The refined best-response correspondence in normal form games

Dieter Balkenborg[*], Josef Hofbauer[†], and Christoph Kuzmics[‡]

June 3, 2009

Abstract

This paper provides an in-depth study of the (most) refined best-reply correspondence introduced by Balkenborg, Hofbauer, and Kuzmics (2009a). We study notions of strict and weak dominance most appropriate to it, its fixed points, and rationalizability based on it, and how these concepts are related to well-known concepts such as, among others, Selten's (1975) trembling-hand perfection, Kalai and Samet's (1984) persistent equilibria, and various refinements of Bernheim's (1984) and Pearce's (1984) notion of rationalizability.

Keywords: best-response correspondence, persistent equilibria, Nash equilibrium refinements, strict and weak dominance, strategic stability
JEL codes: C62, C72, C73

[*]Department of Economics, School of Business and Economics, University of Exeter, D.G.Balkenborg@exeter.ac.uk

[†]Department of Mathematics, University of Vienna, Josef.Hofbauer@univie.ac.at

[‡]Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, c-kuzmics@kellogg.northwestern.edu.

We would like to thank Carlos Alos-Ferrer, Pierpaolo Battigalli, Eddie Dekel, Amanda Friedenberg, Drew Fudenberg, Klaus Ritzberger, Karl Schlag, Mark Voorneveld, and Jörgen Weibull for helpful comments and suggestions.

1 Introduction

The minimal refined best-reply correspondence was introduced by Balkenborg, Hofbauer, and Kuzmics (2009a) in an effort to find the smallest face of the polyhedron of mixed strategy profiles that can be termed evolutionary stable under some reasonable dynamics. The dynamics found is a differential inclusion based on the, so termed, minimal refined best-reply correspondence. In this paper we study this correspondence in detail for all normal form games with a mild restriction.

The paper proceeds as follows. Section 2 defines the very general class of games we study and defines the minimal refined best-reply correspondence. Section 3 defines notions of dominance, termed inferiority, more appropriate for the refined best-reply correspondence. This section also includes two important Theorems, one on the (normal-form) generic equivalence of the best-reply correspondence and the minimal refined best-reply correspondence, and the other a full characterization of the minimal best-reply correspondence in two-player games. The somewhat difficult proves are relegated to appendices. Section 4 places fixed-points of the best-reply correspondence into the literature by demonstrating its relationship to Selten's (1975) trembling-hand perfect equilibria and Kalai and Samet's (1984) persistent equilibria. Section 6 relates rationalizability based on the refined best-reply correspondence to other refinements of Bernheim's (1984) and Pearce's (1984) notion of rationalizability. The paper concludes without conclusion.

2 Preliminaries

Let $\Gamma = (I, S, u)$ be a finite n-player normal form game, where $I = \{1, ..., n\}$ is the set of players, $S = \times_{i \in I} S_i$ is the set of pure strategy profiles, and $u : S \to \mathrm{I\!R}^n$ the payoff function[1]. Let $\Theta_i = \Delta(S_i)$ denote the set of player i's mixed strategies, and let $\Theta = \times_{i \in I} \Theta_i$ denote the set of all mixed strategy profiles. Let $\mathrm{int}(\Theta) = \{x \in \Theta : \; x_{is} > 0 \; \forall s \in S_i \; \forall i \in I\}$ denote the set of all completely mixed strategy profiles.

For $x \in \Theta$ let $\mathcal{B}_i(x) \subset S_i$ denote the set of pure-strategy best-replies to x for player i. Let $\mathcal{B}(x) = \times_{i \in I} \mathcal{B}_i(x)$. Let $\beta_i(x) = \Delta(\mathcal{B}_i(x)) \subset \Theta_i$ denote the set of mixed-strategy best-replies to x for player i. Let $\beta(x) = \times_{i \in I} \beta_i(x)$.

As in Balkenborg, Hofbauer, and Kuzmics (2009a) we shall restrict attention to games with a normal form in which the set of mixed-strategy profiles $\Psi = \{x \in \Theta | \; \mathcal{B}(x) \text{ is a singleton}\}$ is dense in Θ. We denote this class by $\mathcal{G}^*$. This is essentially the restriction to semi-reduced normal form games (See the appendix in Balkenborg, Hofbauer, and Kuzmics (2009a) for a detailed discussion).

For games in $\mathcal{G}^*$ let $\sigma : \Theta \rightrightarrows \Theta$ be the *refined best-reply correspondence* as defined in Balkenborg, Hofbauer, and Kuzmics (2009a) and as follows. For $x \in \Theta$ let

$$\mathcal{S}_i(x) = \{s_i \in S_i | \; \exists \{x_t\}_{t=1}^{\infty} \in \Psi : x_t \to x \wedge \mathcal{B}_i(x_t) = \{s_i\} \; \forall t\}.$$

Then $\sigma_i(x) = \Delta\left(\mathcal{S}_i(x)\right)$ and $\sigma(x) = \times_{i \in I} \sigma_i(x) \; \forall \; x \in \Theta$.

[1]The function u will also denote the expected utility function in the mixed extension of the game Γ.

3 Notions of dominance

This section defines and discusses notions of strict and weak dominance that are appropriate for the refined best-reply correspondence. We term these notions *strict* and *weak inferiority*.[2] They are such that, naturally, every strictly dominated strategy is strictly inferior, every weakly dominated strategy is weakly inferior, and every strictly inferior strategy is weakly inferior.

Definition 1 *A strategy $s_i \in S_i$ is strictly inferior if for every $x \in \Theta$ there is a $t_i \in S_i$ such that $u_i(s_i, x_{-i}) < u_i(t_i, x_{-i})$. A strategy $w_i \in S_i$ is weakly inferior if for every $x \in \Theta$ there is a $t_i \in S_i$, $t_i \neq w_i$ such that $u_i(w_i, x_{-i}) \leq u_i(t_i, x_{-i})$ and $u_i(w_i, y_{-i}) < u_i(t_i, y_{-i})$ for some $y \in \Theta$.*

For games in $\mathcal{G}^*$ a strictly inferior strategy s_i is such that $s_i \notin \mathcal{B}_i(x)$ for any $x \in \Theta$, while a weakly inferior strategy w_i is such that if $w_i \in \mathcal{B}_i(x)$ then $\mathcal{B}_i(x)$ is not a singleton. Note that every game in $\mathcal{G}^*$ has at least one strategy for each player which is not weakly inferior.

While in two-player games a strategy is strictly inferior if and only if it is strictly dominated, for 3 or more player games this is not so for the well known reason that a strategy can be a best reply to correlated strategy profile of the opponents, but never a best reply to any uncorrelated strategy profile (see, e.g., Example 5.7 in Ritzberger (2002) for a strategy which is strictly inferior but not strictly dominated).

While a weakly dominated strategy is weakly inferior, the converse is not necessarily true even in two-player games. This is demonstrated in Lemma 4 below.

Call $x_i \in \Theta_i$ a **robust best reply** against $x \in \Theta$ if x_i is a best reply against all strategy combinations in a neighborhood of x. Call $x_i \in \Theta_i$ a **robust strategy** if x_i is a robust best reply against *some* strategy combination $x \in \Theta$. This terminology is inspired by Okada (1983).

The following lemma, due to Kalai and Samet (1984), is used in the proof of Lemma 2.

