Technische Universität München

Fakultät für Mathematik
Lehrstuhl für Wirtschaftsinformatik und Entscheidungstheorie

Zero-Sum Games in Social Choice and Game Theory

Florian Brandl

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende(r):		Prof. Dr. Peter Gritzmann
Prüfer der Dissertation:	1.	Prof. Dr. Felix Brandt
	2.	Prof. Dr. Herve Moulin
	3.	Prof. Dr. Clemens Puppe

Die Dissertation wurde am 16.01.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 10.06.2018 angenommen.

ZERO-SUM GAMES IN SOCIAL CHOICE AND GAME THEORY

FLORIAN BRANDL

E-MAIL:
f.brandl@mytum.de

This thesis is based on a tasteful style designed by Hans Georg Seedig. It was typeset using LaTeX and the ClassicThesis style by André Miede, combined with the ArsClassica package by Lorenzo Pantieri and some minor modifications. The text is set in Palatino with math in Euler, both due to Hermann Zapf. Headlines are set in Iwona by Janusz M. Nowacki. Most of the graphics were created using Ti*k*Z by Till Tantau.

ABSTRACT

Zero-sum games model the most extreme form of competition among players. When there are only two players, von Neumann's minimax theorem shows that every zero-sum game admits a pair of maximin strategies that achieve unique optimal payoffs. After providing a new proof of the minimax theorem, we derive a set of epistemic conditions that necessitates maximin play. For the special case of *symmetric* zero-sum games, we determine the distribution over supports of maximin strategies in randomly chosen games.

In decision theory, zero-sum games appear as representations of preferences over probabilistic outcomes through skew-symmetric bilinear utility functions. A subdomain of these preferences are preferences based on *pairwise comparisons*, for which one outcome is preferred to another outcome if and only if the former is more likely to yield a more preferred alternative. We show that three impossibility results of collective preference aggregation that obtain on the unrestricted domain cease to hold for preferences based on pairwise comparisons: Arrow's dictatorship theorem, Moulin's incompatibility of Condorcet consistency and resistance to the no-show paradox, and the conflict between consistency with respect to variable electorates and consistency with respect to components of similar alternatives.

PUBLICATIONS

This thesis is based on the following publications and working papers.

[1] The distribution of optimal strategies in symmetric zero-sum games. *Games and Economic Behavior*, 104:674–680, 2017.

[2] Arrovian aggregation of convex preferences. (with F. Brandt), Working paper.

[3] Justifying optimal play via consistency. (with F. Brandt), Working paper.

[4] Welfare maximization entices participation. *Games and Economic Behavior*, 2018 (with F. Brandt and J. Hofbauer), Forthcoming.

[5] Consistent probabilistic social choice. *Econometrica*, 84(5):1839–1880, 2016 (with F. Brandt and H. G. Seedig).

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to a number of people who have contributed to the completion of this thesis.

To Felix for introducing me to the topic of this thesis, for the countless times I had a question and immediately received an answer, for providing a productive and pleasant environment within his group, and for being an inspiring example of someone how is truly dedicated to what he does. I consider myself very lucky to have had Felix as my supervisor.

To my present and former colleagues Christian Geist, Johannes Hofbauer, Christian Saile, Hans Georg Seedig, and Christian Stricker for numerous engaging discussions and for never making work feel like work. Special thanks go to Johannes Hofbauer and Christian Saile for valuable comments on preliminary versions of this thesis.

To my coauthors and friends in academia Haris Aziz, Markus Brill, Manuel Eberl, Rupert Freeman, Paul Harrenstein, Martin Olsen, Dominik Peters, Martin Strobel, Warut Suksompong, and Kavitha Telikepalli for the work we have done together and for the good times we had at conferences.

To Dirk Bergemann, Klaus Nehring, Ariel Procaccia, and Clemens Puppe for their career advice and for invaluable comments on my work. I owe particular thanks to Dirk Bergemann and Ariel Procaccia for enabling me to visit Yale University and Carnegie Mellon University.

To my friends at TUM Christoph, Katharina, Michi, and Patrick for always being approachable in academic and non-academic matters and for providing the necessary distractions over the past years.

To my parents for their unconditional support starting long before I began crafting this thesis.

CONTENTS

1 INTRODUCTION

Game theory and social choice theory are two closely related fields located at the intersection of mathematics and economics. Both address aspects of decision making—individual decision making in the context of game theory and collective decision making in the context of social choice theory. Game theory could be described as the study of strategic interaction of multiple rational and self-interested players. Typically, the actions taken by one player influence not only his own outcome but also the outcome for the other players. Rational players are capable of reasoning about the actions taken by other players while taking into account that other players are capable of this, as well. Self-interest refers to the fact that players seek to get an outcome that is in their own best interest. This does not mean that players are *per se* selfish; whichever benevolent considerations players may undertake can be taken as primitives for their valuations of outcomes.

Conversely, social choice theory studies decision making by a society of agents that strives to make a collectively desirable decision. Alternatively, one can think of this problem from the perspective of an outside social planner who knows the preferences of all agents over all possible outcomes and, based on those, wants to reach a decision that is most desirable for the society. The social planner himself has no preferences over the possible outcomes and is thus impartial. Individual agents are not bound to report their true preferences, however, and may choose not to do so if it enables them to potentially influence the collective decision in their favor. This is where game theory enters social choice theory.

1.1 PREFERENCES OVER UNCERTAIN OUTCOMES

The prevailing model in game theory is that players may choose their actions probabilistically. While social choices are typically assumed to be deterministic, it can also make sense to allow them to randomize over the possible deterministic alternatives in the appropriate context, for example when decisions are low stakes or repeated frequently. In both instances, agents are faced with deciding among uncertain outcomes. Understanding choices by single agents in the absence of strategic interaction is vital for any analysis of strategic or social choice.

The predominant model for preferences over uncertain outcomes are preferences that can be represented by a linear utility function.

Von Neumann and Morgenstern (1944) have shown that these preferences, henceforth called vNM preferences, are characterized by three axioms called continuity, transitivity, and independence. Continuity prescribes that the preference between two outcomes should not be reversed by slight perturbations of the outcomes. Transitivity requires the preference relation to be transitive. Independence prescribes that one outcome should be preferred to another if and only if a coin toss between the former and a third outcome is preferred to a coin toss between the latter and the third outcome when the same coin is used in both cases. Geometrically, independence means that the preference between two outcomes does not change if they are shifted in the same direction by the same magnitude within the simplex of probability distributions. It is important to keep in mind that without further justification, any utility function is merely a representation of ordinal preferences rather than a direct numerical measure for satisfaction. A utility function "[...] is not in itself a basis for numerical comparison of utilities for one person nor of any comparison between different persons" (von Neumann and Morgenstern, 1944).

Experiments have shown systematic violations of both independence and transitivity. The Allais Paradox (Allais, 1953) is perhaps the most famous example for violations of independence. Kahnemann and Tversky (1979) described various further examples, many of which are based on the certainty effect, which also contributes to the Allais paradox. It prescribes that human decision makers experience a greater loss when moving from certainty to almost-certainty than when moving from a moderate chance of winning to a slightly lower chance of winning. Detailed accounts of violations of the independence axiom are provided by Machina (1983), Machina (1989), and McClennen (1988). Even the transitivity axiom, once deemed indispensable for rational decision making, has been subject to criticism (see, e.g., May, 1954; Fishburn, 1970; Bar-Hillel and Margalit, 1988; Fishburn, 1991; Anand, 1993; Anand, 2009). One instance demonstrating violations of transitivity is the preference reversal phenomenon, which describes that a decision maker prefers one uncertain outcome to another uncertain outcome, but exhibits reversed preferences over the certainty equivalents of both outcomes (Grether and Plott, 1979). Still, transitivity has been very persistent in the economics literature. The prime reason for insisting on transitivity is presumably that it guarantees the existence of maximal elements, i.e., undominated outcomes, which are the basis for being able to make sensible choices. When every finite set of outcomes is considered feasible, acyclicity (a weakening of transitivity) of a preference relation is equivalent to the existence of maximal elements. Acyclicity prescribes that all outcomes can be ordered on a line such that, for any two outcomes, the more preferred outcome (if any) is to the left of the less preferred outcome. Hence, acyclicity retains the one-dimensional spirit of transitivity. When the set of outcomes itself is convex, it can be argued that feasible sets should also be convex. Sonnenschein (1971) showed that if a preference relation is continuous and has convex upper contour sets, then every non-empty, compact, and convex set of outcomes admits a maximal element.

Another frequently cited argument in favor of transitivity is the money pump, which demonstrates a situation where an unlimited amount of money is elicited from an agent with intransitive preferences by repeatedly offering him a more preferred outcome in exchange for his current outcome plus a small amount of money. The money pump relies on the possibility to confront an agent with a sequence of choices from a small set of outcomes, however. If the agent was offered to choose from the convex hull of the set of all outcomes that he is offered in the process, he would choose a maximal element and not exchange it for any other outcome from this set later on. This is similar to an argument by Blavatskyy (2006), who argues that repeated choices from small sets should be perceived as one choice from the union of these sets in which case the agent will choose a maximal element from the large set and cannot be exploited by the money pump. Fishburn (1991) objects that the money pump "applies transitive thinking to an intransitive world", since preferences over money are (assumed to be) transitive.

Skew-symmetric bilinear (SSB) utility theory (Fishburn, 1982) can accommodate both, the Allais paradox and the preference reversal phenomenon. An SSB function is a skew-symmetric and bilinear function that maps an ordered pair of outcomes to a real number. A preference relation is represented by an SSB function if this number is positive exactly when the first outcome is preferred to the second outcome. Fishburn (1982) characterized preference relations that can be represented by an SSB function via continuity, convexity, and symmetry. The latter two properties are weakenings of the conjunction of transitivity and independence. A preference relation is convex if, for every outcome, the set of outcomes it is indifferent to is a hyperplane that separates its upper and lower contour set. The symmetry axiom requires that the indifference curves for every triple of outcomes are either parallel or intersect in one point, which may be outside of their convex hull. Due to bilinearity, the SSB value for two outcomes can be derived from the SSB values for pairs of pure outcomes (cf. Fishburn, 1984c). Hence, every SSB function can be represented by a

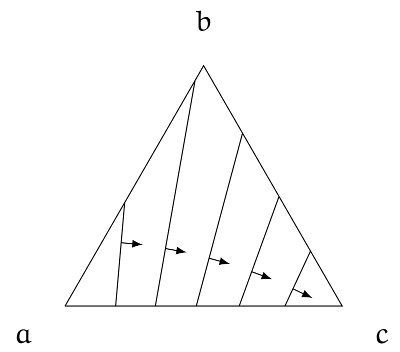


Figure 1.1: Illustration of the preference relation $\succ$ represented by the SSB matrix in Example 1. The arrows represent the normal vectors to the indifference curve of the outcome at the base of the arrow (pointing towards the lower contour set). Each indifference curve separates the corresponding upper and lower contour set.

skew-symmetric matrix.[1] Example 1 shows an SSB function ϕ on the set of probability distributions over $\{a, b, c\}$.

$$\phi = \begin{array}{cc} \begin{array}{ccc} a & b & c \end{array} & \\ \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 2 \\ -2 & -2 & 0 \end{pmatrix} & \begin{array}{c} a \\ b \\ c \end{array} \end{array} \qquad \text{(Example 1)}$$

We write $\underline{x}$ to denote the pure outcome that assigns probability one to $x \in \{a, b, c\}$. According to the preference relation $\succ$ represented by ϕ, $\underline{a}$ is preferred to $\underline{b}$ and to $\underline{c}$ and $\underline{b}$ is preferred to $\underline{c}$. When writing outcomes as convex combinations of pure outcomes, the SSB value between any two outcomes can be conveniently determined by multiplying them to the matrix ϕ from left and right, respectively, e.g., for $p = 1/2\,\underline{a} + 1/2\,\underline{b}$ and $q = 2/3\,\underline{a} + 1/3\,\underline{c}$, we have $\phi(p, q) = p^t \phi q = 2/3$. The minimax theorem guarantees the existence of an outcome p with $p^t \phi \geqslant 0$ and hence, the existence of maximal elements of preference relations satisfying the SSB axioms.

1.2 GAME THEORY

Game theory models strategic interaction among multiple players with possibly conflicting preferences. When players are allowed to randomize over the actions they could take, it is standard to assume that their preferences over outcomes (randomizations over action profiles) satisfy the vNM axioms and can hence be represented by a linear utility function. Equivalently, one could assume that every action profile assigns a payoff to every agent and agents prefer higher expected

1 A matrix $M \in \mathbb{R}^{m \times m}$ is skew-symmetric (or anti-symmetric) if $M = -M^t$.

payoffs. The most severe form of conflict arises when there are only two players with completely opposed preferences. In this case, the preferences of both players can be represented by a single matrix that specifies the payoffs of the first player and, when negated, gives the payoffs of the second player. Hence, the expected payoffs sum up to zero for all randomizations over actions, which coined the term *zero-sum game*. Despite the apparent limitations of this class of games, much of the early work on game theory, including von Neumann and Morgenstern (1944), focuses on zero-sum games (see also Wald, 1945b; Kaplansky, 1945; Bohnenblust et al., 1950; Kuhn and Tucker, 1950). In a zero-sum game every action of the first player corresponds to a row and every action of the second player corresponds to a column of the matrix representing the game. Therefore, the players are subsequently referred to as the row player and the column player. In the game below, the row player has three actions—top (t), middle (m), and bottom (b)—and the column player has two actions—left (l) and right (r).

$$\begin{array}{cc} & \begin{array}{cc} l & r \end{array} \\ \begin{array}{c} t \\ m \\ b \end{array} & \begin{pmatrix} 0 & 1 \\ 2 & 0 \\ 1 & -1 \end{pmatrix} \end{array} \qquad \text{(Example 2)}$$

When the game is played, both players simultaneously choose an action, possibly in a probabilistic way. The strategy of a player is his randomization over actions. The objective of game theory is to provide a basis for players to choose their strategies. For the game in Example 2, it is clear at first glance that the row player should never play b, since playing m yields a higher payoff for every strategy of the column player. Apart from that it is not at all obvious what the players should play, since every action is a best response against some strategy of the other player. Any further recommendations for one player will thus depend on assumptions about the other player.

1.2.1 The Minimax Theorem

If the row player plays t with probability two thirds and m with probability one third, then his expected payoff is two thirds independently of the action chosen by the column player. This strategy is denoted by $p = {}^2/_3\, \underline{t} + {}^1/_3\, \underline{m}$. No other strategy can guarantee a higher expected payoff. Similarly, the column player can guarantee an expected payoff of minus two thirds by playing $q = {}^1/_3\, \underline{l} + {}^2/_3\, \underline{r}$. Observe that the guaranteed expected payoffs of both players add up to 0. This is not a coincidence. Von Neumann (1928) showed that every zero-sum game admits a value, a real number, such that the row player can play a strategy that guarantees himself an expected payoff that is at least as high as the value independently of the strategy of the column player and, additionally, the column player can play a strategy that guarantees himself an expected payoff that is at least as high as the negative of the value independently of the strategy of the row player.[2] This statement is known as the *minimax theorem*; the corresponding strategies are called *maximin strategies*. It has been influential in mathematics far beyond game theory (e.g., in linear programming duality). Its importance for game theory is highlighted by a quote of von Neumann: "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" (Casti, 1996). The proof by von Neumann (1928) was of analytic nature and could be fittingly described as a "tour de force" (Heims, 1980). Subsequent work on the minimax theorem provided simpler proofs (von Neumann and Morgenstern, 1944; Loomis, 1947), generalized it (Wald, 1945a; Fan, 1953), and highlighted its connections to fixed-point theorems (von Neumann, 1937; Fan, 1952). Kjeldsen (2001) gives a thorough discussion of the history of the minimax theorem. It is of practical and conceptual importance that maximin strategies can be found efficiently, i.e., in polynomial time in the size of the game matrix. Otherwise it would seem like a demanding and perhaps unwarranted assumption that players are actually able to play maximin strategies. In the first part of this thesis, we give yet another proof of the minimax theorem. It is purely algebraic for generic games and employs analytical methods only for degenerate games. Readers familiar with basic linear algebra and analysis will be able to follow it easily.

