RESEARCH PAPER NO. 2021R

Axiomatic Equilibrium Selection for Generic Two-Player Games

Srihari Govindan

Robert Wilson

May 2009, Revised May 2011

This work was partially funded by a National Science Foundation grant.

ABSTRACT

We impose three conditions on refinements of the Nash equilibria of finite games with perfect recall that select closed connected subsets, called solutions.

A. Each equilibrium in a solution uses undominated strategies;

B. Each solution contains a quasi-perfect equilibrium;

C. The solutions of a game map to the solutions of an embedded game,

where a game is embedded if each player's feasible strategies and payoffs are preserved by a multilinear map. We prove for games with two players and generic payoffs that these conditions characterize each solution as an essential component of equilibria in undominated strategies, and thus a stable set as defined by Mertens (1989).

0

Contents

AXIOMATIC EQUILIBRIUM SELECTION FOR GENERIC TWO-PLAYER GAMES

SRIHARI GOVINDAN AND ROBERT WILSON

Abstract. We impose three conditions on refinements of the Nash equilibria of finite games with perfect recall that select closed connected subsets, called solutions.

A. Each equilibrium in a solution uses undominated strategies;

B. Each solution contains a quasi-perfect equilibrium;

C. The solutions of a game map to the solutions of an embedded game,

where a game is embedded if each player's feasible strategies and payoffs are preserved by a multilinear map. We prove for games with two players and generic payoffs that these conditions characterize each solution as an essential component of equilibria in undominated strategies, and thus a stable set as defined by Mertens (1989).

1. Introduction

The literature on refinements applies stronger criteria than Nash's [38, 39] definition of equilibrium in a finite game. Among the goals are to exclude use of weakly dominated strategies, and to exclude outcomes inconsistent with backward induction. In the large literature reviewed in [14, 22, 23, 51], most refinements either enforce desirable properties directly or inherit them from equilibria of perturbed games.[1] Kohlberg and Mertens [24] argue that ideally a refinement should derive from axioms adapted from decision theory. They also specify properties that axioms should imply. Mertens [32, 33, 34, 36] shows subsequently that his revised definition of stability has these properties.

We show here that just three properties imply that a refinement selects stable sets for any game in the class consisting of two-player games in extensive form with perfect recall and generic payoffs. These three properties therefore imply the others for such a game—several are listed below in §1.1. We discuss the restriction to this class of games in Section 6.

We assume that for each game a refinement selects closed connected subsets of its Nash equilibria, called *solutions*, having the following properties.[2]

Key words and phrases. game, equilibrium, refinement, axiom, admissibility, backward induction, small worlds, stability. *JEL subject classification:* C72. *Acknowledgement:* This work was funded in part by a grant from the National Science Foundation. *Date:* 22 May 2011.

[1]Examples include sequential rationality as represented by subgame perfection [47] and sequential equilibrium and its extensions [26, 18, 31, 43], perturbed payoffs [12, 24], and perturbed strategies (e.g. equilibria that are perfect [48], quasi-perfect [50], proper [37], or stable [24, 20, 32]).

[2]Solutions are assumed to be sets because Kohlberg and Mertens [24, pp. 1015, 1019, 1029] show that there need not exist a single equilibrium satisfying weaker properties than those invoked here. The technical requirement that a solution is connected excludes the trivial refinement that selects all equilibria. If only

A. Equilibria in solutions use only undominated strategies.
B. Each solution contains a quasi-perfect equilibrium.
C. A solution is immune to embedding the game within any larger game that preserves players' feasible strategies and payoffs.

These properties are assumed for all games, except that (B) presumes perfect recall, not just for those games in the class for which we prove that they imply stability. They are described further in Section 3. Here we call them 'axioms,' and defer to Section 6 a discussion of their potential role in a general axiomatic theory of refinements.

Axiom A adapts the basic axiom of decision theory called admissibility (Luce and Raiffa [29, Axiom 5, p. 287]). Its relevance in game theory is discussed by Kohlberg and Mertens [24, Section 2.7]. Mertens [36, esp. Section 2.1] considers stronger variants, but for two-player games each is equivalent to exclusion of dominated strategies.

Axiom B is a strong version of sequential rationality. A quasi-perfect equilibrium yields a sequential equilibrium for which a player's strategy following an information set is admissible in the continuation. When payoffs are generic, all sequential equilibria are quasi-perfect, and all equilibria in a solution have the same outcome; thus in this case the role of (B) is to ensure that a solution's outcome is consistent with sequential rationality. We do not use sequential equilibrium in (B) because our proof of the main theorem applies (C) to a larger game with nongeneric payoffs.

Axiom C requires that a solution is induced by a solution of any larger game in which the original players retain the same feasible strategies and payoffs. This invariance property is satisfied by Nash equilibria; hence (C) implements the principle that a refinement should inherit invariance properties of equilibria. The practical interpretation is that (C) excludes framing or presentation effects; otherwise, a wider context could influence which equilibria are selected by a refinement. Its technical role is to replace perturbations of the original game by embeddings in larger games. The gist of our main theorem is that, for the assumed class of games, stability against perturbations of strategies is equivalent to stability against embeddings in larger games. This conforms to Kohlberg and Mertens' insistence that a refinement be justified by principles of rationality in the specified game (rather than perturbed versions), provided one accepts that rationality is not influenced by contextual features of any larger game in which it might be embedded.

We establish that these three axioms imply that a solution is a stable set for a game in the assumed class. Our characterization is cast in terms of the 'enabling form' of a game in which two pure strategies of a player are considered equivalent if they exclude the same terminal nodes of the game tree, as allowed by Axiom C. The enabling form is defined in Section 4.3 and explained further in Appendix A. The proof shows that each solution is

a single (possibly unconnected) subset is selected then only the trivial refinement satisfies the conditions invoked by Norde, Potters, Reijnierse, and Vermeulen [40].

an essential component of admissible equilibria.[3] This is the defining property of a stable set of the enabling form.[4] The proof applies only to two-player games with generic payoffs because it is based on a version of Mertens' definition of a stable set, stated in Theorem 5.2, that relies on a lexicographic representation of optimal responses to deviations that takes a simple form for games in this class.

1.1. **Implications of the Theorem.** When payoffs are generic, all equilibria in a connected set yield the same probability distribution over outcomes [26, 24, 9]. Therefore, for economic models and econometric studies that use games in the class considered here, the axioms' chief practical implication is that a predicted outcome distribution should result from equilibria in an essential component of the game's admissible equilibria, and in particular, from the sequential equilibria it necessarily contains. The secondary implication that a solution includes all equilibria in the component is germane only for predicting behavior after deviations from equilibrium play, but it addresses the issue of whether sequential rationality is a relevant decision-theoretic criterion after deviations (Reny [43, 44, 45]). Because we show that each equilibrium in a solution is induced by a quasi-perfect equilibrium in a solution of some larger game in which it is embedded, every equilibrium in a solution is sequentially rational when viewed in a larger game. See [13, §2.3] and Section 3.4 for examples.

Our conclusion that a solution of a game in the assumed class must be a stable set implies the following corollaries. Mertens [32] proves these properties of a stable set for any finite game.

1. *Admissibility and Perfection.* All equilibria in a stable set are perfect, hence admissible.
2. *Backward Induction and Forward Induction.* A stable set includes a proper equilibrium that induces a quasi-perfect 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] A component is a maximal closed connected set, and it is essential if the local projection map has the coincidence property described in footnote 4. For the usual normal form of the game, a solution can be a subset of admissible equilibria in a component for which the subset is equivalent to a component for the enabling form.

[4] See Theorem 5.2 for the characterization of a stable set used here. Mertens' general definition invokes homology theory, which is used here only in Appendix C, Lemma D.10 of Appendix D, and Appendix E. A simplified rendition is that a connected closed set of equilibria is stable if the local projection map, from a connected closed neighborhood in the graph of equilibria over the space of players' strategies perturbed toward mixed strategies, has a point of coincidence with every continuous map having the same domain and range; i.e. the projection has nonzero degree. Govindan and Mertens [8] establish an equivalent definition in terms of players' best-reply correspondences. Stable sets can differ from payoff-essential sets [12] for which it is players' payoffs that are perturbed; a payoff-essential set contains a stable set, but stable sets can be subsets of a payoff-inessential set [22, §13.5].

3. *Invariance, 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.
4. *Player Splitting.* Stable sets are not affected by splitting a player into agents such that no path through the game tree includes actions of two agents.

In spite of these and other desirable properties, Mertens' definition of stability via essentiality of the projection map (see footnote 4) has been a major impediment to justifying it as an economically relevant refinement. For instance, it implies more than the plausible requirement that that there exist nearby equilibria of nearby games obtained by perturbing players' strategies. However, our main theorem shows for the assumed class of games that it is equivalent to the conjunction of Axioms A, B, and C, each of which has economic significance.

1.2. **Synopsis.** Section 2 establishes notation for Section 3, which specifies the axioms and a precise definition of embedding a game in a larger game. Section 4 establishes more notation, defines quasi-perfection and enabling strategies, describes generic payoffs, and identifies the version of stability used here. Section 5.1 provides in Theorem 5.2 a simplified characterization of a stable set that is valid only for two-player games with generic payoffs. This characterization suffices here as the definition for readers new to the subject. Section 5 states and proves the main result, Theorem 5.1. 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 larger game with perfect recall that embeds the given game. Section 6 discusses the axioms and assumptions.

Appendix A illustrates enabling strategies. Appendix B proves the equivalent version of Axiom C stated as Proposition 3.2 in the text. It also proves Proposition 3.3 that Nash equilibria satisfy Axiom C. The last three invoke homology theory: Appendix C proves that stable sets satisfy Axiom C, Appendix D proves Theorem 5.2, and Appendix E establishes existence of a function used in proving Theorem 5.1.

2. Notation

This section provides sufficient notation for statements of the axioms in Section 3. Section 4 introduces additional notation for the theorems in Section 5 and Appendix D.

A typical finite game in extensive form is denoted Γ. 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 Σ_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 Γ assigns to each profile of players' pure strategies the profile of their expected payoffs; equivalently, it is the multilinear (i.e. linear in each player's strategy) function $G : \Sigma \to \mathbb{R}^N$ that to each profile of their mixed strategies assigns 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 $B = \prod_n B_n$ the set of profiles of players' behavioral strategies. Each mixed strategy σ_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 [27]).

As defined by Nash [38, 39], an equilibrium is a profile of players' mixed strategies such that each player's strategy is an optimal reply to other players' strategies. That is, if $\mathrm{BR}_n(\sigma) \equiv \arg\max_{\sigma'_n \in \Sigma_n} G_n(\sigma'_n, \sigma_{-n})$ is player n's best-reply correspondence then $\sigma \in \Sigma$ is an equilibrium iff $\sigma_n \in \mathrm{BR}_n(\sigma)$ 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 closed connected subsets of its Nash equilibria. Each assigned 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. 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 non-excluded 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 strategies. Because Axiom C implies invariance to redundant strategies, later we use equivalence classes of strategies.

3. The Axioms

3.1. **Undominated Strategies.** The first axiom requires simply that no player uses a weakly dominated strategy. 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 consistent with Axiom A we further insist on conditionally admissible continuations from information sets. Here we obtain these properties from quasi-perfect equilibrium. The original definition by van Damme [50] relies on consideration of perturbed strategies. The proof of our main theorem preserves continuity with previous literature by invoking this definition in Section 4.1.

An alternative definition uses its representation as a lexicographic equilibrium [2, 7]. In this decision-theoretic version, each player's behavior is described initially by a finite sequence of mixed strategies, interpreted as other players' alternative hypotheses about his actions. At a subsequent information set that reveals deviation from equilibrium play, hypotheses that fail to explain the deviation are discarded. Considering only a two-player game for simplicity, quasi-perfection requires that a player's continuation is optimal against the remaining subsequence of the other player. Thus, at each information set a player continues with the first strategy in his sequence that reaches that information set, and this continuation must be optimal in reply to the subsequence consisting of those hypotheses about the other player that do not exclude that information set. A lexicographic criterion resolves ties: a tie between two strategies that are equally good against a hypothesis is resolved by the first subsequent hypothesis in the subsequence for which one is a superior reply.

A lexicographic equilibrium is similar in a game with more players. For example, with three players and two hypotheses σ_n^0, σ_n^1 for each player n, player 1 optimizes against the primary hypothesis that 2 and 3 play according to (σ_2^0, σ_3^0), but if not and either 2 or 3 might deviate independently, then against a secondary hypothesis that is an average of (σ_2^1, σ_3^0) and (σ_2^0, σ_3^1), and the tertiary hypothesis that they are playing (σ_2^1, σ_3^1). See [7, Section 2] for the general representation and an example, and the explicit lexicographic representation derived in Section 5.10.

Van Damme shows that a quasi-perfect equilibrium induces a perfect equilibrium of the normal form and a sequential equilibrium of the extensive form. Moreover, by construction a quasi-perfect equilibrium provides for each player an optimal continuation from each of his information sets that is admissible [2]. If payoffs are generic then every sequential equilibrium is extensive-form perfect [26, 3] and quasi-perfect [21, 41]. However, if payoffs are nongeneric (as here in subsection 5.6) then quasi-perfection invokes a stronger form of sequential rationality than in Kreps and Wilson's [26, §4,5] definition of sequential equilibrium. Sequential equilibrium requires only that each player's continuation from an information set is optimal given the belief that the other player continues according to the first strategy in the subsequence enabling the information set to be reached, whereas quasi-perfection requires lexicographic optimality against the entire subsequence so that ties are resolved consistently and conditional admissibility is preserved.

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

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

Axiom A implies that each solution lies in a component of the admissible Nash equilibria, and Axiom B implies that it contains sequential equilibria with admissible continuations. If payoffs are generic then each equilibrium in the solution yields the same distribution over outcomes as its sequential equilibria. The latter property excludes outcomes that rely on non-credible threats, which is the chief motive for backward induction, but as discussed in Section 6.2, here we do not represent backward induction in terms of more primitive axioms.

3.3. **Invariance to Embedding.** 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, we define an embedding using the normal form $G : \Sigma \to \mathbb{R}^N$ of the extensive-form game Γ.

An embedding is described by a larger game $\tilde{G} : \tilde{\Sigma} \times \tilde{\Sigma}_o \to \mathbb{R}^{N \cup O}$ in which game G is embedded, subject to certain restrictions specified below. The larger game $\tilde{G}$ has *insiders* who are the players in N with strategy space $\tilde{\Sigma}$, and a set O of *outsiders* with strategy space $\tilde{\Sigma}_o$, 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 S_n but are equivalent to mixed strategies in Σ_n. The basic requirement is that an embedding should preserve the game among insiders, conditional on outsiders' actions.

These restrictions have a technical formulation. There should exist a multilinear map $f : \tilde{\Sigma} \times \tilde{\Sigma}_o \to \Sigma$ 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 $f_n : \tilde{\Sigma}_n \times \tilde{\Sigma}_o \to \Sigma_n$ such that $f_n(\cdot, \sigma_o) : \tilde{\Sigma}_n \to \Sigma_n$ is surjective for each profile $\sigma_o \in \tilde{\Sigma}_o$ of outsiders' strategies.

A statement of the axiom that uses this technical language could contain unsuspected implications, so after stating the formal definition we provide in Proposition 3.2 an equivalent formulation that is more detailed and more transparent, and that verifies the requisite properties.

Definition 3.1 (Embedding)**.** A game $\tilde{G} : \tilde{\Sigma} \times \tilde{\Sigma}_o \to \mathbb{R}^{N \cup O}$ and maps $f = (f_n)_{n \in N}$, where each map $f_n : \tilde{\Sigma}_n \times \tilde{\Sigma}_o \to \Sigma_n$ is multilinear, *embed* a game $G : \Sigma \to \mathbb{R}^N$ if for each $n \in N$

(a) $f_n(\cdot, \sigma_o) : \tilde{\Sigma}_n \to \Sigma_n$ is surjective for each $\sigma_o \in \tilde{\Sigma}_o$, and
(b) $\tilde{G}_n = G_n \circ f$.

Condition (a) ensures that embedding has no net effect on insiders' feasible strategies, conditional on outsiders' strategies, and condition (b) ensures that there is no net effect on insiders' payoffs.

Hereafter, if $(\tilde{G}, f)$ embeds G then we say that $\tilde{G}$ 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 metagame in extensive form that embeds a game in extensive form is constructed in proving Theorem 5.1, and an example is provided in Section 3.4.

Each multilinear map $f_n : \tilde{\Sigma}_n \times \tilde{\Sigma}_o \to \Sigma_n$ is completely specified by its values at profiles of pure strategies. Let $\hat{f}_n$ be the restriction of f_n to the set $\tilde{S}_n \times \tilde{S}_o$ of profiles of pure strategies of player n and outsiders in O, and let $\hat{f} = (\hat{f}_n)_{n \in N}$. The following proposition, proved in Appendix B, provides an alternative definition of embedding in terms of pure strategies.

Proposition 3.2. *$(\tilde{G}, f)$ embeds G if and only if for each player $n \in N$ there exists $\tilde{T}_n \subseteq \tilde{S}_n$ and a bijection $\pi_n : \tilde{T}_n \to S_n$ such that for each $(\tilde{s}, s_o) \in \tilde{S} \times \tilde{S}_o$ and $\tilde{t}_n \in \tilde{T}_n$:*

(1) $\hat{f}_n(\tilde{t}_n, s_o) = \pi_n(\tilde{t}_n)$, *and*
(2) $\tilde{G}_n(\tilde{s}, s_o) = G_n(\hat{f}(\tilde{s}, s_o))$.

Property (2) assures that insiders' 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 $\tilde{t}_n = \pi^{-1}(s_n) \in \tilde{T}_n$ independently of outsiders' strategies 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 $\hat{f}_n(\tilde{s}_n, s_o) = \sigma_n \notin S_n$ 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 Nash equilibria are not affected by embedding.

Proposition 3.3. *If $(\tilde{G}, f)$ embeds G then the Nash equilibria of G are the f-images of the Nash equilibria of $\tilde{G}$.*

An important corollary is that embedding does not introduce correlation among insiders' strategies. Axiom C is as follows, using Definition 3.1 of embedding.

Axiom C (Invariance to Embedding): If $(\tilde{G}, f)$ embeds G then the f-images of the solutions of $\tilde{G}$ are the solutions of G.

Axiom C is a commutativity property: starting from $\tilde{G}$, the same solutions are obtained whether one first maps $\tilde{G}$ to G via f and then finds solutions of G, or first finds solutions of $\tilde{G}$ and then maps them via f to solutions of G. When f is implicit we abbreviate Axiom C by saying that if $\tilde{G}$ projects to G then $\tilde{G}$'s solutions project to G's solutions.

In view of Proposition 3.3, Axiom C conforms to the principle that a refinement should inherit invariance properties of Nash equilibria. Two special cases are the following.

Small Worlds (Mertens [34]): Suppose $\tilde{\Sigma} = \Sigma$ and f is the identity map. Then Axiom C implies that a solution is not affected by adding players that are dummies with respect to the game G.

Invariance to Redundant Strategies: (Kohlberg and Mertens [24]): Suppose $\tilde{S}_o$ is a singleton and insiders' payoffs and strategies in $\tilde{G}$ differ from G only by treating some mixed strategies in Σ as additional pure strategies in $\tilde{S}$. Then Axiom C implies that solutions depend only on the game's reduced normal form obtained by deleting such redundant pure strategies. In [18] we show for the class of games considered here that Axiom B and Invariance to Redundant Strategies imply forward induction, and provide detailed examples.

The following example illustrates differences among Axiom C and its weaker variants. Suppose an embedding is obtained by adding actions by an outsider that determine which among several identical copies of the game is played by the insiders, and insiders other than player n do not observe this action. If the outsider is Nature then Invariance to Redundant Strategies implies no effect on solutions, whether or not n observes the action. If the action is taken by a strategic player then Small Worlds implies no effect on solutions if player n does not observe the action, whereas Axiom C implies no effect whether or not player n observes the action. Small Worlds can be envisioned as excluding dependence on 'downstream' outsiders who might be affected by insiders' actions, but insiders are not affected by outsiders' actions, whereas Axiom C also allows 'upstream' outsiders whose actions are observed by insiders but have no net effect on insiders' feasible strategies and payoffs.

3.4. **Example of the Application of Axiom C.** To illustrate Axiom C further, we demonstrate its application to an example. This example also provides a sketch of the metagame constructed in Sections 5.4-5.7 of the proof of the main theorem.

We apply Axiom C to the 'Beer-Quiche' signaling game G studied by Cho and Kreps [5] displayed in Figure 1. For this example we show in [18, §2.2] that Axioms B and C imply that the inessential component of Nash equilibria, in which both types S and W of player 1 choose Q, is not a solution. In the essential component, both types W and S of player 1 choose B, and 2 replies to B with R, but if 2 observes the deviation Q then 2 chooses R with any probability $r \leqslant 1/2$.

We show here that Axiom C requires a solution to include every value r^* of r in $(0, 1/2)$, and therefore every $r^* \in [0, 1/2]$ since a solution is a closed set. For each $r^* \in (0, 1/2)$ we define a family of metagames $\tilde{G}^\delta$, where $\delta > 0$ is a small parameter. Their form is sketched

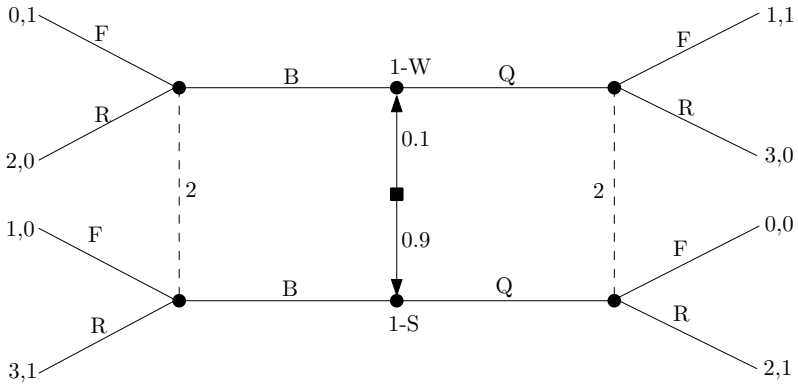


FIGURE 1. The Signaling Game Beer-Quiche

in Figure 2, omitting that in G player 2 responds to B or Q with F or R. Below we use XY to denote a pure strategy of 1 that chooses X if his type is S and chooses Y if his type is W.

In each metagame $\tilde{G}^\delta$, player 1 first chooses whether or not to play his strategy BB. If he decides not to play BB then next outsider 3 chooses one of two pure strategies labeled by $p \in \{0, 1\}$. Player 1, without observing 3's choice, then chooses among his other pure strategies and a redundant strategy, denoted $x(\delta)$ (the following analysis is modified slightly if player 1 observes 3's choice). The exact strategy in the original game of which $x(\delta)$ is a duplicate depends on 3's choice; in particular, it is equivalent to the mixed strategy that assigns probabilities $1 - \delta(9 - 8p)$, δp, and $\delta 9(1 - p)$ to BB, QB, and BQ, respectively. For any one of these strategies, 1's choice is implemented automatically in a copy of G. After 1's choice, 2 moves knowing only the message B or Q sent by 1. Note that after observing Q, 2 knows that 3 chose some p and then 1 chose either the $x(\delta)$-duplicate or some pure strategy other than BB. Finally, if the message was Q then outsider 4 moves and chooses $q \in \{0, 1\}$, knowing only that Q was sent. The outsiders' payoffs are constructed to have the following properties. Outsider 3's payoff is one if $p = q$ and zero otherwise; i.e. 3 wants to mimic 4. Outsider 4 wants to choose $q = 1$ if 2's strategy after observing Q uses $r < r^*$, and to choose $q = 0$ if $r > r^*$ (see Section 5.6).

The metagame $\tilde{G}^\delta$ embeds the original game G, so Axioms B and C require the metagame to have a quasi-perfect equilibrium that maps to a point in the solution of G. In this equilibrium, 1 must choose BB initially, and 2 must reply to B and Q as in an equilibrium in the solution of G. As above, let r be the probability that 2 responds to Q with R. A calculation shows that Player 1 strictly prefers $x(\delta)$ to QB and QQ conditional on either p. Therefore, sequential rationality requires that after rejecting BB, 1 chooses either BQ or $x(\delta)$. It also requires that after observing Q, 2's choice between F and R must be optimal based on the conditional probability of 1's type S derived via Bayes' Rule from some mixture

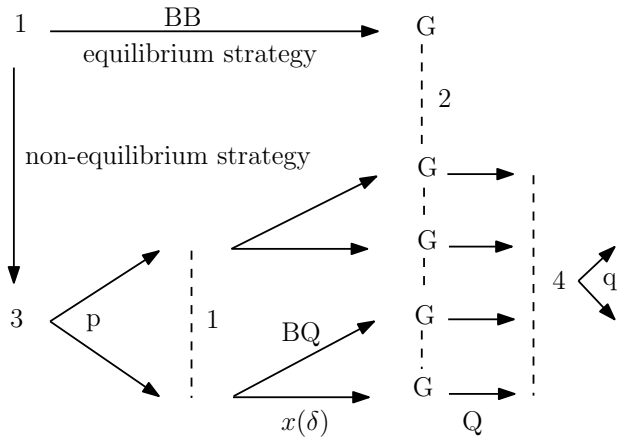


FIGURE 2. Sketch of the Metagame

of the duplicates and BQ. For instance, if 2 expects 3 to choose p and 1 to choose $x(\delta)$ then this conditional probability is p.

