$, construct a CPS $q^*_c$ so that

$$q^*_c(s^*_{c-}|S_{c-}) = 1 - \varepsilon \quad \text{and} \quad q^*_c(r^*_{c-}|S_{c-}) = \varepsilon$$

and, for each $h \in H_c$ with $S_{c-}(h) \cap \{s^*_{c-}, r^*_{c-}\} = \emptyset$, $q^*_c(\cdot|S_{c-}(h)) = p_c(\cdot|S_{c-}(h))$. Note, the unique CPS $q^*_c$ that satisfies these conditions does not strongly believe $[s^*_{c-}]$.

Now observe that we can find some $\tilde{\varepsilon} > 0$ so that for each $\varepsilon \in (0, \tilde{\varepsilon})$ the following holds: If $h \in H_c$ with
Let $s_e^* \in S_e(h)$, then
\[
\sum_{s_{-c} \in S_{-c}} [\pi_e(s_e^*, s_{-c}) - \pi_e(r_e, s_{-c})]q_e^*(s_{-c}|S_{-c}) = (1 - \varepsilon)[\pi_e(s_e^*, s_{-c}^*) - \pi_e(s_e^*, s_{-c})] + \varepsilon[\pi_e(s_e^*, r_{-c}^*) - \pi_e(s_e^*, r_{-c})] > 0
\]
for each $r_e \in S_e(h)$. Thus, $\mathbb{B}R[q_e^*] = [s_e^*]$ for all $\varepsilon \in (0, \bar{\varepsilon})$.

**Case B.** Suppose that there exists some $h^* \in H_e$ so that $s_e^* \in S_e(h^*)$ but $s_{-c}^* \notin S_{-c}(h^*)$. Choose $h^*$ so that, if $S(h^*) \subseteq S(h)$, then $s_{-c}^* \in S_{-c}(h)$. Let $\mu_e^c = p_c(\cdot|S_{-c}(h^*))$ and observe that $\mu_e^c([s_{-c}^*]) = 0$ since $s_{-c}^* \notin S_{-c}(h^*)$. For each $\varepsilon \in (0, 1)$, construct a CPS $q_e^c$ so that
\[
q_e^c(s_{-c}|S_{-c}) = \begin{cases} 
1 - \varepsilon & \text{if } s_{-c} = s_{-c}^* \\
\varepsilon \mu_e^c(s_{-c}) & \text{if } s_{-c} \neq s_{-c}^*.
\end{cases}
\]
and, for each $h \in H_e$, with $S_{-c} \cap \{s_{-c}^*\} \cup \text{Supp} \mu_e^c = \emptyset$, $q_e^c(s_{-c}|S_{-c}(h)) = p_c(\cdot|S_{-c}(h))$. Note, the unique CPS $q_e^c$ that satisfies these conditions does not strongly believe $[s_{-c}^*]$. We show that we can choose $\varepsilon > 0$ so that $\mathbb{B}R[q_e^c] = [s_e^*]$.

**Step 1:** We begin by showing that, for each $r_e \in S_e$, there exists some $\bar{\varepsilon}(r_e) > 0$ so that the following holds: For all $\varepsilon \in (0, \bar{\varepsilon}(r_e))$,
\[
\sum_{s_{-c} \in S_{-c}} [\pi_e(s_e^*, s_{-c}) - \pi_e(r_e, s_{-c})]q_e^c(s_{-c}|S_{-c}) \begin{cases} 
> 0 & \text{if } \zeta(r_e, s_{-c}^*) \neq \zeta(s_e^*, s_{-c}^*) \\
\geq 0 & \text{if } \zeta(r_e, s_{-c}^*) = \zeta(s_e^*, s_{-c}^*). 
\end{cases}
\]  
(5)

First, suppose that $\zeta(r_e, s_{-c}^*) \neq \zeta(s_e^*, s_{-c}^*)$. Then, there exists some $\hat{h}$ so that $(s_e^*, s_{-c}^*), (r_e, s_{-c}^*) \in S(\hat{h})$ and $s_e^*(\hat{h}) \neq r_e(\hat{h})$. Moreover, $p_e(s_{-c}^*|S_{-c}(\hat{h})) = 1$. Thus, applying Lemma C.4, $\pi_e(s_e^*, s_{-c}^*) > \pi_e(r_e, s_{-c}^*)$. It follows that there exists some $\bar{\varepsilon}(r_e) > 0$ so that, for all $\varepsilon \in (0, \bar{\varepsilon}(r_e))$,
\[
\sum_{s_{-c} \in S_{-c}} [\pi_e(s_e^*, s_{-c}) - \pi_e(r_e, s_{-c})]q_e^c(s_{-c}|S_{-c}) = (1 - \varepsilon)[\pi_e(s_e^*, s_{-c}^*) - \pi_e(r_e, s_{-c})] + \varepsilon \sum_{s_{-c} \in S_{-c}} [\pi_e(s_e^*, s_{-c}) - \pi_e(r_e, s_{-c})] \mu_e^c(s_{-c}) > 0
\]

Second, suppose that $\zeta(r_e, s_{-c}^*) = \zeta(s_e^*, s_{-c}^*)$. In this case, $\pi_e(s_e^*, s_{-c}^*) - \pi_e(r_e, s_{-c}^*) = 0$. Moreover, if $s_e^* \in S_e(h^*)$, then $r_e \in S_e(h^*)$. (See Lemma C.5.) Since $s_e^*$ is optimal under $\mu_e^c$ given $S_e(h^*)$, it follows that
\[
\sum_{s_{-c} \in S_{-c}} [\pi_e(s_e^*, s_{-c}) - \pi_e(r_e, s_{-c})] \mu_e^c(s_{-c}) \geq 0.
\]
As such,
\[
\sum_{s_{-c} \in S_{-c}} [\pi_c(s^*_c, s_{-c}) - \pi_c(r_c, s_{-c})]q^*_c(s_{-c}|S_{-c}) = \\
(1 - \varepsilon)[\pi_c(s^*_c, s^*_{-c}) - \pi_c(r_c, s^*_{-c})] + \\
\varepsilon \sum_{s_{-c} \in S_{-c}} [\pi_c(s^*_c, s_{-c}) - \pi_c(r_c, s_{-c})]\mu_c(s_{-c}) \geq 0
\]
for all \(\varepsilon > 0\).

**Step 2:** Take \(\bar{\varepsilon} = \min\{|\varepsilon(r_c) : r_c \in S_c\}\). We will show that \([s^*_c] \subseteq \mathbb{BR}[q^*_c]\) for all \(\varepsilon \in (0, \bar{\varepsilon})\). To do so, begin by noting that Equation (5) holds for all \(r_c \in S_c\), provided \(\varepsilon \in (0, \bar{\varepsilon})\). To complete the argument, it suffices to show that, if \(h \in H_c\) with \(s^*_c \in S_c(h)\) then either: \(q^*_c(\cdot|S_{-c}(h)) = q^*_c(\cdot|S_{-c})\) or \(q^*_c(\cdot|S_{-c}(h)) = p_c(\cdot|S_{-c}(h))\). From this the conclusion will follow.

First, suppose that \(S(h^*) \subseteq S(h)\). In that case, \(q^*_c(\cdot|S_{-c}(h)) = q^*_c(\cdot|S_{-c})\). Second, suppose that \(S(h) \subseteq S(h^*)\). In that case, \(q^*_c(\cdot|S_{-c}(h)) = p_c(\cdot|S_{-c}(h))\). Finally, suppose that \(S(h^*) \cap S(h) = \emptyset\). In that case, \(s^*_c \in S_c(h^*) \cap S_c(h)\) and so, \(S_{-c}(h^*) \cap S_{-c}(h) = \emptyset\). (See Lemma A.2.) From this, \(q^*_c(\text{Supp} \mu^*_c|S_{-c}(h)) = 0\) and so \(q^*_c(\cdot|S_{-c}(h)) = p_c(\cdot|S_{-c}(h))\).

