AXIOMATIC EQUILIBRIUM SELECTION

SRIHARI GOVINDAN AND ROBERT WILSON

Abstract. We apply three axioms adapted from decision theory to refinements of the Nash equilibria of games with perfect recall that select connected closed subsets called solutions.

Undominated Strategies: No player uses a weakly dominated strategy in an equilibrium in a solution.

Backward Induction: Each solution contains a quasi-perfect equilibrium and thus a sequential equilibrium in strategies that provide conditionally admissible optimal continuations from information sets.

Small Worlds: A refinement is immune to embedding a game in a larger game with additional players provided the original players' strategies and payoffs are preserved, i.e. solutions of a game are the same as those induced by the solutions of any larger game in which it is embedded.

For games with two players and generic payoffs, we prove that these axioms characterize each solution as an essential component of equilibria in undominated strategies, and thus a stable set as defined by Mertens (1989).

Date: 23 April 2009.

Key words and phrases. game, equilibrium, refinement, axiom, admissibility, backward induction, small worlds, stability. JEL subject classification: C72.

This work was funded in part by a grant from the National Science Foundation of the United States.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>2</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>3</td>
</tr>
<tr>
<td>2. Notation</td>
<td>5</td>
</tr>
<tr>
<td>3. The Axioms</td>
<td>6</td>
</tr>
<tr>
<td>3.1. Undominated Strategies</td>
<td>6</td>
</tr>
<tr>
<td>3.2. Backward Induction</td>
<td>6</td>
</tr>
<tr>
<td>3.3. Small Worlds</td>
<td>7</td>
</tr>
<tr>
<td>3.4. Summary of the Axioms</td>
<td>9</td>
</tr>
<tr>
<td>4. Additional Notation and Properties</td>
<td>9</td>
</tr>
<tr>
<td>4.1. Payoffs</td>
<td>10</td>
</tr>
<tr>
<td>4.2. Notation for the Extensive Form</td>
<td>10</td>
</tr>
<tr>
<td>4.3. Enabling Strategies</td>
<td>10</td>
</tr>
<tr>
<td>4.4. Stable Sets of Equilibria</td>
<td>11</td>
</tr>
<tr>
<td>4.5. The Associated Pseudomanifold</td>
<td>12</td>
</tr>
<tr>
<td>5. Statement and Proof of the Theorem</td>
<td>13</td>
</tr>
<tr>
<td>5.1. Preliminaries</td>
<td>14</td>
</tr>
<tr>
<td>5.2. A Game with Redundant Strategies</td>
<td>14</td>
</tr>
<tr>
<td>5.3. Extensive Form of the Metagames</td>
<td>15</td>
</tr>
<tr>
<td>5.4. Outsiders' Payoffs in the Metagames</td>
<td>16</td>
</tr>
<tr>
<td>5.5. Outsiders' Strategies in a Quasi-Perfect Equilibrium</td>
<td>16</td>
</tr>
<tr>
<td>5.6. Limits of the Quasi-Perfect Equilibria of the Metagames</td>
<td>18</td>
</tr>
<tr>
<td>5.7. Insiders' Strategies in a Quasi-Perfect Equilibrium, and the induced Lexicographic Probability System.</td>
<td>19</td>
</tr>
<tr>
<td>5.8. Limits of the Lexicographic Probability Systems.</td>
<td>20</td>
</tr>
<tr>
<td>5.9. Final Step of the Proof</td>
<td>23</td>
</tr>
<tr>
<td>6. Concluding Remarks</td>
<td>23</td>
</tr>
<tr>
<td>Appendix A. Enabling Strategies</td>
<td>25</td>
</tr>
<tr>
<td>Appendix B. Proofs of Propositions</td>
<td>27</td>
</tr>
<tr>
<td>B.1. Proof of Proposition 3.3</td>
<td>27</td>
</tr>
<tr>
<td>B.2. Proof of Proposition 3.4</td>
<td>27</td>
</tr>
<tr>
<td>Appendix C. Proof of Theorem 4.1</td>
<td>28</td>
</tr>
<tr>
<td>References</td>
<td>42</td>
</tr>
</tbody>
</table>

## List of Figures

1. A game tree and its pure reduced normal form in which each pure strategy is identified by the terminal nodes it does not exclude. 25

2. The game tree of a signaling game and its pure reduced normal form in which each pure strategy is identified by the terminal nodes it does not exclude. 26
1. Introduction

Nash’s [21, 22] definition of equilibrium is insufficient when applied to a game in extensive form. Even for a game with perfect information, some equilibria use weakly dominated strategies, and some yield outcomes different from the outcome predicted by backward induction. The literature on refinements aims to sharpen Nash’s definition to exclude such equilibria—see [8, 11, 28] for recent surveys.

Kohlberg and Mertens [13] suggest that a refinement should be characterized by axioms adapted from decision theory. They also specify properties that axioms should imply. Subsequently, Mertens [18, 19, 20] defines the set-valued refinement called stability and shows that, among others, it has these properties:\(^1\)

1. Admissibility and Perfection. All equilibria in a stable set are perfect and therefore admissible, i.e. each player’s strategy in each equilibrium is undominated.

2. Backward Induction and Forward Induction. A stable set includes a proper equilibrium that induces a quasi-perfect (hence, sequential) equilibrium in every extensive-form game with perfect recall that has the same normal form. A subset of a stable set survives iterative elimination of weakly dominated strategies and strategies that are inferior replies at every equilibrium in the set.

3. Small Worlds and Decomposition. The stable sets of a game are the projections of the stable sets of any larger game in which it is embedded. The stable sets of the product of two independent games are the products of their stable sets.

For games in extensive form with perfect recall, we propose three axioms that are versions of these properties. We assume that for each game a refinement selects a collection of nonempty, connected, and closed subsets of equilibria, called solutions.\(^2\) The axioms are the following:

A. Undominated Strategies: No player uses a weakly dominated strategy in any equilibrium in a solution.

\(^1\)Mertens considers the graph of equilibria over the space of perturbed games obtained by perturbing players’ strategies toward mixed strategies, and then defines a connected closed set of equilibria as stable if the projection from a neighborhood in this graph is an essential map, i.e. has nonzero degree. See Mertens [18] for further topological aspects of the definition, and Govindan and Mertens [3] for an equivalent definition in terms of players’ best-reply correspondences.

\(^2\)Solutions are assumed to be sets because Kohlberg and Mertens [13, pp. 1015, 1019, 1029] show that there need not exist a single equilibrium satisfying weaker properties than the axioms invoked here. The technical requirement that a solution is connected excludes the trivial refinement that selects all equilibria. If only a single (possibly unconnected) subset is selected then only the trivial refinement satisfies the conditions invoked by Norde, Potters, Reijnierse, and Vermeulen [23]. Kreps and Wilson [15, Theorem 2] show that if payoffs are generic then all equilibria in a connected set yield the same probability distribution over terminal nodes, and thus the same paths of equilibrium play in the extensive form.
B. Backward Induction: A solution contains a quasi-perfect equilibrium.

S. Small Worlds: A solution is immune to embedding the game within larger games that preserve players’ strategies and payoffs.

In Axiom A we invoke only the implication of admissibility that no player uses a weakly dominated strategy. Axiom A is equivalent to admissibility in two-player games. Axiom B insists on inclusion of a quasi-perfect equilibrium because it induces a sequential equilibrium in strategies that provide admissible optimal continuations from players’ information sets.\(^3\) Axiom S ensures that a refinement is not vulnerable to framing effects depending on how the game is presented within wider contexts called metagames. A metagame can include additional players and additional pure strategies, provided these additional features do not alter optimal decisions in the original game. Our version of small worlds is stronger than Mertens’ [20] version because our definition of embedding allows a more general class of metagames.

We prove that these three axioms characterize refinements that select some or all of the stable sets of a game in extensive form with perfect recall, two players, and generic payoffs. Our main theorem shows that the axioms require each solution to be an essential component of admissible equilibria, it is a component only in enabling strategies and which in this case is the defining property of a stable set.\(^4\) For such a game, therefore, axioms A, B, S imply the seemingly stronger properties 1, 2, 3.

When payoffs are generic, all equilibria in a component yield the same probability distribution over outcomes. For economic modeling and econometric studies, therefore, the axioms’ chief implication is that a predicted outcome distribution should result from an essential component of the game’s admissible equilibria. A secondary implication is that a solution must include all equilibria in the component. This is germane for predicting players’ behaviors only after deviations from equilibrium play, but it addresses the issue of whether sequential rationality is a relevant decision-theoretic criterion after deviations (Reny [25]). We show that each equilibrium in a solution is induced by a quasi-perfect equilibrium in a solution of some metagame; therefore, it is sequentially rational when viewed in the wider context of this metagame. See [7, Section 2.3] for an example.

Section 2 establishes notation for Section 3, which specifies Axioms A (undominated strategies), B (backward induction), and S (small worlds), including a precise definition of embedding a game in a larger game. Section 4 establishes further notation and states a key

---

\(^3\)We intend that a revised version will state Axiom B in this form.

\(^4\)A component is a maximal closed connected set, and it is essential if it has the property described in footnote 1.
technical property proved in Appendix C. Section 5 states and proves the main theorem. The proof is constructive in that each equilibrium in a stable set is shown to be induced by a quasi-perfect equilibrium in a solution of a particular metagame with perfect recall (but nongeneric payoffs) that embeds the given game. Section 6 interprets this result and provides concluding remarks.

2. Notation

This section provides sufficient notation for statements of the axioms in Section 3. Section 4 introduces notation needed for the main theorem in Section 5.

A typical game in extensive form is denoted $\Gamma$. Its specification includes a set $N$ of players, a game tree with perfect recall for each player, and an assignment of players’ payoffs at each terminal node of the tree. Let $H_n$ be player $n$’s collection of information sets, and let $A_n(h)$ be his set of feasible actions at information set $h \in H_n$. The specification of the tree can include a completely mixed strategy of Nature.

A player’s pure strategy chooses an action at each of his information sets. Denote $n$’s simplex of mixed strategies by $\Sigma_n$ and interpret its vertices as his set $S_n$ of pure strategies. The sets of profiles of players’ pure and mixed strategies are $S = \prod_n S_n$ and $\Sigma = \prod_n \Sigma_n$. The normal form of $\Gamma$ assigns to each profile of players’ pure strategies the profile of their expected payoffs; equivalently, it is the multilinear function $G : \Sigma \rightarrow \mathbb{R}^N$ that assigns to each profile of their mixed strategies the profile of their expected payoffs.

A player’s behavioral strategy specifies a mixture over his actions at each of his information sets. Let $B_n$ be $n$’s set of behavioral strategies, and let $B = \prod_n B_n$ be the set of profiles of players’ behavioral strategies. Each mixed strategy $\sigma_n$ induces a behavioral strategy $b_n$. Because the game has perfect recall, for each behavioral profile there are profiles of mixed strategies that induce it and yield the same distribution of outcomes (Kuhn [16]).

As defined by Nash [21, 22], an equilibrium is a profile of players’ strategies such that each player’s strategy is an optimal reply to other players’ strategies. That is, if $\text{BR}_n(\sigma) \equiv \arg\max_{\sigma_n \in \Sigma_n} G_n(\sigma_n', \sigma_{-n})$ is player $n$’s best-reply correspondence, then $\sigma \in \Sigma$ is an equilibrium iff $\sigma_n \in \text{BR}_n(\sigma)$ for every player $n$. The analogous definition of equilibrium in behavioral strategies is equivalent for games with perfect recall.

A refinement is a correspondence that assigns to each game a nonempty collection of nonempty connected closed subsets of its equilibria. Each selected subset is called a solution. According to the above definitions, a mixed or behavioral strategy makes choices even at information sets that its previous choices exclude from being reached. In Section 4 we consider pure strategies to be equivalent if they make the same choices at information sets.
they do not exclude. And, we further simplify mixed and behavioral strategies by considering only their induced probability distributions on terminal nodes. The definitions of an equilibrium and a refinement have equivalent statements in terms of these strategy spaces. Each equilibrium in a reduced strategy space corresponds to a set of equilibria as defined above, and analogously for solutions selected by a refinement. The axioms in Section 3 are stated in terms of mixed and behavioral strategies as defined above. One reason for this is that Axiom S implies invariance to redundant strategies, and thus justifies our later use of equivalence classes of strategies.

3. The Axioms

3.1. Undominated Strategies. The first axiom requires that no player uses a weakly dominated strategy. Say that a profile of players’ strategies is undominated if each player’s strategy is undominated.

**Axiom A** (Undominated Strategies): Each equilibrium in a solution is undominated.

3.2. Backward Induction. We interpret sequential equilibrium as the generalization to games with perfect recall of backward induction in games with perfect information, and to be consistent with Axiom A we insist on conditionally admissible optimal continuations from information sets. Here we obtain these properties from a quasi-perfect equilibrium, which requires a consistency property of beliefs that is stronger than required by Kreps and Wilson’s [15] definition of sequential equilibrium. Van Damme [27, p. 8] defines a quasi-perfect equilibrium in terms of players’ behavioral strategies.

**Definition 3.1** (Quasi-Perfect Equilibrium). A profile \( b \in B \) of behavioral strategies is a quasi-perfect equilibrium if it is the limit of a sequence of profiles of completely mixed behavioral strategies for which, for each player \( n \), from each of his information sets, continuation of his strategy \( b_n \) is an optimal reply to every profile in the sequence.

That is, if \( BR_n(\cdot|h) \) is \( n \)'s best-reply correspondence in terms of behavioral strategies that continue from his information set \( h \in H_n \), and \( b_n(h) \) is the continuation of his behavioral strategy \( b_n \) from this information set, then the profile \( b \in B \) is quasi-perfect if

\[
(\forall k) \quad (\forall n \in N, h \in H_n) \quad b_n(h) \in BR_n(\hat{b}^k|h)
\]

for some sequence \( \hat{b}^k \in B \setminus \partial B \) for which \( \hat{b}^k \to b \).

If the mixed-strategy profile \( \sigma \in \Sigma \) induces a behavioral profile \( b \in B \) that is a quasi-perfect equilibrium then we say that \( \sigma \) too is quasi-perfect. Similarly, the justifying sequence \( \hat{b}^k \) can be represented by a sequence \( \hat{\sigma}^k \in \Sigma \setminus \partial \Sigma \) for which \( \hat{\sigma}^k \to \sigma \).
Van Damme shows that a quasi-perfect equilibrium induces a perfect equilibrium of the normal form, and a sequential equilibrium of the extensive form in which all players’ beliefs are induced by the same justifying sequence. Moreover, by construction a quasi-perfect equilibrium provides for each player an admissible optimal continuation from each of his information sets. It can be shown that if payoffs are generic then every sequential equilibrium is quasi-perfect and extensive-form perfect.

The second axiom requires that some equilibrium in a solution is quasi-perfect.

**Axiom B** (Backward Induction): Each solution contains a quasi-perfect equilibrium.

If payoffs are generic then Axioms A and B imply that each solution lies in a component of undominated equilibria, each of which yields the same distribution over outcomes as a sequential equilibrium in the solution.

### 3.3. Small Worlds.

The third axiom requires that a refinement is not affected by extraneous features of wider contexts in which a game is embedded, provided such contexts do not alter players’ feasible strategies and payoffs. An embedding allows the presence of additional players whose actions might provide the original players with additional pure strategies equivalent to mixed strategies in the original game, and thus redundant. For simplicity, to define an embedding we use the normal form $G : \Sigma \rightarrow \mathbb{R}^N$ of the extensive-form game $\Gamma$.

An embedding is described by a ‘larger’ game $\tilde{G} : \tilde{\Sigma} \times \Sigma_o \rightarrow \mathbb{R}^{N \cup o}$ in which game $G$ is ‘embedded,’ subject to certain restrictions specified below. The larger game $\tilde{G}$ has outsiders in a set $o$, in addition to insiders who are the players in $N$, and there can be additional moves by Nature. An insider $n$ can have additional pure strategies in $\tilde{\Sigma}_n$ that are not pure strategies in $\Sigma_n$ but are equivalent to mixed strategies in $\Sigma_n$. The basic requirement is that an embedding should preserve the game among insiders, conditional on actions by outsiders.

These restrictions have a technical formulation. There should exist a multilinear map $f : \tilde{\Sigma} \times \Sigma_o \rightarrow \Sigma$ that is surjective and such that $\tilde{G}_n = G_n \circ f$ for each insider $n$. Moreover, to exclude an embedding from introducing correlation among insiders’ strategies, $f$ should factor into separate multilinear maps $(f_n)_{n \in N}$, where each component is a map $f_n : \tilde{\Sigma}_n \times \Sigma_o \rightarrow \Sigma_n$ such that $f_n(\cdot, \sigma_o)$ maps $\tilde{\Sigma}_n$ surjectively onto $\Sigma_n$ for each profile $\sigma_o \in \Sigma_o$ of outsiders’ strategies.

Admittedly, a statement of the axiom that uses this technical language could contain unsuspected implications. However, after stating the formal definition, we provide in Proposition 3.3 an equivalent formulation that is more detailed and more transparent, and that verifies the requisite properties. Also, Proposition 3.4 applies a precise test of whether the
axiom is correctly stated, namely, a refinement that satisfies the axiom should be immune to the same embeddings that equilibria are.

**Definition 3.2 (Embedding).** A game \( \tilde{G} : \tilde{\Sigma} \times \Sigma_o \rightarrow \mathbb{R}^{N \cup o} \) and a collection of multilinear maps \( f_n : \tilde{\Sigma}_n \times \Sigma_o \rightarrow \Sigma_n \), one for each player \( n \in N \), embed a game \( G : \Sigma \rightarrow \mathbb{R}^N \) if

(a) for each \( \sigma_o \in \Sigma_o \), \( f_n(\cdot, \sigma_o) \) maps \( \tilde{\Sigma}_n \) surjectively onto \( \Sigma_n \), and

(b) \( \tilde{G}_n = G_n \circ f \), where \( f = (f_n)_{n \in N} \).

Condition (a) ensures that embedding has no net effect on an insider’s set of mixed strategies, conditional on outsiders’ strategies, and condition (b) ensures that there is no net effect on any insider’s payoffs.

Hereafter, if \( \tilde{G} \) embeds \( G \) via maps \( f = (f_n) \) then we say that \( (\tilde{G}, f) \) embeds \( G \) and that \( \tilde{G} \) is a metagame for \( G \). We omit description of \( f \) for metagames in extensive form that embed a game in extensive or strategic form. An elaborate example of a metagame in extensive form that embeds a game in extensive form is constructed in proving Theorem 5.1. Note that the property that \( \tilde{G} \) is a metagame for \( G \) does not depend on specifying payoffs for the outsiders; in particular, it depends only on their extensive forms and the insiders’ payoffs in the two games.

A multilinear map \( f_n : \tilde{\Sigma}_n \times \Sigma_o \rightarrow \Sigma_n \) is completely specified by its values at profiles of pure strategies. Let \( \tilde{f}_n \) be the restriction of \( f_n \) to the set \( \tilde{S}_n \times S_o \) of profiles of pure strategies of player \( n \) and the outsiders in \( o \). The following Proposition, proved in Appendix B, provides an alternative definition of embedding in terms of pure strategies.

**Proposition 3.3.** \( \tilde{G} \) embeds \( G \) via a collection of multilinear maps \( f = (f_n)_{n \in N} \) if and only if for each player \( n \) there exists \( \tilde{T}_n \subseteq \tilde{S}_n \) and a bijection \( \pi_n : \tilde{T}_n \rightarrow S_n \) such that for each \( (\tilde{s}, s_o) \in \tilde{S}_n \times S_o \) and \( \tilde{t}_n \in \tilde{T}_n \):

1. \( \tilde{f}_n(\tilde{t}_n, s_o) = \pi_n(\tilde{t}_n) \),
2. \( \tilde{G}_n(\tilde{s}, s_o) = G_n(\tilde{f}(\tilde{s}, s_o)) \), where \( \tilde{f} = (\tilde{f}_n)_{n \in N} \).

Property (2) assures that players’ payoffs from pure strategies of \( G \) are preserved by the metagame \( \tilde{G} \). Hence property (1) assures that each pure strategy \( s_n \in S_n \) is equivalent to some pure strategy \( \tilde{t}_n = \pi_n^{-1}(s_n) \in \tilde{T}_n \), independently of the outsiders’ profile \( s_o \).

Pure strategies in \( \tilde{S}_n \setminus \tilde{T}_n \) are redundant because payoffs from profiles in \( \prod_n \tilde{T}_n \) exactly replicate payoffs from corresponding profiles in \( \prod_n S_n \) for the embedded game \( G \). In particular, if \( \tilde{f}_n(\tilde{s}_n, s_o) = \sigma_n \notin S_n \) then conditional on \( s_o \) the pure strategy \( \tilde{s}_n \) is equivalent for insiders to the mixed strategy \( \sigma_n \in \Sigma_n \).
The next Proposition, proved in Appendix B, verifies that equilibria are not affected by embedding in a metagame.

**Proposition 3.4.** If \((\tilde{G}, f)\) embeds \(G\) then the equilibria of \(G\) are the \(f\)-images of the equilibria of \(\tilde{G}\).

A corollary of Proposition 3.4 is that embedding does not introduce correlation among insiders’ strategies.

Using Definition 3.2 of embedding, the small worlds axiom is the following.