1.2.2 Justification of Maximin Play

We argued that in Example 2 a player who is interested in maximizing his expected payoff should never play b, since m dominates b, i.e., m yields a higher payoff no matter which action is played by the column player. A player who never plays dominated actions is called *rational*. Example 2 shows that rationality is not sufficient to force a player to play a maximin strategy, since all strategies that randomize only over t and m are in accordance with rationality. This begs the question under which assumptions about the opponent a player should play a maximin strategy. Von Neumann and Morgenstern (1944) argue that a player should play a maximin strategy if there is no valid basis for making any assumptions about the other player, since it expresses cautious behavior in the absence of justified alternatives. This however is an informal argument that lacks theoretical defense. It is not

2 Every zero-sum game induces a bilinear function that maps a pair of probability distributions to a real number (the expected payoff of the row player). The original proof by von Neumann (1928) even holds for functions that are quasi-convex in the first argument and quasi-concave in the second argument.

clear why players should expect to face a worst case opponent. Typically it will be reasonable for a player to make some assumptions about his opponent, however, e.g., that he is rational. The discipline that studies game theory in the presence of such beliefs about the opponent is called epistemic game theory (see Perea (2012) for a survey of the corresponding literature).

A model to study epistemic game theory are interactive belief systems due to Harsanyi (1967). They allow to conveniently specify belief hierarchies such as "the row player thinks that the column player thinks that the row player is rational." Central to epistemic justifications for playing maximin strategies is some notion of rationality, e.g., rationality (Aumann and Brandenburger, 1995), mutual knowledge of rationality (Barelli, 2009), or common knowledge of rationality (Aumann and Drèze, 2008).[3] Mutual knowledge of rationality prescribes that players are rational and know that their opponent is rational. They do not need to know that their opponent knows that they are rational, etc. Common knowledge of rationality requires that rationality is known for arbitrarily long belief hierarchies.

The framework for justifying maximin strategies used here is different in that the players' beliefs about other players are reflected by their choices of strategy, e.g., *mutual knowledge of rationality* implies that a player never plays an action that is dominated assuming the other player never plays a dominated action. Other conditions, which in combination with mutual knowledge of rationality will imply maximin play, are formulated in a framework with variable sets of actions. It is assumed that there is a universal set of actions that both players can choose from. A proto game specifies the payoffs for all combinations of actions. In any given situation, only a finite set of actions is feasible for both players. A player has to choose a strategy given a proto game and sets of feasible actions. Only strategies whose support is contained in the set of feasible actions are feasible.[4] Restricting a proto game to sets of feasible actions yields a game in the usual sense.

Independence of infeasible actions prescribes that the strategy chosen by a player must not depend on the payoffs for actions that are not feasible. Independence of infeasible actions is the game theoretic analog of Arrow's independence of infeasible alternatives in social choice theory (Arrow, 1951). In some games there are actions that are indistinguishable from each other in terms of their consequences, i.e., they yield the same payoff against every action of the opponent. Such actions are called clones. If a player was to treat clones differently, this would be solely based on their names. *Consequentialism* requires that the probabilities for non-clones should not depend on how many

3 These authors use a stronger notion of rationality, which prescribes that players choose an action that maximizes their expected payoff given their beliefs about other players.

4 The support of a strategy is the set of actions to which it assigns positive probability.

clones of an action are feasible. Moreover, the overall probability assigned to clones can be distributed arbitrarily among them. The last condition prescribes how players deal with games that they consider equivalent in terms of chosen strategies, i.e., they would choose the same pair of strategies, one as the row player and another one as the column player, in both games. Now assume that a coin is tossed to decide which one of two games that are considered equivalent in this sense is played and a player has to decide on his strategy before knowing the outcome of the coin toss. *Consistency* prescribes that any strategy that he would choose in both games is also chosen prior to the coin toss. This assumes that the game resulting from choosing a strategy before the execution of the coin toss is treated in the same way as the game whose payoffs are the expected payoffs of this randomization. We show that every player whose strategic choices respect mutual knowledge of rationality and consequentialism and are independent of infeasible actions and consistent has to choose maximin strategies.

1.2.3 Random Symmetric Zero-Sum Games

The model of games discussed above assumes that the players know about the payoffs associated with each action profile. This assumption is dropped when studying *random games*, i.e., only a probability distribution over games is known, but the realization of the game is unknown. For example, one could consider a probability distribution over zero-sum games, where all payoffs are drawn from independent normal distributions. Questions about random games typically address determining the distribution of some characteristic of the eventually realized game such as the number of Nash equilibria or the value in the case of zero-sum games. Finding the exact distribution of "good" strategies is typically a very hard task. Hence, a number of authors have studied the distribution of the support of Nash equilibrium strategies or maximin strategies (McLennan, 2005; McLennan and Berg, 2005; Faris and Maier, 1987; Jonasson, 2004). We continue this line of work by determining the distribution of maximin strategies of random *symmetric zero-sum games*. A zero-sum game is symmetric if both players have the same set of actions and if they swap actions then their payoffs are swapped, as well. This is reflected by the fact that the payoff matrix is skew-symmetric.

A very well-known example of a symmetric zero-sum game is "rock, paper, scissors". It is played between two players who simultaneously choose either rock (r), paper (p), or scissors (s) by displaying the chosen object with their hand. The payoffs are determined as follows: rock beats (smashes) scissors, scissors beats (cuts) paper, and paper beats (covers) rock. The winner gets payoff 1 and the loser gets payoff -1. If both players choose the same object, the game is tied and both players get payoff 0. The corresponding payoff matrix is depicted in Example 3.

$$\begin{array}{c} \\ r \\ p \\ s \end{array}\begin{array}{c} \begin{array}{ccc} r & p & s \end{array} \\ \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \end{array} \tag{Example 3}$$

The unique maximin strategy for both players is to play $1/3\,\underline{r} + 1/3\,\underline{p} + 1/3\,\underline{s}$. We show that the probability that a maximin strategy of a randomly drawn symmetric zero-sum game has a given support is 0 if the support has even cardinality and $2^{-(m-1)}$ if the support has odd cardinality, where m is the dimension of the payoff matrix.

Symmetric zero-sum games appear in many areas of natural science such as biology, physics, and chemistry. We only give two examples here. In evolutionary biology, they can be used to model population dynamics among multiple species with actions corresponding to species and payoffs corresponding to the probabilities that an individual from one species "beats" an individual from another species; the probabilities of a maximin strategy specify the fractions of individuals from each species in a stable state. Hence, the support of a maximin strategy corresponds to the set of species that survive in a stable state (see, e.g., Allesina and Levine, 2011; Levine et al., 2017; Grilli et al., 2017). In quantum physics, symmetric zero-sum games can be used to model bosonic systems where different quantum states take the role of actions and the transition probabilities from one state to another form the payoff matrix. Knebel et al. (2015) consider the support of maximin strategies in these games to determine which states become condensates at a macroscopic level.

1.2.4 Normal-Form Games

Games that involve more than two players or that are not zero-sum are not considered in this thesis. The *normal-form representation* of such games specifies a payoff function for every player, which gives his payoff for every combination of actions taken by the group of all players. The most wide-spread solution concept for normal-form games is *Nash equilibrium*, which prescribes a state in which no player has an incentive to unilaterally change his strategy given the strategies of the other players. Nash (1950a) has famously shown that every normal-form game admits at least one Nash equilibrium. For zero-sum games, Nash equilibria correspond to pairs of maximin strategies. Hence, the existence of Nash equilibria guarantees the existence of maximin strategies. In contrast to maximin strategies, Nash equilibrium strategies are not interchangeable, i.e., replacing the strategy of one player in a Nash equilibrium by a strategy he plays in some other Nash equilibrium does not necessarily yield another Nash equilibrium (see also Nash, 1951). Hence, it is highly unclear which strategy a player should choose in case there are multiple Nash equilibria and obtaining a Nash equilibrium requires coordination among the players. This makes the concept of Nash equilibrium questionable from a normative viewpoint.

1.3 SOCIAL CHOICE THEORY

Social choice theory examines the aggregation of the preferences of multiple agents over possible alternatives. Its origins date back at least to the time of the French revolution. The French mathematician Jean Charles de Borda (1784) proposed a voting system that is nowadays known as Borda's rule. Every agent assigns a score to every alternative that equals the number of alternatives that he likes less. The alternatives with the highest accumulated score are implemented as the social choice. Borda's rule was used by the French Academy of Sciences at that time. It is a representative of the class of scoring rules, i.e., rules for which every agent assigns a score to every alternative that is based on the alternative's position in his ranking. Another French mathematician, the Marquis de Condorcet (1785) realized that Borda's rule may fail to choose a *Condorcet winner*—an alternative that is preferred to every other alternative by a majority of agents. In fact, every scoring rule may fail to choose a Condorcet winner (Fishburn, 1973; Young and Levenglick, 1978). The important observation that Condorcet winners may fail to exist is also due to Condorcet. Consider a setting with three alternatives called a, b, and c and three agents called 1, 2, and 3 whose preference relations are shown in Example 4.

1	2	3
a	b	c
b	c	a
c	a	b

(Example 4)

The ith column depicts the preferences of the ith agent, e.g., agent 1 prefers a to b, b to c, and a to c. The collection of all agents' preference relations is called a preference profile. A majority of two agents prefers a to b. Similarly, a majority prefers b to c and c to a. Hence, the majority relation is cyclic and there is no Condorcet winner. The property of choosing a Condorcet winner whenever it exists is called Condorcet consistency. The majority margin of a over b is the number of agents who prefer a to b minus the number of agents who prefer b to a. In Example 4, the majority margin of a

over b, b over c, and c over a is 1 in each case. Example 5 depicts the corresponding matrix of majority margins.

$$\begin{array}{cc} & \begin{array}{ccc} a & b & c \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \end{array} \qquad \text{(Example 5)}$$

The matrix of majority margins is necessarily skew-symmetric. Hence, it can be interpreted as a symmetric zero-sum game, where the set of actions is the set of alternatives. This connection between game theory and social choice theory will play a crucial role for the upcoming analysis.

Much of the modern literature on social choice theory is inspired by Arrow (1951), who reformulated the social choice problem by introducing *social welfare functions (SWFs)*. An SWF maps a preference profile to a collective preference relation. Arrow's conception of SWFs is different from Bergson-Samuelson SWFs (Bergson, 1938), which consider a single preference profile in isolation and derive social utilities for alternatives by adding up individual utilities. His justification for his departure from this model was that "as with any type of behavior described by maximization, the measurability of social welfare need not be assumed; all that matters is the existence of a social ordering satisfying [completeness and transitivity]" (Arrow, 1951).

Arrow showed that every SWF violates at least one of three desirable and seemingly mild properties called Pareto optimality, independence of irrelevant alternatives, and non-dictatorship, when there are at least three alternatives and preferences have to be transitive. *Pareto optimality* prescribes that if all agents prefer one alternative to another, then the first alternative should also be preferred to the second alternative according to the collective preference relation. *Independence of irrelevant alternatives* requires that the collective preference between two alternatives should only depend on the agents' preferences over these two alternatives and not on their preferences over other alternatives. A dictator is an agent whose strict preferences over alternatives feature as strict preferences in collective preference relation, independently of the preferences of the other agents. *Non-dictatorship* prescribes the absence of a dictator. The potential intransitivity of the majority relation can be thought of as the source of Arrow's negative result.

An implicit assumption of Arrow's theorem is that there are no restrictions on the agents' preference relations apart from transitivity. This is referred to as the full domain assumption. It is hard to argue against it without any further knowledge about the set of alternatives. The situation is different if the set of alternatives admits a non-trivial intrinsic structure, however. Just as in the theory of games when moving from actions to strategies, one can consider enlarging the set of possible choices to probability distributions over alternatives, called outcomes, rather than just deterministic alternatives, i.e., pure outcomes. Hence, the set of outcomes has the structure of a unit simplex. Probability distributions over alternatives can be interpreted as fractional allocations of arbitrary divisible goods such as time, money, or probability. In this framework, an SWF is a function mapping the preferences of the agents over all outcomes to a collective preference relation over all outcomes. Arguably, the structure on the set of outcomes should be reflected in the preference relations, which makes Arrow's full domain assumption seem overly restrictive. Most of the social choice literature on aggregation of preferences over probabilistic outcomes has assumed that agents possess vNM preferences (see, e.g., Harsanyi, 1955; Kalai and Schmeidler, 1977a; Hylland, 1980b; Dhillon and Mertens, 1999). As discussed in Section 1.1, experimental evidence suggests that, depending on the context, the vNM preference model may be insufficient to capture the preferences of human decision makers. It is therefore an interesting task to study preference aggregation for other decision theoretic preference models. In this thesis, we revisit different aspects of the social choice problem for SSB preferences.

1.3.1 Arrovian Preference Aggregation

First, we consider SWFs that satisfy Arrow's axioms of Pareto optimality and independence of irrelevant alternatives, henceforth called *Arrovian SWFs*, when individual as well as collective preferences over outcomes satisfy the SSB axioms. Additionally, we consider *anonymity*, a fairness property that is stronger than non-dictatorship and requires that an SWF is invariant under renaming the agents. Arrovian SWFs have been studied by a number of authors when the agents as well as the society possess vNM preferences. Kalai and Schmeidler (1977b) showed that Arrow's impossibility remains valid when there are at least four alternatives and the SWF is continuous. Hylland (1980a) found that this theorem also holds without assuming continuity of the SWF. Various authors have shown variants of this result that differ in whether the input is a preference relation or a utility function and in the exact extension of Arrow's axioms to vNM preferences (Sen, 1970b; Schwartz, 1970; Le Breton, 1986; Mongin, 1994; Dhillon and Mertens, 1997). In contrast, we show that when enlarging the set of feasible preferences from vNM preferences to SSB preferences, there is a unique largest domain of individual preferences for which an anonymous Arrovian SWF exists. In particular, Arrow's impossibility theorem does not hold on this domain. It contains exactly those preference relations that are based on *pairwise comparisons*. One outcome is preferred to another outcome according to pairwise comparisons if the likelihood that the former returns a more preferred alternative is larger than the likelihood that the latter returns a more preferred alternative. This corresponds to the case when all entries in the representing SSB matrix are 1, 0, or -1. Hence, the collective preferences returned by anonymous Arrovian SWFs are completely determined by the agents' preferences over pure outcomes. In particular, this prohibits agents to express different intensities of preference between pairs of pure outcomes. Experiments support that preferences based on pairwise comparisons are indeed exhibited by human decision makers (see, e.g., Butler et al., 2016). Blavatskyy (2006) characterizes preferences based on pairwise comparisons using the fanning-in axiom in addition to Fishburn's SSB axioms.

We go on to show that every Arrovian SWF on the domain of preferences based on pairwise comparisons is affine utilitarian, meaning that the SSB function representing the collective preferences is derived as a weighted sum of the SSB functions representing the agents' preferences. When again requiring anonymity, this characterizes *relative utilitarianism*, i.e., the affine utilitarian SWF with the same positive weight assigned to all agents. Affine utilitarianism is well-studied for the case of vNM preferences over outcomes. Harsanyi (1955) has famously shown that every SWF that maps a profile of vNM preferences to vNM preferences and satisfies Pareto indifference is affine utilitarian. Pareto indifference prescribes that two outcomes that are considered equally good by all agents have to be considered equally good according to the collective preference relation. Harsanyi's statement is single-profile in that the weights of the agents may vary across different preference profiles. Fishburn and Gehrlein (1987) and Turunen-Red and Weymark (1999) have shown that Harsanyi's theorem does not hold when the class of feasible preference relations is enlarged to SSB preferences, even under stronger Pareto-type assumptions. This shows that aggregating SSB preferences is fundamentally different from aggregating vNM preferences. It is therefore remarkable that when additionally assuming anonymity and independence of irrelevant alternatives, one can retrieve affine utilitarianism.

1.3.2 Relative Utilitarian Social Choice

To a certain extent SWFs are more of theoretical than of practical relevance. For most applications of social choice theory, the objective is to choose a collectively most preferred outcome rather than obtaining a collective preference relation. The framework for choosing outcomes should be flexible enough to accommodate for situations where not all outcomes that the agents have preferences over are feasible, e.g., when a group of people decides where to go for lunch, some restaurants may be closed because it is their rest day. Mathematically this is formulated via *social choice functions (SCFs)*, which map a preference profile and a feasible set of outcomes to a set of chosen outcomes. We will assume that the feasibility of outcomes is based on the availability of alternatives and therefore only consider feasible sets which contain all probability distributions over some finite set of alternatives. Note that every SWF induces an SCF which, for every feasible set, chooses the set of maximal elements of the collective preference relation restricted to that set. Since the collective preferences are assumed to satisfy continuity and convexity, the thereby defined SCF always returns at least one outcome and satisfies standard choice consistency conditions due to Sen (1969) and Sen (1971) known as Sen's α and Sen's γ. The SCF that returns the maximal elements of relative utilitarianism on the domain of preferences based on pairwise comparisons is known as *maximal lotteries* (Kreweras, 1965; Fishburn, 1984b). It chooses the outcomes that are preferred to all other outcomes by an expected majority of agents. Hence, maximal lotteries can be seen as an extension of Condorcet's method of choosing Condorcet winners whenever they exist. The classical model of social choice, which requires SCFs to choose from a set of deterministic alternatives, can be embedded in ours by considering SCFs that return *all* outcomes that randomize over some subset of the alternatives. This subset of alternatives is interpreted as the set of chosen alternatives. These SCFs will be called *pure*.