Suppose first that $r > r^*$. Then 4 prefers $q = 0$, so 3 prefers $p = 0$. In this case, neither the duplicate nor BQ has type S sending message Q, so 2's conditional probability of S after observing Q is necessarily zero. But then 2 strictly prefers F, contradicting the supposition that $r > r^*$. Similarly, if $r < r^*$ then 4 prefers $q = 1$, so 3 prefers $p = 1$. In this case, message Q is sent only if 1 chooses either BQ or the duplicate and the duplicate selects QB. But 1's payoff from BQ is $2.8 + 0.2r$, and from the duplicate it is $(1 - \delta)2.9 + \delta(0.2 + 1.8r)$, so if $r < (1-27\delta)/(2-18\delta)$ then 1 prefers the duplicate, implying that 2's conditional probability of S after observing Q is one and therefore 2 strictly prefers R, contradicting the supposition that $r < r^*$. This is indeed the case if δ is sufficiently small because $r^* < 1/2$ and thus $r < r^* < (1 - 27\delta)/(2 - 18\delta)$. Hence the supposition that $r < r^*$ is false for small δ. In sum, for any sufficiently small δ, in the metagame $\tilde{G}^\delta$ there is no quasi-perfect equilibrium for which $r \neq r^*$. In the quasi-perfect equilibrium with $r = r^*$, outsiders 3 and 4 use the mixed strategies (1/2,1/2), 1 chooses $x(\delta)$ with probability 1, and 2 is indifferent whether to choose F or R after observing Q. Therefore $r = r^*$ in the corresponding equilibrium of G.

In [17, §4.2] we sketch how Axiom C is applied to a game with perfect information.

3.5. **Summary of the Axioms.** We study refinements that are independent of embeddings in metagames that preserve insiders' feasible strategies and payoffs for every profile of outsiders' strategies. And, we require that each solution is a closed connected set of undominated equilibria, including one that is quasi-perfect.

4. Additional Notation and Properties

In this section we prepare for the statements and proofs of the theorems in Section 5.

4.1. **Quasi-Perfect Equilibrium.** The proof of Theorem 5.1 invokes van Damme's [50, p. 8] original definition.

Definition 4.1 (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.

Equivalently, 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, for some sequence $\hat{b}^k \in B \setminus \partial B$ converging to b,

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

This definition is equivalent to the one in Section 3.2 as a lexicographic equilibrium. 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 σ 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 $\hat{\sigma}^k$ converges to a mixed strategy that enables the same outcome distribution as σ does.

4.2. **Additional Notation for the Extensive Form.** Let X be the set of nodes in the game tree Γ, and let X_n be the set of nodes where player n moves, partitioned into his information sets $h \in H_n$. For a node $x \in X_n$, $h(x)$ is the unique information set $h \in H_n$ that contains x. For each n and $h \in H_n$, $A_n(h)$ is 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 $a \in A_n(h(x))$ we 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 $\beta(y, s) \equiv \beta_0(y) \prod_n \beta_n(y, s_n)$.

For each node y the function $\beta_n(y, \cdot)$ extends to a function over Σ_n via $\beta_n(y, \sigma_n) = \sum_{s_n \in S_n} \beta_n(y, s_n)\sigma_n(s_n)$ for $\sigma_n \in \Sigma_n$. 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 β to profiles of mixed or behavioral strategies. Given a mixed-strategy profile $\sigma \in \Sigma$, the probability that outcome z results is $\beta(z,\sigma) = \beta_0(z)\prod_n \beta_n(z,\sigma_n)$.

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 σ_n then $\rho_n(\sigma_n) = (\beta_n(z,b_n))_{z\in Z}$. Let P_n be the image of ρ_n, and $P = \prod_n P_n$ the image of ρ. Then P_n is a compact convex polyhedron, called the space of n's enabling strategies in [10]. Each vertex of P_n corresponds to an equivalence class of n's pure strategies that exclude the same outcomes. The vertices of P_n are n's pure strategies in the 'pure reduced normal form' defined by Mailath, Samuelson, and Swinkels [30]; see Appendix A for illustrations.

If $\sigma \in \Sigma$ and $p = \rho(\sigma)$ then the probability of outcome z is $\gamma_z(p) = \beta_0(z)\prod_n p_n(z)$. Thus the function $\gamma : P \to \Delta(Z)$ summarizes the extensive form. The analog of the game Γ's normal form $G : \Sigma \to \mathbb{R}^N$ is the enabling form $\mathcal{G} : P \to \mathbb{R}^N$ that assigns to each profile of enabling strategies the profile of players' expected payoffs, where $\mathcal{G}_n(p) = \sum_z \gamma_z(p)u_n(z)$ if $u_n(z)$ is n's payoff from node z. 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. The axioms have direct analogs in terms of enabling strategies, as shown in [17].

Enabling strategies are sufficient representations for games in extensive form with perfect recall, and minimal representations when payoffs are generic. 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 [10, §5]. Because Axiom C implies Invariance to Redundant Strategies, it is immaterial whether solutions are characterized in terms of mixed 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 5.2 below. Using these features, Section 5 derives the implications of the axioms in terms of enabling strategies.

4.4. **Generic Payoffs.** 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 in that there exists a lower-dimensional subset $U_\circ$ of U such that our results are true for all games in $U \setminus U_\circ$. The set $U_\circ$ includes the nongeneric set described in [9]. Therefore, each game outside $U_\circ$ has finitely many equilibrium outcome distributions, and in particular all

equilibria in a component yield the same distribution over outcomes. In addition, Lemma D.10 requires a characterization of stable sets that in [15] we show is valid for all games outside a lower-dimensional set. Finally, for a game with two players, the constructions in Appendix D rely on certain polyhedra being in general position. Each of these polyhedra, of which there are finitely many, is a set of enabling strategies for one player against which, for strategies of the other player in a certain class, a subset is optimal. Since these are defined by linear equations and inequalities in the payoffs of the other player, the set of games where the arguments fail is a lower-dimensional set.

4.5. **The Version of Stability.** Lastly, we identify the version of stability used in Section 5. Kohlberg and Mertens [24] offered a tentative definition of stability but reported an example for which such a stable set fails to contain a sequential equilibrium. Mertens' [32, 34] revised formulation eliminates this deficiency, but allows variants depending on the coefficient module used for (co)homology. Among these, for p either zero or a prime integer, only the p-stable sets satisfy decomposition and small-worlds. The p-stable sets use the field of integers modulo p as the coefficient module, with the convention that it is the field of rationals when p is zero. We prove in Appendix C that p-stability satisfies Axiom C for any such p. However, in [15] we show for extensive-form games with generic payoffs that if a set is stable for any variant then it is 0-stable.[5] Therefore, hereafter by stability we mean 0-stability unless explicitly stated. With this convention, in the next section we show for any refinement that satisfies the axioms that each solution of a two-player game in extensive form with perfect recall and generic payoffs must be a stable set.

5. Statement and Proof of the Theorem

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 ρ of a solution $\Sigma^* \subset \Sigma$. We say that Q^* is stable if Σ^* is stable.[6]

Theorem 5.1. *If a refinement satisfies Axioms A, B, and C then each solution of a two-player game with perfect recall and generic payoffs is a stable set.*

The proof of Theorem 5.1 occupies the remainder of this section. Hereafter, the set of players is $N = \{1, 2\}$, and when we consider a typical player n the other player is denoted by $m \neq n$.

[5]Even for games with nongeneric payoffs, the 0-stable sets seem right on heuristic grounds because p-stability excludes a set having a neighborhood in the graph of equilibria over perturbed strategies whose projection map has a cycle of order p.

[6]Alternatively, one can apply the definition of stability directly to Q^* as a component of equilibria, represented in terms of enabling strategies, in the graph over the space of perturbations of players' enabling strategies as in footnote 4. As noted by Mertens [34, 36], more generally one can apply analogs of stability and the axioms to games in strategic form for which each player's strategy set is a convex polyhedron and payoffs are multilinear functions defined on the product of players' strategy sets.

5.1. **A Characterization of Stable Sets.** We first derive a characterization of a stable set for a two-player game with generic payoffs. It is simpler than the general definition in Mertens [32] and for readers unfamiliar with homology theory it can be taken as the definition. Mertens' general definition (see footnote 4) requires essentiality of the local projection map from the graph of equilibria over the space of perturbed strategies, whereas Theorem 5.2 below uses the graph of lexicographically optimal replies to deviations from equilibria play. This graph depends only on the given game, rather than perturbed games, and considers only players' preferred responses to deviations.

Let $\bar{\Sigma}^*$ be a component of the equilibria of Γ in terms of mixed strategies, and let $\bar{\Sigma}^*_n$ be the projection of $\bar{\Sigma}^*$ in Σ_n. Also let Σ^* be a component of the undominated equilibria of the game Γ that is contained in $\bar{\Sigma}^*$. Let Q^* be the image of Σ^* under ρ and for each n let Q^*_n be the image of Σ^*_n under ρ_n, i.e. represented in enabling strategies.

By genericity, all equilibria in $\bar{\Sigma}^*$ induce the same distribution over outcomes. Therefore, for each node x, $\beta(x,\sigma)$ is the same for all $\sigma \in \bar{\Sigma}^*$; 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 σ_n of player n in $\bar{\Sigma}^*$. 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 $\bar{\Sigma}^*$.

Let $S^0_n \subset S_n$ be the set of pure strategies s^0_n with the property that, at each information set $h \in H^*_n$ that s^0_n does not exclude, s^0_n prescribes an action in A^*_n. Let $S^1_n = S_n \setminus S^0_n$, i.e. each pure strategy s_n in S^1_n chooses a non-equilibrium action at some information set $h \in H^*_n$ that it does not exclude.

For $i = 0, 1$, let Σ^i_n be the set of mixed strategies whose support is contained in S^i_n. Observe that the support of n's strategy in every equilibrium in $\bar{\Sigma}^*$ is contained in S^0_n and that every strategy in S^0_n is a best reply against every equilibrium in $\bar{\Sigma}^*$. Thus $\bar{\Sigma}^*_n$ is contained in Σ^0_n and $\bar{\Sigma}^* = \bar{\Sigma}^*_1 \times \bar{\Sigma}^*_2$. Hence $\Sigma^* = \Sigma^*_1 \times \Sigma^*_2$ where Σ^*_n is a component of the intersection of $\bar{\Sigma}^*_n$ with the set of undominated strategies.

If S^1_n is empty for each player n then each equilibrium in $\bar{B}^*$ is completely mixed; by genericity, $\bar{B}^*$ is a singleton and its equivalent mixed strategy is stable. Thus if a solution concept satisfying the axioms selects this equilibrium then it is automatically stable. The only interesting case, therefore, is one where S^1_n is nonempty for at least one of the players. To avoid different cases, we assume that S^1_n is nonempty for each n. Along the way we indicate how the proof changes when S^1_n is empty for exactly one player.

Let P^0_n be n's set of enabling strategies in the image of Σ^0_n under ρ_n. Let $Z^1_n \subset Z$ 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 $Z^0_n = Z_n \setminus Z^1_n$. Then P^0_n is the set of $p_n \in P_n$ such that $p_n(z) = 0$ for all $z \in Z^1_n$ and thus P^0_n is a face

of P_n. However, the image P_n^1 of Σ_n^1 under ρ_n need not be a face of P_n. For $i = 0, 1$, let $P^i = P_1^i \times P_2^i$ and define $\mathbb{P} = P^0 \times P^1$.

For each enabling strategy $p_n \in P_n$, let $\Psi_{Z_n^1}(p_n)$ be the projection of p_n to $\mathbb{R}_+^{Z_n^1}$; then $\Psi_{Z_n^1}(p_n) = 0$ iff $p_n \in P_n^0$. Fix a point $\bar{p}_m$ in the interior of P_m and define $\eta_n : P_n \to \mathbb{R}$ by $\eta_n(p_n) = \sum_{z \in Z_n^1} p_0(z)\bar{p}_m(z)p_n(z)$, 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 $\mathcal{H}_n$ be the hyperplane in $\mathbb{R}^{Z_n^1}$ with normal $\left(p_0(z)\bar{p}_m(z)\right)_{z \in Z_n^1}$ and constant ε. Then $\mathcal{H}_n$ separates the origin from $\Psi_{Z_n^1}(P_n^1)$. Let Π_n^1 be the intersection of $\mathcal{H}_n$ with $\Psi_{Z_n^1}(P_n)$. Let $\bar{\pi}_n^1$ be the function from $P_n \setminus P_n^0$ to Π_n^1 that maps each $p_n \notin P_n^0$ to the point $\varepsilon(\eta_n(p_n))^{-1}\Psi_{Z_n^1}(p_n)$.

In the following we invoke lexicographically optimal replies as defined in Blume, Brandenburger, and Dekel [2] and Govindan and Klumpp [7]. Recall that n's strategy σ_n is lexicographically optimal against a sequence $(\sigma_m^k)_{k=1,2,..}$ of m's strategies if any alternative strategy $\hat{\sigma}_n$ that is a better reply to σ_m^k for some k is a worse reply to σ_m^j for some $j < k$.

Given Q^*, let $\mathcal{Q}$ be the set of those $(q^*, (p^0, p^1), \pi^1) \in Q^* \times \mathbb{P} \times \Pi^1$ such that there exist $r^0, \tilde{p}^0 \in P^0$, $r^1 \in P^1$, and for each n scalars $\lambda_n^0, \lambda_n^1, \mu_n^1$ in the interval $(0, 1]$ such that, if

$$q_n^0 = \lambda_n^0 p_n^0 + (1 - \lambda_n^0) r_n^0 \quad \text{and} \quad q_n^1 = (1 - \lambda_n^1)\tilde{p}_n^0 + \lambda_n^1(\mu_n^1 p_n^1 + (1 - \mu_n^1) r_n^1),$$

then for each n:

(i) $\bar{\pi}_n^1(q_n^1) = \pi_n^1$.
(ii) q_n^0, and r_n^0 if $\lambda_n^0 < 1$, are lexicographically optimal replies against (q_m^*, q_m^0, q_m^1).
(iii) If $\mu_n^1 < 1$ then r_n^1 is an optimal reply against q_m^* and lexicographically as good a reply against (q_m^*, q_m^0, q_m^1) as other strategies in P_n^1.

In case S_n^1 is empty (and S_m^1 is not) then the set Π_n^1 is empty so we set $\mathbb{P} = P^0 \times P_m^1$ and points in $\mathcal{Q}$ then have the form $(q^*, (p^0, p_m^1), \pi_m^1)$, and we drop the optimality requirement (iii) for n.

Interpretation. The set $\mathcal{Q}$ is the graph of lexicographically optimal replies to possible deviations from equilibria in Q^*. The formulation appears complicated only because of the need to consider for each player n both enabling strategies p_n^0 and p_n^1 that do and do not adhere to equilibrium play, and also possible mixtures of these with others having the same properties, so that altered probabilities of actions are included. For each equilibrium q^* one considers for each player n a pair (p_n^0, p_n^1) of enabling strategies such that p_n^0 conforms to the equilibrium and p_n^1 deviates. Further, one considers a mixture q_n^0 of p_n^0 and some other conforming strategy r_n^0, and also a mixture q_n^1 of some conforming strategy $\tilde{p}_n^0$ and a mixture of the nonconforming strategy p_n^1 and some other nonconforming strategy r_n^1, where condition (i) requires that q_n^1 yields the specified probabilities π^1 on those terminal nodes excluded by equilibrium behavior. From these strategies one obtains the sequence (q_n^*, q_n^0, q_n^1) of alternative hypotheses about n's strategies, ordered lexicographically. For the

case that the mixtures give positive weight to the alternative conforming strategy r_n^0 and nonconforming strategy r_n^1, one requires in (ii) that r_n^0 is lexicographically optimal against the other's sequence, and in (iii) that r_n^1 is an optimal reply to the equilibrium q^* and lexicographically optimal among nonconforming strategies. The point $(q^*, (p^0, p^1), \pi^1)$ is then in the graph above the point (p^0, p^1) describing the players' primary conforming and nonconforming strategies, if each player n's mixture q_n^0 of p_n^0 and r_n^0 is lexicographically optimal against the other's sequence.

Let $\Psi : \mathcal{Q} \to \mathbb{P}$ be the natural projection, i.e. $\Psi(q^*, (p^0, p^1), \pi^1) = (p^0, p^1)$. Let $\partial\mathcal{Q} = \Psi^{-1}(\partial\mathbb{P})$. The projection map Ψ is *essential* if every continuous map $\phi : \mathcal{Q} \to \mathbb{P}$ has a point of coincidence with Ψ, i.e. $\phi(x) = \Psi(x)$ for some $x \in \mathcal{Q}$.

Theorem 5.2. *$(\mathcal{Q}, \partial\mathcal{Q})$ is a pseudomanifold of the same dimension as $(\mathbb{P}, \partial\mathbb{P})$. Moreover, Q^* is stable if and only if the projection map $\Psi : (\mathcal{Q}, \partial\mathcal{Q}) \to (\mathbb{P}, \partial\mathbb{P})$ is essential.*

This characterization of a stable set is proved in Appendix D.

5.2. **Plan of the Proof.** Next we outline the construction used to prove Theorem 5.1.

Let $\hat{\Sigma} \subset \Sigma$ be a solution of the game Γ and let $\hat{Q} \equiv \rho(\hat{\Sigma})$ be the set of enabling strategies equivalent to $\hat{\Sigma}$. By Axiom A, $\hat{\Sigma}$ is a connected set of equilibria in undominated strategies. Hence it belongs to a component Σ^* of equilibria in undominated strategies and thus $\hat{Q}$ is contained in $Q^* \equiv \rho(\Sigma^*)$. We show that $\hat{Q}$ equals Q^* and is stable. We accomplish this as follows. We take an arbitrary point $q^{0,*} \in Q^*$ and an arbitrary neighborhood $U(q^{0,*})$ of $q^{0,*}$. Then we construct a corresponding sequence of metagames $\tilde{\Gamma}^\delta$ as a parameter δ converges to zero. Using Axioms B and C, for each metagame $\tilde{\Gamma}^\delta$ in the sequence, there exists a quasi-perfect equilibrium whose projection, call it $q^{0,\delta}$, to P is contained in $\hat{Q}$. Take any such sequence of $q^{0,\delta}$ converging to some point $q^{0,0}$ in $\hat{Q}$. We show that: (i) the limit point $q^{0,0}$ is in $U(q^{0,*})$, hence $\hat{Q} = Q^*$; and (ii) the existence of such a sequence converging to a limit point in $U(q^{0,*})$ implies that the projection map Ψ is essential. Then Theorem 5.2 implies that Q^* is stable.

5.3. **Preliminaries.** In this subsection we lay the groundwork for the metagames constructed in the proof.

For the given set Q^* containing the solution $\hat{Q}$, let $(\mathcal{Q}, \partial\mathcal{Q})$ be the associated pseudomanifold constructed in Section 5.1. Let $q^{0,*}$ be an arbitrary point in Q^* and let $U(q^{0,*})$ be a neighborhood of $q^{0,*}$. For each player m, choose a point $p_m^{0,*}$ in the interior of P_m^0 against which $q^{0,*}$ is a best reply and strategies in P_n^1 are inferior replies. Such a choice is possible by genericity of payoffs: the interior of the projection of $\bar{\Sigma}_m^*$, which is the component of m's equilibrium strategies that contains Σ_m^*, belongs to the interior of P_m^0 and all strategies in P_n^1 are inferior replies against every such point. Since $q_n^{0,*}$ belongs to Q_n^*, which consists only of undominated strategies, there exists a point p_m in the interior of P_m against which $q_n^{0,*}$

is a best reply. p_m is equivalent to a completely mixed strategy σ_m in Σ_m. Also, by the genericity of payoffs, we can choose p_m to be such that strategies in P_n^0 that do not belong to the face containing $q_n^{0,*}$ in its interior are strictly inferior replies against p_m. Express σ_m as a convex combination of σ_m^0 and σ_m^1, where for $i = 0, 1$, σ_m^i belongs to the interior of Σ_m^i. Let p_m^0 and $p_m^{1,*}$ be the enabling strategies that are equivalent to σ_m^0 and σ_m^1, respectively. Then p_m^0 and $p_m^{1,*}$ are in the relative interiors of P_m^0 and P_m^1 respectively, and p_m is a convex combination of p_m^0 and $p_m^{1,*}$. It follows that $x^* \equiv (q^{0,*}, (p^{0,*}, p^{1,*}), \pi^{1,*})$ belongs to $\mathcal{Q} \setminus \partial\mathcal{Q}$, where for each n, $\pi_n^{1,*} = \bar{\pi}_n^1(p_n^{1,*})$, and in the definition of $\mathcal{Q}$, q_n^0 is $p_n^{0,*}$, and q_n^1 is p_n, which is a convex combination of p_n^0 and $p_n^{1,*}$.

It follows from our construction in Appendix D that x^* belongs to the interior of a polyhedron of the same dimension as $\mathbb{P}$. Therefore, we can choose a neighborhood $V(x^*)$ of x^* that is homeomorphic to a simplex, is contained in $\mathcal{Q} \setminus \partial\mathcal{Q}$, and is such that the projection onto the first factor is contained in the neighborhood $U(q^{0,*})$, i.e. if $(q^0, (p^0, p^1), \pi^1) \in V(x^*)$ then $q^0 \in U(q^{0,*})$.

In Appendix E we construct a continuous map $g : \mathcal{Q} \to \mathbb{P}$ and a constant $\alpha > 0$ such that $\|g(x) - \Psi(x)\| \leqslant \alpha$ for some $x \in \mathcal{Q}$ only if Ψ is essential and $x \in V(x^*)$. Now extend the map g to the whole of $P^0 \times \mathbb{P} \times \Pi$ in an arbitrary fashion, calling it still g. Also, we now view Ψ as the projection from $P^0 \times \mathbb{P} \times \Pi$ to $\mathbb{P}$. 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 $\mathcal{L}_n^\Pi$ of Π_n^1 such that, letting $\mathcal{L}$ be the resulting multisimplicial subdivision of $P^0 \times \mathbb{P} \times \Pi^1$, g has a multisimplicial approximation $\tilde{g}$ [12, Theorem 6] 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 $\|g(x) - \Psi(x)\| \leqslant \alpha$: this follows from the fact that, since $\tilde{g}$ is a multisimplicial approximation of g, $\tilde{g}(x)$ belongs to the multisimplex that contains $g(x)$ in its interior. An important implication of this observation is that if, in particular, this x also belongs to $\mathcal{Q}$, then Ψ is essential, $x \in V(x^*)$, and $q^0 \in U(q^{0,*})$. Thus, that such a point belongs to $\mathcal{Q}$ will be the final step of the proof in Section 5.12.

As in [12, Appendix B], now take a further polyhedral subdivision $\mathcal{T}$ of $\mathcal{L}$ and let γ be the convex function generated by $\mathcal{T}$, i.e. γ is piecewise linear and linear precisely on each full-dimensional polyhedron of the subdivision.

5.4. **A Game with Redundant Strategies.** In this subsection we construct from Γ a larger game by adding redundant pure strategies that will be the basis for the metagames specified in Section 5.5. Because Axiom C implies Invariance to Redundant Strategies, the solutions of Γ are equivalent to the solutions of this larger game.

For each fixed $\hat{p} = (p^0, p^1) \in \mathbb{P}$ and $\delta \in (0, 1)$, consider the following game $\Gamma(\delta, \hat{p})$. The players act simultaneously with each player choosing a strategy as follows. Unaware of his

opponent's choices, player n initially chooses provisionally some pure strategy $s_n^0 \in S_n^0$, or he rejects all strategies in S_n^0.

- If initially he chooses a strategy s_n^0 then at a subsequent second stage he can retain s_n^0 or revise his choice. If he chooses to revise his choice, then at a third stage the revisions available are the 'duplicate' pure strategies in the set $T_n^0(\delta, p_n^0)$ consisting of all mixed strategies of the form $t_n(\delta, p_n^0) \equiv (1-\delta)t_n + \delta p_n^0$ for some $t_n \in S_n^0$.
- If he rejects all strategies in S_n^0 at the first stage, then at a second stage he can choose among the strategies in $T_n^0(\delta, p_n^0)$ or reject them all.[7] If he rejects them all then at a third stage he chooses 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 $\tilde{S}_n(\delta, \hat{p}_n) \equiv S_n \cup T_n^0(\delta, p_n^0) \cup T_n^1(\delta, p_n^1)$. Thus, game $\Gamma(\delta, \hat{p})$ has the same reduced normal form as Γ.

5.5. **Extensive Form of the Metagames.** Now we specify a family of similar metagames $\tilde{\Gamma}^\delta$, one for each $\delta \in (0,1)$.

Before the insiders play, thirteen outsiders, denoted players o_0 and $o_{n,j}^i$ 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 $o_{n,j}^i$, for $n = 1, 2$, $j = 1, 3$ and $i = 0, 1$, chooses a point in the vertex set W_n^i of $\mathcal{K}_n^i$. Outsider $o_{n,2}^0$ chooses a point in a finite subset $S_n^{0,\delta}$ of P_n^0 chosen such that every point in P_n^0 is within δ of some point in $S_n^{0,\delta}$; outsider $o_{n,2}^1$ chooses a point in a finite subset $S_n^{1,\delta}$ of Π_n^1 such that every point in Π_n^1 is within δ of some point in $S_n^{1,\delta}$.

For outsider $o_{n,1}^i$, each pure strategy v_n^i corresponds to a point in P_n^i denoted $p_n^i(v_n^i)$. Therefore, each mixed strategy $\sigma_{n,1}^i$ corresponds to a point in P_n^i, denoted $p_n^i(\sigma_{n,1}^i)$, which is obtained by taking the appropriate average of the points induced by the pure strategies in the support of $\sigma_{n,1}^i$. Likewise a mixed strategy $\sigma_{n,2}^0$ of $o_{n,2}^0$ corresponds to a point $q_n^0(\sigma_{n,2}^0)$ in P_n^0, and a mixed strategy $\sigma_{n,2}^1$ of $o_{n,2}^1$ corresponds to a point $\pi_n^1(\sigma_{n,2}^1)$ in Π_n^1.

