Justifying Optimal Play via Consistency

Florian Brandl   Felix Brandt
Technische Universität München
{brandlfl,brandtf}@in.tum.de

Developing normative foundations for optimal play in two-player zero-sum games has turned out to be surprisingly difficult, despite the powerful strategic implications of the Minimax Theorem. We characterize maximin strategies by postulating coherent behavior in varying games. The first axiom, called *consequentialism*, states that how probability is distributed among completely indistinguishable actions is irrelevant. The second axiom, *consistency*, demands that strategies that are optimal in two different games should still be optimal when there is uncertainty which of the two games will actually be played. Finally, we impose a very mild *rationality* assumption, which merely requires that strictly dominated actions will not be played. Our characterization shows that a rational and consistent consequentialist who ascribes the same properties to his opponent has to play maximin strategies. This result can be extended to characterize Nash equilibrium in bimatrix games whenever the set of equilibria is interchangeable.

1 Introduction

Two-player zero-sum games, i.e., games in which the interests of both players are diametrically opposed, are the historical point of departure for the theory of games. Perhaps the most central result in game theory, the Minimax Theorem (von Neumann, 1928), has established that zero-sum games admit natural "optimal" strategies, which furthermore correspond to unique payoffs for both players. 25 years after his seminal discovery, John von Neumann wrote that "as far as I can see, there could be no theory of games [. . .] without that theorem [. . .] I thought there was nothing worth publishing until the 'minimax theorem' was proved." (von Neumann and Fréchet, 1953).

When Nash (1950) and others initiated the study of more general, non-cooperative games, it became apparent that much of the elegance of the theory of zero-sum games is lost and no unequivocal notion of optimality pertains. Unlike maximin strategies in zero-sum games, Nash equilibria, their various refinements, as well as their coarsenings such as correlated equilibria require some form of coordination among the players, i.e, the optimality of a strategy depends on the strategies chosen by the other players. In zero-sum games, the set of Nash equilibria is interchangeable, i.e., it is the Cartesian set of pairs of maximin strategies. Hence, the optimality of a player's strategy is independent of the strategy chosen by the other player. Moreover, maximin strategies can be efficiently computed using linear programming while finding Nash equilibria in non-zero-sum games has been shown to be computationally intractable, even when there are only two players (Daskalakis et al., 2009; Chen et al., 2009).

Based on these considerations, a wide-spread sentiment in game theory states that rational players *ought* to play maximin strategies in zero-sum games. However, this conclusion is premature, even under relatively strong interpretations of rationality.[1] The approach we pursue in this paper is based on axioms that require players to behave coherently across varying hypothetical games. Our first axiom, called *consequentialism*, demands that players only care about the payoffs obtained by playing an action, not about its name. More precisely, when there are actions that yield identical payoffs against *every* action of the opponent, so-called clones, the probability assigned to other actions should be independent of the number of clones and the remaining probability can be distributed arbitrarily among the clones. Furthermore, a player's strategy should not depend on clones available to the other player. The second axiom prescribes how players deal with games that they consider strategically equivalent. Consider two games in which both players would choose the same *pair* of strategies, one for the row player and one for the column player. *Consistency* requires that the same pair of strategies should be played in a new game where a coin toss decides which of the two original games is actually being played (the payoffs of the new game are given by the randomization between the payoffs of the two original games). This is in fact the only place in our model where expected payoffs enter the picture. The only rationality assumption we make—our third axiom—requires that a strictly dominated action will not be played.

Our main theorem then shows that every rational and consistent consequentialist has to play maximin strategies if he ascribes the same properties to his opponent (Theorem 2). This result can be extended to characterize Nash equilibrium in bimatrix games whenever the set of equilibria is interchangeable.

[1] Early objections against the normative power of maximin strategies were raised by Bacharach (1987), who wrote that "[the maximin principle] has been much criticized, and the qualified acceptance it has enjoyed owes something to its protective alliance with other elements of von Neumann and Morgenstern's theory. Their own arguments for it were suggestive rather than apodictic. It is claimed to express a rational caution in a situation in which a player has no valid basis for assigning probabilities to his opponent's decision. A second argument is also advanced, unworthy of them and justly attacked by Ellsberg (1956), according to which it is rational for A to choose by supposing he is playing the 'minorant game' associated with the payoff matrix: in this game A chooses first and B second in knowledge of A's choice (so that A is a Stackelberg 'leader'). In *this* situation rock-hard principles of decision under certainty make it rational for A to maximin. But convincing reasons for A to assume that it obtains are missing."

2 Related Work

The question which concrete assumptions lead to equilibrium play has been primarily studied in epistemic game theory, which amends the traditional game-theoretic models by formally modeling the knowledge of individual players. This is achieved using Bayesian belief hierarchies, which consist of a game and a set of types for each player with each type including the action played by this type and a probability distribution over types of the other players, called the belief of this type. In this model, it is not assumed that players actively randomize. Instead, the beliefs about the types of the other players are randomized. Players are rational if they maximize expected payoff given their types and beliefs. Aumann and Brandenburger (1995) have shown that for two-player (not necessarily zero-sum) games the beliefs of every pair of types whose beliefs are mutually known and whose rationality is mutually known constitute a Nash equilibrium.[2] This result extends to games with more than two players if the beliefs are commonly known and admit a common prior. A different perspective was taken by Aumann and Drèze (2008) who, among others things, investigate which payoffs rational players should expect in zero-sum games. Aumann and Drèze showed that if rationality is common knowledge and the beliefs admit a common prior, then the players should expect the value of the game. In their proof, Aumann and Drèze (2008) consider games in which all actions are "doubled", roughly reminiscent of our consequentialism axiom.

Another stream of research has delivered characterizations of the *value* of a zero-sum game using axioms that are not necessarily motivated by decision-theoretic considerations (Vilkas, 1963; Tijs, 1981; Norde and Voorneveld, 2004). Since the row player's payoff will always be the value when both players play maximin strategies, our result can also be interpreted as a characterization of the value of a zero-sum game. Note, however, that characterizations of the value are weaker than characterizations of maximin strategies because the value is devoid of any strategic content. In symmetric zero-sum games, for example, the value is constantly zero but finding or characterizing maximin strategies is non-trivial.

Hart et al. (1994) provide sufficient conditions for agents to evaluate zero-sum games by their value and thereby attempt to motivate playing maximin strategies. In their framework, the players have preferences over game forms satisfying a number of axioms, whose implications are two-fold: first, they guarantee that lotteries over outcomes are evaluated based on their expected utility for some underlying vNM utility function; second, game forms are ranked based on the value of the zero-sum game resulting from them given this utility function. A weakening of the axiom that drives the second implication is called *irrelevance of duplications* and states that the agent is indifferent between game forms that only differ with regards to duplicate rows or columns. It is thus reminiscent of consequen-

[2] In the model of Aumann and Brandenburger (1995) the payoff functions are unknown and the players have beliefs about the payoff functions that may depend on their type. For this model their result additionally requires the payoff functions to be mutually known.

tialism. However, since evaluation of games by their value does not necessitate maximin play, Hart et al. conclude that "our goal has not been completely achieved, in that we have rationalized evaluation of zero-sum games by their value, but we have not proved that 'rationality' implies playing maxmin strategies."