Lemma 1 *Let U be a non-empty open subset of Θ. Then two strategies $x_i, y_i \in \Theta_i$ are own-payoff equivalent (for player i) if and only if $u_i(x_i, z_{-i}) = u_i(y_i, z_{-i})$ for all $z \in U$.*

Lemma 2 *A strategy is robust if and only if it is not weakly inferior.*

Proof: If $s_i \in S_i$ is not weakly inferior, then there exists an $x_{-i} \in \Theta_{-i}$ such that, for all $t_i \in S_i$, $u_i(s_i, x_{-i}) > u_i(t_i, x_{-i})$ or $u_i(s_i, y_{-i}) \geq u_i(t_i, y_{-i})$ holds for all $y_{-i} \in \Theta_{-i}$. In the latter case t_i is weakly dominated by s_i. By continuity s_i is a best reply in a neighborhood of x_{-i} and hence a robust strategy. Conversely, if $s_i \in S_i$ is robust it is a best reply on a non-empty open set in Θ. Any strategy which is not own-payoff equivalent to s_i can, by Lemma 1, be a best reply jointly with s_i only on a closed, nowhere dense set. There are only finitely many pure strategies to consider and any mixed strategy is a best reply only against strategy profiles against which all pure strategies in its support are also best replies. Thus, there exists a non-empty open set in Θ_i such that a strategy of player i is a best response against a strategy profile in this set if and only if it is own-payoff equivalent to s_i. Hence s_i is not weakly inferior. QED

As an immediate consequence of Lemma 2 we have:

Lemma 3 *Let $w_i \in S_i$ be weakly inferior for player i. Then $w_i \notin \mathcal{S}_i(x)$ (i.e., w_i is not a semi-robust best reply) for any $x \in \Theta$.*

[2]Our notions of strict and weak inferiority are motivated by, but not identical to, the notion of inferior choices in Harsanyi and Selten (1988).

In two-player games the following converse statement can be established. Its proof can be found in appendix B.

Theorem 1 *In a two-player game a strategy is a semi-robust best reply, $s_i \in \mathcal{S}_i(x)$, if and only if it is a best reply, $s_i \in \mathcal{B}_i(x)$, and is not weakly inferior.*

For two-player games, in fact, we can completely characterize weakly inferior strategies. The proof is again in appendix B.

Lemma 4 *In a two-player game a strategy is weakly inferior if and only if it is weakly dominated or equivalent to a proper mixture of strategies which are not own-payoff equivalent.*

Theorem 1 implies the following Proposition.

Proposition 1 *Let $\Gamma = (I, S, u) \in \mathcal{G}^*$ be a two-player game with minimal refined best-reply correspondence $\sigma(\Gamma)$. Then there is a game $\Gamma' = (I', S', u') \in \mathcal{G}^*$ with $I' = I$, $S' = S$, and a payoff function $u' : S \to \mathrm{IR}^2$ such that its best-reply correspondence $\beta(\Gamma') \equiv \sigma(\Gamma)$.*

Proof: Let Γ' be such that, for every $i \in I$, every $s_{-i} \in S_{-i}$, and every weakly inferior $w_i \in S_i$, $u'_i(w_i, s_{-i}) = u_i(w_i, s_{-i}) - \delta$ for some $\delta > 0$. Then, for this game Γ' no weakly inferior strategy is ever a best reply. Thus, by Theorem 1, $\sigma(\Gamma) \equiv \sigma(\Gamma') \equiv \beta(\Gamma')$. QED

Proposition 1 is useful as it tells us that in two-player games, the structure of fixed points of σ is the same as the structure of Nash equilibria. In particular, it implies that, in two-player games, there are only finitely many components of fixed points of σ. More precisely, applying the results in Jansen, Jurg, and Vermeulen (2002) we have the following.

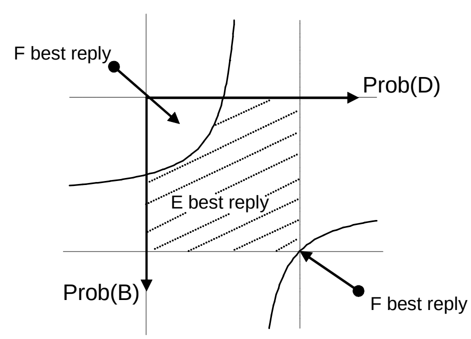
Corollary 1 *Let $\Gamma = (I, S, u) \in \mathcal{G}^*$ be a two-player game with minimal refined best-reply correspondence σ. Then the set of fixed points of σ is the union of finitely many polytopes and hence the union of finitely many connected components.*

	C	D
A	0	0
B	0	0

E

	C	D
A	1	-1
B	-1	0

F



Game 1 and Figure 1: A game where the refined best reply correspondence is not the best reply correspondence of a modified game. In Figure 1 the regions where strategies E and F of player 3 are best replies in this game are indicated in the square of strategy profiles of players 1 and 2. The probability with which player 1 chooses B is indicated vertically downwards in the graph while the probability of player 2 choosing D is indicated horizontally. In the shaded area between the two branches of the hyperbola E is the best reply for player 3, outside it is F. The lower branch of the hyperbola intersects the square only in the point (B, D), indicating that F is a best reply against (B, D), but not semi-robust.

Proposition 1 does not extend to games with three or more players, as the $2 \times 2 \times 2$ game 1 shows. Here and in the following games, player 1 chooses the row, player 2 the column and player 3 the matrix. In this example we specify the payoffs of player 3 only. As indicated in Figure 1, note that against opponent strategy profiles $(1/2A + 1/2B, C)$, $(A, 1/2C + 1/2D)$, and $(2/3A + 1/3B, 2/3C + 1/3D)$ (among others) both E and F are semi-robust best replies. However, against (A, C) F is the only best reply and against (B, D) E is the only semi-robust best reply. Nearby the latter strategy profile there is no open set in the square of the opponents' mixed strategy profiles where F is a best response.

Now assume there exists another game with the same strategies for which the best response mapping for player 3 is identical to the *minimal refined* best response correspondence of the given game. This implies that player 3 must remain indifferent between E and F against the strategy profiles $(1/2A + 1/2B, C)$ $(A, 1/2C + 1/2D)$, and $(2/3A + 1/3B, 2/3C + 1/3D)$. Moreover, F must be a best response against (A, C), but not against (B, D). This implies

$$\begin{aligned}
\frac{1}{2}\left(u_3\left(A,C,E\right) - u_3\left(A,C,F\right)\right) - \frac{1}{2}\left(u_3\left(B,C,E\right) - u_3\left(B,C,F\right)\right) &= 0\\
\frac{1}{2}\left(u_3\left(A,C,E\right) - u_3\left(A,C,F\right)\right) - \frac{1}{2}\left(u_3\left(A,D,E\right) - u_3\left(A,D,F\right)\right) &= 0\\
\frac{4}{9}\left(u_3\left(A,C,E\right) - u_3\left(A,C,F\right)\right) - \frac{2}{9}\left(u_3\left(B,C,E\right) - u_3\left(B,C,F\right)\right) &\\
-\frac{2}{9}\left(u_3\left(A,D,E\right) - u_3\left(A,D,F\right)\right) + \frac{1}{9}\left(u_3\left(B,D,E\right) - u_3\left(B,D,F\right)\right) &= 0
\end{aligned}$$

We conclude that $u_3\left(B,D,E\right) - u_3\left(B,D,F\right) = 0$, and, thus F is a best response against (B, D), a contradiction.

The following theorem says that for generic normal form games the minimal refined best-reply correspondence coincides everywhere with the best-reply correspondence, i.e., $\sigma \equiv \beta$. Remember, however, that many games of interest in $\mathcal{G}^*$ (derived for instance from an extensive form) are not among these generic normal form games.

Theorem 2 *For generic normal form games a strategy is a semi-robust best reply if and only if it is a best reply (i.e., $s_i \in \mathcal{S}_i(x) \Leftrightarrow s_i \in \mathcal{B}_i(x)$). Consequently, the best reply correspondence is minimal (i.e., $x_i \in \sigma_i(x) \Leftrightarrow x_i \in \beta_i(x)$) in these games.*

For the proof see appendix A.

4 Nash equilibrium versus best-reply refinements

This section provides a few results relating fixed points of the (minimal) refined best-reply correspondence to (refinements of) Nash equilibria.

Proposition 2 *Let Γ be a finite two-player game in $\mathcal{G}^*$. Let $x \in \Theta$ be a fixed point of the refined best-reply correspondence σ. Then $x_{iw_i} = 0$ for every weakly inferior $w_i \in S_i$.*

Proof: Immediate from Lemma 3: Let $x \in \sigma(x)$. By Lemma 3 $w_i \notin \mathcal{S}_i(x)$ for any weakly inferior $w_i \in S_i$. But then no $y \in \Theta$ with $y_{iw_i} > 0$ can be in $\sigma(x)$. QED

Selten (1975) introduced the concept of a (trembling-hand) perfect (Nash) equilibrium. A useful characterization of a perfect equilibrium in normal form games is given in the following lemma, which is also due to Selten (1975) (see also Proposition 6.1 in Ritzberger (2002) for a textbook treatment).