A mixed-strategy profile $\tilde{\sigma}_o$ for the outsiders induces a point $(q^0(\tilde{\sigma}_o), (p^0(\tilde{\sigma}_o), p^1(\tilde{\sigma}_o)), \pi^1(\tilde{\sigma}_o))$ in $P^0 \times \mathbb{P} \times \Pi^1$, where for each n and i, $p_n^i(\tilde{\sigma}_o)$ depends on the choice by $o_{n,1}^i$, and $q_n^0(\tilde{\sigma}_o)$ and $\bar{\pi}_n^1(\tilde{\sigma}_o)$ depend on the choices by $o_{n,2}^0$ and $o_{n,2}^1$ respectively.

After each pure-strategy profile $\tilde{s}_o$ of the outsiders there follows a copy of the game $\Gamma(\delta, \hat{p}(\tilde{s}_o))$. That is, if in the profile $\tilde{s}_o$ outsiders $o_{n,1}^i$ 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 $\tilde{\Gamma}^\delta$ are such that the insiders play without knowing which copy of $\Gamma(\delta, \hat{p})$

[7] 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.

they are playing. The sets of 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 about 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 $t_n^i(\delta) \in T_n^i(\delta)$ depends on the choice by outsider $o_{n,1}^i$, which insiders do not observe. Thus in the metagame $\tilde{\Gamma}^\delta$, player n's set of pure strategies, up to duplication of pure strategies, is $\tilde{S}_n(\delta) \equiv S_n \cup T_n^0(\delta) \cup T_n^1(\delta)$.

Proposition 3.2 implies that the metagame $\tilde{\Gamma}^\delta$ embeds Γ. The strategies in S_n are available as pure strategies in $\tilde{S}_n(\delta)$ and the other pure strategies, which belong to $T_n^i(\delta)$, for $i = 0, 1$, implement mixtures in Σ_n that depend on the choices of outsiders $o_{n,1}^i$.

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

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

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

Outsider $o_{n,2}^0$ wants to mimic the actual choice implemented by player n when this choice belongs to P_n^0. Similarly, outsider $o_{n,2}^1$ wants to mimic the π_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 φ_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 $o_{n,2}^i$ chooses a pure strategy $s_{n,2}^{i,\delta}$ and n chooses a pure strategy $\tilde{s}_n$ in $\tilde{S}_n(\delta)$. If $\tilde{s}_n$ 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 $o_{n,1}^i$; and otherwise let $q_n = \tilde{s}_n$. For outsider $o_{n,2}^0$, his payoff is zero if $q_n \notin P_n^0$; otherwise, it is $\xi_n^0(s_{n,2}^{0,\delta}, q_n)$. For outsider $o_{n,2}^1$, his payoff is zero if $q_n \in P_n^0$; otherwise it is $\xi_n^1(s_{n,2}^{1,\delta}, \bar{\pi}_n^1(q_n))$.

The payoff to outsider $o_{n,3}^i$ depends on the choices of all other outsiders. If o_0 chooses a polyhedron T then there exists a unique multisimplex L of $\mathcal{L}$ that contains T. For each vertex w_n^i of W_n^i, and each vertex $\tilde{v}$ of L, let $u_{n,3}^i(T, \tilde{v}, w_n^i) = 1$ 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 $u_{n,3}^i$ extends multilinearly to L and, since L is full-dimensional, to the whole of $P^0 \times \mathbb{P} \times \Pi^1$, denoted still by $u_{n,3}^i(T, \cdot, w_n^i)$. Given an arbitrary mixed strategy of the other players, if o_0 chooses

T and $o^i_{n,3}$ chooses w^i_n then the payoff of $o^i_{n,3}$ is $u^i_{n,3}(T,(p,q),w^i_n)$, where (p,q) is the point in $P^0\times\mathbb{P}\times\Pi^1$ induced by the mixed strategies of the other players.

5.7. **Outsiders' Strategies in a Quasi-Perfect Equilibrium.** In this subsection we derive the relevant features of outsiders' strategies in a quasi-perfect equilibrium.

By Axioms B and C, in the metagame $\tilde{\Gamma}^\delta$ there exists a quasi-perfect equilibrium $\tilde{b}^\delta$ whose equivalent mixed-strategy profile $\tilde{\sigma}^\delta$ belongs to a solution and whose image under the map from the metagame $\tilde{\Gamma}^\delta$ to Γ is a point in the solution $\hat{Q}$ for the original game Γ.

Each player n's strategy in $\tilde{b}^\delta$ necessarily has the following feature. He avoids going to his information set where his choices are among the strategies in $S^1_n(\delta)\cup T^1_n(\delta)$, since each of these strategies chooses a non-equilibrium action at some information set on the equilibrium path. Let $q^{0,\delta}_n$ be n's actual strategy in P^0_n that is implemented by n's strategy in the profile $\tilde{b}^\delta$ in the metagame $\tilde{\Gamma}^\delta$. By construction, $q^{0,\delta}_n$ belongs to $\hat{Q}$.

Let $\tilde{x}^\delta\equiv(\tilde{q}^{0,\delta},(p^{0,\delta},p^{1,\delta}),\tilde{\pi}^{1,\delta})$ be the point in $P^0\times\mathbb{P}\times\Pi^1$ that is induced by the profile $\tilde{\sigma}^\delta_o$ of the outsiders' strategies in the equilibrium $\tilde{\sigma}^\delta$.

Under $\tilde{b}^\delta$, after n has rejected all strategies in S^0_n and $T^0_n(\delta)$, consider the strategy implemented by n. Let $(1-\alpha^{1,\delta}_n)$ be the total probability of choosing a duplicate in $T^1_n(\delta)$ under $\tilde{b}^\delta$. Then n's choice at this information set is equivalent to an enabling strategy in P_n of the form

$$\check{q}^\delta_n\equiv(1-\alpha^{1,\delta}_n)((1-\delta)\check{p}^{0,\delta}_n+\delta p^{1,\delta}_n)+\alpha^{1,\delta}_n r^{1,\delta}_n\,,$$

where: (i) $\check{p}^0_n(\delta)$ is the mixture over strategies t_n such that the strategy $t^1_n(\delta)$ is played with positive probability at this information set; (ii) $p^{1,\delta}_n=p^1_n(\tilde{\sigma}^1_{n,1})$ is the enabling strategy induced by the equilibrium strategy $\tilde{\sigma}^1_{n,1}$ of outsider $o^1_{n,1}$; (iii) $r^{1,\delta}_n$ is the enabling strategy in P^1_n that is obtained from n's actual mixture over strategies in S^1_n if $\alpha^{1,\delta}_n>0$, and is arbitrary otherwise. Let

$$q^{1,\delta}=(\delta(1-\alpha^{1,\delta}_n)p^{1,\delta}_n+\alpha^{1,\delta}_n r^{1,\delta}_n)/(\delta(1-\alpha^{1,\delta}_n)+\alpha^{1,\delta}_n)\,.$$

Then $\bar{\pi}^1_n(\check{q}^\delta_n)=\bar{\pi}^1_n(q^{1,\delta}_n)\equiv\pi^{1,\delta}_n$.

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

Lemma 5.3. *The equilibrium strategies of the outsiders satisfy the following properties.*

1. *For each n, suppose the vertices in the support $W^{i,\delta}_n$ of $o^i_{n,3}$'s equilibrium strategy span a simplex $K^{i,\delta}_n$ of $\mathcal{K}^i_n$. Then $p^{i,\delta}_n$ belongs to $K^{i,\delta}_n$.*
2. *If every polyhedron in the support of o_0's strategy contains $\tilde{x}^\delta$ then, for each n and i, the vertices in $W^{i,\delta}_n$ span a simplex $K^{i,\delta}_n$, and $\tilde{g}_{n,i}(\tilde{x}^\delta)$ belongs to the interior of a simplex $\bar{K}^{i,\delta}_n$ that has $K^{i,\delta}_n$ as a face.*
3. *Every polyhedron in the support of o_0's strategy contains $\tilde{x}^\delta$.*
4. *$\tilde{q}^{0,\delta}_n$ is within δ of $q^{0,\delta}_n$ and $\tilde{\pi}^{1,\delta}_n$ is within δ of $\pi^{1,\delta}_n$.*

Proof of Lemma. Outsider $o^i_{n,1}$ wants to mimic outsider $o^i_{n,3}$. So, if the vertices of $W^{i,\delta}_n$ span a simplex $K^{i,\delta}_n$ then the payoff to $o^i_{n,1}$ from choosing a vertex w^i_n is positive if it belongs to $W^{i,\delta,}_n$ and zero otherwise. Point (1) follows.

Let $\bar{L} = \tilde{L}^0 \times (L^0 \times L^1) \times L^\Pi$ be the unique multisimplex of $\mathcal{L}$ that contains $\tilde{x}^\delta$ in its interior. For each polyhedron T in the support of o_0's strategy, there exists a full-dimensional multisimplex $\hat{L}$ of $\mathcal{L}$ that contains T. Obviously $\hat{L}$ has $\bar{L}$ as a face. $o^i_{n,3}$'s payoff from choosing a strategy w^i_n 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 $\tilde{g}^i_n$ and zero otherwise. Since the image of the vertices of $\bar{L}$ under the coordinate function $\tilde{g}^i_n$ span a simplex $\bar{K}^{i,\delta}_n$, $\tilde{g}^i_n(\tilde{x}^\delta) \in \bar{K}^{i,\delta}_n$ and the vertices of $W^{i,\delta}_n$ span a face of $\bar{K}^{i,\delta}_n$. Therefore, point (2) follows.

For each polyhedron T of $\mathcal{T}$, o_0's payoff from T is $\gamma_T(\tilde{x}^\delta)$ and by construction, $\gamma_T(\tilde{x}^\delta) \leqslant \gamma(\tilde{x}^\delta)$ with the inequality being strict iff $\tilde{x}^\delta$ does not belong to T, which proves (3).

It remains to prove (4). The actual strategy implemented by n is $q^{0,\delta}_n$, which belongs to P^0_n. $o^0_{n,2}$'s payoff function is such that his best replies to $q^{0,\delta}_n$ are the points in $S^{0,\delta}_n$ that are closest to $q^{0,\delta}_n$. Thus $\tilde{q}^{0,\delta}_n$ is within δ of $q^{0,\delta}_n$. Since $q^{0,\delta}_n$ belongs to P^0_n, all of $o^1_{n,2}$'s strategies yield a payoff of zero against $q^{0,\delta}_n$. However, since the behavioral strategy $\tilde{b}^\delta$ is a quasi-perfect equilibrium, there exists a sequence of completely mixed behavioral strategies $\tilde{b}^{\varepsilon,\delta}$ converging to $\tilde{b}^\delta$ against which $o^1_{n,2}$'s equilibrium strategy $\tilde{\sigma}^\delta_{n,2}$ is a best reply. Under the sequence $\tilde{b}^{\varepsilon,\delta}$, there is a positive probability that player n rejects the strategies in $S^0_n \cup T^0_n(\delta)$ and makes a choice among strategies in $S^1_n \cup T^1_n(\delta)$. The fact that $\tilde{\sigma}^\delta_{n,2}$ is optimal against the sequence implies that it is optimal against the limiting choice $\check{q}^\delta_n$ there. Since $\bar{\pi}^1_n(\check{q}^\delta_n) = \pi^{1,\delta}_n$, $o^1_{n,2}$'s best replies are within δ of $\pi^{1,\delta}_n$ and thus $\tilde{\pi}^{1,\delta}_n$ is also within δ of $\pi^{1,\delta}_n$. □

5.8. **Limits of the Quasi-Perfect Equilibria of the Metagames.** In this subsection we derive the limits of the metagames' quasi-perfect equilibria as $\delta \downarrow 0$.

Consider a sequence of δ'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 and let $(\tilde{q}^{0,\delta}, (p^{0,\delta}, p^{1,\delta}), \tilde{\pi}^{1,\delta})$ be the sequence in $P^0 \times \mathbb{P} \times \Pi^1$ induced by the outsiders' strategies.

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

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 δ, $(\tilde{q}^{0,\delta}, (p^{0,\delta}, p^{1,\delta}), \tilde{\pi}^{1,\delta})$ 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,

$x^0 \equiv (q^{0,0}, (p^{0,0}, p^{1,0}), \pi^{1,0})$ belongs to $\bar{L}$ and its image under $\tilde{g}$ belongs to $\bar{K}$; also $(p^{0,0}, p^{1,0})$ belongs to $\bar{K}$.

If we can show that $x^0 \in \mathcal{Q}$, then, by construction, Ψ is essential and $x^0 \in V(x^*)$; therefore $q^{0,0} \in U(q^{0,*})$ and Q^* is stable, which proves the theorem. To show that x^0 belongs to $\mathcal{Q}$, it suffices to prove that $x^\delta \equiv (q^{0,0}, (p^{0,\delta}, p^{1,0}), \pi^{1,0})$ belongs to $\mathcal{Q}$ for all small δ, since $\mathcal{Q}$ is closed. The remainder of the proof establishes this property.

5.9. **Insiders' Strategies in a Quasi-Perfect Equilibrium.** Next we derive the important features of the insiders' strategies in quasi-perfect equilibria of the metagames, and their limits as $\delta \downarrow 0$.

Let $\tilde{b}^{\varepsilon,\delta}$ be a sequence of completely mixed behavioral strategies converging to $\tilde{b}^\delta$ against which for each insider n and each information set of n in $\tilde{\Gamma}^\delta$, his continuation strategy as given by b^δ is optimal. 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 $T_n^0(\delta)$ 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 moving on to a third stage and then choosing a strategy in $T_n^1(\delta) \cup S_n^1$ 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 $S_n^1 \cup T_n^1(\delta)$.

Let $\tilde{\sigma}^{\varepsilon,\delta}$ be a sequence of mixed-strategy profiles in $\tilde{\Gamma}^\delta$ that is equivalent to the sequence $\tilde{b}^{\varepsilon,\delta}$ of behavioral-strategy profiles. For each player n, his strategy $\tilde{\sigma}_n^{\varepsilon,\delta}$ 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 δ 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 $\bar{T}_n^i(\delta)$ be the union over all $w_n^i \in W_n^i$ of the sets $T_n^i(\delta, p_n^i(w_n^i))$. 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 $\bar{S}_n$.

The sequence $\tilde{\sigma}^{\varepsilon,\delta}$ induces a mixed strategy $\bar{\sigma}_n^{\varepsilon,\delta}$ in $\bar{S}_n$ 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 $\tilde{\sigma}_n^{\varepsilon,\delta}(s_n)$; for each $i = 0, 1$, $w_n^i \in W_n^i$ and $t_n \in S_n^0$,

the probability $\bar{\sigma}_n^{\varepsilon,\delta}(t_n^i(\delta, p_n^i(w_n^i)))$ is $\tilde{\sigma}_n^{\varepsilon,\delta}(t_n^i(\delta))\tilde{\sigma}_{o_{n,1}^i}^{\varepsilon,\delta}(w_n^i)$. From player m's perspective it is the sequence $\bar{\sigma}_n^{\varepsilon,\delta}$, or rather its equivalent sequence in P_n, that matters for his choice.

5.10. **The Induced Lexicographic Probability System.** The next step uses these sequences to obtain a representation of the insiders' strategies as a lexicographic probability system.

By Blume, Brandenburger, and Dekel [2, Appendix Proposition 2], we can construct for each player n a lexicographic probability system (LPS) $\bar{\Lambda}_n^\delta = (\bar{\sigma}_n^{0,\delta}, \ldots, \bar{\sigma}_n^{l_n(\delta),\delta})$ over his strategies in $\bar{S}_n$ such that for each ε in a subsequence converging to zero,

$$\bar{\sigma}_n^{\varepsilon,\delta} = (1-\nu_0(\varepsilon))(\bar{\sigma}_n^{0,\delta} + \nu_0(e)((1-\nu_1(\varepsilon))\bar{\sigma}_n^{1,\delta} + \nu_1(\varepsilon)((1-\nu_2(\varepsilon))\bar{\sigma}_n^{1,\delta} + \cdots + \nu_{l_n(\delta)-1}(\varepsilon)\bar{\sigma}_n^{l_n(\delta),\delta}))\,,$$

where $(\nu_0(\varepsilon), \ldots, \nu_{l_n(\delta)-1}(\varepsilon))$ is a sequence in $\mathbb{R}_{++}^{l_n(\delta)}$ converging to the origin. Moreover, $l_n(\delta)$ depends only on the cardinality of $\bar{S}_n(\delta)$, which is independent of δ. Let $l_n^{0,\delta}$ be the first level in $\bar{\Lambda}_n^\delta$ at which some strategy in $\bar{T}_n^0(\delta)$ appears with positive probability. Let $l_n^{1,\delta}$ be the first level of $\bar{\Lambda}_n^\delta$ at which some strategy in $S_n^1 \cup \bar{T}_n^1(\delta)$ appears with positive probability.

From the LPS $\bar{\Lambda}_n^\delta$ construct an LPS $\Lambda^\delta = (q_n^{0,\delta}, \ldots, q_n^{l_n,\delta})$ for Γ where for each l, $q_n^{l,\delta}$ is an enabling strategy in Γ that is equivalent to $\bar{\sigma}_n^{l,\delta}$. $q_n^{0,\delta}$ is a lexicographic best reply against Λ_m^δ. If we let $\bar{\lambda}_n^{i,\delta}$ be the total probability of the strategies in $\bar{T}_n^i(\delta)$ under $\bar{\sigma}_n^{i,\delta}$, then for each $w_n^i \in W_n^i$, the total probability under $\bar{\sigma}_n^{l_n^{i,\delta}}$ of the strategies in $T_n^i(\delta, p_n^i(w_n^i))$ is $\bar{\lambda}_n^{i,\delta}\tilde{\sigma}_{o_{n,1}^i}^\delta(w_n^i)$, where $\sigma_{o_{n,1}^i}^\delta(w_n^i)$ is the probability assigned to vertex w_n^i by outsider $o_{n,1}^i$'s equilibrium strategy $\sigma_{o_{n,1}^i}^\delta$. Since $p_n^i(\tilde{\sigma}_{o_{n,1}^i}^\delta)$ is, by definition, $p_n^{i,\delta}$, level $l_n^{i,\delta}$ is expressible as a convex combination $\lambda_n^{l_n^i,\delta} p_n^{i,\delta} + (1-\lambda_n^{l_n^i,\delta}) r_n^{l_n^i,\delta}$, with $\lambda_n^{l_n^i,\delta} = \bar{\lambda}_n^{i,\delta}\delta$ and $r_n^{l_n^i,\delta} \in P_n$; moreover, $\lambda_n^{0,\delta} > 0$ since by definition $l_n^{0,\delta}$ is the first level l where a strategy in $\bar{T}_n^0(\delta)$ appears in in the support of the $\bar{\sigma}_n^{l,\delta}$. Also, since $l_n^{1,\delta}$ is the first level of Λ^δ that does not have its support in P_n^0, $\bar{\pi}_n^1(q_n^{l_n^1,\delta})$ equals $\pi_n^{1,\delta}$.

We claim that for each $l \leqslant l_n^{0,\delta}$, $q_n^{l,\delta}$ induces the equilibrium outcome against $q_m^{0,\delta}$. Indeed, since $q_m^{0,\delta}$ belongs to Q^*, the strategies in $S_n^0 \cup T_n^0(\delta)$ are optimal against $q_m^{0,\delta}$; also, by quasi-perfection, every strategy s_n^1 in S_n^1, (resp. every strategy $t_n^1(\delta)$ in $T_n^1(\delta)$) such that s_n^1 (resp. $t_n^1(\delta, w_n^1)$ for some w_n^1) appears at a level $l \leqslant l_n^{0,\delta}$ of $\bar{\Lambda}_n^\delta$ must be a best reply to $q_m^{0,\delta}$, since n has to reject the strategies in $T_n^0(\delta)$ before choosing a strategy in $S_n^1 \cup T_n^1(\delta)$. Thus for all $l \leqslant l_n^{0,\delta}$, $q_n^{l,\delta}$ is a best reply to $q_m^{0,\delta}$. Since $q_m^{0,\delta}$ is a lexicographic best reply to Λ_n^δ, $(\tilde{q}_n(\varepsilon), q_m^{0,\delta})$ is an equilibrium of the game Γ for all small ε, where

$$\tilde{q}_n(\varepsilon) = (1-\varepsilon)\; q_n^{0,\delta} + \varepsilon((1-\varepsilon)q_n^{1,\delta} + \varepsilon^2((1-\varepsilon)q_n^{2,\delta} + \varepsilon^3(\cdots + \varepsilon^{l_n^{0,\delta}} q_n^{l_n^{0,\delta}})))\,.$$

By genericity of payoffs, Γ 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) $l_n^{1,\delta} > l_n^{0,\delta}$; (ii) the enabling strategy $r_n^{l_n^0,\delta}$ 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^*.

5.11. **Limit of the Lexicographic Probability System.** Next we characterize the limit of the LPS as $\delta \downarrow 0$.

Take a subsequence of the δ's such that the following properties hold for the associated LPSs Λ_n^δ for each n: (i) $l_n^{i,\delta}$ is independent of δ for each i, call it l_n^i; (ii) for each $l \leqslant l_n^1$, the face of P_n that contains $q_n^{l,\delta}$ in its interior, as well as the strategies for m in S_m that are best replies to $q_n^{l,\delta}$, are independent of δ.

Let $\sigma_n^{l_n^1,\delta}$ be a sequence of strategies in Σ_n that is equivalent to the sequence $q_n^{l_n^1,\delta}$. Again using Blume, Brandenburger, and Dekel [2, Appendix Proposition 2], there is now an LPS $(\sigma_n^{l_n^1,0},\ldots,\sigma_n^{l_n^1,k_n})$ 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 δ's, $\sigma_n^{l_n^1,\delta}$ is expressible as the nested combination

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

This LPS induces an equivalent LPS $(q_n^{l_n^1,0},\ldots,q_n^{l_n^1,k_n})$ in enabling strategies.

Let k_n^1 be the first level k of this LPS where $q_n^{l_n^1,k}$ does not belong to P_n^0. Recall from the previous subsection that $q_n^{l_n^1,\delta}$ is expressible as a convex combination $\lambda_n^{l_n^1,\delta}p_n^{1,\delta}+(1-\lambda_n^{l_n^1,\delta})r_n^{l_n^1,\delta}$ and that $\bar{\pi}_n^1(q_n^{l_n^1,\delta}) = \pi_n^{1,\delta}$. Express $r_n^{l_n^1,\delta}$ as a convex combination $\alpha_n^{0,\delta}r_n^{0,\delta}+\alpha_n^{1,\delta}r_n^{1,\delta}$, where for $i=0,1$, $r_n^{i,\delta} \in P_n^i$. Then $\lambda_n^{l_n^1,\delta}+(1-\lambda_n^{l_n^1,\delta})\alpha_n^{1,\delta} > 0$ since $q_n^{l_n^1,\delta}$ does not belong to P_n^0. Going to a subsequence, let λ_n^1 be the limit of $\lambda_n^{l_n^1,\delta}/(\lambda_n^{l_n^1,\delta}+(1-\lambda_n^{l_n^1,\delta})\alpha_n^{1,\delta})$ and let $r_n^{1,0}$ be the limit of $r_n^{1,\delta}$. Since the limit of $p_n^{1,\delta}$ is $p_n^{1,0}$, we have that $q_n^{l_n^1,k_n^1}$ is expressible as a convex combination $\zeta_n(\lambda_n^1 p_n^{1,0}+(1-\lambda_n^1)r_n^{1,0})+(1-\zeta_n)\check{p}_n^0$ for some $\zeta_n > 0$ and $\check{p}_n^0 \in P_n^0$. Moreover, since $\pi_n^{1,0}$ is the limit of $\pi_n^{1,\delta}$, $\bar{\pi}_n^1(q_n^{l_n^1,k_n^1}) = \pi_n^{1,0}$.

For each δ in the subsequence used above, define now an LPS $\hat{\Lambda}_n^\delta = (\hat{q}_n^{0,\delta},\ldots,\hat{q}_n^{l_n^1+k_n^1+1,\delta})$ as follows: $\hat{q}_n^{0,\delta} = q_n^{0,0}$, $\hat{q}_n^{l,\delta} = q_n^{l-1,\delta}$ if $0 < l \leqslant l_n^1$, and $\hat{q}_n^{l,\delta} = q_n^{l_n^1,l-l_n^1-1}$ otherwise. The strategy $\hat{q}_n^{l,\delta}$ is independent of δ for $l=0$ and $l > l_n^1$. Each level $l < l_n^1+k_n^1+1$ is a strategy in P_n^0. $\hat{q}_n^{l_n^0+1,\delta}$ is a convex combination $\lambda_n^{l_n^0,\delta}p_n^{0,\delta}+(1-\lambda_n^{l_n^0,\delta})r_n^{l_n^0,\delta}$ while $\hat{q}_n^{l_n^1+k_n^1+1,\delta}$ is a convex combination $\zeta_n(\lambda_n^1 p_n^{1,0}+(1-\lambda_n^1)r_n^{1,0})+(1-\zeta_n)\check{p}_n^0$ where $\zeta_n > 0$ and $\bar{\pi}_n^1(\hat{q}_n^{l_n^1+k_n^1,\delta}) = \pi_n^{1,0}$. The next lemma sets out the key properties of $\hat{\Lambda}_n^\delta$ that lead to a conclusion of our proof.

Lemma 5.4. *For all small δ, the LPS $\hat{\Lambda}_n^\delta$ satisfies the following properties.*

(1) *A strategy is at least as good a reply against Λ_n^δ as 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* $q_m^{0,0}$ *and is at least as good a reply lexicographically against* $\hat{\Lambda}_m^\delta$ *as every strategy in* S_n^1.
(3) $\lambda_n^1 > 0$ *and every level* $l < l_n^1 + k_n^1 + 1$ *induces the equilibrium outcome against* $q_m^{0,0}$.
(4) *The strategy* $\hat{q}_n^{l,\delta}$ *for* $l < l_n^0$ *and the strategy* $r_n^{l_n^0,\delta}$ *if* $\lambda_n^{l_n^0,\delta} < 1$ *are lexicographic best replies against* $\hat{q}_m^\delta$.