Our characterization bears some resemblance to a recent characterization of a randomized voting rule that is based on the symmetric zero-sum game given by the pairwise majority margins of the voters' preferences (Brandl et al., 2016). Even though the proofs are quite different, the common key idea is to consider convex combinations of matrices that are permutations of each other. The characterization described in Remark 5 by Brandl et al. uses *population-consistency*, *cloning-consistency*, and *Condorcet-consistency*. Population-consistency corresponds to our consistency, cloning-consistency is slightly weaker than consequentialism, and Condorcet-consistency is stronger than rationality. On top of that, Brandl et al. also need upper hemi-continuity, convexity, and decisiveness of solution concepts, which are not required in our proof (essentially because it directly operates on game matrices rather than preference profiles). The zero-sum games considered arising in the proof by Brandl et al. are always symmetric because they are induced by majority margins.

3 The Model

Let U be an infinite universal set of actions and denote by $\mathcal{F}(U)$ the set of all finite and non-empty subsets of U. For $A, B \in \mathcal{F}(U)$, $M \in \mathbb{Q}^{A\times B}$ is a (two-player zero-sum) *game* with action sets A and B for the row and the column player, respectively. A (mixed) *strategy* for a player with action set $A \in \mathcal{F}(U)$ is a probability distribution over A and thus an element of $\Delta(A) = \{p \in \mathbb{Q}^A_{\geq 0} \colon \sum_{a\in A} p_a = 1\}$.[3] With slight abuse of notation, we will sometimes identify actions with the degenerate strategy that puts all probability on the respective action.

A *solution concept* f is a function that maps a game $M \in \mathbb{Q}^{A\times B}$ with actions sets $A, B \in \mathcal{F}(U)$ to a set of optimal strategies $f(M) \subseteq \Delta(A)$ for the row player. Note that $f(-M^t)$ is the set of optimal strategies for the column player in the game M. A widely accepted solution concept for zero-sum games is the function that returns the set of maximin strategies, i.e., strategies that maximize the minimum expected payoff. Formally, for all $A, B \in \mathcal{F}(U)$ and $M \in \mathbb{Q}^{A\times B}$,

$$\mathit{maximin}(M) = \arg\max_{p\in\Delta(A)} \min_{q\in\Delta(B)} p^t M q. \qquad \text{(maximin strategies)}$$

A strategy $p \in \Delta(A)$ is a *maximin strategy* for the row player in M if $p \in \mathit{maximin}(M)$. Note that the set of maximin strategies is convex, since it is the set of solutions to a linear

[3] Games with rational-valued payoffs admit rational-valued maximin strategies and consequentially also have a rational value. For real-valued payoffs and probabilities, see Remark 4.

program. The Minimax Theorem (von Neumann, 1928) shows that the minimum expected payoff of a maximin strategy for the row player is equal to the negative of the minimum expected payoff of a maximin strategy for the column player. This payoff is called the *value* of the game.

In the following, assumptions about the players' behavior are modeled as properties of the underlying solution concept.

3.1 Consequentialism

The first property we consider, consequentialism, prescribes that the players evaluate actions solely based on their payoffs and disregard the names of the actions. Hence, if two actions of the row player are indistinguishable in the sense that, against every action of the column player, they yield exactly the same payoff—we call such actions clones—, then optimality of a strategy is independent of how probability is distributed among those two actions. Furthermore, optimality of a strategy is independent of the number of clones of an action for the opponent. Formally, for $\hat{A} \subseteq A \in \mathcal{F}(U)$, $\hat{B} \subseteq B \in \mathcal{F}(U)$, $M \in \mathbb{Q}^{A\times B}$, and $\hat{M} \in \mathbb{Q}^{\hat{A}\times\hat{B}}$, we say that $\hat{M}$ is a *reduced form* of M if there exist surjective functions $\alpha\colon A \to \hat{A}$ and $\beta\colon B \to \hat{B}$ such that, for all $(a,b) \in A \times B$, $M_{ab} = \hat{M}_{\alpha(a)\beta(b)}$. Actions in $\alpha^{-1}(a)$ for the same $a \in \hat{A}$ are called *clones* in M, since they yield the same payoff against every action of the column player. A solution concept f satisfies *consequentialism* if

$$f(M) = \bigcup_{\hat{p}\in f(\hat{M})} \{p \in \Delta(A)\colon \sum_{a\in\alpha^{-1}(\hat{a})} p(a) = \hat{p}(\hat{a}) \text{ for all } \hat{a} \in \hat{A}\}, \qquad \text{(consequentialism)}$$

whenever $\hat{M}$ is a reduced form of M through the function α and some arbitrary function β. So the strategies returned for the game M have to be exactly those which put a total probability of $\hat{p}(\hat{a})$ on the clones of $\hat{a}$, where $\hat{p}$ is some strategy returned for $\hat{M}$. In decision theory, consequentialism corresponds to Chernoff's (1954) *Postulate 6* (cloning of player's actions) and *Postulate 9* (cloning of nature's states, which are equivalent to opponent's actions in our model). The latter is also known as *column duplication* (Milnor, 1954) and *deletion of repetitious states* (Arrow and Hurwicz, 1972; Maskin, 1979). In the context of social choice theory, a related condition called *independence of clones* was introduced by Tideman (1987).

Consequentialism implies that permuting rows in the payoff matrix results in the same permutation within the set of returned strategies and that permuting columns in the payoff matrix has no effect on the recommended strategies for the row player. For a formal definition, let $\Pi(A)$ denote the set of all permutations on A for some $A \in \mathcal{F}(U)$. For $x \in \mathbb{Q}^A$ and $\pi \in \Pi(A)$, x_π is the permutation of entries of x with respect to π, i.e., $x_\pi = x \circ \pi^{-1}$. With this definition, $(x_\pi)_{\pi(i)} = x_i$ for all $i \in A$. We extend this notation to sets, so that, for example, $f(M)_\pi$ is the set of strategies obtained by permuting each strategy in $f(M)$ according to π. Similarly, for $A, B \in \mathcal{F}(U)$, $M \in \mathbb{Q}^{A\times B}$, $\pi \in \Pi(A)$, and

$\sigma \in \Pi(B)$, $M_{\pi\sigma} = M \circ (\pi^{-1} \times \sigma^{-1})$. A solution concept f satisfies *permutation invariance* if, for all $A, B \in \mathcal{F}(U)$, $M \in \mathbb{Q}^{A\times B}$, $\pi \in \Pi(A)$, and $\sigma \in \Pi(B)$,

$$f(M_{\pi\sigma}) = f(M)_{\pi}. \qquad \text{(permutation invariance)}$$

Permutation invariance is a classic condition in decision theory (e.g., *Postulate 3* by Chernoff (1954) or Milnor's (1954) *symmetry*) and social choice theory (where it is known as *neutrality*). It is clear from the definitions that consequentialism implies permutation invariance, since $M_{\pi\sigma}$ is a reduced form of M for $\alpha = \pi$ and $\beta = \sigma$, since $M_{ab} = (M_{\pi\sigma})_{\pi(a)\sigma(b)}$ for all $(a, b) \in A \times B$.

