The Effect of Changes in Risk Attitude on Strategic Behavior

Jonathan Weinstein*

Washington University in St. Louis

First draft: October 2013; This version: November 2015

Abstract

We study families of normal-form games with fixed preferences over pure action profiles but varied preferences over lotteries. That is, we subject players’ utilities to monotone but non-linear transformations and examine changes in the rationalizable set and set of equilibria. Among our results: The rationalizable set always grows under concave transformations (risk aversion) and shrinks under convex transformations (risk love). The rationalizable set reaches an upper bound under extreme risk aversion, and lower bound under risk love, and both of these bounds are characterized by elimination processes. For generic two-player games, under extreme risk love or aversion, all Nash equilibria are close to pure and the limiting set of equilibria can be

*I am grateful for conversations and correspondence with Pierpaolo Battigalli, Tilman Borgers, Amanda Friedenberg and Muhamet Yildiz, and seminar participants at University of Arizona, Arizona State, University of Texas, the Stony Brook International Game Theory Festival, and the World Congress of the Econometric Society. An earlier version had the working title “The Behavior of Solution Concepts under Changes in Risk Attitude.”
described using preferences over pure action profiles; under extreme risk love, all rationalizable actions are played in some equilibrium.

1 Introduction

The standard specification of a normal-form game provides utility functions for each player, which represent Von Neumann-Morgenstern (VN-M) preferences on lotteries over action profiles. When specifying a game we may not in fact have full knowledge of players’ preferences over lotteries, but may only know their preferences over pure action profiles. This is in fact the typical situation in experiments on games with monetary payoffs: it is safe enough to assume that players prefer more money to less, but we have little knowledge of their preferences over lotteries. This paper will study the impact on strategic behavior of transformations that preserve preferences over action profiles (sometimes called ordinal preferences) but alter preferences over lotteries. Equivalently, we look at families of games generated by monotone but non-linear transformations of players’ utilities. Within each such family, we examine the varying predictions of each of rationalizability, Nash equilibrium and correlated equilibrium.

Rationalizability, introduced by Bernheim (1984) and Pearce (1984), is a solution concept which characterizes the implications of common knowledge of rationality. For finite normal-form games, it can be defined by iteratively eliminating actions which are dominated by a pure or mixed strategy. Pearce (1984) showed that this is equivalent to a process in which only actions which are best replies to some mixture (possibly correlated) over the opponents’ remaining actions survive.¹

A simple example suffices to show that the set of rationalizable actions can be affected by cardinal preferences over action profiles, even with ordinal preferences fixed:

¹See that paper for references to earlier, closely related results.
In this game, M is not rationalizable for the row player, whom we call Alice; it is dominated by an equal mixture of U and D. However, suppose we subject Alice’s utilities to the monotone transformation $\bar{u} = -e^{-u}$. One interpretation of this transformation is that the payoffs above are monetary, and Alice has gone from being risk neutral to risk averse, with a constant absolute risk aversion (CARA) specification. The new game is:

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
<td>-0.05, 0</td>
<td>-1, 0</td>
</tr>
<tr>
<td>$M$</td>
<td>-0.37, 0</td>
<td>-0.37, 0</td>
</tr>
<tr>
<td>$D$</td>
<td>-1, 0</td>
<td>-0.05, 0</td>
</tr>
</tbody>
</table>

and in this game, M is a best reply to a uniform belief over opposing actions and all actions are rationalizable. This example helps illustrate many aspects of the general theory of the behavior of rationalizability under monotone transformations, found in Section 3:

1. All actions (for both players) except M are best replies to pure actions by the opposing player. Indeed, all except M would survive an iterated procedure where only best replies to pure actions survive each round; we say that such actions are point rationalizable. This process is not sensitive to monotone transformations of either player’s utility, so all point rationalizable actions are rationalizable under all such transformations.

2. Imagine we had introduced the second game first. The action M is then rationalizable, but not point rationalizable. Under a transformation to
the first game, it is not rationalizable. This generalizes: for any ac-

tion which is not point rationalizable, there is a transformation which

makes it non-rationalizable. This result, Proposition 2, provides a com-
plete characterization of the actions which are rationalizable under all

monotone transformations.

3. In the first game, M is not dominated (even weakly) by any pure strat-

ey. It turns out that, as discovered by Borgers (1993), this is sufficient

(and generically necessary) for there to exist a transformation, e.g. the
second game, under which M is a best reply.

4. To get from the first to the second game, we applied a monotone con-
cave transformation to Alice’s utility. Equivalently, Alice became more
risk-averse in the sense that if she prefers a certain outcome to a lot-
tery in the first game, then she does in the second game also. The
rationalizable set grew. This generalizes: a monotone concave trans-
formation always makes the rationalizable set (weakly) larger, and a
convex transformation makes it smaller (Proposition 1). Loosely, com-
promise actions can be more easily rationalized by more risk-averse
players, while risky actions can be rationalized for all levels of risk
aversion because extreme beliefs are allowed.

5. The fact that we were able to use a CARA (exponential) transformation
to make M rationalizable (or not) is also completely general (Proposi-
tions 2 and 3.) Our results specify tight upper and lower bounds for the
rationalizable set as the game undergoes monotone transformations,
and shows that both bounds are achieved under CARA transforma-
tions, or indeed any family of transformations which is “unboundedly
concave” (or convex) as defined later.

After covering the effect of monotone transformations on rationalizabil-
ity (Section 3), we move on to Nash equilibria (Section 4). Here we show,
for generic two-player games, that under extreme risk love or aversion the NE set converges to a set of pure profiles, and we characterize these limiting sets. Interestingly, in either the limiting case of extremely risk-averse, non-VN-M “maximin” preferences, or the analogous limiting case of extreme risk love, equilibria may fail to exist. Here our limiting set is a good candidate to replace the equilibrium concept. We also discuss the behavior of correlated equilibria (Section 5), and for Bayesian games, interim correlated rationalizability (Section 6).

2 Preliminaries

A finite complete-information normal-form game (henceforth “game”) is a triple \( G = (N, A, u) \) where \( N = \{1, \ldots, n\} \) is a finite set of players, \( A = \prod_{i \in N} A_i \) is a finite set of (pure) action profiles, and \( u = (u_1, \ldots, u_n) \) is a profile of VN-M utilities for each player over \( A \), \( u_i : A \rightarrow \mathbb{R} \). Utilities are extended linearly to lotteries over \( A \) as usual. Call a game \( G \) generic if we always have \( u_i(a) \neq u_i(a') \) when \( a \neq a' \). As usual, \( A_{-i} = \prod_{j \neq i} A_j \); also, \( \Delta(A_{-i}) \) is the set of possible beliefs for player \( i \) over opposing action profiles, where correlated beliefs are allowed. Let \( \Sigma_i = \Delta(A_i) \) be the mixed strategies of player \( i \). For a probability distribution \( \pi \), let \( \text{supp(\pi)} \) denote the support of \( \pi \). We call a product set \( B = B_1 \times \ldots \times B_n \), where \( B_i \subseteq A_i \), a block, and the block game \( G_B \) is simply the restriction of \( G \) to action sets \( B_i \).