Proof of Lemma. Suppose s_m is a better reply against $\hat{\Lambda}_n^\delta$ than another strategy t_m. We show that s_m is also a better reply against Λ_n^δ. Let l be the first level of $\hat{\Lambda}_n^\delta$ such that s_m is a better reply against $\hat{q}_m^{l,\delta}$ than t_m. If $l = 0$ then for all small δ, s_m is a better reply against $q_n^{0,\delta}$ since $\hat{q}_n^{0,\delta}$ equals the limit $q_n^{0,0}$ of $q_n^{0,\delta}$; thus against Λ_n^δ, t_m is a worse reply against the very first level. If $0 < l < l_n^1 + 1$ then obviously s_m is better reply against Λ_n^δ than t_m since level l of $\bar{\Lambda}_n^\delta$ corresponds to level $l-1$ of Λ_n^δ. Suppose then that $l_n^1 + 1 \leqslant l \leqslant l_n^1 + k_n^1 + 1$. Then for all small δ, s_m is a better reply against $q_n^{l_n^1,\delta}$ since $q_n^{l_n^1,\delta}$ is a nested combination of $(q_n^{l_n^1,0}, \ldots, q_n^{l_n^1,k_n})$. Thus, s_m is a better reply against Λ_n^δ. This proves (1).

In the game $\tilde{\Gamma}^\delta$ player n, when finally making a choice among the strategies in $S_n^1 \cup T_n^1(\delta)$, would choose a strategy s_n in S_n^1 with positive probability only if it is at least as good a reply against Λ_n^δ as the other strategies in S_n^1. Therefore, such a strategy would show up with positive probability under level l_n^1 of Λ_n^δ (and hence in $\hat{\Lambda}_n^\delta$) only if this is the case. This implies that if $(1 - \mu_n^{l_n^1,\delta})\alpha_n^1$ is positive, then the strategy $r_n^{1,\delta}$ is at least as good a reply against Λ_n^δ as the strategies in S_n^1. Recall that $r_n^{1,0}$ is the limit of $r_n^{1,\delta}$ and λ_n^1 is the limit of $\mu_n^{l_n^1,\delta}/(\mu_n^{l_n^1,\delta} + (1 - \mu^{l_n^1,\delta})\alpha_n^{1,\delta})$. Therefore, by point (1) of this lemma, $r_n^{1,0}$ is at least as good a reply against $\hat{\Lambda}_n^\delta$ as strategies in S_n^1 if $\lambda_n^1 < 1$. To show that it is also a best reply against $q_m^{0,0}$, suppose to the contrary that it is not. Then every strategy in $T_n^1(\delta)$, regardless of outsider $o_{n,1}^1$'s choice, is a better reply against $q_m^{0,\delta}$ than $r_n^{1,\delta}$ for all small δ, since for $\delta = 0$ the strategies in $T_n^1(\delta)$ are in P_n^0, which are best replies to $q_m^{0,0}$. Therefore, quasi-perfection implies that n would prefer to play the strategies in $T_n^1(\delta)$ rather than implementing $r_n^{1,0}$ (or $r_n^{1,\delta}$ when δ is small), which shows that $(1 - \mu^{l_n^1,\delta})\alpha_n^{l,\delta} = 0$ for all small δ and hence that $\lambda_n^1 = 1$. This proves (2).

We turn now to (3). Every strategy $\hat{q}_n^{l,\delta}$ for $l < l_n^1 + k_n^1 + 1$ belongs to P_n^0 and is thus optimal against $q_m^{0,0}$, which belongs to Q_m^*. The strategy $\check{p}_n^0$ (which recall is part of the expression defining $q_n^{l_n^1,k_n^1}$) which is chosen with positive probability is also in P_n^0 and hence optimal against $q_m^{0,0}$. As we saw in the previous paragraph, if $\lambda_n^1 < 1$, the strategy r_n^1 must also be optimal against $q_m^{0,0}$. Obviously $q_m^{0,0}$ is optimal against $\hat{\Lambda}_n^\delta$ since it is the limit of $q_m^{0,\delta}$, which by point (1) is optimal against $\hat{\Lambda}_n^\delta$. Therefore, for all small ε, $(q_m^{0,0}, q_n(\varepsilon))$ is an equilibrium

of Γ, where

$$\tilde{q}_n(\varepsilon) = \left(\sum_{l=0}^{l_n^1+k_n^1} \varepsilon^l \hat{q}_n^{l,\delta} + (1 - 1_{\lambda_n^1>0})\varepsilon^{l_n^1+k_n^1+1} \hat{q}_n^{l_n^1+k_n^1+1} \right) \Big/ \left(\sum_{l=0}^{l_n^1+k_n^1} \varepsilon^l + (1 - 1_{\lambda_n^1>0})\varepsilon^{l_n^1+k_n^1+1} \right),$$

where $1_{\lambda_n^1>0}$ is the indicator function. This is impossible if $\lambda_n^1 = 0$, since $\hat{q}_n^{l_n^1+k_n^1+1}$ is a convex combination of a strategy in P_n^0 and one in P_n^1. Thus $\lambda_n^1 > 0$. Moreover, since this is a continuum of equilibria, genericity implies that all of them induce the same outcome. Therefore, all strategies at levels preceding $l_n^1+k_n^1+1$ induce the equilibrium outcome against $q_m^{0,0}$.

Lastly we prove (4). By the previous paragraph, all strategies in P_n^0 are optimal against $\hat{q}_m^{k,\delta}$ for $k \leqslant l_m^1 + k_m^1$. Thus the optimality of a strategy in P_n^0 depends on how it fares against $\hat{q}_m^{l_m^1+k_m^1+1}$. Obviously every strategy t_n in the support of $q_n^{0,0}$ is optimal against Λ_m^δ for all small δ and is therefore optimal against $\hat{q}_m^{l_m^1+k_m^1+1}$. Now, if a strategy $s_n \in S_n^0$ is not optimal against $\hat{q}_m^{l_m^1+k_m^1+1}$, then for all small δ, the strategy $t_n^0(\delta)$ for some t_n^0 in the support of $q_n^{0,0}$ is a superior reply to $\hat{q}_m^{l_m^1+k_m^1+1}$ regardless of what outsider $o_{n,1}^0$'s choice is. Therefore, in $\tilde{\Gamma}^\delta$, when player n, after making a provisional choice of s_n, reconsiders his decision, he would prefer to play $t_n^0(\delta)$ rather than s_n; moreover, at every information set where he is choosing among the strategies in $T_n^0(\delta)$, he would prefer to play $t_n^0(\delta)$ rather than the duplicate $s_n^0(\delta)$ involving s_n. Hence, the probability of s_n is zero for all $l < l_n^0$ in Λ_n^δ and its probability under $r_n^{l_n^0,\delta}$ is zero as well. Point (4) now follows. □

5.12. **Final Step of the Proof.** The proof can now be completed by invoking the results established above.

Fix a small δ that satisfies the properties enumerated in Lemma 5.4 of the previous subsection. Let

$$\bar{q}_n^0(\varepsilon) = \left(\sum_{l \leqslant l_n^0+1} \varepsilon^l \hat{q}_n^{l,\delta} \right) \Big/ \left(\sum_{l \leqslant l_n^0+1} \varepsilon^l \right), \qquad \bar{q}_n^1(\varepsilon) = \left(\sum_{l=l_n^0+2}^{l_n^1+k_n^1+1} \varepsilon^l \hat{q}_n^{l,\delta} \right) \Big/ \left(\sum_{l=l_n^0+2}^{l_n^1+k_n^1+1} \varepsilon^l \right).$$

Observe that $\bar{q}_n^0(\varepsilon)$ belongs to P_n^0 for all ε and it is a convex combination of $p_n^{0,\delta}$ and a subset R_n^0 of strategies that are at least as good replies against $\hat{\Lambda}_m^\delta$ as other strategies in S_n^0. Likewise $\bar{q}_n^1(\varepsilon)$ is a convex combination of $p_n^{1,0}$, r_n^1 and a point in P_n^0 such that the strategy r_n^1 if it has a positive weight is at least as good a reply against $\hat{\Lambda}_m^\delta$ as other strategies in P_n^1. There exists a small ε such that a strategy s_m is at least as good as a strategy t_m against $\hat{\Lambda}_n^\delta$ iff it is lexicographically at least as good a reply against the LPS $(q^{0,0}, \bar{q}_n^0(\varepsilon), \bar{q}_n^1(\varepsilon))$. Since $\bar{\pi}_n^1(\bar{q}_n^1(\varepsilon)) = \bar{\pi}_n^1(q_n^{1,0}) = \pi^{1,0}$ for all ε, it follows that $(q^{0,0}, (p^{0,\delta}, p^{1,0}), \pi^{1,0})$ belongs to $\mathcal{Q}$. As argued in Sections 5.2 and 5.3, proving that this point belongs to $\mathcal{Q}$ shows that in fact it belongs to $V(x^*)$ and hence that $q^{0,0}$ in $U(q^{0,*})$.

This completes the proof of Theorem 5.1 when S_n^1 is not empty for either player n. In case S_n^1 is empty for exactly one player n, as we said initially in the description of $\mathbb{P}$ and $\mathcal{Q}$, we do not have the factor P_n^1 or Π_n^1. In the family of games $\Gamma(\delta, \hat{p})$, player n decides provisionally in the first stage on the strategy in S_n to play and in the second stage gets to execute it or switch to playing a strategy in $T_n(\delta, \hat{p})$. In the metagame, we do not have outsiders $o_{n,j}^1$ for $j = 1, 2, 3$. The rest of the proof is essentially the same modulo these provisions.

6. Discussion

The characterizations of stable sets in Theorems 5.1 and 5.2 apply only to a two-player game for which payoffs are generic in an extensive form with perfect recall. For any game, Axioms A, B, C are implied by stability, but they do not suffice to characterize stability for games with more players and/or nongeneric payoffs. Our contribution here is limited to identifying conditions minimally sufficient to characterize a prominent refinement for one rich class of games. We hope nevertheless that Theorem 5.1 is a useful first step toward a theory of equilibrium refinement using axioms adapted from decision theory. Theorem 5.2 has technical interest because it shows for the assumed class of games that stability with respect to perturbed strategies is equivalent to an analogous stability with respect to lexicographically optimal replies to deviations from equilibria of the given game.

We conclude by discussing embedding, the axioms, and our assumptions.

6.1. **Embedding.** First we address a referee's concern that "because of correlations, analysis of a game is not invariant to its context." An influential example is by Ely and Peski [6]; see Liu [28] and Sadzik [46] for further analysis. Suppose it is common knowledge that the game between players 1 and 2 is one of the two games L and R shown below in normal form, and these two games are equally likely.

L:

1 | 2:	d	e	f
a	1,1	-10,-10	-10,0
b	-10,-10	1,1	-10,0
c	0,-10	0,-10	0,0

R:

1 | 2:	d	e	f
a	-10,-10	1,1	-10,0
b	1,1	-10,-10	-10,0
c	0,-10	0,-10	0,0

Consider two versions. In version I, each player simply maximizes his expected payoff. Then 1 and 2 choose c and f in the unique Nash equilibrium, and their payoffs are (0,0). In version II, suppose that 1 and 2 observe private signals $s \in \{A, B\}$ and $t \in \{D, E\}$, respectively, about which game is played, for which the joint distribution is shown below.

L:

s | t:	D	E
A	1/4	0
B	0	1/4

R:

s | t:	D	E
A	0	1/4
B	1/4	0

In this version it remains common knowledge that the games are equally likely (and for each signal realization, each player's hierarchy of beliefs about the other's beliefs is the same

as before). But now there are two other Nash equilibria and these achieve payoffs (1,1); implicitly, the players coordinate perfectly. In one, player 1 chooses a or b as his signal A or B, and 2 chooses d or e as his signal is D or E.

The referee interprets the difference between versions I and II as possibly due to contextual features. For instance, the use of version II's more detailed elaboration of the common knowledge that the two games are equally likely could stem from exogenous factors that determine whether the signals are observed or deemed relevant — and this might impair our 'practical interpretation' in Section 1 that Axiom C excludes presentation effects. Even so, quite apart from our proposed interpretation, the difference between the two versions is immaterial to our results. Proposition 3.3 implies that version II is not a metagame for version I since their Nash equilibria differ.

To appreciate further that embedding does not introduce correlation, consider a correlated equilibrium of a game G. Suppose that Nature (or an outsider) uses a probability distribution P to choose a vector $t \in T = \prod_{n\in N} T_n$ of signals and then to each player n reports only t_n. In the induced larger game, $\tilde{\Sigma} = \prod_{n\in N} \Sigma_n^{T_n}$, and $\tilde{G}_n(\tilde{\sigma}) = \sum_{t\in T} G_n((\tilde{\sigma}_m(t_m))_{m\in N})P(t)$. In the special case that P is a correlated equilibrium as in Aumann [1], each $T_n = S_n$, and n receives a signal s_n only if his pure strategy s_n is an optimal reply. Assume therefore that each S_n is restricted to those pure strategies in the support of P. Then the maps $f_n(\tilde{\sigma}_n) = \sum_{t\in T} \tilde{\sigma}_n(t_n)P(t)$ to marginal distributions are surjective and thus satisfy condition (a) in Definition 3.1 of an embedding, but Proposition 3.3 implies that condition (b) is also satisfied only if the Nash equilibria of $\tilde{G}$ map to the Nash equilibria of G. Hence $\tilde{G}$ qualifies as a metagame only if the marginal distributions derived from P form a Nash equilibrium of G; i.e. the 'correlation' is nil.

6.2. **Alternative Axioms.** Next we mention alternative versions of the axioms.

Axiom A. An alternative to Axiom A requires instead that solutions are minimal sets satisfying Axioms B and C. It can be shown that Axioms B and C and minimality imply that solutions are connected sets of admissible equilibria, hence stable.

Axiom B. In [17, 16] we show that a sequential equilibrium suffices in Axiom B for perfect-information and signaling games. We conjecture that Axiom B can be weakened here to require only that each solution contains a sequential equilibrium. This suffices in our article [18] on forward induction, but we foresee a much longer proof here because the metagames constructed in the present proof have nongeneric payoffs and therefore their sequential equilibria need not be quasi-perfect.

Exclusion of non-credible threats is an oft-cited justification for selecting sequential equilibria; and to exclude 'threatening with beliefs,' for selecting those whose outcomes satisfy forward induction. Here, the assumed genericity of payoffs implies that all equilibria in a

solution have the same outcome, so Axiom B serves to exclude outcomes that depend on non-credible threats, and in combination with Axiom C it requires outcomes to satisfy forward induction [18].

Nevertheless, we agree with a referee that Axiom B will not be appropriate for a genuine axiomatic development because it invokes one refinement (quasi-perfection) to derive another (stability). We hope that eventually sequential rationality will be derived from axioms more primitive than the lexicographic optimality described in Section 3.2. Of equal interest would be axioms that imply backward induction or dynamic programming algorithms; cf. Kohlberg and Mertens [24, Section 2.6] and Hillas and Kohlberg [22, Section 10]. Brandenburger and Friedenberg [4, p. 6] define backward induction via a property requiring (loosely) that "the solution on the whole tree should be included in the solution on what is left after replacing a subtree with what the solution allows on the subtree." They show that sequential equilibrium satisfies this property, but quasi-perfect equilibrium need not for some games with nongeneric payoffs because admissibility can be violated 'on what is left.' This suggests that for games with nongeneric payoffs it will be desirable to use sequential equilibrium in Axiom B.

Axiom C. Axiom C requires that solutions of metagames map to solutions of the embedded game, and also that, for each solution and each metagame, there exists a solution of the metagame whose image is the given solution of the embedded game. The second part can be weakened to require only that the image is contained in the given solution. This weaker version of Axiom C is satisfied by the weaker version of stability called metastability [15], which differs from stability chiefly in implying only a relaxed form of the decomposition property described in Section 1.1. In fact, our proof of Theorem 5.1 uses only the second part of this weaker version of Axiom C, but it suffices because metastability and stability coincide for games with generic payoffs in an extensive-form with perfect recall.

A stronger version of Axiom C recasts it as preservation of best-reply correspondences, rather than feasible strategies and payoffs.

6.3. **The Restriction to Two Players.** Our assumption that there are only two players is convenient in Theorem 5.2 because with generic payoffs it enables a simpler characterization of stability than Mertens' general definition. But it is crucial to our proof of Theorem 5.1. This can be seen in the game Beer-Quiche in Section 3.4 when one considers the two agents of player 1 to be distinct players. For the resulting three-player game, the present line of proof, applied to the inessential component of equilibria in which both agents choose Q, fails to exclude player 2 from assigning low probability to agent S after observing the deviation B. For Beer-Quiche with three players, this deficiency can be remedied by applying Axiom C to a larger class of embeddings in which one requires only that insiders' best-reply correspondences are preserved by the mapping from a metagame. We conjecture that Theorem 5.2 can be generalized to many players, and then a similar proof of Theorem

5.1 will be possible using a stronger version of Axiom C that requires solutions to inherit further invariance properties of Nash equilibria.

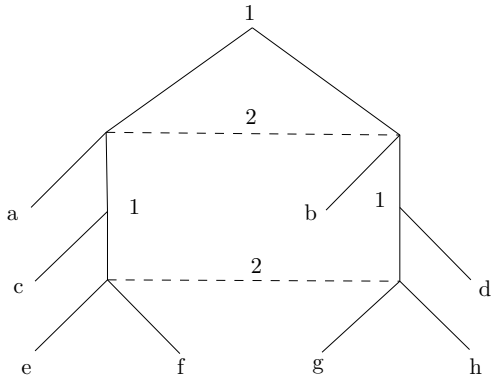
6.4. **The Restriction to Generic Payoffs.** A referee rightly criticizes our reliance on generic payoffs, because several refinements were motivated by deficiencies revealed by examples with nongeneric payoffs. We acknowledge that our assumption of generic payoffs limits the significance of Theorem 5.2, and therefore also Theorem 5.1 whose present proof relies on it. But an axiomatic theory of refinement can proceed with progressively more general assumptions. We hope it is useful as an initial step to establish a minimal set of axioms defined for all games with perfect recall that identify stability as the implied refinement for two-player games with extensive-form payoffs outside a lower-dimensional set.

It is true that some solutions of games with nongeneric payoffs are limits of solutions of nearby games with generic payoffs, but this expedient does not identify all solutions. Further work is required to find axioms that characterize all solutions of general games.

6.5. **Conclusion.** We view Axiom C as a main contribution of this article. It suggests that requiring refinements to inherit invariance properties of Nash equilibria could enable axiomatic characterizations for games more general than the restrictive class considered here. Although a stronger version might be needed for general games, Axiom C adheres to Kohlberg and Mertens' dictum that a refinement should be justified by rational behavior in the given game, to which they add and we endorse, and in those games to which it is equivalent due to the invariance properties invoked. It will be valuable to identify other or stronger invariance properties than the exclusion of presentation effects embodied in Axiom C. And, Axiom C redirects studies of refinements from properties inherited from perturbed games to properties preserved in larger games that embed the given game.

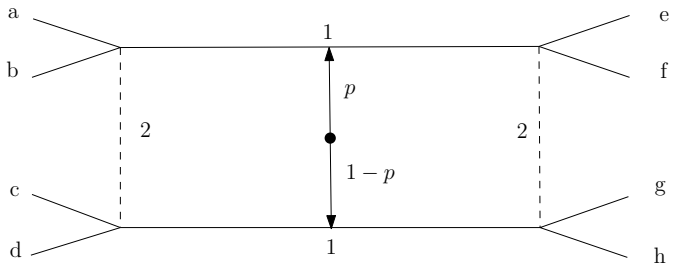
Appendix A. Enabling Strategies

In the normal-form representation of a game in extensive form, a player's pure strategy specifies the actions chosen at 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 3 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 [30]. 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



1 — 2:	ab	cdeg	cdfh
ac	a	c	c
aef	a	e	f
bd	b	d	d
bgh	b	g	h

FIGURE 3. A game tree and its pure reduced normal form in which each pure strategy is identified by the terminal nodes it does not exclude.



1—2:	aceg	acfh	bdeg	bdfh
abcd	ac	ac	bd	bd
abgh	ag	ah	bg	bh
efcd	ec	fc	ed	fd
efgh	eg	fh	eg	fh

FIGURE 4. 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.

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 4 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. For instance, the outcome of 1's strategy *abcd* and 2's strategy *aceg* is a with probability p and c with probability $1 - p$.

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 3 and 4. Enabling strategies are minimal representations of strategic behavior in games with perfect recall.[8]

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 $\mathcal{G}_n(p) = \sum_z \gamma_z(p) u_n(z)$. The game Γ is therefore summarized by the multilinear function $\mathcal{G} : P \to \mathbb{R}^N$ that assigns to each profile of players' enabling strategies their expected payoffs. This summary specification is called the *enabling form* of the game.

Appendix B. Proofs of Propositions

B.1. **Proof of Proposition 3.2.** A multilinear map $f_n : \tilde{\Sigma}_n \times \tilde{\Sigma}_o \to \Sigma_n$ 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 $\tilde{S}_n \times \tilde{S}_o$ of profiles of pure strategies.

Proposition 3.2. $\tilde{G}$ 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 $\pi_n : \tilde{T}_n \to S_n$ such that for each $(\tilde{s}, s_o) \in \tilde{S} \times \tilde{S}_o$ and $\tilde{t}_n \in \tilde{T}_n$:

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

Proof. Suppose we have a game $\tilde{G} : \tilde{\Sigma} \times \tilde{\Sigma}_o \to \mathbb{R}^{N \cup O}$ and a collection of multilinear maps $f_n : \tilde{\Sigma}_n \times \tilde{\Sigma}_o \to \Sigma_n$, 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 σ_o, $f_n(\cdot, \sigma_o)$ is surjective because it maps the face spanned by $\tilde{T}_n$ homeomorphically onto Σ_n. Also, condition (2) and multilinearity of each f_n imply that $\tilde{G} = G \circ f$. According to Definition 3.1, therefore, $(\tilde{G}, f)$ embeds G.

Now suppose that $(\tilde{G}, f)$ embeds G. Let σ_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 Σ_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 $s_o \in \tilde{S}_o$. 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

[8] Mertens [34, p. 554] introduces the mapping of mixed strategies to induced distributions on terminal nodes. Koller and Megiddo [25] call them realization plans.

each s_o, $\sigma_o(s_o)$ is the probability of s_o under σ_o. Therefore, since σ_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 π_n be the associated bijection. Define $\hat{f}_n : \tilde{S}_n \times \tilde{S}_o \to \Sigma_n$ 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. □

B.2. **Proof of Proposition 3.3.**

Proposition 3.3. If $(\tilde{G}, f)$ embeds G then the Nash equilibria of G are the f-images of the Nash 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 3.1 an embedding. Using condition (b),

$$G_n(\tau_n, \sigma_{-n}) = G_n(f(\tilde{\tau}_n, \tilde{\sigma}_{-n}, \sigma_o)) = \tilde{G}_n(\tilde{\tau}_n, \tilde{\sigma}_{-n}, \sigma_o) \leqslant \tilde{G}_n(\tilde{\sigma}, \sigma_o) = G_n(f(\tilde{\sigma}, \sigma_o)) = G_n(\sigma)\,,$$

where the inequality obtains because $(\tilde{\sigma}, \sigma_o)$ is an equilibrium of $\tilde{G}$. Hence σ is an equilibrium of G.

Conversely, suppose σ is an equilibrium of G. For each n, let π_n be the bijection given by Proposition 3.2. Let $\tilde{\sigma}_n$ be the strategy for insider n in $\tilde{G}$ defined by $\tilde{\sigma}_n(\tilde{t}_n) = \sigma_n(\pi_n(\tilde{t}_n))$ for $\tilde{t}_n \in \tilde{T}_n$ and $\tilde{\sigma}_n(\tilde{s}_n) = 0$ for $\tilde{s}_n \notin \tilde{T}_n$. Since f_n is multilinear, by condition (1) of Proposition 3.2, $f_n(\tilde{\sigma}_n, \cdot) = \sigma_n$ and thus $f(\tilde{\sigma}, \cdot) = \sigma$. Hence, it suffices to show that there exists a strategy profile σ_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 σ_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:

$$\tilde{G}_n(\tilde{s}_n, \tilde{\sigma}_{-n}, \sigma_o) = G_n(f_n(\tilde{s}_n, \sigma_o), \sigma_{-n}) \leqslant G_n(\sigma) = G_n(f(\tilde{\sigma}, \sigma_o)) = \tilde{G}_n(\tilde{\sigma}, \sigma_o)\,,$$

where the first and second equalities use the property $f(\tilde{\sigma}, \cdot) = \sigma$ established above, and the inequality obtains because σ is an equilibrium of G. □

Appendix C. Stability Satisfies Axiom C

As in Mertens [34], we consider versions of stability according to whether Čech cohomology has coefficients in $\mathbb{Z}_p$ for some prime p, or in the rationals $\mathbb{Z}_0 \equiv \mathbb{Q}$. Essentiality of maps here means essentiality with respect to Čech cohomology. Also, the concept of p-essentiality comes from Mertens [35]. The theorem below considers any finite game G.