**Step 3:** We now show that, for all \(\varepsilon \in (0, \bar{\varepsilon})\), \(\mathbb{BR}[q^*_c] \subseteq [s^*_c]\). To see this, fix some \(r_c \not\in [s_c]\). Then there exists some \(h \in H_c \cup \{\phi\}\) so that \(s_c, r_c \in S_c(h)\) and
\[
\sum_{s_{-c} \in S_{-c}} [\pi_c(s^*_c, s_{-c}) - \pi_c(r_c, s_{-c})]p_c(s_{-c}|S_{-c}(h)) > 0.
\]
(See Lemma C.4.) If \(q^*_c(\cdot|S_{-c}(h)) = p_c(\cdot|S_{-c}(h))\), then certainly \(r_c \not\in \mathbb{BR}[q^*_c]\). If \(q^*_c(\cdot|S_{-c}(h)) \neq p_c(\cdot|S_{-c}(h))\), then \(S(h^*) \subseteq S(h)\). In that case,
\[
\sum_{s_{-c} \in S_{-c}} [\pi_c(s^*_c, s_{-c}) - \pi_c(r_c, s_{-c})]p_c(s_{-c}|S_{-c}(h)) = \pi_c(s^*_c, s^*_{-c}) - \pi_c(r_c, s^*_{-c}) > 0.
\]
Thus, \(\zeta(s^*_c, s^*_{-c}) \neq \zeta(r_c, s^*_{-c})\) and, so, by Equation (5) \(r_c \not\in \mathbb{BR}[q^*_c]\). □

**Appendix D** Analyzing Games

**D.1 Repeated Prisoner’s Dilemma**

It will be convenient to adopt the following national convention: Fix some player \(c\) and write \((\phi, (\alpha^1_c, \alpha^2_{-c})) \in V^2\) for a second-period history (resp. \((\phi, (\alpha^1_c, \alpha^2_{-c}), (\alpha^2_c, \alpha^1_{-c})) \in V^3\) for a third-period history). Note, when we write \((\alpha^i_c, \alpha^j_{-c})\) we refer to vector that first specifies a \(t\)-period action of player \(c\) and then specifies a \(t\)-period action of player \(-c\). So, when we fix some player \(c\) and look at the history \((\phi, (C, D))\), we mean the history where, in the first period, player \(c\) chooses \(C\) and player \(-c\) chooses \(D\) in the first period. Write \(EFR^c\) for the set of \(s_c\) so that \(s_c(v) = D\) for each \(v \in V^3\) with \(s_c \in S_c(v)\).
**D.1.1 Unique Best Responses**

We begin by recording results about best responses. This will serve three purposes. First, we will use the results to establish that the game is generic. Second, we will use the results to establish that $\hat{\text{EFR}} = \text{EFR}_c$, a claim made in the text. Third, we will use the results to prove Proposition 7.1.

**Lemma D.1.** Let $s^*_c \in \hat{\text{EFR}}_c$ be such that one of the following conditions hold:

(i) $s^*_c(\phi) = C$ and $s^*_c(\phi, (C, C)) = D$,

(ii) $s^*_c(\phi) = D$ and $s^*_c(\phi, (D, D)) = D$,

(iii) $s^*_c(\phi) = D$, $s^*_c(\phi, (D, D)) = C$ and $s^*_c(\phi, (D, C)) = D$, or

(iv) if $\zeta + \gamma > 2\delta$, $s^*_c(\phi) = C$, $s^*_c(\phi, (C, C)) = C$ and $s^*_c(\phi, (C, D)) = D$.

Then there exists some $p_c \in C(S_{-c}, E_c)$ and some $s^*_{-c} \in \hat{\text{EFR}}_{-c}$ so that (i) $\mathbb{B}[p_c] = [s^*_c]$, and (ii) $p_c$ strongly believes $[s^*_{-c}]$.

**Lemma D.2.** Let $s^*_c \in \hat{\text{EFR}}_c$ be such that one of the following conditions hold:

(i) $s^*_c(\phi) = C$ and $s^*_c(\phi, (C, C)) = C$, or

(ii) $s^*_c(\phi) = D$, $s^*_c(\phi, (D, C)) = C$ and $s^*_c(\phi, (D, D)) = C$.

Then there exists some $p_c \in C(S_{-c}, E_c)$ so that $\mathbb{B}[p_c] = [s^*_c]$.

As a corollary of these Lemmata, we have that the following:

**Corollary D.1.**

(i) The 3-repeated Prisoner’s Dilemma is generic.

(ii) $\text{EFR}_c = \hat{\text{EFR}}_c$.

**Proof.** It is immediate that $\text{EFR}_c \subseteq \hat{\text{EFR}}_c$. Thus, both parts follow if, for each strategy $s_c \in \hat{\text{EFR}}_c$, there is a CPS $p_c$ so that $\mathbb{B}[p_c] = [s_c]$. This follows from Lemma D.1(i)-(ii)-(iii) and Lemma D.2(i)-(ii). 

The proofs of Lemmata D.1-D.2 proceed as follows: We begin by fixing a strategy $s^*_c$ and showing that $\pi_c(s^*_c, s^*_{-c}) \geq \pi_c(r_c, s^*_{-c})$ for each $r_c \in \hat{\text{EFR}}_c$ with $\zeta(s^*_c, s^*_{-c}) \neq \zeta(r_c, s^*_{-c})$. Next, we observe that there is exactly one $v^* \in V^2$ so that $s^*_c \in S_c(v^*)$ and $s^*_{-c} \notin S_{-c}(v^*)$. (This follows since each player has two actions.) We then fix a strategy $r^*_c \in S_{-c}(v^*)$ and show that, for each $r_c \in \hat{\text{EFR}}_c \cap S_c(v^*)$ with $\zeta(s^*_c, r^*_c) \neq \zeta(r_c, r^*_c)$, $\pi_c(s^*_c, r^*_c) \geq \pi_c(r_c, r^*_c)$. Moreover, for each $r_c \in \hat{\text{EFR}}_c \setminus [s^*_c]$, either $\pi_c(s^*_c, s^*_{-c}) > \pi_c(r_c, s^*_{-c})$ or $\pi_c(s^*_c, r^*_c) > \pi_c(r_c, r^*_c)$.\(^{17}\) Now consider a CPS $p_c$ so that

(i) $p_c([s^*_{-c}] | S_{-c}(v)) = 1$, if $s^*_{-c} \in S_{-c}(v)$, and

(ii) $p_c([r^*_{-c}] | S_{-c}(v)) = 1$, if $s^*_{-c} \notin S_{-c}(v)$ but $r^*_{-c} \in S_{-c}(v)$.

\(^ {17}\)A detail in how this is implemented in the proof: For each $r_c \in \hat{\text{EFR}}_c \setminus [s^*_c]$ either $\zeta(s^*_c, s^*_{-c}) \neq \zeta(r_c, s^*_{-c})$ or $\zeta(s^*_c, r^*_c) \neq \zeta(r_c, r^*_c)$. This, again follows from the fact that each player has two actions.
(All other choices are arbitrary.) Then (i) $\mathbb{B[R]}[p_c] = [s^*_c]$, and (ii) $p_c$ strongly believes $[s^*_c]$. In the specific case of Lemma D.1, we will be choosing $s^*_{r_c} \in \mathbb{EFR}_{1-c}$, allowing for a stronger conclusion.

**Remark D.1.** There is one case of interest not covered by Lemma D.1: namely $\zeta + \gamma = 2\delta$. In that case, there will be one other strategy—called $r_c$ in the proof—so that $\pi(s^*_c, s^*_{r_c}) = \pi(r_c, s^*_{r_c})$ and $r_c \not\in S_c(v^*)$. In that case, under the construction $\mathbb{B[R]}[p_c] = [s^*_c] \cup [r_c]$. We will make use of that fact.

We now complete the non-common features of the proofs.

**Proof of Lemma D.1(i).** Let $s^*_{r_c}$ be a strategy with $s^*_{r_c}(\phi) = C$, $s^*_{r_c}(\phi, (C, C)) = C$, $s^*_{r_c}(\phi, (D, C)) = D$ and, for each $v \in V^3$, $s^*_{r_c}(v) = D$. Note, $s^*_{r_c} \in \mathbb{EFR}_{1-c}$. Observe that $\pi_c(s^*_{r_c}, s^*_{r_c}) = \kappa + \zeta + \delta$. Fix some $r_c \in \mathbb{EFR}_c$ with $\zeta(r_c, s^*_{r_c}) \neq \zeta(s^*_{r_c}, s^*_{r_c})$. There are three possible cases:

- $r_c(\phi) = D$ and $r_c(\phi, (D, C)) = C$: $\pi_c(r_c, s^*_{r_c}) = \zeta + \gamma + \delta$.
- $r_c(\phi) = D$ and $r_c(\phi, (D, C)) = D$: $\pi_c(r_c, s^*_{r_c}) = \zeta + 2\delta$, and
- $r_c(\phi) = C$ and $r_c(\phi, (C, C)) = C$: $\pi_c(r_c, s^*_{r_c}) = 2\kappa + \delta$.

In each case, $\pi_c(s^*_{r_c}, s^*_{r_c}) > \pi_c(r_c, s^*_{r_c})$.

There is a single history $v^* \in V^2$ so that $s^*_{r_c} \in S_c(v^*)$ but $s^*_{r_c} \not\in S_{-c}(v^*)$, namely $v^* = (\phi, (C, D))$. We will choose distinct strategies $r^*_{r_c}$ based on whether $s^*_{r_c}(v^*) = C$ or $s^*_{r_c}(v^*) = D$.

First, suppose $s_c(v^*) = C$. Let $r^*_{r_c}$ be such that $r^*_{r_c}(\phi) = D$ and, for each $v \in V^2 \cup V^3$, $r^*_{r_c}(v) = C$ if and only if $v = ((\alpha^{-1}, \alpha^{-1}-1))$ with $\alpha^{-1} = C$. Then, $\pi_c(s^*_{r_c}, r^*_{r_c}) = \gamma + \kappa + \zeta$. Fix $r_c \in \mathbb{EFR}_c \cap S_c(v^*)$ with $\zeta(r_c, r^*_{r_c}) \neq \zeta(s^*_{r_c}, r^*_{r_c})$. Then $r_c(v^*) = D$ and so $\pi_c(r_c, r^*_{r_c}) = \gamma + \kappa + \zeta < s^*_{r_c}(r^*_{r_c})$.