**Axiom S** (Small Worlds): If \((\tilde{G}, f)\) embeds \(G\) then the \(f\)-images of the solutions that a refinement selects for \(\tilde{G}\) are the solutions selected for \(G\).

In view of Proposition 3.4, this axiom is an instance of the general principle that a refinement should inherit invariance properties of equilibria.

Two special cases of Axiom S are the following.

**Invariance:** Suppose \(\Sigma_o\) and \(o\) are singletons and insiders’ payoffs and strategies in \(\tilde{\Sigma}\) differ from \(\Sigma\) only by treating some mixed strategies in \(\Sigma\) as additional pure strategies in \(\tilde{S}\). Then Axiom S implies that solutions depend only on a game’s reduced normal form obtained by deleting such redundant pure strategies.

**Mertens’ Small Worlds Axiom** [20]: Suppose \(\tilde{\Sigma} = \Sigma\). Then Axiom S implies that a solution does not depend on the presence of outsiders, i.e. solutions of the original game are the projections of the solutions of metagames obtained by adding dummy players.

### 3.4. Summary of the Axioms.

We study refinements that are independent of embeddings in metagames that, for each profile of outsiders’ strategies, preserve the strategies and payoffs of the game among insiders. And, we require that each of their solutions is a closed connected set of undominated equilibria, including one that is quasi-perfect. In particular, a solution of a metagame must contain a quasi-perfect equilibrium whose image is in the corresponding solution of the embedded game.

### 4. Additional Notation and Properties

In the sequel we consider only a game \(\Gamma\) in extensive form with perfect recall, two players, and generic payoffs. In this section we prepare for the statement and proof of the main theorem in Section 5.
4.1. **Payoffs.** Let $Z$ be the set of terminal nodes of the game tree. Players’ payoffs are given by a point $u$ in $U = \mathbb{R}^N \times Z$, where $u_n(z)$ is the payoff to player $n \in N$ at terminal node $z \in Z$. We assume that payoffs are generic, i.e. $u \in U \setminus U_0$ where the excluded set $U_0$ is a lower-dimensional set derived in [4].

4.2. **Notation for the Extensive Form.** The set of players is $N = \{1, 2\}$, typically represented as a player $n$ and the other player $m \neq n$. Let $X$ be the set of nodes in the game tree. Let $X_n$ be the set of nodes where player $n$ moves, partitioned into his information sets $h \in H_n$. For a node $x \in X_n$ we write $h(x)$ for the unique information set $h \in H_n$ that contains $x$. For each $n$ and $h \in H_n$, let $A_n(h)$ be the set of actions available to player $n$ at $h$. Assume that actions at all information sets are labeled differently, and let $A_n$ be the set of all actions for player $n$.

Node $x$ precedes another node $y$, written $x \prec y$, if $x$ is on the unique path from the root of the tree to $y$. For a node $x \in X_n$ and $a \in A_n(h(x))$ write $(x, a) \prec y$ if $x \prec y$ and the path from the root of the tree to $y$ requires player $n$ to choose $a$ at $h(x)$. If $(x, a) \prec y$ and $x$ and $y$ belong to $n$’s information sets $h$ and $h'$, respectively, then every node in $h'$ follows some node in $h$ by the choice of $a$, so we write $(h, a) \prec h'$.

The set of pure strategies of player $n$ is the set $S_n$ of functions $s_n : H_n \to A_n$ such that $s_n(h) \in A_n(h)$ for all $h \in H_n$. For each $n$, $s_n \in S_n$ and $y \in X$, let $\beta_n(y, s_n)$ be the probability that $s_n$ does not exclude $y$, i.e. $\beta_n(y, s_n) = 1$ if for each $(x, a) \prec y$ with $x \in X_n$, $s_n(h(x)) = a$, and otherwise $\beta_n(y, s_n) = 0$. By perfect recall, if $y \in X_n$ then $\beta_n(y', s_n) = \beta_n(y, s_n)$ for all $y' \in h(y)$ and we write $\beta_n(h(y), s_n)$ for this probability. Likewise, for any node $y$ we write $\beta_0(y)$ for the probability that Nature does not exclude $y$. Then for a profile $s \in S$ the probability that node $y$ is reached is $\beta(y, s) \equiv \beta_0(y)\beta_1(y, s_1)\beta_2(y, s_2)$.

Recall that $\Sigma_n$ is the set of mixed strategies of player $n$. For each node $y$ the function $\beta_n(y, \cdot)$ extends to a function over $\Sigma_n$ via $\beta_n(y, \sigma_n) = \sum_{s_n \in S_n} \beta_n(y, s_n)\sigma_n(s_n)$ for $\sigma_n \in \Sigma_n$. Recall also that $B_n$ is the set of behavioral strategies of player $n$. For each $b_n \in B_n$, $\beta_n(y, b_n)$ is the product of $b_n$’s probabilities of $n$’s actions on the path to $y$.

Similarly extend $\beta$ to profiles of mixed or behavioral strategies. Given a mixed-strategy profile $\sigma \in \Sigma$, the probability that outcome $z$ results is $\beta(z, \sigma) = \beta_0(z)\beta_1(z, \sigma_1)\beta_2(z, \sigma_2)$.

4.3. **Enabling Strategies.** For each player $n$ define $\rho_n : \Sigma_n \to [0, 1]^Z$ by the formula $\rho_n(\sigma_n) = (\beta_n(z, \sigma_n))_{z \in Z}$, and let $\rho = (\rho_n)_{n \in N}$. Similarly, if $b_n \in B_n$ is the behavioral strategy induced by $\sigma_n$ then $\rho_n(\sigma_n) = (\beta_n(z, b_n))_{z \in Z}$. Let $P_n$ be the image of $\rho_n$, and $P = \prod_n P_n$ the image of $\rho$. Then $P_n$ is a compact convex polyhedron, called the space of $n$’s enabling
strategies in [5]. Each vertex of $P_n$ corresponds to an equivalence class of $n$’s pure strategies that exclude the same sets of outcomes.\(^5\)

If $\sigma \in \Sigma$ and $p = \rho(\sigma)$ then the probability of outcome $z$ is $\gamma_z(p) = \beta_0(z)p_1(z)p_2(z)$. Thus the function $\gamma : P \to \Delta(Z)$ summarizes the extensive form. The analog of the game $\Gamma$’s normal form $G : \Sigma \to \mathbb{R}^N$ is the enabling form $G : P \to \mathbb{R}^N$ that assigns to each profile of enabling strategies the profile of players’ expected payoffs, where $G_n(p) = \sum_z \gamma_z(p)u_n(z)$. Note that $\gamma$ and $G$ are multilinear functions. From players’ best-reply correspondences in terms of enabling strategies one obtains the definition of equilibrium in enabling strategies. To each equilibrium in enabling strategies there correspond families of outcome-equivalent equilibria in behavioral and mixed strategies. Axioms A and S have direct analogs in terms of enabling strategies, as shown in [10].

For games in extensive form with perfect recall, enabling strategies are minimal representations. For example, using perfect recall, by working backward in the induced tree of a player’s information sets, from his enabling strategy one can construct the corresponding behavioral strategy at his information sets that his prior actions do not exclude. Because Axiom S implies Invariance, it is immaterial whether solutions are characterized in terms of mixed, behavioral, or enabling strategies. We use enabling strategies here because induced distributions over outcomes are multilinear functions of enabling strategies, like they are for mixed strategies but unlike the nonlinear dependence on behavioral strategies. Also, the dimensions of the spaces of enabling and behavioral strategies are the same, which is important for the technical property established in Theorem 4.1 below. Using these aspects, Section 5 derives the implications of the axioms in terms of enabling strategies.

4.4. Stable Sets of Equilibria. Now let $\Sigma^*$ be a component of the equilibria of $\Gamma$ in terms of mixed strategies, and let $\Sigma^*_n$ be the projection of $\Sigma^*$ in $\Sigma_n$.

By genericity, all equilibria in $\Sigma^*$ induce the same distribution over outcomes. Therefore, for each node $x$, $\beta(x, \sigma)$ is the same for all $\sigma \in \Sigma^*$; in particular, if $x$ belongs to information set $h \in H_n$ and $h$ is on an equilibrium path then $\beta_n(h, \sigma_n)$ is the same for every equilibrium strategy $\sigma_n$ of player $n$ in $\Sigma^*$. We therefore denote these probabilities by $\beta_n^*(x)$ and $\beta_n^*(h)$. Let $H^*_n$ be the collection of information sets $h \in H_n$ of player $n$ such that $\beta_n^*(h) > 0$ and let $A^*_n$ be the set of actions at information sets in $H^*_n$ that are chosen with positive probability by the equilibria in $\bar{B}^*$, where $\bar{B}^*$ is the set of profiles of behavioral strategies induced by equilibria in $\Sigma^*$.

\(^5\)The vertices are $n$’s pure strategies in the ‘pure reduced normal form’ defined by Mailath, Samuelson, and Swinkels [17]. See Appendix A for illustrations.
Let \( S_n^0 \subset S_n \) be the set of pure strategies \( s_n^0 \) with the property that, at each information set \( h \in H_n^* \) that \( s_n^0 \) does not exclude, \( s_n^0 \) prescribes an action in \( A_n^* \). Let \( S_n^1 = S_n \setminus S_n^0 \), i.e. each pure strategy \( s_n \) in \( S_n^1 \) chooses a non-equilibrium action at some information set \( h \in H_n^* \) that it does not exclude.

For \( i = 0, 1 \), let \( \Sigma_i^1 \) be the set of mixed strategies whose support is contained in \( S_i^1 \). Observe that the support of \( n \)'s strategy in every equilibrium in \( \Sigma_1^* \) is contained in \( S_n^0 \) and that every strategy in \( S_n^0 \) is a best reply against every equilibrium in \( \Sigma_1^* \). Thus \( \Sigma_n^* \) is contained in \( \Sigma_n^0 \) and \( \Sigma^* = \Sigma_1^* \times \Sigma_2^* \).

Let \( P_n^0 \) be the image of \( \Sigma_n^0 \) under \( \rho_n \). Let \( Z_n^1 \subset Z \) be the set of terminal nodes \( z \) such that \( (h, a) \prec z \) for some \( h \in H_n^* \) and \( a \notin A_n^* \). Let \( Z_n^0 = Z \setminus Z_n^1 \). Then \( P_n^0 \) is the set of \( p_n \in P_n \) such that \( p_n(z) = 0 \) for all \( z \in Z_n^1 \) and thus \( P_n^0 \) is a face of \( P_n \). However, the image \( P_n^1 \) of \( \Sigma_1^1 \) under \( \rho_n \) need not be a face of \( P_n \). For \( i = 0, 1 \), let \( P_i = P_i^1 \times P_i^2 \) and define \( \mathbb{P} = P^0 \times P^1 \).

For each \( p_n \in P_n \), let \( \psi_{Z_n^1}(p_n) \) be the projection of \( p_n \) to \( \mathbb{R}^Z_n \); then \( \psi_{Z_n^1}(p_n) = 0 \) iff \( p_n \in P_n^0 \). Fix a point \( \bar{p}_n \) in the interior of \( P_n \) and define \( \eta_n : P_n \to \mathbb{R} \) by \( \eta(p_n) = \sum_{z \in Z_n^1} p_0(z)\bar{p}_n(z)p_n(z) \), where \( p_0 \) is Nature’s enabling strategy. Then \( \eta(p_n) = 0 \) iff \( p_n \in P_n^0 \).

Choose \( \varepsilon > 0 \) such that \( \eta_n(p_n) > \varepsilon \) for all \( p_n \in P_n^1 \). Let \( H \) be the hyperplane in \( \mathbb{R}^{Z_n^1} \) with normal \( (p_0(z)\bar{p}_n(z))_{z \in Z_n^1} \) and constant \( \varepsilon \). \( H \) separates the origin from \( \psi_{Z_n^1}(P_n^1) \). Let \( \Pi_n^1 \) be the intersection of \( H \) with \( \Pi_n \). Then there is a function \( \pi_n^1 \) from \( P_n \setminus P_n^0 \) to \( \Pi_n^1 \) that maps each \( p_n \notin P_n^0 \) to the point \( \varepsilon(\eta_n(p_n))^{-1}\psi_{Z_n^1}(p_n) \).

Now let \( \Sigma^* \) be a component of the undominated equilibria of the game \( \Gamma \) that is contained in \( \Sigma^*_1 \). Since \( \Sigma^*_1 = \Sigma^*_1 \times \Sigma^*_2 \), \( \Sigma^*_1 \) is expressible as a product \( \Sigma_1^* \times \Sigma_2^* \), where \( \Sigma_1^* \) is a component of the intersection of \( \Sigma_1^* \) with the set of undominated strategies. For each \( n \), let \( Q_n^* \) be the image of \( \Sigma_n^* \) under \( \rho_n \), and \( Q^* \) the image of \( \Sigma^* \) under \( \rho \), i.e. represented in enabling strategies.

Hereafter we say that \( Q^* \) is stable if \( \Sigma^* \) is a stable set of equilibria.\(^6\)

4.5. **The Associated Pseudomanifold.** Let \( Q \) be the set of \( (q^*, (p_0^0, p_1^1), \pi^1) \in Q^* \times \mathbb{P} \times \Pi^1 \) such that there exist \( r_0^0, r_1^0 \in P_n^0, r_1^1 \in P_n^1 \), and for each \( n \) scalars \( \lambda_n^0, \lambda_n^1, \mu_n^1 \) in the interval \((0, 1]\) such that, if

\[
q_0^0 = \lambda_n^0 p_0^0 + (1 - \lambda_n^0) r_0^0 \quad \text{and} \quad q_1^1 = (1 - \lambda_n^1) p_0^1 + \lambda_n^1 (\mu_n^1 p_n^1 + (1 - \mu_n^1) r_0^1),
\]

then for each \( n \):\(^7\)

\[
(1) \quad \pi^1_n(q_1^1) = \pi^1_n.
\]

\(^6\)Alternatively, as noted by Mertens [20], one can apply the definition of stability directly to \( Q^* \) as a component of equilibria, represented in terms of enabling strategies, over the space of perturbations of players’ enabling strategies.

\(^7\)Lexicographically optimal replies are defined by Blume, Brandenburger, and Dekel [1] and Govindan and Klumpp [2].
(ii) \( q_n^0 \), and \( r_n^0 \) if \( \lambda_n^0 < 1 \), are lexicographically as good replies against \((q_m^0, q_m^0, q_m^1)\) as the strategies in \( S_n^0 \).

(iii) If \( \mu_n^1 < 1 \) then \( r_n^1 \) is a best reply against \( q_n^0 \) and lexicographically as good a reply against \((q_m^0, q_m^0, q_m^1)\) as other strategies in \( P_n^1 \).

Let \( \psi : Q \to \mathbb{P} \) be the natural projection, i.e. \( \psi(q^*, (p^0, p^1), \pi^1) = (p^0, p^1) \). The following technical property is proved in Appendix C.

**Theorem 4.1.** \((Q, \partial Q)\) is a pseudomanifold of the same dimension as \((\mathbb{P}, \partial \mathbb{P})\). Moreover, the projection map \( \psi : (Q, \partial Q) \to (\mathbb{P}, \partial \mathbb{P}) \) is essential iff \( Q^* \) is stable.

5. **Statement and Proof of the Theorem**

We now state and prove the main theorem.

**Theorem 5.1.** Axioms A, B, S imply that each solution is an essential component of the game’s undominated equilibria.

**Proof.** For the proof we assume that a solution is represented in terms of enabling strategies, i.e. \( Q^* \subset P \) is a solution iff it is the image under \( \rho \) of a solution \( \Sigma^* \subset \Sigma \).

Let \( \deg(\psi) \) be the degree of the projection map \( \psi : (Q, \partial P) \to (\mathbb{P}, \partial \mathbb{P}) \). Let \( q^{0,*} \) be an arbitrary point in \( Q^* \). For each \( m \), choose a point \( p_m^{0,*} \) in the interior of \( P_m^0 \) against which \( q^{0,*} \) is a best reply. Such a choice is possible by genericity of payoffs: the interior of the projection of \( \Sigma_m^* \) belongs to the interior of \( P_m^0 \) and all strategies in \( P_m^1 \) are inferior replies against every interior point. Since \( q_n^{0,*} \) belongs to \( Q_n^* \), which consists only of undominated (hence admissible) strategies, there exists a point \( p_m \) in the interior of \( P_m \) against which \( q_n^{0,*} \) is a best reply. Expressing \( p_m \) as a convex combination of points \( p_m^0 \) and \( p_m^1 \) in the relative interiors of \( P_m^0 \) and \( P_m^1 \) respectively, it follows that \( x^* \equiv (q_n^{0,*}, (p_n^0, p_n^1), \pi_n^1) \) belongs to \( Q \setminus \partial Q \), where for each \( n \), \( \pi_n^{1,*} = \pi_n^1(p_n^{1,*}) \)—and taking, for \( i = 0, 1 \), \( p_n^i \) to be \( p_n^{i,*} \), in the definition of \( Q \).

For any neighborhood \( U(q_0^{0,*}) \) of \( q_0^{0,*} \) we can choose a neighborhood \( U(x^*) \) of \( x^* \) that is homeomorphic to a simplex, is contained in \( Q \setminus \partial Q \) and such that the projection onto the first factor is contained in the neighborhood \( U(q_0^{0,*}) \), i.e. if \( (q^0, (p^0, p^1), \pi^1) \in U(x^*) \) then \( q^0 \in U(q_0^{0,*}) \). For the proof we construct a metagame such that: (i) if \( \deg(\psi) \) is zero then there is no quasi-perfect equilibrium of the metagame projecting to an equilibrium in \( Q^* \); and (ii) if \( \deg(\psi) \) is nonzero then all its quasi-perfect equilibria that project to a point in \( Q^* \), project to a point in \( U(q_0^{0,*}) \). Since \( U(q_0^{0,*}) \) is an arbitrary neighborhood of the arbitrary point \( q_0^{0,*} \) in \( Q^* \), this proves the theorem.
5.1. Preliminaries. Fix some \( \bar{p} \) in the interior of \( \mathbb{P} \) and define \( \iota : \partial \mathbb{P} \to \partial \mathbb{P} \) as follows: \( \iota(p) \) is the unique point in the boundary of the form \( \lambda \bar{p} + (1 - \lambda)\bar{p} \) for \( \lambda < 0 \). Then \( \iota \) has no fixed point. Let \( g_{\partial Q} : \partial Q \to \mathbb{P} \) be the function \( \iota \circ \psi_{\partial Q} \), where \( \psi_{\partial Q} \) is the restriction of \( \psi \) to \( \partial Q \). Then \( g \) has no point of coincidence with \( \psi \), i.e., for each \( x \in \partial Q \), \( g(x) \neq \iota(\psi(x)) \). \( \psi_{\partial Q} \), and hence also \( g_{\partial Q} \), has the same degree as \( \psi \). If this degree is zero, then \( g_{\partial Q} \) can be extended to a map \( g : Q \to \partial \mathbb{P} \). If this degree is nonzero, then there exists a map \( g_{\partial U(x^*)} \) that has the same degree and is such that the two maps \( g_{\partial Q} \) and \( g_{\partial U(x^*)} \) can be extended to a map from \( Q \setminus (U(x^*) \setminus \partial U(x^*)) \) to \( \mathbb{P} \); furthermore, by mapping points in \( U(x^*) \) to \( \mathbb{P} \) in way that extends \( g_{\partial U(x^*)} \), we now obtain a map \( g : Q \to \mathbb{P} \) such that all its points of coincidence with \( \psi \), of which there is at least one, are contained in \( U(x^*) \setminus \partial U(x^*) \).

By construction there exists \( \alpha > 0 \) such that for all \( x \in Q \), \( \|g(x) - \psi(x)\| \leq \alpha \) iff \( \psi \) has a nonzero degree and \( x \in U(x^*) \). Choose a triangulation \( \mathcal{K}_n^i \) of \( P_n^i \) for each \( n \) and \( i = 0, 1 \) such that the diameter of each simplex is no more than \( \alpha/2 \). For each \( n \), there exists for each \( i \) a triangulation \( \mathcal{L}_n^i \) of \( P_n^i \) and a triangulation \( \Pi_n^i \) of \( \Pi_1^i \) such that, letting \( \mathcal{L} \) be the resulting multisimplicial subdivision of \( P_n \), \( g \) has a multisimplicial approximation \( \tilde{g} \) [9, Appendix B] with the triangulation of the range given by \( \mathcal{K} = \prod_{n,i} \mathcal{K}_n^i \). Observe that if for some \( x = (q^0, p^0, p^1, \pi^1) \) there exists a multisimplex \( K \) that contains \( \psi(x) \) and \( \tilde{g}(x) \) then, since \( \tilde{g} \) is a multisimplicial approximation of \( g \), that multisimplex is a face of one that contains \( g(x) \) and thus \( \|g(x) - \psi(x)\| \leq \alpha \); in particular, if \( x \) also belongs to \( Q \), then \( x \in U(x^*) \), and thus \( q^0 \in U(q^{0,*}) \). As in [9, Theorem B.2], now take a further polyhedral subdivision \( \mathcal{T} \) of \( \mathcal{L} \) and let \( \gamma \) be the convex function generated by \( \mathcal{T} \), i.e. \( \gamma \) is piecewise linear and linear precisely on each full-dimensional polyhedron of the subdivision.