For any \( C \subseteq A_{-i} \), \( \Delta(C) \subseteq \Delta(A_{-i}) \) are those beliefs assigning positive probability only to elements of \( C \). We define (correlated) rationalizability by iterated elimination of never-best-replies:\(^2\)

\(^2\)Bernheim and Pearce originally defined rationalizability as constrained to independent beliefs over opposing actions, a significant constraint when \( n > 2 \). In recent years allowing correlation has become more popular, in part due to the arguments of Aumann (1987), and we do so throughout. Here, and elsewhere, we suppress the dependence of the solution concept on the game \( G \) when the intended game is clear in context, but include it, writing \( S_\infty^G \), when necessary.
\[ S_i^0 \equiv A_i \]
\[ S_i^n = \bigcup_{\sigma_{-i} \in \Delta(S_{i-1}^{n-1})} BR_i(\sigma_{-i}), \ n \geq 1 \]
\[ S_i^\infty \equiv \bigcap_{n=0}^{\infty} S_i^n \]

where \( BR_i(\sigma_{-i}) \) is the set of pure best replies to a belief \( \sigma_{-i} \in \Delta(A_{-i}) \). Here Pearce’s Lemma tells us that this procedure is equivalent (step-by-step) to eliminating actions which are strictly dominated by some pure or mixed strategy.

We define \( P^k \) to be actions surviving \( k \) rounds of pure best-reply; i.e. in each round, only actions which are a best reply to a profile of remaining pure strategies survive. Symbolically,

\[ P_i^0 \equiv A_i \]
\[ P_i^n = \bigcup_{a_{-i} \in P_{i-1}^n} BR_i(a_{-i}), \ n \geq 1 \]
\[ P_i^\infty \equiv \bigcap_{n=0}^{\infty} P_i^n \]

It is immediate that \( P^n \subseteq S^n \) for all \( n \in \mathbb{Z} \cup \infty \). As Bernheim (1984), we call \( P^\infty \) the point-rationalizable actions.

The following notion was introduced by Borgers (1993): An action \( a_i \in A_i \) is \( B \)-dominated\(^3\) if for all non-empty subsets \( T \subseteq A_{-i} \), there is an \( a_i^T \in A_i \) such that for all \( a_{-i} \in T \), \( u(a_i^T, a_{-i}) \geq u(a_i, a_{-i}) \), with strict inequality for some \( a_{-i} \in T \). Being \( B \)-dominated is implied by strict dominance by a pure action, and implies weak dominance by a pure action (the case \( T = A_{-i} \)),

\(^3\)Borgers calls this simply “dominated.”
but is not equivalent to either of these. Though a weakly dominated action can be a best reply, a $B$-dominated action cannot. To see this, let $a_i$ be any $B$-dominated action. It cannot be a best reply to any belief $\sigma_{-i} \in \Delta(A_{-i})$, because if we let $T$ be the support of $\sigma_{-i}$, $a_i^T$ will outperform $a_i$ against $\sigma_{-i}$. Let $B_i^k \subseteq A_i$ be actions surviving $k$ rounds of iterated $B$-dominance, and $B_i^\infty \equiv \bigcap_{n=0}^\infty B_i^n$. Iterating the argument above shows that $S^n \subseteq B^n$ for all $n \in \mathbb{Z} \cup \infty$.

We write $G \sim G'$ if $G$ and $G'$ have identical player and action sets, and players have identical ordinal preferences, i.e. identical preferences over pure action profiles. Now, we have argued earlier that $P^n(G) \subseteq S^n(G) \subseteq B^n(G)$, and it is immediate from the definitions that $P^k$ and $B^k$ only depend on ordinal preferences, i.e. they are constant on each equivalence class under $\sim$. So, we have the sandwich relationship:\(^5\)

\[ \forall G' \sim G, \forall n \in \mathbb{Z} \cup \infty : P^n(G) \subseteq S^n(G') \subseteq B^n(G) \]

Borgers showed that the upper bound is tight, in the sense that every action in $B^n(G)$ can be rationalized by some beliefs consistent with common knowledge of ordinal preferences. We will show, in Proposition 3, that it is tight in the slightly stronger sense that there is a specific complete-information game $G' \sim G$ such that $S^\infty(G') = B^\infty(G)$. The constructed $G'$ can be described simply as any transformation of $G$ in which players are sufficiently risk-averse. In a complementary fashion, Proposition 2 states that any $G'$ in which players are sufficiently risk-loving achieves the lower bound $S^\infty(G') = P^\infty(G)$.

To formalize the ideas of increasing risk aversion or risk love, we introduce

---

\(^4\)For generic games as defined above, $B$-dominance is equivalent to dominance (strict or weak) by a pure action.

\(^5\)In fact, this relationship holds for a weaker notion of similarity, in which ordinal preferences are fixed conditional on each opposing profile. That is, let $(N, A, u) \approx (N, A, u')$ if $\forall i, a_i, a_i' : u_i(a_i, a_{-i}) > u_i(a_i', a_{-i}) \Leftrightarrow u_i'(a_i, a_{-i}) > u_i'(a_i', a_{-i})$. Then $P^k$ and $B^k$ are well-defined on equivalence classes under $\approx$, so the sandwich relationship here, as well as Propositions 2 and 3, are still true with $\approx$ replacing $\sim$. 

the following definitions:

Fixing $G = (N, A, u)$, let $G^r = (N, A, u^r), r \in (0, \infty)$ be a family of games with $G^r \sim G$. We say that this family is unboundedly convex if:

1. For any $r > s$ and all $i \in N$, $u^r_i = f_{i,r,s} \circ u^s_i$ for some increasing, convex $f_{i,r,s}$.

2. For any $\pi, \pi' \in \Delta(A)$ with $\max_{a \in \text{supp}(\pi)} u_i(a) > \max_{a \in \text{supp}(\pi')} u_i(a)$, $u^r_i(\pi) > u^r_i(\pi')$ for all sufficiently large $r$.

Condition 2 is equivalent to the condition that for every $i \in N$ and $a, a', a'' \in A$ with $u_i(a) > u_i(a') > u_i(a'')$:

$$\lim_{r \to \infty} \frac{u^r_i(a') - u^r_i(a'')}{u^r_i(a) - u^r_i(a')} = 0$$

In particular, letting $u^r_i(a) = e^{ru_i(a)}$ produces such a family. In words, condition 2 says that for large enough $r$, preferences are dominated by the likelihood of the largest prize.

Similarly, we say that the family $G^{-r} = (N, A, u^{-r}), r \in (0, \infty)$ with $G^{-r} \sim G$ is unboundedly concave if:

1. For any $r > s$ and all $i \in N$, $u^{-r}_i = f_{i,r,s} \circ u^{-r}_i$ for some increasing, concave $f_{i,r,s}$.

2. For any $\pi, \pi' \in \Delta(A)$ with $\min_{a \in \text{supp}(\pi)} u_i(a) < \min_{a \in \text{supp}(\pi')} u_i(a)$, $u^{-r}_i(\pi) > u^{-r}_i(\pi')$ for all sufficiently large $r$.