Second, suppose $s_c(v^*) = D$. Let $r^*_{r_c}$ be such that $r^*_{r_c}(v) = D$ for each $v \in V$. Then, $\pi_c(s^*_{r_c}, r^*_{r_c}) = \gamma + 2\delta$. Fix $r_c \in \mathbb{EFR}_c \cap S_c(v^*)$ with $\zeta(r_c, r^*_{r_c}) \neq \zeta(s^*_{r_c}, r^*_{r_c})$. Then $r_c(v^*) = C$ and so $\pi_c(r_c, r^*_{r_c}) = 2\gamma + \delta < s^*_{r_c}(r^*_{r_c})$. ■

**Remark D.2.** The proof of Lemma D.1(i) establishes that we can take $s^*_{r_c}$ so that $s^*_{r_c}(\phi) = C$, $s^*_{r_c}(\phi, (C, C)) = C$, $s^*_{r_c}(\phi, (D, C)) = D$ and, for each $v \in V^3$, $s^*_{r_c}(v) = D$.

**Proof of Lemma D.1(ii).** Let $s^*_{r_c}$ be a strategy with $s^*_{r_c}(v) = D$ for each $v \in V$. Note, $s^*_{r_c} \in \mathbb{EFR}_{1-c}$. Observe that $\pi_c(s^*_{r_c}, s^*_{r_c}) = 3\delta$. Fix some $r_c \in \mathbb{EFR}_c$ with $\zeta(r_c, s^*_{r_c}) \neq \zeta(s^*_{r_c}, s^*_{r_c})$. Then, either (i) $r_c(\phi) = C$ or (ii) $r_c(\phi) = D$ and $r_c(\phi, (D, D)) = C$. As such, $\pi_c(r_c, s^*_{r_c}) \leq 2\delta + \gamma < \pi_c(s^*_{r_c}, s^*_{r_c})$.

There is a single history $v^* \in V^2$ so that $s^*_c \in S_c(v^*)$ but $s^*_c \not\in S_{-c}(v^*)$, namely $v^* = (\phi, (D, C))$. We will choose distinct strategies $r^*_{r_c}$ based on whether $s^*_{c}(v^*) = C$ or $s^*_{c}(v^*) = D$.

First, suppose $s_c(v^*) = C$. Let $r^*_{r_c}$ be such that (i) $r^*_{r_c}(\phi) = C$, (ii) for each $v \in V^2$, $r^*_{r_c}(v) = C$ if and only if $v = ((\gamma, \gamma, \gamma))$, and (iii) for each $v \in V^3$, $r^*_{r_c}(v) = C$ if and only if $v = ((\gamma, \gamma, \gamma), (C, C), \gamma)$. Note, $\pi_c(s^*_{c}, r^*_{c}) = 2\zeta + \kappa$. Fix $r_c \in \mathbb{EFR}_c \cap S_c(v^*)$ with $\zeta(r_c, r^*_{r_c}) \neq \zeta(s^*_{c}, r^*_{r_c})$. Then $r_c(v^*) = D$ and so $\pi_c(r_c, r^*_{r_c}) = 2\zeta + \kappa < \pi_c(s^*_{c}, r^*_{r_c})$.

Second, suppose $s_c(v^*) = D$. Let $r^*_{r_c}$ be such that (i) $r^*_{r_c}(\phi) = C$, (ii) for each $v \in V^2 \cup V^3$, $r^*_{r_c}(v) = D$. Note, $\pi_c(s^*_{c}, r^*_{r_c}) = \zeta + 2\delta$. Fix $r_c \in \mathbb{EFR}_c \cap S_c(v^*)$ with $\zeta(r_c, r^*_{r_c}) \neq \zeta(s^*_{c}, r^*_{r_c})$. Then $r_c(v^*) = C$ and so $\pi_c(r_c, r^*_{r_c}) = \zeta + \gamma + \delta < \pi_c(s^*_{c}, r^*_{r_c})$. ■

33
Remark D.3. The proof of Lemma D.1(ii) establishes that we can take $s^*_c$ so that $s^*_c(v) = D$ for each $v \in V$.

Proof of Lemma D.1(iii). Let $s^*_c$ be a strategy with $s^*_c(\phi) = C$ and, $v \in V^2 \cup V^3$, $s^*_c(v) = D$. Note, $s^*_c \in \widehat{EFR}_{-c}$. Observe that $\pi_c(s^*_c, s^*_c) = \gamma + 2\delta$. Fix some $r_c \in \widehat{EFR}_{-c}$ with $\zeta(r_c, s^*_c) \neq \zeta(s^*_c, s^*_c)$. There are three possible cases:

- $r_c(\phi) = D$ and $r_c(\phi, (D, C)) = C$: $\pi_c(r_c, s^*_c) = 2\delta + \gamma$,
- $r_c(\phi) = C$ and $r_c(\phi, (C, C)) = D$: $\pi_c(r_c, s^*_c) = 3\delta + \gamma$,
- $r_c(\phi) = C$ and $r_c(\phi, (C, D)) = C$: $\pi_c(r_c, s^*_c) = \gamma + \delta$.

In each case, $\pi_c(s^*_c, s^*_c) > \pi_c(r_c, s^*_c)$.

There is a single history $v^* \in V^2$ so that $s^*_c \in S_c(v^*)$ but $s^*_c \notin S_{-c}(v^*)$, namely $v^* = (\phi, (D, D))$. Let $r^*_c$ be such that (i) $r^*_c(\phi) = D$, (ii) for each $v \in V^2$, $r^*_c(v) = C$ if and only if $v = (\phi, (D, \cdot))$, and (iii) for each $v \in V^3$, $r^*_c(v) = C$ if and only if $v = (\phi, (\cdot, \cdot, (C, \cdot)))$. Note, $\pi_c(s^*_c, r^*_c) = \gamma + \delta + \zeta$. Fix $r_c \in \widehat{EFR}_{-c} \cap S_c(v^*)$ with $\zeta(r_c, r^*_c) \neq \zeta(s^*_c, r^*_c)$. Then $r_c(v^*) = D$ and so $\pi_c(r_c, r^*_c) = 2\delta + \gamma < \pi_c(s^*_c, r^*_c)$.

Remark D.4. The proof of Lemma D.1(iii) establishes that we can take $s^*_c$ so that $s^*_c(v) = D$ for each $v \in V^2 \cup V^3$.

Proof of Lemma D.1(iv). Let $s^*_c$ be a strategy with (i) $s^*_c(\phi) = D$, (ii) for each $v \in V^2$, $s^*_c(v) = C$ if and only if $v = (\phi, (C, \cdot))$, and (iii) for each $v \in V^3$, $s^*_c(v) = D$. Note, $s^*_c \in \widehat{EFR}_{-c}$. Observe that $\pi_c(s^*_c, s^*_c) = \gamma + \zeta + \delta$. Fix some $r_c \in \widehat{EFR}_{-c}$ with $\zeta(r_c, s^*_c) \neq \zeta(s^*_c, s^*_c)$. There are three possible cases:

- $r_c(\phi) = D$ and $r_c(\phi, (D, C)) = C$: $\pi_c(r_c, s^*_c) = 2\delta + \gamma$,
- $r_c(\phi) = D$ and $r_c(\phi, (D, C)) = D$: $\pi_c(r_c, s^*_c) = 3\delta + \gamma$,
- $r_c(\phi) = C$ and $r_c(\phi, (C, D)) = C$: $\pi_c(r_c, s^*_c) = \gamma + \delta$.

In each case, $\pi_c(s^*_c, s^*_c) > \pi_c(r_c, s^*_c)$. (Here we use the fact that $\gamma + \zeta > 2\delta$)

There is a single history $v^* \in V^2$ so that $s^*_c \in S_c(v^*)$ but $s^*_c \notin S_{-c}(v^*)$, namely $v^* = (\phi, (C, C))$. Let $r^*_c$ be such that (i) $r^*_c(\phi) = C$, (ii) for each $v \in V^2$, $r^*_c(v) = C$ if and only if $v = (\phi, (C, \cdot))$, and (iii) for each $v \in V^3$, $r^*_c(v) = C$ if and only if $v = (\phi, (\cdot, \cdot, (C, \cdot)))$. Note, $\pi_c(s^*_c, r^*_c) = 2\delta + \zeta$. Fix $r_c \in \widehat{EFR}_{-c} \cap S_c(v^*)$ with $\zeta(r_c, r^*_c) \neq \zeta(s^*_c, s^*_c)$. Then $r_c(v^*) = D$ and so $\pi_c(r_c, r^*_c) = \gamma + \delta < \pi_c(s^*_c, r^*_c)$.

Remark D.5. Fix $s^*_c(\phi) = C$, $s^*_c(\phi, (C, C)) = C$ and $s^*_c(\phi, (C, D)) = D$. Suppose $\gamma + \zeta = 2\delta$. Then the proof of Lemma D.1(iv) establishes the following: Then there exists some CPS $p_c$ and some $s^*_c \in \widehat{EFR}_{-c}$ so that (i) $[s^*_c] \subseteq \mathbb{B}[p_c]$, and (ii) $p_c$ strongly believes $[s^*_c]$. 