Theorem C.1. *Axiom C is satisfied by p-stability for all p either prime or zero.*

Proof. Suppose that the game G is embedded in the game $\tilde{G}$ via the multilinear maps $f = (f_n)_{n\in N}$. As in Section 3.2, the strategy sets of the insiders $n \in N$ in G and $\tilde{G}$ are $\Sigma = \prod_n \Sigma_n$ and $\tilde{\Sigma} = \prod_n \tilde{\Sigma}_n$, respectively, and the strategy set of the outsiders in $\tilde{G}$ is $\tilde{\Sigma}_o$. Each map $f_n : \tilde{\Sigma}_n \times \tilde{\Sigma}_o \to \Sigma_n$ is surjective for each $\sigma_o \in \Sigma_o$, and $\tilde{G}_n = G_n \circ f$. We show that for each p, either prime or zero, the p-stable sets of G are exactly the projections of the p-stable sets of $\tilde{G}$.

We begin with some notation. For each $\varepsilon \in [0,1]$ let P_ε be the set $[0,\varepsilon] \times \Sigma$ with ∂P_ε its topological boundary. Each $(\varepsilon,\tau) \in P_1$ defines a perturbed game $G(\varepsilon,\tau)$ where the strategy sets are as in G but where the payoffs from a strategy profile σ are the payoffs from the perturbed profile $(1-\varepsilon)\sigma + \varepsilon\tau$ in G. Let $\mathcal{N} \subseteq P_1 \times \Sigma$ be the graph of the Nash equilibrium correspondence over P_1, i.e. $\mathcal{N}$ is the set of $(\varepsilon,\tau,\sigma)$ such that σ is an equilibrium of $G(\varepsilon,\tau)$. Let Ψ be the natural projection from $\mathcal{N}$ to P_1. For each ε and each subset X of $\mathcal{N}$, let $(X_\varepsilon, \partial X_\varepsilon) \equiv (\Psi^{-1}(P_\varepsilon) \cap X, \Psi^{-1}(\partial P_\varepsilon) \cap X)$. The sets $\tilde{P}_1$ and $\tilde{\mathcal{N}}$ and also the map $\tilde{\Psi}$ are likewise defined for the game $\tilde{G}$. Throughout, H refers to Čech cohomology with coefficients in $\mathbb{Z}_p$.

Let $\tilde{\Sigma}^*$ be a stable set of $\tilde{G}$. Then there exists a subset $\tilde{X}$ of $\tilde{\mathcal{N}}$ such that:

(1) $\tilde{\Sigma}^* = \{\, (\tilde{\sigma},\tilde{\sigma}_o) \mid (0,(\tilde{\tau},\tilde{\tau}_o),(\tilde{\sigma},\tilde{\sigma}_o)) \in \tilde{X} \text{ for some } (0,\tilde{\tau},\tilde{\tau}_o) \in \tilde{P}_1 \,\}$.
(2) For each neighborhood $\tilde{V}$ of $\tilde{X}_0$ in $\tilde{X}$ there exists a connected component of $\tilde{V} \setminus \partial\tilde{X}_1$ whose closure is a neighborhood of $\tilde{X}_0$.
(3) $\tilde{\Psi}^* : H^*(\tilde{P}_\varepsilon, \partial\tilde{P}_\varepsilon) \to H^*(\tilde{X}_\varepsilon, \partial\tilde{X}_\varepsilon)$ is nonzero for some $\varepsilon \in (0,1]$.

Define $h : \tilde{X} \to \mathcal{N}$ by $h(\varepsilon,(\tilde{\tau},\tilde{\tau}_o),(\tilde{\sigma},\tilde{\sigma}_o)) = (\varepsilon, f(\tilde{\tau},(1-\varepsilon)\tilde{\sigma}_o + \varepsilon\tilde{\tau}_o), f(\tilde{\sigma},(1-\varepsilon)\tilde{\sigma}_o + \varepsilon\tilde{\tau}_o))$. h is easily seen to be a well-defined map. Let X be the image of $\tilde{X}$ under h, and Σ^* the image of $\tilde{\Sigma}^*$ under f. Then X and Σ^* satisfy the corresponding versions of properties (1) and (2) above. Thus, to show that Σ^* is p-stable, it remains to show that $\Psi^* : H^*(P_\varepsilon, \partial P_\varepsilon) \to H^*(X_\varepsilon, \partial X_\varepsilon)$ is nonzero for the ε for which property (3) above holds for $\tilde{\Psi}^*$.

Since $\tilde{\Psi}$ is essential, it is p-essential. Also, letting $\tilde{\tau}_o^\circ$ be the profile of uniform mixtures for the outsiders, the map $\alpha : \tilde{P}_\varepsilon \to P_\varepsilon$ defined by $\alpha(\varepsilon,\tilde{\tau},\tilde{\tau}_o) = (\varepsilon, f(\tilde{\tau},\tilde{\tau}_o^\circ))$ is p-essential. Therefore, $\alpha \circ \tilde{\Psi}$ is p-essential, i.e. $\tilde{\Psi}^* \circ \alpha^* : H^*(P_\varepsilon, \partial P_\varepsilon) \to H^*(\tilde{X}_\varepsilon, \tilde{A})$ is nonzero, where $\tilde{A} = (\alpha \circ \Psi)^{-1}(\partial P_\varepsilon)$. For each n, s_n, $\tilde{\tau}_n$, if $f_{n,s_n}(\tilde{\tau}_n, \tilde{\tau}_o^\circ) = 0$ then $f_{n,s_n}(\tilde{\tau}_n, \tilde{\tau}_o) = 0$ for all $\tilde{\tau}_o$. Therefore, $\alpha \circ \tilde{\Psi}$ is linearly homotopic to $\Psi \circ h$ as maps from $(\tilde{X}_\varepsilon, \tilde{A})$ to $(P_\varepsilon, \partial P_\varepsilon)$. Hence this last map is essential. Since $\tilde{A} \subseteq (\Psi \circ h)^{-1}(\partial P_\varepsilon)$, $\Psi \circ h$ is essential as a map from $(\tilde{X}_\varepsilon, (\Psi \circ h)^{-1}(\partial P_\varepsilon))$, which implies that Ψ is essential and completes the proof that the f-image of a p-stable set of $\tilde{G}$ is a p-stable set of G.

For the other direction we follow the proof in Mertens [34, p. 556]. We begin with a preliminary lemma.

Lemma C.2. *Let $(B, \partial B)$ be the ball-pair whose dimension is $dim(\tilde{\Sigma}) - dim(\Sigma)$. There exists a map $h : \Sigma \times B \times \tilde{\Sigma}_o \to \tilde{\Sigma}$ such that for each $(\sigma, \tilde{\sigma}_0) \in \Sigma \times \tilde{\Sigma}_o$, $f \circ h(\sigma, B, \tilde{\sigma}_o) = \sigma$; furthermore, if $\sigma \in \Sigma \setminus \partial\Sigma$ then h maps $B \setminus \partial B$ homeomorphically onto the set of $\tilde{\sigma} \in \tilde{\Sigma} \setminus \partial\tilde{\Sigma}$ such that $f(\tilde{\sigma}, \tilde{\sigma}_o) = \sigma$.*

Proof of Lemma. For each n, let $K_n = \dim(\tilde{\Sigma}_n) - \dim(\Sigma_n)$. We show that there exists a map $h_n : \Sigma_n \times [0,1]^{K_n} \times \tilde{\Sigma}_o \to \tilde{\Sigma}_n$ such that for each $(\sigma, \tilde{\sigma}_o)$, $f_n \circ h_n(\sigma_n, [0,1]^{K_n}, \tilde{\sigma}_o) = \sigma_n$; furthermore, if $\sigma_n \in \Sigma_n \setminus \partial\Sigma_n$ then h_n maps the interior of $[0,1]^{K_n}$ homeomorphically onto the set of $\tilde{\sigma}_n \in \tilde{\Sigma}_n \setminus \partial\tilde{\Sigma}_n$ such that $f_n(\tilde{\sigma}_n, \tilde{\sigma}_o) = \sigma_n$. The lemma follows, since f is a product of the f_n's.

Fix a player n. Since $(\tilde{G}, f)$ embeds G, there exists a face $\tilde{\Sigma}_n^0$ of $\tilde{\Sigma}_n$ such that for each $\tilde{\sigma}_o$, $f_n(\cdot, \tilde{\sigma}_o)$ is a linear homeomorphism between $\tilde{\Sigma}_n^0$ and Σ_n, with the homeomorphism being independent of $\tilde{\sigma}_o$. Therefore, we can view Σ_n as this face $\tilde{\Sigma}_n^0$. As there is nothing to prove if $K_n = 0$, assume that $K_n > 0$. Take now a sequence $\tilde{\Sigma}_n^0 \subsetneq \cdots \subsetneq \tilde{\Sigma}_n^{K_n}$, where for each $0 < k \leqslant K_n$, $\tilde{\Sigma}_n^{k-1}$ is a maximal proper face of $\tilde{\Sigma}_n^k$, and $\tilde{\Sigma}_n^{K_n} = \tilde{\Sigma}_n$. For each $0 < k \leq K_n$, construct a function $f_n^k : \tilde{\Sigma}_n^k \times \tilde{\Sigma}_o \to \tilde{\Sigma}_n^{k-1}$ as follows. For each $\tilde{\sigma}_o$, let $f_n^k(\tilde{\sigma}_n, \tilde{\sigma}_o)$ equal: (a) $f_n(\tilde{\sigma}_n, \tilde{\sigma}_o)$ if $\tilde{\sigma}_n$ is the unique vertex of $\tilde{\Sigma}_n^k$ that is not contained in $\tilde{\Sigma}_n^{k-1}$; (b) $\tilde{\sigma}_n$ if $\tilde{\sigma}_n \in \tilde{\Sigma}_n^{k-1}$; (c) the value obtained by a linear interpolation of the values of f_n^k on Σ_n^{k-1} and the vertex outside $\tilde{\Sigma}_n^k$. Then $f_n(\tilde{\sigma}_n, \tilde{\sigma}_o) = f_n^1 \circ \cdots \circ f_n^{K_n}(\tilde{\sigma}_n, \tilde{\sigma}_o)$; and, for each k, $f_n^k \circ \cdots \circ f_n^{K_n}(\tilde{\sigma}_n, \tilde{\sigma}_o)$ belongs to the interior of $\tilde{\Sigma}_n^{k-1}$ if $\tilde{\sigma}_n$ belongs to the interior of $\tilde{\Sigma}_n$. Therefore, the lemma is proved if we can show that for each $0 < k \leq K_n$, there is a map $h_n^k : \tilde{\Sigma}_n^{k-1} \times [0,1] \times \tilde{\Sigma}_o \to \tilde{\Sigma}_n^k$ such that for each $(\tilde{\sigma}_n^{k-1}, \tilde{\sigma}_o) \in \tilde{\Sigma}_n^{k-1} \times \tilde{\Sigma}_o$: (i) $f_n^k \circ h_n^k(\tilde{\sigma}_n^{k-1}, [0,1], \tilde{\sigma}_o) = \tilde{\sigma}_n^{k-1}$; (ii) if $\tilde{\sigma}_n^{k-1} \in \tilde{\Sigma}_n^{k-1} \setminus \partial\tilde{\Sigma}_n^{k-1}$, then it maps $([0,1], \{0,1\})$ homeomorphically onto the projection to $(\tilde{\Sigma}_n^k, \partial\tilde{\Sigma}_n^k)$ of $(f_n^k)^{-1}(\tilde{\Sigma}_n^k \times \{\tilde{\sigma}_o\})$.

Now fix $0 < k \leqslant K_n$. Let $\tilde{s}_n^k$ be the vertex of $\tilde{\Sigma}_n^k$ that is not contained in $\tilde{\Sigma}_n^{k-1}$. For each $(\tilde{\sigma}_n^{k-1}, \tilde{\sigma}_o)$, let $\tilde{\tau}_n^k(\tilde{\sigma}_n^{k-1}, \tilde{\sigma}_o)$ be unique point whose $\tilde{s}_n^k$-th coordinate is the largest among all those points $\tilde{\sigma}_n \in \tilde{\Sigma}_n^k$ for which $f_n^k(\tilde{\sigma}_n, \tilde{\sigma}_o) = \tilde{\sigma}_n^{k-1}$. $\tilde{\tau}_n^k(\tilde{\sigma}_n^{k-1}, \tilde{\sigma}_o)$ is clearly a continuous function, lies in $\partial\tilde{\Sigma}_n^k$, and lies outside $\tilde{\Sigma}_n^{k-1}$ if $\tilde{\sigma}_n^{k-1} \notin \partial\tilde{\Sigma}_n^{k-1}$. We can now define $h_n^k : \tilde{\Sigma}_n^{k-1} \times [0,1] \times \tilde{\Sigma}_o$ as follows: $h_n^k(\tilde{\sigma}_n^{k-1}, t, \tilde{\sigma}_o) = (1-t)\tilde{\sigma}_n^{k-1} + t\tilde{\tau}_n^k(\tilde{\sigma}_n^{k-1}, \tilde{\sigma}_o)$. One sees easily that h_n^k has the desired properties. □

Let Σ^* be a p-stable set of G. We construct a p-stable set $\tilde{\Sigma}^*$ of $\tilde{G}$ such that $f(\tilde{\Sigma}^*) = \Sigma^*$. As in Mertens [34] it is sufficient to prove this result when Σ^* is semi-algebraic and has a semi-algebraic germ $X \subseteq \mathcal{N}$; that is, X satisfies the corresponding versions of properties (1)-(3) of the first part of the proof of this theorem, but where property (2) is stronger: for each small ε, $X_\varepsilon \setminus \partial X_\varepsilon$ is connected and dense in X_ε. Let $\varepsilon \in (0,1)$ be such that the projection $\Psi : (X_\varepsilon, \partial X_\varepsilon) \to (P_\varepsilon, \partial P_\varepsilon)$ is essential. Define $\bar{P}_\varepsilon \equiv P_\varepsilon \times B \times \tilde{\Sigma}_o$ and let $\partial\bar{P}_\varepsilon$ be its topological

boundary. Let $\bar{X}_\varepsilon \subset \bar{P}_\varepsilon \times \Sigma \times B$ be the set of $(\varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu)$ such that $(\varepsilon', \tau, \sigma) \in X_\varepsilon$. Then the projection $\bar{\Psi} : (\bar{X}_\varepsilon, \partial\bar{X}_\varepsilon) \to (\bar{P}_\varepsilon, \partial\bar{P}_\varepsilon)$ is essential, where $\partial\bar{X}_\varepsilon = \bar{\Psi}^{-1}(\partial\bar{P}_\varepsilon)$.

Let $\tilde{\mathcal{G}}_o$ be the space of all payoff functions for the outsiders when the strategy space is $\tilde{\Sigma} \times \tilde{\Sigma}_o$. Let $\tilde{\Gamma}_o = \tilde{\mathcal{G}}_o \times \bar{P}_\varepsilon \times \Sigma \times B$ and let $\hat{E} \subseteq \tilde{\Gamma}_o \times \tilde{\Sigma}_o$ be the set of $(\tilde{G}'_o, \varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu, \tilde{\sigma}_o)$ such that: $\tilde{\sigma}_o$ is an equilibrium of the game among the outsiders that is induced by $\tilde{G}'_o$ when the insiders play $(1-\varepsilon')h(\sigma_o, \mu, (1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o) + \varepsilon' h(\tau, \lambda, (1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o)$ and the strategies of the outsiders are perturbed towards $\tilde{\tau}_o$ by ε'. Let $\hat{\alpha} : (\hat{E}, \partial\hat{E}) \to (\tilde{\Gamma}_o, \partial\tilde{\Gamma}_o)$ be the natural projection, where $\partial\tilde{\Gamma}_o$ is the boundary of $\tilde{\Gamma}_o$ and $\partial\hat{E}$ is the inverse image of this set under the projection. Let $\tilde{\mathcal{G}}^*_o$ be the one-point compactification of $\tilde{\mathcal{G}}_o$ obtained by adding the point ∞. Denote by $\tilde{\Gamma}^*_o$ the set $\tilde{\mathcal{G}}^*_o \times \bar{P}_\varepsilon \times \Sigma \times B$ and let $\partial\tilde{\Gamma}^*_o$ be the boundary of $\tilde{\Gamma}^*_o$. Then $(\tilde{\Gamma}^*_o, \partial\tilde{\Gamma}^*_o)$ is a compact manifold with boundary. Let $\hat{E}^*$ be the quotient space of $\hat{E} \cup (\{\infty\} \times \bar{P}_\varepsilon \times \Sigma \times B \times \tilde{\Sigma}_o)$ obtained by treating for each fixed $(\varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu)$ all points of the form $(\infty, \varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu, \tilde{\sigma}_o)$ as being equivalent. Then $\hat{\alpha}$ extends to a projection map from $(\hat{E}^*, \partial\hat{E}^*)$ to $(\tilde{\Gamma}^*_o, \partial\tilde{\Gamma}^*_o)$, still denoted by $\hat{\alpha}$, where $\partial\hat{E}^*$ is the inverse image of $\partial\tilde{\Gamma}^*_o$ under the projection.

Lemma C.3. *There exists a homeomorphism $\hat{\beta} : (\tilde{\Gamma}^*_o, \partial\tilde{\Gamma}^*_o) \to (\hat{E}^*, \partial\hat{E}^*)$ such that $\hat{\alpha} \circ \hat{\beta}$ is homotopic to the identity map.*

Proof of Lemma. The construction in Kohlberg and Mertens [24, Theorem 1] extends here to prove the lemma. We sketch the changes required. As there, parameterize a game $\tilde{G}'_o$ as a pair $(\hat{G}'_o, \hat{g}')$, where (a) $\hat{g}'$ is a vector that for each outsider i and pure strategy $\tilde{s}_i$ of i gives the payoff to this strategy against the uniform strategy of his opponents; and (b) for each pure strategy profile $(\tilde{s}_i, \tilde{s}_{-i})$, $\hat{G}'_{oi}(\tilde{s}_i, \tilde{s}_{-i}) = \hat{G}'_i(\tilde{s}_i, \tilde{s}_{-i}) - \hat{g}'_{i,\tilde{s}_i}$.

We first describe the map β^{-1}. Given a point $(\hat{G}'_o, \hat{g}', \varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu, \tilde{\sigma}_o)$ in $\hat{E}$, map it to $(\hat{G}'_o, \hat{z}', \varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu)$ where for each i and $\tilde{s}_i$, $\hat{z}'_{i,\tilde{s}_i} = (1-\varepsilon')\tilde{\sigma}_{i,\tilde{s}_i} + \varepsilon'\tilde{\tau}_{i,\tilde{s}_i} + v_i$ where v_i is the payoff to i from the profile $(((1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o, (1-\varepsilon')h(\sigma, \mu, (1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o) + \varepsilon' h(\tau, \lambda, (1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o))$.

The map β can now be computed as follows. Given $(\hat{G}'_o, \hat{z}', \varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu)$, compute first the equilibrium strategy $\tilde{\sigma}_o$ by computing for each i the point $\hat{\sigma}_o$ in the perturbed simplex $(1-\varepsilon')\tilde{\Sigma}_i + \varepsilon'\tilde{\tau}_i$ that is closest to $\hat{z}'_i$ (in the ℓ_2-norm) and then letting $\tilde{\sigma}_i = (\hat{\sigma}_i - \varepsilon'\tilde{\tau}_i)/(1-\varepsilon')$. Then compute $\hat{g}'$ to be the unique point such that in the game $(\hat{G}'_o, g')$, for each i and each pure strategy $\tilde{s}_i$, the payoff to player i from each strategy of the form $(1-\varepsilon')\tilde{s}_i + \varepsilon'\tilde{\tau}_i$ against $(\hat{\sigma}_o, (1-\varepsilon')h(\sigma, \mu, \hat{\sigma}_o) + \varepsilon' h(\tau, \lambda, \hat{\sigma}_o))$ is $z_{i,\tilde{s}_i} - \hat{\sigma}_{i,\tilde{s}_i}$.

One sees easily that β and β^{-1} are inverses of each other as maps between $\hat{E}$ and $\tilde{\Gamma}_o$. Moreover, β extends to $\tilde{\Gamma}^*_o$ in the obvious way, as does β^{-1} from $\hat{E}^*$ to $\tilde{\Gamma}^*_o$. Also, $\hat{E}^* \circ \beta$ is linearly homotopic to the identity as in [24, p. 1022]. □

The above lemma implies that the projection map $\hat{\alpha}$ is p-essential. Now let $\hat{Y}_\varepsilon \subset \hat{E}$ be the fibered product of $\hat{\alpha}$ and the map from $\bar{X}_\varepsilon$ to $\tilde{\Gamma}_o$ that sends each $(\varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu)$ to $(\tilde{G}_o, \varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu)$, where $\tilde{G}_o$ is the payoff function of the outsiders in the game $\tilde{G}$. Then the projection map from $(\hat{Y}_\varepsilon, \partial\hat{Y}_\varepsilon)$ to $(\bar{X}_\varepsilon, \partial\bar{X}_\varepsilon)$ is essential, where $\partial\hat{Y}_\varepsilon$ is the inverse image of $\bar{X}_\varepsilon$ under this projection. Hence the map $\hat{\Psi} : (\hat{Y}_\varepsilon, \partial\hat{Y}_\varepsilon) \to (\bar{P}_\varepsilon, \partial\bar{P}_\varepsilon)$ is essential. As in Mertens [34] we can extract a subset $\hat{X}_\varepsilon$ of $\hat{Y}_\varepsilon$ such that: (a) the projection of $\hat{X}_o$ to Σ is Σ^*; (b) the restriction of $\hat{\Psi}$ to $\hat{X}_\varepsilon$ is essential; (c) for all $0 < \varepsilon' \leqslant \varepsilon$, $X_{\varepsilon'} \setminus \partial X_{\varepsilon'}$ is connected and dense in $X_{\varepsilon'}$.

Let $\iota : (\bar{P}_\varepsilon, \partial\bar{P}_\varepsilon) \to (\tilde{P}_\varepsilon, \partial\tilde{P}_\varepsilon)$ be the map $\iota(\varepsilon', \tau, \lambda, \tilde{\tau}_o) = (\varepsilon', h(\tau, \lambda, \tilde{\tau}_o^0), \tilde{\tau}_o))$, where $\tilde{\tau}_o^0$ is the uniform-strategy profile of the outsiders. Since ι maps the interior of P_ε homeomorphically onto the interior of $\tilde{P}_\varepsilon$ (by Lemma C.2), ι and hence also $\iota \circ \hat{\Psi}$ are essential. Define $\hat{\phi} : (\hat{X}_\varepsilon, \partial\hat{X}_\varepsilon) \to (\tilde{P}_\varepsilon, \partial\tilde{P}_\varepsilon)$ by $\hat{\phi}(\varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \mu, \tilde{\sigma}_o) = (\varepsilon', h(\tau, \lambda, (1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o), \tilde{\tau}_o)$. Then $\hat{\phi}$ is homotopic to $\iota \circ \hat{\alpha}$ using the homotopy that sends $(\varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \tilde{\sigma}_o, \mu)$ to $(\varepsilon', h(\tau, \lambda, (1-t)((1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o) + t\tilde{\tau}_o^0), \tilde{\tau}_o)$ as t goes from 0 to 1. Hence $\hat{\phi}$ is essential as well.

Define $\hat{h} : (\hat{X}_\varepsilon, \partial X_\varepsilon) \to (\tilde{\mathcal{N}}_\varepsilon, \partial\tilde{\mathcal{N}}_\varepsilon)$ by

$$\hat{h}(\varepsilon', \tau, \lambda, \tilde{\tau}_o, \sigma, \tilde{\sigma}_o, \mu) = (\varepsilon', h(\tau, \lambda, (1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o), \tilde{\tau}_o, h(\sigma, \mu, (1-\varepsilon')\tilde{\sigma}_o + \varepsilon'\tilde{\tau}_o), \tilde{\sigma}_o)\,.$$

Letting $\tilde{X}_\varepsilon = \hat{h}(\hat{X}_\varepsilon)$, we have that $\hat{\phi} = \tilde{\Psi} \circ \hat{h}$. Therefore, $\tilde{\Psi}$ is essential. Also, for each $\varepsilon' \in (0, \varepsilon)$, $\tilde{X}_{\varepsilon'} \setminus \partial\tilde{X}_{\varepsilon'}$ is homeomorphic to $\hat{X}_{\varepsilon'} \setminus \partial\hat{X}_{\varepsilon'}$ by Lemma C.2. Hence $\tilde{X}$ is a germ for $\tilde{X}_0$. Since the image under f of $\tilde{X}_0$ is exactly Σ^*, the proof is complete. □

Appendix D. Proof of Theorem 5.2

Theorem 5.2. $(\mathcal{Q}, \partial\mathcal{Q})$ is a pseudomanifold of the same dimension as $(\mathbb{P}, \partial\mathbb{P})$. Moreover, Q^* is stable if and only if the projection map $\Psi : (\mathcal{Q}, \partial\mathcal{Q}) \to (\mathbb{P}, \partial\mathbb{P})$ is essential.

Proof. As assumed in Section 4.4, the proof invokes genericity of payoffs by assuming that certain points and polyhedra, identified as they arise during the proof, are in general position.

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 (in P_n) of the strategies in T_n. For simplicity, we write $d(T_n)$ for $d(P_n(T_n))$. See Section 5.1 for additional notation used below.

For any vertex π_n^1 of Π_n^1, let $S_n^1(\pi_n^1)$ be the set of pure strategies that map to π_n^1 under $\bar{\pi}_n^1$. For any face Ψ_n^1 of Π_n^1, let $S_n^1(\Psi_n^1)$ be the set of pure strategies s_n^1 such that $\bar{\pi}_n^1(s_n^1) \in \Psi_n^1$.