3.2 Consistency

The next property is based on consistent behavior in related games. To this end, we consider two games that are strategically related with respect to some solution concept in the sense that for each player, the sets of recommended strategies overlap. Now consider a situation where a coin toss decides which of $\hat{M}$ and $\bar{M}$ is played. Consistency prescribes that, since there is a strategy that is optimal for the row player in both games *after* the resolution of the coin toss, this strategy should also be optimal *prior* to the coin toss. The expected payoffs of the latter game are given by the convex combination of $\hat{M}$ and $\bar{M}$ with respect to some $\lambda \in [0, 1]$, representing the probability of "heads". So, formally, a solution concept f satisfies *consistency* if, for all $A, B \in \mathcal{F}(U)$, $\lambda \in [0, 1] \cap \mathbb{Q}$, and $\hat{M}, \bar{M} \in \mathbb{Q}^{A\times B}$ such that $f(\hat{M}) \cap f(\bar{M}) \neq \emptyset$ and $f(-\hat{M}^t) \cap f(-\bar{M}^t) \neq \emptyset$,

$$f(\hat{M}) \cap f(\bar{M}) \subseteq f(\lambda\hat{M} + (1-\lambda)\bar{M}). \qquad \text{(consistency)}$$

Note that the antecedent of consistency involves both the row and the column player's strategies and therefore reflects the perspective of an outside observer. It is possible to define a stronger notion of consistency by proclaiming that two games are strategically related whenever f returns the same strategies for the row player but not necessarily for the column player. However, this notion is too demanding and violated by maximin strategies. In fact, it is incompatible with our other axioms (cf. Remark 2). To the best of our knowledge, consistency has not been considered in decision theory or game theory before.[4] Chernoff's (1954) *Postulate 9* is related in that it also requires that the optimality of an action in two different games implies its optimality in a third combined game. However, the combined game is obtained by taking the union of actions rather than convex combinations of payoff matrices. Convex combinations of actions (or states) have been considered for various decision-theoretic axioms (see, e.g., Chernoff, 1954; Milnor, 1954; Gilboa and Schmeidler, 2003). Consistency is related to Shapley's (1953) additivity

[4]Peleg and Tijs (1996) have introduced an unrelated consistency condition which concerns games with a varying number of players and can be used to characterize Nash equilibrium (see, also Norde et al., 1996)

axiom (which he calls *law of aggregation*) in the characterization of the Shapley value. The analogue of consistency for social choice theory is the central property in a number of important axiomatic characterizations (see, e.g., Smith, 1973; Young, 1975; Young and Levenglick, 1978; Brandl et al., 2016).

3.3 Rationality

The third condition, rationality, prescribes that a player does not play actions that are never best responses. It is well-known that an action is never a best response if and only if it is (strictly) dominated. For $A, B \in \mathcal{F}(U)$ and $M \in \mathbb{Q}^{A\times B}$, an action $a \in A$ dominates another action $a' \in A$ in M if a yields higher payoff than a' against every action of the column player, i.e., $M_{ab} > M_{a'b}$ for all $b \in B$. In the sequel, $dom(M)$ denotes the set of dominated rows in M. Clearly, a player who seeks to maximize his payoff should never play a dominated action, since there is another action that yields higher payoff independently of what the opponent does. Hence, a reasonable solution concept should not assign positive probability to a dominated action. We merely require that dominated actions should not be played with probability one.[5] A solution concept f is called *rational* if, for all $A, B \in \mathcal{F}(U)$ and $M \in \mathbb{Q}^{A\times B}$,

$$dom(M) \cap f(M) = \emptyset. \qquad \text{(rationality)}$$

Note that our notion of rationality does not rely on expected payoffs or any assumptions about the other player. It is equivalent to Milnor's (1954) *strong domination*, Maskin's (1979) *Property (5)*, and weaker than Chernoff's (1954) *Postulate 2*.

4 The Result

We first show that the solution concept that returns all maximin strategies satisfies all of the properties defined in the previous section. While we only show that *maximin* satisfies our weak notion of rationality, it is also compatible with the much stronger common knowledge of rationality assumption (see Footnote 7).

Theorem 1. *The solution concept maximin satisfies consequentialism, consistency, and rationality.*

Proof. To show that *maximin* satisfies consequentialism, let $\hat{A} \subseteq A \in \mathcal{F}(U)$, $\hat{B} \subseteq B \in \mathcal{F}(U)$, $M \in \mathbb{Q}^{A\times B}$, and $\hat{M} \in \mathbb{Q}^{\hat{A}\times\hat{B}}$ such that $\hat{M}$ is a reduced form of M through α and β. We have to show that

$$maximin(M) = \bigcup_{\hat{p}\in maximin(\hat{M})} \{p \in \Delta(A) \colon \sum_{a\in\alpha^{-1}(\hat{a})} p(a) = \hat{p}(\hat{a}) \text{ for all } \hat{a} \in \hat{A}\}.$$

[5] In fact, the notion of rationality used in the proof of Theorem 2 is even weaker because we only need rationality in 2×2 games.

First observe that the number of clones of an action for the column player does not influence the set of maximin strategies for the row player. Hence, we may assume without loss of generality that $\hat{B} = B$ and β is the identity function. For $p \in \mathit{maximin}(M)$, let $\hat{p} \in \Delta(\hat{A})$ such that $\hat{p}(\hat{a}) = \sum_{a \in \alpha^{-1}(\hat{a})} p(a)$ for all $\hat{a} \in \hat{A}$. Since, for every $\hat{a} \in \hat{A}$, all actions in $\alpha^{-1}(\hat{a})$ are clones, it follows that $\hat{p}^t \hat{M} = p^t M$, which shows that $\hat{p}$ is a maximin strategy and proves the inclusion from left to right.

For the other inclusion, let $p \in \bigcup_{\hat{p} \in \mathit{maximin}(\hat{M})} \{p \in \Delta(A) \colon \sum_{a \in \alpha^{-1}(\hat{a})} p(a) = \hat{p}(\hat{a}) \text{ for all } \hat{a} \in \hat{A}\}$. Then, there is $\hat{p} \in \mathit{maximin}(\hat{M})$ such that, for all $\hat{a} \in \hat{A}$, $\sum_{a \in \alpha^{-1}(\hat{a})} p(a) = \hat{p}(\hat{a})$. As before, we have that $p^t M = \hat{p}^t \hat{M}$, which shows that $p \in \mathit{maximin}(M)$.

To show that *maximin* satisfies consistency, let $A, B \in \mathcal{F}(U)$ and $\hat{M}, \bar{M} \in \mathbb{Q}^{A \times B}$ such that

$$p \in \mathit{maximin}(\hat{M}) \cap \mathit{maximin}(\bar{M}), \text{ and}$$
$$q \in \mathit{maximin}(-\hat{M}^t) \cap \mathit{maximin}(-\bar{M}^t).$$

Let $\lambda \in [0,1] \cap \mathbb{Q}$ and $M = \lambda \hat{M} + (1-\lambda)\bar{M}$. For the values $\hat{v}, \bar{v} \in \mathbb{Q}$ of $\hat{M}$ and $\bar{M}$, it follows that $p^t \hat{M} q \geq \hat{v}$ and $p^t \bar{M} q \geq \bar{v}$. Hence, $p^t M q \geq \lambda \hat{v} + (1-\lambda)\bar{v}$. For q, we have that $p^t \hat{M} q \leq \hat{v}$, and $p^t \bar{M} q \leq \bar{v}$. Hence, $p^t M q \leq \lambda \hat{v} + (1-\lambda)\bar{v}$. This implies that $\lambda \hat{v} + (1-\lambda)\bar{v}$ is the value of M. Thus, $p \in \mathit{maximin}(M)$.

To see that *maximin* satisfies rationality, let $A, B \in \mathcal{F}(U)$, $M \in \mathbb{Q}^{A \times B}$, and $q \in \Delta(B)$. Assume for contradiction that $a \in \mathit{maximin}(M)$ for some $a \in \mathit{dom}(M)$ and let a be dominated by some action $\hat{a}$. Clearly, $\hat{a}^t M q > a^t M q$, which contradicts the assumption that $a \in \mathit{maximin}(M)$.

□

4.1 Characterization of Maximin Strategies

Our main theorem shows that every solution concept that satisfies consequentialism, consistency, and rationality has to return maximin strategies. Together with Theorem 1, this implies that *maximin* is the coarsest solution concept satisfying these properties.