Lemma 5 *A (possibly mixed) strategy profile $x \in \Theta$ is a (trembling-hand) perfect (Nash) equilibrium if there is a sequence $\{x_t\}_{t=1}^{\infty}$ of completely mixed strategy profiles (i.e., each $x_t \in int(\Theta)$) such that x_t converges to x and $x \in \beta(x_t)$ for all t.*

Not every fixed point of σ is necessarily a trembling-hand perfect equilibrium, even in 2-player games. To see this consider Game 2, taken from Hendon, Jacobson, and Sloth (1996). For this game σ and β are identical. The mixed strategy profile $x^* = ((0, 1/2, 1/2); (1/2, 0, 1/2))$ is a Nash equilibrium, hence a fixed point of β, hence of σ, which, as Hendon, Jacobson, and Sloth (1996) point out is not perfect. Indeed, while the two pure strategies in the support of x_2^*, i.e., strategies D and F are not weakly dominated, the mixture x_2^* is weakly dominated by the pure strategy E. As weakly dominated strategies in a 2-player game cannot be perfect (see e.g., Theorem 3.2.2 in van Damme (1991)), x^* is not perfect.

	D	E	F
A	0,0	0,1	0,0
B	2,0	2,1	0,2
C	0,2	0,1	2,0

Game 2: A game in which a fixed point of σ is not perfect.

	D	E	F
A	2,2	1,2	1,2
B	2,1	2,2	0,0
C	2,1	0,0	2,2

Game 3: A game in which a perfect equilibrium (and, in fact, KM-stable equilibrium) is not a fixed point of σ.

Proposition 3 *Let Γ be a 2-player game in $\mathcal{G}^*$. Then every pure fixed-point, $s \in S$, of the refined best-reply correspondence, σ, is a perfect equilibrium.*

Proof: Every pure fixed point of σ is undominated by Proposition 2. In two-player games every undominated Nash equilibrium is (trembling-hand) perfect. QED

The converse of Proposition 3 is not true. Consider Game 3. In this game strategy A (and similarly D) is equivalent to the mixture of pure strategies B and C (E and F respectively). However, A is a best-reply only on a thin set of mixed-strategy profiles. In fact, A is best against any $x \in \Theta$ in which $x_{2E} = x_{2F}$, the set of which is a thin set. Thus, this game is in $\mathcal{G}^*$. In this game (A, D) constitutes a perfect equilibrium. In fact every mixed strategy profile $((\alpha, \frac{1-\alpha}{2}, \frac{1-\alpha}{2}); (\alpha, \frac{1-\alpha}{2}, \frac{1-\alpha}{2}))$ is a strictly perfect equilibrium, and hence, constitutes a singleton KM-stable set, where KM-stable is (strategically) stable in the sense of Kohlberg and Mertens (1986) (a minimal set satisfying their Property S). See next section for a definition. But none of these equilibria, except the one with $\alpha = 0$, are fixed points of σ, due to the fact that A (and D) is only best on a thin set (it is a weakly inferior strategy).

Proposition 3 cannot be generalized to general n-player games. To see this consider the following immediate characterization of fixed points of σ. For $x_i \in \Theta_i$ let $C(x_i) = \{s_i \in S_i | x_{is_i} > 0\}$ denote the carrier (or support) of x_i.

Lemma 6 *Strategy profile $x \in \Theta$ satisfies $x \in \sigma(x)$ if and only if for all $i \in I$ and for all $s_i \in C(x_i)$ there is an open set $U^{s_i} \subset \Theta$, with x in the closure of U^{s_i}, such that $s_i \in \mathcal{B}_i(y)$ for all $y \in U^{s_i}$.*

Suppose $x \in \sigma(x)$. Consider player i. Then for all $s_i \in C(x_i)$ let U^{s_i} denote this open set in which s_i is best. Now if $\bigcap_{i\in I}\bigcap_{s_i\in C(x_i)} U^{s_i} \neq \emptyset$, then x is also trembling-hand perfect. However, this is not necessarily the case. We already saw this for the 2-player Game 2. In the fixed point of σ, $x^* = ((0,1/2,1/2);(1/2,0,1/2))$, player 2 uses his pure strategies D and F only. D is best in the open set $U^D = \{x \in \Theta | x_{1C} > \frac{1}{2}\}$, while F is best in the open set $U^F = \{x \in \Theta | x_{1B} > \frac{1}{2}\}$. There are no bigger open set with the same property. Yet the two sets have an empty intersection. Hence, x^* is not perfect.

Balkenborg, Hofbauer, and Kuzmics (2009a) prove that CURB sets (Basu and Weibull (1991)) based on σ give rise to absorbing retracts (Kalai and Samet (1984)) and minimal such sets give rise to persistent retracts. One might think that fixed points of σ will have some relation to persistent equilibria (Nash equilibria in a persistent retract, Kalai and Samet (1984)). This is not true, though. Note first that the mixed equilibrium in the coordination game is not persistent and is a fixed point of σ. Consider Game 4 taken from Kalai and Samet (1984). The equilibrium (B,D,E) is perfect and proper but not persistent as Kalai and Samet (1984) point out. It is also a fixed point of σ. To see this, note that E is weakly dominant for player 3 and that B and D are best (for players 1 and 2, respectively) against all nearby strategy profiles in which player 2 chooses strategy C with smaller probability than player 3 chooses F and player 1 chooses A with smaller probability than player 3 chooses F.

	C	D
A	1,1,1	0,0,0
B	0,0,0	0,0,0

E

	C	D
A	0,0,0	0,0,0
B	0,0,0	1,1,0

F

Game 4: A game in which a pure fixed point of σ is not persistent.

Game 5, taken from Kalai and Samet (1984), demonstrates that there are persistent equilibria which are not fixed points of σ. The strategy profile (A,C,E) is persistent (see Kalai and Samet (1984)) but is not a fixed point of σ. To see this note that player 1's strategy A is never best for nearby strategy profiles. The one pure strategy combination (of players 2 and 3) against which A is better than B is (D,F) which for nearby (to (A,C,E)) strategy profiles will always have lower probability than the outcomes (C,F) and (D,E) against which B is better than A.

	C	D
A	0,0,0	0,0,1
B	0,1,0	1,0,1

E

	C	D
A	0,1,0	1,0,0
B	1,0,1	0,1,0

F

Game 5: A game in which a pure persistent equilibrium is not a fixed point of σ.

5 σ-CURB sets and strategic stability

As already mentioned Balkenborg, Hofbauer, and Kuzmics (2009a) prove that CURB sets (Basu and Weibull (1991)) based on σ give rise to absorbing retracts (Kalai and Samet (1984)) and minimal such sets give rise to persistent retracts. This equivalence allows us to provide a relatively simple proof of the fact that every persistent retract contains a KM-stable set.

A set $R \subset S$ is a **strategy selection** if $R = \times_{i\in I} R_i$ and $R_i \subset S_i$, $R_i \neq \emptyset$ for all i. For a strategy selection R let $\Theta(R) = \times_{i\in I}\Delta(R_i)$ denote set of independent strategy mixtures of the pure strategies in R. A set $\Psi \subset \Theta$ is a **face** if there is a strategy selection R such that $\Psi = \Theta(R)$. Note that $\Theta = \Theta(S)$. Note also that $\beta(x) = \Theta(\mathcal{B}(x))$ and $\sigma(x) = \Theta(\mathcal{S}(x))$.

A strategy selection R is a **σ-CURB set** if $\mathcal{S}(\Theta(R)) \subset R$. It is a **tight σ-CURB set** if, in addition $\mathcal{S}(\Theta(R)) \supset R$, and, hence, $\mathcal{S}(\Theta(R)) = R$. It is a **minimal σ-CURB set** if it does not properly contain another σ-CURB set.