An important property of SCFs, especially in the context of voting, is that the agents have an incentive to submit their preferences and thus do not abstain from the aggregation process. However, for some SCFs an agent can obtain a more preferred outcome by abstaining, which is called the *no-show paradox* (Fishburn and Brams, 1983). Moulin (1988) showed that every Condorcet consistent pure SCF suffers from the no-show paradox. Following his terminology, an SCF that does not suffer from the no-show paradox entices *participation*. We show that when the agents have SSB preferences, the SCF induced by relative utilitarianism entices participation. This SCF even satisfies a very strong notion of participation, called utilitarian participation, which prescribes that no group of agents can abstain from the aggregation process and thereby obtain an outcome that yields higher accumulated utility for the abstaining group of agents. According to this notion, some agents may even be worse off by abstaining if their loss in utility is compensated by other abstaining agents' gains. An SCF is homogeneous if adding the same number of clones (agents with the same preferences) of every agent to the electorate does not change the social choice and weakly utilitarian if it chooses pure maximal outcomes according to relative utilitarianism whenever they exist. We go on to show that every SCF that is homogeneous and weakly utilitarian has to choose maximal elements according to relative utilitarianism. This result requires that the domain of preferences is sufficiently rich, i.e., it has to be closed under reversal of preferences and it has to be possible for agents to have an arbitrary pure outcome as their most preferred outcome.[5]

These results also apply to the domain of preferences based on pairwise comparisons. In this case, relative utilitarianism induces maximal lotteries and hence, maximal lotteries entices utilitarian participation. Since maximal lotteries is Condorcet consistent, it can be seen as a possible resolution of Moulin's no-show paradox. Additionally, our second result implies that every homogeneous and Condorcet consistent SCF that entices utilitarian participation has to choose maximal lotteries.

1.3.3 Consistent Social Choice

When considering choice consistency for single agents, it is typically defined as consistency with respect to variable feasible sets. In the social choice context, one can also consider choice consistency with respect to variable electorates, i.e., variable sets of agents. For example, if there are two disjoint electorates, each of which chooses the same outcome, then the union of both electorates should also choose this outcome. This condition is called *population consistency*. Alike conditions were considered by Smith (1973), Young (1974b), and Fine and Fine (1974). Reinforcement, a strengthening of population consistency, is known to be the characterizing feature of scoring rules (Young, 1975).

The second type of consistency considered here was introduced by Tideman (1987), who gives the following illustrative example. "When I was 12 years old I was nominated to be treasurer of my class at school. A girl named Michelle was also nominated. I relished the prospect of being treasurer, so I made a quick calculation and nominated Michelle's best friend, Charlotte. In the ensuing election I received 13 votes, Michelle received 12, and Charlotte received 11, so I became treasurer." If one assumes for simplicity that every student either preferred both girls to become treasurer over Tideman or preferred Tideman to become treasurer over both girls, then by introducing Charlotte as a third option, Tideman altered the result in this favor, since Michelle would have received 23 votes in a head-to-head comparison. In this example, the two girls can be thought of as variants of one alternative. It is likely that they have the same relation to all other alternatives according to the agents' preference relations. An SCF should arguably take the structure of a preference profile introduced by variants of one alternative into account. A set of alternatives is a *component* in a preference profile if it is an interval in every agent's preference relation over alternatives. Alternatives within a component can be thought of as of variants, called *clones*, of an arbitrary representative from the component. Taking components

5 The latter condition is called minimal richness by Puppe (2016).

into account, the choice from the entire feasible set can be decomposed into two choices. First, all clones of the representative are disregarded and a choice is made from the remaining set of alternatives. *Cloning consistency* prescribes that the probabilities assigned to alternatives outside of the component should then be the same as when the choice is made from the entire feasible set. In particular, their probability is not influenced by the presence of clones or the agents' preferences over the clones. Cloning consistency however neglects the information contained in the preferences over clones. Thus, a second choice is made from the component disregarding all other alternatives. In addition to cloning consistency, *composition consistency* prescribes that the probability assigned to alternatives inside the component should be directly proportional to their probabilities when the choice is made from the entire feasible set. Composition consistency was introduced by Laffond et al. (1996) for pure SCFs. Our definition of composition consistency boils down to theirs for pure SCFs.

We consider consistent SCFs on the domain of preferences based on pairwise comparisons. Our first result shows that population consistency is incompatible even with cloning consistency for pure SCFs. When allowing for non-pure SCFs, population consistency and composition consistency uniquely characterize maximal lotteries. Both results operate under the assumption that SCFs additionally satisfy standard normative properties, which might well be part of the definition of SCFs, and mild regularity conditions.

The results detailed above illustrate that the domain of preferences based on pairwise comparisons allows to combine properties that are incompatible on the unrestricted domain. In combination with anonymity, it turns Arrow's impossibility theorem into a characterization of relative utilitarianism. Moreover, maximal lotteries, the SCF that chooses the maximal elements of relative utilitarianism on the domain of preferences based on pairwise comparisons, satisfies Condorcet consistency, participation, population consistency, and composition consistency. This is in contrast to a number of results that have shown the incompatibility of these properties on the unrestricted domain.

ORGANIZATION OF THIS THESIS

We start by defining notation and decision theoretic concepts that are used throughout this thesis in Chapter 2. The remainder is divided into two parts. The first part studies maximin strategies in zero-sum games. After formally introducing zero-sum games in Chapter 3, we derive a proof of the minimax theorem in Chapter 4. We proceed in Chapter 5 with a set of behavioral assumptions about players that lead to maximin play. This part concludes by studying randomly chosen symmetric zero-sum games in Chapter 6.

The second part examines social choice problems on domains of SSB preferences over outcomes. The necessary social choice specific concepts are introduced in Chapter 7. In Chapter 8, we establish that anonymous Arrovian aggregation necessitates preferences based on pairwise comparisons and relative utilitarianism. In the remaining two chapters, we study the social choice function that chooses maximal elements according to relative utilitarianism. Its resistance to strategic abstention is examined in Chapter 9. Chapter 10 concludes with a characterization of maximal lotteries using population consistency and composition consistency.

The two parts are independent in that each of them can be understood separately.

UNDERLYING ARTICLES

This thesis is based on the publications and working papers listed on page vii. Chapter 5 is based on [3], Chapter 6 is based on [1], Chapter 8 is based on [2], Chapter 9 is based on [4], and Chapter 10 is based on [5].

EXCLUDED ARTICLES

Apart from the articles listed on page vii, my work contributed to the following publications and working papers. Their results have been excluded, since they did not fit the theme of this thesis.

[1] On the tradeoff between efficiency and strategyproofness. *Games and Economic Behavior*, 110:1–18, 2018 (with H. Aziz, F. Brandt, and M. Brill).

[2] Fractional hedonic games. 2018 (with H. Aziz, F. Brandt, P. Harrenstein, M. Olsen, and D. Peters), Working paper.

[3] Proving the incompatibility of efficiency and strategyproofness via SMT solving. *Journal of the ACM*, 65(2), 2018 (with F. Brandt, M. Eberl, and C. Geist).

[4] Popular matchings with multiple partners. In *Proceedings of the 37th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, Leibniz International Proceedings in Informatics (LIPIcs), 19:1–19:15. LZI, 2018 (with T. Kavitha).

[5] Random assignment with optional participation. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 326–334. IFAAMAS, 2017 (with F. Brandt and J. Hofbauer).

[6] An axiomatic characterization of the Borda mean rule. 2017 (with D. Peters), Working paper.

[7] Proving the incompatibility of efficiency and strategyproofness via SMT solving. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, 116–122. AAAI Press, 2016 (with F. Brandt and C. Geist).

[8] The impossibility of extending random dictatorship to weak preferences. *Economics Letters*, 141:44–47, 2016 (with F. Brandt and W. Suksompong).

[9] Universal Pareto dominance and welfare for plausible utility functions. *Journal of Mathematical Economics*, 60:123–133, 2015 (with H. Aziz and F. Brandt).

[10] Strategic abstention based on preference extensions: Positive results and computer-generated impossibilities. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, 18–24. AAAI Press, 2015 (with F. Brandt, C. Geist, and J. Hofbauer).

[11] Incentives for participation and abstention in probabilistic social choice. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 1411–1419. IFAAMAS, 2015 (with F. Brandt and J. Hofbauer).

[12] Fractional hedonic games: Individual and group stability. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 1219–1227. IFAAMAS, 2015 (with F. Brandt and M. Strobel).

[13] On the incompatibility of efficiency and strategyproofness in randomized social choice. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*, 545–551. AAAI Press, 2014 (with H. Aziz and F. Brandt).

2 PRELIMINARIES

We start by introducing some notation that will be used in various different contexts throughout this thesis. For $n \in \mathbb{N}$, we define $[n] = \{1, \dots, n\}$ as the set of all natural numbers from 1 to n. Given some set U, $\mathbb{R}^U$ denotes the set of all real-valued sequences indexed by U that are eventually 0. By $\mathcal{F}(U)$ we denote the set of all finite and non-empty subsets of U. For a vector $x \in \mathbb{R}^U$, $x_+ = \{i \in U \colon x_i > 0\}$ and $x_- = \{i \in U \colon x_i < 0\}$ denote the sets of elements of U for which x is positive and negative, respectively. The support $\operatorname{supp}(x) = x_+ \cup x_-$ is the set of all elements of U corresponding to non-zero entries of x. For $A \subseteq U$, $x_A = (x_i)_{i \in A}$ is the restriction of x to A. Similarly, for $M \in \mathbb{R}^{U \times U}$ and $A, B \subseteq U$, $M_{AB} = (M_{ij})_{i \in A, j \in B}$ is the restriction of M to the submatrix induced by A and B. We write $M_A = M_{AA}$ for short.

For $X \subseteq \mathbb{R}^U$, the *convex hull* $\operatorname{conv}(X)$ is the set of all convex combinations of elements of X, i.e.,

$$\operatorname{conv}(X) = \left\{ \lambda_1 x^1 + \dots + \lambda_k x^k \colon x^i \in X, \lambda \in \mathbb{R}^k_{\geqslant 0}, \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

If $X = \operatorname{conv}(X)$, X is *convex*. The *affine hull* $\operatorname{aff}(X)$ is the set of all affine combinations of elements of X, i.e.,

$$\operatorname{aff}(X) = \left\{ \lambda_1 x^1 + \dots + \lambda_k x^k \colon x^i \in X, \lambda \in \mathbb{R}^k, \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

X is an *affine subspace* if $X = \operatorname{aff}(X)$. We say that $x^1, \dots, x^k \in \mathbb{R}^U$ are *affinely independent* if, for all $\lambda \in \mathbb{R}^k$ with $\sum_{i=1}^k \lambda_i = 0$, $\sum_{i=1}^k \lambda_i x^i = 0$ implies $\lambda = 0$. The *dimension* $\dim(X)$ of an affine subspace X is $k - 1$, where k is the maximal number of affinely independent vectors in X. If this number is unbounded, X has infinite dimension. The dimension of a set $X \subseteq \mathbb{R}^U$ is defined as the dimension of $\operatorname{aff}(X)$. The *linear hull* $\operatorname{lin}(X)$ is the set of all linear combinations of elements of X, i.e.,

$$\operatorname{lin}(X) = \left\{ \lambda_1 x^1 + \dots + \lambda_k x^k \colon x^i \in X, \lambda \in \mathbb{R}^k \right\}.$$

The 1-*norm* of $x \in \mathbb{R}^U$ is denoted by $\|x\|$, i.e., $\|x\| = \sum_{i \in U} |x_i|$. By $B_\epsilon(x) = \{y \in \mathbb{R}^U \colon \|x - y\| < \epsilon\}$ we denote the open ϵ-*ball* around $x \in \mathbb{R}^U$. For $X, Y \subseteq \mathbb{R}^U$, the *interior* of X in Y is $\operatorname{int}_Y(X) = \{x \in X \colon B_\epsilon(x) \cap Y \subseteq X$ for some $\epsilon > 0\}$. We say that X is open in Y if $X = \operatorname{int}_Y(X)$. Similarly, X is closed in Y if the complement of X is open in Y. The *closure* of X in Y, $\operatorname{cl}_Y(X)$, is the intersection of all sets that are closed in Y and contain X. X is *dense* in Y if $\mathrm{cl}_Y(X) = Y$. Alternatively, X is dense at $y \in \mathbb{R}^U$ if, for every $\epsilon > 0$, there is $x \in X$ such that $\|x - y\| < \epsilon$. X is dense in Y if X is dense at y for every $y \in Y$.

For $A, B \subseteq U$, $\Pi_U(A, B)$ denotes the set of all permutations on U that map A to B. If $A = B$, we write $\Pi_U(A)$ for short. Whenever U is clear from the context, the subscript will be omitted. For $x \in \mathbb{R}^U$ and $\pi \in \Pi(U)$, 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 $i \in U$. Similarly, for $M \in \mathbb{R}^{U \times U}$ and $\pi, \sigma \in \Pi(U)$, $M_{\pi\sigma} = M \circ (\pi^{-1} \times \sigma^{-1})$ and M_π is short for $M_{\pi\pi}$.

For $A \subseteq U$, $\Delta(A)$ denotes the set of all probability distributions on U whose support is finite and contained in A. Sometimes the set of rational-valued probability distributions $\Delta^{\mathbb{Q}}(A) = \Delta(A) \cap \mathbb{Q}^U$ is used. For $i \in U$, $\underline{i}$ denotes the one-point measure on i, i.e., $\underline{i}$ assigns probability 1 to i. For $A \in \mathcal{F}(U)$, $\mathrm{uni}(A) \in \Delta(U)$ denotes the probability distribution that distributes probability uniformly over A. Let $A, B \subseteq U$ such that $A \cap B = \{j\}$. For $p \in \Delta(A)$, $q \in \Delta(B)$, and $i \in U$, let

$$(p \times_j q)_i = \begin{cases} p_i & \text{if } i \in U \setminus B, \\ p_j q_i & \text{if } i \in B. \end{cases}$$

The operations defined here for a given type of object extend to sets of objects of the same type by applying them to every element in the set.

2.1 DECISION THEORETIC FUNDAMENTALS

A preference relation $\succ$ for a decision maker (or agent) on a set of outcomes $\Delta(U)$ is an asymmetric relation on $\Delta(U)$.[6] For $p, q \in \Delta(U)$, we will write $p \sim q$ if neither $p \succ q$ nor $q \succ p$ and $p \succsim q$ if either $p \succ q$ or $p \sim q$. We say that p is preferred to, indifferent to, and weakly preferred to q if $p \succ q$, $p \sim q$, and $p \succsim q$, respectively. Hence, the weak preference relation $\succsim$ is complete. For $p \in \Delta(U)$, the lower contour set $L(p) = \{q \in \Delta(U) \colon p \succ q\}$ and the upper contour set $U(p) = \{q \in \Delta(U) \colon q \succ p\}$ of p collect all outcomes that are less preferred and more preferred than p. The indifference set $I(p) = \{q \in \Delta(U) \colon p \sim q\}$ contains all outcomes that p is indifferent to. For $\pi \in \Pi_U$, we define the permutation of $\succ$ with respect to π, $\succ^\pi$, such that $p_\pi \succ^\pi q_\pi$ if and only if $p \succ q$ for all $p, q \in \Delta(U)$. The restriction of $\succ$ to a set of outcomes $X \subseteq \Delta(U)$ is $\succ|_X = \succ \cap (X \times X)$.