Hereafter, unless explicitly stated, we represent players’ pure and mixed strategies in terms of the induced enabling strategies; thus, pure strategies in \( S_n \) are represented as vertices of \( P_n \) and mixed strategies are mixtures of these vertices.

5.2. A Game with Redundant Strategies. We first construct a game with redundant pure strategies that is the basis for the metagame specified later. For each fixed \( \bar{p} = (p^0, p^1) \in \mathbb{P} \) and \( \delta \in (0, 1) \), consider the following game \( \Gamma(\delta, \bar{p}) \). Each player \( n \) chooses a strategy in the following manner (and unaware of his opponent’s choices). Initially, player \( n \) provisionally chooses a pure strategy \( s^0_n \in S_0^n \), or he rejects all strategies in \( S_0^n \).

- If initially he chooses a strategy \( s^0_n \) then at a subsequent second stage he can retain \( s^0_n \) or revise his choice. At this second stage the revisions available are ‘duplicate’ pure strategies in a set \( \mathcal{T}_n^0(\delta, p^0_n) \), each of which is a mixed strategy of the form \( t_n(\delta, p^0_n) \equiv (1 - \delta)t_n + \delta p^0_n \) for some \( t_n \in S_0^n \).
• If he rejects all strategies in $S_n^0$ then at a second stage he can choose among the strategies in $T_n^0(\delta, p_n^0)$ or reject them all.\footnote{It would have sufficed, at this stage, to give player $n$ the option of playing just the strategy $p_n^0$ instead of all the strategies in $T_n^0(\delta, p_n^0)$, which we do only for economy in notation.} If he rejects them all then at a third stage he chooses one among the pure strategies in $S_n^1 \cup T_n^1(\delta, p_n^1)$, where each strategy in $T_n^1(\delta, p_n^1)$ is a duplicate of the form $t_n(\delta, p_n^1) \equiv (1 - \delta)t_n + \delta p_n^1$ for some $t_n \in S_n^0$.

In $\Gamma(\delta, \hat{p})$ the set of $n$’s pure strategies is $\hat{S}_n(\delta, \hat{p}_n) \equiv S_n \cup T_n^0(\delta, p_n^0) \cup T_n^1(\delta, p_n^1)$. Thus, game $\Gamma(\delta, \hat{p})$ has the same reduced normal form as $\Gamma$.

5.3. Extensive Form of the Metagames. Now we specify a family of similar metagames $\hat{\Gamma}^\delta$, one for each $\delta > 0$.

Before the insiders play, thirteen outsiders, denoted players $o_0$ and $o_{n,i,j}$ for $n = 1, 2$, $i = 0, 1$ and $j = 1, 2, 3$, move simultaneously. Outsider $o_0$ chooses a full-dimensional polyhedron $T$ of the polyhedral complex $\mathcal{T}$. Outsider $o_{n,i,j}$, for $n = 1, 2$, $j = 1, 3$ and $i = 0, 1$, chooses a point in the vertex set $V_n^i$ of $\mathcal{K}^1_n$. Outsider $o_{n,2}$ chooses a point in a finite subset $S_{n,\delta}^0$ of $P_n^0$ chosen such that every point in $P_n^0$ is within $\delta$ of some point in $S_{n,\delta}^0$; outsider $o_{n,2}$ chooses a point in a finite subset $S_{n,\delta}^1$ of $\Pi_n^1$ such that every point in $\Pi_n^1$ is within $\delta$ of some point in $S_{n,\delta}^1$.

For outsider $o_{n,1}$, each pure strategy $v_n^i$ corresponds to a point in $P_n^i$ denoted $p_n^i(v_n^i)$ and each mixed strategy $\sigma_n^i$ corresponds to a point in $P_n^i$ denoted $p_n^i(\sigma_n^i)$. Likewise a mixed strategy $\sigma_{n,2}^0$ of $o_{n,2}$ corresponds to a point $q_n^0(\sigma_{n,2}^0)$ in $P_n^0$, and a mixed strategy $\sigma_{n,2}^1$ of $o_{n,2}$ corresponds to a point $\pi_n^1(\sigma_{n,2}^1)$ in $\Pi_n^1$.

A mixed-strategy profile $\sigma_o$ for the outsiders induces a point $(q_n^0(\sigma_o), (p_n^0(\sigma_o), p_n^1(\sigma_o)), \pi_n^1(\sigma_o))$ in $P_n^0 \times \mathbb{P} \times \Pi_n^1$, where for each $n$ and $i$, $p_n^i(\sigma_o)$ depends on the choice by $o_{n,1}$, and $q_n^0(\sigma_o)$ and $\pi_n^1(\sigma_o)$ depend on the choices by $o_{n,2}$ respectively.

After each pure-strategy profile $\tilde{s}_n$ of the outsiders there follows a copy of the game $\Gamma(\delta, \hat{p}(\tilde{s}_n))$. That is, if in the profile $\tilde{s}_n$ outsiders $o_{n,1}$ choose points $v_n^i$, then there follows a copy of $\Gamma(\delta, \hat{p})$ after these choices, where for each $n$, $\hat{p}_n = (p_n^0(v_n^0), p_n^1(v_n^1))$. However, the information sets in $\hat{\Gamma}^\delta$ are such that the insiders play without knowing which copy of $\Gamma(\delta, \hat{p})$ they are playing. The duplicate strategies available are therefore now denoted by $T_n^0(\delta)$ and $T_n^1(\delta)$, omitting the reference to $p_n^0$ and $p_n^1$, since the insiders are uninformed of which mixtures were implemented by outsiders. Put differently, for $i = 0, 1$ and $t_n \in S_n^0$, the exact duplicate strategy implemented by choosing $t_n(\delta) \in T_n^i(\delta)$ depends on the choice by the outsider $o_{n,1}$ that insiders do not observe. Thus in the metagame $\hat{\Gamma}^\delta$, player $n$’s set of pure strategies, up to duplication of pure strategies, is $\hat{S}_n(\delta) \equiv S_n \cup T_n^0(\delta) \cup T_n^1(\delta)$. 
That the metagame $\tilde{\Gamma}^\delta$ embeds $\Gamma$ follows from Proposition 3.3. The strategies in $S_n$ are available as pure strategies in $\tilde{S}_n(\delta)$; the other pure strategies, which belong to $T_n^i(\delta)$, for $i = 0, 1$, implement mixtures that are dependent on the choice of outsiders $o_{-1}^i$.

5.4. Outsiders’ Payoffs in the Metagames. Next we describe the outsiders’ payoffs in each metagame $\tilde{\Gamma}^\delta$.

The payoffs to $o_0$ depend on the choices of all outsiders except outsiders $o_{-3}^i$ for $i = 0, 1$ and $n = 1, 2$. Recall that the convex function $\gamma$ is linear over each full-dimensional polyhedron $T$ of $T$. This linear function extends uniquely to a linear function $\gamma_T$ over $P^0 \times P \times \Pi^1$. Every mixed strategy profile of the other outsiders induces a unique point $(\tilde{q}, \tilde{p}, \pi^1) \in P^0 \times P \times \Pi^1$ and $o_0$’s payoff from choosing $T$ is $\gamma_T(\tilde{q}, \tilde{p}, \pi^1)$.

 Outsider $o_{-1}^i$ wants to mimic $o_{-3}^i$. In particular, if he chooses $v_n^i$ and $o_{-1}^i$ chooses $w_n^i$ then his payoff is one if $v_n^i = w_n^i$ and zero otherwise.

 Outsider $o_{-2}^0$ wants to mimic the actual choice implemented by player $n$ when this choice belongs to $P_n^0$. Similarly, outsider $o_{-2}^1$ wants to mimic the $\pi_n^1$ implied by $n$’s choice when he plays a strategy in $P_n^1$. Specifically, for $i = 0, 1$, let $\varphi_n^i : \mathbb{R}^{Z_n^i} \to \mathbb{R}$ be the function given by $\varphi_n^i(r) = \sum_{z \in Z_n^i} r_z^2$. For each $r \in \mathbb{R}^{Z_n^i}$, let $\xi_n^i(r, \cdot)$ be the affine approximation to $\varphi_n^i$ at $r$, i.e. for each $r' \in \mathbb{R}^{Z_n^i}$, $\xi_n^i(r, r') = \sum_z (r_z^2 + 2r_z(r'_z - r_z))$. Suppose now that $o_{-2}^0$ chooses a pure strategy $s_{-2}^{0,0}$ and $n$ chooses a pure strategy $s_n$ in $\tilde{S}_n(\delta)$. If $s_n$ is in $T_n^0(\delta)$ or $T_n^1(\delta)$ then let $q_n$ be the actual strategy in $P_n$ that is implemented based on $p_n^0(v_n^0)$ or $p_n^1(v_n^1)$ where for each $i, v_n^i$ is the choice of outsider $o_{-1}^i$; and otherwise let $q_n = s_n$. For outsider $o_{-2}^0$, his payoff is zero if $q_n \notin P_n^0$, otherwise, it is $\xi_n^0(s_{-2}^{0,0}, q_n)$. For outsider $o_{-2}^1$, his payoff is zero if $q_n \in P_n^0$; otherwise it is $\xi_n^1(s_{-2}^{0,1}, \pi_n^1(q_n))$.

 The payoff to outsider $o_{-3}^i$ depends on the choices of all other outsiders. If $o_0$ chooses a polyhedron $T$ then there exists a unique multisimplex $L$ of $L$ that contains $T$. For each vertex $w_n^i$ of $V_n^i$, and each vertex $\tilde{v}$ of $L$, let $u_n^i(T, \tilde{v}, w_n^i) = 1$ if $w_n^i$ is the image of $\tilde{v}$ under $\tilde{g}_n^i$ and zero otherwise, where $\tilde{g}_n^i$ is the $(n, i)$-th coordinate map of $\tilde{g}$. The function $u_n^i$ extends multilinearly to $L$ and, since $L$ is full-dimensional, to the whole of $P^0 \times P \times \Pi^1$, denoted still by $u_n^i(T, \cdot, w_n^i)$. Given an arbitrary mixed strategy of the other players, if $o_0$ chooses $T$ and $o_{-2}^i$ chooses $w_n^i$ then the payoff of $o_{-2}^i$ is $u_n^i(T, (p, q), w_n^i)$, where $(p, q)$ is the point in $P^0 \times P \times \Pi^1$ induced by the mixed strategies of the other players.

5.5. Outsiders’ Strategies in a Quasi-Perfect Equilibrium. If $Q_*$ is a solution of $\Gamma$ then, by Axioms B and S, in the metagame $\tilde{\Gamma}^\delta$ there exists a quasi-perfect equilibrium $\tilde{b}^\delta$.
whose equivalent mixed-strategy profile $\bar{\sigma}^\delta$ belongs to a solution and whose image under the map from the metagame $\bar{\Gamma}^\delta$ to $\Gamma$ is a point $q^*,\delta$ in $Q^*$.

Each player $n$’s strategy in $\bar{b}^\delta$ necessarily has the following feature. He avoids going to his information set where his choices are among the strategies in $S_n^1(\delta) \cup T_n^1(\delta)$, since each of these strategies chooses a non-equilibrium action at some information set on the equilibrium path. Let $q_n^0,\delta$ be $n$’s actual strategy in $P_n^0$ that is implemented by $n$’s strategy in the profile $\bar{b}^\delta$ in the metagame $\bar{\Gamma}^\delta$. By construction, $q_n^0,\delta$ belongs to $Q_n^*$. Let $(q_n^0,\delta, p_n^0,\delta, \pi_n^1,\delta)$ be the point in $P^0 \times \mathbb{P} \times \Pi^1$ that is induced by the profile $\bar{\sigma}_n^\delta$ of the outsiders’ strategies in the equilibrium $\bar{\sigma}^\delta$.

Under $\bar{b}^\delta$, after $n$ has rejected all strategies in $S_n^0$ and $T_n^0(\delta)$, consider the strategy implemented by $n$. Let $(1 - \alpha_n^1,\delta)$ be the total probability of choosing a duplicate in $T_n^1(\delta)$ under $\bar{b}^\delta$. Then $n$’s choice at this information set is equivalent to an enabling strategy in $P_n$ of the form

$$q_n^1,\delta = (1 - \alpha_n^1,\delta)((1 - \delta)p_n^0,\delta + \delta p_n^1,\delta) + \alpha_n^1,\delta r_n^1,\delta,$$

where: (i) $p_n^0,\delta$ is the mixture over strategies $t_n$ such that the strategy $t_n^1(\delta)$ is played with positive probability at this information set; (ii) $p_n^1,\delta = p_n^i(\bar{\sigma}_n^i,\delta)$ is the enabling strategy induced by outsider $o_n^i$; (iii) $r_n^1,\delta$ is the enabling strategy in $P_n^1$ that is obtained from $n$’s actual mixture over strategies in $S_n^1$ if $\alpha_n^1,\delta > 0$, and is arbitrary otherwise. Let

$$q_n^1,\delta = (\delta(1 - \alpha_n^1,\delta) + \alpha_n^1,\delta)^{-1}(\delta(1 - \alpha_n^1,\delta)p_n^1,\delta + \alpha_n^1,\delta r_n^1,\delta).$$

Then $\pi_n^1(q_n^1,\delta) = \pi_n^1(q_n^1,\delta) = \pi_n^1,\delta$.

The following lemma characterizes the important aspects of the outsiders’ equilibrium strategies.

**Lemma 5.2.** The equilibrium strategies of the outsiders satisfy the following properties.

1. For each $n$, suppose the vertices in the support $V_n^i,\delta$ of $o_{n,3}^i$’s equilibrium strategy span a simplex $K_n^i,\delta$ of $\mathcal{K}_n^i$. Then $p_n^i,\delta$ belongs to $K_n^i$.

2. If every polyhedron in the support of $o_0$’s strategy contains $x$ then, for each $n$ and $i$, the vertices in $V_n^i,\delta$ span a simplex $K_n^i,\delta$, and $\bar{g}_{n,0}(x)$ belongs to the interior of a simplex $\bar{K}_n^i,\delta$ that has $K_n^i,\delta$ as a face.

3. Every polyhedron in the support of $o_0$’s strategy contains $x$.

4. $\bar{g}_n^0,\delta$ is within $\delta$ of $q_n^0,\delta$ and $\pi_n^1,\delta$ is within $\delta$ of $\pi_n^1,\delta$.

**Proof of Lemma.** Outsider $o_{n,1}^i$ wants to mimic outsider $o_{n,3}^i$. So, if the vertices of $V_n^i,\delta$ span a simplex $K_n^i,\delta$, then the payoff to $o_{n,1}^i$ from choosing a vertex $w_n^i$ is positive if it belongs to $V_n^i,\delta$, and zero otherwise. Point (1) follows.
Let \( \bar{L} = L^0 \times (\tilde{L}^0 \times L^1) \times L^\Pi \) be the unique multisimplex of \( \mathcal{L} \) that contains \( x \) in its interior. For each polyhedron \( T \) in the support of \( o_0 \)'s strategy, there exists a full-dimensional multisimplex \( \bar{L} \) of \( \mathcal{L} \) that contains \( T \). Obviously \( \bar{L} \) has \( \bar{L} \) as a face. \( \tilde{o}_{n,3}'s \) payoff from choosing a strategy \( w_n^i \) if \( o_0 \) chooses such a \( T \), and given the strategies of the other outsiders, is positive if it is the image of a vertex of \( \bar{L} \) under \( \tilde{g} \) and zero otherwise. Since the image of \( \bar{L} \) under the coordinate function \( \tilde{g}_n^i \) is a simplex \( \bar{K}_n^{i,\delta} \), \( \tilde{g}_n^i(x) \in \bar{K}_n^{i,\delta} \) and the vertices of \( V_{n,i}^{\delta} \) span a face of \( \bar{K}_n^{i,\delta} \). Therefore, point (2) follows.

For each polyhedron \( T \) of \( \mathcal{T} \), the payoff from \( T \) is \( \gamma_T(x) \) and by construction, \( \gamma_T(x) \leq \gamma(x) \) with the inequality being strict iff \( x \) does not belong to \( T \), which proves (3).

The actual strategy implemented by \( n \) is \( q_n^{0,\delta} \), which belongs to \( P_n^0 \). \( q_{n,2}^{0,\delta} \)'s payoff function is such that his best replies to \( q_n^{0,\delta} \) are the points in \( S_n^{1,\delta} \) that are closest to \( q_n^{0,\delta} \). Thus \( q_n^{0,\delta} \) is within \( \delta \) of \( q_n^{0,\delta} \).

Since \( q_n^{0,\delta} \) belongs to \( P_n^0 \), all of \( o_{n,2}^{1,\delta} \)'s strategies yield a payoff of zero against \( q_n^{0,\delta} \). However, since the behavioral strategy \( \tilde{b}^{\delta} \) is a quasi-perfect equilibrium, there exists a sequence of completely mixed behavioral strategies \( \tilde{b}^{e,\delta} \) converging to \( \tilde{b}^{\delta} \) against which \( o_{n,2}^{1,\delta} \)'s equilibrium strategy \( \tilde{\sigma}_{n,2}^{e,\delta} \) is a best reply. Under the sequence \( \tilde{b}^{e,\delta} \), there is a positive probability that player \( n \) moves into the third stage of his choice and makes a choice among strategies in \( S_n^{1} \cup T_n^{1}(\delta) \). The fact that \( \tilde{\sigma}_{n,2}^{e,\delta} \) is optimal against the sequence implies that it is optimal against the limiting choice \( \tilde{q}^{\delta}_n \) there. Since \( \pi_n^{1}(\tilde{q}^{\delta}_n) = \pi_n^{1,\delta} \), \( o_{n,2}^{1,\delta} \)'s best replies are within \( \delta \) of \( \pi_n^{1,\delta} \).

\[ \square \]

5.6. Limits of the Quasi-Perfect Equilibria of the Metagames. Consider a sequence of \( \delta \)'s converging to zero and a corresponding sequence \( \tilde{b}^{\delta} \) of quasi-perfect equilibria in solutions of the metagames \( \tilde{\Gamma}^{\delta} \). Let \( \tilde{\sigma}^{\delta} \) be an equivalent sequence of mixed strategies. By passing to a subsequence, we can assume that the simplex \( \tilde{K}^{i,\delta}_n \) spanned by the support \( V_n^{i,\delta} \) of \( o_{n,3}^{i,\delta} \)'s equilibrium strategy, as well as the simplex \( \tilde{K}^{i,\delta}_n \) is constant across all \( \delta \)'s for each \( n \) and \( i \). Denote this common simplex by \( \bar{K}_n^{i,\delta} \) and let \( K \) be the corresponding multisimplex.

Let \( (p^{0,\delta}, p^{1,\delta}), \bar{q}^{0,\delta}, \bar{\pi}^{0,\delta}, \bar{\pi}^{1,\delta} \) be the corresponding limits of \( (p^{0,\delta}, p^{1,\delta}), q_{n,i}^{0,\delta}, \pi_{n,i}^{0,\delta}, \pi_{n,i}^{1,\delta} \). Let \( q_n^{0,\delta} \) and \( \pi_n^{1,\delta} \) be the limits of \( q_n^{0,\delta} \) and \( \pi_n^{1,\delta} \). By properties 1-3 of the previous lemma, for each \( \delta \), and \( n, i, \) there exists a simplex \( \bar{K}_n^{i,\delta} \) of \( \mathcal{K}_n \) that contains both \( \tilde{g}_{n,i}(q_{n,i}^{0,\delta}, (p^{0,\delta}, p^{1,\delta}), \pi_{n,i}^{1,\delta}) \) and \( p_n^{1,\delta} \), with the former belonging to its interior. By property 4 of the previous lemma, \( \bar{q}_n^{0,\delta} = q_n^{0,\delta} \) and \( \bar{\pi}_n^{1,\delta} = \pi_n^{1,\delta} \).

By passing to a subsequence, we can assume that there exist multisimplices \( \bar{L} \) of \( \mathcal{L} \) and \( \bar{K} \) of \( \mathcal{K} \) such that for all \( \delta \), \( (\bar{q}_n^{0,\delta}, (p^{0,\delta}, p^{1,\delta}), \bar{\pi}_n^{1,\delta}) \) belongs to the interior of \( \bar{L} \) and its image under \( \tilde{g} \) belongs to the interior of \( \bar{K} \)—hence \( \bar{K} \) also contains \( (p^{0,\delta}, p^{1,\delta}) \). Going to the limit,
(q^{0,0}, (p^{0,0}, p^{1,0}, \pi^{1,0}) belong to \bar{L} and its image under \tilde{g} belongs to \bar{K}. Since \bar{L} and \bar{K} are multisimplices and \tilde{g} is a multisimplicial map, we have that for all \delta, x^{\delta} \equiv (q^{0,0}, (p^{0,\delta}, p^{1,0}), \pi^{1,0}) belongs to \bar{L}, and \tilde{g}(x^{\delta}), (p^{0,\delta}, p^{1,0}) \in \bar{K}. If we show that for small \delta, x^{\delta} belongs to \mathcal{Q}, then the degree of \psi is nonzero, x^{\delta} \in U(x^{\ast}), hence q^{0,0} \in U(q^{0,\ast}) and the theorem is proved.