Definition 2 *Let $\Gamma = (I, S, u)$ be a normal form game. For $i \in I$ let $\eta_i : S_i \to \mathbb{R}$ be such that*

$$\eta_i(s_i) > 0 \ \forall s_i \in S_i$$

and

$$\sum_{s_i \in S_i} \eta_i(s_i) < 1.$$

Then $\eta = (\eta_1, ..., \eta_n)$ is a **tremble**. *Let $\Theta_i(\eta) = \{x \in \Theta_i | x_i(s_i) \geq \eta_i(s_i) \ \forall s_i \in S_i\}$. Then $\Gamma(\eta) = (I, \Theta(\eta), u)$ is the η-***perturbed** *game.*

The following defines property S of Kohlberg and Mertens (1986) without the requirement of being a subset of the set of Nash equilibria, before defining Kohlberg and Mertens's (1986) concept of strategic stability.

Definition 3 *Let Γ be a finite normal form game. Let $Q \subset \Theta$ be a closed subset of the set of mixed strategy profiles. Q is* **prestable** *if for all $\epsilon > 0$ there is a $\delta > 0$ such that for all trembles η with $\eta_i(s_i) < \delta$ for all $s_i \in S_i$ and for all $i \in I$ there is a Nash equilibrium, x^η, of the perturbed game $\Gamma(\eta)$ such that $||x^\eta - x|| < \epsilon$ for some $x \in Q$. Such a set Q is* **KM-stable** *if it is prestable and does not properly contain another prestable set.*

Note that minimality requires that a KM-stable set consists of exclusively perfect equilibria.

Proposition 4 *Let $\Gamma = (I, S, u)$ be a normal form game. Every σ-CURB set of Γ contains a KM-stable set.*

Proof: It is sufficient to show that a σ-CURB set is prestable. Let R be a σ-CURB set. Fix a tremble η and the associated perturbed game $\Gamma(\eta)$ with the set of restricted strategy profiles $\Theta(\eta)$. Define $\Theta^*(R) = \{x \in \Theta(\eta) | x_{is} = \eta_{is}$ if $s \notin R_i\}$, a compact and convex subset of $\Theta(\eta)$. For $x \in \Theta^*(R)$ let $\sigma^*(x) = \{y \in \Theta(\eta) | y_{is} = \eta_{is}$ if $s \notin \mathcal{S}_i(x)\}$. Thus, σ^* is an upper hemi-continuous correspondence defined on a convex compact set. By Kakutani's fixed point theorem σ^* has a fixed point. By the assumption that R is a σ-CURB set and the fact that σ is upper hemi-continuous, we have that there is a neighborhood U of $\Theta(R)$ such that $\sigma(U) \subset \Theta(R)$. Thus, as long as η is sufficiently close to zero, such that $\Theta^*(R) \subset U$, this fixed point of σ^* is a Nash equilibrium of the perturbed game. Thus, every sufficiently close perturbed game has a Nash equilibrium close to the σ-CURB set. QED

Note that no such result holds for the two preliminary notions of strategic stability also put forward in Kohlberg and Mertens (1986). Both hyper-stable and fully stable sets can sometime include pure weakly dominated strategies. A minimal σ-CURB set cannot.

Still, given the interpretation of Balkenborg, Hofbauer, and Kuzmics (2009a) that σ-Curb sets are asymptotically stable sets under the refined best-reply dynamics, this result is reminiscent to the result of Swinkels (1993) that every convex asymptotically stable set of states under some reasonable deterministic dynamics, in which every Nash equilibrium is stationary, contains a hyper-stable set. Of course, not every Nash equilibrium is stationary under the refined best-reply dynamics.

6 $\boldsymbol{\sigma}$-Rationalizability

For $A \subset \Theta$ let $\mathcal{B}_i(A) = \{s_i \in S_i | s_i \in \mathcal{B}_i(x) \text{ for some } x \in A\}$ denote the set of all pure best-replies for player i to all strategy profiles in set A. Let $\beta_i(A) = \Delta(\mathcal{B}_i(A))$ denote the convex hull of this set $\mathcal{B}_i(A)$. Let $\beta(A) = \times_{i \in I} \beta_i(A)$. For $k \geq 2$ let $\beta^k(A) = \beta(\beta^{k-1}(A))$. For $A = \Theta$, $\beta^k(A)$ is a decreasing sequence, and we denote $\beta^\infty(\Theta) = \bigcap_{k=1}^{\infty} \beta^k(\Theta)$. A pure strategy profile $s \in S$ is **rationalizable** if it is an element of the strategy selection $R \subset S$ which satisfies $\Theta(R) = \beta^\infty(\Theta)$ (Bernheim (1984) and Pearce (1984); see also Ritzberger (2002), Definition 5.3 for a textbook treatment).

The same can be done with the refined best-reply correspondence σ. For $A \subset \Theta$ let $\mathcal{S}_i(A) = \{s_i \in S_i : s_i \in \mathcal{S}_i(x) \text{ for some } x \in A\}$ denote the set of all pure refined best-replies for player i to all strategy profiles in set A. Let $\sigma_i(A) = \Delta(\mathcal{S}_i(A))$. Let $\sigma(A) = \times_{i \in I} \sigma_i(A)$. For $k \geq 2$ let $\sigma^k(A) = \sigma(\sigma^{k-1}(A))$. For $A = \Theta$, $\sigma^k(A)$ is again a decreasing sequence, and we denote $\sigma^\infty(\Theta) = \bigcap_{k=1}^{\infty} \sigma^k(\Theta)$. A pure strategy profile $s \in S$ is **$\boldsymbol{\sigma}$-rationalizable** if it is an element of the strategy selection $R \subset S$ which satisfies $\Theta(R) = \sigma^\infty(\Theta)$.

Using Theorem 2 we obtain the result that for generic normal form games σ-rationalizable is the same as just rationalizable.

Proposition 5 *In generic normal form games a strategy is σ-rationalizable if and only if it rationalizable.*

Proof: This follows immediately from Theorem 2. QED

For general games in $\mathcal{G}^*$, by the fact that $\sigma(x) \subset \beta(x)$ for all $x \in \Theta$, we obviously have that the set of σ-rationalizable strategies is a subset of the set of rationalizable strategies. However, we can say more. Suppose we want to compare σ-rationalizability with the Dekel-Fudenberg, or $S^\infty W$, procedure of Dekel and Fudenberg (1990). We shall call strategies that survive this elimination procedure **Dekel-Fudenberg rationalizable**.[3] This procedure performs one round of deletion of all weakly dominated pure strategies, and then, iteratively, all strictly dominated ones. It is immediate that any σ-rationalizable strategy is Dekel-Fudenberg rationalizable. There is an obvious strengthening of the Dekel-Fudenberg procedure. In the first round instead of eliminating weakly dominated strategies we can eliminate weakly inferior strategies. In later stages instead of eliminating strictly dominated strategies, we can eliminate strictly inferior strategies. Let us call pure strategies which survive this latter elimination procedure **strong Dekel-Fudenberg rationalizable**.

Strong DF-rationalizability can be seen as original rationalizability for a game in which payoffs are perturbed slightly. Let $\tilde{\Gamma} = (I, S, \tilde{u})$ denote the game derived from Γ by defining $\tilde{u}_i(s_i, s_{-i}) = u_i(s_i, s_{-i}) - \delta$, for a fixed positive δ, if $s_i \in S_i$ is weakly, and not strictly, inferior in Γ and $\tilde{u}_i(s_i, s_{-i}) = u_i(s_i, s_{-i})$ otherwise. Every pure strategy which is weakly inferior in Γ is, therefore, strictly inferior in $\tilde{\Gamma}$. Let $\tilde{\beta}$ denote the best-reply correspondence of $\tilde{\Gamma}$. Then we have the following lemma.

[3] See Börgers (1994) and Brandenburger (1992) for epistemic conditions under which DF rationalizability is obtained.