Typically, preference relations are assumed to satisfy some notion of rationality. This is captured by restricting the set of feasible pref-

6 A relation $\succ$ is asymmetric if, for all $p, q \in \Delta(U)$, $p \succ q$ implies not $q \succ p$.

erence relations. For an abstract set of outcomes, any two preference relations that only differ by renaming the outcomes are equally reasonable. When considering preference relations on $\Delta(U)$, it seems desirable to take into account the structure of $\Delta(U)$. A well-known example of restricted preferences over $\Delta(U)$ is linear expected utility theory due to von Neumann and Morgenstern (1953), which is based on three axioms called transitivity, independence, and continuity. Transitivity requires that $\succ$ is a weak order on $\Delta(U)$, i.e.,

$$\succ \text{ and } \sim \text{ are transitive.} \qquad \text{(transitivity)}$$

Independence prescribes that the preference between two outcomes does not change if they are both shifted in the same direction and by the same magnitude within $\Delta(U)$. Formally, $\succ$ satisfies independence if, for all $p, q, r \in \Delta(U)$ and $\lambda \in (0, 1)$,

$$p \succ q \text{ if and only if } \lambda p + (1-\lambda)r \succ \lambda q + (1-\lambda)r. \quad \text{(independence)}$$

Lastly, a preference relation is continuous if it prohibits preference reversals under small perturbations of outcomes. A preference relation $\succ$ is continuous if, for all $p, q, r \in \Delta(U)$,

$$p \succ q \succ r \text{ implies } \lambda p + (1-\lambda)r \sim q \text{ for some } \lambda \in (0, 1). \quad \text{(continuity)}$$

Transitivity, independence, and continuity are called vNM axioms in the sequel; a vNM preference relation is a preference relation satisfying the vNM axioms. It can be shown that a preference relation $\succ$ satisfies the vNM axioms if and only if there exists a linear function $u \colon \Delta(U) \to \mathbb{R}$ such that, for all $p, q \in \Delta(U)$, $p \succ q$ exactly when $u(p) > u(q)$ (see, e.g., Fishburn, 1988). For an extensive discussion of vNM utility theory, we refer to Karni (2014).

Another standard assumption is that preferences are convex. We will use convexity as defined by Fishburn (1982).[7] A preference relation $\succ$ is convex if, for all $p, q, r \in \Delta(U)$ and $\lambda \in (0, 1)$,

$$\begin{aligned} &p \succ q \text{ and } p \succsim r \text{ imply } p \succ \lambda q + (1-\lambda)r, \\ &q \succ p \text{ and } r \succsim p \text{ imply } \lambda q + (1-\lambda)r \succ p, \text{ and} \qquad \text{(convexity)} \\ &p \sim q \text{ and } p \sim r \text{ imply } p \sim \lambda q + (1-\lambda)r. \end{aligned}$$

Equivalently, one could require that the indifference set of an outcome p is the intersection of $\Delta(U)$ with a hyperplane through p; the upper and lower contour sets are the intersection of $\Delta(U)$ with the corresponding half spaces. Note that convexity implies that upper contour sets, lower contour sets, and indifference sets are convex. Moreover, upper contour sets and lower contour sets are open and indifference sets are closed. Proofs of these statements appear in

7 This notion of convexity is called dominance by Fishburn (1982).

Lemmas 8.4, 8.6, and 8.7. This notion of convexity is rather strong. Weaker notions of convexity only require that upper contour sets are convex or that lower contour sets are convex or both.

Transitivity of preference relations is frequently used to guarantee the existence of maximal elements. The following theorem by Sonnenschein (1971) shows that this assumption is to some extent made out of technical convenience rather than necessity, since continuity and convexity already suffice to assure that maximal elements exist within compact and convex sets.[8]

Proposition 2.1 *(Sonnenschein, 1971)*
Let $\succ$ is a continuous and convex preference relation. Then, $\max_{\succ} X \neq \emptyset$ for every non-empty, compact, and convex set $X \subseteq \Delta(U)$.

Arguably the prime reason for requiring the existence of maximal elements is that a decision maker should be able to choose an outcome from some feasible set of outcomes. If a feasible set does not admit a maximal element, than the decision maker will not be able to make a satisfactory choice, since whichever outcome he chooses, there will always be a preferred outcome. For a finite abstract set of outcomes, there is no basis for assuming that some sets should not be feasible. Sen (1969) and Sen (1971) has shown that in this setting, two intuitive choice consistency conditions are equivalent to choosing maximal elements according to an acyclic relation. These conditions are known as *Sen's* α (or *contraction*) and *Sen's* γ (or *expansion*).[9] Contraction requires that if an outcome is chosen from some set, then it should also be chosen from any subset thereof that it is contained in. This condition is satisfied when choosing maximal elements without imposing any restrictions on $\succ$. Expansion prescribes that an outcome that is chosen from two sets X and Y should also be chosen from their union $X \cup Y$. Since we are only interested in choosing from convex sets, we strengthen this condition by taking the convex hull $\mathrm{conv}(X \cup Y)$ in the consequence. If $\succ$ is convex, then $\max_{\succ}$ satisfies this notion of expansion. To see this, consider $X, Y \subseteq \Delta(U)$ and assume that $p \in \max_{\succ} X \cap \max_{\succ} Y$. Then, $p \succsim q$ for all $q \in X \cup Y$ and, since $\succ$ satisfies convexity, we have $p \succsim q$ for all $q \in \mathrm{conv}(X \cup Y)$. Thus, $p \in \max_{\succ} \mathrm{conv}(X \cup Y)$. Sen's proof can even be adapted to show that *every* choice function satisfying contraction and expansion is of the form $\max_{\succ}$ for some $\succ$ with convex weak lower contour sets. The preference relation $\succ$ on $\Delta(\{a, b\})$ that has $p \succ q$ if $p_a > q_a$ for

8 Sonnenschein only required that upper contour sets are convex and that lower contour sets are open. The latter assumption is weaker than continuity when the set of alternatives is finite. See Bergstrom (1992) and Llinares (1998) for a discussion of Sonnenschein's and related results.

9 Sen's α can be traced back to Chernoff (1954) and Nash (1950b), where it is called *independence of irrelevant alternatives* (not to be confused with Arrow's IIA). We refer to Monjardet (2008) for more details.

all $p, q \in \Delta(\{a, b\})$ except that $\underline{a} \sim \underline{b}$ has convex weak lower (and weak upper) contour sets. The choice function that chooses maximal elements of compact and convex subsets of $\Delta(\{a, b\})$ according $\succ$ satisfies contraction and expansion. However, $\succ$ does not satisfy convexity, since $I(\underline{a}) = I(\underline{b}) = \{\underline{a}, \underline{b}\}$ is not convex. Hence, not every choice behavior that satisfies contraction and expansion can be expressed by choosing maximal elements according to a convex preference relation.

2.2 SSB UTILITY THEORY

When requiring preference relations to satisfy a certain set of axioms, it is desirable to have a mathematically compact way of representing preference relations that satisfy these axioms. Fishburn (1988, p. 85) showed that a preference relation $\succ$ on $\Delta(U)$ is continuous and convex if and only if it can be represented by a non-transitive convex utility function, i.e., a function $\phi\colon \Delta(U) \times \Delta(U) \to \mathbb{R}$ that is sign skew-symmetric and linear in the first argument.[10] A function $\phi\colon \Delta(U) \times \Delta(U) \to \mathbb{R}$ represents a preference relation $\succ$ whenever $p \succ q$ if and only if $\phi(p, q) > 0$ for all $p, q \in \Delta(U)$. Observe that this does not rule out the possibility that $|\phi(p, q)| > |\phi(q, p)|$ for two outcomes p and q. Informally this means that the magnitude of preference between p and q depends on the order in which they are compared. To prevent this, Fishburn (1982) additionally requires preference relations to satisfy the following symmetry axiom. A preference relation $\succ$ satisfies symmetry if, for all $p, q, r \in \Delta(U)$ and $\lambda \in (0, 1)$,

$$\text{if } p \succ q \succ r,\ p \succ r, \text{ and } q \sim 1/2\, p + 1/2\, r, \text{ then}$$
$$[\lambda p + (1-\lambda) r \sim 1/2\, p + 1/2\, q \text{ if and only if } \lambda r + (1-\lambda) p \sim 1/2\, r + 1/2\, q]. \tag{Symmetry}$$

The implications of the symmetry axiom can be expressed as follows. Continuity and convexity imply that, for every triple of outcomes, the indifference curves within their convex hull are straight lines. Symmetry prescribes that, either all these indifference curves are parallel or intersect in one point (which may be outside of their convex hull). In the first case, the preferences over their convex hull can be represented by a linear utility function.

Fishburn (1984c) himself states "I am a bit uncertain as to whether this should be regarded more as a convention than a testable hypothesis – much like the asymmetry axiom [...], which can almost be thought of as a definitional characteristic of strict preference." We will frequently consider preference relations that satisfy continuity,

10 A function $\phi\colon \Delta(U) \times \Delta(U) \to \mathbb{R}$ is sign skew-symmetric if, for all $p, q \in \Delta(U)$, $\phi(p, q) > 0$ if and only if $\phi(q, p) < 0$.

convexity, and symmetry. The set of all such preference relations on $\Delta(U)$ is henceforth denoted by $\mathcal{R}$.

The addition of symmetry implies that every preference relation in $\mathcal{R}$ can be represented by a skew-symmetric and bilinear (SSB) function $\phi\colon \Delta(U) \times \Delta(U) \to \mathbb{R}$ (Fishburn, 1982).[11] Moreover, ϕ is unique up to multiplication by a positive scalar. Hence, any two SSB functions that only differ by multiplication with a positive scalar represent the same preference relation. Every preference relation $\succ \in \mathcal{R}$ other than complete indifference can thus be associated with a unique normalized SSB function on $\Delta(U) \times \Delta(U)$ whose largest positive value is equal to 1. By Φ we denote the set of all SSB functions that are normalized in this way. For $\succ \in \mathcal{R}$, $\phi^{\succ} \in \Phi$ denotes the normalized SSB function representing $\succ$. Note that the value of an SSB function ϕ is maximized for a pair of vertices of $\Delta(U)$, since it is linear in both arguments. For two SSB functions ϕ and $\hat{\phi}$ we write $\phi \equiv \hat{\phi}$ if $\phi = \alpha\hat{\phi}$ for some $\alpha > 0$. Since the set of outcomes $\Delta(U)$ only contains outcomes with finite support, $\phi(p, q)$ can be written as a convex combination of the values of ϕ for pure outcomes (Fishburn, 1984c). Thus, for all $p, q \in \Delta(U)$,

$$\phi(p, q) = \sum_{a,b \in U} p_a q_b \phi(\underline{a}, \underline{b}).$$

Consequently, every SSB function ϕ can be represented by a skew-symmetric matrix $M \in \mathcal{M}$ and *vice versa*, where $\phi(\underline{a}, \underline{b}) = M_{ab}$ for all $a, b \in U$. Then, we have that $\phi(p, q) = p^t M q$.

The conjunction of transitivity and the independence axiom implies both convexity and symmetry. Remarkably, the independence axiom in addition to continuity and convexity is enough to guarantee that a preference relation can be represented by a vNM utility function (Fishburn, 1982, Proposition 1). Hence, in the presence of continuity and convexity, the independence axiom implies transitivity. In this case ϕ is additively separable, i.e., $\phi(p, q) = u(p) - u(q)$ for all $p, q \in \Delta(U)$ for some linear utility function u representing $\succ$. On the other hand, transitive relations in $\mathcal{R}$ are exactly those which can be represented by a weighted linear (WL) utility function as introduced by Chew (1983).[12] For independently distributed outcomes (as considered in this thesis), SSB utility theory coincides with regret theory as introduced by Loomes and Sugden (1982) (see also Loomes and Sugden, 1987; Blavatskyy, 2006).

Through the representation of $\succ \in \mathcal{R}$ by a skew-symmetric matrix, it becomes apparent that the minimax theorem implies the exis-

11 A function $\phi\colon \Delta(U) \times \Delta(U) \to \mathbb{R}$ is skew-symmetric if $\phi(p, q) = -\phi(q, p)$ for all $p, q \in \Delta(U)$. ϕ is bilinear if it is linear in both arguments.

12 A WL function is characterized by a linear utility function and a non-negative weight function. The utility of an outcome is the utility derived from the linear utility function weighted according to the weight function. Thus, WL functions are more general than linear utility functions, as every linear utility function is equivalent to a WL function with constant weight function. See also Fishburn (1983).

$$\phi = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{ccc} a & b & c \\ \end{array}$$
$$\phi = \begin{array}{cc} & \begin{array}{ccc} a & \;b & \;c \end{array} \\ & \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix} \begin{array}{c} a \\ b \\ c \end{array} \end{array}$$

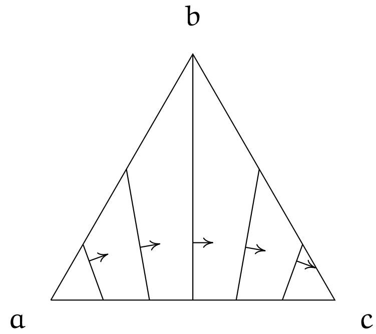


Figure 2.1: Illustration of PC preferences when preferences on pure outcomes are given by the transitive relation $\underline{a} \succ \underline{b} \succ \underline{c}$. The arrows represent the normal vectors to the indifference curves of the outcome at the base of the arrow (pointing towards the lower contour set). Each indifference curve separates the corresponding upper and lower contour set.

tence of maximal elements of $\succ$ on $\Delta(A)$ for every $A \in \mathcal{F}(U)$. This was noted by Fishburn (1984c, Theorem 4) and already follows from Proposition 2.1. Fishburn (1984c) goes on to show that choosing maximal elements of $\succ$ from feasible sets satisfies contraction and expansion. As discussed before, this even holds for arbitrary convex relations.[13]

We will be frequently interested in a particular subclass of SSB functions. An SSB function $\phi \in \Phi$ is based on pairwise comparisons if $\phi(\underline{a}, \underline{b}) \in \{-1, 0, 1\}$ for all $a, b \in U$. The set of all SSB functions based on pairwise comparisons is denoted by Φ^{PC}. A preference relation $\succ \in \mathcal{R}$ is based on pairwise comparisons if its SSB representation is based on pairwise comparisons. Preference relations based on pairwise comparisons with be called PC preferences for short. The set $\mathcal{D}^{PC}$ collects all such preference relations. From the SSB representation it can be seen that, for all $\succ \in \mathcal{D}^{PC}$ and $p, q \in \Delta(U)$,

$$p \succ q \quad \text{if and only if} \quad \sum_{\underline{a} \succ \underline{b}} p_a q_b > \sum_{\underline{b} \succ \underline{a}} p_b q_a.$$

The sum on the left hand side of the inequality is the probability that p yields a better alternative than q. Analogously, the sum on the right hand side is the probability that q yields a better alternative than p. Hence, p is preferred to q by pairwise comparison if the probability that p yields a better alternative is larger than for q. Note that $\succ$ is completely determined by the preferences over pure outcomes, which establishes a one-to-one correspondence between ϕ^{PC} and the set of asymmetric relations on U.

PC preferences appear in earlier as well as more recent literature on decision theory (Blyth, 1972; Packard, 1982; Blavatskyy, 2006).[14] In

13 Fishburn (1984c) defines expansion without taking the convex closure of the union of two feasible sets, which results in a weaker notion of expansion.

14 PC preferences are referred to the rule of expected dominance by Packard (1982) and preference for the most probable winner by Blavatskyy (2006).

$$\phi = \begin{array}{c c} & \begin{array}{cccc} a & b & c & d \end{array} \\ & \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix} \begin{array}{c} a \\ b \\ c \\ d \end{array} \end{array} \qquad \begin{aligned} p &= \underline{c} \\ q &= {}^{2}/_{5}\,\underline{a} + {}^{3}/_{5}\,\underline{d} \\ r &= {}^{3}/_{5}\,\underline{b} + {}^{2}/_{5}\,\underline{d} \end{aligned}$$

Figure 2.2: Illustration of PC preferences when preferences over pure outcomes are given by the transitive relation $\underline{a} \succ \underline{b} \succ \underline{c} \succ \underline{d}$. The left-hand side shows the corresponding SSB function. The preferences between the three outcomes p, q, and r defined in the table on the right-hand side are cyclic: $\phi(p, q) = 3/5 - 2/5 = 1/5 > 0$, $\phi(q, r) = 4/25 > 0$, and $\phi(r, p) = 1/5 > 0$. Hence, $p \succ q \succ r \succ p$.

the social choice literature, Pareto efficiency, strategyproofness, and participation of social choice functions with respect to these preferences were studied (Aziz et al., 2015; Aziz et al., 2018; Brandl et al., 2018).

Figure 2.1 illustrates PC preferences for three transitively ordered alternatives.[15] Blavatskyy (2006) gives an axiomatic characterization using Fishburn's SSB axioms and an additional axiom called *fanning-in*, which essentially prescribes that indifference curves fan in at a certain rate (see Figure 2.1). As a corollary of Theorem 8.1, fanning-in is implied by Fishburn's SSB axioms and Arrow's axioms. Blavatskyy cites extensive experimental evidence for the fanning-in of indifference curves. Recent evidence for preferences based on pairwise comparison has been provided by Butler et al. (2016).

For at least four alternatives, PC preferences can be cyclic even when the preferences over pure outcomes are transitive. This phenomenon, known as the *Steinhaus-Trybula paradox*, is illustrated in Figure 2.2 (see, e.g., Steinhaus and Trybula, 1959; Blyth, 1972; Packard, 1982; Rubinstein and Segal, 2012; Butler et al., 2016). Butler et al. (2016) have conducted an extensive experimental study of the Steinhaus-Trybula paradox and found significant evidence for PC preferences.