Let H_n^0 be the set of information sets $h_n \in H_n \setminus H_n^*$ of player n such that at the last information set $h_n' \in H_n^*$ that precedes h_n, the action there leading to h_n belongs to A_n^*, which is the set of his equilibrium actions. If a subset T_n^0 of S_n^0 is such that $P_n(T_n^0)$ contains an equilibrium in $\bar{Q}_n^* \equiv \rho_n(\bar{\Sigma}_n)$, then for each first information set $h_n \in H_n^0$ there exists a strategy in T_n^0 that enables h_n. Let $H_n^1 = H_n \setminus (H_n^* \cup H_n^0)$.

Lemma D.1. *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 $\bar{Q}_n^*$. 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 Π_n^* be the projection of P_n^0 to $\mathbb{R}^{Z^*}$, where Z^* is the set of terminal nodes reached with positive probability under the equilibria in $\bar{Q}^*$. $Z^* = Z_n^0 \cap Z_m^0$. Consequently, the payoff to a strategy s_n^0 against p_m^0 depends on s_n^0 only through its projection to Π_n^*. Since the projection of q_n^*—which is at least as good a reply as the other strategies in S_n^0 against p_m^0—belongs to the relative interior of Π_n^*, the strategies in S_n^0 are equally good replies against p_m^0. □

Lemma D.2. *Suppose that T_n^0 is a subset of S_n^0 such that $P_n(T_n^0)$ is a face of P_n^0 containing an equilibrium strategy in $\bar{Q}_n^*$. For each $t_n \in S_n^0 \setminus T_n^0$ there exists an information set $h_n \in H_n^0$ that t_n enables and where the action chosen by it is avoided by all $t_n^0 \in T_n^0$ that enable h_n.*

Proof. Since $P_n(T_n^0)$ is a face of P_n^0, which is itself a face of P_n, there exists a linear function $f : \mathbb{R}^Z \to \mathbb{R}$ that is zero on $P_n(T_n^0)$ and negative everywhere else on P_n. Fix p_m in the interior of P_m and define a payoff function $\tilde{u}_n$ for n by the equation: $p_0(z)p_m(z)\tilde{u}_n(z) = f(e_z)$ where e_z is the z-th unit vector in $\mathbb{R}^Z$. Then when n's payoff function is $\tilde{u}_n$, the strategies in $P_n(T_n^0)$ are the best replies against p_m. Take $t_n \notin T_n^0$. Since it is suboptimal against p_m, there exists an information set h_n where the action a chosen by t_n is suboptimal. Then h_n does not belong to H_n^*: indeed, since t_n belongs to S_n^0, a is an equilibrium action if $h_n \in H_n^*$; and since $P_n(T_n^0)$ contains an equilibrium strategy, there exists a strategy in T_n^0 that enables such an h_n and chooses a. Let h_n' be the last information set preceding h_n that belongs to H_n^*. Then obviously the action chosen by t_n at h_n' is an equilibrium action, as $t_n \in S_n^0$, and thus h_n belongs to H_n^0. Any strategy in T_n^0 that enables h_n avoids a, which proves the lemma. □

For the next lemma, it is worth recapitulating the exact definition of the set Π_n^1. Recall from Section 5.1 that we fix a completely mixed enabling strategy $\bar{p}_m$ for player m and compute for each p_n the total probability $\eta(p_n)$ of reaching a terminal node in Z_n^1 under $(\bar{p}_m, p_n)$. $\mathcal{H}_n$ is a hyperplane in $\mathbb{R}^{Z_n^1}$ that separates the projection $\Psi_{Z_n^1}(P_n^1)$ of P_n^1 to $\mathbb{R}^{Z_n^1}$ from the origin of $\mathbb{R}^{Z_n^1}$ and that has $\left(p_0(z)\bar{p}_m(z)\right)_{z \in Z_n^1}$ as its normal and some $\varepsilon > 0$ as its constant. The function $\bar{\pi}_n^1$ maps each point $p_n \in P_n \setminus P_n^0$ to $\varepsilon(\eta(p_n))^{-1}\Psi_{Z_n^1}(p_n) \in \mathcal{H}_n$.

Lemma D.3. *For a strategy s_n of S_n^1, $\bar{\pi}_n^1(s_n)$ is a vertex of Π_n^1 iff there exists a unique information set $h_n \in H_n^*$ with the property that s_n enables h_n and chooses a non-equilibrium action there.*

Proof. Let $s_n \in S_n^1$ be a pure strategy satisfying the condition of the lemma. We prove by contradiction that $\bar{\pi}_n^1(s_n)$ is a vertex of Π_n^1. Therefore suppose to the contrary that $\bar{\pi}_n^1(s_n)$, which equals $\varepsilon(\eta(s_n))^{-1}\Psi_{Z_n^1}(s_n)$, is not a vertex of Π_n^1. We can express $\bar{\pi}_n^1(s_n)$ as a unique

convex combination $\sum_{j=1}^{J} \lambda^j \pi_n^{1,j}$ where $J > 1$ and for each $1 \leqslant j \leqslant J$, $\pi_n^{1,j}$ is a vertex of Π_n^1. For each j, since $\pi_n^{1,j}$ is a vertex of Π_n^1, we can express $\pi_n^{1,j}$ as $\varepsilon(\eta(s_n^{1,j}))^{-1}\Psi_{Z_n^1}(s_n^{1,j})$ for some $s_n^{1,j}$ in S_n^1. Since $\lambda^j > 0$, $s_n^{1,j}$ cannot choose a non-equilibrium action at any $h_n' \neq h_n$ in H_n^* that it enables; since it belongs to $S_n^{1,j}$ it must therefore enable h_n; and it cannot choose a different non-equilibrium action from s_n at h_n. Observe now that the probability $\eta(s_n)$ and $\eta(s_n^{1,j})$ for all j are equal and exactly the probability that Nature and the strategy $\bar{p}_m$ do not exclude h_n. Therefore, $\Psi_{Z_n^1}(s_n) = \sum_j \lambda^j \Psi_{Z_n^1}(s_n^{1,j})$. Modify $s_n^{1,j}$ to a strategy $t_n^{1,j}$ so that at every information set other than the successors to h_n, $t_n^{1,j}$ agrees with s_n, and at the successors to h_n it agrees with $s_n^{1,j}$. It is now clear that when viewed as enabling strategies, $s_n = \sum_j \lambda^j t_n^{1,j}$ and thus s_n is a convex combination of the strategies $t_n^{1,j}$. But that is a contradiction since s_n is a vertex of P_n and all the strategies $t_n^{1,j}$ are different from one another and from s_n as they induce the points $\pi_n^{1,j}$ that are different from one another and from $\bar{\pi}_n^1(s_n)$ in $\mathcal{H}_n$.

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 $s_n^{1,k}$ 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 $\Psi_{Z_n^1}(s_n) = \sum_k \Psi_{Z_n^1}(s_n^{1,k})$. Therefore, $\bar{\pi}_n^1(s_n)$ cannot be a vertex of Π_n^1. □

For each $T_n^0 \subseteq S_n^0$ 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 there exists a strategy $t_n^0 \in T_n^0$ that agrees with s_n^0 at all information sets in H_n^* and H_n^0 that s_n^1 enables, except those in H_n^* where s_n^1 chooses a nonequilibrium action. For a face Ψ_n^1 of Π_n^1, let $S_n^1(T_n^0; \Psi_n^1) = S_n^1(\Psi_n^1) \cap S_n^1(T_n^0)$ and $T_n \equiv T_n^0 \cup S_n^1(T_n^0; \Psi_n^1)$. For notational simplicity, we refer to $S_n^1(T_n^0; \Psi_n^1)$ as T_n^1. The following lemma provides an important feature of the set T_n.

Lemma D.4. *The strategies in T_n are the vertices of a face of P_n whose dimension is $d(T_n) \equiv d(T_n^0) + d(\Psi_n^1) + 1$.*

Proof. Let $\hat{T}_n^1$ be the set of strategies t_n^1 in T_n^1 such $\bar{\pi}_n^1(t_n^1)$ is a vertex of Ψ_n^1. We will first show that every $t_n \in T_n \setminus (T_n^0 \cup \hat{T}_n^1)$ is affinely dependent on the strategies in $T_n^0 \cup \hat{T}_n^1$. Let $t_n^1 \in T_n \setminus (T_n^0 \cup \hat{T}_n^1)$. By Lemma D.3, there exist information sets $h_n^1, \ldots h_n^K$, $K > 1$, in H_n^* such that for each k, t_n^1 chooses a non-equilibrium action a^k at h_n^k, and at each other $h_n \in H_n^*$ it chooses an equilibrium action. Fix $t_n^0 \in T_n^0$ that agrees with t_n^1 everywhere except at the information sets h_n^k, and their successors, for each k. For each k let $t_n^{1,k}$ 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 $t_n^{1,k}$ belongs to $\hat{T}_n^1$ by Lemma D.2 and also, $t_n^1 = \sum_k t_n^{1,k} - (K-1)t_n^0$. Thus, t_n^1 is an affine combination of the strategies in $T_n^0 \cup \hat{T}_n^1$.

For each $j = 0, \ldots, d(\Psi_n^1)$, pick a strategy $t_n^{1,j} \in S_n^1(T_n^0)$ such that $\bar{\pi}_n^1(t_n^{1,j})$ is a vertex of Ψ_n^1. Let $\tilde{T}_n^1$ be the collection of these strategies. We show that strategies in $\hat{T}_n^1 \setminus \tilde{T}_n^1$ are now affinely dependent on the strategies in $T_n^0 \cup \tilde{T}_n^1$. Fix $t_n^1 \in \hat{T}_n^1 \setminus \tilde{T}_n^1$. By Lemma D.3 there exists a unique information set $h_n \in H_n^*$ enabled by t_n^1 and where it chooses a non-equilibrium action. By construction of $\tilde{T}_n^1$, there exists a subset $(t_n^{1,j})_{j=1}^J$ of $\tilde{T}_n^1$ such that $\bar{\pi}_n^1(t_n^1)$ is expressible as an affine combination $\sum_j \lambda^j \bar{\pi}(t_n^{1,j})$ with $\lambda^j \neq 0$ for all j. For each j, $s_n^{1,j}$ enables h_n and chooses a at h_n; at all other information sets in H_n^* it chooses an equilibrium action. Let $t_n^0 \in T_n^0$ be a strategy that agrees with t_n^1 everywhere except at h_n and its successors. For each j, let $\tilde{t}_n^{0,j}$ be a strategy in T_n^0 that agrees with $t_n^{1,j}$ everywhere except at h_n and its successors. Modify $\tilde{t}_n^{0,j}$ to a strategy that agrees with $\tilde{t}_n^{0,j}$ everywhere except at h_n and its successors, where it agrees with t_n^0. By Lemma D.2, $t_n^{0,j}$ belongs to T_n^0 for each j. Now $t_n^1 = t_n^0 + \sum_j \lambda^j (t_n^{1,j} - t_n^{0,j})$ and is affinely dependent on the strategies in $T_n^0 \cup \tilde{T}_n^1$.

It follows now from the above arguments that the affine space A spanned by $T_n^0 \cup \tilde{T}_n^1$ 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 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 $Q_n \neq P_n(T_n)$. There exists a point p_n in the relative of interior of $P_n(T_n)$ and Q_n. Therefore p_n can be expressed as a convex combination of the vertices of Q_n in two different ways: (a) $\sum_i \lambda^i t_n^{0,i} + \sum_j \lambda^j t_n^{1,j}$, where the $t_n^{0,i}$'s are in T_n^0 and the $t_n^{1,j}$'s are in T_n^1; (b) $\sum_i \mu^i t_n^{0,i} + \sum_j \mu^j t_n^{1,j} + \sum_k \mu^k t_n^{2,k}$, where now the $t_n^{2,k}$'s are the vertices of Q_n that are not in T_n. Consider one of the $t_n^{2,k}$'s. If it belongs to $S_n^0 \setminus T_n^0$ then by Lemma D.2 there is an information set h_n in H_n^0 that is enabled by $t_n^{2,k}$ where the action chosen by h_n is avoided all strategies in T_n^0 that enable it; by the definition, the strategies in T_n^1 avoid this action as well. This implies under the expression in (b) that the nodes following this action are assigned a positive probability, but not under (a), which is impossible. If $t_n^{2,k}$ belongs to $S_n^1 \setminus T_n^1$ then it must belong to $S_n^1(\Psi_n^1)$ since otherwise under (b) $\bar{\pi}_n^1(p_n) \notin \Psi_n^1$. Since $s_n \notin T_n^1$ there exists an information set $h_n \in H_n^0$ enabled by $t_n^{2,k}$ where the continuation strategy of $t_n^{2,k}$ coincides with that of some $t_n \in S_n^0 \setminus T_n^0$ but not for any $s_n \in T_n^0$. As in the previous case, this too is impossible. □

One corollary of the above result obtains when we take T_n^1 to be S_n^1 and Ψ_n^1 to be Π_n^1. The dimension of P_n is $d(P_n^0) + d(\Pi_n^1) + 1$. Observe that P_n^1 is a face of P_n iff $\Psi_{Z_n^1}(P_n^1)$ is a face of $\Psi_{Z_n^1}(P_n)$, which is equivalent to saying that Π_n^1 is homeomorphic to $\Psi_{Z_n^1}(P_n^1)$. Thus, the dimension of P_n^1 equals $d(\Pi_n^1)$ if P_n^1 is a face of P_n and otherwise it equals the dimension of P_n.

Fix $T \equiv ((T_1^0, \Psi_1^1), (T_2^0, \Psi_2^1))$, where for each n, $P_n(T_n^0) \cap Q_n^*$ nonempty and Ψ_n^1 is a (possibly empty) face of Π_n^1. Let $A_n(T)$ be the set of points in $P_n(T_n^0)$ such that the strategies in T_m are

all best replies. Let $\tilde{T}_n^0$ be the unique subset of T_n^0 such that the interior of $A_n(T)$ is contained in the interior of $P_n(\tilde{T}_n^0)$. Let $\tilde{d}_n^*(T)$ be the dimension of $A_n(T)$. Since $P_m(T_m^0)$ contains an equilibrium strategy for m, and likewise for n, $A_n(T)$ is a subset of $A_n^*(T) = P_n(\tilde{T}_n^0) \cap \bar{Q}_n^*$. Moreover, since the equilibrium strategy for m in $P_m(T_m^0)$ is undominated, this last set $A_n^*(T)$ is the set $P_n(\tilde{T}_n^0) \cap Q_n^*$ as well.

Lemma D.5. *If T_m^1 is nonempty then $A_n(T)$ is a proper face of $A_n^*(T)$.*

Proof. This follows from the genericity of payoffs. Fix $t_m^1 \in T_m^1$. There exists an information set $h_m \in H_m^*$ where it chooses a non-equilibrium action a. If the path from each (x, a), for $x \in h_m$ that is reached under the equilibrium outcome, does not pass through an information set $h_n \in H_n^0$, then a would be suboptimal against every equilibrium $\bar{Q}_n^*$ and $A_n(T)$ would be empty. Thus, there exists a first information set $h_n \in H_n^0$ and nodes $x \in H_m^*$ and $y \in h_n$ such that $(x, a) \prec y$ and x is reached under the equilibrium outcome. Because $A_n^*(T)$ is nonempty, there is a strategy $t_n^0 \in \tilde{T}_n^0$ that enables h_n. Clearly, there must be multiple such strategies that differ in the continuation from h_n, again by genericity. Perturbing the probabilities of the terminal nodes following y does not affect the payoffs to strategies in T_m^0 but they affect the payoff to t_m^1. Hence $A_n(T)$ is a proper face of $A_n^*(T)$. □

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

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

Lemma D.6. *Suppose T_m^1 is nonempty. $B_n(T)$ is a polyhedron of dimension $d_n^*(T) + d(T_n^0) - \tilde{d}_n^*(T)$. 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 $r_m \in \bar{T}_m^1$ such that for each $p_n^0 \in B_n'(T)$, r_m is an equally good reply against every point of the form $\lambda p_n^0 + (1-\lambda) r_n^0$, for $0 \leqslant \lambda < 1$, in $B_m^*(T)$; moreover, in this case, letting $\check{R}_m^1$ be the set of such r_m, for any $r_m \in \bar{R}_m^1$, if r_m is a best reply against a point in $B_n^*(T)$ then every point in $\check{R}_m^1$ 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. Let $\tilde{P}_n^0$, $\tilde{P}_n(T_n^0)$, and $\tilde{B}_n^*(T)$ be the convex cones spanned by P_n^0, $P_n(T_n^0)$, and $B_n^*(T)$ respectively. Let $\xi : P_n^0 \times \tilde{P}_n(T_n^0) \to \tilde{P}_n^0$ be the function $\xi(p_n^0, \tilde{r}_n^0) = p_n^0 + \tilde{r}_n^0$. Then for each $\tilde{p}_n^0$, $\xi^{-1}(\tilde{p}_n^0)$ is a set of dimension $d(T_n^0)$. Hence the dimension of $\hat{B}(T) \equiv \xi^{-1}(\tilde{B}_n^*(T))$ is $d(T_n^0) + d_n^*(T) + 1$. Obviously $\hat{B}(T)$ is a polyhedron. For each face $\hat{B}_n'(T)$ of $\hat{B}_n(T)$ and for all points $(p_n^0, \tilde{r}_n^0) \in \hat{B}_n'(T)$ at least one of the following holds: (i) p_n^0 belongs to the boundary of P_n^0; (ii) there exists a strategy $t_m \in \bar{T}_m^1$ such that $\xi(p_n^0, \tilde{r}_n^0)$ belongs to the convex cone spanned by the face of $B_n^*(T)$ where this strategy t_m is an equally good reply; (iii) $\tilde{r}_n^0$ belongs to the boundary of $\tilde{P}_n(T_n^0)$.

Observe now that $B_n(T)$ is the projection of $\hat{B}_n(T)$ onto the first factor. Obviously it is a polyhedron. For each $p_n^0 \in B_n(T)$ and each $\tilde{r}_n^0$ such that $(p_n^0, \tilde{r}_n^0) \in \hat{B}_n(T)$, $(p_n^0, \tilde{r}_n^0 + \lambda r_n^0) \in \hat{B}_n(T)$ for all $r_n^0 \in B_n^*(T) \cap P_n(T_n^0)$ and $\lambda \geqslant 0$. If the set $B_n^*(T)$ is in generic position (i.e. if the payoffs are in generic position), then for each $p_n^0 \in B_n(T)$, there exists $\tilde{r}_n^0$ such that all points in $\hat{B}_n(T)$ that project to p_n^0 can be expressed in the form $(p_n^0, \tilde{r}_n^0 + \lambda r_n^0)$ for some $r_n^0 \in B_n^*(T) \cap P_n(T_n^0)$ and $\lambda \geqslant 0$. Since the set $B_n^*(T) \cap P_n(T_n^0)$ is the intersection of $P_n(T_n^0)$ with the affine space spanned by $A_n(T)$, the dimension of $B_n(T)$ is as asserted. The enumerated properties of $B_n(T)$ now follow directly from the corresponding points above; only the last part of property (iii) needs a proof. Suppose r_n^0 belongs to a proper face $P_n(R_n^0)$ of $P_n(T_n^0)$ and $A_n(T)$ is not contained in $P_n(R_n^0)$. If $(p_n^0, \lambda r_n^0)$ belongs to $\hat{B}_n(T)$, then so does $(p_n^0, \lambda r_n^0 + \bar{r}_n^0)$ for $\bar{r}_n^0 \in A_n(T) \setminus P_n(R_n^0)$ and $\lambda r_n^0 + \bar{r}_n^0$ does not belong to the convex cone generated by $P_n(R_n^0)$. □

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 $\hat{T}_m^1$ and $S_n^0 \setminus T_n^0$. By Lemma D.4, the dimension of the face spanned by T_m is $d(T_m^0) + d(\Psi_m^1) + 1$. By genericity of payoffs, the dimension of $C_n^*(T)$ is therefore $d(P_n) - d(T_m^0) - d(\Psi_m^1) - 1$.

Let $C_n(T)$ be 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 $\sum_i \mu^i p_n^i \in C_n^*(T)$, $\sum_i \mu^i = 1$, $\mu^1 > 0$, and $\bar{\pi}_n^1(\sum_i \mu^i p_n^i) = \pi_n^1$.

Lemma D.7. *The set $C_n(T)$ is a polyhedron of dimension $d(P_n^0) + d(P_n^1) + d_n(\Psi_n^1) - d(T_m^0) - d(\Psi_m^1) - d_n^*(T)$. On each maximal proper face C' of $C_n(T)$, exactly one of the following holds for all (p_n^1, π_n^1) in C'. If $q_n \in C^*(T)$ 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 $\bar{\pi}_n^1(\sum_i \mu^i p_n^i) = \pi_n^1$, 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 $\hat{T}_m^i$ 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^0_m and s_m span a face of P^0_m of which $P_m(T^0_m)$ is a maximal proper face; and if $i = 1$, Ψ^1_m and $\bar\pi^1_n(s_m)$ span a face of Π^1_m of which Ψ^1_m is a maximal proper face.

(3) *there exists a maximal proper face Ψ' of Ψ_n such that $\bar\pi^1_n(p^2_n)$ belongs to Ψ'.*

Proof. We show that $C_n(T)$ is a polyhedron of the stated dimension. Since the construction is similar to that in the previous lemma, the enumerated properties can be proved just as before. Let P^2_n be the convex hull of P^0_n and $P_n(T_n)$. Using Lemma D.4, the dimension of P^2_n is $d(P^0_n) + d(\Psi^1_n) + 1$. Let $\tilde P^2_n$ and $\tilde P_n$ be the convex cones spanned by P^2_n and P_n, respectively. Define $\xi : P^1_n \times \tilde P^2_n \to \tilde P_n$ by $\xi(p^1_n, \tilde p^2_n) = p^1_n + \tilde p^2_n$. Then for each $\tilde p_n$ in the interior of $\tilde P_n$, $\xi^{-1}(\tilde p_n)$ is a set of dimension $d(P^1_n) + d(P^0_n) + d(\Psi^1_n) + 1 - d(P_n)$. Letting $\tilde C^*_n(T)$ be the convex cone spanned by $C^*_n(T)$, the dimension of $\hat C_n(T) \equiv \xi^{-1}(\tilde C^*_n(T))$ is $d(P^0_n) + d(P^1_n) + d(\Psi^1_n) + 1 - d(T^0_m) - d(\Psi^1_m)$. The function $\bar\pi^1_n$ extends to $\tilde P \setminus \{\,0\,\}$. $C_n(T)$ is the image of $\hat C_n(T)$ under the function $\chi : P^1_n \times \tilde P^2_n$ given by $\chi(p^1_n, \tilde p^2_n) = (p^1_n, \bar\pi^1_n(p^1_n + \tilde p^2_n))$ and is thus a polyhedron. As will be shown in the course of the proof of the next lemma, $C^*_n(T) \cap P^2_n \subset P^0_n$. Therefore, $C^*_n(T) \cap P^2_n$ is the intersection of P^0_n with the affine space spanned by $B^*_n(T)$. For each $(p^1_n, \tilde p^2_n) \in \hat C_n(T)$, the point $(p^1_n, \tilde p^2_n + \mu q^0_n)$ belongs to $\hat C_n(T)$ for all $\mu > 0$ and $q^0_n \in C^*_n(T) \cap P^2_n$, and has the same image under χ as $(p^1_n, \tilde p^2_n)$. Moreover, if $C^*_n(T)$ is in general position then for each (p^1_n, π^1_n) in $C_n(T)$ every point in its inverse image under χ is expressible in this form. Therefore, the dimension of $C_n(T)$ is as given. □

Let $\mathcal{T}$ be the collection of T's such that $A_n(T)$, $B_n(T)$ and $C_n(T)$ are nonempty for each n. 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 $\mathcal{Q}(T) = \mathcal{Q}_1(T) \times \mathcal{Q}_2(T)$.

Lemma D.8. *(q^*, p^0, p^1, π^1) belongs to $\mathcal{Q}$ iff it belongs to $\mathcal{Q}(T)$ for some $T \in \mathcal{T}$.*

Proof. Suppose for each n that $q^*_n \in A_n(T), p^0_n \in B_n(T), (p^1_n, \pi^1_n) \in C_n(T)$ for some T. Choose $r^0_n \in P_n(T^0_n)$ and λ^0_n such that $q^0_n \equiv (1 - \lambda^0_n)p^0_n + \lambda^0_n r^0_n \in B^*_n(T)$. Also, fix $\tilde p^0_n$, $r^2_n \in P_n(T^1_n)$, $\mu^0_n, \mu^1_n, \mu^2_n$ such that $q^1_n \equiv \mu^0_n \tilde p^0_n + \mu^1_n p^1_n + \mu^2_n r^2_n$ belongs to $C^*_n(T)$. Fix points $\bar q^0_n$ and $\bar q^1_n$ in the interior of $A_n(T)$ and $B^*_n(T)$ for each n and consider for each $0 < \varepsilon < 1$, the LPS $(q^*, \tilde q^0(\varepsilon), \tilde q^1(\varepsilon))$ where for each n, $\tilde q^0_n(\varepsilon) = (1-\varepsilon)\bar q^0_n + \varepsilon q^0_n$; and $\tilde q^1_n(\varepsilon) = (1-\varepsilon)\bar q^0_n + \varepsilon \bar q^1_n(\varepsilon) + \varepsilon^2 q^1_n$. The strategies in T_n are equally good replies to $q^*_m, \tilde q^0_m(\varepsilon), \tilde q^1_m(\varepsilon)$ for all ε. We show that these strategies are lexicographic best replies to $(q^*, \tilde q^0(\varepsilon), \tilde q^1(\varepsilon))$ for all small ε, which proves that (q^*, p^0, p^1, π^1) belongs to $\mathcal{Q}$.