Lemma 7 *For $\tilde{\Gamma}$ and $\tilde{\beta}$ defined as above we have $\sigma(x) \subset \tilde{\beta}(x)$ for all $x \in \Theta$.*

Proof: Follows immediately from Lemma 3. QED

The refined best-reply set $\sigma(x)$ may, for some games Γ and some $x \in \Theta$, be a proper subset of $\tilde{\beta}(x)$. To see this consider Game 6, taken from van Damme (1991), Figure 2.2.1; see also exercise 6.10 in Ritzberger (2002). In this game strategies D and F are strictly inferior for players 2 and 3, respectively. There are no strategies which are weakly but not strictly inferior. Hence, $\tilde{\beta}(x) = \beta(x)$ for any $x \in \Theta$. Player 1's strategy B is (the unique) best strategy when players 2 and 3 play D and F, respectively. Both A and B are best when players 2 and 3 play C and E, respectively. However, for any (mixed) strategy profile, $y \in \Theta$ in which players 2 and 3 play close to C and E, A is the unique best reply. Hence, $B \notin \mathcal{S}_1(x)$ for any $x \in \Theta$ for which $x_{2C} = 1$ and $x_{3E} = 1$. Therefore, $\mathcal{S}_1(x) = \{A\}$ is indeed a proper subset of $\mathcal{B}_1(x) = \{A, B\}$ for any such x, and, hence, $\sigma(x)$ is a proper subset of $\tilde{\beta}(x)$ for any such x. In fact, this game is usually used to illustrate that in 3-player games an undominated Nash equilibrium, (B, C, E), need not be perfect, as is indeed the case here.

	C	D
A	1,1,1	1,0,1
B	1,1,1	0,0,1

E

	C	D
A	1,1,0	0,0,0
B	0,1,0	1,0,0

F

Game 6: A game in which for some $x \in \Theta$, $\sigma(x)$ is a proper subset of $\tilde{\beta}(x)$.

Let $\tilde{\beta}^\infty$ be defined analogously to β^∞. Thus a pure strategy $s \in S$ is strong Dekel-Fudenberg rationalizable if it is an element of the strategy selection $R \subset S$ which satisfies $\Theta(R) = \tilde{\beta}^\infty(\Theta)$.

Proposition 6 *Let $\Gamma \in \mathcal{G}^*$. Every σ-rationalizable strategy for Γ is strong Dekel-Fudenberg rationalizable.*

In two-player games the converse is also true.

Proposition 7 *In a two-player game the σ-rationalizable strategies are the strategies that survive one round of elimination of weakly inferior pure strategies followed by the iterated elimination of strictly dominated strategies.*

Proof: This follows from Theorem 1 and the fact that in two-player games strictly inferior strategies are the same as strictly dominated strategies. QED

Note that Proposition 7 indeed implies that, for two-player games, a strategy is σ-rationalizable if and only if it is DF-rationalizable. In fact for two-player games without pure strategies which are equivalent to a proper mixture of two or more pure strategies we have that, by Lemma 4, a strategy is σ-rationalizable if and only if it survives the DF-procedure.

Game 6 illustrates that, in games with more than two players, the set of σ-rationalizable strategies, here $\{A\} \times \{C\} \times \{E\}$, may well be a proper subset of the set of DF-rationalizable strategies, here $\{A, B\} \times \{C\} \times \{E\}$.

There are a variety of refinements of the concept of (uncorrelated) rationalizability of Bernheim (1984) and Pearce (1984). The ones we are aware of are **cautious rationalizability** (Pearce (1984)), **perfect rationalizability** (Bernheim (1984)), **proper rationalizability** (Schuhmacher (1999)), **trembling-hand perfect rationalizability**, and **weak perfect rationalizability** (both Herings and Vannetelbosch (1999)).

Herings and Vannetelbosch (1999) study the relationship between all these concepts. They find that perfect and proper rationalizability both imply weakly perfect rationalizability and provide counter-examples to every other possible set-inclusion. We do not want to go into the various definitions here now, but will just point out how these concepts are related to σ-rationalizability as defined in this paper.

In Game 3 all of the above refinements of rationalizability yield the whole strategy set, while σ-rationalizability leads to the smaller set $\{B, C\} \times \{E, F\}$. In Game 8, trembling hand perfect rationalizability yields, with $\{A\}\times\{D\}$, a subset of the set of σ-rationalizable strategies, $\{A, B\}\times \{D, E\}$. In the game, derived from Game 8 by replacing C and F with strictly dominated strategies, and not changing the payoffs other strategies obtain against C and F, the set of cautiously rationalizable strategies, $\{A\} \times \{D\}$, is a proper subset of the set of σ-rationalizable strategies, again given by $\{A, B\} \times \{D, E\}$. In Game 7, the semi-reduced normal form of a centipede game, the set of properly rationalizable strategies, $\{A\} \times \{F\}$, is smaller than the set of σ-rationalizable strategies, $\{A, B, C\}\times\{D, F\}$. While we thus have no systematic relationship between the concepts of cautious, trembling hand perfect, proper, and σ-rationalizability, it may well be the case that perfect and weakly perfect rationalizability, both as defined in Herings and Vannetelbosch (1999), are, sometimes strictly, weaker criteria than σ-rationalizability. This issue is open.

	D	E	F
A	3,0	3,0	3,0
B	4,3	1,2	1,2
C	4,3	0,1	2,4

Game 7: The semi-reduced normal form game of a centipede game .

To illustrate that σ-rationalizability does not always allow the iterated deletion of weakly dominated strategies, unlike trembling-hand perfect rationalizability, consider Game 8 from Samuelson (1992). In this game strategies C and F are weakly dominated, and, hence, not σ-rationalizable. In the semi-reduced game without strategies C and F, strategies B and E are now weakly dominated, and, hence, not trembling-hand perfect rationalizable. However, B (the analogue holds for E) is a best reply against completely mixed strategy profiles close to D, in which the weight on F is greater than the weight on E. Hence, $\mathcal{S}_1(x) = \{A, B\}$ for any such $x \in \Theta$. Hence, B is σ-rationalizable.

	D	E	F
A	1,1	1,1	2,1
B	1,1	0,0	3,1
C	1,2	1,3	1,1

Game 8: A game in which the set of σ-rationalizable strategies includes an iteratively weakly dominated strategy.

In some special contexts σ-rationalizability does allow the iterated deletion of weakly dominated strategies. See Balkenborg, Hofbauer, and Kuzmics (2009b).

A On the generic equivalence of best responses and semi-robust best responses

This section provides a proof of Theorem 2. Note first that generic normal form games are in the class $\mathcal{G}^*$ where all semi-robust responses are pure strategies. The proof of Theorem 2 is organized in a number of steps: We will first fix some notations for the mappings and various submanifolds to be considered. Step 1 argues that the embedding of the uncorrelated strategy combinations into the set of beliefs has nice differentiability properties. Step 2 invokes the transversality theorem (see Guillemin and Pollack (1974)) to show that for generic payoff functions we obtain the needed transversality conditions.[4] Step 3 argues that we can restrict attention to completely mixed strategy combinations of the opponents. If the player is indifferent between several of his strategies against a given completely mixed strategy combination, step 4 shows how we can construct an arbitrarily nearby strategy combination, against which the player strictly prefers a given one among these strategies. Step 5 completes the argument.

For any finite set M let $\mathbb{R}^M$ be the vector space of all mappings $y : M \to \mathbb{R}$. The dimension of $\mathbb{R}^M$ is the number of elements in M.

Let $q_i : \prod_{j \neq i} \mathbb{R}^{S_j} \to \mathbb{R}^{S_{-i}}$ be the mapping defined by

$$\left(q_i\left(x_{-i}\right)\right)\left(s_{-i}\right) := \prod_{j \neq i} x_j\left(s_j\right) \ .$$

q_i describes the first step in the computation mentioned above.

While x_{-i} denotes the usual strategy combination of the opponents, we define χ_{-i} to describe a "correlated strategy of the opponents", i.e., a belief over S_{-i}.