5.7. Insiders’ Strategies in a Quasi-Perfect Equilibrium, and the induced Lexicographic Probability System. Let \tilde{b}^{\varepsilon,\delta} be a sequence of completely mixed behavioral strategies converting to \tilde{b}^{\delta} against which for each insider \( n \) and each information set of \( n \) in \( \Gamma^{\delta} \), his continuation strategy is optimal. We now study the important features of both the sequence and its limit. If \( n \) chooses \( s_{n} \in S_{n}^{0} \) in the first stage, then in the second stage he has the option of revising this strategy to play something in \( T_{n}^{0}(\delta) \). Therefore, quasi-perfection implies that player \( n \) will end up implementing \( s_{n} \) with positive probability in \( \tilde{b}_{n}^{\delta} \) only if this strategy is at least as good a reply against the sequence \( \tilde{b}^{\varepsilon,\delta} \) as the strategies in \( T_{n}^{0}(\delta) \). Likewise, player \( n \) has the option of playing a strategy in \( T_{n}^{1}(\delta) \) before he decides to play a strategy in \( S_{n}^{1} \) and even when he makes a choice among these strategies, he has the option of choosing a strategy in \( T_{n}^{1}(\delta) \). Therefore, at the information set that follows his choice of avoiding strategies in \( S_{n}^{0} \), \( \tilde{b}_{n}^{\delta} \) assigns a positive probability to choosing a strategy in \( T_{n}^{0}(\delta) \cup S_{n}^{1} \) only if one of these strategies is at least as good a reply against the sequence \( \tilde{b}^{\varepsilon,\delta} \) as all the strategies in \( T_{n}^{0}(\delta) \). Furthermore, at the information set obtained after \( n \) avoids strategies in \( S_{n}^{0} \cup T_{n}^{0}(\delta), \tilde{b}_{n}^{\delta} \) assigns a positive probability to a strategy \( s_{n} \) in \( S_{n}^{1} \) only if \( s_{n} \) is at least as good a reply against the sequence \( \tilde{b}^{\varepsilon,\delta} \) as all the strategies in \( T_{n}^{1}(\delta) \) and \( S_{n}^{1} \).

Let \( \bar{\sigma}^{\varepsilon,\delta} \) be a sequence of mixed strategy profiles that is equivalent to the sequence \( \tilde{b}^{\varepsilon,\delta} \) of behavioral strategy profiles. For each player \( n \), his strategy \( \bar{\sigma}_{n}^{\varepsilon,\delta} \) in the sequence is a mixture over his pure strategy set \( \tilde{S}_{n} = S_{n} \cup T_{n}^{0}(\delta) \cup T_{n}^{1}(\delta) \). However, the implications of \( n \)'s strategy (for \( m \)'s choices) depend on the choices of the outsiders through their implications for strategies in \( T_{n}^{0}(\delta) \) and \( T_{n}^{1}(\delta) \). Each strategy \( t_{n}^{i}(\delta) \) plays \( t_{n} \) with probability \( (1 - \delta) \) and with probability \( \delta \) plays a strategy in \( P_{n}^{i} \) that is determined by \( o_{n,1}^{i} \)'s strategy. In order to fully capture the impact that \( o_{n,1}^{i} \) has on \( t_{n}^{i}(\delta) \), let \( W_{n}^{i} \) be the vertex set of \( \bar{K}_{n}^{i} \). Let \( \bar{T}_{n}^{i}(\delta) \) be the union over all \( w_{n}^{i} \in W_{n}^{i} \) of the sets \( T_{n}(\delta, p_{n}^{i}(w_{n}^{i})) \). Let \( \bar{S}_{n} = S_{n} \cup \bar{T}_{n}^{0}(\delta) \cup \bar{T}_{n}^{1}(\delta) \) and let \( \bar{\Sigma}_{n} \) be the set of mixtures over \( \bar{S}_{n} \).

Since \( W_{n}^{i} \) is the pure strategy set of outsider \( o_{n,1}^{i} \), \( \bar{\sigma}_{n}^{\varepsilon,\delta} \) has \( W_{n}^{i} \) as its support for all \( \varepsilon \) in the sequence. The sequence \( \bar{\sigma}^{\varepsilon,\delta} \) induces a mixed strategy \( \bar{\sigma}_{n}^{\varepsilon,\delta} \) in \( \tilde{S}_{n} \) for each \( n \) as follows. For each \( s_{n} \in S_{n} \), the probability \( \bar{\sigma}_{n}^{\varepsilon,\delta}(s_{n}) \) of \( s_{n} \) is \( \bar{\sigma}^{\varepsilon,\delta} \). For each \( i = 0, 1 \), \( w_{n}^{i} \in W_{n}^{i} \) and \( t_{n} \in T_{n}^{0} \),
the probability \( \tilde{\sigma}^\varepsilon_n(t_n^i(\delta, p_n^i(w_n^i))) = \tilde{\sigma}^\varepsilon_n(t_n^i(\delta)\sigma_{\varepsilon,1}^\varepsilon_n(w_n^i)) \). From player \( m \)'s perspective it is sequence \( \sigma_{n,1}^\varepsilon \), or rather its equivalent sequence in \( P_n \), that matters for his choice.

Using the method in Blume, Brandenburger, and Dekel [1, Appendix], from the sequence \( \tilde{\sigma}^\varepsilon_n \) construct for each player \( n \) a lexicographic probability system (LPS) \( \tilde{\Lambda}_n^\delta \) = \( (\tilde{\sigma}_{n,0}^\varepsilon, \ldots, \tilde{\sigma}_{n,m}^\varepsilon(\delta,\delta)) \) over his strategies in \( S_n \) such that for each \( \varepsilon \) in a subsequence,

\[
\tilde{\sigma}_{n}^\varepsilon(t_n^i(\delta, p_n^i(w_n^i))) = \frac{1}{1 + \nu_0(\varepsilon)}(\tilde{\sigma}_{n,0}^\varepsilon(t_n^i(\delta))\sigma_{\varepsilon,1}^\varepsilon(\delta,\delta) + \nu_1(\varepsilon)((1 - \nu_1(\varepsilon))\tilde{\sigma}_{n,1}^\varepsilon(t_n^i(\delta)) + \cdots + \nu_{n-1}(\varepsilon)\tilde{\sigma}_{n,n}^\varepsilon(t_n^i(\delta))),
\]

where \( (\nu_0(\varepsilon), \ldots, \nu_{n-1}(\varepsilon)) \) is a sequence in \( \mathbb{R}_{+}^{n} \) converging to the origin.

Quasi-perfection implies that \( \tilde{\sigma}_{n}^\varepsilon \) is equivalent to the enabling strategy \( q_{n,0}^\varepsilon \) in \( \Gamma \) and it is a lexicographic best reply to \( \tilde{\Lambda}_n^\delta \). Let \( l_{n,0}^\delta \) be the first level in \( \tilde{\Lambda}_n^\delta \) at which some strategy in \( T_{n}^\delta(\delta) \) appears with positive probability. Let \( l_{n,0}^\delta \) be the first level of \( \tilde{\Lambda}_n^\delta \) at which some strategy in \( S_n^1 \cup T_{n}^1(\delta) \) appears with positive probability.

From the LPS \( \tilde{\Lambda}_n^\delta \), construct an LPS \( \Lambda_n^\delta = (q_{n,0}^\delta, \ldots, q_{n,m}^\delta) \) for \( \Gamma \) where for each \( l \), \( q_{n,l}^\delta \) is an enabling strategy in \( \Gamma \) that is equivalent to \( \tilde{\sigma}_{n}^\varepsilon \). \( q_{n,l}^\delta \) is a lexicographic best reply against \( \Lambda_n^\delta \). For each \( w_n^i \in W_n^i \), the total probability under \( \sigma_{n,l}^{\varepsilon}(\delta,\delta) \) of the strategies in \( T_{n}^l(\delta, p_n^i(w_n^i)) \) is \( \sigma_{n,l}^{\varepsilon}(\delta,\delta) \). Since \( p_n^i(\sigma_{n,l}^{\varepsilon}(\delta,\delta)) \) is, by definition, \( p_n^i(\delta,\delta) \), we have that level \( l_{n,0}^\delta \) is expressible as a convex combination of \( \lambda_{l_{n,0}^\delta}^\varepsilon p_n^i + (1 - \lambda_{l_{n,0}^\delta}^\varepsilon) r_{n,0}^\delta l_{n,0}^\delta \) with \( \lambda_{l_{n,0}^\delta}^\varepsilon > 0 \) and \( r_{n,0}^\delta > 0 \).

We claim that for each \( l \leq l_{n,0}^\delta \), \( q_{n,l}^\delta \) induces the equilibrium outcome against \( q_{n,0}^\delta \). Indeed, since \( q_{n,0}^\delta \) belongs to \( Q^* \), and the strategies in \( T_{n}^0(\delta) \) always implement some strategy in \( P_n^0 \) regardless of the choice of outsider \( o_{n,0,1}^\delta \), all the strategies in \( T_{n}^0(\delta) \) are optimal against \( q_{n,0}^\delta \). By quasi-perfection, this then implies that every strategy in \( S_n \) that appears at a level \( l \leq l_{n,0}^\delta \) of \( \tilde{\Lambda}_n^\delta \) must be a best reply to \( q_{n,0}^\delta \). Thus for all \( l \leq l_{n,0}^\delta \), \( q_{n,l}^\delta \) is a best reply to \( q_{n,0}^\delta \). Since \( q_{n,0}^\delta \) is a lexicographic best reply to \( \Lambda_n^\delta \), \( (\tilde{q}_n(\varepsilon), q_{n,0}^\delta) \) is an equilibrium of the game \( \Gamma \) for all small \( \varepsilon \), where

\[
\tilde{q}_n(\varepsilon) = (1 - \varepsilon) q_{n,0}^\delta + \varepsilon ((1 - \varepsilon) q_{n,1}^\delta + \varepsilon^2 ((1 - \varepsilon) q_{n,2}^\delta + \varepsilon^3 (\cdots \varepsilon_{n-1}^{\varepsilon} q_{n,n}^\delta)).
\]

By genericity, \( \Gamma \) has finitely many equilibrium outcomes, so each of these equilibria induces the same equilibrium outcome—hence the claim follows. Three implications of this claim are: (i) \( l_{n,0}^\delta > l_{n,0}^\delta \); (ii) the enabling strategy \( r_{n,0}^\delta l_{n,0}^\delta \) in the previous paragraph belongs to \( P_n^0 \); (iii) all levels up to \( l_{n,0}^\delta \) prescribe the same mixture at each information set on the equilibrium path and differ only at information sets excluded by (all of) \( m \)'s equilibrium strategies in \( Q_n^* \), which is the \( \rho_m \)-image of the component \( \Sigma_n^* \) of equilibria.

5.8. Limits of the Lexicographic Probability Systems. Take a subsequence of the \( \delta \)'s such that the following properties hold for the associated LPSs \( \Lambda_n^\delta \) for each \( n \): (i) \( l_{n,0}^\delta \) is
independent of $\delta$ for each $i$; (ii) for each $l \leq l_1^{1,\delta}$, the face of $P_n$ that contains $q_l^{1,\delta}$ in its interior as well the strategies for $m$ that are best replies to $q_l^{1,\delta}$ are independent of $\delta$.

Let $\sigma_n^{l_1,\delta}$ be a sequence of strategies in $\Sigma_n$ that is equivalent to the sequence $q_n^{l_1,\delta}$. Using BBD again, for a subsequence of $\sigma_n^{l_1,\delta}$ there is now an LPS $(\sigma_n^{l_1,0}, \ldots, \sigma_n^{l_1,k_n})$ and a sequence $(\mu_0(\delta), \ldots, \mu_{k_n-1}(\delta)) \in \mathbb{R}^{k_n}$ converging to zero such that for a subsequence of $\delta$’s, $\sigma_n^{l_1,\delta}$ is expressible as the nested combination

$$\sigma_n^{l_1,\delta} = (1-\mu_0(\delta))(\sigma_n^{l_1,0} + \mu_0(\delta)((1-\mu_1(\delta))\sigma_n^{l_1,1} + \mu_1(\delta)((1-\mu_2(\delta))\sigma_n^{l_1,2} + \cdots + \mu_{k_n-1}(\delta)\sigma_n^{l_1,k_n}))$$

This LPS induces an equivalent LPS $(q_n^{l_1,0}, \ldots, q_n^{l_1,k_n})$ in enabling strategies.

Let $k_1^{1}$ be the first level $k$ of this LPS where $q_n^{l_1,k}$ does not belong to $P_n^0$. Recall from the previous section that $q_n^{l_1,\delta}$ is expressible as a convex combination $\lambda_n^{l_1,\delta}p_n^{1,\delta} + (1 - \lambda_n^{l_1,\delta})r_n^{1,\delta}$. Express $r_n^{1,\delta}$ as a convex combination $\alpha_n^{0,\delta}r_n^{0,\delta} + \alpha_n^{1,\delta}r_n^{1,\delta}$, where for $i = 0, 1$, $r_n^{i,\delta} \in P_n^i$. Let $\lambda_n^{1}$ be the limit of $(\lambda_n^{l_1,\delta} + (1 - \lambda_n^{l_1,\delta})\alpha_n^{k_n,\delta})^{-1}\lambda_n^{l_1,\delta}$ and let $r_n^{1,0}$ be the limit of $r_n^{1,\delta}$. Since the limit of $p_n^{1,\delta}$ is $p_n^{0,0}$, we have that $q_n^{l_1,k_n}$ is expressible as a convex combination $\zeta_n(\lambda_n^{1}p_n^{0,0} + (1 - \lambda_n^{1})r_n^{1,0}) + (1 - \zeta_n)p_n^{0,0}$ for some $\zeta_n > 0$ and $p_n^{0,0} \in P_n^0$.

For each $\delta$ in the subsequence used above, define now an LPS $\hat{\Lambda}_n^{\delta} = (\hat{q}_n^{0,0}, \ldots, \hat{q}_n^{l_1^{1,\delta}+k_n,\delta})$ as follows: $\hat{q}_n^{0,0} = q_n^{0,0}$; $\hat{q}_n^{l_1,\delta} = q_n^{l_1,\delta}$ if $0 < l < l_1^{1}$; $\hat{q}_n^{l_1^{1,\delta}+l,\delta} = q_n^{l_1^{1,\delta}+l,\delta}$ otherwise. The next claim sets out the key properties of $\hat{\Lambda}_n^{\delta}$ that will lead to a conclusion of our proof.

**Claim 5.3.** For all small $\delta$, the LPS $\hat{\Lambda}_n^{\delta}$ satisfies the following properties.

1. A strategy is at least as good a reply against $\Lambda_n^{\delta}$ than another only if it is at least as good a reply against $\hat{\Lambda}_n^{\delta}$.
2. If $\lambda_n^{1} < 1$, then the strategy $r_n^{1,0}$ is a best reply to $q_n^{0,0}$ and is at least good a reply lexicographically against $\hat{\Lambda}_n^{\delta}$ as every strategy in $S_n^{1}$.
3. $\lambda_n^{1} > 0$ and every level $l < l_1^{1} + k_1^{1}$ induces the equilibrium outcome against $q_n^{0,0}$.
4. The strategy $q_n^{l,\delta}$ for $l < l_0^{1}$ and the strategy $r_n^{l_1^{1,\delta}}$ if $\lambda_n^{l_1^{1,\delta}} < 1$ are lexicographic best replies against $\hat{q}_n^{l_1^{1,\delta}}$.

**Proof of the Claim.** Suppose $s_m$ is a better reply against $\hat{\Lambda}_n^{\delta}$ than another strategy $t_m$. We will show that $s_m$ is a better reply against $\Lambda_n^{\delta}$. Let $l$ be the first level of $\Lambda_n^{\delta}$ such that $s_m$ is a better reply against $\hat{q}_n^{l,\delta}$ than $t_m$. If $l < l_1^{1}$, then obviously $s_m$ is better reply against $\Lambda_n^{\delta}$ than $t_m$. Suppose then that $l_1^{1} \leq l \leq l_1^{1} + k_1^{1}$. Then for all small $\delta$, $s_m$ is a better reply against $q_n^{l_1^{1,\delta}}$ since $q_n^{l,\delta}$ is a nested combination of $(q_n^{l_1^{1,0}}, \ldots, q_n^{l_1^{1,k_n}})$. Thus, $s_m$ is a better reply against $\Lambda_n^{\delta}$. 

AXIOMATIC EQUILIBRIUM SELECTION 21
In the game $\tilde{\Gamma}^\delta$, player $n$ when finally making a choice among the strategies in $S^1_n \cup T^1_n(\delta)$ would choose a strategy $s_n$ in $S^1_n$ with positive probability only if it is at least good a reply against $\Lambda^\delta_n$ as the strategies the other strategies in $S^1_n$. Therefore, such a strategy would show up with positive probability under level $l^1_n$ of $\tilde{\Lambda}^\delta_n$ (and hence in $\tilde{\Lambda}^\delta_n$) only if this is the case. This implies that if $(1 - \mu_n^{l^1_n,\delta}) \alpha_n^1$ is positive, then the strategy $r_n^{1,\delta}$ is at least as good a reply against $\Lambda^\delta_n$ as the strategies in $S^1_n$. $r_n^{1,0}$ is the limit of $r_n^{1,\delta}$ and $\lambda^1_n$ is the limit of $(\mu_n^{l^1_n,\delta} + (1 - \mu_n^{l^1_n,\delta}) \alpha_n^1) \mu_n^{l^1_n,\delta}$. Therefore, by point (1) of this claim, $r_n^{1,0}$ is at least good a reply against $\tilde{\Lambda}^\delta_n$ as strategies in $S^1_n$ if $\lambda^1_n < 1$. To show that it is also a best reply against $q_n^{0,0}$, suppose to the contrary that it is not. Then every strategy in $T^1_n(\delta)$ regardless of outsider $a_{n,1}$’s choice is a better reply against $q_n^{0,\delta}$ for all small $\delta$, since for $\delta = 0$, these strategies are in $P^0_n$, which are best replies to $q_n^{0,0}$. Therefore, quasi-perfection implies that $n$ would prefer to play the strategies in $T^1_n(\delta)$ rather than implementing $r_n^{1,0}$ (or $r_n^{1,\delta}$ when $\delta$ is small), which shows that $(1 - \mu_n^{l^1_n,\delta}) \alpha_n^1 = 0$ for all small $\delta$ and hence that $\lambda^1_n = 1$.

We turn now to point (3). Every strategy $\tilde{q}_m^{l^1_n,\delta}$ for $l < l^1_n + k^1_n$ belongs to $P^0_n$ and is thus optimal against $q_m^{0,0}$, which belongs to $Q^*_n$. The strategy $\tilde{p}_n^0$ (which recall is part of the expression defining $q_m^{l^1_n,k^1_n}$ which is chosen with positive probability is also in $P^0_n$ and hence optimal against $q_m^{0,0}$. As we saw at the end of the last paragraph, the strategies in $T^1_n(\delta)$ are optimal against $q_m^{0,0}$. Thus, if $\mu_n < 1$, the strategy $r_n^{1,0}$ must also be optimal against $q_m^{0,0}$. Obviously $q_m^{0,0}$ is optimal against $\tilde{\Lambda}^\delta_n$ since it is the limit of $q_m^{0,\delta}$, which by point (1) is optimal against $\tilde{\Lambda}^\delta_n$. Therefore, for all small $\varepsilon$, $(q_m^{0,0}, q_n(\varepsilon))$ is an equilibrium of $\Gamma$, where

$$\tilde{q}_n(\varepsilon) = \left( \sum_{l=0}^{l^1_n + k^1_n - 1} \varepsilon^l + (1 - 1_{\lambda_n > 0}) \varepsilon^{l^1_n + k^1_n} \right)^{-1} \sum_{l=0}^{l^1_n + k^1_n - 1} \varepsilon^l q_m^{l^1_n + k^1_n} + (1 - 1_{\lambda_n > 0}) \varepsilon^{l^1_n + k^1_n} q_n^{l^1_n + k^1_n}.$$  

This is impossible if $\lambda_n = 0$, since $q_m^{l^1_n + k^1_n}$ is a convex combination of a strategy in $P^0_n$ and one in $P^1_n$. If $\lambda_n = 0$, this implies that all strategies at levels preceding $l^1_n + k^1_n$ induce the equilibrium outcome against $q_m^{0,0}$.