The key insight to prove this statement is that, while consequentialism seems very innocuous on its own, it becomes quite powerful when used in conjunction with consistency. In particular, taking the convex combination of game matrices that are permutations of each other can be used to generate games with cloned actions, which eventually rules out solution concepts other than *maximin* (see Remark 1 for examples).

The high-level structure of the proof is as follows. If one of the players does not play a maximin strategy, their strategies p and q do not constitute a Nash equilibrium and one of the players, say the row player, does not best-respond to his opponent's strategy. By linearity of the payoff function, the row player has a pure best response, say $\hat{a}$, against q, which, by consequentialism, may be assumed to be *outside* the support of p. We will construct a game in which the row player only has two actions and does not best-respond. To this end, we introduce a number of clones for each action in the support of p proportional to the probability assigned to it. Since then, by consequentialism, the row player plays the uniform strategy over all those clones (among other strategies), permutation invariance implies that he also plays the uniform strategy in all games that result from permuting these actions, while the column player invariably plays q. Taking the uniform convex combination of all these permuted games yields a game in which all clones of actions in the support of p are clones of each other. Moreover, consistency implies that the row player plays the uniform strategy over these clones. An application of consequentialism then yields a game where the row player plays a pure strategy, say a_1, which is not a best response.

In the remainder of the proof, we modify this game further to obtain a 2×2 game, in which the row player plays a dominated action. We start by cloning $\hat{a}$ (which is still a best response) a number of times. By consequentialism, this does not change either players' strategy. Then we apply a similar procedure to the actions outside the support of the row player's strategy as we did before to those inside the support, i.e., we take the uniform convex combination of all games resulting from permuting these actions, which yields a game in which they are all clones of each other and consequentialism allows us to contract them into one action, say a_2. In the resulting game, the row player plays the pure strategy a_1, the column player plays q, and, by cloning $\hat{a}$ sufficiently many times, a_2 is a best response while a_1 is not. By a number of similar steps, we can construct a game such that the column player also only has two actions, say b_1 and b_2 (where b_1 corresponds to actions in the support of q and b_2 corresponds to actions outside the support of q), and plays b_1 with probability one.

In this game, a_2 yields higher payoff against b_1 than a_1, since a_2 is a best response against b_1 while a_1 is not. However, a_2 may not dominate a_1 because of their payoffs against b_2. To address this, we introduce a number of clones of b_1 and, by consequentialism, may assume that the strategies remain unchanged. Then when taking the uniform convex combination of all games resulting from permuting b_2 and all clones of b_1 (except for b_1 itself), we obtain a game in which, by consistency, the row player plays a_1 but a_2 dominates a_1 (provided we cloned b_1 sufficiently many times). This contradicts rationality and concludes the proof.

Theorem 2. *If a solution concept f satisfies consequentialism, consistency, and rationality, then $f \subseteq$ maximin.*

Proof. For $A, A' \in \mathcal{F}(U)$ with $A' \subseteq A$, let $\text{uni}_A(A')$ denote the strategy in $\Delta(A)$ that randomizes uniformly over A'.

Now, assume for contradiction that $f \not\subseteq$ *maximin*, i.e., there are $A, B \in \mathcal{F}(U)$ and $M \in \mathbb{Q}^{A \times B}$ such that $f(M) \not\subseteq$ *maximin*(M). Let $v \in \mathbb{Q}$ be the value of M, $p \in f(M) \setminus$ *maximin*(M), and $q \in f(-M^t)$. If $p^t M q < v$, there is $a \in A$ such that $a^t M q > p^t M q$. If $p^t M q \geq v$, there is $b \in B$ such that $p^t M b < p^t M q$. In any case, (p, q) is not a Nash equilibrium of M. By symmetry of the roles of the row player and the column player, we may assume without loss of generality that the former case obtains, i.e., that the row player does not play a best response to the column player's strategy. Let $\hat{a} \in A$ be a best response to q, so that for all $a \in A$, $\hat{a}^t Mq \geq a^t Mq$, with a strict inequality for at least one $a \in \text{supp}(p)$.

First we show that we may assume without loss of generality that $\hat{a} \notin \text{supp}(p)$ and $\text{supp}(q) \subsetneq B$. If $\hat{a} \in \text{supp}(p)$, let $\hat{a}' \in U \setminus A$ and $M' \in \mathbb{Q}^{(A\cup\{\hat{a}'\})\times B}$ such that M is a reduced form of M' with $\alpha(\hat{a}) = \alpha(\hat{a}') = \hat{a}$, $\alpha(a) = a$ for all $a \in A \setminus \{\hat{a}\}$, and β equal to the identity function. Consequentialism implies that $\{p' \in \Delta(A) \colon \sum_{a\in\alpha^{-1}(\hat{a})} p'(a) = p(\hat{a}) \text{ for all } \hat{a} \in \hat{A}\} \subseteq f(M')$. In particular, $p'' \in f(M')$, where $p''(a) = p(a)$ for all $a \in A$ (and hence, $p''(\hat{a}') = 0$). Note that $p'' \notin \mathit{maximin}(M')$, since *maximin* satisfies consequentialism by Theorem 1. Then, $\hat{a}' \notin \text{supp}(p'')$ and, for all $a \in A \cup \{\hat{a}'\}$, $(\hat{a}')^t M'q \geq a^t M'q$, with a strict inequality for at least one $a \in \text{supp}(p'')$. The argument for showing that we may assume $\text{supp}(q) \subsetneq B$ is analogous and is therefore omitted. So from now on, we assume that $\hat{a} \notin \text{supp}(p)$ and $\text{supp}(q) \subsetneq B$.

The next step is to construct a game $\bar{M}$, where the row player only has two different types of actions and puts all probability on actions that are not best responses to the strategy of the column player. Let $\delta \in \mathbb{Q}$ be the greatest common divisor of $\{p_a \colon a \in A\}$, which exists, since f is assumed to map to $\Delta(A) \subseteq \mathbb{Q}^A$.[6] For all $a \in A \setminus \{\hat{a}\}$, let $m_a = \max\{1, {}^{p_a}/_{\delta}\}$ and $A_a \in \mathcal{F}(U)$ such that $|A_a| = m_a$, $A_a \cap A = \{a\}$, and all A_a are pairwise disjoint. Let $\epsilon_1 = \hat{a}^t Mq - p^t Mq$ and $\epsilon_2 = p^t Mq - v_{\min}$, where $v_{\min} = \min_{a\in A} a^t Mq$. Observe that $\epsilon_1 > 0$ by the choice of $\hat{a}$. Now let $k = |A \setminus (\text{supp}(p) \cup \{\hat{a}\})|$ and $A_{\hat{a}} \in \mathcal{F}(U)$ such that $|A_{\hat{a}}| > k{}^{\epsilon_2}/_{\epsilon_1}$, $A_{\hat{a}} \cap A = \{\hat{a}\}$, and $A_{\hat{a}} \cap A_a = \emptyset$ for all $a \in A \setminus \{\hat{a}\}$. Let $\hat{A} = \bigcup_{a\in A} A_a$ and $\hat{M} \in \mathbb{Q}^{\hat{A}\times B}$ such that M is a reduced form of $\hat{M}$ with α such that $\alpha^{-1}(a) = A_a$ for all $a \in A$ and β equal to the identity function. By application of consequentialism to $\hat{M}$ and M with α and β as above, it follows that $\text{uni}_{\hat{A}}(\bar{A}) \in f(\hat{M})$ and $q \in f(-\hat{M}^t)$, where $\bar{A} = \bigcup_{a\in\text{supp}(p)} A_a$. Let $\bar{\Pi} \subseteq \Pi(\hat{A})$ be the set of permutations $\pi \in \Pi(\hat{A})$ such that $\pi(\bar{A}) = \bar{A}$ (and thus, $\pi(\hat{A}\setminus\bar{A}) = \hat{A}\setminus\bar{A}$). Note that $\hat{A}\setminus\bar{A}$ is non-empty, since $\hat{a} \notin \text{supp}(p)$ by assumption. Since f satisfies permutation invariance, it follows that $\text{uni}_{\hat{A}}(\bar{A}) \in f(\hat{M}_{\pi,\text{id}})$ and $q \in f(-(\hat{M}_{\pi,\text{id}})^t)$ for all $\pi \in \bar{\Pi}$. Let $\bar{M} = {}^{1}/_{|\bar{\Pi}|} \sum_{\pi\in\bar{\Pi}} \hat{M}_{\pi,\text{id}}$. Consistency implies that $\text{uni}_{\hat{A}}(\bar{A}) \in f(\bar{M})$ and $q \in f(-\bar{M}^t)$. Observe that $\bar{A}$ and $\hat{A}\setminus\bar{A}$ are sets of clones in $\bar{M}$. The following calculation shows that the number of clones of $\hat{a}$ in $\hat{M}$ was chosen large enough such that when the column player plays q in $\bar{M}$, the expected payoff of actions in $\hat{A}\setminus\bar{A}$ is larger than that of actions in $\bar{A}$. By construction of $\bar{M}$, we have that, for all $a \in \bar{A}$ and