34
**Proof of Lemma D.2(i).** Construct a strategy \( s^*_c \) with \( s^*_c(\phi) = C \) and, for each \( v \in V^2 \cup V^3 \), \( s^*_c(v) = C \) if and only if \( v = (\cdot, (\alpha_{c}^{-1}, \alpha_{c}^{-1}^{-1})) \) with \( \alpha_{c}^{-1} = C \). Observe that \( \pi_c(s^*_c, s^*_c) = 2\kappa + \zeta \). Fix some \( r_c \in \overline{\text{EFR}_c^1} \) with \( \zeta(r_c, s^*_c) \neq \zeta(s^*_c, s^*_c) \). There are three possible cases:

- \( r_c(\phi) = D \) and \( r_c(\phi, (D, C)) = C \): \( \pi_c(r_c, s^*_c) = 2\zeta + \gamma \),

- \( r_c(\phi) = D \) and \( r_c(\phi, (D, C)) = D \): \( \pi_c(r_c, s^*_c) = \zeta + 2\delta \), and

- \( r_c(\phi) = C \) and \( r_c(\phi, (C, C)) = D \): \( \pi_c(r_c, s^*_c) = \kappa + \zeta + \delta \).

In each case, \( \pi_c(s^*_c, s^*_c) > \pi_c(r_c, s^*_c) \). (Here we use the assumption that \( 2\kappa > \zeta + \gamma \).

There is a single history \( v^* \in V^2 \) so that \( s^*_c \in S_c(v^*) \) but \( s^*_c \notin S_{-c}(v^*) \), namely \( v^* = (\phi, (C, D)) \). We will choose distinct strategies \( r^*_c \) based on whether \( s^*_c(v^*) = C \) or \( s^*_c(v^*) = D \).

First, suppose \( s_c(v^*) = C \). Let \( r^*_c \) be a strategy so that \( r^*_c(\phi) = D \) and, for each \( v \in V^2 \cup V^3 \), \( r^*_c(v) = s^*_c(v) \). Then, \( \pi_c(s^*_c, r^*_c) = \gamma + \kappa + \zeta \). Fix \( r_c \in \overline{\text{EFR}_c^1} \cap S_c(v^*) \) with \( \zeta(r_c, r^*_c) \neq \zeta(s^*_c, r^*_c) \). Then \( r_c(v^*) = D \) and so \( \pi_c(r_c, r^*_c) = 2\zeta + \delta < \pi_c(s^*_c, r^*_c) \).

Second, suppose \( s_c(v^*) = D \). Let \( r^*_c \) be a strategy so that, for each \( v \in V \), \( r^*_c(v) = D \). Then, \( \pi_c(s^*_c, r^*_c) = \gamma + 2\delta \). Fix \( r_c \in \overline{\text{EFR}_c^1} \cap S_c(v^*) \) with \( \zeta(r_c, r^*_c) \neq \zeta(s^*_c, r^*_c) \). Then \( r_c(v^*) = C \) and so \( \pi_c(r_c, r^*_c) = 2\gamma + \delta < \pi_c(s^*_c, r^*_c) \).  

**Proof of Lemma D.2(ii).** Construct a strategy \( s^*_c \) with (i) \( s^*_c(\phi) = C \), (ii) \( s^*_c(\phi, (\alpha_{c}^{-1}, \cdot)) = C \) if and only if \( \alpha_{c}^{-1} = D \), and (iii) \( s^*_c(\cdot, (\alpha_{c}^{-1}^{-1}, \cdot)) = C \) if and only if \( \alpha_{c}^{-1} = C \). Observe that \( \pi_c(s^*_c, s^*_c) = 2\zeta + \kappa \). Fix some \( r_c \in \overline{\text{EFR}_c^1} \) with \( \zeta(r_c, s^*_c) \neq \zeta(s^*_c, s^*_c) \). There are three possible cases:

- \( r_c(\phi) = D \) and \( r_c(\phi, (D, C)) = D \): \( \pi_c(r_c, s^*_c) = 2\zeta + \delta \),

- \( r_c(\phi) = C \) and \( r_c(\phi, (C, C)) = C \): \( \pi_c(r_c, s^*_c) = \kappa + \gamma + \zeta \), and

- \( r_c(\phi) = C \) and \( r_c(\phi, (C, C)) = D \): \( \pi_c(r_c, s^*_c) = 2\kappa + \zeta \).

In each case, \( \pi_c(s^*_c, s^*_c) > \pi_c(r_c, s^*_c) \).

There is a single history \( v^* \in V^2 \) so that \( s^*_c \in S_c(v^*) \) but \( s^*_c \notin S_{-c}(v^*) \), namely \( v^* = (\phi, (D, D)) \). Let \( r^*_c \) be a strategy so that (i) \( r^*_c(\phi) = D \), (ii) \( r^*_c(\phi, (\alpha_{c}^{-1}, \cdot)) = C \) if and only if \( \alpha_{c}^{-1} = D \), and (iii) \( r^*_c(\cdot, (\alpha_{c}^{-1}, \cdot)) = C \) if and only if \( \alpha_{c}^{-1} = C \). Then, \( \pi_c(s^*_c, r^*_c) = \delta + \kappa + \zeta \). Fix \( r_c \in \overline{\text{EFR}_c^1} \cap S_c(v^*) \) with \( \zeta(r_c, r^*_c) \neq \zeta(s^*_c, r^*_c) \). Then \( r_c(v^*) = D \) and so \( \pi_c(r_c, r^*_c) = 2\delta + \zeta < \pi_c(s^*_c, r^*_c) \).  

**D.1.2 m-BRS’s**

Corollary D.1 establishes part (i) of Proposition 7.1: \( S^1_c = \overline{\text{EFR}_c^1} = \overline{\text{FR}_c^1} \). To show part (ii), it will be useful to have the following Lemmata.

**Lemma D.3.** Fix a strategy \( s_c \in \text{EFR}_c^1 \) so that

(i) \( D \in s_c(\{(\phi, (s_c(\phi), C)), (\phi, (s_c(\phi), D))\}) \) and

(ii) if \( 2\delta > \zeta + \gamma \), then \( s_c(\phi) = C \) implies \( s_c(\phi, (C, C)) = D \).

Then there exists some 2-BRS \((Q^0, Q^1, Q^2)\) so that \( s_c \in Q^2 \subseteq S^2_c \).
Proof. Fix a strategy $s_c$ as in the statement of the Lemma. By Corollary D.1, $s_c \in \overline{EFR}_c^1$. Moreover, by Lemma D.1, Remark D.5, and Corollary D.1, there exists some $s_{-c} \in \overline{EFR}_{-c}^1 = \overline{EFR}_{-c}^1$ and a CPS $p_c$ that strongly believes $[s_{-c}]$ so that $[s_c] \subseteq \mathbb{B}[p_c]$. Construct $Q^1_c = Q^2_c = \mathbb{B}[p_c]$, $Q^1_{-c} = [s_{-c}]$, and $Q^2_{-c} = \emptyset$. Applying Lemmata D.1-D.2, $(S,Q^1)$ is a 1-BRS. As such, $(S,Q^1,Q^2)$ is a 2-BRS with $s_c \in Q^2_c \subseteq S^2_c$.

Lemma D.4. Fix an $(m+1)$-BRS $(Q^0,\ldots,Q^m,Q^{m+1})$, where $m \geq 1$. If $v \in V^2$ with $Q^m_{-c} \cap S_{-c}(v) \neq \emptyset$ then, for each $s_c \in Q^{m+1}_c \cap S_c(v)$, $s_c(v) = D$.

Proof. Fix $v \in V^2$ with $Q^m_{-c} \cap S_{-c}(v) \neq \emptyset$ and $s_c \in Q^{m+1}_c \cap S_c(v)$. Then there exists $\mu \in \Delta(S_c(v))$ with (i) $\pi(s_c,\mu) \geq \pi(r_c,\mu)$ for each $r_c \in S_c(v)$, and (ii) $\mu(Q^m_{-c}) = 1$. Note, since $m \geq 1$, each $s_{-c} \in Q^m_{-c} \cap S_{-c}(v)$ has $s_{-c}(v') = D$ for all $v' \in V^3$ with $s_{-c} \in S_{-c}(v')$. Thus, $s_c(v) = D$. □

Proof of Proposition 7.1(ii). Lemma D.3 implies that every strategy specified in the result is contained in $S^2_c$. So we focus on the converse.

Fix a strategy $s_c \in \overline{EFR}_c^1 = \overline{EFR}_c^1$. Suppose there is a 2-BRS, viz. $(Q^0,Q^1,Q^2)$, with $s_c \in Q^2_c$. This implies that there exists some CPS $p_c$ so that $s_c \in \mathbb{B}[p_c]$ and $p_c$ strongly believes $Q^2_{-c}$. Thus, $Q^2_{-c} \neq \emptyset$. As such, there is some $v \in V^2$ with $s_c \in S_c(v)$ and $Q^1_{-c} \cap S_{-c}(v) \neq \emptyset$. By Lemma D.4, $s_c(v) = D$. So, either $s_c(\phi,(s_c,\phi),(C)) = D$ or $s_c(\phi,(s_c,\phi),(D)) = D$.