Observe first that for all ε, a strategy in $S^0_n \setminus T^0_n$ is an equally good reply against q^*_m and $\tilde q^0_m(\varepsilon)$ as strategies in T^0_n, by Lemma D.1, and no better a reply against $\tilde q^1_m(\varepsilon)$ by construction of $C^*_n(T)$. Now for a strategy s^1_n in S^1_n, consider the strategy t^1_n that agrees with s^1_n at every information set except that starting at each first information set $h_n \in H^0_n$ that s^1_n enables, t^1_n agrees with some t^0_n in T^0_n. Since strategies in T^0_n are at least as good as the other strategies in S^0_n, clearly t^1_n is at least as good a reply against $(q^*_m, \tilde q^0_m(\varepsilon), \tilde q^1_n(\varepsilon))$ as s^1_n. Observe now

that t_n^1 belongs to $S_n^1(T_n^0)$. If it belongs to $\hat{T}_n^1$, then it is an equally good reply as strategies in T_n against q_m^* and $\tilde{q}_m^0(\varepsilon)$ and no better reply a reply against $\tilde{q}_m^1(\varepsilon)$ by definition. If it belongs to $\bar{T}_n^1$ then it is an equally good reply to q_m^*, no better reply against $\tilde{q}_m^0(\varepsilon)$ for all ε, and a strictly worse reply against $\tilde{q}_m^1(\varepsilon)$ for all small ε, again by construction, since it is an inferior reply against $\bar{q}_n^1$ which belongs to the interior of $B_n^*(T)$. Finally, if it belongs to $S_n^1(T_n^1) \setminus (\hat{T}_n^1 \cup \bar{T}_n^1)$ then it is no better a reply against q_m^* and an inferior reply to $\tilde{q}_m^0(\varepsilon)$ for all small ε, since it is an inferior reply to $\bar{q}_n^0$ by construction of $A_n(T)$. Thus the strategies in T are lexicographic best replies to $(q^*, \tilde{q}^0(\varepsilon), \tilde{q}^1(\varepsilon))$ for all small ε.

Before proceeding to prove the converse, we use the above argument to show that the intersection of $C_n^*(T)$ with the convex hull P_n^2 of P_n^0 with $P_n(T_n)$ is in fact the intersection F of P_n^0 with the affine space spanned by $B_n^*(T)$—a fact that was asserted, but not proved, in the course of the proof of the previous lemma. Take a point q_n^1 in $C_n^*(T) \cap P^2$. If it belongs to P_n^0, then in fact it belongs to F by the definitions of $B_n^*(T)$ and $\hat{T}_m^1$. If it does not belong to P_n^0, then it assigns a positive weight to some strategy $s_n^1 \in T_n^1$. The above argument applied when using this q_n^1 shows that q_m^* is a best reply to $(q_n^*, \tilde{q}_n^0(\varepsilon), \tilde{q}_n^1(\varepsilon))$ and the strategies in T_n and S_n^0 are best replies to q_m^*. Observe now that $\tilde{q}_n^1(\varepsilon)$ is a convex combination of strategies in S_n^0 and T_n^1. Therefore, for all small δ, $((1-\delta-\delta^2)q_n^* + \delta\tilde{q}_n^0(\varepsilon) + \delta^2\tilde{q}_n^1(\varepsilon), q_m^*)$ is an equilibrium if ε is small as well. But these points induce different outcomes because $\tilde{q}_n^1(\varepsilon)$ has a non-equilibrium strategy, namely one in T_n^1, in its support, which is impossible. Thus, $C_n^*(T) \cap P_n^2 = F$ as claimed.

Returning to the proof of this lemma, 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 $q_n^1 = \mu_n^0\tilde{p}_n^0 + \mu_n^1 p_n^1 + \mu_n^2 r_n^2$ where $(1-\lambda_n^0)q_n^* + \lambda_n^0 r_n^0$ is a best reply against (q^*, q^0, q^1); 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 $(1-\lambda_n^0)q_n^* + \lambda_n^0 r_n^0$ in its interior. 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 a first information set $h_n \in H_n^0$ then the choices from there on prescribed by t_n^0 coincide with the choices dictated by some vertex of Q_n^0 or Q_n^1 that enables h_n. Observe that each $t_n^0 \in T_n^0$ is optimal against (q^*, q^0, q^1). If $\mu_n^2 > 0$, let Ψ_n^1 be the face of Π_n^1 that contains $\bar{\pi}_n^1(r_n^2)$ in its interior; otherwise let Ψ_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: pick a point $\check{r}_m^0$ in the face $\tilde{Q}_m^0$ of Q_m^0 containing q_m^* in its interior such that the strategies in Q_n^0 are equally good replies, 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. 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 $(1-\nu_m^i)q_m^i + \nu_m^i \check{r}_m^i$. 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 an information set h_n that is enabled by s_n is enabled by some vertex of either Q_n^0 or Q_n^1 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 H_n^0; then obviously some strategy in Q_n^0 or Q_n^1 enables h_n and chooses a_n, by the definition of T_n^0. 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_n^1 enables it and chooses this action, since otherwise s_n enables a terminal node that is excluded by all strategies in Q_n^1, contradicting the assumption that $\bar{\pi}_n^1(s_n) \in \Psi_n^1$. Thus s_n is a best reply and (q^*, q^0, q^1) belongs to $\mathcal{Q}(T)$. □

Lemma D.9. *$\mathcal{Q}$ is a pseudomanifold of dimension $\hat{d} \equiv d(\mathbb{P})$.*

Proof. Each $\mathcal{Q}(T)$ is a polyhedron of dimension $\hat{d}$. By the previous lemma $\mathcal{Q} = \cup_{T\in\mathcal{T}}\mathcal{Q}(T)$. Therefore, $\mathcal{Q}$ has dimension $\hat{d}$. To show that $\mathcal{Q}$ is a pseudomanifold, we establish three facts for each $(q^*, (p^0, p^1), \pi^1)$ that belongs to some $\mathcal{Q}(T)$: (1) if $(q^*, (p^0, p^1), \pi^1)$ belongs to the interior of $\mathcal{Q}(T)$, then it does not belong to the interior of $\mathcal{Q}(R)$ for $R \neq T$; (2) if $(q^*, (p^0, p^1), \pi^1)$ is a generic point in a maximal proper face $\mathcal{Q}'$ of $\mathcal{Q}(T)$, then it does not belong to $\mathcal{Q}(R)$ for any $R \neq T$ if $(p^0, p^1) \in \partial\mathbb{P}$, and it belongs to the boundary of $\mathcal{Q}(R)$ for exactly one other $R \neq T$ if $(p^0, p^1) \notin \partial\mathbb{P}$; moreover in the latter case it belongs to the interior of a maximal proper face of this $\mathcal{Q}(R)$ as well; (3) given $T, R \in \mathcal{T}$, there exists a finite chain $T = T(0), \ldots, T(k) = R$ such that for each $0 \leq j \leq k-1$, $\mathcal{Q}(T(j)) \cap \mathcal{Q}(T(j+1))$ 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 $x = (q^*, (p^0, p^1), \pi^1) \in \mathcal{Q}(T)$. For each n, choose $q_n^0 \equiv (1-\lambda_n^0)p_n^0 + \lambda_n^0 r_n^0$ in $B_n^*(T)$ and $q_n^1 \equiv \mu_n^0 \tilde{p}_n^0 + \mu_n^1 p_n^1 + \mu_n^2 r_n^1 \in C_n^*(T)$.

We start with (1). Suppose now that x belongs to the interior of $\mathcal{Q}(R)$ for some R. We show that $R = T$. Since x belongs to the interior of $\mathcal{Q}(T)$, we can assume that every strategy in $S_m^0 \setminus T_m^0$ is inferior to q_n^1. Let s_m be a strategy in $S_m^0 \setminus T_m^0$. Since s_m is an inferior reply against q_m^1 compared to the strategies in T_n^0, by Lemma D.2 there exists an information set $h_n \in H_n^0$ that is enabled by s_n where the action chosen by s_n is suboptimal and different from the action chosen by every $t_n \in T_n^0$ that enables h_n. But the posterior belief over the terminal nodes following h_n computed from q_m^1 can be computed from π_m^1. This implies that for any q_m' such that $\bar{\pi}_m^1(q_m') = \pi_m^1$, s_n is an inferior strategy. Therefore, (p_m^1, π_m^1) cannot belong to $C_m(R)$ unless $R_m^0 \subseteq T_m^0$. Moreover, if $R_m^0 \subsetneq T_m^0$, then it cannot belong to the interior of $C_m(R)$, since strategies in $T_m^0 \setminus R_m^0$ are also optimal. Thus, if x belongs to the interior of $\mathcal{Q}(R)$, $R_m^0 = T_m^0$ for each m. If Ψ_m^1 is empty, this implies that $R_m = T_m$. Suppose now that Ψ_m^1 is nonempty. Since x is in the interior we can assume that $\bar{\pi}_m^1(p_m^1) \neq \pi_m^1$ and

that $\bar{\pi}_n(r^1_m)$ is in the interior of Ψ^1_m. Observe that $\bar{\pi}^1_n(r^1_m)$ can be computed uniquely from p^1_m and π^1_m by taking the line segment from $\bar{\pi}^1_m(p^1_m)$ through π^1_m and computing the boundary point of this line. This implies that if x belongs to the interior of $\mathcal{Q}(R)$, then $R_m = (T^0_m, \Psi^1_m)$. Thus $R = T$.

We turn to point (2). Suppose now x belongs to the relative interior of a maximal proper face of $\mathcal{Q}(T)$ and that (p^0, p^1) belongs to $\partial\mathbb{P}$. Then if it belongs to another $\mathcal{Q}(R)$ it cannot be in the interior and must belong to the boundary. The arguments of the previous paragraph apply to show that x does not belong to the interior of a maximal face of $\mathcal{Q}(R)$: indeed, it relied on strategies in $S^0_m \setminus T^0_m$ being inferior to q^1_n, and q^0_n (resp. q^1_n) not belonging to the boundary of $B_n(T)$ (resp. $C_n(T)$). Thus x must belong to a face of dimension at most $\hat{d} - 2$. The set of such points in this maximal proper face of $\mathcal{Q}(T)$ then has dimension at most $\hat{d} - 2$, i.e. it is nongeneric.

Suppose that x belongs to the relative interior of a maximal proper face of $\mathcal{Q}(T)$ but that (p^0, p^1) is in the interior in $\mathbb{P}$. Then for exactly one n, just one of the following hold: (2a) q^*_n belongs to the boundary of $A_n(T)$; (2b) p^0_n belongs to the boundary of $B_n(T)$; (2c) (p^1_n, π^1_n) belongs to the boundary of $C_n(T)$. We start with (2c). By the properties we proved for $C_n(T)$ in Lemma D.7, and since $(p^0, p^1) \notin \partial\mathbb{P}$, either property (ii) or property (iii) of that lemma holds. Under property (ii) x belongs to the boundary of $\mathcal{Q}(R)$ where $R = ((R^0_m, \Phi^1_m), (T_n, \Psi^1_n))$ is defined as follows. If the strategy r^i_m identified there belongs to S^0_m, then R^0_m is the vertex set of the face spanned by r^i_m and T^0_m, while $\Phi^1_m = \Psi^1_n$; if the strategy r^i_m belongs to $S^1_m(T^0_m; \Psi^1_m)$, then $R^0_m = T^0_m$ and Φ^1_m is a face of Π^1_m that has Ψ^1_m as a maximal face with $\pi^1_m(r^1_m) \in \Phi^1_m \setminus \Psi^1_m$. Under property (iii) x belongs to the boundary of $\mathcal{Q}(R)$ where $R = ((T^0_n, \Phi^1_n), (T^0_m, \Psi^1_m))$ where Φ^1_n is the maximal proper face of Ψ^1_n identified there.

Suppose x satisfies (2b). Then by the properties we proved for $B_n(T)$ in Lemma D.6, either property (ii) or property (iii) of that lemma holds. Under property (ii) let $\tilde{R}^1_m$ be the set of strategies in $\tilde{T}^1_m$ that are now best replies against q^0_n. Let $\tilde{\Phi}^1_m$ be the smallest face of Π^1_m that contains Ψ^1_m and the vectors $\pi^1_m(\tilde{r}^1_m)$ for $\tilde{r}^1_m \in \tilde{R}^1_m$. Then the strategies in T^0_m and $S^1_m(\Phi^1_m)$ are equally good replies against q^0_n. 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 Φ^1_m of $\tilde{\Phi}^1_m$ that has Ψ^1_m a maximal proper face, choose a strategy $r_m(\Phi^1_m)$ that maps to a vertex of Φ^1_m that is not contained in Ψ^1_m. 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 $r_m(\Phi^1_m)$ are also best replies has dimension $d(P_n) - d(T^0_m) - d(\Psi^1_m) - 3$ or less. Therefore, for R and R' of the form $((T^0_n, \Psi^1_n), (T^0_m, \Phi^1_m))$ the set of (p^1_n, π^1_n) that lies in the intersection $C_n(T) \cap C_n(R) \cap C_n(R')$ is at most $d^0_n + d^1_n + d_n(\Psi^1_n) - d^*_n(T) - 1$ or less. This implies that generic (p^1_n, π^1_n) in $C_n(T)$

belongs to at most one of these sets. Moreover, if x belongs to $\mathcal{Q}(R)$, then it belongs to the boundary of $\mathcal{Q}(R)$: indeed the point ϕ^1_m in Φ^1_m such that π^1_m is a convex combination of $\bar{\pi}^1_m(p^1_m)$ and ϕ^1_m is uniquely determined, as we argued above; since this point belongs to Ψ^1_m, which is a face of Φ^1_m, x indeed belongs to the boundary of $\mathcal{Q}(R)$ if it belongs to $\mathcal{Q}(R)$. To finish the proof of this case, we now show that x belongs to at least one $\mathcal{Q}(R)$. Take an $\tilde{r}^1_m$ that yields the highest payoff against q^1_n among the strategies in $\tilde{R}^1_m$. If this payoff is higher than the payoff to the strategies in T^0_m, pick a point $\bar{q}^0_n$ in the interior of $B^*_n(T)$ and replace q^1_n with $\tilde{q}^1_n(\varepsilon) \equiv (1-\varepsilon)q^1_n + \varepsilon\bar{q}^0_n$ where ε is the unique number where the strategies T^0_m and $\tilde{r}^1_m$ are equally good replies; then x belongs to some $\mathcal{Q}(R)$ that has $\pi^1_m(\tilde{r}^1_m)$ as an extra vertex. If the payoff to $\tilde{r}^1_m$ is lower, take a point $\bar{q}^0_n$ in P^0_n against which the strategies in T^0_m are equally good and worse than the strategies in $\tilde{R}^1_m$ and repeat the argument to show that x belongs to $\mathcal{Q}(R)$.

Finally, suppose that x satisfies (2a). Let A'_n be a maximal proper face of $A_n(T)$ that contains q^*_n in its interior. Let $\check{R}^1_m$ be the set of strategies $\check{r}_m$ in $S^1_m(T^0_m) \setminus (T^1_m \cup \hat{T}^1_m \cup \bar{T}^1_m)$ that are best replies against all the points in A'_n. Observe that $\check{R}^1_m$ is nonempty if q^*_n belongs to the interior of $P_n(\tilde{T}^0_n)$. Indeed, in this case, the interior of A'_n which is a face of $A_n(T)$ is contained in the interior of $P_n(\tilde{T}^0_n)$, which implies that some strategy in S^1_m is now optimal against every point in this face. Let $\mathcal{R}^0_n$ be the set of subsets R^0_n of T^0_n such $P_n(R^0_n)$ is a maximal proper face of $P_n(T^0_n)$ and $P_n(R^0_n) \cap Q^*_n = A'_n$. Observe that $\mathcal{R}^0_n$ is nonempty if q^*_n belongs to the boundary of $P_n(\tilde{T}^0_n)$.

Let $\check{\Phi}^1_m$ be the smallest face of Π^1_m that contains Ψ^1_m and the vectors $\pi^1_m(\check{r}^1_m)$ for $\check{r}^1_m \in \check{R}^1_m$. For each face Φ^1_m of $\check{\Phi}^1_m$ that has Ψ^1_m as a maximal proper face, choose a strategy $\check{r}^1_m(\Phi^1_m)$ that maps to a vertex of Φ^1_m that is not contained in Ψ^1_m. Take R and R' of the form $((T^0_n, \Psi^1_n), (T^0_m, \Phi^1_m))$ where Φ^1_m has Ψ^1_m as a maximal face. Since A'_n is a face of $A_n(T)$, $\tilde{d}^*_n(R) = \tilde{d}^*_n(R') = \tilde{d}^*_n(T) - 1$. Since the strategies in $\check{R}^1_m$ are inferior replies to points in the interior of $A_n(T)$, $B^*_n(R)$ if nonempty has dimension $d^*_n(T)-1$. Therefore, if $B_n(R) \neq B_n(R')$, their intersection with $B_n(T)$ has codimension 1 in $B_n(T)$ and a generic x cannot belong to two of these sets at once. On the other hand, if $B_n(R) = B_n(R')$ then an argument similar to that under case (2b) shows that generic $(p^1_n, \pi^1_n) \in C_n(T)$ can belong to at most one of these sets, $C_n(R)$ and $C_n(R')$. Hence a generic x belongs to at most one of these sets. Likewise, for R of the form $((R^0_n, \Psi^1_n), (T^0_m, \Psi^1_m))$ with $R^0_n \in \mathcal{R}^0_n$ and R' of the form $((R'^0_n, \Psi^1_n), (T^0_m, \Psi^1_m))$ or $((T^0_n, \Psi^1_n), (T^0_m, \Phi^1_m))$, the intersection of $B_n(T)$ with $B_n(R)$ and $B_n(R')$ has codimension at least one. Thus x belongs to at most one set $\mathcal{Q}(R)$. To finish the proof of this part, we show that it belongs to at least one such set.

Suppose that the interior of A'_n is contained in the interior of $P_n(\tilde{T}^0_n)$. Let $\check{r}^1_m(\Phi^1_m)$ be a strategy that is a lexicographic best reply to (q^0_n, q^1_n) among the strategies in this class. If $\check{r}^1_m(\Phi^1_m)$ is a lexicographic weakly better (resp. strictly inferior) reply against (q^0_n, q^1_n) we

choose a point $\bar{q}_n^0$ in the interior of $P_n(\tilde{T}_n^0)$ against which the strategy $\check{r}_m^1(\Phi_m^1)$ is at least as good as the other points in this class and inferior (resp. superior) to strategies in T_m^0. The strategies in T_m^0 and $S_m(T_m^0;\Phi_m^1)$ are now equally good replies against some average of q_n^0 and $\bar{q}_n^0$, as well as some average of q_n^1 and $\bar{q}_n^0$. The point x then belongs to $((T_m^0,\Phi_m^1),(T_n^0,\Psi_n^1))$.

Suppose now that q_n^* belongs to the boundary of $P_n(\tilde{T}_n^0)$, then $\mathcal{R}_n^0$ is nonempty. There exists $\check{R}_n^0$ in $\mathcal{R}_n^0$ and a point $\tilde{q}_n^0 \in B_n(T)$ that is a convex combination of p_n^0 and some point in $P_n(\check{R}_n^0)$. If the strategies in $\check{R}_m^1$ are weakly inferior against $\tilde{q}_n^0$ to the strategies in T_m, then x belongs $Q((\check{R}_n^0,\Psi_n^1),(T_m^0,\Psi_m^1))$. Otherwise, as above, we can replace $\tilde{q}_n^0$ by a convex combination $\bar{q}_n^0$ of $\tilde{q}_n^0$ with a point in the interior of $A_n(T)$ and replace q_n^1 with an a point $\bar{q}_n^1$ that is a convex combination of q_n^1 with either $\tilde{q}_n^0$ or a point in the interior of $A_n(T)$ depending on whether the strategy $\check{r}_m^1(\Phi_m^1)$ that is superior to the strategies in T_m against $\tilde{q}_n^0$ is inferior or weakly superior in comparison against q_n^1. The points $\bar{q}_n^0$ and $\bar{q}_n^1$ belong to $B_n(T')$ and $C_n(T')$ respectively, where $T'=((T_n,\Psi_n^1),(T_m,\Phi_m^1)$ and thus x belongs to $Q(T')$.

We turn now to (3). Given $\mathcal{Q}(T)$ and $\mathcal{Q}(R)$ 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 $\tilde{T}_n^0\neq T_n^0$ for some n, let $T=T(0),\ldots,T(k)$ be a sequence where for each $j>0$, $T(j)=((T_n^0(j),\Psi_n^1),(T_m^0(j),\Psi_n^1))$ with $P_n(T_n^0(j))\times P_m(T_m^0(j))$ being a maximal proper face of $P_n(T_n^0(j-1))\times P_m(T_m^0(j-1))$, and $T_n(k)=\tilde{T}_n^0$ for each n. This sequence generates a sequence of polyhedra $\mathcal{Q}(T)=\mathcal{Q}(T(0)),\ldots,\mathcal{Q}(T(k))$ where for each $j>0$, the intersection of $\mathcal{Q}(T(j))$ with $\mathcal{Q}(T(j-1))$ is contained in a maximal proper face of each and has a nonempty interior. After this operation we have $\mathcal{Q}(\tilde{T}(k))$, where $\tilde{T}(k)=((\tilde{T}_n^0,\Psi_n^1),(\tilde{T}_m^0,\Psi_m^1))$. In case Ψ_n^1 is nonempty for some n, let $\Psi^1=(\Psi_n^1(0),\Psi_m^1(0)),\ldots,(\Psi_n^1(l),\Psi_m^1(l))=(\emptyset,\emptyset)$ be a sequence such that for each $1\leqslant j\leq l$, $\Psi_m^1(j)\times\Psi_m^1(j)$ is a maximal proper face of $\Psi_n^1(j-1)\times\Psi_m^1(j-1)$, and $\Psi^1(l)=\emptyset$. This way we can connect $\mathcal{Q}(\tilde{T}(k))$ with $\mathcal{Q}(\tilde{T})$ 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_n^* is connected for each n, there exists a sequence $\tilde{T}^0=(S_n^0(0),S_m^0(0)),\ldots,(S_n^0(l),S_m^0(l))=\tilde{R}_n^0$ where for each $1\leqslant j\leqslant l$, 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_n^* intersects the interior of the set $P_n(S_n^0(j))$. This generates a sequence $\mathcal{Q}^0,\ldots\mathcal{Q}^j$, where $\mathcal{Q}^j=((S_n^0(j),\emptyset),(S_m^0(j),\emptyset))$. Thus we have constructed a sequence of sets in $\mathcal{T}$ that connect T and R. □

This concludes the proof of the first statement in the theorem. Next we prove the second statement, invoking now the original definition of a stable set in Mertens [32].

Lemma D.10. *Q^* is a stable set if and only if the projection map $\Psi:(\mathcal{Q},\partial\mathcal{Q})\to(\mathbb{P},\partial\mathbb{P})$ is essential.*

Proof of Lemma. Let $Y = [0,1] \times P$. For each $0 < \varepsilon \leqslant 1$, let $Y_\varepsilon = [0,\varepsilon] \times P$ and let ∂Y_ε be the boundary of Y_ε. 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 $\mathcal{E}$ be the closure of the set of (ε, p, q) such that $(\varepsilon, p) \in Y_1 \setminus \partial Y_1$ and q is a perturbed equilibrium of $G(\varepsilon, p)$. Let θ be the projection map from $\mathcal{E}$. For each subset E of $\mathcal{E}$ and each $0 < \varepsilon$, let $(E_\varepsilon, \partial E_\varepsilon)$ be $E \cap \theta^{-1}(P_\varepsilon, \partial P_\varepsilon)$.

In [15] we show that there exists $0 < \bar{\varepsilon} \leqslant 1$ and a finite number of subsets $E^1, \ldots, E^K$ of $\mathcal{E}$ such that for each $0 < \varepsilon \leqslant \bar{\varepsilon}$: (i) $(E^k_\varepsilon, \partial E^k_\varepsilon)$ is a pseudomanifold (in fact an orientable semi-algebraic homology manifold) of dimension $d(P)+1$ for each k; (ii) $E^k_\varepsilon \cap E^j_\varepsilon \subset \theta^{-1}(\partial P_1)$ for $k \neq j$; (iii) $\cup_k E^k_\varepsilon = \mathcal{E}_\varepsilon$. We will assume that $\bar{\varepsilon}$ is small enough such that for each player n, and each $(\varepsilon, p, q) \in E_{\bar{\varepsilon}}$, if a strategy s_n is optimal against a strategy q in Γ, then it is optimal against some point in $\bar{Q}^*_n$, the component of equilibria containing Q^*_n.

One could define the set of perturbations for the normal form of the game and consider the graph of the equilibria over this space. In [11] we show that there exists a neighborhood of Σ^* that is disjoint from the other components of Γ and an $\varepsilon > 0$ such that the set of ε-perfect equilibria in this neighborhood (viewed as points in the graph of equilibria) is connected. The corresponding set of ε-perfect equilibria in enabling strategies is therefore connected. Thus there exists some k such that $E^k_0 = \{\, 0 \,\} \times Q^*$ and for each $j \neq k$, $E^j_0 \cap (\{\, 0 \,\} \times Q^*)$ is empty. For simplicity in notation we refer to this E^k as simply E. According to Mertens' [32] definition, Q^* is stable iff the projection θ from E_ε to Y_ε is cohomologically essential for some (and then all smaller) $0 < \varepsilon \leqslant \bar{\varepsilon}$. Moreover, since E_ε is a pseudomanifold, θ is cohomologically essential iff it is essential in homotopy [33, Theorem, Section 4E]. By [15, Lemmas A.3, A.4], this is equivalent to saying that Ψ is essential in the sense we have used it in Section 5.