15 For three alternatives PC preferences as depicted in Figure 2.1 can be represented by a WL function with utility function $u(a) = u(b) = 1$ and $u(c) = 0$ and weight function $w(a) = 0$ and $w(b) = w(c) = 1$.

Part I

ZERO-SUM GAMES

3 GAME THEORETIC FUNDAMENTALS

3.1 ZERO-SUM GAMES

Studying strategic interaction requires a model that allows to specify the rules of the underlying situation. The representation of such a situation, called a game, needs to be rich enough to allow us to express a sufficient variety of rules and to capture the behavior of the players. On the other hand, a concise representation typically simplifies the analysis. For the purpose of this thesis, it is sufficient to consider two-player zero-sum games represented by a single matrix. However, our model enriches the usual formulation in that it allows to relate games on different sets of actions to each other. To this end, let U be the set of all actions that a player could conceivably take. The strategy of a player consists of choosing a randomization over actions, i.e., an element of $\Delta(U)$. For $i \in U$, $\underline{i}$ is called a pure strategy. A *proto game* $M \in \mathbb{R}^{U \times U}$ specifies the payoff of the row player for every combination of pure strategies, i.e., for $i, j \in U$, M_{ij} is the payoff of the row player if the row player plays $\underline{i}$ and the column player plays $\underline{j}$. The payoff of the column player is the negative of the payoff of the row player. However, not all actions are feasible in every situation. A two-player zero-sum game M_{AB} is obtained by restricting M to sets of feasible actions $A, B \in \mathcal{F}(U)$. Two-player zero-sum games will be simply referred to as *games*. For two strategies $p, q \in \Delta(U)$, one for the row player and one for the column player, the expected payoff for the row player is $p^t M q$. The expected payoff of the column player is the negative thereof.

The objective of game theory is to provide a formal basis for decision making in the presence of strategic interaction. This is typically formalized via *solution concepts*, which in our framework map a proto game and a pair of feasible sets of actions to a non-empty set of strategies for the row player. Formally, a solution concept is a function $f \colon \mathbb{R}^{U \times U} \times \mathcal{F}(U)^2 \to 2^{\Delta(U)} \setminus \{\emptyset\}$ such that $f(M, A, B) \subseteq \Delta(A)$ for all $M \in \mathbb{R}^{U \times U}$ and $A, B \in \mathcal{F}(U)$. The last part of the definition ensures that infeasible actions are assigned probability zero. Note that $f(-M^t, B, A)$ is the set of strategies recommended for the column player. This definition of solution concept enables us to relate strategic choices for different sets of feasible actions to each other. A widely accepted solution concept are maximin strategies (*MS*), which maximize the minimum expected payoff. Formally,

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

A strategy p is a maximin strategy for the row player in the game M_{AB} if $p \in MS(M, A, B)$. Note that the set of maximin strategies is convex, since it is the set of solutions to a linear 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 Chapter 4 we give a new and simple proof of the minimax theorem.

The following two lemmas state useful facts about games, which will be applied in the proofs of our results. Raghavan (1994) showed that every feasible action of the row player that yields the same payoff as a maximin strategy against all maximin strategies of the column player is played with positive probability in some maximin strategy of the row player. This is known as the *equalizer theorem*.

Proposition 3.1 *(Raghavan, 1994)*
Let $M \in \mathbb{R}^{U \times U}$, $A, B \in \mathcal{F}(U)$, $p \in MS(M, A, B)$, and $i \in A$. If $(Mq)_i = p^t M q$ for all $q \in MS(-M^t, B, A)$, then there is $\hat{p} \in MS(M, A, B)$ with $\hat{p}_i > 0$.

Following Harsanyi (1973a), a Nash equilibrium (p, q) is quasi-strict if every action of the row player that is outside the support of p yields strictly less expected payoff against q than every action in the support of p (and similarly for the column player).[16] It is a well-known fact that if all equilibria of a game are quasi-strict, then there is a unique equilibrium. For the case of zero-sum games, Lemma 3.2 generalizes this observation by showing that an equilibrium is quasi-strict if and only if it is in the relative interior of the set of equilibria. The proof of Lemma 3.2 makes use of the equalizer theorem.

Lemma 3.2
Let $M \in \mathbb{R}^{U \times U}$ and $A, B \in \mathcal{F}(U)$. Then, (p, q) is a quasi-strict equilibrium if and only if $(p, q) \in \text{relint}(MS(M, A, B) \times MS(-M^t, B, A))$.

Proof. Let $v = p^t M q \in \mathbb{R}$ be the value of M_{AB} and $S = MS(M, A, B) \times MS(-M^t, B, A)$. Now let $(p, q) \in \text{relint}(S)$ and assume for contradiction that (p, q) is not quasi-strict. Then, without loss of generality, there is $i \in A \setminus \text{supp}(p)$ such that $(Mq)_i = v$. Proposition 3.1 implies that there is at least one quasi-strict equilibrium $(\hat{p}, \hat{q}) \in S$. Since

16 Harsanyi originally used the term quasi-strong equilibrium, which was referred to as quasi-strict equilibrium in subsequent work to avoid confusion with Aumann's notion of strong equilibrium (Aumann, 1959).

$(p,q) \in \text{relint}(S)$, there is $\epsilon > 0$ such that $(p^\epsilon, q^\epsilon) = ((1+\epsilon)p - \epsilon\hat{p}, (1+\epsilon)q - \epsilon\hat{q}) \in S$. Then,

$$(Mq^\epsilon)_i = (1+\epsilon)\underbrace{(Mq)_i}_{=v} - \epsilon\underbrace{(M\hat{q})_i}_{<v} > v,$$

which contradicts $(p^\epsilon, q^\epsilon) \in S$.

Now let $(p,q) \in S \setminus \text{relint}(S)$ and assume for contradiction that (p,q) is quasi-strict. Let $(\hat{p},\hat{q}) \in \text{relint}(S)$. If $\text{supp}(\hat{p}) \not\subseteq \text{supp}(p)$, there is $i \in \text{supp}(\hat{p}) \setminus \text{supp}(p)$ such that $(Mq)_i < v$, since (p,q) is quasi-strict. Hence, $\hat{p}^t Mq < v$, which contradicts $(\hat{p}, q) \in S$. Thus, we get $\text{supp}(\hat{p}) \subseteq \text{supp}(p)$ and, similarly, $\text{supp}(\hat{q}) \subseteq \text{supp}(q)$. This implies that there is $\epsilon > 0$ such that $(p^\epsilon, q^\epsilon) \in \Delta(A) \times \Delta(B)$ and, for all $i \in A \setminus \text{supp}(p)$ and $j \in B \setminus \text{supp}(q)$,

$$(Mq^\epsilon)_i = (1+\epsilon)\underbrace{(Mq)_i}_{<v} - \epsilon(M\hat{q})_i \leqslant v, \text{ and}$$

$$((p^\epsilon)^t M)_j = (1+\epsilon)\underbrace{(p^t M)_j}_{>v} - \epsilon(\hat{p}^t M)_i \geqslant v.$$

Hence, $(p^\epsilon, q^\epsilon) \in S$, which contradicts $(p,q) \in S \setminus \text{relint}(S)$. □

Lemma 3.2 implies that if both player have a unique maximin strategy, then this pair is a quasi-strict equilibrium.

3.2 SYMMETRIC ZERO-SUM GAMES

A game is symmetric if both players have the same set of feasible actions and swapping actions results in a swap of the payoffs. Formally, a proto game $M \in \mathbb{R}^{U \times U}$ is symmetric if M is skew-symmetric, i.e., $M = -M^t$. The set of all symmetric proto games is denoted by $\mathcal{M}$. If $M \in \mathcal{M}$ and $A \in \mathcal{F}(U)$, M_A is a symmetric game. Symmetry implies that the sets of recommended strategies coincide for both players. In particular, the sets of maximin strategies are the same for both players. For symmetric games, we will hence simply use the term maximin strategy without referring to a specific player. Observe that no player can guarantee an expected payoff of more than 0, since both players get expected payoff 0 if they play the same strategy. The minimax theorem implies that symmetric games have value 0. Symmetric games can be represented as weighted digraphs with actions corresponding to vertices and payoffs corresponding to weights of edges.

It will be useful to keep in mind that skew-symmetric matrices of odd size cannot have full rank. For $M \in \mathcal{M}$ and $A \in \mathcal{F}(U)$,

$$\det(M_A) = \det(M_A^t) = \det(-M_A) = (-1)^{|A|}\det(M_A).$$

Hence, $\det(M_A) = 0$ if A has odd cardinality. This implies that the rank of a skew-symmetric matrix is even. The functions ρ_A, $A \subseteq U$, defined below are partial reflections on $\mathcal{M}$. In the graphical representation, ρ_A inverts all edges between A and $U \setminus A$. Formally, for $A \subseteq U$, let $\rho_A \colon \mathcal{M} \to \mathcal{M}$ such that, for all $i, j \in U$,

$$(\rho_A(M))_{ij} = \begin{cases} M_{ij} & \text{if } i, j \in A \text{ or } i, j \in U \setminus A, \text{ and} \\ -M_{ij} & \text{otherwise.} \end{cases}$$

Similarly, for $x \in \mathbb{R}^U$, let

$$(\rho_A(x))_i = \begin{cases} -x_i & \text{if } i \in A, \text{ and} \\ x_i & \text{otherwise.} \end{cases}$$

Partial reflections have been considered in the context of tournament games by Fisher and Ryan (1995), who used the term "flip operators". Some of their properties listed here have already been stated in Fisher and Ryan (1995, Lemma 1). Observe that, for all $A, B \subseteq U$, we have $\rho_A \circ \rho_B = \rho_{A \Delta B}$, where Δ denotes the symmetric difference of A and B. Furthermore, $\rho_A = \rho_{U \setminus A}$ for all $A \subseteq U$. As a consequence, $(\{\rho_A \colon A \subseteq U\}, \circ)$ is a group with neutral element $\rho_\emptyset$ where every element is self-inverse. The fact that $A \Delta B = B \Delta A$ implies that this group is Abelian. The following lemma shows that ρ_A commutes with matrix-vector multiplication, which will be useful to prove the minimax theorem for symmetric games and to determine the distribution of maximin strategies in symmetric games.

Lemma 3.3

Let $M \in \mathcal{M}$, $x \in \mathbb{R}^U$, and $A \subseteq U$. Then,

$$\rho_A(M)\rho_A(x) = \rho_A(Mx).$$

Proof. This is readily checked by verifying the following sequence of equalities:

$$\begin{aligned} \rho_A(M)\rho_A(x) &= \begin{pmatrix} M_A & -M_{A,U\setminus A} \\ -M_{U\setminus A,A} & M_{U\setminus A} \end{pmatrix} \cdot \begin{pmatrix} -x_A \\ x_{U\setminus A} \end{pmatrix} \\ &= \begin{pmatrix} -(Mx)_A \\ (Mx)_{U\setminus A} \end{pmatrix} = \rho_A(Mx). \end{aligned}$$

□

4 A PROOF OF THE MINIMAX THEOREM

> 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.
>
> J. von Neumann

The minimax theorem states that in every zero-sum game the minimum expected payoff of a maximin strategy for the row player is the same as the negative of the minimum expected payoff of a maximin strategy for the column player. Equivalently, it shows that the value of a game is well-defined. The original proof of the minimax theorem by von Neumann (1928) used methods from functional analysis and was quite elaborate. We give a much simpler proof that only uses basic linear algebra and analysis. The theorem is first proven for symmetric games and then generalized to arbitrary games using a symmetrization procedure due to Gale et al. (1950). This shows that proving the minimax theorem for symmetric games is not essentially easier than proving it for arbitrary games. The idea to invoke symmetrization to prove the minimax theorem is not new and has been exploited previously by Gale et al. (1950). Their proof is purely algebraic and relies on a little known theorem of the alternative by Stiemke (1915).

Since games are assumed to be zero-sum, the column player maximizes his minimal expected payoff if he minimizes the maximal expected payoff of the row player. Hence, the minimax theorem can be stated as follows.

THEOREM 4.1

Let $M \in \mathbb{R}^{U \times U}$ and $A, B \in \mathcal{F}(U)$. Then,

$$\max_{p \in \Delta(A)} \min_{q \in \Delta(B)} p^t M q = \min_{q \in \Delta(B)} \max_{p \in \Delta(A)} p^t M q.$$

Proof. The proof goes along the following lines. We first prove the theorem for an arbitrary symmetric game M_A. To this end, we consider the set of core vectors of submatrices of M that are induced by odd-sized subsets of A. (Here we use the fact that skew-symmetric matrices of odd size cannot have a trivial core.) The partial reflection of every such vector that maps it to a non-negative vector (normalized to unit sum) is a maximin strategy of the corresponding partial reflection of M. For generic games, these vectors are pairwise distinct and are maximin strategies for different partial reflections of M. Since the number of odd-sized subsets of A is the same as the number of different partial reflections of M, we can conclude that the partial reflection of one of these vectors is a maximin strategy of M_A. This proves the existence of maximin strategies for generic symmetric games. A simple analytic argument extends this conclusion to all symmetric games. The symmetrization procedure by Gale et al. (1950) allows us to connect maximin strategies in symmetric games to maximin strategies in arbitrary games.

Let $M \in \mathcal{M}$ and $A \in \mathcal{F}(U)$. Without loss of generality, it is assumed that $M_{ij} = 0$ for all $(i, j) \notin A \times A$. Note that, by skew-symmetry of M, $\max_{p \in \Delta(A)} \min_{q \in \Delta(A)} p^t M q \leqslant 0$, since $p^t M p = 0$ for all $p \in \Delta(A)$, and

$$\min_{q \in \Delta(A)} \max_{p \in \Delta(A)} p^t M q = - \max_{q \in \Delta(A)} \min_{p \in \Delta(A)} q^t M p.$$

Hence, it suffices to show that there is $p \in \Delta(A)$ such that $p^t M \geqslant 0$.

Denote by $\mathcal{F}_{even}(A)$ the set of subsets of A of even cardinality and by $\mathcal{F}_{odd}(A)$ the set of subsets of A of odd cardinality. First consider the case when M_S has full rank for every $S \in \mathcal{F}_{even}(A)$. For every $S \in \mathcal{F}_{odd}(A)$, let $p^S \in \mathbb{R}^U \setminus \{0\}$ such that $(p^S_S)^t M_S = 0$, $\mathrm{supp}(p^S) \subseteq S$, and $\|p^S\| = 1$. The vectors p^S exist, since skew-symmetric matrices of odd size cannot have full rank. Note that the support of p^S is S for every $S \in \mathcal{F}_{odd}(A)$, as otherwise there would be a submatrix of M_A of even size that does not have full rank. For the same reason, $((p^S)^t M)_i \neq 0$ for all $i \in A \setminus S$. Consider the function $f \colon \mathcal{F}_{odd}(A) \to 2^A/\cdot^c$, where $2^A/\cdot^c$ denotes the quotient space of 2^A with respect to the complement operation, defined as follows:

$$f(S) = [((p^S)^t M)_- \cup p^S_-].$$

By definition of f, we have that

$$\rho_{p^S_-}(p^S)^t \rho_{f(S)}(M) =$$

$$\begin{array}{c} \\ \begin{pmatrix} p^S_{p^S_+} \\ -p^S_{p^S_-} \\ 0 \\ 0 \end{pmatrix}^t \end{array} \begin{array}{c} \begin{array}{cccc} p^S_+ & p^S_- & ((p^S)^t M)_+ & ((p^S)^t M)_- \end{array} \\ \begin{pmatrix} & - & & - \\ - & & - & \\ & - & & - \\ - & & - & \end{pmatrix} \end{array} \geqslant 0.$$

The above matrix depicts $\rho_{f(S)}(M)$. Blank cells mark submatrices that remain unchanged by $\rho_{f(S)}$ and minus signs correspond to submatrices whose entries are negated by $\rho_{f(S)}$. The column labels denote the corresponding partition of the columns; the partition of the rows is the same as for the columns. It can then be observed that $\rho_{p^S_-}(p^S)$ guarantees a payoff of at least 0 for the row player in $\rho_{f(S)}(M)$ when only actions in A are feasible.