For the remainder of the proof, suppose that $s_c(\phi) = C$, $s_c(\phi,(C,C)) = C$, and $s_c(\phi,(C,D)) = D$. It suffices to show that, for each $s_{-c} \in Q^1_{-c}$, $s_{-c}(\phi) = D$. If so, then

$$\sum_{s_{-c} \in S_{-c}} \pi_c(s_{-c},s_{-c})p(\{s_{-c}\} \mid S_{-c}) \leq \gamma + \zeta + \delta.$$  

But there is an alternate strategy $r_c$ with

$$\sum_{s_{-c} \in S_{-c}} \pi_c(r_{-c},s_{-c})p(\{s_{-c}\} \mid S_{-c}) \geq 3\delta.$$  

(In particular, take $r_c(v) = D$ for all $v \in V$.) Since $s_c \in \mathbb{B}[p_c]$, $\zeta + \gamma \geq 2\delta$.

We now show that, for each $s_{-c} \in Q^1_{-c}$, $s_{-c}(\phi) = D$: Since $s_c$ is a best response under $p_c(\cdot \mid S_{-c}(\phi,(C,C)))$, there must be some $s_{-c}$ and $v \in V^3$ with $p_c(\{s_{-c}\} \mid S_{-c}(\phi,(C,C))) > 0$, $s_{-c}(v) = C$, and $s_{-c} \in S_{-c}(v)$. This implies that $p_c(Q^1_{-c} \mid S_{-c}(\phi,(C,C))) < 1$. Given that $p_c$ strongly believes $Q^1_{-c}$, $Q^1_{-c} \cap S_{-c}(\phi,(C,C)) = \emptyset$; that is, each $s_{-c} \in Q^1_{-c}$ specifies $s_{-c}(\phi) = D$. □

Lemma D.5. Fix a strategy $s_c \in S^2_c$ with $s_c(\phi) = D$ and $s_c(\phi,(D,D)) = D$. Then, for each $m$, $s_c \in S^m_c$.

Proof. Let $s_{-c}$ be such that $s_{-c}(v) = D$ for all $v \in V$. Lemma D.1 and Remark D.3 imply that $[s_c] \times [s_{-c}]$ is an EFBRS. Define $Q^1 = \cdots = Q^m = [s_c] \times [s_{-c}]$ and note that $(S,Q^1,\ldots,Q^m)$ is a $m$-BRS. □

Proof of Proposition 7.1(iii). Fix a 3-BRS $(Q^0,Q^1,Q^2)$ and some $s_c \in Q^3$. Then there exists some CPS $p_c$ so that $[s_c] \subseteq \mathbb{B}[p_c]$ and $p_c$ strongly believes $Q^2_{-c}$. Suppose, contra hypothesis, that $s_c(\phi) = C$. Then $Q^2_c \neq \emptyset$. Moreover, for each $s_{-c} \in Q^2_{-c}$ and each $v \in V^2 \cup V^3$ with $(s_c,s_{-c}) \in S(v)$, $s_c(v) = s_{-c}(v) = D$. (See Lemma D.4.) So, for each $s_{-c} \in Q^2_{-c}$, either $\pi_c(s_c,s_{-c}) = \kappa + 2\delta$ or $\pi_c(s_c,s_{-c}) = \gamma + 2\delta$. 36
Since $p(Q_2^2 \mid S_{-c}) = 1$, there exists some $q \in [0, 1]$ with

$$\sum_{s_{-c} \in S_{-c}} \pi_c(s_c, s_{-c}) p(\{s_{-c}\} \mid S_{-c}) = q \kappa + (1 - q) \gamma + 2\delta.$$ 

Consider an alternate strategy $r_c$ with $r_c(v) = D$ for each $v \in V$. Observe that

$$\sum_{s_{-c} \in S_{-c}} \pi_c(r_c, s_{-c}) p(\{s_{-c}\} \mid S_{-c}) \geq q \zeta + (1 - q) \delta + 2\delta.$$

This contradicts $[s_c] \subseteq \mathbb{B}[p_c]$.

Now, fix some $s_c \in S_c^2$ with $s_c(\phi) = D$. We will show that there exists a 3-BRS $(Q^0, Q^1, Q^2, Q^3)$ with $s_c \in Q_n^2$. If $s_c(\phi, (D, D)) = D$, this follows from Lemma D.5. So suppose that $s_c(\phi, (D, D)) = C$. It will be convenient to define strategies $r_c$ and $s_{-c}$. In particular, take $r_c \in \overline{EFR}_c$ so that (i) $r_c(\phi) = C$, (ii) $r_c(\phi, (C, C)) = C$, and (iii) $r_c(\phi, (C, D)) = D$. Take $s_{-c} \in \overline{EFR}_{-c}$ so that (i) $s_{-c}(\phi) = C$ and (ii) for each $v \in V^2 \cup V^3$, $s_{-c}(v) = D$. Set $Q^1_c = [s_c] \cup \{r_c\}$ and $Q^2_c = Q^3_c = [s_c]$. Set $Q^1_{-c} = Q^2_{-c} = [s_{-c}]$ and $Q^3_{-c} = \emptyset$. By Lemmata D.1-D.2 and Remark D.2, $(S, Q^1, Q^2, Q^3)$ is a 3-BRS.

**Proof of Proposition 7.1(iv).** Immediate from Lemmata D.4-D.5. ■

### D.2 Centipede

Throughout this sub-appendix, fix an $m$-BRS $(Q^0, Q^1, \ldots, Q^m)$ of the Centipede game. We will show that $Q^m_n \subseteq \overline{EFR}_n^m$. We begin with the following observation:

**Observation D.1.** Observe that $[\text{in}]_\ell \cap Q^1_\ell = \emptyset$ and so $Q^1_\ell \times Q^m_{-\ell} \subseteq \overline{EFR}_\ell^1 \times \overline{EFR}_{-\ell}^1$.

**Lemma D.6.** One of the following must hold:

(i) $[\text{in}]_{-\ell} \cap Q^2_{-\ell} = \emptyset$, or

(ii) $[\text{out}, |V|]_\ell \cap Q^1_\ell = \emptyset$ and $|V| = 3$.

**Proof.** First, suppose that $[\text{out}, |V|]_\ell \subseteq Q^1_\ell$. In that case, any CPS strongly believes $Q^1_\ell$ must assign probability 1 to $[\text{out}, |V|]_\ell$ at node $|V| - 1$. (This uses Observation D.1, i.e., the fact that $[\text{in}]_\ell \cap Q^1_\ell = \emptyset$.) Thus, $[\text{in}]_{-\ell}$ is not a sequential best response at node $|V| - 1$. From this $[\text{in}]_{-\ell} \cap Q^2_{-\ell} = \emptyset$.

Second, suppose that $[\text{out}, |V|]_\ell \cap Q^1_\ell = \emptyset$. Let $p_{-\ell}$ be a CPS that strongly believes $Q^1_\ell$ and note that $p_{-\ell}(\cdot | S_\ell)$ must assign probability 1 to

$$\{s_\ell : s_\ell(v) = \text{out}_v \text{ for some } v \leq |V| - 2\}.$$ 

(That is, *ex ante*, $p_{-\ell}$ assigns probability 1 to the game ending at some node $v \leq |V| - 2$, independent of the strategy that $-\ell$ plays.) If $|V| \geq 4$, then there is some node $\tilde{v} \leq |V| - 3$ at which $-\ell$ moves and $p_{-\ell}(\text{out}, \tilde{v} + 1| | S_\ell(\tilde{v})) = 1$. Thus, at node $\tilde{v}$, $[\text{out}, \tilde{v}]_{-\ell}$ is a unique best response. So certainly $[\text{in}]_{-\ell} \cap Q^2_{-\ell} = \emptyset$.

**Lemma D.7.** Fix some $m = 3, \ldots, |V| - 1$. If $m$ is odd then either

(i) $[\text{out}, |V| + 3 - m]_\ell \cap Q^m_\ell = \emptyset$, or
(ii) \([\text{out}, |V| + 2 - m]_{-\ell} \cap Q_{\ell}^{m-1} = \emptyset\) and \(|V| \leq m + 1\).

And, if \(m\) is even then either

(i) \([\text{out}, |V| + 3 - m]_{-\ell} \cap Q_{\ell}^m = \emptyset\), or

(ii) \([\text{out}, |V| + 2 - m]_{-\ell} \cap Q_{\ell}^{m-1} = \emptyset\) and \(|V| \leq m + 1\).

**Proof.** We show the base cases of \(m = 3, 4\). The inductive step simply repeats those arguments up to relabelling. Note, since \(|V| - 1 \geq m \geq 3, |V| \geq 4\). So, by Lemma D.6, \([\text{in}]_{-\ell} \cap Q_{\ell}^2 = \emptyset\). We repeatedly use this fact below.

\(m = 3\): Throughout, we suppose that \([\text{out}, |V|]_{\ell} \subseteq Q_{\ell}^1\). (If not, then we are done.) From this, Lemma D.6 gives that \([\text{in}]_{-\ell} \cap Q_{\ell}^2 = \emptyset\). We divide the argument into two cases.

First, suppose that \([\text{out}, |V| - 1]_{-\ell} \subseteq Q_{\ell}^2\). In that case, any CPS strongly believes \(Q_{\ell}^2\) must assign probability 1 to \([\text{out}, |V| - 1]_{-\ell}\) at node \(|V| - 2\). (This uses the fact that \([\text{in}]_{-\ell} \cap Q_{\ell}^2 = \emptyset\).) Thus, \([\text{out}, |V|]_{\ell}\) is not a best response at node \(|V| - 2\). From this \([\text{out}, |V|]_{\ell} \cap Q_{\ell}^3 = \emptyset\).