[6] The greatest common divisor of a set of rational numbers $\{x_1, \ldots, x_k\}$ is defined as $\max\{\delta \in \mathbb{Q} \colon {}^{x_i}/_{\delta} \in \mathbb{N} \text{ for all } i \in \{1, \ldots, k\}\}$.

$a' \in \hat{A} \setminus \bar{A}$,

$$\begin{aligned}(a')^t \bar{M} q - a^t \bar{M} q &\geq \frac{1}{|A_{\hat{a}}| + k} \left(|A_{\hat{a}}|(\hat{a}^t M q - p^t M q) - k(p^t M q - v_{\min})\right) \\ &> \frac{1}{|A_{\hat{a}}| + k} \left(\frac{k(p^t M q - v_{\min})}{\hat{a}^t M q - p^t M q}(\hat{a}^t M q - p^t M q) - k(p^t M q - v_{\min})\right) \\ &= 0,\end{aligned} \tag{1}$$

where the first inequality follows from the definition of $\bar{M}$, the second inequality follows from the definition of $A_{\hat{a}}$, and the last equality follows from basic algebra.

Now we apply a similar construction to $\bar{M}$ to construct at game $\tilde{M}$ in which the column player also only has two different types of actions. Let $\tau \in \mathbb{Q}$ be the greatest common divisor of $\{q_b \colon b \in B\}$. For all $b \in B$, let $m_b = \max\{1, q_b/\tau\}$ and $B_b \in \mathcal{F}(U)$ such that $|B_b| = m_b$, $B_b \cap B = \{b\}$, and all B_b are pairwise disjoint. Let $\hat{B} = \bigcup_{b \in B} B_b$ and $\mathring{M} \in \mathbb{Q}^{\hat{A} \times \hat{B}}$ such that $\bar{M}$ is a reduced form of $\mathring{M}$ for α equal to the identity function and β such that $\beta^{-1}(b) = B_b$ for all $b \in B$. By two applications of consequentialism (once to $\mathring{M}$ and $\bar{M}$ and once to $-\mathring{M}^t$ and $-\bar{M}^t$), it follows that $\mathrm{uni}_{\hat{A}}(\bar{A}) \in f(\mathring{M})$ and $\mathrm{uni}_{\hat{B}}(\bar{B}) \in f(-\mathring{M}^t)$, where $\bar{B} = \bigcup_{b \in \mathrm{supp}(q)} B_b$. Let $\bar{\Sigma} \subseteq \Pi(\hat{B})$ be the set of permutations $\sigma \in \Pi(\hat{B})$ such that $\sigma(\bar{B}) = \bar{B}$. Since f satisfies permutation invariance, it follows that $\mathrm{uni}_{\hat{A}}(\bar{A}) \in f(\mathring{M}_{\mathrm{id},\sigma})$ and $\mathrm{uni}_{\hat{B}}(\bar{B}) \in f(-(\mathring{M}_{\mathrm{id},\sigma})^t)$ for all $\sigma \in \bar{\Sigma}$. Let $\tilde{M} = 1/|\bar{\Sigma}| \sum_{\sigma \in \bar{\Sigma}} \mathring{M}_{\mathrm{id},\sigma}$. Consistency implies that $\mathrm{uni}_{\hat{A}}(\bar{A}) \in f(\tilde{M})$ and $\mathrm{uni}_{\hat{B}}(\bar{B}) \in f(-\tilde{M}^t)$. Observe that $\bar{A}$ and $\hat{A} \setminus \bar{A}$ are sets of clones in $\tilde{M}$ and $\bar{B}$ and $\hat{B} \setminus \bar{B}$ are sets of clones in $-\tilde{M}^t$ and that $\hat{B} \setminus \bar{B} \neq \emptyset$ since $\mathrm{supp}(q) \subsetneq B$ by assumption. So both players have exactly two different types of actions in $\tilde{M}$. We will use this fact to construct a game M^2 in which both players only have two actions, one of which they play with probability one, and the row player does not play a best response to the strategy of the column player. To this end, let $a_1 \in \bar{A}$, $a_2 \in \hat{A} \setminus \bar{A}$, $b_1 \in \bar{B}$, $b_2 \in \hat{B} \setminus \bar{B}$, and $M^2 \in \mathbb{Q}^{\{a_1,a_2\} \times \{b_1,b_2\}}$ such that M^2 is a reduced form of $\tilde{M}$ for α and β such that $\alpha^{-1}(a_1) = \bar{A}$, $\alpha^{-1}(a_2) = \hat{A} \setminus \bar{A}$, $\beta^{-1}(b_1) = \bar{B}$, and $\beta^{-1}(b_2) = \hat{B} \setminus \bar{B}$. Consequentialism implies that $a_1 \in f(M^2)$ and $b_1 \in f(-(M^2)^t)$. Moreover, $a_2^t M^2 b_1 > a_1^t M^2 b_1$ by (1).

The last step is to use M^2 to construct a game $\tilde{M}^2$ in which both players have two actions and the row player plays a dominated action with probability one. To this end, let $\kappa_1 = a_2^t M^2 b_1 - a_1^t M^2 b_1$ and $\kappa_2 = a_1^t M b_2 - a_2^t M b_2$. Observe that $\kappa_1 > 0$ by (1). Let $B_{b_1} \in \mathcal{F}(U)$ such that $|B_{b_1}| > \kappa_2/\kappa_1 + 1$ and $B_{b_1} \cap \{b_1, b_2\} = \{b_1\}$. Moreover, let $\bar{B}^2 = B_{b_1} \cup \{b_2\}$ and $\hat{M}^2 \in \mathbb{Q}^{\{a_1,a_2\} \times \bar{B}^2}$ such that M^2 is a reduced form of $\hat{M}^2$ for α equal to the identity function and β such that $\beta^{-1}(b_1) = B_{b_1}$. Consequentialism implies that $a_1 \in f(\hat{M}^2)$ and $b_1 \in f(-(\hat{M}^2)^t)$. Now let $\bar{\Sigma}^2 = \{\sigma \in \Pi(\bar{B}^2) \colon \sigma(b_1) = b_1\}$. It follows from permutation invariance that, for all $\sigma \in \bar{\Sigma}^2$, $a_1 \in f(\hat{M}^2_{\mathrm{id},\sigma})$ and $b_1 \in f(-(\hat{M}^2_{\mathrm{id},\sigma})^t)$. Let $\bar{M}^2 = 1/|\bar{\Sigma}^2| \sum_{\sigma \in \bar{\Sigma}^2} \hat{M}^2_{\mathrm{id},\sigma}$. Consistency implies that $a_1 \in f(\bar{M}^2)$ and $b_1 \in f(-(\bar{M}^2)^t)$.