It is now sufficient to prove that Ψ is essential iff the projection θ from E_ε to Y_ε is essential for all small ε. For each n, $\tilde{P}_n \equiv [0,1] \times \mathbb{P}_n$ and define $\chi_n : \tilde{P}_n \to P_n$ by $\chi_n(\lambda_n, p^0_n, p^1_n) = (1-\lambda_n)p^0_n + \lambda_n p^1_n$. Then we have that $((\tilde{P}_n, \partial\tilde{P}_n), (P_n, \partial P_n), \chi_n)$ is a ball-bundle. Let χ be the product map $\chi_1 \times \chi_2$; then $((\tilde{P}, \partial\tilde{P}), (P, \partial P), \chi)$ is a ball-bundle too. Let $\hat{Y}_\varepsilon = [0,\varepsilon] \times \tilde{P}$. Then χ induces a map $h_Y : \hat{Y}_1 \to Y_1$ by $h(\varepsilon, \lambda, p^0, p^1) = (\varepsilon, \chi(\lambda, p^0, p^1))$. Now $((\hat{Y}_\varepsilon, \partial\hat{Y}_\varepsilon), (Y_\varepsilon, \partial Y_\varepsilon), h_Y)$ is a ball-bundle. Let $\tilde{E}$ be the set of all $((\varepsilon, \lambda, p^0, p^1), q) \in \hat{Y}_1 \times P$ such that $h_E(\varepsilon, \lambda, p^0, p^1, q) \equiv (\varepsilon, \chi(\lambda, p^0, p^1), q) \in E$. Then also $((\tilde{E}_\varepsilon, \partial\tilde{E}_\varepsilon), (E_\varepsilon, \partial E_\varepsilon), h_E)$ is a ball-bundle. Moreover, letting $\tilde{\theta}$ be the projection from $\tilde{E}$ to $\mathbb{P}$, we have that $h_Y \circ \tilde{\theta} = \theta \circ h_E$. Therefore, by the Thom Isomorphism Theorem, θ is essential iff $\tilde{\theta}$ is; cf. [33, Appendix IV.3].

Let $\hat{E}$ 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 $\hat{E}$ to $\tilde{E}$ induces an isomorphism of their cohomology groups. Let $\hat{\theta}$ be the projection from $\hat{E}$ to $\hat{Y}$. 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 $D^0_1, \ldots, D^0_l$, and for each D^0_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. Since the sets D_i are semi-algebraic, if necessary by decomposing them into smaller sets, we can assume that the closure of each of these sets is homeomorphic to a simplex. There now exists an i, say 1, such that the closure D_i of D^0_i is homeomorphic to a 3-simplex and contains $[0, \tilde{\varepsilon}] \times \{\,(0,0)\,\}$ for some $\tilde{\varepsilon} < \bar{\varepsilon}$.

Let $(\hat{Y}(D_1), \partial\hat{Y}(D_1)) \equiv (D_1, \partial D_1) \times (\mathbb{P}, \partial\mathbb{P})$ and let $\hat{E}(D_1)$ be the closure of the inverse image of $D_1 \times F_1$ 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 $(\hat{E}_\varepsilon(D_1), \partial\hat{E}_\varepsilon(D_1))$ is an orientable homology manifold (and hence a pseudomanifold) for all small ε, where E_ε is the inverse image under $\theta E_\varepsilon(D_1)$ of the points $(\varepsilon', \lambda) in D_1$ with $e' \leqslant \varepsilon$. The set $\hat{E}$ was constructed from the set E, which is an orientable homology manifold, by constructions involving ball bundles and homeomorphisms. Thus, $\hat{E}$ is an orientable homology manifold. So our claim is proved if we show that $\hat{E}(D_1) \setminus \partial\hat{E}(D_1)$ is path-connected. There exist $0 < \hat{\varepsilon}' < \tilde{\varepsilon}$ and integers $r_n > 1$ for each n such that $(\varepsilon, \lambda) \in D_1$ if $0 < \varepsilon < \hat{\varepsilon}'$ and $0 \leq \lambda_n \leq \varepsilon^{r_n - 1}$. Now given $0 < \hat{\varepsilon} \leqslant \hat{\varepsilon}'$ and given two points $x(0)$ and $x(1)$ in $\hat{E}_{\hat{\varepsilon}}(D_1) \setminus \partial\hat{E}_{\hat{\varepsilon}}(D_1)$, connect them by a semi-algebraic curve $x(t') = (\varepsilon(t'), \lambda(t'), p^0(t'), p^1(t'), q(t'), \pi^1(q(t')))$ in $\hat{E}_{\bar{\varepsilon}} \setminus \partial\hat{E}_{\bar{\varepsilon}}$ 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 $q(t')$. The correspondence from $[0,1]$ to $\bar{Q}^*$ that assigns to each t' the set of q^* such that $r(t')$ is a best reply to q^* is a nonempty, compact convex valued, and upper semi-continuous correspondence. Therefore, there exists a path $((t'(t), q^*(t))$ in the graph of this correspondence with $t'(0) = 0$ and $t'(1) = 1$. We will now view the path $x(t')$ as the path $x(t) \equiv x(t'(t))$.

Choose a positive ε such that $\varepsilon + \varepsilon^{r_n} < \hat{\varepsilon}$. For each n, modify $x_n(t)$ to the vector

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

where

$$\tilde{\lambda}_n(t) = \frac{\varepsilon^{r_n}\varepsilon(t)}{\varepsilon + \varepsilon^{r_n}}\lambda_n(t)\,,$$

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

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

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 $\hat{x}(t; s)$ defined as follows. For $s \in [0,1]$, let $k_n(s) = \min(1, 2s)r_n$ and

then:

$$\hat{x}_n(t;s) = ((2s-1)^+\varepsilon + \varepsilon^{k_n(s)}\varepsilon(t), \hat{\lambda}(t;s), p^0(t;s), p^1(t), \hat{q}(t;s), \pi^1(t))\,,$$

where

$$\hat{\lambda}_n(t) = \frac{\varepsilon^{k_n(s)}\varepsilon(t)\lambda_n(t)}{(2s-1)^+\varepsilon + \varepsilon^{k_n(s)}\varepsilon(t)}\,,$$

$$\tilde{p}^0_n(t;s) = \frac{(2s-1)^+\varepsilon q^*_n(t) + \varepsilon^{r_n}\varepsilon(t)(1-\lambda_n(t))p^0_n(t)}{(2s-1)^+\varepsilon + \varepsilon^{k_n(s)}\varepsilon(t)(1-\lambda_n(t))}\,,$$

$$\tilde{q}_n(t;s) = (1-(2s-1)^+\varepsilon-\varepsilon^{k_n(s)}\varepsilon(t))r_n(t)+(2s-1)^+\varepsilon q^*_n(t)+\varepsilon^{k_n(s)}\varepsilon(t)((1-\lambda_n(t))p^0_n(t)+\lambda_n(t)p^1(t))\,.$$

Thus $\hat{E}(D_1)\setminus\partial\hat{E}(D_1)$ is connected and hence an orientable homology manifold of dimension $d(\mathbb{P})+3$. $\hat{Y}(D_1)$ is a full-dimensional subset of $Y_{\bar{\varepsilon}}$. Therefore $\hat{\theta}$ is essential iff $\hat{\theta}(D_1)$ is essential.

Since $(\hat{E}_{\hat{\varepsilon}}(D_1), \partial\hat{E}_{\hat{\varepsilon}}(D_1))$ is an orientable homology manifold, $(F_1, \partial F_1)$ is now an orientable homology manifold of dimension $d(\mathbb{P})$ and hence an orientable pseudomanifold. Moreover, for each $(\varepsilon,\lambda)\in D^0_1$, with $0<\varepsilon\leqslant\hat{\varepsilon}$, letting $(\hat{E}_{\varepsilon,\lambda}, \partial\hat{E}_{\varepsilon,\lambda}) \equiv h_1^{-1}(\{\,\varepsilon,\lambda\,\}\times(F_1,\partial F_1))$ we have that $\hat{\theta}$ is essential iff $\hat{\theta}_{\varepsilon,\lambda}$, the projection map $(\hat{E}_{\varepsilon,\lambda}, \partial\hat{E}_{\varepsilon,\lambda})\to(\mathbb{P},\partial\mathbb{P})$, is essential.

Let L be the set of $(\varepsilon,\lambda)\in D_1$ such that $\lambda_n=\varepsilon^r$ for some $r\geqslant r_n$ for each n. Let $\hat{E}(L)$ be the closure of the inverse image of $L\setminus\{\,(0,0)\,\}$ under h_1. Let $\partial\hat{E}(L)$ be the inverse image of $L\times\partial\mathbb{P}$ under the projection map $\theta(L)$ from $\hat{E}(L)$ to $L\times\mathbb{P}$. For each $(\varepsilon,\lambda)\in L$, $(E_{\varepsilon,\lambda}, \partial E_{\varepsilon,\lambda})$ is $(\theta(L))^{-1}(\{\,(\varepsilon,\lambda)\,\}\times(\mathbb{P},\partial\mathbb{P}))$. Our next objective is to show that $(\mathcal{Q},\partial\mathcal{Q})$ equals the set of $(q^*,(p^0,p^1),\pi^1)$ such that $(0,0,(p^0,p^1),q^*,\pi^1)$ belongs to $E_{0,0}(L)$. Given $(0,0,(p^0,p^1),q^*,\pi^1)$ in $\hat{E}(L)$ there exists a sequence $(\varepsilon(k),\lambda(k),(p^0,p^1)(k),q(k),\pi^1(k))$ in $\hat{E}(L)\setminus E_{0,0}$ converging to it. For each n and k, we can express $q_n(k)$ as $(1-\varepsilon(k))((1-\mu^1_n)q^0_n(k)+\mu^1_n r^1_n(k))+\varepsilon(k)(\lambda(k)p^0_n(k)+(1-\lambda(k))p^1_n(k))$, where μ^1_n and $\lambda(k)$ converge to zero, $q^0_n(k)$ belongs to P^0_n and converges to q^*_n, and $r^1_n(k)$ belongs to P^1_n. Also $q^0_n(k)$ and $r^1_n(k)$ if $\mu^1_n>0$ are best replies to $q(k)$ for all k. By going to a subsequence, the faces to whose interior $q^0_n(k)$ and $r^0_n(k)$ belong are constant for all k. The sequence generates for each n an LPS $\Lambda_n=(\bar{q}^0_n,\ldots,\bar{q}^l_n)$ where $\bar{q}^0_n=q^*_n$. As in the proof of Theorem 5.1, there exists a level l^i_n for each $i=0,1$ that is expressible as a convex combination of p^0_n and another strategy. Since $\lambda(k)$ converges to zero, $l^0_n<l^1_n$. As in the proof of Theorem 5.1, we can prove that $\bar{q}^l_n$ belongs to $\bar{Q}^*_n$ for each $l<l^0_n$ and if we express $\bar{q}^{l^0_n}_n=\nu^0_n p^0_n+(1-\nu^0_n)r^0_n$, then the strategies $\bar{q}^l_n$ for $l<l^0_n$ and r^0_n if $\nu^0<1$ are lexicographic best replies against Λ_m. Likewise, if we express $\bar{q}^{l^1_n}_n$ as $\bar{\nu}^0_n\tilde{p}^0_n+\bar{\nu}^1_n p^1_n+\bar{\nu}^2_n r^2_n$, then the strategy r^2_n if $\bar{\nu}^2_n>0$ is a lexicographic best reply against Λ_m. As in the proof of Theorem 5.1 in Section 5.8, 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 $\mathcal{Q}$.

Given $(q^*,(p^0,p^1),\pi^1)$ in $\mathcal{Q}$, it belongs to $\mathcal{Q}(T)$ for some T. There exist $q^0_n=(1-\lambda^0_n)p^0_n+\lambda^0_n r^0_n\in B^*_n(T)$ and $q^1_n=\mu^0_n\tilde{p}^0_n+\mu^1_n p^1_n+\mu^2_n r^2_n\in C^*_n(T)$. As in the proof of Lemma D.8, we can

assume that the strategies in T_m are best replies against the LPS (q_n^*, q_n^0, q_n^1). For all small ε, choose $\alpha_n(\varepsilon)$ such that $\varepsilon = \varepsilon^{r+1}(1/\mu_n^1)(\mu_n^0+\mu_n^1)+\alpha_n(\varepsilon)(1-\lambda_n^0)$, where r satisfies the property in the first line of the previous paragraph. Then for all small ε, $(\varepsilon, (\varepsilon^r, \varepsilon^r), (p^0(\varepsilon), p^1), q(\varepsilon))$ belongs to $\hat{E}(L)$, where for each n:

$$p_n^0(\varepsilon) = \frac{\varepsilon^{r+1}(\mu_n^0/\mu_n^1)\tilde{p}_n^0 + \alpha_n(\varepsilon)(1-\lambda_n^0)p_n^0}{\varepsilon^{r+1}(\mu_n^0/\mu_n^1) + \alpha_n(\varepsilon)(1-\lambda_n^0)}$$

and $q_n(\varepsilon) = (1-\alpha_n(\varepsilon)-\varepsilon^{r+1}(1/\mu_n^1))q_n^* + \alpha_n(\varepsilon)q^0 + \varepsilon^{r+1}(1/\mu_n^1)q_n^1$.

For each (ε, λ), $(E_{\varepsilon,\lambda}, \partial E_{\varepsilon,\lambda})$ 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 $(\mathcal{Q}, \partial\mathcal{Q})$. 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 smaller) $(\varepsilon,\lambda) \in L$. The projection map $\hat{\theta}_{0,0}$ is just the map Ψ. Hence Q^* is stable iff Ψ is essential. □

In case S_n^1 is empty, the construction is modified as follows. We can omit the sets $C_n(T)$ and $B_m(T)$ from the description of $\mathcal{Q}(T)$. In the last lemma above, the vector λ is now just a number, one for player m. The simplex D constructed there is 2-dimensional and contains a curve L of the form $\lambda = \varepsilon^r$. The rest of the proof is essentially the same.

This concludes the proof of the Theorem. □

Appendix E. Construction of the Map g

Here we construct the continuous map $g : \mathcal{Q} \to \mathbb{P}$ and constant $\alpha > 0$, such that $\|g(x) - \Psi(x)\| \leqslant \alpha$ for some $x \in \mathcal{Q}$ only if Ψ is essential and $x \in V(x^*)$, that were invoked in Section 5.3.

If the projection map Ψ is inessential then there exists a continuous map $g : \mathcal{Q} \to \mathbb{P}$ that has no point of coincidence with g. Therefore, there exists $\alpha > 0$ such that $\|\Psi(x) - g(x)\| > \alpha$ for all $x \in \mathcal{Q}$.

Suppose now that Ψ is essential. Since $\mathcal{Q}$ is a pseudo-manifold of the same dimension as $\mathbb{P}$, essentiality of Ψ in the sense we have defined it in Section 5 is equivalent to essentiality of Ψ in cohomology [33, Theorem, Section 4][15, Lemmas A.3, A.4], i.e. $\Psi^* : H^d(\mathbb{P}, \partial\mathbb{P}) \to H^d(\mathcal{Q}, \partial\mathcal{Q})$ is nonzero. Moreover, letting d be the dimension of $\mathbb{P}$ and $\Psi_{\partial\mathcal{Q}}$ the restriction of Ψ to $\partial\mathcal{Q}$, the degree of Ψ equals $\delta^* \circ \Psi_{\partial\mathcal{Q}}^*(1)$, where 1 is the generator of $H^{d-1}(\partial\mathcal{Q}) \approx \mathbb{Z}$ and δ^* is the coboundary operator.

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 $\lambda p + (1-\lambda)\bar{p}$ for $\lambda < 0$. ι is a homeomorphism without a fixed point. Let $g_{\partial\mathcal{Q}} : \partial\mathcal{Q} \to \partial\mathbb{P}$ be the function $\iota \circ \Psi_{\partial\mathcal{Q}}$, where $\Psi_{\partial\mathcal{Q}}$ is the restriction of Ψ to $\partial\mathcal{Q}$.

Then $g_{\partial\mathcal{Q}}$ has no point of coincidence with $\Psi_{\partial\mathcal{Q}}$, i.e. for each $x \in \partial\mathcal{Q}$, $g_{\partial\mathcal{Q}}(x) \neq \Psi(x)$. Also, since ι is a homeomorphism, $\delta^* \circ g_{\partial\mathcal{Q}}$ is nonzero.

Construct a continuous map $g_{\partial V(x^*)}$ from $\partial V(x^*)$ to $\partial\mathbb{P}$ such that $\delta^* \circ g^*_{\partial\mathcal{Q}\cup\partial V(x^*)}(1) = 0$ in $H^d(Q \setminus (V(x^*) \setminus \partial V(x^*)), \partial\mathcal{Q} \cup \partial V(x^*))$. By the Hopf Extension Theorem [49, Corollary 8.1.18], the two maps $g_{\partial\mathcal{Q}}$ and $g_{\partial V(x^*)}$ can be extended to a continuous map from $\mathcal{Q} \setminus (V(x^*) \setminus \partial V(x^*))$ to $\mathbb{P}$; furthermore, by mapping points in $V(x^*)$ to $\mathbb{P}$ in a way that extends $g_{\partial V(X^*)}$, we obtain a map $g : \mathcal{Q} \to \mathbb{P}$ such that all its points of coincidence with Ψ, of which there is at least one, are contained in $V(x^*) \setminus \partial V(x^*)$. There now exists $\alpha > 0$ such that $\|\Psi(x) - g(x)\| > \alpha$ for all $x \notin V(x^*) \setminus \partial V(x^*)$.

References

[1] Aumann, R. (1974): Subjectivity and Correlation in Randomized Strategies, Journal of Mathematical Economics, 1, 67-96.

[2] Blume, L., A. Brandenburger, and E. Dekel (1991): Lexicographic Probabilities and Equilibrium Refinements, Econometrica, 59, 81-98.

[3] Blume, L. and W. Zame (1994): The Algebraic Geometry of Perfect and Sequential Equilibrium, Econometrica, 62, 783-794.

[4] Brandenburger, A., and Friedenberg, A. (2011): Are Admissibility and Backward Induction Consistent?, New York University.

[5] Cho, I., and D. Kreps (1987): Signalling Games and Stable Equilibria, Quarterly Journal of Economics, 102, 179-221.

[6] Ely, J., and M. Peski (2006): Hierarchies of Belief and Interim Rationalizability, Theoretical Economics 1, 1965.

[7] Govindan, S., and T. Klumpp (2002): Perfect Equilibrium and Lexicographic Beliefs, International Journal of Game Theory, 31, 229-243.

[8] Govindan, S., and J.-F. Mertens (2003): An Equivalent Definition of Stable Equilibria, International Journal of Game Theory, 3, 339-357.

[9] Govindan, S., and R. Wilson (2001): Direct Proofs of Generic Finiteness of Nash Equilibrium Outcomes, Econometrica, 69, 765-769.

[10] Govindan, S., and R. Wilson (2002): Structure Theorems for Game Trees, Proceedings of the National Academy of Sciences USA, 99, 9077-9080. URL: www.pnas.org/cgi/reprint/99/13/9077.pdf

[11] Govindan, S., and R. Wilson (2002): Maximal Stable Sets of Two-Player Games, International Journal of Game Theory, 30, 557-566.

[12] Govindan, S., and R. Wilson (2005): Essential Equilibria, Proceedings of the National Academy of Sciences USA, 102, 15706-15711. URL: www.pnas.org/cgi/reprint/102/43/15706.pdf.

[13] Govindan, S., and R. Wilson (2006): Sufficient Conditions for Stable Equilibria, Theoretical Economics, 1, 167-206.

[14] Govindan, S., and R. Wilson (2008): "Nash Equilibrium, refinements of," entry in S. Durlauf and L. Blume (eds.), The New Palgrave Dictionary of Economics, 2nd Edition. New York: Palgrave Macmillan. URL: gsbapps.stanford.edu/researchpapers/library/RP1897.pdf

[15] Govindan, S., and R. Wilson (2008): Metastable Equilibria, Mathematics of Operations Research, 33, 787-820.

[16] Govindan, S., and R. Wilson (2008): Axiomatic Theory of Equilibrium Selection in Signaling Games with Generic Payoffs, Research Paper 2000, Stanford Business School, Stanford CA.
URL: gsbapps.stanford.edu/researchpapers/library/RP2000.pdf

[17] Govindan, S., and R. Wilson (2009): Axiomatic Theory of Equilibrium Selection in Games with Games with Two Players, Perfect Information, and Generic Payoffs, Research Paper 2008, Stanford Business School, Stanford CA. URL: gsbapps.stanford.edu/researchpapers/library/RP2008.pdf
[18] Govindan, S., and R. Wilson (2009): On Forward Induction, Econometrica, 77, 1-28.
[19] Harsanyi, J. (1967, 1968): Games with Incomplete Information Played by "Bayesian" Players, Parts I, II, III, Management Science, 14, 159-182, 320-334, 486-502.
[20] Hillas, J. (1990): On the Definition of the Strategic Stability of Equilibria, Econometrica, 58, 1365-1390.
[21] Hillas, J., T. Kao, and A. Schiff (2002): A Semi-Algebraic Proof of the Generic Equivalence of Quasi-Perfect and Sequential Equilibria. University of Auckland, mimeo.
[22] Hillas, J., and E. Kohlberg (2002): Conceptual Foundations of Strategic Equilibrium, in R. Aumann and S. Hart (eds.), Handbook of Game Theory with Economic Applications, III, Chapter 42, 1597-1663. New York: Elsevier.
[23] Kohlberg, E. (1990): Refinement of Nash Equilibrium: The Main Ideas, in T. Ichiishi, A. Neyman, and Y. Tauman (eds.), Game Theory and Applications. San Diego: Academic Press.
[24] Kohlberg, E., and J.-F. Mertens (1986): On the Strategic Stability of Equilibria, Econometrica, 54, 1003-1037.
[25] Koller, D., and N. Megiddo (1992): The Complexity of Two-Person Zero-Sum Games in Extensive Form, Games and Economic Behavior, 4, 528-552.
[26] Kreps, D., and R. Wilson (1982): Sequential Equilibria, Econometrica, 50, 863-894.
[27] Kuhn, H. (1953): Extensive Games and the Problem of Information, in H. Kuhn and A. Tucker (eds.), Contributions to the Theory of Games, II, 193-216. Princeton: Princeton University Press. Reprinted in H. Kuhn (ed.), Classics in Game Theory, Princeton University Press, Princeton, New Jersey, 1997.
[28] Liu, Q. (2009): On Redundant Types and Bayesian Formulation of Incomplete Information, J. Economic Theory, 144: 2115-2145.
[29] Luce, R.D., and H. Raiffa (1957): Games and Decisions. New York: John Wiley and Sons, Inc.
[30] Mailath, G., L. Samuelson, and J. Swinkels (1993): Extensive Form Reasoning in Normal Form Games, Econometrica, 61, 273-302.
[31] McLennan, A. (1985): Justifiable Beliefs in Sequential Equilibrium, Econometrica, 53, 889-904.
[32] Mertens, J.-F. (1989): Stable Equilibria–A Reformulation, Part I: Definition and Basic Properties, Mathematics of Operations Research, 14, 575-625.
[33] Mertens, J.-F. (1991): Stable Equilibria–A Reformulation, Part II: Discussion of the Definition and Further Results, Mathematics of Operations Research, 16, 694-753.
[34] Mertens, J.-F. (1992a): The Small Worlds Axiom for Stable Equilibria, Games and Economic Behavior, 4, 553-564.
[35] Mertens, J-.F. (1992b): Essential Maps and Manifolds, Proceeedings American Mathematical Society, 115, 513-525.
[36] Mertens, J.-F. (1993): Ordinality in Non Cooperative Games, International Journal of Game Theory, 32, 387-430.
[37] Myerson, R. (1978): Refinement of the Nash equilibrium concept. International Journal of Game Theory, 7, 73-80.
[38] Nash, J. (1950): Equilibrium Points in N-Person Games, Proceedings of the National Academy of Sciences USA, 36, 48-49.
[39] Nash, J. (1951): Non-Cooperative Games, Annals of Mathematics, 54, 286-295.
[40] Norde, H., J. Potters, H. Reijnierse, D. Vermeulen (1996): Equilibrium Selection and Consistency, Games and Economic Behavior, 12, 219-225.
[41] Pimienta, C., and J. Shen (2010): On the Equivalence between (Quasi-) Perfect and Sequential Equilibria, University of New South Wales, mimeo.
[42] Polak, B. (1999), Epistemic Conditions for Nash Equilibrium and Common Knowledge of Rationality, Econometrica, 67, 673-676.
[43] Reny, P. (1992): Backward Induction, Normal Form Perfection and Explicable Equilibria, Econometrica, 60, 627-649.

[44] Reny, P. (1992): Rationality in Extensive-Form Games, Journal of Economic Perspectives, 6, 103-118.
[45] Reny, P. (1993): Common Belief and the Theory of Games with Perfect Information, Journal of Economic Theory, 59, 257-274.
[46] Sadzik, T. (2010): Beliefs Revealed in Bayesian-Nash Equilibrium, New York University.
[47] Selten, R. (1965): Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragetragheit. Zeitschrift fur die gesamte Staatswissenschaft, 121, 301-324 and 667-689.
[48] Selten, R. (1975): Reëxamination of the perfectness concept for equilibrium points in extensive games. International Journal of Game Theory, 4, 25-55.
[49] Spanier, E. (1966): Algebraic Topology, New York: McGraw-Hill. Reprinted New York: Springer-Verlag, 1989.
[50] van Damme, E. (1984): A Relation between Perfect Equilibria in Extensive Form Games and Proper Equilibria in Normal Form Games, International Journal of Game Theory, 13, 1-13.
[51] van Damme, E. (2002): Strategic Equilibrium, in R. Aumann and S. Hart (eds.), Handbook of Game Theory, III, 1523-1596. New York: Elsevier.

Department of Economics, University of Rochester, Rochester, NY 14627, USA.
E-mail address: srihari-govindan@uiowa.edu

Stanford Business School, Stanford, CA 94305-7298, USA.
E-mail address: rwilson@stanford.edu