Second, suppose that \([\text{out}, |V| - 1]_{-\ell} \cap Q_{\ell}^2 = \emptyset\). Thus,

\([\text{out}, |V| - 1]_{-\ell} \cup [\text{in}]_{-\ell} \cap Q_{\ell}^2 = \emptyset\).

So, any CPS \(p_{\ell}\) that strongly believes \(Q_{\ell}^2\) must have

\(p_{\ell}(\{s_{-\ell} : s_{-\ell}(v) = \text{out}_v\text{ for some } v \leq |V| - 3\} | S_{-\ell}) = 1\).

(That is, \(\text{ex ante}\), \(p_{\ell}\) assigns probability 1 to the game ending at some node \(v \leq |V| - 3\), independent of the strategy that \(\ell\) plays.) If \(|V| \geq 5\), then there is some node \(\hat{v} \leq |V| - 4\) at which \(\ell\) moves and \(p_{\ell}([\text{out}, \hat{v} + 1]_{-\ell} | S_{-\ell}(\hat{v})) = 1\). Thus, at node \(\hat{v}\), \([\text{out}, \hat{v}]_{\ell}\) is a unique best response. So certainly \([\text{out}, |V|]_{\ell} \cap Q_{\ell}^3 = \emptyset\).

\(m = 4\): Throughout, we suppose that \([\text{out}, |V| - 1]_{-\ell} \subseteq Q_{\ell}^2\). (If not, then we are done.) From this, the base case of \(m = 3\) gives that \([\text{out}, |V|]_{\ell} \cap Q_{\ell}^3 = \emptyset\). We divide the argument into two cases.

First, suppose that \([\text{out}, |V| - 2]_{\ell} \subseteq Q_{\ell}^3\). In that case, any CPS strongly believes \(Q_{\ell}^3\) must assign probability 1 to \([\text{out}, |V| - 2]_{\ell}\) at node \(|V| - 3\). (This uses the fact that \([\text{out}, |V|]_{\ell} \cup [\text{in}]_{\ell} \cap Q_{\ell}^3 = \emptyset\).) Thus, \([\text{out}, |V| - 1]_{-\ell}\) is not a best response at node \(|V| - 3\). From this \([\text{out}, |V| - 1]_{-\ell} \cap Q_{\ell}^4 = \emptyset\).

Second, suppose that \([\text{out}, |V| - 2]_{\ell} \cap Q_{\ell}^3 = \emptyset\). Thus,

\([\text{out}, |V| - 2]_{\ell} \cup [\text{out}, |V|]_{\ell} \cup [\text{in}]_{\ell} \cap Q_{\ell}^3 = \emptyset\).

So, any CPS \(p_{-\ell}\) that strongly believes \(Q_{\ell}^4\) must have

\(p_{-\ell}(\{s_{-\ell} : s_{-\ell}(v) = \text{out}_v\text{ for some } v \leq |V| - 4\} | S_{-\ell}) = 1\).

(That is, \(\text{ex ante}\), \(p_{-\ell}\) assigns probability 1 to the game ending at some node \(v \leq |V| - 4\), independent of the strategy that \(-\ell\) plays.) If \(|V| \geq 6\), then there is some node \(\hat{v} \leq |V| - 5\) at which \(-\ell\) moves and \(p_{-\ell}([\text{out}, \hat{v} + 1]_{-\ell} | S_{-\ell}(\hat{v})) = 1\). Thus, at node \(\hat{v}\), \([\text{out}, \hat{v}]_{-\ell}\) is a unique best response. So certainly \([\text{out}, |V| - 2]_{-\ell} \cap Q_{\ell}^4 = \emptyset\). \(\blacksquare\)
Corollary D.2. If $|V| = m$, then either $Q^{[V]}_a = [\text{out}, 1]_a$ or $Q^{[V]}_a = \emptyset$.

Proof. We show the result for $|V|$ odd. (The case of $|V|$ even is analogous.) If $Q^{[V]-2}_a \not\in \{[\text{out}, 1]_a, \emptyset\}$, the claim is immediate. So suppose otherwise. By Observation D.1, $[\text{in}]_a \not\in Q^{[V]-2}_a$. By Lemma D.6, for each $m \leq |V| - 2$ odd, $[\text{out}, |V| + 3 - m]_a \cap Q^{[V]-2}_a = \emptyset$. So,

$$Q^{[V]-2}_a \in \{[\text{out}, 1]_a \cup [\text{out}, 3]_a, [\text{out}, 3]_a\}.$$ 

In either of these cases, $Q^{[V]-1}_b \in \{[\text{out}, 2]_b, \emptyset\}$. From this, it follows that $Q^{[V]}_a \in \{[\text{out}, 1]_a, \emptyset\}$. ■

Appendix E  Proofs for Section 8

E.1 Canonical CPS

We begin with a mathematical step useful in several results below. Given a strategy $s^*_c$ and an array $p_c \in A(S_{-c}, E_c)$ with $s^*_c \in \mathbb{BR}[p_c]$, we can construct a canonical CPS. Under that CPS, $s^*_c$ remains a sequential best response. Moreover, the CPS preserves strong belief.

Fix a strategy $s^*_c$ and some array $p_c \in A(S_{-c}, E_c)$ with $s^*_c \in \mathbb{BR}[p_c]$. We inductively construct the canonical CPS for $(s^*_c, p_c)$, viz. $q_c = (q_c(\cdot|S_{-c}(h)) : h \in H_c \cup \{\phi\})$ as follows: Let $H^0_c = H_c \cup \{\phi\}$. Choose $h^0 = \phi \in H^0_c$ and observe that $S_{-c}(\phi) = S_{-c}$. Set $q_c(\cdot|S_{-c}) = p_c(\cdot|S_{-c})$. Define $\overline{H}^0_c$ to be the set of $h \in H_c$ so that $S_{-c}(h) \subseteq S_{-c}$ and $q_c(S_{-c}(h)|S_{-c}) > 0$. For each $h \in \overline{H}^0_c$, set

$$q_c(s_{-c}|S_{-c}(h)) = \frac{q_c(s_{-c}|S_{-c})}{q_c(S_{-c}(h)|S_{-c})}$$

for all $s_{-c} \in S_{-c}(h)$. Note, $h^0 \in \overline{H}^0_c$.

Assume the sets $H^k_c$ and $\overline{H}^k_c$ have been defined. Set $H^{k+1}_c = H^k_c \setminus \overline{H}^k_c$. If $H^{k+1}_c = \emptyset$, then we are done. If not, choose some $h^{k+1} \in H^{k+1}_c$ that satisfies the following requirements:

(i) Either $s^*_c \in S_c(h^{k+1})$ or, for all $h \in H^{k+1}_c$, $s^*_c \not\in S_c(h)$.

(ii) There is no $h \in H^{k+1}_c$ so that $S_{-c}(h^{k+1}) \subseteq S_{-c}(h)$.

(iii) If $h \in H^{k+1}_c$ with $S_{-c}(h^{k+1}) = S_{-c}(h)$, then either $S_c(h) \subseteq S_c(h^{k+1})$ or $S_c(h) \cap S_c(h^{k+1}) = \emptyset$.

Set $q_c(\cdot|S_{-c}(h^{k+1})) = p_c(\cdot|S_{-c}(h^{k+1}))$. Define $\overline{H}^{k+1}_c$ to be the set of $h \in H^{k+1}_c$ so that $S_{-c}(h) \subseteq S_{-c}(h^{k+1})$ and $q_c(S_{-c}(h)|S_{-c}(h^{k+1})) > 0$. For each $h \in \overline{H}^{k+1}_c$, set

$$q_c(s_{-c}|S_{-c}(h)) = \frac{q_c(s_{-c}|S_{-c}(h))}{q_c(S_{-c}(h)|S_{-c}(h^{k+1}))}$$

for all $s_{-c} \in S_{-c}(h)$.

It might be useful to recap the construction: We begin by identifying information sets $h^0, h^1, \ldots, h^K$. Refer to these as basic information sets. (Note, they depend on both $p_c$ and $s^*_c$.) We set $q_c(\cdot|S_{-c}(h^k))$ to coincide with the original CPS $p_c(\cdot|h^k)$. For any non-basic information set $h$, there is exactly one basic information $h^k$ so that $S_{-c}(h) \subseteq S_{-c}(h^k)$ and $q_c(S_{-c}(h)|S_{-c}(h^k)) > 0$. Thus, we construct the belief $q_c(\cdot|S_{-c}(h))$ from $q_c(\cdot|S_{-c}(h^k))$ by conditioning on $S_{-c}(h)$. The construction obviously yields a CPS.
Lemma E.1. Fix a strategy $s^*_c$ and some array $p_c \in A(S_{-c}, E_c)$ with $s^*_c \in \mathbb{B}[q_c]$. Let $q_c \in C(S_{-c}, E_c)$ be the canonical CPS for $(s^*_c, p_c)$.

(i) $[s^*_c] \subseteq \mathbb{B}[q_c]$, and

(ii) if $p_c$ strongly believes $E_{-c}$, then $q_c$ strongly believes $E_{-c}$.