In words, Condition 2 says that for large enough $r$, preferences are dominated by the smallest possible prize in each lottery, approaching a “max-min” criterion. Condition 2 is equivalent to the condition that for every $i \in N$ and $a, a', a'' \in A$ with $u_i(a) > u_i(a') > u_i(a'')$:

$$\lim_{r \to \infty} \frac{u^r_i(a) - u^r_i(a')}{u^r_i(a') - u^r_i(a'')} = 0$$
In particular, letting \( u'_i(a) = -e^{-ru_i(a)} \) produces such a family. If the original game \( G \) has monetary payoffs, these utilities are of course well-known as representing Constant Absolute Risk Aversion (CARA) with parameter \( r \). The convex case, \( u'_i(a) = e^{ru_i(a)} \), could correspondingly be termed constant absolute risk love. I suggest that the reader keep these examples in mind when interpreting the coming results.

Finally, the phrase “for all sufficiently large \( r \),” occasionally abbreviated to “for large \( r \),” means that there is an \( R \) such that the statement in question is true for all \( r > R \).

3 Behavior of Rationalizability

We discussed the intuition for the following result in the introduction:

**Proposition 1.** Let \( G = (N, A, u) \) be a finite complete-information normal-form game, and let \( G' = (N, A, v) \) where \( v_i = f_i \circ u_i \) for some strictly increasing, convex functions \( f_i : \mathbb{R} \to \mathbb{R} \). Then \( S^k(G') \subseteq S^k(G) \) for each \( k \in \mathbb{N} \cup \{\infty\} \). The reverse inclusion holds for increasing, concave \( f_i \).

**Proof:** The case \( k = 0 \) is trivial. Suppose the claim is true for \( k \), but false for \( k + 1 \), so that we have \( a_i \in S^k_i(G') - S^{k+1}_i(G) \). Since \( S^{k+1}_i(G') \subseteq S^k(G) \subseteq S^k(G') \), we have \( a_i \in S^k_i(G) - S^{k+1}_i(G) \), so by the equivalence of rationalizability with iterated dominance, there must be a mixture \( \sigma_i \in \Delta(S^k_i(G)) \) with \( u_i(\sigma_i, a_{-i}) > u_i(a_i, a_{-i}) \) for all \( a_{-i} \in S^{k+1}_{-i}(G) \). Then convexity and monotonicity of \( f_i \) give

\[ v_i(\sigma_i, a_{-i}) \geq f(u_i(\sigma_i, a_{-i})) > f(u_i(a_i, a_{-i})) = v_i(a_i, a_{-i}) \]

for all \( a_{-i} \in S^k_{-i}(G) \), a fortiori for all \( a_{-i} \in S^k_{-i}(G') \). This contradicts \( a_i \in S^{k+1}_i(G') \) and completes the induction step. The case \( k = \infty \) follows immediately from the finite case.
The concave case follows from the fact that the inverse\(^6\) of an increasing, concave function is increasing and convex. □

In contemporaneous work, Battigalli et. al. (2015) study the effects of both risk attitude and ambiguity attitude on the set of justifiable actions in decision problems. Their main results focus on a single player. In other respects they are more general forms of Proposition 1, concerning a concave transformation of the DM’s attitude towards ambiguity (in the smooth model) or risk, and including the case of infinite action spaces with appropriate continuity and compactness assumptions. As here, the results are a consequence of Pearce’s lemma, which they extend to a more general form needed in their setting. Pearce’s proof of this key lemma uses the minimax theorem for finite games, and by maintaining the necessary assumptions on infinite games for which generalized minimax results hold, Battigalli et. al. are able to extend Pearce’s lemma and hence Proposition 1 to their setting.

The next proposition shows that for extreme risk love, the elimination process for rationalizability becomes precisely the process for point rationalizability. That is, in our notation, \(S^k\) becomes \(P^k\). The proof will use the characterization of \(S^\infty\) via iterated dominance; our strategy is to show that for a very “risk-loving” player, a uniform mixture over all actions dominates any action which is not a best reply to any pure strategy. Viewed through the lens of best replies, the intuition is that extremely risk-loving players essentially use a “maximax” criterion: they will always “bet” on a single action profile by the other players and play a pure best-reply to that profile.

**Proposition 2.** Let \(G = (N, A, u)\) be a game and \(G^r \sim G\) be any unboundedly convex family of transformations of \(G\). Then:

1. As \(r \to \infty\), \(S^\infty(G^r)\) shrinks monotonically to \(P^\infty(G)\).

\(^6\)The inverse may not be defined on all of \(\mathbb{R}\), but will be defined on the full interval of payoffs in our finite game, and this suffices.
2. \[ \bigcap_{G' \sim G} S^\infty(G') = P^\infty(G) \]

**Proof:** Monotonicity is immediate from Proposition 1. We will show that for sufficiently large \( r \), \( S^k(G^r) = P^k(G) \) for all \( k \in \mathbb{N} \cup \{\infty\} \). Fix an action \( a_i \notin P^\infty_i(G) \) and suppose it is eliminated at step \( k + 1 \), i.e. \( a_i \in P^k_i(G) - P^{k+1}_i(G) \). This means that for every \( a_{-i} \in P^k_i(G) \), there is an action \( a_i(a_{-i}) \in P^k_i(G) \) with \( u_i(a_i(a_{-i}), a_{-i}) > u_i(a_i, a_{-i}) \). Let \( \sigma_i \) be a mixture putting equal weight on each action in \( P^k_i(G) \). Then by the definition of unbounded convexity, for each \( a_{-i} \in P^k_i(G) \) we have, for sufficiently large \( r \), \( u^r_i(\sigma_i, a_{-i}) > u^r_i(a_i, a_{-i}) \). Since the game is finite, we can find \( r \) large enough for every such \( a_{-i} \in P^k_i(G) \). Again by finiteness, for \( r \) large enough, all \( a_i \notin P^\infty_i(G) \) will be dominated at the appropriate stage. We have just showed that for sufficiently large \( r \), only pure best replies survive at each stage of the rationalizability process in \( G^r \), which is the claim \( S^k(G^r) = P^k(G) \) for finite \( k \). The case \( k = \infty \) follows immediately.

In part 2, the inclusion \( P^\infty(G) \subseteq \bigcap_{G' \sim G} S^\infty(G') \) is immediate, since best replies to pure action profiles are preserved under monotone transformations, and the opposite inclusion is implied by part 1. □

Next, we establish the upper bound on the actions that may be rationalizable in some transformed game. As mentioned earlier, this upper bound is achieved in the limit as the players become extremely risk averse.\(^7\)

**Proposition 3.** Let \( G = (N, A, u) \) be a game and \( G^{-r} \sim G \) be any unboundedly concave family of transformations of \( G \). Then:

1. As \( r \to \infty \), \( S^\infty(G^{-r}) \) grows monotonically to \( B^\infty(G) \).

2. \[ \bigcup_{G' \sim G} S^\infty(G') = B^\infty(G) \]

\(^7\)The \( n = 1 \) version of this result, without mention of monotonicity, was discovered in contemporaneous work by Gafarov-Salcedo (2014).
Proof: Monotonicity is immediate from Proposition 1. We established the inclusion $S^\infty(G') \subseteq B^\infty(G)$ for any $G' \sim G$ in the preliminaries. Since the game is finite, we must have some $k$ such that $B^k(G) = B^\infty(G)$, so that $B^\infty(G)$ is a fixed point of the $B$-elimination process. We will show that for some $r$, $B^\infty(G)$ is also a fixed point of the process defining $S^\infty(G^r)$, which implies $B^\infty(G) \subseteq S^\infty(G^r)$ and hence the desired results.