Observe that all actions in $\bar{B}^2 \setminus \{b_1\}$ are clones in $-(\bar{M}^2)^t$ and that, for all $b \in \bar{B}^2 \setminus \{b_1\}$,

$$\begin{aligned}
a_2^t \bar{M}^2 b - a_1^t \bar{M}^2 b &= \frac{1}{|B_{b_1}|} \left((|B_{b_1}| - 1)(a_2^t M^2 b_1 - a_1^t M^2 b_1) + a_2^t M^2 b_2 - a_1^t M^2 b_2 \right) \\
&> \frac{1}{|B_{b_1}|} \left(\frac{a_1^t M^2 b_2 - a_2^t M^2 b_2}{a_2^t M^2 b_1 - a_1^t M^2 b_1} (a_2^t M^2 b_1 - a_1^t M^2 b_1) + a_2^t M^2 b_2 - a_1^t M^2 b_2 \right) \\
&= 0,
\end{aligned}$$

where the equality follows from the definition of $\bar{M}^2$, the inequality follows from the cardinality lower bound on B_{b_1}, and the last equality follows from basic algebra. Recall that, by definition of $\bar{M}^2$, $a_2^t \bar{M}^2 b_1 > a_1^t \bar{M}^2 b_1$. Hence, a_2 dominates a_1 in $\bar{M}^2$. Lastly, let $\tilde{M}^2 \in \mathbb{Q}^{\{a_1,a_2\} \times \{b_1,b_2\}}$ such that $\tilde{M}^2$ is a reduced form of $\hat{M}^2$ for α equal to the identity function and β such that $\beta^{-1}(b_2) = \bar{B}^2 \setminus \{b_1\}$. Consequentialism implies that $a_1 \in f(\tilde{M}^2)$ and $b_1 \in f(-(\tilde{M}^2)^t)$. However, by construction of $\tilde{M}^2$, a_2 dominates a_1 in $\tilde{M}^2$. This contradicts the rationality of the row player. □

4.2 Illustrative Example

To illustrate the proof construction, consider a variant of Rock-Paper-Scissors (or Roshambo), in which each of the traditional three actions is available as a "positive" or "negative" version. The relationship between the positive actions is cyclical as usual whereas that between the negative ones is reversed. Every positive action is beaten by its negative, but beats the other two negative actions. The corresponding payoff matrix M is depicted in Figure 1a. The unique maximin strategy of this game puts probability $1/3$ on each positive action. None of the actions is dominated and therefore playing any of the negative actions is rational even in the presence of common knowledge of rationality.[7] For example, the row player could justifiably play $\bar{r}$ if he believes that his opponent plays r.

Theorem 2 shows that rationality suffices to single out the maximin strategy if both players' are consequentialists and act consistently. Assume for contradiction that both players, instead of playing the maximin strategy, for example play p and q which randomize uniformly over each player's first and fourth actions, respectively. Consequentialism implies that both players still do so if there were 4 clones of the second row. The resulting game $\hat{M}$ is depicted in Figure 1b. Since consequentialism implies permutation invariance, the players would still play p and q if the rows corresponding to these two actions were permuted or the remaining rows were permuted. By consistency, they would still play the same strategy if one of the permutations of $\hat{M}$ was chosen uniformly at random. The game $\bar{M}$ resulting from having to choose a strategy before knowing the outcome of this randomization is

[7] In general, common knowledge of rationality in two-player games is equivalent to the condition that players only assign positive probability to actions that survive the iterated elimination of dominated actions (Pearce, 1984; Bernheim, 1984).

$$\begin{array}{cc}
 & \begin{array}{cccccc} 1/2 & & & 1/2 & & \end{array} \\
\begin{array}{c} 1/2 \\ \\ \\ 1/2 \\ \\ \\ \end{array} & \left(\begin{array}{ccc|ccc}
0 & -1 & 1 & -1 & 1 & 1 \\
1 & 0 & -1 & 1 & -1 & 1 \\
-1 & 1 & 0 & 1 & 1 & -1 \\
\hline
1 & -1 & -1 & 0 & 1 & -1 \\
-1 & 1 & -1 & -1 & 0 & 1 \\
-1 & -1 & 1 & 1 & -1 & 0
\end{array}\right)
\end{array}$$

(a) Original game M.

$$\begin{array}{cc}
 & \begin{array}{cccccc} 1/2 & & & 1/2 & & \end{array} \\
\begin{array}{c} 1/2 \\ \\ \\ \\ \\ \\ 1/2 \\ \\ \\ \end{array} & \left(\begin{array}{cccccc}
0 & -1 & 1 & -1 & 1 & 1 \\
\hline
1 & 0 & -1 & 1 & -1 & 1 \\
1 & 0 & -1 & 1 & -1 & 1 \\
1 & 0 & -1 & 1 & -1 & 1 \\
1 & 0 & -1 & 1 & -1 & 1 \\
\hline
-1 & 1 & 0 & 1 & 1 & -1 \\
1 & -1 & -1 & 0 & 1 & -1 \\
-1 & 1 & -1 & -1 & 0 & 1 \\
-1 & -1 & 1 & 1 & -1 & 0
\end{array}\right)
\end{array}$$

(b) Game $\hat{M}$ with 4 clones of the second row.

$$\begin{array}{cc}
 & \begin{array}{cccccc} 1/2 & & & 1/2 & & \end{array} \\
\begin{array}{c} 1/2 \\ \\ \\ \\ \\ \\ 1/2 \\ \\ \\ \end{array} & \left(\begin{array}{cccccc}
1/2 & -1 & 0 & -1/2 & 1 & 0 \\
\hline
1/7 & 1/7 & -4/7 & 5/7 & -4/7 & 4/7 \\
1/7 & 1/7 & -4/7 & 5/7 & -4/7 & 4/7 \\
1/7 & 1/7 & -4/7 & 5/7 & -4/7 & 4/7 \\
1/7 & 1/7 & -4/7 & 5/7 & -4/7 & 4/7 \\
1/7 & 1/7 & -4/7 & 5/7 & -4/7 & 4/7 \\
\hline
1/2 & -1 & 0 & -1/2 & 1 & 0 \\
\hline
1/7 & 1/7 & -4/7 & 5/7 & -4/7 & 4/7 \\
1/7 & 1/7 & -4/7 & 5/7 & -4/7 & 4/7
\end{array}\right)
\end{array}$$

(c) Uniform randomization $\bar{M}$ over all $\hat{M}_{\pi,\mathrm{id}}$.