We show that f is injective. Let $S, T \in \mathcal{F}_{odd}$ such that $S \neq T$ and assume for contradiction that $f(S) = f(T) = X$. Then, we have that

$$\rho_{p^S_-}(p^S)^t \rho_X(M) \geqslant 0 \qquad \text{and} \qquad \rho_{p^T_-}(p^T)^t \rho_X(M) \geqslant 0.$$

Since M is skew-symmetric and $\mathrm{supp}(\rho_{p^S_-}(p^S)) = \mathrm{supp}(p^S) = S$ and $\mathrm{supp}(\rho_{p^T_-}(p^T)) = \mathrm{supp}(p^T) = T$, it follows that

$$(\rho_{p^S_-}(p^S)^t \rho_X(M))_{S\cup T} = 0 \qquad \text{and} \qquad (\rho_{p^T_-}(p^T)^t \rho_X(M))_{S\cup T} = 0.$$

The fact that ρ_X is self-inverse and Lemma 3.3 imply that

$$(\rho_{X\Delta p^S_-}(p^S)^t M)_{S\cup T} = 0 \qquad \text{and} \qquad (\rho_{X\Delta p^T_-}(p^T)^t M)_{S\cup T} = 0.$$

Hence, $M_{S\cup T}$ contains an even-sized square submatrix that does not have full rank, which contradictions our initial assumption. Since $|\mathcal{F}_{odd}| = |2^A/\cdot^c| = 2^{|A|-1}$, it follows that f is bijective and, in particular, surjective. Thus, for every $X \in 2^A$, there is $S \in \mathcal{F}_{odd}(A)$ such that $\rho_{p^S_-}(p^S)^t \rho_X(M) \geqslant 0$. In particular, there is $S \in \mathcal{F}_{odd}(A)$ such that $\rho_{p^S_-}(p^S)^t M \geqslant 0$. Since $\rho_{p^S_-}(p^S) \geqslant 0$ and $\|p^S\| = 1$ by assumption, $\rho_{p^S_-}(p^S)$ is a maximin strategy of M.

Next, we show that arbitrary symmetric games have value 0. To this end, let $M \in \mathcal{M}$ and $A \in \mathcal{F}(U)$. Again, we may assume without loss of generality that $M_{ij} = 0$ for all $(i,j) \notin A \times A$. For every $S \in \mathcal{F}_{even}(A)$, the set of matrices in $\mathcal{M}$ such that the submatrix induced by S does not have full rank is nowhere dense in $\mathcal{M}$. Since the union of finitely many nowhere dense sets is nowhere dense, the set of matrices such that every square submatrix with rows and columns in A of even size has full rank is the complement of a nowhere dense set and thus dense in $\mathcal{M}$. Hence, we can find a sequence $(M^i)_{i\in\mathbb{N}} \subseteq \mathcal{M}$ such that M^i converges to M and M^i_S has full rank for every $S \in \mathcal{F}_{even}(A)$ and $i \in \mathbb{N}$. We know from before, that M^i_A has value 0, i.e., there is $p^i \in \Delta(A)$ such that $(p^i)^t M^i \geqslant 0$, for all $i \in \mathbb{N}$. The sequence $(p^i)_{i\in\mathbb{N}}$ admits a convergent subsequence $(p^{i_j})_{j\in\mathbb{N}}$. Denote by $p \in \Delta(A)$ its limit point. Then,

$$p^t M = \lim_{j\to\infty} (p^{i_j})^t M^{i_j} \geqslant 0.$$

Hence, M has value 0.

Lastly, we use the previously obtained conclusion for symmetric games to prove the minimax theorem for arbitrary games. Let $M \in \mathbb{R}^{U\times U}$ and $A, B \in \mathcal{F}(U)$. Without loss of generality, we may assume that all entries of M_{AB} are positive, since adding the same constant to all entires of M_{AB} does not change the set of maximin strategies. We invoke a symmetrization procedure by Gale et al. (1950) that relates maximin strategies of an arbitrary game to maximin strategies of a symmetric game. Assume without loss of generality that A and B are disjoint and let $h \in U \setminus (A \cup B)$. Let $\hat{M} \in \mathbb{R}^{U \times U}$ such that

$$\hat{M}_{A \cup B \cup \{h\}} = \begin{array}{c} \\ \\ \\ \end{array} \begin{array}{ccc} A & B & h \end{array}$$

$$\hat{M}_{A \cup B \cup \{h\}} = \begin{pmatrix} 0 & M_{AB} & -\mathbb{1} \\ -M_{BA}^t & 0 & \mathbb{1} \\ \mathbb{1}^t & -\mathbb{1}^t & 0 \end{pmatrix} \begin{array}{c} A \\ B \\ h \end{array}$$

The symbol $\mathbb{1}$ denotes a column vector of appropriate size with all entries equal to 1. The matrix $\hat{M}_{A \cup B \cup \{h\}}$ corresponds to a symmetric game where both players' choice of strategy can be decomposed into a two stage decision process: first, they can choose to play the original game as either the row player or the column player or they can play a additional action called "hedging"; second, if they choose to play the original game, they have to choose an action from either A or B. Since all payoffs in M_{AB} are assumed to be positive (so that the payoff of the row player is guaranteed to be positive when playing M_{AB}), the first choice corresponds to a variant of "rock, paper, scissors". From before, we know that $\hat{M}$ has value 0, i.e., there is $\hat{p} \in \Delta(A \cup B \cup \{h\})$ such that $(\hat{p}^t \hat{M})_{A \cup B \cup \{h\}} \geqslant 0$. Let $\hat{p} = (\alpha p, \beta q, \gamma)$ for some $p \in \Delta(A)$, $q \in \Delta(B)$, and $(\alpha, \beta, \gamma) \in \Delta(\{A, B, h\})$. By case analysis, it can be seen that $\hat{p}_a > 0$ and $\hat{p}_b > 0$ for some $a \in A$ and $b \in B$ and $\hat{p}_h > 0$. This implies that all of α, β, γ are strictly positive. Thus,

$$-\beta(q^t M^t)_A + \gamma \geqslant 0 \qquad \text{and} \qquad \alpha(p^t M)_B - \gamma \geqslant 0,$$

or, equivalently,

$$(Mq)_A \leqslant \gamma/\beta \qquad \text{and} \qquad (p^t M)_B \geqslant \gamma/\alpha.$$

Multiplication of $\hat{p}$ with the column of $\hat{M}$ corresponding to h yields that $\alpha = \beta$. Thus, M_{AB} has value γ/α and maximin strategies p and q. $\square$

5 JUSTIFICATION OF MAXIMIN PLAY

> To achieve a meaningful extension of von Neumann's value, we must take into account the interactive nature of games: that the players are rational, and reason about each other.
>
> R. J. Aumann and J. H. Drèze

Much of game theory is concerned with the analysis of equilibrium concepts. Typically questions such as the existence and uniqueness of equilibria or the computational complexity of finding an equilibrium are addressed. A wide range of different equilibrium notions has been proposed, perhaps most notably the notions of Nash equilibrium (Nash, 1950a) (and various refinements thereof) and correlated equilibrium (Aumann, 1974) (a coarsening of Nash equilibrium). For general normal-form games, both of these notions require some form of coordination among the players, i.e., optimality of a strategy is subject to the strategies chosen by other players. In the absence of a central coordination device, it is however unclear why a player should expect other players to play certain strategies. Even if it is agreed that the players' strategies should form, say, a Nash equilibrium, ambiguity remains as to which Nash equilibrium should be obtained in case there are multiple. In any case, even finding a Nash equilibrium has been shown to be complete for the complexity class PPAD (Daskalakis et al., 2009) even when there are only two players (Chen et al., 2009). Hence, for reasonably large games it may be intractable to even find a Nash equilibrium.

For two-player zero-sum games the situation is different in at least two respects. The set of Nash equilibria is precisely the set of pairs of maximin strategies. This set of pairs of strategies is Cartesian, i.e., the cross product of two sets of strategies. Hence, in two-player zero-sum games, optimality of a player's strategy does not depend on the strategy played by the other player.[17] Secondly, maximin strategies can be computed in polynomial time in the size of the game.

Still, it remains to be answered why players should play maximin strategies. On the surface this seems like a problem of decision making under uncertainty, where the uncertainty comes from the fact that a player does not know the strategy of his opponent. This however

17 Note that the definition of a solution concept for two-player zero-sum games in Chapter 3 already rules out dependence of the choice of strategy by a player on the choice of strategy by the other player.

is different from nature probabilistically choosing some event, since a player's choice of strategy is influenced by reasoning about his opponent and reasoning about the reasoning about his opponent and so on. Aumann and Drèze (2008) argue that games given by a set of actions for each player and the corresponding payoffs are underspecified in the sense that it is not possible to give a definite recommendation on this basis. Epistemic game theory aims to put games into context by assuming that the players have knowledge apart from the specification of the game. For example, a player may know that his opponent is rational and possibly even that his opponent knows that he himself is rational and possibly he may even have higher order beliefs. Such belief hierarchies are typically formulated via *interactive belief systems* as introduced by Harsanyi (1967) (see also Aumann and Brandenburger, 1995; Perea, 2007; Aumann and Drèze, 2008; Perea, 2012). A belief system consists 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. This model does not assume that players actively randomize, only the beliefs about the types of the other players are randomized.

Aumann and Brandenburger (1995) show 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.[18] Here rationality prescribes that the action chosen by a type has to maximize his *expected* payoff given his belief. This result extends to arbitrary games if the beliefs are commonly known and admit a common prior. Perea (2012) gives a set of assumptions about the players' beliefs that imply that the beliefs of a player constitute a Nash equilibrium but argues that some of these assumptions are unrealistic, e.g., they require that players believe that their opponents hold correct beliefs about other players. Hence, Nash equilibrium is at least questionable from an epistemic perspective for general normal-form games. A different viewpoint was taken by Aumann and Drèze (2008) who argue which payoff a rational player should expect. If there is reason for players to play maximin strategies in two-player zero-sum games, then they should expect the value of the game. Rationality or even common knowledge of rationality are not sufficient to derive this conclusion. However, Aumann and Drèze (2008) showed that if rationality is common knowledge and the beliefs admit a common prior, then the players should expect the value of the game.

Common knowledge assumptions in game theory have been criticized for not adequately modeling reality. The Wilson doctrine (Wil-

18 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.

son, 1987) states that effort should be made to derive results that do not require common knowledge assumptions. Steps in that direction were taken by Barelli (2009), Hellman (2013), and Bach and Tsakas (2014) who showed that the results of Aumann and Brandenburger (1995) and Aumann and Drèze (2008) still hold with weakened common knowledge assumptions.

As there is serious doubt about the significance of any strategic advice for general normal-form games, our considerations are restricted to two-player zero-sum games. Furthermore, the model studied here is different in that it is not based on interactive belief systems but solution concepts, which map a proto game and sets of feasible actions to a set of "good" strategies. In particular, it is assumed that players may implement randomized strategies. This framework allows to relate the choice of strategy for different feasible sets of actions to each other. Assumptions about other players are captured by requiring that certain strategies must or must not be chosen.

The equivalent of Arrow's independence of irrelevant alternatives in social choice theory in our model, called *independence of infeasible actions*, prescribes that the strategy of a player should not depend on the payoffs for infeasible actions. This condition seems basic enough that one might consider making it implicit in the definition of a solution concept, since any information about infeasible actions does not make the situation strategically different.

A notable difference to the afore-mentioned work is that we do not require players to maximize their expected payoff subject to some probabilistic belief about the other player's strategy. Our rationality assumptions are phrased purely in deterministic terms. A player is called rational if he never plays actions that are strictly dominated, i.e., the strategies he chooses assign probability zero to such actions. If a rational player knows that his opponent is rational, he will never play actions that are dominated given that his opponent never plays dominated actions. This condition is called *mutual knowledge of rationality*. It is weaker than common knowledge of rationality, which is obtained if belief in rationality is assumed for arbitrary long sequences. Common knowledge of rationality is equivalent to the condition that players only assign positive probability to rationalizable actions, i.e., actions that survive the process of iterated elimination of dominated actions (Bernheim, 1984; Pearce, 1984; Tan and da Costa Werlang, 1988). Gintis (2009) argues that common knowledge of rationality is too strong, since it is not implied by any set of plausible epistemic conditions.

The third condition prescribes how players deal with games that they consider equivalent in terms of chosen strategies, i.e., they would choose the same *pair* of strategies, one for each player, in both games. Now assume that a coin is tossed to decide which of two equivalent games is played and a player has to decide on his strategy before knowing the outcome of the coin toss. *Consistency* prescribes that any strategy that he would choose in both games is also chosen prior to the coin toss. This assumes that the game resulting from choosing a strategy before the execution of the coin toss is treated in the same way as the game whose payoffs are the expected payoffs of the randomization. This is in fact the only place where expected payoffs enter the picture. One could define a stronger notion of consistency, called strong consistency, by assuming that a player considers two games equivalent if he plays the same strategy as the row player in both games. This however neglects the interactive nature of game theory. The same strategy may be chosen in different games based on different strategic reasonings. When a new game is obtained by randomization over others, a different reasoning may apply. Equivalence as in the definition of consistency requires that the same strategy is chosen in different games based on the same reasoning, assuming that the reasoning of a player is based on his own strategic choices.

Lastly, players are assumed to be consequentialists in the sense that only the payoffs of an action are relevant to them, not the name of the action. If there are actions that yield the same payoff against every action of the opponent, so-called clones, the probability assigned to other actions should be independent of which of the clones are feasible, as long as at least one of them is feasible; the remaining probability can be assigned arbitrarily to the clones. Additionally, a player's strategy must not depend on the feasibility of clones for his opponent. This condition is called *consequentialism*. The idea to relate a game to one that contains clones of an action has also been used by Aumann and Drèze (2008). Their main result shows that any rational expectation for a player is identical to his expected payoff in a correlated equilibrium of the game that contains two clones of each action. This exploits the fact that the expectation of a player does not depend on the feasibility of clones of actions that he does not play.

We show that choosing all maximin strategies satisfies independence of infeasible actions, mutual knowledge of rationality, consistency, and consequentialism (Theorem 5.2). Our main result states that a player who adheres to all of these axioms has to choose maximin strategies (Theorem 5.3). The conditions are formally presented in Section 5.1. Section 5.2 gives the proofs of the main results. Section 5.3 concludes with a number of remarks about the result.

5.1 INDEPENDENCE, RATIONALITY, CONSISTENCY, AND CONSEQUENTIALISM

The strategies chosen by a player for different games can be summarized by a solution concept. Assumptions about the player are phrased in terms of properties of this solution concept. For games that contain irrational payoffs, it is not clear that the payoff matrix can be specified in a finite amount of space. Additionally, strategies that involve irrational probabilities may not be implementable in practice. Because of these concerns, we only consider games from $\mathbb{Q}^{U \times U}$ and solution concepts that map to strategies from $\Delta^{\mathbb{Q}}(U)$ for the purpose of this chapter. Observe that games with rational-valued payoffs admit rational-valued maximin strategies and consequentially also have a rational-valued value. Since we consider axioms that relate choices for different feasible sets of actions to each other, we assume that the set of all conceivable actions U is infinite. The proof of Theorem 5.3, the main result of this chapter, crucially relies on the assumptions that strategies are rational-valued and that U is infinite.

The first property prescribes that only the payoffs for feasible actions should be taken into account; the payoffs for infeasible actions are irrelevant. A solution concept f satisfies independence of infeasible actions if, for all $M, \hat{M} \in \mathbb{Q}^{U \times U}$ and $A, B \in \mathcal{F}(U)$,

$$f(M, A, B) = f(\hat{M}, A, B) \text{ whenever } M_{AB} = \hat{M}_{AB}. \tag{IIA}$$

An action dominates another action if it yields more payoff against every action of the opponent. For $M \in \mathbb{Q}^{U \times U}$, $A, B \in \mathcal{F}(U)$, and $i, i' \in U$, i is dominated by i' with respect to B if $M_{ij} < M_{i'j}$ for all $j \in B$. In the sequel, $D(M, A, B)$ denotes the set of actions in A that are dominated by some other action in A with respect to B. Note that domination is preserved when B is replaced by one of its subsets. It is never advisable for a player to play a dominated action, since there is an action that is preferable independently of which action is played by the other player. Hence, every reasonable solution concept should assign probability zero to all dominated actions. A solution concept f is rational if, for all $M \in \mathbb{Q}^{U \times U}$ and $A, B \in \mathcal{F}(U)$, $p_i = 0$ for all $i \in D(M, A, B)$ and $p \in f(M, A, B)$. Rationality is probably one of the most uncritical assumptions one can make about self-interested players, since it does not rely on expected payoffs or any assumptions about the other player. In fact, it seems basic enough that a player can safely assume that his opponent is rational. If a player knows that his opponent is rational, then he knows that the latter will not play dominated actions. But then, if the player is himself rational, he should not play actions that are dominated with respect to the set of undominated actions of his opponent. This assumption is called mutual knowledge of rationality. Formally, a solution concept f satisfies mutual knowledge of rationality if, for all $M \in \mathbb{Q}^{U \times U}$, $A, B \in \mathcal{F}(U)$,

$$p_i = 0 \text{ for all } i \in D(M, A, B \setminus D(-M^t, B, A)) \text{ and } p \in f(M, A, B). \tag{mutual knowledge of rationality}$$

The rationale underlying mutual knowledge of rationality could be applied further to obtain arbitrarily long chains of the form "a player knows that his opponent knows ... that he is rational", where "he" may be either player. If these beliefs obtain for arbitrary long chains, rationality is common knowledge. As argued before, common knowledge of rationality is a rather demanding assumption.