We prove point (4). By the previous paragraph, all strategies in $P^0_n$ are optimal against $q_m^{0,0}$ for $k < l^1_n + k^1_n$. Thus the optimality of a strategy in $P^0_n$ depends on how it fares against $q_m^{l^1_n + k^1_n}$. If a strategy $s_n \in S_m^n$ is not optimal against this strategy, then for all small $\delta$, it is not optimal against $\Lambda^\delta_n$. But that implies that for all small $\delta$, some strategy in $T^0_n(\delta)$ is a superior reply to $q_m^{l^1_n,k^1_n}$ regardless of what outsider $o_{n,1}$’s choice is. Therefore, in the quasi-perfect equilibrium of $\Gamma^\delta$, player $n$ when given a choice of actually implementing $s_n$
will prefer some strategy in $T^0_n(\delta)$, i.e., this strategy’s probability is zero for all $l \leq t_n^0$. Point (4) now follows. □

5.9. Final Step of the Proof. Fix small $\delta$ that satisfies the conditions of the claim from the previous subsection. Let

$$\bar{q}^0_n(\varepsilon) = \left( \sum_{l \leq t_n^0} \varepsilon^l \right)^{-1} \sum_{l \leq t_n^0} \varepsilon^l \bar{q}^{l,\delta}_n,$$

$$\bar{q}^1_n(\varepsilon) = \left( \sum_{l^0+1 \leq t_n^1+k_n^1} \varepsilon^l \right)^{-1} \sum_{l^0+1 \leq t_n^1+k_n^1} \varepsilon^l \bar{q}^{l,\delta}_n.$$ 

Observe that $\bar{q}^0_n(\varepsilon)$ belongs to $P^0_n$ for all $\varepsilon$ and it is a convex combination of $p_n^{0,\delta}$ and a subset $R^0_n$ of strategies that are at least as good replies against $\Lambda_n^\delta$ as other strategies in $S_n^0$. Likewise $\bar{q}^1_n(\varepsilon)$ is a convex combination of $p_n^1$, $r_n^1$ and a point in $P^0_n$ such that the strategy $r_n^1$ if it has a positive weight is at least as good a reply against $\Lambda_n^\delta$ as other strategies in $P^1_n$. There exists a small $\varepsilon$ such that a strategy $s_m$ is at least as good as a strategy $t_m$ against $\Lambda_n^\delta$ iff it is lexicographically at least as good a reply against the LPS $(q^0, \bar{q}^0_n(\varepsilon), \bar{q}^1_n(\varepsilon))$. Since $\pi_n^1(\bar{q}^1_n(\varepsilon)) = \pi_n^1(q^{1,0}) = \pi^{1,0}$ for all small $\varepsilon$, it follows that $(q^0, (p^{0,\delta}, p^{1,0}), \pi^{1,0})$ belongs to $Q$. As we noted earlier, proving that this point belongs to $Q$ shows that in fact it belongs $U(x^*)$ and hence that $q^0$ in $U(q^{0,*})$, which completes the proof. □

6. Concluding Remarks

The characterization in this paper is a step toward a theory of equilibrium refinement using axioms adapted from decision theory. Theorem 5.1 is confined to games in extensive form with perfect recall, two players, and generic payoffs, but it suggests that an extension to more general games might be possible. We hope also that Axiom B can be weakened to require only that a solution contains a sequential equilibrium for which each player’s strategy is conditionally admissible from each of his information sets — thus keeping the axioms entirely within standard decision theory.

Previously, some proposed refinements selected equilibria with one or more desirable properties, like admissibility, subgame perfection, or sequential rationality. Other proposed refinements derived some properties from limits of equilibria of perturbed games, such as perfect, quasi-perfect, and proper equilibria. In the latter approach, a key step forward was Kohlberg and Mertens’ [13] argument that an axiomatic development requires set-valued refinements. Their program achieved remarkable success with Mertens’ [18] definition of a stable set, which has the desirable properties 1, 2, 3 listed in Section 1 and others too, such as ordinality and immunity to splitting players into agents.
However, an axiomatic theory of refinement should be based on basic principles of rational behavior in the game at hand, as in decision theory. This precludes reliance on perturbed games obtained by perturbing players’ sets of feasible strategies. The challenge, therefore, has been to establish why consideration of perturbed games yields the requisite decision-theoretic properties.

Our answer here begins with Axiom S, which is a generalization of Kohlberg and Mertens’ [13] invariance criterion, and Mertens’ [20] small worlds criterion, as explained in Section 3.3. Absent a strong invariance property like Axiom S, a refinement is vulnerable to ‘framing effects’ depending on wider contexts in which the given game might be embedded. In decision theory, such effects were examined by Savage [26], and in cognitive psychology they play a prominent role in studies of decisions by subjects in experiments, as for instance by Kahneman and Tversky [12] who emphasize subjects’ tendency to be influenced by an exogenously specified status quo, or other reference point affecting perceptions of gains and losses, entitlements due to initial endowments, etc. For a theory of thoroughly rational behavior, however, an axiom should exclude framing effects. Axiom S does this by requiring a solution of a game to be consistent with the solution of any metagame in which it is imbedded. As shown in Proposition 3.4, it is already true of any equilibrium that it is consistent with an equilibrium of any metagame in which the game is imbedded. Axiom S merely extends to refinements this fundamental invariance property of equilibria.

Our answer continues with the proof of Theorem 4.1 in Appendix B. There it is shown that a set $Q^*$ of equilibria in enabling strategies is stable iff the corresponding projection map from the pseudomanifold $Q$ to the space $P$ of enabling strategies is essential. Using this key property, the proof of the main theorem shows that for each equilibrium in a component of undominated equilibria there exists a corresponding metagame for which the equilibrium is the image of a quasi-perfect equilibrium in the metagame if and only if the projection map is essential. Hence Axioms A, B, S imply that a solution is a stable set, and conversely due to Mertens’ previous proofs. The answer to the ‘why’ question above is thus that, given Axioms A and B, stability with respect to perturbed games is equivalent to an analogous ‘stability’ with respect to embeddings in metagames, as required by Axiom S.

Theorem 5.1 establishes that a solution of a game must be a component of undominated equilibria, and that the component must be essential. Because payoffs are assumed to be generic, all equilibria in the component yield the same distribution of outcomes and thus the same paths of equilibrium play. Therefore the main implication for selection is that a predicted outcome distribution must result from equilibria in an essential component of undominated equilibria, and in particular, from the sequential equilibria it necessarily contains.
The secondary implication is that other equilibria in the component, which are relevant only off the paths of equilibrium play, are actually sequentially rational with respect to beliefs induced by some metagame. Thus, even though common knowledge of rationality precludes a deviation in the given game, after a deviation there are continuations that are sequentially rational if viewed in the context of some metagame that embeds the given game.

APPENDIX A. ENABLING STRATEGIES

In the normal-form representation of a game in extensive form, a player’s pure strategy specifies the action chosen at each of his information sets in the game tree. However, outcomes are not affected by a strategy’s actions at information sets excluded by his previous actions. One therefore considers equivalence classes of pure strategies. Say that two pure strategies are outcome equivalent if the sets of terminal nodes they do not exclude are the same. For instance, the game in Figure 1 is shown on the left side in extensive form and on the right side in the ‘pure reduced normal form’ (PRNF) introduced by Mailath, Samuelson, and Swinkels [17]. In the PRNF, each outcome-equivalent class of player 1’s pure strategies is identified by the terminal nodes it does not exclude, as indicated by labels of rows along the left side; and each equivalence class of player 2’s pure strategies is identified by the terminal nodes it does not exclude, as indicated by labels of columns along the top. Because this game has no moves by Nature, each row and column determine a unique outcome that is the intersection of the row and column labels, shown as the corresponding entry in the matrix.

A similar example is shown in Figure 2 for the game tree of a signaling game. In this case, each profile of pure strategies determines a pair of outcomes such that the first or second outcome occurs depending on whether Nature’s initial move is up or down.
Say that a terminal node that is not excluded is an enabled outcome. A pure strategy of a player enables outcome \( z \) if it chooses all his actions on the path to \( z \). A player’s mixed strategy randomizes over his pure strategies, whereas a behavioral strategy randomizes over actions at each of his information sets. A strategy of either kind induces a probability distribution over outcome-equivalent classes of his pure strategies, and thus a distribution over enabled outcomes. Such a distribution is called an enabling strategy. A point \( p_n \in [0, 1]^Z \) is an enabling strategy for player \( n \) if it is the distribution over enabled outcomes induced by some mixed or behavioral strategy, i.e., \( p_n(z) \) is the mixed strategy’s probability of those pure strategies that enable outcome \( z \). The vertices of the polyhedron \( P_n \) of \( n \)'s enabling strategies correspond to outcome-equivalent classes of \( n \)'s pure strategies in the PRNF, as in Figures 1 and 2. Enabling strategies are minimal representations of strategic behavior in games with perfect recall.

Let \( p_*(z) \) be the probability that Nature’s strategy enables outcome \( z \), which is 1 if Nature has no moves. Then for each profile \( p \in P = \prod_n P_n \) of players’ enabling strategies, the probability that outcome \( z \) results is \( \gamma_z(p) = p_*(z) \prod_n p_n(z) \), because Nature and the players randomize independently. The extensive form is therefore summarized by the multilinear function \( \gamma : P \to \Delta(Z) \subset \mathbb{R}^Z \) that assigns to each profile of players’ enabling strategies a distribution over terminal nodes, including the effect of Nature’s enabling strategy. Player \( n \)'s expected payoff is \( G_n(p) = \sum_z \gamma_z(p)u_n(z) \). The game \( \Gamma \) is therefore summarized by the multilinear function \( G : P \to \mathbb{R}^N \) that assigns to each profile of players’ enabling strategies their expected payoffs. This summary specification is called the enabling form of the game.

\[\text{Figure 2. The game tree of a signaling game and its pure reduced normal form in which each pure strategy is identified by the terminal nodes it does not exclude.}\]
Appendix B. Proofs of Propositions

B.1. Proof of Proposition 3.3. A multilinear map \( f_n : \tilde{\Sigma}_n \times \Sigma_o \rightarrow \Sigma_n \) is completely specified by its values at profiles of pure strategies. We use \( \hat{f}_n \) to denote the restriction of \( f_n \) to the set \( \tilde{S}_n \times S_o \) of profiles of pure strategies.

Proposition B.1. \( \tilde{G} \) embeds \( G \) via a collection of multilinear maps \( f = (f_n)_{n \in N} \) if and only if for each player \( n \) there exists \( \tilde{T}_n \subseteq \tilde{S}_n \) and a bijection \( \pi_n : \tilde{T}_n \rightarrow S_n \) such that for each \((\tilde{s}, s_o) \in \tilde{S} \times S_o \) and \( \tilde{t}_n \in \tilde{T}_n \):

1. \( \hat{f}_n(\tilde{t}_n, s_o) = \pi_n(\tilde{t}_n), \)
2. \( \hat{G}_n(\tilde{s}, s_o) = G_n(\hat{f}(\tilde{s}, s_o)) \), where \( \hat{f} = (\hat{f}_n)_{n \in N} \).

Proof. Suppose we have a game \( \tilde{G} : \tilde{\Sigma} \times \Sigma_o \rightarrow \mathbb{R}^{N \cup o} \) and a collection of multilinear maps \( f_n : \tilde{\Sigma}_n \times \Sigma_o \rightarrow \Sigma_n \), one for each \( n \in N \), such that conditions (1) and (2) of the proposition are satisfied. Then, by condition (1) and multilinearity of \( f_n \) for each \( n \), for each fixed \( \sigma_o \), \( f_n(\cdot, \sigma_o) \) is surjective because it maps the face spanned by \( \tilde{T}_n \) homeomorphically onto \( \Sigma_n \). Also, condition (2) and multilinearity of each \( f_n \) imply that \( \tilde{G} = G \circ \hat{f} \). According to Definition 3.2, therefore, \( (\tilde{G}, f) \) embeds \( G \).

Now suppose that \( (\tilde{G}, f) \) embeds \( G \). Let \( \sigma_o \) be a profile of completely mixed strategies for outsiders. Because \( f_n \) is multilinear it induces a linear mapping \( f_n(\cdot, \sigma_o) \) from \( \tilde{\Sigma}_n \) to \( \Sigma_n \) that, by the definition of an embedding, is surjective. Hence, for each \( s_n \in S_n \) there exists a pure strategy \( \tilde{t}_n(s_n) \) in \( \tilde{S}_n \) that is mapped to \( s_n \) by this linear map. We claim that \( f_n(\tilde{t}_n(s_n), s_o) = s_n \) for all \( s_o \in S_o \). Indeed, observe that \( f_n(\tilde{t}_n(s_n), \sigma_o) = \sum s_o f_n(\tilde{t}_n(s_n), s_o)\sigma_o(s_o) \), where for each \( s_o, \sigma_o(s_o) \) is the probability of \( s_o \) under \( \sigma_o \). Therefore, since \( \sigma_o \) is completely mixed, if \( f_n(\tilde{t}_n(s_n), s_o) \neq s_n \) for some \( s_o \) then \( f_n(\tilde{t}_n(s_n), \sigma_o) \), which is an average of values at vertices of \( S_o \), cannot be \( s_n \). Thus, \( f_n(\tilde{t}_n(s_n), s_o) = s_n \) for all \( s_o \). Let \( \tilde{T}_n \subset \tilde{S}_n \) be a collection comprising a different pure strategy \( \tilde{t}_n(s_n) \) for each \( s_n \in S_n \) and let \( \pi_n \) be the associated bijection. Define \( \hat{f}_n : \tilde{S}_n \times S_o \rightarrow \Sigma_n \) by \( \hat{f}_n(\tilde{s}_n, s_o) = f_n(\tilde{s}_n, s_o) \). Then conditions (1) and (2) of the proposition are satisfied. \( \square \)

B.2. Proof of Proposition 3.4.

Proposition B.2. If \( (\tilde{G}, f) \) embeds \( G \) then the equilibria of \( G \) are the \( f \)-images of the equilibria of \( \tilde{G} \).

Proof. Suppose \( (\tilde{\sigma}, \sigma_o) \) is an equilibrium of \( \tilde{G} \) and let \( \sigma = f(\tilde{\sigma}, \sigma_o) \). For any insider \( n \) and his strategy \( \tau_n \in \Sigma_n \) there exists \( \tilde{\tau}_n \in \tilde{\Sigma}_n \) such that \( f_n(\tilde{\tau}_n, \sigma_o) = \tau_n \) because \( f_n(\cdot, \sigma_o) \) is surjective.
by condition (a) of Definition 3.2 an embedding. Using condition (b),
\[ G_n(\tau_n, \sigma_{-n}) = G_n(f(\tau_n, \tilde{\sigma}_{-n}, \sigma_o)) = \hat{G}_n(\tau_n, \tilde{\sigma}_{-n}, \sigma_o) \leq \hat{G}_n(\tilde{\sigma}, \sigma_o) = G_n(f(\tilde{\sigma}, \sigma_o)) = G_n(\sigma), \]
where the inequality obtains because \((\tilde{\sigma}, \sigma_o)\) is an equilibrium of \(\hat{G}\). Hence \(\sigma\) is an equilibrium of \(G\).

Conversely, suppose \(\sigma\) is an equilibrium of \(G\). For each \(n\), let \(\pi_n\) be the bijection given by Proposition 3.3. Let \(\tilde{\sigma}_n\) be the strategy for insider \(n\) in \(\tilde{G}\) defined by \(\tilde{\sigma}_n(\tilde{t}_n) = \sigma_n(\pi_n(\tilde{t}_n))\). Since \(f_n\) is multilinear, by condition (1) of Proposition 3.3, \(f_n(\tilde{\sigma}_n, \cdot) = \sigma_n\) and thus \(f(\tilde{\sigma}, \cdot) = \sigma\). Hence, it suffices to show that there exists a strategy profile \(\sigma_o\) for outsiders such that \((\tilde{\sigma}, \sigma_o)\) is an equilibrium of \(\tilde{G}\). By fixing the profile of insiders’ strategies to be \(\tilde{\sigma}\) one induces a game among outsiders. Let \(\sigma_o\) be an equilibrium of this induced game among outsiders. To see that \((\tilde{\sigma}, \sigma_o)\) is an equilibrium of \(\tilde{G}\), observe that for each pure strategy \(\tilde{s}_n\) of an insider \(n\):
\[ \hat{G}_n(\tilde{s}_n, \tilde{\sigma}_{-n}, \sigma_o) = G_n(f_n(\tilde{s}_n, \sigma_o), \sigma_{-n}) \leq G_n(\sigma) = G_n(f(\tilde{\sigma}, \sigma_o)) = \hat{G}_n(\tilde{\sigma}, \sigma_o), \]
where the first and second equalities use the property \(f(\tilde{\sigma}, \cdot) = \sigma\) established above, and the inequality obtains because \(\sigma\) is an equilibrium of \(G\).

\[\square\]

Appendix C. Proof of Theorem 4.1

For any set \(X\), we write \(d(X)\) for its dimension. For any subset \(T_n\) of \(S_n\), let \(P_n(T_n)\) be the convex hull of the strategies in \(T_n\). For simplicity, we write \(d(T_n)\) for \(d(P_n(T_n))\).

**Theorem C.1.** \((Q, \partial Q)\) is a pseudomanifold of the same dimension as \((\mathbb{P}, \partial \mathbb{P})\). Moreover, \(\psi : (Q, \partial Q) \rightarrow (\mathbb{P}, \partial \mathbb{P})\) is an essential map iff \(Q^*\) is stable.

Proof. For any vertex \(\pi^1_n\) of \(\Pi^1_n\), let \(S_n^1(\pi^1_n)\) be the set of pure strategies that map to \(\pi^1_n\). For any face \(\Psi^1_n\) of \(\Pi^1_n\), let \(S_n^1(\Psi^1_n)\) be the set of strategies \(s_n^1\) such that \(\pi^1_n(s_n^1) \in \Psi^1_n\).

Let \(H^0_n\) be the set of information sets \(h_n\) of player \(n\) such that its immediate predecessor, if any, is on the equilibrium path and player \(n\)’s choice there leading to \(h_n\) belongs to \(A^*_n\). If a subset \(T^0_n\) of \(S^0_n\) is such that \(P_n(T^0_n)\) contains an equilibrium in \(Q^*_n\), then for each \(h_n \in H^0_n\) there exists a strategy in \(T^0_n\) that enables \(h_n\).

Let \(H^1_n\) be the set of information sets \(h_n\) of player \(n\) such that its immediate predecessor is on the equilibrium path and player \(n\)’s action there that leads to \(h_n\) is not in \(A^*_n\). Observe that an information set does not belong to or follow an information set in \(H^0_n\) or \(H^1_n\) iff it belongs to \(H^*_n\).
Lemma C.2. Suppose $T_n^0$ is a subset of strategies in $S_n^0$ such that $P_n(T_n^0)$ contains an equilibrium $q_n^*$ in $Q_n^*$. If the strategies in $T_n^0$ are at least as good replies against $p_m^0 \in P_m^0$ as other strategies in $S_n^0$, then all the strategies in $S_n^0$ are equally good replies against $p_m^0$.

Proof. Let $\Pi_n$ be the projection of $P_n^0$ to $\mathbb{R}^{Z_n^*}$, where $Z_n^*$ is the set of terminal nodes reached with positive probability under the equilibria in $\bar{Q}^*$. Then, $Z_n^* = Z_n^0 \cap Z_n^m$. Consequently, the payoff to a strategy $s_n^0$ against $p_m^0$ depends on $s_n^0$ only through its projection to $\Pi_n^*$.

For the next lemma, it is worth recapitulating the exact definition of the set $\Pi_n^1$. Recall from Section 4 that we fix a completely mixed enabling strategy $\bar{p}_m$ for player $m$ and compute for each $p_n$, the probability $\psi_{Z_n^1}(p_n)$ is the projection of $p_n$ to $\mathbb{R}^{Z_n^1}$ and $\eta(p_n)$ is the total probability of reaching a terminal node in $Z_n^1$ under $(\bar{p}_m, p_n)$, and $H$ separates $\psi_{Z_n^1}(P_n^1)$ from the origin of $\mathbb{R}^{Z_n^1}$. The function $\pi_n^1$ maps each point $p_n \in P_n \setminus P_n^0$ to $\varepsilon(\eta(p_n))^{-1} p_n$.

Lemma C.4. For a strategy $s_n \in S_n^1$, $\pi_n^1(s_n)$ is a vertex of $\Pi_n^1$ iff there exists a unique information set $h_n \in H_n^*$ that $s_n$ enables and where it chooses a non-equilibrium action.

Proof. Let $s_n \in S_n^1$ be a pure strategy satisfying the condition of our lemma. We prove by contradiction that $\pi_n^1(s_n)$ is a vertex of $\Pi_n^1$. Suppose, therefore, to the contrary that $\pi_n^1(s_n)$,
which equals $\varepsilon(\eta(p_n))^{-1}p_n$ is not a vertex of $\Pi_n^1$. We can express $\pi_n^1(s_n)$ as a unique convex combination $\sum_{j=1}^J \lambda^j \pi_n^1j$ where $J > 1$ and for each $1 \leq j \leq J$, $\pi_n^1j$ is a vertex of $\Pi_n^1$. For each $j$, since $\pi_n^1j$ is a vertex of $\Pi_n^1$, we can express $\pi_n^1j$ as $\varepsilon(\eta(p_n))^{-1}s_n^1j$ for some $s_n^1j$ in $S_n^1$. Since $s_n^1j$ cannot choose a non-equilibrium action at any $h_n' \neq h_n$ in $H_n^*$ that it enables; it belongs to $S_n^1j$ it must therefore enable $h_n$; and it cannot choose a different non-equilibrium action from $s_n$ at $h_n$. Modify $s_n^1j$ to a strategy $t_n^1j$ so that at every information set other than the successors to $h_n$, $t_n^1j$ agrees with $s_n$ and at the successors to $h_n$ it agrees with $t_n^1j$. It is now clear that when viewed as enabling strategies, $s_n = \sum_j \lambda^j s_n^1j$ and thus $s_n$ is a convex combination of the strategies $s_n^1j$.