Suppose to the contrary, that $B^\infty(G)$ is not a fixed point for any $r$. Let $\tilde{G}^r$ be the game with payoffs as in $G^r$ and action sets restricted to those in $B^\infty(G)$, so that no action is $B$-dominated in $\tilde{G}^r$. By monotonicity there must be an action $\hat{a}_i \in B^\infty_i(G)$ which is never a best reply in $\tilde{G}^r$ for any $r$, no matter how large. Applying Pearce’s lemma, for each $r > 0$ let $\pi^r$ be a mixture on $B^\infty_i(G)$ which strictly dominates $\hat{a}_i$ in $\tilde{G}^r$. The fact that $\hat{a}_i$ is not $B$-dominated in $\tilde{G}^r$ means that there is a non-empty $T \subseteq B^\infty_i(G)$ such that every $a'_i \in B^\infty_i(G)$ either satisfies (i) $u(a'_i, a_{-i}) = u(\hat{a}_i, a_{-i})$ for all $a_{-i} \in T$ or (ii) $u(a'_i, a_{-i}) < u(\hat{a}_i, a_{-i})$ for some $a_{-i} \in T$. By strict dominance, the mixture $\pi^r$ cannot have any actions of type (i); by throwing out the actions of type (i) and renormalizing, we can obtain for each $r$ a mixture $\phi^r$ which assigns positive mass only to actions $a'_i$ satisfying (ii) and satisfies $u_i^r(\phi^r, a_{-i}) > u_i^r(\hat{a}_i, a_{-i})$ for all $a_{-i} \in T$.

Then, fixing any $\varepsilon \in (0, 1/|A_i|)$, there must be some $a'_i \in B^\infty_i(G)$ and $a_{-i} \in T$ for which $\phi^r(a'_i) > \varepsilon$ for an unbounded set of $r$ and $u_i(a'_i, a_{-i}) < u_i(\hat{a}_i, a_{-i})$. Let $\bar{a}$ be player $i$’s most preferred outcome, and let $\mu$ be the lottery with $\mu(\bar{a}) = 1 - \varepsilon, \mu(a'_i, a_{-i}) = \varepsilon$. Clearly $u_i^r(\phi^r, a_{-i}) < u_i^r(\mu)$ for the unbounded set of $r$ where $\phi^r(a'_i) > \varepsilon$. Also, by definition of unbounded concavity, for all sufficiently large $r$ we have $u_i^r(\mu) < u_i^r(\hat{a}_i, a_{-i})$. Then for some $r$, $u_i^r(\phi^r, a_{-i}) < u_i^r(\hat{a}_i, a_{-i})$, contradicting the definition of $\phi^r$. □

As discussed in the previous section, this result includes the main result of Borgers (1993). Our proof is made more straightforward by focusing on a family of concave transformations and exploiting Pearce’s Lemma, while Borgers explicitly constructs a transformation and belief to to make an arbi-
trary action in $B^\infty$ a best reply.

4 Behavior of Nash Equilibrium

The next natural line of inquiry concerns the behavior of the set of Nash equilibria as we apply monotone transformations to a game. Here also, there are some striking regularities when we look at a family of convex or concave transformations (such as the exponential family). Restricting attention to generic two-player games, we will be able to state a precise description of the limiting set of equilibria under extreme risk love (Proposition 4) and extreme risk aversion (Proposition 5). Please note that while our results show existence of these limits, if we allow preferences to go to the extreme of infinite risk aversion (max-min preferences), then Nash equilibria may not exist. Propositions 4 and 5 do not hold for more than two players; see counterexamples and discussion in the appendix.

In order pin down the limiting behavior of the set of Nash equilibria under extreme risk love or aversion, we will ask which “blocks” $S \times T$ are possible supports of Nash equilibria, where $S \subseteq A_1, T \subseteq A_2$. For the case of risk-love, the key will be:

**Definition 1.** A block $S \times T$ has the point best-reply property if every action in $S$ is a best-reply (in the full game) to some action in $T$ and vice-versa.

In generic games, the point best-reply property implies that $|S| = |T|$; indeed, the players’ best-reply correspondences must give (possibly distinct) bijections between $S$ and $T$. Notice that $P^\infty(G)$ will always have the point best-reply property, and a singleton block has the property if and only if it is a pure NE.

**Proposition 4.** Fix a generic two-player game $G$, an unboundedly convex family of transformations $G^r$, and subsets $S \subseteq A_1, T \subseteq A_2$. 
1. If $S \times T$ has the point best-reply property, then for all sufficiently large $r$ there is a NE of $G^r$ with support $S \times T$. As $r \to \infty$, all equilibria $(\sigma_1^r, \sigma_2^r)$ of $G^r$ with support $S \times T$ converge to the pure profile $(\tilde{a}_1, \tilde{a}_2)$ where

$$
\tilde{a}_1 = \arg\min_{a_1 \in S} \left[ \max_{a_2 \in T} u_2(a_1, a_2) \right]
$$

$$
\tilde{a}_2 = \arg\min_{a_2 \in T} \left[ \max_{a_1 \in S} u_1(a_1, a_2) \right]
$$

2. If $S \times T$ lacks the point best-reply property, then for all sufficiently large $r$ there is no NE of $G^r$ with support $S \times T$.

The second item follows from the same argument as the proof of Proposition 2; indeed, if an $a_1$ is not a best reply to any $a_2 \in T$, then for large $r$ it is not a best reply in $G^r$ to any belief over $T$. The first item, existence, is shown with the aid of a general lemma on existence of Nash equilibrium with given supports. An intuition for our identification of the limiting equilibria is that in the limit, each player is mainly interested in chasing the largest payoff possible, so to make (say) Player 2 indifferent, Player 1 must offer a very high probability that the least enticing of his best responses is optimal. All of these results are restricted to the two-player case, for reasons discussed in the proofs, in the appendix. For more than two players, necessity of the point best-reply property follows by the same logic, but sufficiency does not hold: See in the appendix a generic three-player example where every action is a point best-reply but there is no totally mixed equilibrium for large $r$, as well as an example where mixed equilibria do not approach pure profiles for large $r$.

Keeping our focus on generic two-player games, a similar result obtains for extreme risk aversion, with a slightly more complicated statement. It turns out that for extremely risk-averse players, what kills an equilibrium on $S \times T$ is the presence of an action $a_1$ which is not the worst reply among $S \cup \{a_1\}$ to any $a_2 \in T$ (or the analogue with roles reversed.) Indeed, one direction of our
proof verifies that for a sufficiently risk-averse player, such an $a_1$ dominates a uniform mixture over $S$ when Player 2 is restricted to $T$, preventing any full-support mixture on $S$ from being a best-reply. Accordingly, we make these definitions:

**Definition 2.** A never-worst deviation from a block $S \times T$ is an $a_1 \in A_1$ such that

$$\forall a_2 \in T : \exists a'_1 \in S : u_1(a_1, a_2) > u_1(a'_1, a_2)$$

or an $a_2$ with roles reversed. We say $S \times T$ allows no safety if there is no never-worst deviation from $S \times T$.