Two games are considered equivalent with respect to some solution concept if it returns the same strategy from the perspective of both players. For a solution concept f, $\hat{M}, \bar{M} \in \mathbb{Q}^{U \times U}$, and $A, B \in \mathcal{F}(U)$, $\hat{M}_{AB}$ and $\bar{M}_{AB}$ are called f-equivalent if $f(\hat{M}, A, B) \cap f(\bar{M}, A, B) \neq \emptyset$ and $f(-\hat{M}^t, B, A) \cap f(-\bar{M}^t, B, A) \neq \emptyset$. Consistency prescribes that any strategy that is chosen in two f-equivalent games is also chosen in any game that is derived by randomizing over these two games. Formally, a solution concept f satisfies consistency if, for all $\hat{M}, \bar{M} \in \mathbb{Q}^{U \times U}$, $A, B \in \mathcal{F}(U)$, and $\lambda \in [0,1] \cap \mathbb{Q}$ such that $\hat{M}_{AB}$ and $\bar{M}_{AB}$ are f-equivalent,

$$f(\hat{M}, A, B) \cap f(\bar{M}, A, B) \subseteq f(M, A, B), \qquad \text{(consistency)}$$

where $M = \lambda\hat{M} + (1-\lambda)\bar{M}$.

In some games, a player can only distinguish two actions by their names but not by their payoffs, i.e., both actions yield the same payoff independently of the action of the other player. Such actions are called clones. Formally, two actions $i, i' \in U$ are clones in $M \in \mathbb{Q}^{U \times U}$ if $M_{ij} = M_{i'j}$ for all $j \in U$. Let $\hat{A}, \hat{B}, C, D \in \mathcal{F}(U)$ such that $\hat{A} \cap C = \{a\}$, $A = \hat{A} \cup C$ and $\hat{B} \cap D = \{b\}$, $B = \hat{B} \cup D$. A solution concept f satisfies consequentialism if, for all $M \in \mathbb{Q}^{U \times U}$ such that all actions in C are clones in M and all actions in D are clones in $-M^t$,

$$f(M, \hat{A}, \hat{B}) \times_a \Delta^{\mathbb{Q}}(C) = f(M, A, B). \qquad \text{(consequentialism)}$$

Hence, if consequentialism obtains, probability among clones can be distributed arbitrarily. The probability assigned to other actions is not affected by the feasibility of clones of a feasible action. Moreover, a player's strategy is not influenced by the feasibility of clones for his opponent.

A related condition called neutrality prescribes that renaming the actions for the row player results in the same renaming in the set of chosen strategies. Renaming the actions for the column player has no effect on the choices of the row player. Renaming actions corresponds to permuting rows and columns in the payoff matrix. A solution concept f satisfies neutrality if, for all $M \in \mathbb{Q}^{U \times U}$, $A, B \in \mathcal{F}(U)$, and $\pi, \sigma \in \Pi(U)$,

$$f(M_{\pi\sigma}, \pi(A), \sigma(B)) = f(M, A, B)_{\pi}. \qquad \text{(neutrality)}$$

The following lemma shows that consequentialism in conjunction with independence of infeasible alternatives implies neutrality. This implication is driven by the part of consequentialism that prescribes that the probability assigned to non-clones does not depend on the number of feasible clones.

LEMMA 5.1
Every solution concept that satisfies IIA and consequentialism satisfies neutrality.

Proof. We only prove the case when σ is the identity. The full statement can be shown by applying the construction below to the column player instead of the row player. Let $M \in \mathbb{Q}^{U \times U}$, $A, B \in \mathcal{F}(U)$, and $\pi \in \Pi(U)$. Let $p^A \in f(M, A, B)$ and $A = \{a_1, \ldots, a_m\}$. Since U is infinite, there is $\hat{A} = \{\hat{a}_1, \ldots, \hat{a}_m\} \in \mathcal{F}(U)$ such that $\hat{A} \cap (A \cup \pi(A)) = \emptyset$. Now let $\hat{M} \in \mathbb{Q}^{U \times U}$ such that $M_{AB} = \hat{M}_{AB}$ and $\{a_i, \hat{a}_i\}$ is a set of clones for all $i \in [m]$. By IIA, we have that $p^A \in f(\hat{M}, A, B)$. We now apply consequentialism to a_i and $\{a_i, \hat{a}_i\}$ for all $i \in [m]$. This implies that

$$f(\hat{M}, A, B) \times_{a_1} \Delta^{\mathbb{Q}}(\{a_1, \hat{a}_1\}) \cdots \times_{a_m} \Delta^{\mathbb{Q}}(\{a_m, \hat{a}_m\}) = f(\hat{M}, A \cup \hat{A}, B).$$

Hence, $p^{\hat{A}} \in f(\hat{M}, A \cup \hat{A}, B)$, where $p^{\hat{A}}_{\hat{a}_i} = p^A_{a_i}$ for all $i \in [m]$. Finally, let $\bar{M} \in \mathbb{Q}^{U \times U}$ such that $\bar{M}_{\hat{A}B} = \hat{M}_{\hat{A}B}$ and $\{\pi(a_i), \hat{a}_i\}$ is a set of clones for all $i \in [m]$. By IIA, we have that $p^{\hat{A}} \in f(\bar{M}, \hat{A}, B)$. As before, it follows from consequentialism that $p^{\pi(A)} \in f(\bar{M}, \pi(A), B)$, where $p^{\pi(A)}_{\pi(a_i)} = p^{\hat{A}}_{\hat{a}_i}$ for all $i \in [m]$. Notice that $p^{\pi(A)} = p^A_\pi$. Since $\bar{M}_{\pi(A),B} = (M_{\pi\sigma})_{\pi(A),B}$ by construction of $\bar{M}$, IIA implies that $p^A_\pi \in f(M_{\pi\sigma}, \pi(A), B)$. Thus, $f(M, A, B)_\pi \subseteq f(M_{\pi\sigma}, \pi(A), B)$. Equality follows from application of the above to $M_{\pi\sigma}$ and π^{-1}. $\square$

5.2 CHARACTERIZATION OF MAXIMIN STRATEGIES

Theorem 5.2 shows that choosing maximin strategies satisfies all of the above defined properties. The proof for mutual knowledge of rationality can be extended to show that *MS* is even compatible with common knowledge of rationality. On the other hand, it can be shown that *MS* violates strong consistency (cf. Remark 5.2).

THEOREM 5.2
MS satisfies IIA, mutual knowledge of rationality, consistency, and consequentialism.

Proof. The fact that *MS* satisfies IIA is clear by definition.

To show that *MS* satisfies mutual knowledge of rationality, let $M \in \mathbb{Q}^{U \times U}$, $A, B \in \mathcal{F}(U)$, and $v \in \mathbb{Q}$ be the value of M_{AB}. Let $p \in MS(M, A, B)$ and $q \in MS(-M^t, B, A)$. If $q_b > 0$ for some $b \in D(-M^t, B, A)$, then let $\hat{b} \in B$ be an action that dominates b with respect to A in $-M^t$ and $\hat{q} \in \Delta^{\mathbb{Q}}(B)$ be the strategy that is identical to q except that $\hat{q}_b = 0$ and $\hat{q}_{\hat{b}} = q_{\hat{b}} + q_b$. Then, $p^t M\hat{q} < p^t Mq = v$, which contradicts the assumption that p is a maximin strategy for the row player. Hence, $q_b = 0$ for all $b \in D(-M^t, B, A)$. If $p_a > 0$ for some $a \in D(M, A, B \setminus D(-M^t, B, A))$, the same construction can be used to obtain $\hat{p} \in \Delta^{\mathbb{Q}}(A)$ such that $\hat{p}^t M q > p^t M q = \nu$, which contradicts the assumption that q is a maximin strategy for the column player.

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

$$MS(\hat{M}, A, B) \cap MS(\bar{M}, A, B) \neq \emptyset, \text{ and}$$
$$MS(-\hat{M}^t, B, A) \cap MS(-\bar{M}^t, B, A) \neq \emptyset.$$

Let $p \in MS(\hat{M}, A, B) \cap MS(\bar{M}, A, B)$, $\lambda \in [0,1] \cap \mathbb{Q}$, and $M = \lambda\hat{M} + (1-\lambda)\bar{M}$. For the values $\hat{v}, \bar{v} \in \mathbb{Q}$ of $\hat{M}_{AB}$ and $\bar{M}_{AB}$, it follows that $p^t\hat{M} \geqslant \hat{v}$ and $p^t\bar{M} \geqslant \bar{v}$. Hence, $p^t M \geqslant \lambda\hat{v} + (1-\lambda)\bar{v}$. For $q \in MS(-\hat{M}^t, B, A) \cap MS(-\bar{M}^t, B, A)$ we have that $\hat{M}q \leqslant \hat{v}$, and $\bar{M}q \leqslant \bar{v}$. Hence, $Mq \leqslant \lambda\hat{v} + (1-\lambda)\bar{v}$. This implies that $\lambda\hat{v} + (1-\lambda)\bar{v}$ is the value of M_{AB}. Thus, $p \in MS(M, A, B)$.

Lastly we show that *MS* satisfies consequentialism. It is not hard to see 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 only consider the case when the set of clones for the column player is trivial, i.e., $D = \{b\}$. Let $M \in \mathbb{Q}^{U \times U}$, $\hat{A}, B, C \in \mathcal{F}(U)$ such that $\hat{A} \cap C = \{a\}$, $A = \hat{A} \cup C$, and all actions in C are clones in M. We have to show that

$$MS(M, \hat{A}, B) \times_a \Delta^{\mathbb{Q}}(C) = MS(M, A, B).$$

Let $p \in MS(M, \hat{A}, B) \times_a \Delta^{\mathbb{Q}}(C)$. Then, there are $\hat{p} \in MS(M, \hat{A}, B)$ and $\bar{p} \in \Delta^{\mathbb{Q}}(C)$ such that $\hat{p} \times_a \bar{p} = p$. Since all actions in C are clones in M, it follows that $\hat{p}^t M = p^t M$. Hence $p \in MS(M, A, B)$, which implies that $MS(M, \hat{A}, B) \times_a \Delta^{\mathbb{Q}}(C) \subseteq MS(M, A, B)$. To prove the other inclusion, let $p \in MS(M, A, B)$ and $\hat{p} \in \Delta^{\mathbb{Q}}(\hat{A})$ such that $\hat{p}_a = \sum_{c \in C} p_c$ and $\hat{p}_c = p_c$ for all $c \in A \setminus C$. Again, $\hat{p}^t M = p^t M$, since all actions in C are clones in M. Hence, $\hat{p} \in MS(M, \hat{A}, B)$ which implies that $MS(M, A, B) \subseteq MS(M, \hat{A}, B) \times_a \Delta^{\mathbb{Q}}(C)$. □

Our main theorem shows that every solution concept that satisfies IIA, mutual knowledge of rationality, consistency, and consequentialism has to choose maximin strategies. Together with Theorem 5.2, this implies that *MS* is the largest solution concept satisfying these properties.

THEOREM 5.3
If a solution concept f satisfies IIA, mutual knowledge of rationality, consistency, and consequentialism, then $f \subseteq MS$.

Proof. Assume for contradiction that $f \not\subseteq MS$, i.e., there are $M \in \mathbb{Q}^{U \times U}$ and $A, B \in \mathcal{F}(U)$ such that $f(M, A, B) \not\subseteq MS(M, A, B)$. Let

$v \in \mathbb{Q}$ be the value of M_{AB}. Let $p \in f(M, A, B) \setminus MS(M, A, B)$ and $q \in f(-M^t, B, A)$. If $p^t Mq < v$, there is $a \in A$ such that $\underline{a}^t Mq > p^t Mq$. If $p^t Mq \geqslant v$, there is $b \in B$ such that $p^t M\underline{b} < p^t Mq$. In any case, (p, q) is not a Nash equilibrium of M_{AB}. By symmetry of the roles of the row player and the column player, we may assume without loss of generality that there is $\hat{b} \in B$ such that $p^t M\underline{\hat{b}} < p^t Mq$.

Let δ be the greatest common divisor of $\{p_a : a \in A\}$, which exists, since f is assumed to map to $\Delta^{\mathbb{Q}}(U)$. For all $a \in A$, 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 $\hat{A} = \bigcup_{a\in A} A_a$ and $\hat{M} \in \mathbb{Q}^{U\times U}$ such that $\hat{M}_{AB} = M_{AB}$ and, for all $a \in A$, A_a is a set of clones in $\hat{M}$. By application of consequentialism to a and A_a for all $a \in A$, it follows that $\text{uni}(\bar{A}) \in f(\hat{M}, \hat{A}, B)$ and $q \in f(-\hat{M}^t, B, \hat{A})$, where $\bar{A} = \bigcup_{a\in \text{supp}(p)} A_a$. Let $\hat{\Pi} \subseteq \Pi(\bar{A})$ be the set of permutations $\pi \in \Pi(\bar{A})$ such that $\pi(a) = a$ for all $a \in U \setminus \bar{A}$. Since, by Lemma 5.1, f satisfies neutrality, it follows that $\text{uni}(\bar{A}) \in f(\hat{M}_{\pi,\text{id}}, \hat{A}, B)$ and $q \in f(-\hat{M}^t_{\text{id},\pi}, B, \hat{A})$ for all $\pi \in \hat{\Pi}$. Let $\bar{M} = {}^1/_{|\bar{A}|!} \sum_{\pi\in\hat{\Pi}} \hat{M}_{\pi,\text{id}}$. Consistency implies that $\text{uni}(\bar{A}) \in f(\bar{M}, \hat{A}, B)$ and $q \in f(-\bar{M}^t, B, \hat{A})$. Observe that $\bar{A}$ is a set of clones in $\bar{M}$. By construction of $\bar{M}$,

$$\text{uni}(\bar{A})^t \bar{M}\underline{\hat{b}} = p^t M\underline{\hat{b}} < p^t Mq = \text{uni}(\bar{A})^t \bar{M}q.$$

Now let $M^{\bar{A}} \in \mathbb{Q}^{U\times U}$ such that, for all $a \in \bar{A}$ and $b \in B$,

$$M^{\bar{A}}_{ab} = \begin{cases} 1 & \text{if } a \in \bar{A}, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $\bar{A}$ is a set of clones in $M^{\bar{A}}$ and all actions in $\hat{A} \setminus \bar{A}$ are dominated by all actions in $\bar{A}$ with respect to B. Hence, (mutual knowledge of) rationality and consequentialism imply that $\text{uni}(\bar{A}) \in \Delta^{\mathbb{Q}}(\bar{A}) = f(M^{\bar{A}}, \hat{A}, B)$. Moreover, B is a set of clones in $-(M^{\bar{A}})^t$. Thus, consequentialism implies that $q \in \Delta^{\mathbb{Q}}(B) = f(-(M^{\bar{A}})^t, B, \hat{A})$.