To prove Lemma E.1, it will be useful to have the following:

Lemma E.2. If $h \in \mathcal{T}_c^k$ and $s^*_c \in S_c(h)$, then $S(h) \subseteq S(h^k)$.

**Proof.** Fix $h \in \mathcal{T}_c^k$ with $s^*_c \in S_c(h)$. Then, by construction, $s^*_c \in S_c(h) \cap S_c(h^k) \neq \emptyset$. Suppose, contra hypothesis, that $S(h)$ is not contained in $S(h^k)$. By perfect recall, either $S(h^k) \subseteq S(h)$ or $S(h) \cap S(h^k) = \emptyset$. First, assume that $S(h^k) \subseteq S(h)$. Again employing perfect recall,

$S(h^k) = S_c(h^k) \times S_{-c}(h^k) \subseteq S_c(h) \times S_{-c}(h) = S(h)$.

Using the fact that $S_{-c}(h^k) \subseteq S_{-c}(h)$ and Property (ii) of the construction, $S_{-c}(h^k) = S_{-c}(h)$. So, $S_c(h^k) \subseteq S_c(h)$. But then, by Property (iii) of the construction, $S_c(h) \cap S_c(h^k) = \emptyset$, a contradiction. Second, assume that $S(h) \cap S(h^k) = \emptyset$. Since $h \in \mathcal{T}_c^k$, $\emptyset \neq S_c(h) \subseteq S_{-c}(h^k)$. It follows from Lemma A.2 that $S_c(h) \cap S_c(h^k) = \emptyset$, a contradiction. ■

**Proof of Lemma E.1.** First, we show that $s^*_c \in \mathbb{B}[q_c]$. (That implies $[s^*_c] \subseteq \mathbb{B}[q_c]$.) Toward that end, fix some $h \in H_c$ with $s^*_c \in S_c(h)$. Observe that there exist a $k$ so that $h \in \mathcal{T}_c^k$, i.e., there exists a basic $h^k$ so that $q_c(\cdot]\mid S_{-c}(h))$ is derived from $p_c(\cdot]\mid S_{-c}(h^k))$ by conditioning. (Note, $h$ may well be $h^k$.) By construction, $s^*_c$ is optimal under $q_c(\cdot]\mid S_{-c}(h^k))$ given all strategies in $S_c(h^k)$. It follows from Lemmata E.2-A.3 that $s^*_c$ is optimal under $q_c(\cdot]\mid S_{-c}(h))$ given all strategies in $S_c(h)$.

Second, we show that, if $p_c$ strongly believes $E_{-c}$, then $q_c$ strongly believes $E_{-c}$. Fix an information set $h \in H_c$ so that $E_{-c} \cap S_{-c}(h) \neq \emptyset$. There exists some $h^k \in H_c$ so that $S_{-c}(h) \subseteq S_{-c}(h^k)$, $p_c(S_{-c}(h)\mid S_{-c}(h^k)) > 0$ and, for every $s_{-c} \in S_{-c}(h)$,

$q_c(s_{-c}\mid S_{-c}(h)) = \frac{p_c(s_{-c}\mid S_{-c}(h^k))}{p_c(S_{-c}(h)\mid S_{-c}(h^k))}$.

Since $S_{-c}(h) \subseteq S_{-c}(h^k)$, $E_{-c} \cap S_{-c}(h^k) \neq \emptyset$. If $p_c$ strongly believes $E_{-c}$ then $p(E_{-c}\mid S_{-c}(h^k)) = 1$ and so $q(E_{-c}\mid S_{-c}(h)) = 1$. ■

### E.2 Generic Games: No Relevant Ties

We first observe that a game can satisfy NRT, even though it is non-generic. We then give two classes of NRT games (one a subclass of the other) that are generic.

**Example E.1.** The game in Figure E.1 satisfies no relevant ties. Yet it is not generic: $D$ is a sequential best response under $p_a$ if and only if $p_a(L|S_b) = p_a(R|S_b) = 1/2$. Thus, $\mathbb{B}[p_a] = \{U, M, D\}$ and there is no $q_a$ with $\mathbb{B}[q_a] = \{D\}$. ■

Note, in Example E.1, $D$ is justifiable, but not optimal under any CPS that involves point beliefs.
Definition E.1. Given a conditional probability space \((\Omega, \mathcal{E})\), call a CPS \(p \in \mathcal{C}(\Omega, \mathcal{E})\) degenerate if, for each conditioning event \(E\), there exists some \(\omega \in E\) with \(p(\omega|E) = 1\).

So, a CPS is degenerately justifiable if each conditional belief is a point belief.

Definition E.2. Call a game degenerately justifiable if, whenever \(s_c\) is justifiable, there exists some degenerate CPS \(p_c \in \mathcal{C}(S_{-c}, \mathcal{E}_c)\) so that \(s_c \in BR[p_c]\).

Example E.1 is not degenerately justifiable.

Proposition E.1. A degenerately justifiable game that satisfies NRT is generic.

Proof. Fix a degenerately justifiable game satisfying NRT and a justifiable strategy \(s_c\). Then there exists a degenerate CPS \(p_c \in \mathcal{C}(S_{-c}, \mathcal{E}_c)\) so that \(s_c \in BR[p_c]\). We will show that, if \(r_c \not\in [s_c]\), then \(r_c \not\in BR[p_c]\).

Fix some \(r_c \not\in [s_c]\). Then there exists some \(h \in H_c\) with \(s_c, r_c \in S_c(h)\) and \(s_c(h) \neq r_c(h)\). Let \(s_{-c} \in S_{-c}(h)\) with \(p_c(s_{-c}|S_{-c}(h)) = 1\). Since \(s_c\) is a sequential best response under \(p_c\), \(\pi_c(s_c, s_{-c}) \geq \pi_c(r_c, s_{-c})\).

But, since \(\zeta(s_c, s_{-c}) \neq \zeta(r_c, s_{-c})\), NRT implies \(\pi_c(s_c, s_{-c}) > \pi_c(r_c, s_{-c})\). Thus, \(r_c \not\in BR[p_c]\).

Lemma E.3. A perfect-information game satisfying NRT is degenerately justifiable.

Proof. Let \(s_c\) be a justifiable strategy. Then, by Lemma 1.2.1 in Ben Porath (1997), for each \(S_{-c}(h) \in \mathcal{E}_c\) with \(s_c \in S_c(h)\), we can find some \(s_{-c}^h \in S_{-c}(h)\) so that \(\pi_c(s_c, s_{-c}^h) \geq \pi_c(r_c, s_{-c}^h)\) for all \(r_c \in S_c(h)\). Use the collection \((s_{-c}^h : h \in H_c \cup \{\phi\})\) to form a degenerate array \(p_c\) with \(s_c \in BR[p_c]\). Then, the canonical CPS \(q_c\) is degenerate and, by Lemma E.1, \(s_c \in BR[q_c]\).

The following is now immediate from Proposition E.1 and Lemma E.3.

Proposition E.2. A perfect-information game satisfying no relevant ties is generic.

E.3 Generic Games: No Relevant Convexities

Given a strategy \(s_c^*\) and an array \(p_c \in \mathcal{A}(S_{-c}, \mathcal{E}_c)\) with \(s_c^* \in BR[p_c]\), we can construct a canonical CPS that preserves strong belief and, moreover, \(s_c^*\) remains a sequential best response. (See Lemma E.1.) We now show, if the game satisfies NRC, then we can choose the CPS so that the set of best responses is simply \([s_c^*]\). This proves Proposition 8.2 and establishes that a game that satisfies NRC is generic.

Say \((s_c^*, p_c)\) satisfies Property [*] if the following holds:
Property [*]: For each \( h \in H_c \) with \( s^*_c \in S_c(h) \), if \( r_c \in S_c(h) \) is optimal under \( p_c(\cdot|S_c(h)) \) among strategies in \( S_c(h) \), then \( \zeta(s_c, s_{-c}) = \zeta(r_c, s_{-c}) \) for all \( s_{-c} \in \text{Supp} p_c(\cdot|S_c(h)) \).

**Lemma E.4.** Fix a strategy \( s^*_c \) and some array \( p_c \in \mathcal{A}(S_{-c}, \mathcal{E}_c) \) with \( s^*_c \in \mathcal{B}[p_c] \). Suppose \( (s^*_c, p_c) \) satisfies Property [*]. Then, the canonical CPS for \( (s^*_c, p_c) \), viz. \( q_c \in \mathcal{C}(S_{-c}, \mathcal{E}_c) \), satisfies:

(i) \( [s^*_c] = \mathcal{B}[q_c] \), and

(ii) if \( p_c \) strongly believes \( E_{-c} \), then \( q_c \) strongly believes \( E_{-c} \).