Notice that an $a_1 \in S$ is a never-worst deviation from $S \times T$ exactly when it is not a point worst-reply in the block game $G_{S \times T}$. In particular, under genericity, if $|S| > |T|$ there will be a never-worst deviation $a_1 \in S$. Therefore, when the block $S \times T$ allows no safety, we must have $|S| = |T|$, and the “worst-reply correspondences” of the block game must each form a bijection between $S$ and $T$. Observe also that a singleton block allows no safety exactly when it is a pure NE.

**Proposition 5.** Fix a generic two-player game $G$, an unboundedly concave family of transformations $G^{-r}$, and subsets $S \subseteq A_1, T \subseteq A_2$.

1. If $S \times T$ allows no safety, then for all sufficiently large $r$ a NE of $G^{-r}$ with support $S \times T$ exists. As $r \to \infty$, the equilibria $(\sigma_1^{-r}, \sigma_2^{-r})$ of $G^{-r}$ with supports $S, T$ converge to $(\tilde{a}_1, \tilde{a}_2)$ where

$$\tilde{a}_1 := \text{argmax}_{a_1 \in S} \left[ \min_{a_2 \in T} u_2(a_1, a_2) \right]$$

$$\tilde{a}_2 := \text{argmax}_{a_2 \in T} \left[ \min_{a_1 \in S} u_1(a_1, a_2) \right]$$

2. Otherwise, for all sufficiently large $r$ no NE of $G^{-r}$ with support $S \times T$ exists.
Let $NE(G)$ be the set of NE profiles of $G$, and apply the Hausdorff metric\(^8\) to compact sets of (mixed) profiles. From Propositions 4 and 5 we can immediately deduce the following:

**Proposition 6.** For every generic two-player game $G$, the sets $NE(G^r), NE(G^{-r})$ each converge to a set of pure profiles as $r \to \infty$. The limiting sets are the same for all $G' \sim G$ and do not depend on the particular families of transformations $G^r, G^{-r}$.

That is, for extreme risk love or extreme risk aversion, the set of equilibria approaches a set of extreme points in the space of mixed strategies. These sets can easily be found by applying Propositions 4 and 5: first check each square block $S \times T$ to see if it has an equilibrium in the limit, then each proposition tells us which extreme point is the limiting equilibrium in this block. The criteria determining these limits depend only on ordinal data, i.e. on preferences over pure action profiles.

Since pure NE are preserved by monotone transformations, we know that in a game with no pure NE, the limit points will not be pure NE, as in this example, a generic form of Matching Pennies:

<table>
<thead>
<tr>
<th></th>
<th>$L$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
<td>2, −2</td>
<td>−1, 1</td>
</tr>
<tr>
<td>$D$</td>
<td>−2, 2</td>
<td>1, −1</td>
</tr>
</tbody>
</table>

This generic 2x2 game has no pure equilibria, and one mixed, for any monotone transformation. The limiting NE are $NE(G^r) \to \{(U, R)\}, NE(G^{-r}) \to \{(D, R)\}$ as $r \to \infty$. Note that for max-min (non-VN-M) decision-makers, the limiting case of risk aversion, there is no equilibrium, because no mixture makes either player indifferent. The limit of $NE(G^{-r})$ is a good candidate for a solution concept for this case. The limiting case of extreme risk love, decision makers who believe they will receive their maximum possible payoff,

---

\(^8\)The Hausdorff distance between two sets is the supremum of the distances from each point in each set to the nearest point in the other set.
or “wishful thinkers,” are studied in Yildiz (2007). He finds that in examples such as this one (generic two-player games with no pure NE), even the analogue of rationalizability is empty: no profile is consistent with common knowledge of wishful thinking. The limit of $NE(G^r)$ is a good candidate for a solution concept in this case.

The fact that equilibria converge to a non-equilibrium may appear to violate upper hemicontinuity of the equilibrium correspondence. It does not, because payoffs are unbounded and the games $G^r$ do not have a convergent subsequence. If we normalized payoffs to have range $[0, 1]$ after each transformation, then upper hemicontinuity would apply: we would find that for each $\varepsilon$, $(U, R)$ is a $\varepsilon$-equilibrium of the normalized version of $G^r$ for large enough $r$, and the payoff matrix would converge to:

<table>
<thead>
<tr>
<th></th>
<th>$L$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
<td>1, 0</td>
<td>0, 0</td>
</tr>
<tr>
<td>$D$</td>
<td>0, 1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

where $(U, R)$ is an equilibrium. Note, though, that the other pure NE of this limit game, $(U, L)$, is not a limit point of equilibria of $G^r$, and hence one cannot identify the limiting set simply by finding NE of the limit game; this is a standard failure of lower hemicontinuity of the equilibrium correspondence.

If pure NE exist, of course they will be among the limit points, but there may be others, as in this example:

<table>
<thead>
<tr>
<th></th>
<th>$L$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
<td>3, 1</td>
<td>0, 0</td>
</tr>
<tr>
<td>$D$</td>
<td>2, 2</td>
<td>1, 3</td>
</tr>
</tbody>
</table>

This generic 2x2 game, $G$, has two pure equilibria, $(U, L)$ and $(D, R)$, and one mixed. Any monotone transformation, of course, preserves the pure equi-

---

9Upper hemicontinuity tells us that the graph of the correspondence from games to NE is closed, or in other words that given a convergent sequence of NE of a convergent sequence of games, the limit of the NE is a NE of the limiting game.
libria and also has one mixed equilibrium. $NE(G^r) \rightarrow \{(U, L), (D, R), (U, R)\}$, $NE(G^{-r}) \rightarrow \{(U, L), (D, R), (D, L)\}$ as $r \rightarrow \infty$. Proposition 4 tells us that in the limit of extreme risk love, in any full-support equilibrium, each player almost always chooses the action which gives the opponent the lowest possible maximum (i.e. the minmaxing action if only pure actions are allowed.) Conversely, in the limit of extreme risk aversion, the weight goes to the action which maximizes the opponent’s minimum payoff.

5 Correlated Equilibria and Joint Coherence

Let $J(G)$ denote the set of strategy profiles which are played with positive probability in some correlated equilibrium (CE) of $G$. By Aumann (1987) we know that these are the profiles which are consistent with common knowledge of rationality together with a common prior. The added power of the common-prior assumption gives a refinement of rationalizability; we will shortly see an example where the refinement is strict. Nau and McCardle (1990) called $J(G)$ the “jointly coherent” profiles and showed that players who do not always play profiles in $J(G)$ are susceptible to a form of arbitrage against the group of players: an outside observer can find a set of bets, each accepted by some player, which gives him positive payoff whenever an element of $A - J(G)$ is played and zero payoff on $J(G)$. Their result is another argument for $J(G)$ as a natural notion of joint rationality. Unlike rationalizability, joint coherence may grow under monotone convex transformations of the players’ utility functions, as the following example shows. We call this game “Matching Pennies with Entry Fee”:

<table>
<thead>
<tr>
<th></th>
<th>$H$</th>
<th>$T$</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>10, −10</td>
<td>−10, 10</td>
<td>0, 1</td>
</tr>
<tr>
<td>$T$</td>
<td>−10, 10</td>
<td>10, −10</td>
<td>0, 1</td>
</tr>
<tr>
<td>$S$</td>
<td>1, 0</td>
<td>1, 0</td>
<td>1, 1</td>
</tr>
</tbody>
</table>
The intended interpretation of the game is that two players walk into a casino with $1 each, and may either play Safe and keep their dollar, or each pay the casino $1 for the privilege of playing Matching Pennies against each other for a $10 stake. Although all strategies are rationalizable here, the only jointly coherent profile is \((S, S)\). That is, the unique CE puts probability 1 on \((S, S)\). This game gives an economic intuition for the impact of the common-prior assumption which is embedded in equilibrium concepts: players with a common prior cannot both believe they will win the zero-sum block game on \(\{H, T\} \times \{H, T\}\), an argument similar to those for no-trade theorems.\(^{10}\)

However, consider the game \(G^1\) formed by applying the monotone convex transformation \(f(u) = e^u\) to both players’ utilities. There is a Nash equilibrium of \(G^1\) in which Alice mixes equally U and D and Bob mixes equally L and R, showing that \(J(G) \subset J(G^1)\). Indeed, risk-loving players may, even in equilibrium, pay a casino a premium to play zero-sum games such as poker against one another. In fact, \(J(G^1) = A\), i.e. all profiles are jointly coherent\(^{11}\). Though this example is non-generic, its behavior is robust to small payoff perturbations. For generic games, Proposition 7 generalizes the example. It says that for extreme risk love, joint coherence and rationalizability coincide, which we can interpret epistemically to mean that the common-prior assumption has no bite. An intuitive justification is that extreme risk love makes one a “wishful thinker” who chases the highest possible payoff, hence a restriction to identical beliefs becomes irrelevant – only the support of beliefs matters.

**Proposition 7.** For any generic two-player game \(G\) and unboundedly convex transformations \(G^r\), for all sufficiently large \(r\) we have \(J(G^r) = S^\infty(G^r) = \)

\(^{10}\)This example has the same properties as one used in Bernheim (1984) to demonstrate that some rationalizable profiles are not played in any Nash equilibrium. I transformed the payoffs to give the game a better story. Nau-McCardle noted the uniqueness of CE as well as NE in Bernheim’s example.

\(^{11}\)The full-support distribution \(p\) with \(p(M, L) = p(M, R) = p(U, C) = p(D, C) = \varepsilon^2\), \(p(M, C) = \varepsilon\), and \(p(U, L) = p(U, R) = p(D, L) = p(D, R) = .25 - \varepsilon/4 - \varepsilon^2\) is a CE for small enough \(\varepsilon > 0\).
Proof: The latter equality is exactly Proposition 2. Now apply Proposition 4 to the block $P^\infty(G^r)$ of $G^r$, which has the point best-reply property. It guarantees an equilibrium $\sigma$ of $G^r$ with support $P^\infty(G^r)$ for large $r$. □

The proof only uses Nash equilibria, so in the proposition the extra freedom of CE over NE in the definition of $J(G)$ was not actually needed.

Now, because the set of CE is convex, we cannot hope for Proposition 6 to extend perfectly to correlated equilibria. Any time there are multiple pure NEs, for instance, the set of CEs will include the convex hull of these NEs for all $r$. However, we do conjecture, again for generic $G$, that the the polytope of CEs of $G^r$, or $G^{-r}$, converges in Hausdorff topology to a convex hull of pure profiles as $r \to \infty$. This conjecture is not resolved at this time. An equivalent statement of this conjecture is that for each profile $a$, the maximum weight assigned to $a$ in any CE of $G^r$ goes to 0 or 1 as $r \to \infty$.

6 Incomplete information on Cardinal Payoffs

To this point, we have been analyzing complete-information games. The natural interpretation of the complete-information approach is that the exact preferences are commonly known to all the players, but not to a modeler who tries to predict their behavior knowing only preferences over pure action profiles. Perhaps a more natural assumption would be that the players themselves know only each others preferences over action profiles and not over lotteries. It happens that this generalization does not affect the upper or lower bounds on the rationalizable set.

To address this point we will need to introduce monotone elaborations of a complete-information game. These are Bayesian games in which preferences over pure action profiles are fixed. Formally:
Definition: Given a complete-information game $G = (N, A, u)$, a Bayesian game $\hat{G} = (N, A, \hat{u}, \Theta, T, \kappa)$ is a monotone elaboration of $G$ if for all $i \in N, a, a' \in A$ and $\theta \in \Theta$,

$$u_i(a) > u_i(a') \iff \hat{u}_i(a, \theta) > \hat{u}_i(a', \theta)$$

Our solution concept for incomplete-information games will be interim correlated rationalizability (ICR), as introduced in Dekel, Fudenberg and Morris (2007). A definition of this concept, and the proof of the following proposition, are found in Appendix B.

**Proposition 8.** For any monotone elaboration $\hat{G}$ of $G$ and every type $t_i$ of every player $i$, $P_i^\infty(G) \subseteq ICR_{\hat{G}}(t_i) \subseteq B_i^\infty(G)$

### 7 Appendix A

Additional notation: Given a subset $T$ of $A_{-i}$, write $\sigma_i >_T \sigma'_i$ if $u_i(\sigma_i, a_{-i}) \geq u_i(\sigma'_i, a_{-i})$ for all $a_{-i} \in T$, with strict inequality for some $a_{-i} \in T$; that is, if $\sigma_i$ weakly dominates $\sigma'_i$ when opposing profiles are restricted to $T$.

The following lemma, used in proofs of Propositions 4 and 5, is Corollary 2 in Weinstein (2015):

**Lemma 1.** A two-player game has a Nash equilibrium with supports $S \subseteq A_1, T \subseteq A_2$, if and only if there is no pair $\sigma_1 \in \Delta(A_1), \sigma'_1 \in \Delta(S)$ such that $\sigma_1 >_T \sigma'_1$ and no pair $\sigma_2 \in \Delta(A_2), \sigma'_2 \in \Delta(T)$ such that $\sigma_2 >_S \sigma'_2$. When such pairs do exist, they can be chosen with disjoint support, i.e. $\sigma_1(a_1)\sigma'_1(a_1) = 0$ for all $a_1$.

No analogous lemma holds when $n > 2$; the proof relies on the fact that fixing $S$ and $T$, the problem of finding an equilibrium with support $S \times T$ is a linear programming problem. When $n > 2$, products of probabilities are involved and the technique breaks down. Our proofs use $n = 2$ in other
places as well. Counterexamples to Propositions 4 and 5 for \( n > 2 \) are at the end of this section.