$$\begin{array}{cc}
 & \begin{array}{cccccc} 1/2 & & & 1/2 & & \end{array} \\
\begin{array}{c} 1/2 \\ \\ \\ \\ \\ \\ 1/2 \\ \\ \\ \end{array} & \left(\begin{array}{c|cc|c|cc}
0 & 0 & 0 & 0 & 0 & 0 \\
\hline
3/7 & -3/28 & -3/28 & 3/7 & -3/28 & -3/28 \\
3/7 & -3/28 & -3/28 & 3/7 & -3/28 & -3/28 \\
3/7 & -3/28 & -3/28 & 3/7 & -3/28 & -3/28 \\
3/7 & -3/28 & -3/28 & 3/7 & -3/28 & -3/28 \\
3/7 & -3/28 & -3/28 & 3/7 & -3/28 & -3/28 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 \\
\hline
3/7 & -3/28 & -3/28 & 3/7 & -3/28 & -3/28 \\
3/7 & -3/28 & -3/28 & 3/7 & -3/28 & -3/28
\end{array}\right)
\end{array}$$

(d) Uniform randomization $\tilde{M}$ over all $\bar{M}_{\mathrm{id},\sigma}$.

$$\begin{array}{cc}
 & \begin{array}{cc} 1 & \end{array} \\
\begin{array}{c} 1 \\ \\ \end{array} & \begin{pmatrix} 0 & 0 \\ 3/7 & -3/28 \end{pmatrix}
\end{array}$$

(e) Game M^2.

$$\begin{array}{cc}
 & \begin{array}{ccc} 1 & & \end{array} \\
\begin{array}{c} 1 \\ \\ \end{array} & \begin{pmatrix} 0 & 0 & 0 \\ 3/7 & 3/7 & -3/28 \end{pmatrix}
\end{array}$$

(f) Game $\hat{M}^2$.

$$\begin{array}{cc}
 & \begin{array}{ccc} 1 & & \end{array} \\
\begin{array}{c} 1 \\ \\ \end{array} & \begin{pmatrix} 0 & 0 & 0 \\ 3/7 & 9/56 & 9/56 \end{pmatrix}
\end{array}$$

(g) Uniform randomization $\bar{M}^2$ over all $\hat{M}^2_{\mathrm{id},\sigma}$.

$$\begin{array}{cc}
 & \begin{array}{cc} 1 & \end{array} \\
\begin{array}{c} 1 \\ \\ \end{array} & \begin{pmatrix} 0 & 0 \\ 3/7 & 9/56 \end{pmatrix}
\end{array}$$

(h) Game $\tilde{M}^2$.

Figure 1: Payoff matrices for the variant of Rock-Paper-Scissors and the corresponding games constructed in the proof of Theorem 2 under the (incorrect) assumption that players randomize uniformly between the first and the fourth action. Probabilities of the player's strategies are denoted as row and column headers (zeros omitted).

depicted in Figure 1c. Similarly, the players would still play p and q if the columns of $\bar{M}$ were permuted likewise and one of those games was chosen uniformly at random.

In the resulting game $\tilde{M}$ (see Figure 1d), both players only have two different types of actions; the actions within the support of p and q, respectively, and the remaining actions. Moreover, since the second row was cloned sufficiently often in $\hat{M}$, the actions in the support of p yield less expected payoff against q than the remaining actions. By consequentialism, the game $\tilde{M}$ can be reduced to the game M^2 shown in Figure 1e, where both players only have two actions, the first of which is played with probability one. Observe that the first row is not a best response to the first column in M^2. Yet, the first row is *not* dominated by the second row. Now consider the game $\hat{M}^2$ depicted in Figure 1f with two clones of the first column. Consequentialism implies that both players can still play their first actions in $\hat{M}^2$. Again by consequentialism, they would still do so if a coin was tossed to decide if $\hat{M}^2$ was played or the game that results from it by permuting the last two columns. The game $\bar{M}^2$ resulting from having to choose a strategy before the coin toss is depicted in Figure 1g. Consistency implies that both players would play their first action in $\bar{M}^2$. By construction of $\bar{M}^2$, the last two columns are clones in $\bar{M}^2$. Hence, a final application of consequentialism implies that the row player would play the first row in the game $\tilde{M}^2$ depicted in Figure 1h. However, the first row is dominated by the second row in $\tilde{M}^2$, which contradicts the rationality of the row player.

4.3 Extension to Non-Zero-Sum Games

When applied to two-player normal-form (henceforth bimatrix) games, our axioms imply that every strategy profile composed of one recommended strategy for each player constitutes a Nash equilibrium. Since Nash equilibria are in general not interchangeable, we cannot characterize the solution concept that returns all Nash equilibria of a given game. However, we can characterize solution concepts that map to an interchangeable subset of Nash equilibria. Such subsets always exist because any set consisting of only one equilibrium is interchangeable. This effectively shifts the burden of solving the equilibrium selection problem to the solution concept (for instance based on Schelling's focal points).

An alternative interpretation of this result is obtained when restricting the set of considered bimatrix games to the set of games in which equilibria are interchangeable. Nash (1951) referred to these games as *solvable* bimatrix games. The most natural subclass within this class is formed by zero-sum games, but it also contains Moulin and Vial's (1978) *strategically zero-sum* games and all games that admit a unique equilibrium. Within the class of solvable games, Nash equilibrium satisfies consistency.

Interestingly, there are bimatrix games that admit a unique Nash equilibrium whose strategies do not coincide with maximin strategies (Aumann and Maschler, 1972). In these games, our axioms characterize equilibrium rather than maximin play. This is because solution concepts, as we defined them, give recommendations based on the same assumptions for both players. Recommending maximin strategies to both players in these games seems intuitively inconsistent (and, in fact, it does violate our formal consistency axiom).[8]

Yet another way to extend our result to bimatrix games is obtained by changing the model such that solution concepts return *strategy profiles* rather than strategies for just one player. In this model, consistency requires that the intersection of strategy profiles chosen by f in two games has to be a subset of strategy profiles chosen by f in any convex combination of these two games. It can then be concluded that all returned strategy profiles for a given bimatrix game are Nash equilibria. However, in this model, strategic advice given to one player is not independent of the advice given to the other player.

Extensions to games with more than two players are not straightforward because these games may not have equilibria with rational probabilities even if all payoffs are rational-valued (see Remark 4).

5 Concluding Comments

We conclude this paper with a number of remarks.

Remark 1 (Independence of axioms). All properties in Theorem 2 are required to derive the conclusion. The trivial solution concept that always returns all strategies violates rationality but satisfies consistency and consequentialism. The solution concept *maximax* that returns all randomizations over rows that contain a maximal entry of the game matrix violates consistency but satisfies the remaining properties. To see this, consider the following games where the row player can play either top or bottom ($A = \{t, b\}$) and the column player can play either left, middle, or right ($B = \{l, m, r\}$).

$$\hat{M} = \begin{pmatrix} 5 & 1 & 0 \\ 4 & 4 & 0 \end{pmatrix} \quad \bar{M} = \begin{pmatrix} 1 & 5 & 0 \\ 4 & 4 & 0 \end{pmatrix} \quad M = {}^1\!/_2\,\hat{M} + {}^1\!/_2\,\bar{M} = \begin{pmatrix} 3 & 3 & 0 \\ 4 & 4 & 0 \end{pmatrix}$$

Then, $maximax(\hat{M}) = maximax(\bar{M}) = \{t\}$ and $maximax(-\hat{M}^t) = maximax(-\bar{M}^t) = \{r\}$, and consistency would imply that $maximax(M) = \{t\}$. However, $maximax(M) = \{b\}$.