Let L_i be the vector space of all linear mappings

$$v_i : \mathbb{R}^{S_{-i}} \to \mathbb{R}^{S_i}.$$

If $\chi_{-i} \in \mathbb{R}^{S_{-i}}$ is a belief and $s_i \in S_i$ a pure strategy of player i, then $\left(v_i\left(\chi_{-i}\right)\right)\left(s_i\right)$ is the expected payoff for player i. v_i describes for every s_i the second step in the computation of the expected payoff. Any $v_i \in L_i$ corresponds uniquely to a payoff function

$$u_i : S \to \mathbb{R}$$

in the standard notation (and this relation is a homeomorphism).

For $T_i \subseteq S_i$ set $Z_i\left(T_i\right) = \{z \in \mathbb{R}^{S_i} \mid \forall s_i, t_i \in T_i : z\left(s_i\right) = z\left(t_i\right)\}$. Let $X_j := \{x_j \in \mathbb{R}^{S_j} \mid \sum_{s_j \in S_j} x_j\left(s_j\right) = 1\}$ for $j \neq i$ and $X_{-i} := \prod_{j \neq i} X_j$.

For $T_j \subseteq S_j$ $(j \neq i)$ and $T_{-i} := \prod_{j \neq i} T_j$ set

$$X_j\left(T_j\right) := \{x_j \in X_j \mid \forall s_j \in T_j : x_j\left(s_j\right) = 0\}$$

[4] This transversality theorem is a straightforward consequence of Sard's theorem. If one assumes an algebraic map and uses in its proof in Guillemin and Pollack (1974) the algebraic version of Sard's theorem in Bochnak, Coste, and Roy (1998) one obtains a stronger version of the transversality theorem where the conclusion of the theorem holds outside a lower dimensional semi-algebraic set.

and

$$X_{-i}(T_{-i}) := \prod_{j \neq i} X_j(T_j) \ .$$

The sets $\Theta_{-i} \cap X_{-i}(T_{-i})$ describe the various faces of the polyhedron Θ_{-i}.

Step 1: For all T_{-i} the mapping $q_i : X_{-i}(T_{-i}) \to \mathbb{R}^{S_{-i}} \setminus \{0\}$ is a diffeomorphism onto its image (in particular $q_i(X_{-i}(T_{-i}))$ is a closed submanifold of $\mathbb{R}^{S_{-i}} \setminus \{0\}$).

Proof: $X_{-i}(T_{-i})$ is a closed affine submanifold in $\prod_{j \neq i} \left(\mathbb{R}^{S_j} \setminus \{0\}\right)$. It is straightforward to check that

$$q_{i|X_{-i}(T_{-i})} : X_{-i}(T_{-i}) \to \mathbb{R}^{S_{-i}} \setminus \{0\}$$

is well defined, is one-to-one, maps $X_{-i}(T_{-i})$ to a closed set, and has a derivative $dq_i|_{x_{-i}}$ with maximal rank everywhere.[5]

Step 2: Let $Z \subseteq \mathbb{R}^{S_i}$ and $X \subseteq \mathbb{R}^{S_{-i}} \setminus \{0\}$ be submanifolds. Then for almost every $v_i \in L_i$ the mapping $v_i|_X : X \to \mathbb{R}^{S_{-i}} \setminus \{0\}$ is transversal to Z.

Proof: The family of linear maps L_i defines a mapping

$$\begin{aligned} V_i : L_i \times \mathbb{R}^{S_{-i}} &\to \mathbb{R}^{S_i} && (1) \\ (v_i, \chi_{-i}) &\mapsto v_i(\chi_{-i}) \ . && (2) \end{aligned}$$

The derivative of V_i at (v_i, χ_{-i}) can be computed as

$$\begin{aligned} dV_i|_{(v_i,\chi_{-i})} : T_{v_i}L_i \times T_{\chi_{-i}}\mathbb{R}^{S_{-i}} \cong L_i \times \mathbb{R}^{S_{-i}} &\to \mathbb{R}^{S_i} && (3) \\ (\nu_i, \xi_{-i}) &\mapsto \nu_i(\chi_{-i}) + v_i(\xi_{-i}) \, . && (4) \end{aligned}$$

If $\chi_{-i} \neq 0$ we can find for every $\zeta_i \in \mathbb{R}^{S_i}$ some $\nu_i \in L_i$ with $\nu_i(\chi_{-i}) = \zeta_i$. Then $dV_i|_{(v_i,\chi_{-i})}(\nu_i, 0) = \zeta_i$.

Because for $\chi_{-i} \in X$ the tangent space $T_{\chi_{-i}}X \subseteq \mathbb{R}^{S_{-i}}$ contains the 0-vector, $dV_i|_{(v_i,\chi_{-i})} : T_{v_i}L_i \times T_{\chi_{-i}}X \to \mathbb{R}^{S_i}$ is surjective. Thus $V_i : L_i \times X \to \mathbb{R}^{S_i}$ is transversal to Z and our claim follows from the transversality theorem.

By step 1 and step 2 almost every $v_i \in L_i$ satisfies:

$\otimes$ For all subsets $T_i \subseteq S_i$ $(1 \leq i \leq n)$ the mapping $(v_i \circ q_i)|_{X_{-i}(T_{-i})}$ is transversal to $Z_i(T_i)$.

For given v_i define $Y(T_i) = \{x_{-i} \in X_{-i} \mid (v_i \circ q_i)(x_{-i}) \in Z(T_i)\}$. $Y(T_i) \cap \Theta_{-i}$ is the set of strategy combinations of the opponents such that player i is indifferent between the strategies in T_i (i.e., they give the same payoff).

Step 3: Suppose v_i satisfies $\otimes$. For $T_i \subseteq S_i$ let $x_{-i} \in Y(T_i) \cap \Theta_{-i}$ and let O_{-i} be a neighborhood of x_{-i}. Then $O_{-i} \cap Y(T_i)$ contains a point in the interior of Θ_{-i}.

[5] The result is well known, e.g., in algebraic geometry: q_i defines the so-called Segre-embedding. The result is needed in algebraic geometry to show that the product of projective spaces can itself be embedded into a projective space, i.e., is projective-algebraic.

Proof: Suppose x_{-i} is in the boundary of Θ_{-i}. For each $j \neq i$ define $T_j := \{s_j \in S_j \mid x_j(s_j) = 0\}$. If $T_j = \emptyset$, x_j is in the relative interior of Θ_j. By assumption T_j is not empty for at least one opponent j. Fix $j* \neq i$ with $T_{j*} \neq \emptyset$ and $t_{j*} \in T_{j*}$. Set $\tilde{T}_j := T_j$ for $i \neq j \neq j*$ and $\tilde{T}_{j*} := T_{j*} \setminus \{t_{j*}\}$. We show that $O_{-i} \cap Y(T_i)$ contains some $y_{-i} \in \Theta_{-i} \cap X\left(\tilde{T}_{-i}\right)$ such that $\tilde{T}_j = \{s_j \in S_j \mid y_j(s_j) = 0\}$ for all $j \neq i$. In other words: y_{-i} is in the relative interior of the face $\Theta_{-i} \cap X\left(\tilde{T}_{-i}\right)$. The claim then follows by induction.

The transversality conditions imply that the submanifolds $X_{-i}(T_{-i})$ and $Y(T_i) \cap X_{-i}\left(\tilde{T}_{-i}\right)$ meet transversely in $X_{-i}\left(\tilde{T}_{-i}\right)$. (See Guillemin and Pollak [1974, exercise 2.3.7.]) Since $X_{-i}(T_{-i})$ has codimension 1 in $X_{-i}\left(\tilde{T}_{-i}\right)$ it follows that $X_{-i}\left(\tilde{T}_{-i}\right) \cap Y(T_i) \cap \{y_{-i} \mid y_{j*}(t_{j*}) > 0\} \cap O_{-i}$ intersects the relative interior of $X_{-i}\left(\tilde{T}_{-i}\right) \cap \Theta_{-i}$.