**Proof of Proposition 4:**

**Claim 1, existence:** Suppose \( S \times T \) has the point best-reply property. We show that for sufficiently large \( r \), the second condition in Lemma 1 holds in \( G^r \), whence the lemma yields an equilibrium. Fix any \( a_1 \in S \); it is a best response to some \( a_2 \in T \). Let \( a_1' \) be the second-best response to \( a_2 \), and \( a_1'' \) the worst response. Fix \( \varepsilon \in (0, 1/|S|) \). By definition of unboundedly convex, for all sufficiently large \( r \) we have

\[
\varepsilon u_1^r(a_1, a_2) + (1 - \varepsilon)u_1^r(a_1'', a_2) > u_1^r(a_1', a_2) \tag{1}
\]

Suppose \( \sigma_1 >_T \sigma_1' \) with \( \sigma_1 \in \Delta(A_1), \sigma_1' \in \Delta(S) \) having disjoint support, as in the lemma. Suppose \( \sigma_1'(a_1) \geq \varepsilon \). Disjoint support implies \( \sigma_1(a_1) = 0 \), so we have

\[
u_1^r(\sigma_1', a_2) \geq \varepsilon u_1^r(a_1, a_2) + (1 - \varepsilon)u_1^r(a_1'', a_2) > u_1^r(a_1', a_2) \geq u_1^r(\sigma_1, a_2)
\]

contradicting weak dominance. So \( \sigma_1'(a_1) < \varepsilon \). Letting \( r \) be large enough that (1) holds for all \( a_1 \in S \), \( \sigma_1'(a_1) < \varepsilon < 1/|S| \) for all \( a_1 \), so \( \sum_{a_1 \in S} \sigma_1'(a_1) < 1 \), a contradiction. Also take \( r \) large enough for the analogous conditions to hold with roles reversed, and we have proven the existence for such \( r \) of an equilibrium with support \( S \times T \).

**Convergence of equilibria of \( G^r \):** Let \( k = |S| = |T| \). Fix \( \varepsilon > 0 \), define \( \bar{a}_2 = \arg\min_{a_2 \in T} \left[ \max_{a_1 \in S} u_1(a_1, a_2) \right] \) as in the proposition, and let \( \bar{a}_1 \) be the best reply to \( \bar{a}_2 \). Consider the set \( P_\varepsilon := \{ \sigma_2 \in \Delta(A_2) : \sigma_2(\bar{a}_2) \leq 1 - \varepsilon \} \). We claim that for all sufficiently large \( r \), in \( G^r \) action \( \bar{a}_1 \) is not a best reply to any belief in \( P_\varepsilon \). To prove this secondary claim: Viewed as a subset of \( \mathbb{R}^k \), \( P_\varepsilon \) is bounded and defined by finitely many linear inequalities, hence \( P_\varepsilon \) is a polytope and is the convex hull of its finite set of extreme points. These
extreme points of $S$ are\textsuperscript{12} the $k - 1$ points where $\sigma_2(\hat{a}_2) = 1$, for some $\hat{a}_2 \neq \tilde{a}_2$, and the $k - 1$ points where $\sigma_2(\hat{a}_2) = \varepsilon, \sigma_2(\tilde{a}_2) = 1 - \varepsilon$, for some $\hat{a}_2 \neq \tilde{a}_2$. Consider the mixed strategy $\sigma_1$ with $\sigma_1(\hat{a}_1) = 1/(k-1)$ for each of the $k - 1$ actions $\hat{a}_1 \neq \bar{a}_1$. Let $\hat{\sigma}_2 \in P_\varepsilon$ be one of the extreme points of $P_\varepsilon$, so $\hat{\sigma}_2(\hat{a}_2) \geq \varepsilon$ for some $\hat{a}_2 \neq \tilde{a}_2$. Let $\hat{a}_1$ be the best response to $\hat{a}_2$. Now

$$u_1^r(\sigma_1, \hat{\sigma}_2) > u_1^r(\bar{a}_1, \hat{\sigma}_2)$$ \hspace{1cm} (2)

for all sufficiently large $r$, because the right-hand lottery has support contained in $\{(\bar{a}_1, \hat{a}_2), (\bar{a}_1, \tilde{a}_2)\}$, both of which have lower utility than $(\hat{a}_1, \hat{a}_2)$ which is contained in the left-hand lottery. Thus, since the set of extreme points $\hat{\sigma}_2$ is finite, for large enough $r$ the inequality (2) holds for all extreme points $\hat{\sigma}_2$ of $S$. We then conclude by linearity that (2) holds for all $\hat{\sigma}_2$ in the convex hull of these extreme points, which is all of $P_\varepsilon$. Then $\bar{a}_1$ is not a best reply to any belief in $P_\varepsilon$, as desired. This implies that, given $\varepsilon$, for sufficiently large $r$ any equilibrium $(\sigma_1, \sigma_2)$ with support $S \times T$ must have $\sigma_2 \notin P_\varepsilon$, which together with the corresponding claim for $\sigma_1$ gives the desired result.

Claim 2, non-existence: Suppose $a_1 \in S$ is not a best reply to any $a_2 \in T$. Let $\sigma_1$ be an equal mixture of all best replies to elements of $T$. We claim $\sigma_1 >_T a_1$ in $G^r$ for all sufficiently large $r$. Indeed, for $a_2 \in T$,

$$u^r(\sigma_1, a_2) > u^r(a_1, a_2)$$

for sufficiently large $r$, because the left-hand lottery has a higher maximum. The claim follows by finiteness of $T$, and the desired result by the lemma.

□

Proof of Proposition 5:

\textsuperscript{12}This characterization of the extreme points can be verified by straightforward linear algebra: they are the points where $k - 1$ of the defining inequalities bind. The only important thing about the extreme points for the proof is that each assigns mass at least $\varepsilon$ to some $\hat{a}_2 \neq \tilde{a}_2$. 

23
Claim 1, existence: Suppose $S \times T$ admits no safety. We show that for sufficiently large $r$, the second condition in Lemma 1 holds in $G^{-r}$, whence the lemma yields an equilibrium. Fix any $a_1 \in A_1$; it is not a never-worst deviation, so for some $a_2 \in T$, $u(a_1, a_2) < u(a'_1, a_2)$ for all $a'_1 \in S - \{a_1\}$. Fix $\varepsilon \in (0, 1/|A_1|)$. Suppose $\sigma_1 >_T \sigma_1'$ with $\sigma_1 \in \Delta(A_1), \sigma_1' \in \Delta(S)$ having disjoint support, as in the lemma. Suppose $\sigma_1(a_1) \geq \varepsilon$. Disjoint support implies $\sigma_1(S - \{a_1\}) = 1$. Similarly to the previous proof, we can deduce from unbounded concavity that

$$u^*_i(\sigma_1, a_2) < u^*_i(\sigma_1', a_2)$$

for large enough $r$, contradicting weak dominance. So $\sigma_1(a_1) < \varepsilon$. Letting $r$ be large enough that $\sigma_1(a_1) < \varepsilon < 1/|A_1|$ for all $a_1 \in A_1$ gives a contradiction, so for large $r$ no pair $\sigma_1, \sigma_1'$ with $\sigma_1 >_T \sigma_1'$ exists. Also take $r$ large enough for the analogous conditions to hold with roles reversed, and we have proven the existence for such $r$ of an equilibrium with support $S \times T$.

Convergence of equilibria of $G^{-r}$: For any game $G$, define $-G$ as the game with identical action sets and all utilities $u_i$ replaced by $-u_i$. Also, for any $S \subseteq A_1, T \subseteq A_2$, let $G_{S \times T}$ be the block game with actions restricted to $S \times T$. Notice that if $G^{-r}$ is an unboundedly concave family of transformations of $G$, $-[G^{-r}]$ is an unboundedly convex family of transformations of $-G$. Any equilibrium of $G^{-r}$ with support $S \times T$ makes players indifferent among all actions in the block, so it is also a completely mixed equilibrium of $-[G^{-r}]_{S \times T}$. Applying Proposition 4 to the completely mixed equilibria of $-[G^{-r}]_{S \times T}$ tells us that they converge to the pure strategy profile described there. It remains only to observe that

$$\arg\min_{a_1 \in S} \left[ \max_{a_2 \in T} -u_2(a_1, a_2) \right] = \arg\max_{a_1 \in S} \left[ \min_{a_2 \in T} u_2(a_1, a_2) \right]$$

$$\arg\min_{a_2 \in T} \left[ \max_{a_1 \in S} -u_1(a_1, a_2) \right] = \arg\max_{a_2 \in T} \left[ \min_{a_1 \in S} u_1(a_1, a_2) \right]$$

24
as desired.