The solution concept *average* that returns all randomizations over rows with the highest average payoffs (i.e., it best-responds to a uniform strategy by the opponent) violates consequentialism but satisfies the remaining properties. Violations of consequentialism arise from deleting cloned columns as in the following example (again, $A = \{t, b\}$) and $B = \{l, m, r\}$).

$$M = \begin{pmatrix} 0 & 2 & 2 \\ 3 & 0 & 0 \end{pmatrix} \qquad \hat{M} = \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix}$$

Clearly, $average(M) = \{t\}$ but $average(\hat{M}) = \{b\}$, whereas consequentialism implies that the same strategy must be played in both games.

[8] In Aumann and Maschler's 1972 introductory example, maximin strategies violate consistency because taking the uniform convex combination of the original game and the one in which both rows are permuted yields a game in which the column player's maximin strategy of the original game is dominated.

Remark 2 (Strong consistency). The solution concept *maximin* violates the stronger notion of consistency where it is only required that $f(\hat{M}) \cap f(\bar{M}) \neq \emptyset$, but not necessarily that $f(-\hat{M}^t) \cap f(-\bar{M}^t) \neq \emptyset$. Consider the games $\hat{M}$ and $\bar{M}$ in $\mathbb{Q}^{A \times B}$ where the row player can play either top or bottom ($A = \{t, b\}$) and the column player can play either left or right ($B = \{l, r\}$).

$$\hat{M} = \begin{pmatrix} 0 & 4 \\ 2 & 0 \end{pmatrix} \quad \bar{M} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \quad M = {}^1\!/_2\, \hat{M} + {}^1\!/_2\, \bar{M} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$$

The unique maximin strategy in $\hat{M}$ and $\bar{M}$ is ${}^1\!/_3\, t + {}^2\!/_3\, b$. But in the game M, which results from randomizing uniformly over $\hat{M}$ and $\bar{M}$, the unique maximin strategy is to play t with probability one. In particular, ${}^1\!/_3\, t + {}^2\!/_3\, b$ is not a maximin strategy in M. Notice that the maximin strategies in $-\hat{M}^t$ and $-\bar{M}^t$ are different.

As a consequence, Theorem 2 turns into an impossibility theorem when replacing consistency with strong consistency.

Remark 3 (Symmetric games). For symmetric games, the strong notion of consistency discussed in Remark 2 is equivalent to consistency and is hence satisfied by *maximin*. Theorem 2 remains valid within the domain of symmetric games. This requires modifying the proof such that all constructed games are symmetric.[9] More precisely, $\hat{M}$ has to be defined such that A_a is a set of clones for the row player and for the column player for all $a \in A$. The game $\bar{M}$ can be defined by summing over all $\hat{M}_{\pi\pi}$, where π ranges over the same set of permutations as in the original proof. The step of constructing the game $\tilde{M}$ is not necessary. Instead, the game M^2 can be obtained directly from $\bar{M}$ by applying consequentialism. It follows directly from symmetry of M^2 that a_2 dominates a_1, which makes the remainder of the proof disposable.

Remark 4 (Real-valued payoffs). The proof of Theorem 2 relies on the fact that payoffs and probabilities are rational-valued (actually, games with irrational payoffs are only problematic as far as they do not admit rational-valued maximin strategies). We believe that Theorem 2 could be extended to real-valued payoffs at the expense of a significantly more complicated framework (games with a continuum of actions and compact measurable subsets of actions as feasible sets) or additional technical axioms (upper hemi-continuity of solution concepts and the requirement that rational-valued strategies are returned in a dense subset of games).

Acknowledgements

This material is based on work supported by the Deutsche Forschungsgemeinschaft under grant BR 2312/11-1. The authors thank Martin Bullinger, Paul Harrenstein, Dominik Pe-

[9] For the illustrative example in Section 4.2, the standard proof is used even though the game is symmetric.

ters, and William S. Zwicker for helpful comments. The authors are furthermore indebted to the editor and two anonymous referees for their constructive feedback.

References

Kenneth J. Arrow and Leonid Hurwicz. An optimality criterion of decision-making under ignorance. In C. F. Carter and J. L. Ford, editors, *Uncertainty and Expectations in Economics*. Basil Blackwell, 1972.

Robert J. Aumann and Adam Brandenburger. Epistemic conditions for Nash equilibrium. *Econometrica*, 63(5):1161–1180, 1995.

Robert J. Aumann and Jacques H. Drèze. Rational expections in games. *American Economic Review*, 98(1):72–86, 2008.

Robert J. Aumann and Michael Maschler. Some thoughts on the minimax principle. *Management Science*, 18(5):54–63, 1972.

Michael Bacharach. Zero-sum games. In J. Eatwell, M. Milgate, and P. Newman, editors, *The New Palgrave: A Dictionary of Economics*, volume 2. MacMillan, 1987.

B. Douglas Bernheim. Rationalizable strategic behavior. *Econometrica*, 52(4):1007–1028, 1984.

Florian Brandl, Felix Brandt, and Hans Georg Seedig. Consistent probabilistic social choice. *Econometrica*, 84(5):1839–1880, 2016.

Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM*, 56(3), 2009.

Herman Chernoff. Rational selection of decision functions. *Econometrica*, 22(4):422–443, 1954.

Constantinos Daskalakis, Paul Goldberg, and Christos H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

Itzhak Gilboa and David Schmeidler. A derivation of expected utility maximization in the context of a game. *Games and Economic Behavior*, 44(1):172–182, 2003.

Sergiu Hart, Salvatore Modica, and David Schmeidler. A Neo^2 Bayesian foundation of the maxmin value for two-person zero-sum games. *International Journal of Game Theory*, 23(4):347–358, 1994.

Eric Maskin. Decision-making under ignorance with implications for social choice. *Theory and Decision*, 11(3):319–337, 1979.

John Milnor. Games against nature. In *Decision Processes*, chapter 4, pages 49–59. Wiley, 1954.

Hervé Moulin and Jean-Philippe Vial. Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon. *International Journal of Game Theory*, 7(3–4):201–221, 1978.

John F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences (PNAS)*, 36:48–49, 1950.

John F. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

Henk Norde and Mark Voorneveld. Characterizations of the value of matrix games. *Mathematical Social Sciences*, 48(2):193–206, 2004.

Henk Norde, Jos Potters, Hans Reijnierse, and Dries Vermeulen. Equilibrium selection and consistency. *Games and Economic Behavior*, 12(2):219–225, 1996.

David G. Pearce. Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52(4):1029–1050, 1984.

Bezalel Peleg and Stef H. Tijs. The consistency principle for games in strategic form. *International Journal of Game Theory*, 25(1):13–34, 1996.

Lloyd S. Shapley. A value for n-person games. *Annals of Math Studies*, 28:307–317, 1953.

John H. Smith. Aggregation of preferences with variable electorate. *Econometrica*, 41(6): 1027–1041, 1973.

T. Nicolaus Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, 1987.

Stef H. Tijs. A characterization of the value of zero-sum two-person games. *Naval Research Logistics Quarterly*, 28(1):153–156, 1981.

Èduardas I. Vilkas. Axiomatic definition of the value of a matrix game. *Theory of Probability and its applications*, 8(3):304–307, 1963.

John von Neumann. Zur Theorie der Gesellschaftspiele. *Mathematische Annalen*, 100(1): 295–320, 1928.

John von Neumann and Maurice Fréchet. Communication on the Borel notes. *Econometrica*, 21(1):124–127, 1953.

H. Peyton Young. Social choice scoring functions. *SIAM Journal on Applied Mathematics*, 28(4):824–838, 1975.

H. Peyton Young and Arthur Levenglick. A consistent extension of Condorcet's election principle. *SIAM Journal on Applied Mathematics*, 35(2):285–300, 1978.