**Proof.** By Lemma E.1, it suffices to show that \( \mathcal{B}[q_c] \subseteq [s^*_c] \). Fix some \( r_c \in \mathcal{B}[q_c] \backslash [s^*_c] \). Then there is an information set \( h \in H_c \) so that \( s^*_c, r_c \in S_c(h) \) and \( s^*_c(h) \neq r_c(h) \). Let \( k \) be such that \( h \in H^k_c \) and note that \( r_c \) is an optimal under \( q_c(\cdot|S_c(h^k)) = p_c(\cdot|S_c(h^k)) \) given \( S_c(h^k) \). Fix some \( s_{-c} \in S_{-c}(h) \subseteq S_{-c}(h^k) \) such that \( q_c(s_{-c}|S_c(h)) > 0 \). Observe that \( \zeta(s^*_c, s_{-c}) \neq \zeta(r_c, s_{-c}) \) and \( p_c(s_{-c}|S_c(h^k)) > 0 \). This contradicts the fact that \( (s^*_c, p_c) \) satisfies Property [*].

**Lemma E.5.** Suppose \( \Gamma \) satisfies NRC. Let \( s^*_c \) and \( p_c \in \mathcal{A}(S_{-c}, \mathcal{E}_c) \) be such that \( s^*_c \in \mathcal{B}[p_c] \). Then there exists an array \( \hat{p}_c \in \mathcal{A}(S_{-c}, \mathcal{E}_c) \) so that

(i) \( (s^*_c, \hat{p}_c) \) satisfies Property [*],

(ii) \( s^*_c \in \mathcal{B}[\hat{p}_c] \), and

(iii) \( p_c \) strongly believes \( E_{-c} \) if and only if \( \hat{p}_c \) strongly believes \( E_{-c} \).

**Proof.** For each \( h \in H_c \) with \( s^*_c \in S_c(h) \), we can choose \( \hat{p}_c(\cdot|S_c(h)) \) so that (a) \( r_c \in S_c(h) \) is optimal under \( \hat{p}_c(\cdot|S_c(h)) \) amongst all strategies in \( S_c(h) \) if and only if \( r_c \) supports \( s^*_c \) given \( \text{Supp} p_c(\cdot|S_c(h)) \), \( h \), and (b) \( \text{Supp} \hat{p}_c(\cdot|S_c(h)) = \text{Supp} p_c(\cdot|S_c(h)) \). (See Lemmata D.2-D.3-D.4 in Brandenburger, Friedenberg and Keisler, 2008.) Requirement (i) follows from the construction and NRC. Requirements (ii)-(iii) follow immediately from the construction.

**Proof of Proposition 8.2.** Immediate from Lemma E.4-E.5.

**Corollary E.1.** If a game satisfies NRC, then it is generic.

**Proof.** Fix some \( s_c \in \mathcal{B}[p_c] \) for some \( p_c \in \mathcal{C}(S_{-c}, \mathcal{E}_c) \). By Lemmata E.4-E.5, there exists some \( q_c \in \mathcal{C}(S_{-c}, \mathcal{E}_c) \) with \( [s_c] = \mathcal{B}[q_c] \).

One implication of Proposition 8.2 is that, when NRC is satisfied, we can forgo using CPS’s and focus on arrays. This would not be the case absent NRC. The central difficulty comes from condition (BRP.3) of the \( m \)-BRS. Specifically, begin with a decreasing sequence of product sets \( (Q^0, \ldots, Q^{m-1}, Q^m) \). In addition, suppose that \( s_c \in Q^m_c \) so that, for some array \( p_c \in \mathcal{A}(S_{-c}, \mathcal{E}_c) \), conditions (BRP.1)-(BRP.2)-(BRP.3) are satisfied. The canonical CPS \( q_c \in \mathcal{C}(S_{-c}, \mathcal{E}_c) \) will satisfy conditions (BRP.1)-(BRP.2). But, condition (BRP.3) may fail. (An example is available upon request.)

A second implication of Proposition 8.2 is that we can use a weaker maximality criteria. This need not hold in a perfect-information game satisfying NRT, despite the fact that such games are generic.

**Example E.2.** The game in Figure E.2 is a perfect-information game satisfying NRT.\(^{18}\) As such, it is generic. But, the conclusion of Proposition 8.2 does not hold. To see this, let \( (Q^0, Q^1, Q^2) \) be a decreasing

\(^{18}\)A three-player version appears in Battigalli (1997).
sequence of product sets, so that
\[ Q_1^a \times Q_1^b = \{O, L'R'', R'L'', R''R''\} \times \{o, lr', rr''\} \]

and
\[ Q_2^a \times Q_2^b = \{O\} \times \{o\}. \]

Observe \( Q_1^1 \) corresponds to the set of strategies that survive one round of EFR. Thus, \((Q^0, Q^1)\) is a 1-BRS.

We will argue that \((Q^0, Q^1, Q^2)\) satisfies the requirements of Proposition 8.2, but is not a 2-BRS.

Note, \( O \) is a unique sequential best response under a CPS that \textit{ex ante} assigns probability 1 to \( lr' \) and strongly believes \( Q_1^b \). An array of Bob that strongly believes \( Q_2^a \) must assign zero probability to \( L'R'' \) conditional upon Bob’s first information set being reached. Thus, \( o \) is a sequential best response under an array \( p_b \) that strongly believes \( Q_1^a \) if and only if, conditional upon Bob’s first information set being reached, the array assigns probability \( 2/3 : 1/3 \) to \( L'R'' : R'L'' \). So, \((Q^0, Q^1, Q^2)\) satisfies the requirements of Proposition 8.2. But, it is not a 2-BRS: \( \mathbb{BR}[p_b] = \{o, lr', rr''\} \) is not contained in \( Q_2^b \).

\[ \square \]

**E.4 Termination of the Procedure**

Call \( Q = (Q^0, Q^1, \ldots) \) a \textbf{BR-sequence} if, for each \( m, (Q^0, \ldots, Q^m) \) is an \( m \)-BRS.

**Proof of Proposition 8.1.** For each BR-sequence \( Q = (Q^0, Q^1, \ldots) \), there is a finite \( M(Q) \) so that, \( Q^{M(Q)} = Q^{M(Q)+1} \). Choose the lowest such \( M(Q) \) and note

\[ M(Q) \leq \begin{cases} 2 \min\{|S_a|, |S_b|\} - 1 & \text{if } |S_a| \neq |S_b|, \\ 2 \min\{|S_a|, |S_b|\} - 2 & \text{if } |S_a| = |S_b|. \end{cases} \]

Take \( M \) to be the maximum of all such \( M(Q) \) and observe that it, too, is less than or equal to \( 2 \min\{|S_a|, |S_b|\} - 43 \).
1 (resp. $2 \min\{|S_a|, |S_b|\} - 2$) if $|S_a| \neq |S_b|$ (resp. $|S_a| = |S_b|$).

Now note that $S^M = S^\infty$: Certainly $S^\infty \subseteq S^M$. Fix $s \in S^M$ and note that there exists some $Q = (Q^0, Q^1, \ldots)$ with $s \in Q^M$ and $Q^M \subseteq Q$ for some EFBRS $Q$. So $s \in Q \subseteq S^\infty$.  

E.5 Computing $m$-BRS’s

Suppose $\Gamma$ satisfies NRC. Proposition 8.2 offers an alternate approach to computing $m$-BRS’s, one that makes use of arrays. This allows us to use the simplex algorithm to search for appropriate beliefs.

Fix some $(Q^0, \ldots, Q^{m-1}, Q^m)$ where $(Q^0, \ldots, Q^{m-1})$ is an $(m - 1)$-BRS. Fix also some $h \in H_c$ and write $n(h) = \max\{n : Q^n_c \cap S_{-c}(h) \neq \emptyset\}$. Then, enumerate $Q^n_c(h) \cap S_{-c}(h) = \{s^1_{-c}, \ldots, s^K_{-c}\}$ and $S_c(h) = \{s^1_c, \ldots, s^L_c\}$.

Say a strategy $s_c \in Q^m_{-c}$ passes the test at $h$ if either $s_c \not\in S_c(h)$ or, there exists non-negative numbers $\mu^1, \ldots, \mu^K$ with $\sum_{k=1}^K \mu^k = 1$, so that $s_c$ maximizes $\sum_{k=1}^K \pi_c(s_c, s^k_{-c}) \mu^k$ among all strategies in $S_c(h) = \{s^1_c, \ldots, s^L_c\}$. A strategy $s_c$ passes the test if it passes the test at each $h \in H_c$.

The simplex algorithm can be used to determine if $s_c$ passes the test at $h$. Specifically, when $s_c \in S_c(h)$, the problem is equivalent to choosing $(\mu^1, \ldots, \mu^K, \tau^1, \ldots, \tau^L)$ to solve

$$\text{Maximize } \sum_{k=1}^K \pi_c(s_c, s^k_{-c}) \mu^k$$
$$\text{subject to } \sum_{k=1}^K [\pi_c(s_c, s^k_{-c}) - \pi_c(s^l, s^k_{-c})] \mu^k + \tau^l = 0 \text{ for each } l = 1, \ldots, L$$
$$\mu^1 + \mu^2 + \cdots + \mu^K = 1$$
$$(\mu^1, \ldots, \mu^K, \tau^1, \ldots, \tau^L) \geq (0, \ldots, 0).$$

We can apply the simplex algorithm to this linear programming problem. The algorithm terminates by either (a) concluding that there is no feasible solution, (b) providing an optimal solution, or (c) concluding that the objective function is unbounded over the feasible region. (See Chapter 2 in Bradley, Hax and Magnanti, 1977.) In the first scenario $s_c$ fails the test and in the latter two scenarios $s_c$ passes the test.

References