To prove the other way around, suppose $s_n$ is a strategy that, at a collection $h_n^k$ for $k = 1, \ldots, K$ of at least two information sets in $H_n^*$, chooses a non-equilibrium action. For each $k$, choose a strategy $s_n^1k$ in $S_n^1$ that enables $h_n^k$, agrees with $s_n$ there and at all its successors, but at other $h_n \in H_n^*$, chooses an equilibrium action. Then $\psi_{Z_n^1}(s_n) = \sum_k \psi_{Z_n^1}(s_n^1)$. Therefore, $\pi_n^1(s_n)$ cannot be a vertex of $\Pi_n^1$.

For each $T_n^0 \subset S_n^0$ such that $P_n(T_n^0)$ is a face of $P_n^0$ and contains an equilibrium strategy for $n$, let $S_n^1(T_n^0)$ be the subset of $S_n^1$ consisting of strategies $s_n^1$ such that for each $h_n^0 \in H_n^0$ that $s_n^1$ does not exclude, $s_n^1$ agrees with some $t_n^0 \in T_n^0$ at $h_n$ and all its successors, where $t_n^0$ has the property that it does not exclude $h_n^0$. For a face $\Psi_n^1$ of $\Pi_n^1$, let $S_n^1(T_n^0, \Psi_n^1) = S_n^1(\Psi_n^1) \cap S_n^1(T_n^0)$. The following lemma provides an important feature of the set $S_n^1(T_n^0, \Psi_n^1)$.

**Lemma C.5.** The strategies in $T_n \equiv T_n^0 \cup S_n^1(T_n^0, \Psi_n^1)$ are the vertices of a face of $P_n$ whose dimension is $d(T_n) \equiv d(T_n^0) + d(\Psi_n) + 1$.

**Proof.** Let $\tilde{T}_n^1$ be the set of strategies $t_n^1$ in $T_n^1$ such $\tilde{\pi}_n^1(t_n^1)$ is a vertex of $\Psi_n^1$. We will first show that every $t_n^1 \in T_n \setminus (T_n^0 \cup \tilde{T}_n^1)$ is affinely dependent on the strategies in $T_n^0 \cup \tilde{T}_n^1$.

Let $t_n^1 \in T_n \setminus (T_n^0 \cup \tilde{T}_n^1)$. Let $h_n^1, \ldots, h_n^K$, $K > 1$, be information sets in $H_n^*$ such that for each $k$ $t_n^1$ chooses a nonequilibrium action $a_k$ at $h_n^k$. Fix $t_n^0 \in T_n^0$ that enables each of these information sets. For $k$, let $t_n^1k$ be the strategy that agrees with $t_n^1$ at $h_n^k$ and its successors, but everywhere else agrees with $t_n^0$. Each $t_n^1k$ belongs to $\tilde{T}_n^1$ and also, $t_n^1 = \sum_k t_n^1k - (K-1)t_n^0$. Thus, $t_n^1$ is an affine combination of the strategies $T_n^0 \cup \tilde{T}_n^1$.

For each $j = 0, \ldots, d(\Psi_n^1)$, pick a strategy $t_n^1j$ in $S_n^1(T_n^0)$ such that $\pi_n^1(t_n^1j)$ is a vertex of $\Psi_n^1$. Let $\tilde{T}_n^1$ be the collection of these strategies. We will show that strategies in $\tilde{T}_n^1 \setminus \tilde{T}_n^1$ are now affinely dependent on the strategies in $T_n^0 \cup \tilde{T}_n^1$. Fix $t_n^1 \in \tilde{T}_n^1 \setminus \tilde{T}_n^1$. There exists a unique information set $h_n \in H_n^*$ enabled by $t_n^1$ and where it chooses a nonequilibrium action. By construction of $\tilde{T}_n^1$, there exists a subset $(t_n^1j)_j$ of $\tilde{T}_n^1$ such that $\pi_n^1(t_n^1)$ is expressible as
a convex combination \( \sum_{j \in J} \lambda^j \pi(t^{1,j}_n) \) with \( \lambda^j > 0 \) for all \( j \in J \). For each \( j \), \( s^{1,j}_n \) enables \( h_n \) and chooses \( a \) at \( h_n \). Let \( t^0_n \) be a strategy that agrees with \( t^1_n \) everywhere except at \( h_n \) and its successors. For each \( j \in J \), let \( t^{0,j}_n \) be a strategy that agrees with \( t^{1,j}_n \) everywhere except at \( h_n \) and its successors, where it agrees with \( t^0_n \). Then \( t^1_n = t^0_n + \sum_{j \in J} \lambda^j (t^{1,j}_n - t^{0,j}_n) \) and is affinely dependent on the strategies in \( T^1_n \).  

It follows now from the above arguments that the affine space \( A \) spanned by \( T^0_n \cup \tilde{T}^1_n \) contains \( P_n(T_n) \) and that the dimension of \( P_n(T_n) \) is as stated. To finish the proof of the lemma, we will show that \( P_n(T_n) \) is a face of \( P_n \). Let \( Q_n \) be the smallest face of \( P_n \) that contains \( P_n(T_n) \). Suppose \( Q_n \neq P_n(T_n) \). Then there exists a point \( p_n \) in the relative interior of \( P_n(T_n) \) and \( Q_n \). \( p_n \) can therefore be expressed as a convex combination of the vertices of \( Q_n \) in two different ways: (a) \( \sum_i \lambda^i t^{0,i}_n + \sum_j \lambda^j t^{1,j}_n \), where the \( t^{0,i}_n \)'s are in \( T^0_n \) and the \( t^{1,j}_n \)'s are in \( T^1_n \); (b) \( \sum_i \mu^i t^{0,i}_n + \sum_j \mu^j t^{1,j}_n + \sum_k \mu^k t^{2,k}_n \), where now the \( t^{2,k}_n \)'s are the vertices of \( Q_n \) that are not in \( T_n \). Consider one of \( t^{2,k}_n \)'s. If it belongs to \( S^0_n \setminus T^0_n \), then there is an information set \( h_n \) in \( H^0 \) or a successor to one that is enabled by \( t^{2,k}_n \) where the action chosen by \( h_n \) is avoided all strategies in \( T^0_n \) that enable it: that implies under the expression in (b) the nodes following this action are assigned a positive probability but not under (a), which is impossible. If \( t^{2,k}_n \) belongs to \( S^1_n \setminus T^1_n \), then there exists an information set \( h_n \in H^0 \) enabled by \( t^{2,k}_n \) where the continuation strategy of \( t^{2,k}_n \) coincides with that of some \( t_n \in S^0_n \setminus T^0_n \) but not for any \( s_n \in T^0_n \). As in the previous case, this too is impossible.  

\[ \square \]

Let \( T \) be the set of pairs \( T \equiv ((T^0_1, \Psi^1_1), (T^0_2, \Psi^1_2)) \) of subsets such that for each \( n \): (i) \( \emptyset \neq T^0_n \subseteq T_n \), \( P_n(T^0_n) \) is a face of \( P_n \) and intersects \( Q^*_n \); (ii) \( \Psi^1 \) is a (possibly empty) face of \( \Pi^*_n \); (iii) there exist strategies in \( P_m(T^0_m) \) and in the interior of \( P^0_m \) and \( P_m \) against which all the strategies in \( T_n \equiv T^0_n \cup S^1_n(T^0_n; \Psi^1_n) \) are best replies. For simplicity in notation, we sometimes refer to \( S^1_n(T^0_n; \Psi^1_n) \) as \( T^1_n \).

Fix \( T \in T \). Let \( A_n(T) \) be the set of points in \( P_n(T^0_n) \) such that the strategies in \( T_m \) are all best replies. Let \( \tilde{T}^0_n \) be the unique subset of \( T^0_n \) such that the interior of \( A_n(T) \) is contained in the interior of \( P_n(\tilde{T}^0_n) \). Let \( \tilde{d}^*_n(T) \) be the dimension of \( A_n(T) \).

Let \( \tilde{T}^1_m \) be the set of strategies \( s_m \) in \( S^1_m(T^0_m) \setminus T^1_m \) such that: (i) \( s_m \) is an equally good reply against every point in \( P^0_n \) to which the strategies in \( T_m \) are equally good replies. Let \( \tilde{T}^1_m \) be the set of strategies \( s_m \) in \( S^1_m(T^0_m) \setminus (\tilde{T}^1_m \cup T^1_m) \) that are best replies against every point in \( A_n(T) \).

Let \( d^*_n(T) \) be the dimension of the set \( B^*_n(T) \) of points in \( P^0_n \) against which the strategies in \( T_m \) are equally good replies and at least as good replies as strategies in \( \tilde{T}^1_m \). If \( T^1_m \) is empty,
let $B_n(T) = P_n^0$. Otherwise, let $B_n(T)$ be the the closure of the set of points in the interior of $P_n^0$ that are of the form $\lambda q_n^0 + (1 - \lambda) r_n^0$ such that $\lambda \geq 1$, $q_n^0 \in B_n^*(T)$, and $r_n^0 \in P_n(T_n^0)$.

**Lemma C.6.** Suppose $T_m^1$ is nonempty. $B_n(T)$ is a polyhedron of dimension $d_n^*(T) + d_n(T_n^0) - \tilde{d}_n^*(T)$. Each maximal face $B_n'(T)$ of $B_n(T)$ satisfies exactly one of the following:

1. The relative interior of $B_n'(T)$ is contained in the relative interior of a maximal proper face of $P_n^0$.
2. There exists a strategy $r_m \in \tilde{T}_m^1$ such that for each $p_n^0 \in B_n'(T)$, $r_m$ is an equally good reply against every point of the form $\lambda p_n^0 + (1 - \lambda) r_n^0$ in $B_n^*(T)$; moreover, in this case, letting $\tilde{R}_m^1$ be the set of such $r_m$, for any $r_m \in \tilde{R}_m^1$, if $r_m$ is a best reply against a point in $B_n^*(T)$ then every point in $\tilde{R}_m^1$ is also a best reply against this point.
3. There exists a maximal proper face of $P_n(T_n^0)$, say $P_n(R_n^0)$, such that for each $q_n \in B_n^*(T)$, $r_n \in P_n(T_n^0)$ and $\lambda > 1$, if $\lambda q_n + (1 - \lambda) r_n$ belongs to $B_n'(T)$, then $r_n$ belongs to $P_n(R_n^0)$; moreover, $A_n(T)$ is contained in $P_n(R_n^0)$.

**Proof.** By definition, for every $\tilde{r}_m \in \tilde{T}_m^1$, there exists $q_n \in B_n^*(T)$ against which $\tilde{r}_m$ is inferior to strategies in $T_m$. Therefore, there exists $q_n \in B_n'(T)$ against which all the strategies in $\tilde{T}_m^1$ are inferior replies. Since there exists a point in $P_n^0 \setminus \partial P_n^0$ against which the strategies in $T_m$ are best replies, we can assume that $q_n$ is in the interior of $P_n^0$. For $r_n^0 \in P_n(T_n^0) \setminus \partial P_n(T_n^0)$ and $\lambda$ that is close to, but strictly greater than 1, $\lambda q_n + (1 - \lambda) r_n$ belongs to the interior of $B_n(T)$; as well, it belongs to the interior of the affine space spanned by $B_n^*(T)$ and $P_n(T_n^0)$. This last-mentioned affine space has dimension $d_n^*(T) + d_n(T_n^0) - \tilde{d}_n^*(T)$, since the intersection of $B_n^*(T)$ and $P_n(T_n^0)$ is exactly $A_n(T)$. Therefore, $B_n(T)$ has dimension $d_n^*(T) + d_n^0(T_n) - \tilde{d}_n^*(T)$, as asserted. We now show that $B_n(T)$ is actually a polyhedron whose maximal faces satisfy the properties of the lemma.

Choose a basis $q_{n,i}$, $i = 0, \ldots, d_n^*(T)$ for the affine space spanned by $B_n^*(T)$ such that $q_{n,i}$ belongs to $B_n^*(T) \setminus A_n(T)$ for $i \leq d_n^*(T) - \tilde{d}_n^*(T) - 1$ and it belongs to $A_n(T)$ otherwise. Choose vectors $q_{n,i}$ for $i = d_n^*(T) + 1, \ldots, d_n^*(T) + d_n^0(T_n) - \tilde{d}_n^*(T)$ such that the vectors $q_{n,i}$ for $i \geq d_n^*(T) - \tilde{d}_n^*(T)$ span the affine space generated by $P_n(T_n^0)$. Let $\tilde{B}_n(T)$ be the set of pairs $(\lambda, \mu) \in \mathbb{R}^{d_n^*(T)} \times \mathbb{R}^{d_n^0(T_n)}$ such that $\lambda - \bar{\mu} = 1$, $\lambda \geq 1$, where $\hat{\lambda} = \sum_i \lambda_{n,i}$ and $\bar{\mu} = \sum_i \mu_{n,i}$. For each $(\lambda, \mu) \in \tilde{B}_n(T)$, let $q_n(\lambda) = \sum_i \lambda_{n,i} q_{n,i}$, $r_n(\mu) = \sum_i \mu_{n,i} r_{n,i}$, let $h(\lambda, \mu) = q_n(\lambda) - r_n(\lambda)$. Let $\tilde{B}_n(T)$ be the set of $(\lambda, \mu)$ such that: (i) $r_n(\mu)$ and $h_n(\lambda, \mu)$ belong to the convex cone spanned by $P_n(T_n^0)$ and $P_n^0$ respectively—which immediately implies that $q_n(\lambda)$, as a linear combination of $r_n(\mu)$ and $h_n(\lambda, \mu)$ also belongs to the convex cone of $P_n^0$; (ii) the strategies in $T_m$ are equally good replies and better than strategies in $\tilde{T}_m^1$. $B_n(T)$ is now the image of $\tilde{B}_n(T)$ under $h$. $\tilde{B}_n(T)$ and $B_n(T)$ are now easily seen to be polyhedra.
For a maximal proper face $B'_n(T)$ of $B_n(T)$, exactly one of the following holds uniformly for all $p_n$ in the interior of $B'_n(T)$ and for each $(\lambda, \mu)$ in $h^{-1}(p_n)$: (i) $p_n$ is on the boundary; (ii) some strategy $s^*_n \in T'_m$ is a best reply against $q_n(\mu)$; (iii) $r(\mu)$ belongs to the boundary of $P^0_n(T^0_n)$; or (iv) $\sum \lambda_i = 1$.

For condition (i), genericity of payoffs implies that the boundary has to be a maximal proper face.

In the case of (ii), there exists a maximal proper face of $B^*_n(T)$ that is not contained in $\partial P^0_n$ and is such that for every $p_n \in B_n(T)$, the intersection of the convex hull of $p_n$ and $P_n(T^0_n)$ is contained in this face. Let $\tilde{R}^1_m$ be the set of points that are best replies against every point in this face. Then point (ii) holds.

As for (iii), by genericity again, if $r(\mu)$ belongs to a face $P_n(R^0_n)$ of $P_n(T^0_n)$, then this face is maximal. Suppose $A_n(T)$ is not contained in $P_n(R^0_n)$. Pick $q_n \in B^*_n(T)$ and $r_n \in P_n(R^0_n)$ such that $\lambda_n q_n + (1 - \lambda_n) r_n$ belongs to the relative interior of $B'_n(T)$. Pick an $r^*_n \in A_n(T) \setminus P_n(R^0_n)$; $q'_n \equiv (1 - \alpha_n) q_n + \alpha r^*_n = (1 - \alpha_n) (\lambda^{-1}(p_n - (1 - \lambda) r_n)) + \alpha r^*_n$ belongs to $B^*_n(T)$. Since $r^*_n \notin P_n(R^0_n)$, we could not have a boundary point in case (iii) if $A_n(T)$ is not contained in the face $P_n(R^0_n)$.

There remains to show that condition (iv) is impossible. Condition (iv) implies that $B'_n(T) = B^*_n(T)$. Let $\tilde{B}_n(T)$ be the set of points $p^0_n$ in $P^0_n$ such that: (i) the strategies in $T^0_m$ are equally good replies; (ii) the strategies in $T^1_m$ are equally good replies and yield at least as much as those in $\tilde{T}^1_m$. $\tilde{B}_n(T)$ is then a polyhedron which can be expressed as the union of two sets $B^+_n(T)$ and $B^-_n(T)$ where $B^+_n(T)$ (resp., $B^-_n(T)$) is the set of points in $\tilde{B}_n(T)$ where the strategies in $T^0_m$ do at least as well as (resp., no better than) strategies in $T^1_m$. Each of these two sets is a full-dimensional subset of $\tilde{B}_n(T)$ and $B^*_n(T)$ is the intersection of the two sets and has one dimension less.

We claim now that there exist points $r^+_n$ and $r^-_n$ in $P_n(\tilde{T}^0_n)$ that belong to the relative interiors of $B^+_n(T)$ and $B^-_n(T)$, respectively. Indeed, points in $A_n(T)$ belong to the interior $P_n(\tilde{T}^0_n)$ and to both $B^+_n(T)$ and $B^-_n(T)$. Therefore, the only way we cannot construct such points $r^+_n$ and $r^-_n$ is if $A_n(T)$ is full-dimensional, i.e. it is the whole face $P_n(\tilde{T}^0_n)$. But that is impossible: a strategy $r_n \in A_n(T)$ prescribes particular probabilities for terminal nodes not excluded by some $t^1_m \in T^1_m$ but excluded by strategies in $T^0_m$. Perturbing these probabilities clearly changes the payoffs to $t^1_m$. Thus, there exist points $r^+_n$ and $r^-_n$ with the properties specified.

Now pick a point $p^+_n$ in $B^+_n(T)$. For some $\lambda \in (0, 1)$, $\lambda p^+_n + (1 - \lambda) r^-_n$ belongs to $B^*_n(T)$. Thus, $B^+_n(T)$ belongs to $B_n(T)$. A similar argument proves that $B^-_n(T)$ belongs to $B_n(T)$ as well. Therefore, $\tilde{B}_n(T)$ belongs to $B_n(T)$, which proves that $B^*_n(T)$ could not be a maximal face of $B_n(T)$. \qed
Let \( C_n(T) \) be the closure of the set of \( q_n \) in the interior of \( P_n \) such that the strategies in \( T_m \) are all equally good replies and at least as good as strategies in \( \tilde{T}_m^1 \) and \( S_n^0 \setminus T_n^0 \). By Lemma C.5, the dimension of the face spanned by \( T_m \) is \( d(T_m^0) + d(\Psi_n^1) + 1 \). By genericity of payoffs, the dimension of \( C_n(T) \) is therefore \( d(P_n) - d(T_m^0) - d(\Psi_n^1) - 1 \).

Let \( C_n(T) \) be the closure of the set of \((p_n^1, \pi_n^1) \in P_n^1 \times \Pi_n^1 \) such that there exist \( p_n^0 \in P_n^0 \), \( p_n^2 \in P_n(T_n^1) \), and \( \mu \in \mathbb{R}_{++}^3 \), such that \( \sum_i \mu^i p_n^i \in C_n^*(T) \), \( \sum_i \mu^i = 1 \), and \( \tilde{\pi}_n^1(\sum_i \mu^i p_n^i) = \pi_n^1 \).

**Lemma C.7.** The set \( C_n(T) \) has dimension \( d_n^0 + d_n^1 + d_n(\Psi_n^1) - d(T_m^0) - d(\Psi_n^1) - d_n^*(T) \). On each maximal proper face \( C' \) of \( C_n(T) \), exactly one of the following holds for all \((p_n^1, \pi_n^1) \) in \( C' \). If \( q_n \in C_n^*(T) \) is of the form \( \sum_i \mu^i p_n^i \) for some \( p_n^0 \in P_n^0 \) and \( p_n^2 \in P_n(T_n^1) \), and \( \pi_n^1(\sum_i \mu^i p_n^i) = \pi_n^1 \), then:

1. \( p_n^1 \) belongs to a maximal proper face of \( P_n^1 \);
2. for \( i = 0 \) or \( i = 1 \), but not both, there exists \( s_m \in S_m^i \setminus T_m^i \), which actually belongs to \( \tilde{T}_m^i \) if \( i = 1 \), that is as good a reply as points in \( T_m \) against \( q_n \); moreover, in this case, if \( i = 0 \), \( T_m^0 \) and \( s_m \) span a face of \( P_m^0 \) of which \( P_m(T_m^0) \) is a maximal proper face; and if \( i = 1 \), \( \Psi_n^1 \) and \( \pi_n^1(s_m) \) span a face of \( \Pi_m^1 \) of which \( \Psi_n^1 \) is a maximal proper face.
3. there exists a maximal proper face \( \Psi' \) of \( \Psi_n \) such that \( \tilde{\pi}_n^1(p_n^2) \) belongs to \( \Psi' \).

**Proof.** For \( i = 0, 1 \), let \( p_n^{i,j} \), for \( j = 0, \ldots, d_n^i \), be a set of enabling strategies in \( P_n^i \) that span \( P_n^i \) and such that for \( j = 0, \ldots, d(T_m^0) \), \( p_n^{0,j} \) belongs to \( P_n(T_n^0) \). Let \( p_n^{2,j} \), \( j = 0, \ldots, d(\Psi_n^1) \) be a set of enabling strategies in \( T_n^1 \) such that the \( \pi_n^1(p_n^{2,j}) \)'s span \( \Psi_n^1 \). Observe that the strategies \( p_n^{0,j} \), \( j \leq d(T_n) \), and the \( p_n^{2,j} \)'s span the affine space that is generated by \( P_n(T_n) \); therefore, every \( p_n^i \) can be expressed as an affine combination of these strategies. In particular, if \( q_n \in C_n^*(T) \), then \( q_n \) can be expressed as an affine combination of all the \( p_n^{i,j} \)'s.