**Claim 2, non-existence:** Suppose \( a_1 \in A_1 \) a never-worst deviation from \( S \times T \). Let \( \sigma'_1 \) be an equal mixture of all elements of \( S - \{a_1\} \). We claim \( a_1 >_T \sigma'_1 \) in \( G^{-r} \) for all sufficiently large \( r \). Indeed, for all \( a_2 \in T \),

\[
u^{-r}(a_1, a_2) > \nu^{-r}(\sigma'_1, a_2)
\]

for sufficiently large \( r \), because the right-hand lottery has a worst outcome which is less than the left-hand-side, by definition of never-worst. The claim follows by finiteness of \( T \), and the desired result by the lemma.

\[\square\]

The following is a three-player counterexample to Proposition 4: a generic three-player game where every action is a point best-reply, but there is no completely mixed equilibrium in any monotone convex transformation:

<table>
<thead>
<tr>
<th>( In )</th>
<th>( L )</th>
<th>( R )</th>
<th>( Out )</th>
<th>( L )</th>
<th>( R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( U )</td>
<td>4, 1, 3</td>
<td>2, 3, 1</td>
<td>( U )</td>
<td>8, 5, 7</td>
<td>6, 7, 8</td>
</tr>
<tr>
<td>( D )</td>
<td>1, 4, 5</td>
<td>3, 2, 2</td>
<td>( D )</td>
<td>5, 8, 4</td>
<td>7, 6, 6</td>
</tr>
</tbody>
</table>

In fact, any such transformation has a unique NE where Player 3 chooses Out for sure. To see this, first check that Players 1 and 2 must mix in any equilibrium (for instance, if P1 plays U, P2 must play R, whence P1 plays D, etc.) Next, P1 must play U at least half the time; if not, P2 strictly prefers L in both matrices, because for instance \( pf(1) + (1-p)f(4) > pf(3) + (1-p)f(2) \) for any \( p < .5 \) and convex \( f \). Similarly, P2 must play R at least half the time. It is now easy to check that P3 strictly prefers Out.

For a corresponding example for Proposition 5, simply negate all payoffs; now every action is a point worst-reply, so the full block allows no safety, but there can be no fully mixed equilibrium in any concave transformation, since negating would give a convex transformation of the diagrammed game.

Also, to show that Proposition 6 fails when \( n > 2 \), here is a generic three-player game which, under the transformations \( u \mapsto e^{-ru} \), has NE which do not
approach a pure profile:

\[
\begin{array}{c|cc|c|cc}
 & L & R & & L & R \\
\hline
U & 4, 6, 6 & 1, 5, 2 & U & 8, 2, 5 & 5, 1, 1 \\
D & 3, 7, 7 & 2, 8, 3 & D & 7, 3, 8 & 6, 4, 4 \\
\end{array}
\]

Here, in \( G^r \) any profile where \( p(U) = e^{2r}/(e^{2r} + 1), p(L) = 1/(e^{2r} + 1) \) is an equilibrium, for any strategy of P3. One reaction to this example is that a stronger notion of genericity might restore convergence to pure profiles; for instance, we are capitalizing on the fact that P1 and P2’s best replies are not affected by P3’s action, in the original game or any exponential transformations. Notice, though, that many perturbations of the payoffs in the matrix could be undone by a different family of convex transformations, so any successful strengthening of genericity would have to include a restriction on the transformations. Note that neither of these examples is a counterexample to the conjecture on CE stated at the end of Section 5, which to my knowledge may be true for all \( n \).

8 Appendix B

As per Dekel, Fudenberg and Morris (2007), we define the ICR sets for a Bayesian game \((N, A, \hat{u}, \Theta, T, \kappa)\) as follows. For each \( i \) and \( t_i \), set \( ICR_i^0(t_i) = A_i \). Then define \( ICR_i^n(t_i) \) iteratively by letting the set of admissible “conjectures” be

\[
C_n(t_i) \equiv \{ \nu \in \Delta(\Theta \times T_{-i} \times A_{-i}) : \text{marg}_{\Theta \times T_{-i}} \nu = \kappa_{t_i}, \nu(\{ (\theta, t_{-i}, a_{-i}) : a_{-i} \in ICR_{-i}^{n-1}(t_{-i}) \}) = 1 \}
\]

and then letting

\[
ICR_i^n(t_i) = \bigcup_{\nu \in C_n(t_i)} BR_i(\text{marg}_{\Theta \times A_{-i}} \nu)
\]
Proof of Proposition 8: First, we show $P_i^\infty(G) \subseteq ICR\hat{G}(t_i)$. We suppress the $\hat{G}$ in future notation. It suffices to show that if for some $k$, $P_i^k(G) \subseteq ICR_i^k(t_i)$, the same is true for $k + 1$. Let $a_i \in P_i^{k+1}(G)$. Then $a_i \in BR(a_{-i})$ for some pure profile $a_{-i} \in P_{-i}^k(G)$. By inductive hypothesis, $a_{-i} \in ICR_i^k(t_{-i})$ for all $t_{-i}$, implying that $\kappa_i \times \delta_{a_{-i}}$ is an admissible conjecture, where $\delta$ denotes a Dirac measure. Furthermore, because $a_i$ is a best response to $a_{-i}$ for every $\theta$, the sure-thing principle tells us $a_i \in BR_i(\text{marg}_{\Theta}(\kappa_i \times \delta_{a_{-i}}))$ as desired.

Next, we show $ICR_i^\infty(G) \subseteq B_i^\infty(G)$. Following a similar idea, we show inductively that $ICR_i^k(G) \subseteq B_i^k(G)$ for every $i, t_i$ and $k$. Suppose, towards a contradiction, that the first violation of this claim is $a_i \in B_i^k(G) - B_i^{k+1}(G)$ and $a_i \in ICR_i^{k+1}(t_i)$. Let $\nu \in C_{k+1}(t_i)$ be the conjecture which rationalizes $a_i$. By inductive hypothesis, $\text{marg}_{A_{-i}}(B_i^k) = 1$. By definition of the B-process, there is then an action $a'_i$ such that $u(a'_i, a_{-i}) \geq u(a_i, a_{-i})$ for all $a_{-i} \in \text{supp}\text{marg}\nu$ with strict inequality for some such $a_{-i}$. Then also

$$\text{marg}_{\Theta \times A_{-i}}(\nu(\theta, a_{-i})\tilde{u}(a'_i, a_{-i}, \theta)) \geq \text{marg}_{\Theta \times A_{-i}}(\nu(\theta, a_{-i})\tilde{u}(a'_i, a_{-i}, \theta))$$

for every pair $(\theta, a_{-i})$, and again inequality must be strict somewhere, but this contradicts $a_i \in BR_i(\text{marg}_{\Theta \times A_{-i}}(\nu))$.

References