Step 4: Suppose v_i satisfies $\otimes$. For $T_i \subseteq S_i$ with $\#T_i \geq 2$ let $x_{-i} \in Y(T_i)$ be in the interior of Θ_{-i} and let O_{-i} be a neighborhood of x_{-i}. Then we can find for every $s_i \in T_i$ some $y_{-i} \in O_{-i} \cap \Theta_{-i}$ such that

$$(v_i \circ q_i)(y_{-i})(s_i) > (v_i \circ q_i)(y_{-i})(t_i) \qquad \text{for all} \quad t_i \in T_i \setminus \{s_i\}.$$

Proof: Because $v_i \circ q_i : X_{-i} \to \mathbb{R}^{S_i}$ is transversal to both $Z(T_i)$ and $Z(T_i \setminus \{s_i\})$ it follows that $v_i \circ q_i : Y(T_i \setminus \{s_i\}) \to Z(T_i \setminus \{s_i\})$ is transversal to $Z(T_i)$. From this we can deduce the existence of a tangent vector $\xi \in T_{x_{-i}}(Y(T_i \setminus \{s_i\}))$ with $d\lambda_{|x_{-i}}(\xi) = 1$, where λ is the function

$$\begin{aligned} \lambda : Y_i(T_i \setminus \{s_i\}) \cap X_{-i} &\to \mathbb{R} && (5)\\ y_{-i} &\to (v_i \circ q_i)(y_{-i})(s_i) - (v_i \circ q_i)(y_{-i})(t_i) && (6) \end{aligned}$$

defined for arbitrary but fixed $t_i \in T_i \setminus \{s_i\}$. We can therefore select a differentiable curve

$$c : (-\epsilon, \epsilon) \to Y_i(T_i \setminus \{s_i\})$$

with $c(0) = x_{-i}$ and $(\lambda \circ c)'(0) = 1$. For sufficiently small $0 < \gamma < \epsilon \quad y_{-i} := c(\gamma)$ has the required properties.

Step 5: Suppose s_i is a pure best response against x_{-i}. For every neighborhood O_{-i} of x_{-i} the continuity of the payoff function and the two steps above can be used to find $y_{-i} \in O_{-i}$ such that $s_i \in T_i$ is the unique best response against y_{-i}. Shrinking the open sets we can find a sequence of such y_{-i}'s converging to x_{-i}. Continuity yields an open set around each element in the sequence, where s_i is the unique best response. s_i is the unique best response on the union of these sets, which is again open. Thus s_i is a semi-robust best response against x_{-i}. *QED*

B Refined best replies in two-player games

We will restrict attention to the best replies of player 1. Suppose player 2 has $K \geq 2$ strategies $s_2^1, \cdots, s_2^K$. It will be convenient to identify the mixed strategies $x_2 \in \Theta_2$ with the vectors

$$x_2 = \left(x_2^1, x_2^2, \cdots, x_2^{K-1}\right) \in \mathbb{R}^{K-1} \tag{7}$$

for which $x_2^k \geq 0$ for all $1 \leq k \leq K-1$ and $x_2^K := 1 - \sum_{k=1}^{K-1} x_2^k \leq 0$. Notice that the zero vector corresponds to pure strategy s_2^K.

We define the function $f : \mathbb{R}^{K-1} \rightarrow \mathbb{R}$ by

$$f(x_2) = \begin{cases} \max_{s_1 \in S_1} u_1(s_1, x_2) & \text{for} \quad x_2 \in \Theta_2 \\ +\infty & \text{else} \end{cases} \tag{8}$$

Because u_1 is linear in x_2, f is, in the terminology of Rockafellar (1970) a proper convex polyhedral function. Each strategy $x_1 \in \Theta_1$ defines an affine function $a : \mathbb{R}^{K-1} \rightarrow \mathbb{R}$ by $a(x_2) = u_1(x_1, x_2)$, which, for all $x_2 \in \Theta_2$, satisfies the inequality $a(x_2) \leq f(x_2)$ and $a(x_2) = f(x_2)$ holds if and only if $x_1 \in \beta_1(x_2)$.

For a strategy $x_1 \in \Theta_1$ we define the set

$$G(x_1) = \{(x_2, \alpha) \in \Theta_2 \times \mathbb{R} \mid x_1 \in \beta(x_2) \text{ and } \alpha = u_1(x_1, x_2)\} \tag{9}$$

and the set $H(x_1) = \{x_2 \in \Theta_2 \mid x_1 \in \beta(x_2)\}$, the projection of $G(x_1)$ onto Θ_2. $H(x_1)$ is the region where x_1 is a best reply. x_1 is *robust* if $H(x_1)$ has non-empty interior $H^\circ(x_1)$. We can now describe the epigraph

$$F = \left\{(x_2, \alpha) \in \mathbb{R}^{K-1} \times \mathbb{R} \mid f(x_2) \leq \alpha\right\} \tag{10}$$

of f as follows: F is a polyhedral convex set whose compact faces are precisely the sets $G(x_1)$ with $x_1 \in \Theta_1$. The non-compact faces are of the form $F \cap (\Theta_1' \times \mathbb{R})$, where Θ_1' is a face of Θ_1.

Because $H(x_1)$ is a convex polyhedron the closure of $H^\circ(x_1)$ is $H(x_1)$ if $H^\circ(x_1)$ is non-empty. Using lemma 2 this implies immediately Theorem 1.

To prove Lemma 4 we study the conjugate function $f^* : \mathbb{R}^{K-1} \rightarrow \mathbb{R}$ of f defined by

$$f^*(x_2^*) = \sup_{x_2 \in \mathbb{R}^{K-1}} \{x_2^* \bullet x_2 - f(x_2)\} = \max_{x_2 \in \Theta_2} \{x_2^* \bullet x_2 - f(x_2)\} < \infty, \tag{11}$$

where $x_2^* \bullet x_2$ denotes the usual scalar product $\sum_{k=1}^{K-1} x_2^{*k} x_2^k$. As shown for any convex polyhedral function in Rockafellar (1970), the conjugate is again a convex polyhedral function and one has $f^{**}(x_2) = f(x_2)$.

Any two strategies $x_1, x_1' \in \Theta_1$ define the same affine function if and only if the two strategies are equivalent. Without loss of generality we can thus identify Θ_1 with a subset of the affine functions on $\mathbb{R}^{K-1}$.

Any vector (x_2^*, α) with $x_2^* \in \mathbb{R}^{K-1}$ and $\alpha \in \mathbb{R}$ defines one and only one affine function on $\mathbb{R}^{K-1}$ by

$$a(x_2) = -\alpha + \sum_{k=1}^{K-1} x_2^{*k} x_2^k \tag{12}$$

We will identify affine functions with such vectors. For instance, $e = (1, \ldots, 1)$ corresponds to the function $-x_2^{K+} = -1 + \sum_{k=1}^{K-1} x_2^k$ which assigns 0 to the first $K-1$ pure strategies and -1 to the last pure strategy of player 2.

Let F^* be the epigraph of f^*.