For \( i = 0, 2 \), let \( D^i \) be the convex cone spanned by the vectors \( p_n^{i,j} \). Let \( \tilde{P} \) be the convex cone spanned by \( P \). Define \( \xi : D^0 \times P^1 \times D^2 \to \mathbb{R}^z \) by \( \xi(r_n^0, p_n^1, r_n^2) = r_n^0 + p_n^1 + r_n^2 \). The inverse image of each interior point of \( \tilde{P} \) is \( d_n^0 \) \( \tilde{d}_n^1 = d(\Psi_n^1) - 1 \). Let \( \tilde{C}(T) \) be the convex cone spanned by \( C_n^*(T) \). Then \( \tilde{D}(T) = \xi^{-1}(\tilde{C}(T)) \) has dimension \( d_n^0 + d_n^1 + d(\Psi_n^1) - d(T_m^0) - d(\Psi_n^1) + 1 \). Let \( \tilde{C}(T) \) be the projection of \( \tilde{D}(T) \) to \( P^1 \times D^2 \). Observe that for each \( (r_n^0, p_n^1, r_n^2) \in \tilde{D}(T) \), each \( p_n^0 \in B_n^1(T) \), and each \( \nu > 0 \), \( (r_n^0 + \nu p_n^1, p_n^1, r_n^2) \in \tilde{D}(T) \). Therefore, the dimension of \( \tilde{C}(T) \) is \( d_n^0 + d_n^1 + d_n(\Psi_n^1) - d_n(T_m^0) - d_n(\Psi_n^1) - d_n^*(T) \). The function \( \tilde{\pi}_n^1 \) extends to \( \tilde{P} \setminus \{ 0 \} \), and \( (p_n^1, \pi_n^1) \) belongs to \( C_n(T) \) iff it belongs to the closure of the set of \( (p_n^1, \tilde{\pi}_n^1(p_n^1 + r_n^2)) \) such that \( (p_n^1, r_n^2) \) belongs to \( \tilde{C}(T) \). Therefore, \( C(T) \) is a polyhedron whose dimension is as stated. The rest of the properties can be proved in a manner analogous to the previous result and is therefore omitted.
For each $T \in \mathcal{T}$, let $\mathcal{Q}_n(T) = A_n(T) \times B_n(T) \times C_n(T)$ for each $n$ and let $\mathcal{Q}(T) = \mathcal{Q}_1(T) \times \mathcal{Q}_2(T)$.

Against each $(q_n^*, q_n^0, q_n^1)$ in the interior of $A_n(T) \times B_n^*(T) \times C_n^*(T)$, the strategies in $T_m$ are lexicographic best replies. Therefore, it is true for points in the closure as well, which we record below.

**Lemma C.8.** The strategies in $T_m$ are lexicographic best replies against every point $(q_n^*, q_n^0, q_n^1) \in A_n(T) \times B_n^*(T) \times C_n^*(T)$.

**Lemma C.9.** $(q^*, p^0, p^1, \pi^1)$ belongs to $\mathcal{Q}$ iff it belongs to $\mathcal{Q}(T)$ for some $T \in \mathcal{T}$.

**Proof.** Suppose for each $n$, $q_n^* \in A_n(T), p_n^0 \in B_n(T), (p_n^1, \pi_n^1) \in C_n(T)$ for some $T$. Choose $r_n^0 \in P_n(T_n^0)$ and $\lambda_n^0$ such that $q_n^* = (1 - \lambda_n^0)p_n^0 + \lambda_n^0 r_n^0 \in B_n^*(T)$. Also, fix $p_n^0, r_n^2 \in P_n(T_n^1)$, $\mu_n^0, \mu_n^1, \mu_n^2$ such that $q_n^1 = \mu_n^0 p_n^0 + \mu_n^1 p_n^1 + \mu_n^2 r_n^2 \in B_n^*(T)$. By the preceding lemma, the strategies in $T_m$ are lexicographic best replies against $(q_n^*, q_n^0, q_n^1)$, which proves that $(q^*, (p^0, p^1), \pi^1)$ belongs to $\mathcal{Q}$.

To prove the other way around, suppose $(q^*, (p^0, p^1), \pi^1)$ belongs to $\mathcal{Q}$. Let $q_n^0 = (1 - \lambda_n^0)p_n^0 + \lambda_n^0 r_n^0$ and let $q_n^1 = \mu_n^0 p_n^0 + \mu_n^1 p_n^1 + \mu_n^2 r_n^2$ where $(1 - \lambda_n^0)q_n^* + \lambda_n^0 r_n^0$ is at least as good a reply against $(q^*, q^0, q^1)$ than strategies in $S_n^0$; and $r_n^2$, if $\mu_n^2 > 0$, is a best reply against $q_n^*$ and a weakly better reply against $(q^*, q^0, q^1)$ than all the strategies in $P_n^1$. Let $Q_n^0$ be the face of $P_n^0$ that contains $(1 - \lambda_n^0)q_n^* + \lambda_n^0 r_n^0$. Let $Q_n^1$ be the face of $P_n^1$ that contains $r_n^2$ in its interior if $\mu_n^2 > 0$. Let $T_n^0$ be the set of strategies $t_n$ in $S_n^0$ such that if $t_n$ enables $h_n \in H_n^0$, then it coincides with the choices dictated by some vertex of $Q_n^0$ or $Q_n^1$ that enables $h_n$. If $\mu_n^2 > 0$, let $\Psi_n^1$ be the face of $\Pi_n^1$ that contains $\pi_n^1(r_n^2)$ in its interior; otherwise let $\Psi_n^1$ be the empty set.

We can now assume without loss of generality that the strategies in $Q_n^0$ and $Q_n^1$ are equally good replies against $(q_m^*, q_m^0, q_m^1)$ and, hence, best replies. Indeed, if for $i$ either 0 or 1, if the strategies in $Q_n^1$ do not yield the same payoff against $q_m^i$ as those in $Q_n^0$, modify $q_m^i$ as follows: if the strategies in $Q_n^0$ do better than strategies in $Q_n^1$, then pick a point $r_m^0$ in the face containing $q_m^*$ in its interior, such that the strategies in $Q_n^1$ do strictly better than the strategies in $Q_n^0$ and at least as well as the other strategies in $S_n^1$; if the strategies in $Q_n^0$ do worse pick the point $r_m^1$ to be one where the strategies in $Q_n^0$ do strictly better than the strategies in $Q_n^1$, which in turn are at least as good as the strategies in $S_n^1$. There exists a unique $\nu_m^i \in [0, 1]$ such that the strategies in $Q_n^0$ and $Q_n^1$ are now equally good replies against $(1 - \nu_m^i)q_m^i + \nu_m^i r_m^i$. Thus, our assumption is without loss of generality.

Since the strategies in $Q_n^0$ and $Q_n^1$ are best replies. There remains to show that every strategy in $S_n(T_n^0; \Psi_n^1)$ is a best reply against $(q_m^*, q_m^0, q_m^1)$. Fix $s_n \in S_n(T_n^0; \Psi_n^1)$. To show
that it is a best reply it is sufficient to show that for an information set \( h_n \) that is enabled by \( s_n \) is enabled by some vertex of either \( Q^0_n \) or \( Q^1_n \) and that this vertex agrees with \( s_n \)’s choice \( a_n \) there. Suppose that this \( h_n \) is in \( H^*_n \) and \( a_n \in A^*_n \), or \( h_n \) belongs to or follows some information set in \( H^0_n \), then obviously some strategy in \( Q^0_n \) or \( Q^1_n \) enables \( h_n \) and chooses \( a_n \), by the definition of \( T^0_n \). If \( h_n \in H^*_n \) and \( a_n \notin A^*_n \) or \( h_n \) follows some information set in \( H^*_n \) by the choice of a non-equilibrium action, then some strategy in \( Q^1_n \) enables it and chooses this action, since otherwise \( s_n \) enables a terminal node that is excluded by all strategies in \( Q^1_n \), contradicting the assumption that \( \pi^1_n(s_n) \in \Psi^1_n \). Thus \( s_n \) is a best reply and \((q^*,q^0,q^1)\) belongs to \( Q(T) \).  

\[ \text{Lemma C.10. } Q \text{ is a pseudomanifold of dimension } \hat{d} \equiv d(P). \]

\[ \text{Proof. } \]Each \( Q(T) \) is a polyhedron of dimension \( \hat{d} \). By the previous lemma \( Q \) equals \( \cup_{T \in T} Q(T) \). To show that \( Q \) is a pseudomanifold, we establish three facts for each \((q^*,(p^0,p^1),\pi^1)\) that belongs to some \( Q(T) \): (1) if \((q^*,(p^0,p^1),\pi^1)\) belongs to the interior of \( Q(T) \), then it does not belong to the interior of \( Q(R) \) for \( R \neq T \); (2) if \((q^*,(p^0,p^1),\pi^1)\) is a generic point in a maximal proper face \( Q' \) of \( Q(T) \), then it does not belong to \( Q(R) \) for any \( R \neq T \) if \((p^0,p^1) \in \partial \mathbb{P} \), and it belongs to the boundary of \( Q(R) \) for exactly one other \( R \neq T \) if \((p^0,p^1) \notin \partial \mathbb{P} \); (3) given \( T,R \in T \), there exists a finite chain \( T = T(0), \ldots, T(k) = R \) such that for each \( 0 \leq j \leq k-1 \), \( Q(T(j)) \cap Q(T(j+1)) \) is a subset of a maximal proper face of each and has a nonempty interior in this face.

Fix \( T = (T^0,\Psi^1) \) and \( x = (q^*,(p^0,p^1),\pi^1) \in Q(T) \). For each \( n \), choose \( q^0_n \equiv (1-\lambda^0_n)p^0_n + \lambda^0_n r^0_n \in B^*_n(T) \) and \( q^1_n \equiv \mu^0_n p^0_n + \mu^1_n p^1_n + \mu^2_n r^1_n \in C^*_n(T) \).

We start with (1). Suppose now that \( x \) belongs to the interior of \( Q(R) \) for some \( R \). We will show that \( R = T \). Since \( x \) belongs to the interior of \( Q(T) \), we can assume that every strategy in \( S^0_m \setminus T^0_m \) is inferior to \( q^1_m \). Let \( s_m \) be a strategy in \( S^0_m \setminus T^0_m \). Since \( s_m \) is an inferior reply against \( q^1_m \) compared to the strategies in \( T^0_m \), there exists an information set \( h_n \) enabled by \( s_n \) where the action chosen by \( s_n \) is different from the action chosen by every \( t_n \in T^0_n \) that enables \( h_n \). Clearly \( h_n \notin H^*_n \); therefore, it belongs to \( H^0_n \). But the posterior belief over the terminal nodes following \( h_n \) computed from \( q^1_m \) can be computed from \( \pi^1_m \). This implies that for any \( q^1_m \) such that \( \pi^1_m(q^1_m) = \pi^1_m \), \( s_n \) is an inferior strategy to some \( t_n \in T^0_n \). This implies that \((p^1_m,\pi^1_m)\) cannot belong to \( C_m(R) \) unless \( R^0_m \subsetneq T^0_m \). Moreover, if \( R^0_m \subsetneq T^0_m \), then it cannot belong to the interior of \( C_m(R) \), since strategies in \( T^0_m \setminus R^0_m \) are also optimal. Thus, if \( x \) belongs to the interior of \( Q(R) \), \( R^0_m = T^0_m \) for each \( m \). If \( \Psi^1_m \) is empty, this implies that \( R_m = T_m \). Suppose now that \( \Psi^1_m \) is nonempty. Since \( x \) is in the interior we can assume we can assume that \( \tilde{p}^1_m (p^1_m \neq \pi^1_m) \) and that \( \tilde{\pi}_n(r^1_m) \) is in the interior of \( \Psi^1_m \). Observe that
\( \tilde{\pi}_n^1(r_m^1) \) can be computed uniquely from \( p_n^1 \) and \( \pi_m^1 \) by taking the line segment from \( \tilde{\pi}_n^1(p_m^1) \) through \( \pi_m^1 \) and computing the boundary point of this line. This implies that \( x \) belongs to the interior of \( Q(R) \), then \( R_n = (T_m^0, \Psi_m^1) \). Thus \( R = T \).

We turn to point (2). Suppose now that \((p^0, p^1) \in \partial P \) and \( x \) is generic, which it is if it belongs to the relative interior of a maximal proper face of \( Q(T) \). Then the arguments of the previous paragraph apply to show that \( x \) does not belong to another \( Q(R) \). Indeed, it relied on strategies in strategies in \( S_m^0 \) being inferior to \( q_n^1 \) and \( q_n^1 \) (resp. \( q_n^1 \)) not belong to the boundary of \( B_n(T) \) (resp. \( C_n(T) \)). So, suppose that \((p^0, p^1) \) is in the interior in \( P \). Then for exactly one \( n \), just one of the following hold: (2a) \( q_n^1 \) belongs to the boundary of \( A_n^*(T) \); (2b) \( p_n^0 \) belongs to the boundary of \( B_n(T) \); (2c) \( (p_n^1, \pi_n^1) \) belongs to the boundary of \( C_n(T) \). We start with (2c). By the properties we proved for \( C_n(T) \) in Lemma C.7, and since \((p^0, p^1) \notin \partial P \), either property (ii) or property (iii) of that lemma holds. Under property (ii) \( x \) belongs to the boundary of \( Q(R) \) where \( R = ((R_m^0, \Phi_m^1), (T_n, \Psi_n^1)) \) is defined as follows. If the strategy \( r_m^i \) identified there belongs to \( S_m^0 \), then \( R_m^0 \) is the vertex set of the face spanned by \( r_m^i \) and \( T_m^0 \), while \( \Phi_m^1 = \Psi_m^1 \); if the strategy \( r_m^i \) belongs to \( S_n^1(T_m^0, \Psi_m^1) \), then \( R_m^0 = T_m^0 \) and \( \Phi_m^1 \) is a face of \( \Pi_m^1 \) that has \( \Psi_m^1 \) as a maximal face with \( \pi_m^1(r_m^i) \in \Phi_m^1 \setminus \Psi_m^1 \). Under property (iii) \( x \) belongs to the boundary of \( Q(R) \) where \( R = ((T_n^0, \Phi_n^1), (T_m^0, \Psi_m^1)) \) where \( \Phi_m^1 \) is the maximal proper face of \( \Psi_m^1 \) identified there.

Suppose \( x \) satisfies (2b). Then by the properties we proved for \( B_n(T) \) in Lemma C.6, either property (ii) or property (iii) of that lemma holds. Under property (ii) let \( \tilde{R}_m^1 \) be the set of strategies in \( \tilde{T}_m^1 \) that are now best replies against \( q_n^0 \). Let \( \tilde{\Phi}_m^1 \) be the smallest face of \( \Pi_m^1 \) that contains \( \Psi_m^1 \) and the vectors \( \pi_m^1(\tilde{r}_m^1) \) for \( \tilde{r}_m^1 \in \tilde{R}_m^1 \). Let \( \tilde{R}_m = (T_m^0, \tilde{\Phi}_m^1) \). Then the strategies in \( P_n(\tilde{R}_m) \) are equally good replies against \( q_n^0 \). Moreover by the genericity of \( x \), if one of these strategies is a best reply against a point in \( B_n^*(T) \) then all these points are best replies as well. For each face \( \Phi_m^1 \) of \( \tilde{\Phi}_m^1 \) that has \( \Psi_m^1 \) a maximal proper face, choose a strategy \( r_m(\Phi_m^1) \) that maps to a vertex of \( \Phi_m^1 \) that is not contained in \( \Psi_m^1 \). The set of points in \( B_n^*(T) \) against which the strategies in \( \tilde{R}_m \) are as good replies as \( T_m \) has dimension \( d_n^*(T) - 1 \). However, the set of points in \( C_n^*(T) \) where two or more of these strategies \( r_m(\Phi_m^1) \) are also best replies has dimension \( d(P_n) - d(T_m^0) - d(\Psi_m^1) - 3 \) or less. Therefore, for \( R \) and \( R' \) of the form \((T_n^0, \Psi_n^1), (T_m^0, \Phi_m^1))\) the set of \((p_n^1, \pi_n^1) \) that lies in the intersection \( C_n(T) \cap C_n(R) \cap C_n(R') \) is at most \( d_n^0 + d_n^1 + d_n(\Psi_n^1) - d_n^1(T) - 1 \) or less. This implies that generic \((p_n^1, \pi_n^1) \) in \( C_n(T) \) belongs to the boundary of exactly one of these sets.

Suppose finally \( x \) satisfies (2a). Let \( A_n^1 \) be a maximal proper face of \( A_n(T) \) that contains \( q_n^* \) in its interior. Let \( \tilde{R}_m^1 \) be the set of strategies \( \tilde{r}_m \) in \( S_m^1(T_m^0) \setminus (T_m^0 \cup \tilde{T}_m^1 \cup \tilde{T}_m^1) \) that are best replies against all the points in \( A_n^1 \) and such that there exists \( R = ((T_m^0, \Phi_m^1), (T_m^0, \Psi_m^1)) \in T \).
where $S_m(T^0_m; \Phi^1_m)$ contains $\tilde{r}_m$. Let $R_n^0$ be the set of subsets $R_n^0$ of $T_n^0$ such $P_n(\tilde{R}_n^0)$ is a maximal proper face of $P_n(T_n^0)$ and $P_n(\tilde{R}_n^0) \cap Q^*_n = A'_n$. Observe that for each $R_n^0 \in R_n^0$, $R = ((T^0_m, \Psi^1_m), (R_n^0, \Psi^1_n))$ belongs to $T$.

Let $\tilde{r}^1_m$ be the smallest face of $\Pi^1_n$ that contains $\Psi^1_m$ and the vectors $\pi^1_m(\tilde{r}^1_m)$ for $\tilde{r}^1_m \in \tilde{R}_n^0$. For each face $\Phi^1_m$ of $\tilde{r}^1_m$ that has $\Psi^1_m$ as a maximal proper face, choose a strategy $\tilde{r}^1_m(\Phi^1_m)$ that maps to a vertex of $\Phi^1_m$ that is not contained in $\Psi^1_m$. Take $R$ and $R'$ of the form $((T^0_n, \Psi^1_n), (T^0_m, \Phi^1_m))$. Since $A'_n$ is a face of $A_n(T)$, $d^*_n(R) = d^*_n(R') = d^*_n(T) - 1$. If $B^*_n(R) = B^*_n(R')$, then as in the previous paragraph, $(p^1_n, \pi^1_n) \in C_n(T)$ can belong to at most one of these sets, $C_n(R)$ and $C_n(R')$, if it is generic. If on the other hand $B^*_n(R) \neq B^*_n(R')$, then a generic point in $B^*_n(T)$ belongs to at most one of these sets. Hence a generic point in $B_n(T)$ belongs to at most one of these sets. Likewise, for each $\tilde{T}$ of the form $((\tilde{T}_n^0, \Psi^1_n), (T^0_m, \Psi^1_m))$, a point in $Q(T)$ belongs to at most one $\tilde{T}$. Finally whether a generic point in $Q(T)$ belongs to some $Q(R)$ or a $\tilde{T}$ depends on whether for that point there exists a corresponding point $B_n(T) \times C_n(T)$ where the strategies in $R$ are equally good replies or not.

If $p^0_n$ and $p^1_n$ are generic, then there exists a unique $\Phi^1_n$ such that $\tilde{r}^1_m(\Phi^1_m)$ is a lexicographic better reply to $(q^0_n, q^1_n)$ than the other strategies in this class. If $\tilde{r}^1_m(\Phi^1_m)$ is a lexicographic (weakly) better reply against $(q^0_n, q^1_n)$ or (resp. $q^*_n$ belongs to the interior of $P_n(\tilde{T}^0_n)$ and it is a worse reply against $(q^0_n, q^1_n)$) we choose a point $\hat{q}^0_n$ in the interior of $P_n(\tilde{T}^0_n)$ against which the strategy $\tilde{r}^1_m(\Phi^1_m)$ is at least as good as the other points in this class and inferior (resp. superior) to strategies in $T^0_m$. The strategies in $T^0_m$ and $S_m(T^0_m, \Phi^1_m)$ are now equally good replies against some average of $q^0_n$ and $\hat{q}^0_n$, as well as some average of $q^1_n$ and $\hat{q}^1_n$. The point $x$ now belongs to $((T^0_m, \Phi^1_m), (T^0_m, \Psi^1_m))$. If the strategy $\tilde{r}^1_m(\Phi^1_m)$ is a lexicographically inferior reply against $(q^0_n, q^1_n)$ and $q^*_n$ belongs to the boundary of $P_n(\tilde{T}^0_n)$, then $\tilde{T}^0_n$ is nonempty. If $p^0_n$ is generic, then there exists a unique set $\tilde{T}^0_n$ such that $q^0_n$ is a convex combination of a point in $P_n(\tilde{T}^0_n)$ and $p^0_n$. Then $x$ belongs to $((T^0_n, \Psi^1_n), (T^0_m, \Psi^1_m))$. This concludes the proof of point (2).

We turn now to point (3). Given $Q(T)$ and $Q(R)$ for $T = (T^0, \Psi^0) \neq R = (R^0, \Phi^1)$, we will first construct a sequence $T = T(1), \ldots, T(k) = \tilde{T}$ where $\tilde{T} = ((\tilde{T}_n^0, \emptyset), (\tilde{T}_m^0, \emptyset))$. And, likewise one from $R$ to $\tilde{R}$. Then we will show how to construct a sequence from $\tilde{T}$ to $\tilde{R}$.

In case $\tilde{T}^0_n \neq T^0_n$ for some $n$, let $T = T(0), \ldots, T(k)$ be a sequence where for each $j > 0$, $T(j) = ((T^0_n(j), \Psi^1_n), (T^0_m(j), \Psi^1_m))$ with $P_n(T^0_n(j)) \times P_m(T^0_m(j))$ being a maximal proper face of $P_n(T^0_n(j-1)) \times P_m(T^0_m(j-1))$, and $T_n(k) = \tilde{T}_n^0$ for each $n$. This sequence generates a sequence of polyhedra $Q(T) = Q(T(0), \ldots, Q(T(k))$ where for each $j > 0$, the intersection of $Q(T(j))$ with $Q(T(j-1))$ is contained in a maximal proper face of each and has a nonempty interior.
After this operation we have \( Q(\tilde{T}(k)) \), where \( \tilde{T}(k) = ((\tilde{T}_n^0, \Psi^1_n), (\tilde{T}_m^0, \Psi^1_m)) \). In case \( \Psi^1_n \) is nonempty for some \( n \), let \( \Psi^1 = (\Psi^1_n(0), \Psi^1_m(0)), \ldots, (\Psi^1_n(l), \Psi^1_m(l)) = (\emptyset, \emptyset) \) be a sequence such that for each \( 1 \leq j \leq l \), \( \Psi^1_m(j) \times \Psi^1_m(j) \) is a maximal proper face of \( \Psi^1_n(j-1) \times \Psi^1_m(j-1) \), and \( \Psi^1(l) = \emptyset \). This way we can connect \( Q(\tilde{T}(k)) \) with \( Q(\tilde{T}) \) where \( \tilde{T} = ((\tilde{T}_n^0, \emptyset), (\tilde{T}_m^0, \emptyset)) \).

Now we show how to connect \( \tilde{T} \) with \( \tilde{R} \) for two sets \( T, R \in \mathcal{T} \). Because \( Q^* \) is connected for each \( n \), there exists a sequence \( \tilde{T}_n^0 = (S_n^0(0), S_m^0(0)), \ldots, (S_n^0(l), S_m^0(l)) = \tilde{R}_n^0 \) where for each \( 1 \leq j \leq l \), either \( P_n(S_n^0(j)) \times P_n(S_m^0(j)) \) is either a maximal proper face of \( P_n(S_n^0(j+1)) \) or vice versa and for each \( n \), \( Q^* \) intersects the interior of the set \( P_n(S_n^0(j)) \). This generates a sequence \( Q^0, \ldots, Q^l \), where \( Q^l = ((S_n^0(j), \emptyset), (S_m^0(j), \emptyset)) \). Thus we have constructed a sequence of sets in \( \mathcal{T} \) that connect \( T \) and \( R \).

**Lemma C.11.** \( \psi : (Q, \partial Q) \to (\mathbb{P}, \partial \mathbb{P}) \) is essential iff \( Q^* \) is stable.

**Proof.** Let \( Y = [0, 1] \times P \). For each \( 0 < \varepsilon \leq 1 \), \( Y_\varepsilon = [0, \varepsilon] \times P \) and let \( \partial Y_\varepsilon \) be the boundary of \( Y_\varepsilon \). Each \((\varepsilon, p) \in Y\) defines a strategic game \( G(\varepsilon, p) \) where the strategy set is \( P \) but where the payoff from an enabling strategy profile \( q \) is the payoff in \( G \) from the profile \((1-\varepsilon)q + \varepsilon p\). If \( q \) is an equilibrium of \( G(\varepsilon, p) \), we say that \((1-\varepsilon)q + \varepsilon p\) is a perturbed equilibrium of \( G(\varepsilon, p) \).

Let \( E \) be the closure of the set of \((\varepsilon, p, q)\) such that \((\varepsilon, p) \in Y_1 \setminus \partial Y_1 \) and \( q \) is a perturbed equilibrium of \( G(\varepsilon, p) \). Let \( \theta \) be the projection map from \( E \). For each subset \( E \) of \( E \) and each \( 0 < \varepsilon, \), let \((E_\varepsilon, \partial E_\varepsilon) = \theta^{-1}(P_\varepsilon, \partial P_\varepsilon) \cap E \).

By Govindan and Wilson [9] there exists \( 0 < \varepsilon \leq 1 \), and a finite number of subsets \( E^1, \ldots, E^K \) of \( E \) such that for each \( 0 < \varepsilon \leq \varepsilon \): (i) \((E_\varepsilon^k, \partial E_\varepsilon^k) \) is a pseudomanifold (in fact a homology manifold) of dimension \( d(P) + 1 \) for each \( k \); (ii) \( E_\varepsilon^k \cap E_\varepsilon^j \subset \partial \phi^{-1}(\partial P_1) \) for \( k \neq j \); (iii) \( \cup_k E_\varepsilon^k = E_\varepsilon \).

Govindan and Wilson [6] show that there exists a neighborhood of \( \Sigma^* \) and an \( \varepsilon > 0 \) such that the set of \( \varepsilon \)-perfect equilibria in this neighborhood is connected. The image of this set of \( \varepsilon \)-perfect equilibria under \( \rho \) is therefore connected. Thus, there exists some \( k \) such that \( E_0^k = \{0\} \times Q^* \). \( Q^* \) is stable iff the projection \( \theta \) from \( E^k \) to \( Y_\varepsilon \) is essential. For simplicity in notation we refer to this \( E^k \) as simply \( E \).

It is now sufficient to prove that \( \psi \) is essential iff the projection \( \theta \) from \( E \) to \( E \) is essential. For each \( n \), \( \tilde{P}_n \equiv [0, 1] \times \mathbb{P}_n \) and define \( \chi_n : \tilde{P}_n \to P_n \) by \( \chi_n(\lambda_n, p_n^0, p_n^1) = (1 - \lambda_n)p_n^0 + \lambda_n p_n^1 \). Then we have that \(((\tilde{P}_n, \partial \tilde{P}_n), (P_n, \partial P_n), \chi_n)\) is a ball-bundle. Let \( \chi \) be the product map \( \chi_1 \times \chi_2 \). \(((\tilde{P}, \partial \tilde{P}), (P, \partial P), \chi)\) is a ball-bundle too. Let \( \tilde{Y}_\varepsilon = [0, \varepsilon] \times \tilde{P} \). \( \chi \) then induces a map \( h_Y : \tilde{Y}_1 \to Y_1 \) by \( h(\varepsilon, \lambda, p^0, p^1) = (\varepsilon, \chi(\lambda, p^0, p^1)) \). \(((\tilde{Y}_\varepsilon, \partial \tilde{Y}_\varepsilon), (Y_\varepsilon, \partial Y_\varepsilon), h_Y)\) is a ball-bundle now. Let \( \tilde{E} \) be the set of all \(((\varepsilon, \lambda, p^0, p^1), q) \in \tilde{Y}_1 \times P \) such that \( h_E(\varepsilon, \lambda, p^0, p^1, q) \equiv (\varepsilon, f(\lambda, p^0, p^1), q) \in E \). \(((\tilde{E}_\varepsilon, \partial \tilde{E}_\varepsilon), (E_\varepsilon, \partial E_\varepsilon), h_E)\) is a ball-bundle as well. Moreover, letting
be the projection from $E$ to $P$, we have that $h_Y \circ \tilde{\theta} = \theta \circ h_E$. Therefore, by the Thom Isomorphism Theorem (cf. Appendix IV.3, [19]) $\theta$ is essential iff $\tilde{\theta}$ is.

Let $\tilde{E}$ be the closure of the set of $(\varepsilon, \lambda, p^0, p^1, q, \pi^1(q))$ such that $(\varepsilon, \lambda, p^0, p^1, q) \in \tilde{E}$ and $\lambda \neq 0$. By the strong excision property, the natural projection $\hat{\phi}$ from $\tilde{E}$ to $\hat{E}$ induces an isomorphism of their cohomology groups. Let $\hat{\theta}$ be the projection from $\hat{E}$ to $\hat{Y}_\varepsilon$. Then $\hat{\theta} = \tilde{\theta} \circ \hat{\phi}$. Therefore, $\tilde{\theta}$ is essential iff $\hat{\theta}$ is.

Let $\eta : \hat{E} \to \mathbb{R}^3_+$ be the projection map $\eta(\varepsilon, \lambda, (p^0, p^1), q, \pi^1) = (\varepsilon, \lambda)$. Let $D = \eta(\hat{E})$. By the generic local triviality theorem, there exists a partition of $D$ into a finite number of connected subsets $D^0_1, \ldots, D^0_t$, and for each $D_i$ a semi-algebraic fibre pair $(F_i, \partial F_i)$, a homeomorphism $h_i : D^0_i \times (F_i, \partial F_i) \to (\eta^{-1}(D^0_i), \eta^{-1}(D^0_i) \cap \partial \hat{E})$ such that $\eta \circ h_i$ is the projection from $D^0_i \times F_i$ to $D^0_i$. There now exists an $i$, say 1, such that the closure $D_1$ of a subset of $D^0_1$ is homeomorphic to a 3-simplex and contains $[0, \varepsilon] \times \{0\}$ for some $\varepsilon < \varepsilon$.

Let $(\hat{Y}(D_1), \partial \hat{Y}(D_1)) \equiv (D_1, \partial D_1) \times (P, \partial P)$ and let $\hat{E}(D^1)$ be the closure of the inverse image of $D^1_1 \times F_1$ under $h_1$. Let $\hat{\theta}(D_1) : (\hat{E}(D_1), \partial \hat{E}(D_1)) \to (\hat{Y}(D_1), \partial \hat{Y}(D_1))$ be the projection. We claim that $(\hat{E}(D_1), \partial \hat{E}(D_1))$ is a pseudomanifold. Since $\hat{E}$ is a pseudomanifold, this claim is proved if we can show that $\hat{E}(D_1) \setminus \partial \hat{E}(D_1)$ is path-connected. There exist $0 < \varepsilon < \varepsilon$ and integers $n > 1$ for each $n$ such that $(\varepsilon, \lambda) \in D_1$ if $0 < \varepsilon < \varepsilon$ and $0 \leq \lambda_n \leq \varepsilon^{n-1}$. Now given two points $x(0)$ and $x(1)$ in $\hat{E}_\varepsilon(D_1) \setminus \partial \hat{E}_\varepsilon(D_1)$, connect them by a curve $x(t) = (\varepsilon(t), \lambda(t), p^0(t), p^1(t), q(t), \pi^1(q(t)))$ in $\hat{E}_\varepsilon \setminus \partial \hat{E}_\varepsilon$ as $t$ goes from 0 to 1. For each $t$, express $q(t)$ as $\varepsilon(t)(\lambda(t)p^0(t) + (1 - \lambda(t))p^1(t)) + (1 - \varepsilon(t)r(t)$ where $r(t)$ is a best reply to $q(t)$. For each $t$ we can now find a point $q^*(t) \in Q^*$ such that $r(t)$ is a best reply against $q^*(t)$. Choose a positive $\varepsilon$ such that $\varepsilon + \varepsilon^n < \varepsilon$. For each $n$, modify $x_n(t)$ to the vector

$$\tilde{x}_n(t) = (\varepsilon + \varepsilon^n \varepsilon(t), \tilde{\lambda}(t), \tilde{p}^0(t), p^1(t), \tilde{q}(t))$$

where $\tilde{\lambda}_n(t) = \varepsilon^n \varepsilon(t)\lambda_n(t)(\varepsilon + \varepsilon^n)^{-1}$,

$$\tilde{p}^0_n(t) = (\varepsilon + \varepsilon^n \varepsilon(t)(1 - \lambda(t)))^{-1}(\varepsilon q^*(t) + \varepsilon^n \varepsilon(t)(1 - \lambda(t))p^1(t)),$$

$$\tilde{q}_n(t) = (1 - \varepsilon^n)q^*(t) + \varepsilon^n(1 - \varepsilon(t)r(t) + \varepsilon^n \varepsilon(t)(\lambda(t)p^0(t) + (1 - \lambda(t))p^1(t)).$$

Then $\tilde{x}(t)$ belongs to $\hat{E}(D^1) \setminus \partial \hat{E}(D^1)$ for all $t$. Moreover, for $t = 0, 1$, $x(t)$ and $\tilde{x}(t)$ can now be connected by a path $\tilde{x}; s$ defined as follows. For $s \in [0, 1]$, let $k_n(s) = \min(1, 2s)r_n$ and then:

$$\tilde{x}_n(t; s) = ((2s - 1)^{+} \varepsilon + \varepsilon^{k_n(s)} \varepsilon(t), \tilde{\lambda}(t; s), \tilde{p}^0(t), \tilde{p}^1(t; s), \tilde{q}(t))$$
Then for all small $\varepsilon$

$$p_n^1(t; s) = ((2s - 1)^\varepsilon + (\varepsilon k_n(s) \varepsilon(t) (1 - \lambda_n(t)))^{-1} (s \varepsilon q_n^*(t) + \varepsilon r_n \varepsilon(t) (1 - \lambda_n(t)) p_n^1(t)),$$

$$q_n(t; s) = (1 - \varepsilon k_n(s)) q_n^*(t) + \varepsilon k_n(s) (1 - \varepsilon(t)) r(t) + \varepsilon k_n(s) \varepsilon(t) (\lambda_n(t)p^0(t) + (1 - \lambda_n(t))p_n^1(t))$$

Thus, $\hat{E}(D_1) \setminus \partial E(D_1)$ is connected and hence a pseudomanifold of dimension $d(\mathbb{P}) + 3$. $\hat{Y}(D_1)$ is a full-dimensional subset of $Y^\varepsilon$. Therefore $\hat{\theta}$ is essential iff $\hat{\theta}(D_1)$ is essential.

Since $(\hat{E}(D^1), \partial \hat{E}(D^1))$ is a pseudomanifold, $(F_1, \partial F_1)$ is a pseudomanifold of dimension $d(\mathbb{P})$. Moreover, for each $(\varepsilon, \lambda) \in D^0_1$, letting $(\hat{E}_{\varepsilon, \lambda}, \partial \hat{E}_{\varepsilon, \lambda}) \equiv h_{\varepsilon, \lambda}^{-1}(\{\varepsilon, \lambda\} \times (F^1, \partial F^1))$ we have that $\hat{\theta}$ is essential iff $\hat{\theta}_{\varepsilon, \lambda}$, the projection map $(\hat{E}_{\varepsilon, \lambda}, \partial \hat{E}_{\varepsilon, \lambda}) \rightarrow (\mathbb{P}, \partial \mathbb{P})$, is essential.

Let $L$ be the closure of the set of $(\varepsilon, \lambda) \in D_1$ such that $\lambda_n = \varepsilon^r$ for some $r \geq r_n$ for each $n$. Let $\hat{E}(L)$ be the closure of the inverse image of $L \setminus \{(0, 0)\}$ under $h^1$. Let $\partial E(L)$ be the inverse image of $\partial \hat{E}$ under the projection map $\theta(L)$ from $\hat{E}(L)$. For each $(\varepsilon, \lambda) \in L$, $(E_{\varepsilon, \lambda}, \mathcal{E}_{\varepsilon, \lambda})$ is $(\theta(L))^{-1}(\{\varepsilon, \lambda\} \times (\mathbb{P}, \partial \mathbb{P}))$. Our next objective is to show that $(Q, \partial Q)$ equals the set of $(q^*, (p^0, p^1), \pi^1)$ such that $(0, 0, (p^0, p^1), q^*, \pi^1)$ belongs to $\hat{E}_{0, 0}^1(L)$. Given $(0, 0, (p^0, p^1), q^*, \pi^1)$ in $\hat{E}^1$ there exists a sequence $(\varepsilon(k), \lambda(k), (p^0, p^1)(k), q(k), \pi^1(k))$ in $\hat{E}^1$. For each $n$ and $k$, we can express $q_n(k)$ as $(1 - \varepsilon(k))((1 - \mu_n^0(q_n^0(k) + \mu_n^1 r_n^1(k)) + \varepsilon(k)(\lambda(k)p_n^0(k) + (1 - \lambda(k)p_n^1(k)))$, where $\mu_n^1$ and $\lambda(k)$ converge to zero, $q_n^0(k)$ belongs to $P_n^0$ and converges to $q_n^*$, and $r_n^1(k)$ belongs to $P_n^1$. Also $q_n^0(k)$ and $r_n^1(k)$ if $\mu_n^2$ vanish are best replies to $q(k)$ for all $k$. By going to a subsequence, the faces to which $q_n^0(k)$ and $r_n^1(k)$ belong are constant for all $k$. The sequence generates for each $n$ an LPS $\Lambda_n = (q_n^0, \ldots, q_n^1)$ where $\bar{q}_n^0 = q_n^*$. As in the proof of the main theorem, there exists a level $t_n^1$ for each $i = 0, 1$ that is expressible as a convex combination of $t_n^0$ and another strategy. Since $\lambda(k)$ converges to zero, $l_n^0 < l_n^1$. As in the proof of the main theorem, we can prove that $\bar{q}_n^0$ belongs to $Q^*$ for each $l < l_n^0$ and if we express $\bar{q}_n^0 = \nu_n^0 p_n^0 + (1 - \nu_n^0) r_n^0$, then the strategies $\bar{q}_n^1$ for $l < l_n^0$ and $r_n^0$ if $\nu_0 < 1$ are lexicographic best replies against $\Lambda_m$. Likewise, if we express $q_n^{l_2}$ as $\bar{q}_n^{l_2} + \nu_n^{l_2} r_n^1 + \bar{q}_n^{l_2} r_n^2$, then the strategy $r_n^2$ if $\nu_2 > 0$ is a lexicographic best reply against $\Lambda_m$. As in the proof of the main theorem (Section 5.6) we can now write down an LPS $(q^*, \bar{q}^0(\varepsilon), \bar{q}^1(\varepsilon))$ to show that $(q^*, (p^0, p^1), \pi^1)$ belongs to $Q$.

Given $(q^*, (p^0, p^1), \pi^1)$ in $Q$, it belongs to $Q(T)$ for some $T$. There exist $q_n^0 = (1 - \lambda_n^0)p_n^0 + \lambda_n^0 r_n^0$ and $q_n^1 = m_n^0 p_n^0 + m_n^1 p_n^1 + m_n^2 r_n^2$ such that the strategies in $T_m$ are best replies against the LPS $(q^*, q_n^0, q_n^1)$. For all small $\varepsilon$, choose $\alpha(\varepsilon)$ such that $\varepsilon^r \mu_1 = (\varepsilon^r \mu_1 + \varepsilon^r \mu_0 + \alpha(\varepsilon)(1 - \lambda_n^0)p_n^0)^r$.

Then for all small $\varepsilon$, $(\varepsilon, \lambda(\varepsilon), (p^0, p^1)(\varepsilon), q(\varepsilon))$ belongs to $\hat{E}^1$ where:

$$\lambda(\varepsilon) = \frac{\varepsilon^r \mu_1}{\varepsilon^r \mu_1 + \varepsilon^r \mu_0 + \alpha(\varepsilon)(1 - \lambda_n^0)}.$$
and \((p^0, p^1)(\varepsilon) = (p^0(\varepsilon), p^1)\) with for each \(n\),

\[p^0_n(\varepsilon) = \frac{\varepsilon^r \mu_0 \tilde{p}^0_0 + \alpha(\varepsilon)(1 - \lambda^0_n)p^0_n}{\varepsilon^r \mu^0 + \alpha(\varepsilon)}\]

and \(\tilde{q}(\varepsilon) = (1 - \alpha(\varepsilon) - \varepsilon^r)q^* + \alpha(\varepsilon)q^0 + \varepsilon^r q^1\).

For each \((\varepsilon, \lambda)\), \((E_{\varepsilon, \lambda}, \partial E_{\varepsilon, \lambda})\) is a pseudomanifold. Indeed for \((\varepsilon, \lambda) \neq (0, 0)\) this follows from the fact that this pair is homeomorphic to \((F_1, \partial F_1)\), which is a pseudomanifold; for \((0, 0)\), this follows from the fact that \((E_{0,0}, \partial E_{0,0})\) is homeomorphic to \((Q, \partial Q)\). The inclusion map \((E_{\varepsilon, \lambda}, \partial E_{\varepsilon, \lambda})\) induces an isomorphism of the \(d(\mathbb{P})\)-th cohomology groups. Thus \(\hat{\theta}_{0,0}\) is essential iff \(\hat{\theta}_{\varepsilon, \lambda}\) is essential for some (and then all) \((\varepsilon, \lambda) \in L\). The projection map \(\hat{\theta}_{0,0}\) is just the map \(\psi\). Hence \(Q^*\) is stable iff \(\psi\) is essential.

This concludes the proof of the Theorem.

\[\blacksquare\]

REFERENCES


DEPARTMENT OF ECONOMICS, UNIVERSITY OF IOWA, IOWA CITY IA 52242, USA.
E-mail address: srihari-govindan@uiowa.edu

STANFORD BUSINESS SCHOOL, STANFORD, CA 94305-5015, USA.
E-mail address: rwilson@stanford.edu