Lemma 8 *F^* is a polyhedral convex set generated by extreme points x_1 which are robust best replies in Θ_1 and the directions*

$$-e_k = \left(-e_k^1, \ldots, -e_k^K\right) \in \mathbb{R}^K \text{ with } e_k^l = \begin{cases} -1 & \text{for} \quad k = l \\ 0 & \text{else} \end{cases} \tag{13}$$

for $k = 1, \ldots, K-1$ *and*

$$e = (1, \ldots, 1) \in \mathrm{IR}^K \tag{14}$$

Proof: By definition $(x_2^*, \alpha^*) \in F^*$ if and only if $\alpha^* \geq x_2^* \bullet x_2 - f^*(x_2)$ for all $x_2 \in \Theta_2$. $v \in \mathrm{IR}^K$ is a direction in F^* if and only if there exists $(x_2^*, \alpha^*) \in F^*$ such that all vectors $(x_2^*, \alpha^*) + \lambda v$ with $\lambda \geq 0$ are in F^*. We can write $v = -\sum_{k=1}^{K-1} \rho_k e_k + \rho_K e$ with $\rho_1, \ldots, \rho_K \in \mathrm{IR}$ since $-e_1, \ldots, -e_{K-1}, e$ form a vector basis of IR^K. We must show that v is a direction in F^* if and only if all ρ_i are non-negative. Suppose that v is a direction in F^*. Let $x_2 = (0, \ldots, 0) \in \Theta_2$. The condition that $(x_2^*, \alpha^*) + \lambda v \in F^*$ for all $\lambda \geq 0$ yields for this x_2 that $\alpha^* + \lambda\rho_K \geq -f(x_2)$ holds for all $\lambda \geq 0$. This can be true only if $\rho_K \geq 0$. For $e_k \in \Theta_2$ $(1 \leq k \leq K-1)$ we obtain similarly $\alpha^* + \lambda\rho_K \geq x_2^{*k} - \lambda\rho_k + \lambda\rho_K - f(e_k)$ for all $\lambda \geq 0$, which can hold only if $\rho_k \geq 0$. Thus only positive combinations of $-e_1, \ldots, -e_{K-1}, e$ can be directions in F^*. For every combination $v = -\sum_{k=1}^{K-1} \rho_k e_k + \rho_K e$ with $\rho_1, \ldots, \rho_K \geq 0$, every $\lambda \geq 0$, every $(x_2^*, \alpha^*) \in F^*$ and every $x_2 \in \Theta_2$ we have conversely

$$\alpha^* + \lambda\rho_K \geq x_2^{*\prime} x_2 - \sum_{k=1}^{K-1} \lambda\rho_k x_2^k + \lambda\rho_K - f(x_2) \tag{15}$$

which proves that v is a direction in F^*.

We have characterized the directions of F^* and must now determine the extremal points of F^*. Suppose $(\hat{x}_2^*, \hat{\alpha}^*)$ is an extremal point. Because F^* has only finitely many extremal points, these are exposed points by Straszewick's theorem (Theorem 18.6 in Rockafellar (1970)). Therefore we can find $x_2 \in \Theta_2$ such that the hyperplane $\{x_2^* \bullet x_2 = f(x_2)\}$ is a supporting hyperplane which meets F^* only in $(\hat{x}_2^*, f^*(\hat{x}_2^*))$. Because F^* has only finitely many extreme points and directions there exists an open neighborhood U of x_2 in Θ_2 for which the hyperplanes $\{x_2^* \bullet y_2 = f(y_2)\}$ are for all $y_2 \in U$ supporting hyperplanes which intersect F^* only in $(\hat{x}_2^*, f^*(\hat{x}_2^*))$. This implies that the graph of the affine function $(\hat{x}_2^*, f^*(\hat{x}_2^*))$ intersects F in a $K-1$ dimensional face. It is therefore identical to a affine function defined by a strategy x_1 in Θ_1 for which $H(x_1)$ is full dimensional. Given our identification, $(\hat{x}_2^*, f^*(\hat{x}_2^*))$ is consequently a robust strategy in Θ_1, which was to be shown. QED

The lemma implies that all extreme points and hence all the points in the compact faces of F^* are in Θ_1.

However, no points on the compact faces of F^* apart from the extreme points are robust strategies. To see this, notice that a proper mixture $x_1 = \sum_{l=1}^{L} \rho_l x_{1k}$ $(L > 2, \rho_l > 0, \sum_{l=1}^{L} \rho_l = 1)$ of non-equivalent robust strategies in Θ_1 is not robust. Otherwise there would be an open set in Θ_2 on which x_1 and hence all strategies x_{1k} were best replies. They would yield identical payoffs on an open set and were hence (by Lemma 1) all equivalent, contradicting the assumption.

Finally, consider a strategy in Θ_1 which is not on a compact face of F^*. It can be written as $x_1' = x_1 - \sum_{k=1}^{K} \rho_k e_k + \rho_K e$ where x_1 is on one of the compact faces of F^* and, hence, in Θ_1, and the ρ_k are all non-negative and at least some of them are strictly positive. We obtain

$$u_1(x_1', x_2) = u_1(x_1, x_2) - \sum_{k=1}^{K-1} \rho_k x_2^k - \rho_K \left(1 - \sum_{k=1}^{K-1} x_2^k\right) \leq u_1(x_1, x_2), \tag{16}$$

where this inequality holds as a strict one for the k-th pure strategy of player 2 whenever $\rho_k > 0$. Thus x_1' is weakly dominated. It is not a robust strategy because it is a best reply only on a proper face of Θ_1 (see Pearce (1984)).

In summary, the only robust strategies in Θ_1 are the extreme point of F^*. All other strategies are proper mixtures of non-equivalent robust strategies or are weakly dominated and therefore not robust. In particular, we have proved Lemma 4.

References

Balkenborg, D., J. Hofbauer, and C. Kuzmics (2009a): "Refined best-response correspondence and dynamics," mimeo.

——— (2009b): "The refined best-response correspondence in extensive form games," mimeo.

Basu, K., and J. W. Weibull (1991): "Strategy subsets closed under rational behavior," *Economics Letters*, 36, 141–46.

Bernheim, B. D. (1984): "Rationalizable strategic behavior," *Econometrica*, 52, 1007–29.

Bochnak, J., M. Coste, and M. Roy (1998): *Real Algebraic Geometry.* Springer.

Börgers, T. (1994): "Weak dominance and approximate common knowledge," *Journal of Economic Theory*, 64, 265–276.

Brandenburger, A. (1992): "Economic Analysis of Markets and Games," in *Economic Analysis of Markets and Games*, ed. by P. Dasgupta, D. Gale, O. Hart, and E. Maskin. MIT Press, Cambridge.

Dekel, E., and D. Fudenberg (1990): "Rational behavior with payoff uncertainty," *Journal of Economic Theory*, 52, 243–67.

Guillemin, V., and A. Pollack (1974): *Differential Topology.* Prentice Hall, Englewood Cliffs, N.J.

Harsanyi, J., and R. Selten (1988): *A general theory of equilibrium selection in games.* MIT Press, Cambridge, Massachusetts.

Hendon, E., H. J. Jacobson, and B. Sloth (1996): "Fictitious play in extensive form games," *Games and Economic Behavior*, 15, 177–202.

Herings, P. J.-J., and V. J. Vannetelbosch (1999): "Refinements of rationalizability for normal-form games," *International Journal of Game Theory*, 28, 53–68.

Jansen, M., P. Jurg, and D. Vermeulen (2002): "On the set of equilibria of a bimatrix game: A survey," in *Chapters in Game Theory*, ed. by P. Borm, and H. Peters, pp. 121–142. Kluwer, Amsterdam.

Kalai, E., and D. Samet (1984): "Persistent equilibria in strategic games," *International Journal of Game Theory*, 13, 129–44.

Kohlberg, E., and J.-F. Mertens (1986): "On the strategic stability of equilibria," *Econometrica*, 54, 1003–37.

Okada, A. (1983): "Robustness of equilibrium points in strategic games," Working Paper B-137, Tokyo Center for Game Theory.

Pearce, D. G. (1984): "Rationalizable strategic behavior and the problem of perfection," *Econometrica*, 52, 1029–51.

Ritzberger, K. (2002): *Foundations of Non-Cooperative Game Theory*. Oxford University Press.

Rockafellar, R. T. (1970): *Convex Analysis*. Princeton University Press, Orinceton, New Jersey.

Samuelson, L. (1992): "Dominated strategies and common knowledge," *Games and Economic Behavior*, 4, 284–313.

Schuhmacher, F. (1999): "Proper rationalizability and backward induction," *International Journal of Game Theory*, 28, 599–615.

Selten, R. (1975): "Re-examination of the perfectness concept for equilibrium points in extensive games," *International Journal of Game Theory*, 4, 25–55.

Swinkels, J. M. (1993): "Adjustment dynamics and rational play in games," *Games and Economic Behavior*, 5, 455–84.

van Damme, E. E. C. (1991): *Stability and Perfection of Nash Equilibria*. Springer-Verlag, Berlin, Heidelberg.