Let $\lambda \in (0, {}^1/_{2(\max_{a\in\hat{A}, b\in B} |\bar{M}_{ab}|+1)})$ and $\mathring{M} = \lambda\bar{M} + (1-\lambda)M^{\bar{A}}$. Consistency implies that $\text{uni}(\bar{A}) \in f(\mathring{M}, \hat{A}, B)$ and $q \in f(-\mathring{M}^t, B, \hat{A})$. By the choice of λ, all actions in $\bar{A}$ dominate all actions in $\hat{A} \setminus \bar{A}$ with respect to B in $\mathring{M}$, i.e., $\hat{A} \setminus \bar{A} \subseteq D(\mathring{M}, A, B)$. Also, $\bar{A}$ is a set of clones in $\mathring{M}$, which implies that $\mathring{M}_{ab} = \mathring{M}_{a'b}$ for all $a, a' \in \bar{A}$ and $b \in B$. Moreover, $\text{uni}(\bar{A})^t \mathring{M}\underline{\hat{b}} < \text{uni}(\bar{A})^t \mathring{M}q$. Thus, there is $b \in \text{supp}(q)$ such that $\mathring{M}_{a\hat{b}} < \mathring{M}_{ab}$ for all $a \in \hat{A}$. Hence, $b \in D(-\mathring{M}^t, B, \hat{A} \setminus D(\mathring{M}, \hat{A}, B))$. This contradicts mutual knowledge of rationality. □

The following example illustrates (slightly simplified) the proof of Theorem 5.3. Recall the game of "rock, paper, scissors", where paper beats rock, rock beats scissors, and scissors beats paper and all other combinations are ties (cf. Section 1.2). The set of actions is $\{r, p, s\}$ for both players. We consider an extended variant of this game with three additional actions $\bar{r}, \bar{p}$, and $\bar{s}$ that represent the "negatives" of r, p, and s ($A = B = \{r, p, s, \bar{r}, \bar{p}, \bar{s}\}$). The comparisons among the

$$\begin{array}{c} \begin{array}{cccccc} r & p & s & \bar{r} & \bar{p} & \bar{s} \end{array} \\ \left(\begin{array}{ccc|ccc} 0 & -1 & 1 & -1 & 1 & 1 \\ 1 & 0 & -1 & 1 & -1 & 1 \\ -1 & 1 & 0 & 1 & 1 & -1 \\ \hline 1 & -1 & -1 & 0 & 1 & -1 \\ -1 & 1 & -1 & -1 & 0 & 1 \\ -1 & -1 & 1 & 1 & -1 & 0 \end{array}\right) \end{array}$$

(a) The original game M.

$$\begin{array}{c} \begin{array}{cccccc} r & p & s & \bar{r} & \bar{p} & \bar{s} \end{array} \\ \begin{pmatrix} 1 & -1 & -1 & 0 & 1 & -1 \\ & & & & & \\ & & & & & \\ 0 & -1 & 1 & -1 & 1 & 1 \\ & & & & & \\ & & & & & \end{pmatrix} \end{array}$$

(b) The game $M_{(r,\bar{r}),id}$ with r and $\bar{r}$ relabeled for the row player.

$$\begin{array}{c} \begin{array}{cccccc} r & p & s & \bar{r} & \bar{p} & \bar{s} \end{array} \\ \begin{pmatrix} 1/2 & -1 & 0 & -1/2 & 1 & 0 \\ & & & & & \\ & & & & & \\ 1/2 & -1 & 0 & -1/2 & 1 & 0 \\ & & & & & \\ & & & & & \end{pmatrix} \end{array}$$

(c) The uniform randomization $\bar{M}$ over M and $M_{(r,\bar{r}),id}$.

$$\begin{array}{c} \begin{array}{cccccc} r & p & s & \bar{r} & \bar{p} & \bar{s} \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

(d) The game $M^{\{r,\bar{r}\}}$.

$$\begin{array}{c} \begin{array}{cccccc} r & p & s & \bar{r} & \bar{p} & \bar{s} \end{array} \\ 1/8 \left(\begin{array}{cccccc} 7 & 4 & 6 & 5 & 8 & 6 \\ \hline 2 & 0 & -2 & 2 & -2 & 2 \\ -2 & 2 & 0 & 2 & 2 & -2 \\ \hline 7 & 4 & 6 & 5 & 8 & 6 \\ \hline -2 & 2 & -2 & -2 & 0 & 2 \\ -2 & -2 & 2 & 2 & -2 & 0 \end{array}\right) \end{array}$$

(e) The $(1/4, 3/4)$-randomization $\mathring{M}$ over $\bar{M}$ and $M^{\{r,\bar{r}\}}$.

Figure 5.1: Payoff matrices for the variant of "rock, paper, scissors" with action sets $A = B = \{r, p, s, \bar{r}, \bar{p}, \bar{s}\}$ and the corresponding games constructed in the proof of Theorem 5.3. Blank spaces in a matrix denote entries that are identical to the corresponding entries in the preceding matrix.

negatives are exactly reversed. Every non-negative action is beaten by its negative, but beats the other two negative actions. Since the set of feasible actions is A for both players throughout this example, we omit restriction to the set of feasible actions and call this game M. The corresponding payoff matrix is depicted in Figure 5.1(a).

For this larger set of actions, the unique maximin strategy is still to randomize uniformly over r, p, and s. However, rationality or even common knowledge of rationality are not enough to rule out *any* other strategies, since no action is dominated (not even weakly). In contrast, Theorem 5.3 shows that even mutual knowledge of rationality suffices to single out the maximin strategy if a player's choices additionally satisfy IIA, consistency, and consequentialism. Assume that such a player instead of playing the maximin strategy randomizes uniformly over r and $\bar{r}$ in M. Since the game is symmetric, i.e., $M = -M^t$, he would also randomize uniformly over r and $\bar{r}$ when playing $-M^t$. IIA and consequentialism imply neutrality by Lemma 5.1. Hence, the player would also randomize uniformly over r and $\bar{r}$ if their labels were swapped. The resulting game $M_{(r,\bar{r}),id}$ is depicted in Figure 5.1(b). Again, neutrality implies that he would also do so in $-M^t_{id,(r,\bar{r})}$. Now imagine that a fair coin is tossed. If it shows "heads", M is played, if it shows "tails", $M_{(r,\bar{r}),id}$ is played. The resulting game $\bar{M}$ is depicted in Figure 5.1(c). A player whose choices are consistent will still randomize uniformly over r and $\bar{r}$ even when he has to decide on a strategy before knowing the outcome of the coin toss, i.e., if he is playing the game $\bar{M}$. Now consider a game in which r and $\bar{r}$ beat each of p, $\bar{p}$, s, and $\bar{s}$ and all other combinations tie the game. This game called $M^{\{r,\bar{r}\}}$ is depicted in Figure 5.1(d). Rationality and consequentialism imply that the player would randomize uniformly over r and $\bar{r}$ in $M^{\{r,\bar{r}\}}$ (among other strategies). All actions are clones in $-(M^{\{r,\bar{r}\}})^t$. Hence, consequentialism implies that the player would also randomize uniformly over r and $\bar{r}$ in $-(M^{\{r,\bar{r}\}})^t$ (among other strategies). Now another coin is tossed. If it shows "heads", $\bar{M}$ is played and if it shows "tails", $M^{\{r,\bar{r}\}}$ is played. Assume that the coin is biased towards "tails", which is the outcome with a probability of 3/4. The game $\mathring{M}$ resulting from having to choose a strategy before the coin toss is depicted in Figure 5.1(e). Notice that r and $\bar{r}$ dominate all other actions in $\mathring{M}$. If a player knows that his opponent is rational, i.e., only randomizes over r and $\bar{r}$ when playing $\mathring{M}$, then r and $\bar{r}$ are dominated by p in $-\mathring{M}^t$. Consistency however implies that the player will randomize uniformly over r and $\bar{r}$ in $-\mathring{M}^t$, which contradicts mutual knowledge of rationality.

5.3 CONCLUDING REMARKS

We conclude this chapter with a number of remarks.

Remark 5.1 *(Independence of axioms)*

Mutual knowledge of rationality, consistency, and consequentialism are required to derive the conclusion of Theorem 5.3. The trivial solution concept that always chooses all strategies over feasible actions violates (mutual knowledge of) rationality but satisfies IIA, consistency, and consequentialism. The solution concept that chooses all maximin strategies of the game that results from taking the third power of all entries in the payoff matrix violates consistency but satisfies the remaining properties. The solution concept that chooses all strategies that result from squaring the probabilities in maximin strategies (and then normalizing to unit sum) violates consequentialism but satisfies the remaining properties. It is open whether IIA is also required.

Remark 5.2 *(Strong consistency)*
Choosing maximin strategies violates the stronger notion of consistency that is obtained if two games $\hat{M}_{AB}$ and $\bar{M}_{AB}$ are f-equivalent whenever $f(\hat{M}, A, B) \cap f(\bar{M}, A, B) \neq \emptyset$. Consider the games $\hat{M}_{AB}$ and $\bar{M}_{AB}$ where the row player can play either top or bottom ($A = \{t, b\}$) and the column player can play either left or right ($B = \{l, r\}$).

$$\hat{M}_{AB} = \begin{pmatrix} 0 & 4 \\ 2 & 0 \end{pmatrix} \quad \bar{M}_{AB} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \quad M_{AB} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$$

The unique maximin strategy in $\hat{M}_{AB}$ and $\bar{M}_{AB}$ is $1/3\,\underline{t} + 2/3\,\underline{b}$. But in the game M_{AB}, which results from randomizing uniformly over $\hat{M}_{AB}$ and $\bar{M}_{AB}$, the unique maximin strategy is to play t with probability one. In particular, $1/3\,\underline{t} + 2/3\,\underline{b}$ is not a maximin strategy in M_{AB}. Notice that the maximin strategies in $-\hat{M}^t_{AB}$ and $-\bar{M}^t_{AB}$ are different.

Remark 5.3 *(Symmetric games)*
For symmetric games the strong notion of consistency discussed in Remark 5.2 is equivalent to consistency and is hence satisfied by *MS*. Theorem 5.3 remains valid within the domain of symmetric games. This requires modifying the proof such that all constructed games are symmetric. 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 $\hat{M}_\pi$, where π ranges over the same set of permutations as in the original proof. Lastly, $M^{\hat{A}}$ has to be defined as a symmetric game in which all actions in $\hat{A} \cup \{\hat{b}\}$ dominate all actions in $A \setminus (\hat{A} \cup \{\hat{b}\})$ with respect to A for both players.

Remark 5.4 *(Normal-form games)*
Theorem 5.3 can be extended to normal-form games when considering solution concepts that choose a set of strategies for *every* player. In this framework, one can conclude that every tuple of chosen strategies has to be a Nash equilibrium. Two normal-form games are considered f-equivalent if the sets of strategies chosen by f in both games intersect for all players.

6 RANDOM SYMMETRIC ZERO-SUM GAMES

> Possessed of natural interest because of their special character, symmetric games are given additional importance by the computational procedures which are discussed by G. W. Brown and J. von Neumann in their contribution to this Study.
>
> D. Gale, H. W. Kuhn, and A. W. Tucker

This chapter studies the properties of randomly drawn symmetric zero-sum games. More precisely, we determine the distribution over supports of maximin strategies in randomly drawn symmetric games when the set of feasible actions A is fixed. It is shown that, for every set of actions $S \subseteq A$, the probability that a randomly drawn symmetric game admits a maximin strategy with support S is $2^{-(|A|-1)}$ if S has odd cardinality and 0 otherwise. In particular, this probability only depends on the parity of S. This stems from the fact that a generic skew-symmetric matrix of even size has full rank, while a skew-symmetric matrix of odd size cannot have full rank (cf. Section 3.2). For the proof of this result we assume that the distribution over games is symmetric and regular. A distribution is *symmetric* if it is invariant under partial reflections (cf. Section 3.2). Symmetry is for example satisfied by all distributions that arise from drawing payoffs from independent distributions that are symmetric about 0, i.e., distributions with even density function. A distribution is *regular* if a randomly chosen game almost surely admits a unique maximin strategy. We assume throughout that games are drawn from a symmetric and regular distribution.

Related questions have been studied for various classes of games. Wilson (1971) showed that the number of Nash equilibria is finite and odd for almost all n-player normal-form games. A different proof of the same statement was given by Harsanyi (1973b). McLennan (2005) derived a formula for the expected number of Nash equilibria in which the players' strategies have given supports in normal-form games. His model assumes that the payoffs of all players are independent and distributed uniformly over the unit sphere. If games are distributed such that Nash equilibria are almost surely unique, the expected number of Nash equilibria with given supports is equal to the probability that a game admits a Nash equilibrium with these supports. Thus, our result can also be phrased as determining the expected number of Nash equilibria with given supports. Follow-up work by McLennan and Berg (2005) has derived a formula for the expected number of Nash equilibria of a random two-player normal-form game. Similar to McLennan (2005), they assume that the payoffs of both players are drawn independently from a uniform distribution on the unit sphere.

For two-player zero-sum games every convex combination of Nash equilibria is again a Nash equilibrium. Hence there is either a unique Nash equilibrium or infinitely many. However, Wilson's theorem does not imply that Nash equilibria are almost surely unique in low dimensional subclasses of normal-form games, e.g., zero-sum games, symmetric zero-sum games, or tournament games.[19] Fisher and Ryan (1992) showed that every tournament game admits a unique maximin strategy, and hence a unique Nash equilibrium. This result was generalized by Laffond et al. (1997) to symmetric games where all off-diagonal payoffs are odd integers and by Le Breton (2005) to symmetric games where all off-diagonal payoffs satisfy a more general congruency condition. Closest to our result is the unpublished work of Roberts (2004), who proved the same formula that is derived here for a somewhat less general class of distributions over symmetric games. His Theorem 1 assumes that the payoffs are drawn from independent and identical distributions that are symmetric about 0. However, for the proof of this statement he only requires that the distribution over games is absolutely continuous and symmetric in our sense. The result presented here is more general in that it weakens absolute continuity to regularity.

For not necessarily symmetric games the situation is less clear. Experiments by Faris and Maier (1987) suggest that the support size of a maximin strategy of a game chosen uniformly at random approximately follows a binomial distribution that chooses half of the actions in expectation. Jonasson (2004) showed that maximin strategies are almost surely unique if the payoffs follow continuous, independent, and identically distributed random variables that are symmetric about 0. Moreover, he proved that the expected fraction of actions in the support of a maximin strategy is close to $1/2$ when the number of actions goes to infinity. Roberts (2006) considered games where payoffs follow independent and identical Cauchy distributions. Remarkably, he derives a closed form formula for the probability that a pair of maximin strategies of a random game has given supports.

The proof of our main result (Theorem 6.5) is divided into three statements. In Lemma 6.2 we determine the probability that a maximin strategy puts positive probability on all feasible actions, i.e., the probability that a game admits a totally mixed maximin strategy. Lemma 6.3 establishes that if a distribution over games is symmetric and regular on a given set of feasible actions, then it is also symmet-

19 Tournament games are symmetric games in which all off-diagonal payoffs are either 1 or -1.

ric and regular on any subset of these actions. As a consequence of these two statements we get the probability that a game on a subset of feasible actions admits a totally mixed maximin strategy. Lastly, in Lemma 6.4 we determine the probability that a maximin strategy for a subset of feasible actions is a maximin strategy for the entire set of feasible actions. The probability that a game admits a maximin strategy with a given support may then be derived easily. In this sense the structure of the proof is very similar to McLennan's (2005) argument.

Finally, we argue that symmetric and regular distributions occur naturally. For example, if the payoffs of the game follow independent normal distributions, the distribution of games is symmetric and regular. More generally, every absolutely continuous distribution is regular (cf. Remark 6.3). As noted before, every tournament game admits a unique maximin strategy. Thus, the uniform distribution over all tournament games of a given size is symmetric and regular. As a consequence, Theorem 6.5 implies a result of Fisher and Reeves (1995), who determine the probability that the maximin strategy of a tournament game drawn uniformly at random has support size k.

6.1 THE DISTRIBUTION OF MAXIMIN STRATEGIES

Let $A \in \mathcal{F}(U)$ be some fixed set of feasible actions and $S \subseteq A$ be arbitrary. To simplify the presentation, we introduce notation for particular classes of games. A strategy is *totally mixed* if all feasible actions are played with positive probability. The set of all symmetric proto games for which the game induced by S has a totally mixed maximin strategy is denoted by $\mathcal{M}^S$, i.e.,

$$\mathcal{M}^S = \{M \in \mathcal{M} \colon \text{there is } p \in MS(M, S) \text{ with } \operatorname{supp}(p) = S\}.$$

Let $\bar{\mathcal{M}}^S$ be the set of all symmetric proto games for which the game induced by A admits a maximin strategy with support S, i.e.,

$$\bar{\mathcal{M}}^S = \{M \in \mathcal{M} \colon \text{there is } p \in MS(M, A) \text{ with } \operatorname{supp}(p) = S\}.$$

Since every maximin strategy of a game on A is also a maximin strategy of the game induced by its support, $\bar{\mathcal{M}}^S$ is a subset of $\mathcal{M}^S$. Lastly, the set of symmetric proto games for which the game induced by S has multiple maximin strategies is denoted by $\hat{\mathcal{M}}^S$, i.e.,

$$\hat{\mathcal{M}}^S = \{M \in \mathcal{M} \colon |MS(M, S)| > 1\}.$$

We assume that proto games are drawn from a probability distribution $\mathcal{X}$. By X we denote a random variable with distribution $\mathcal{X}$, i.e., $X \sim \mathcal{X}$. For a set of proto games $\mathcal{M}' \subseteq \mathcal{M}$, let $P_X(\mathcal{M}')$ be the probability that a realization of X is in $\mathcal{M}'$. To establish our results, we require that $\mathcal{X}$ satisfies two regularity conditions. A distribution $\mathcal{X}$ is symmetric on S if it is invariant under ρ_T for every $T \subseteq S$, i.e.,

$$P_X(\mathcal{M}') = P_X(\rho_T(\mathcal{M}')) \text{ for every } T \subseteq S \text{ and } \mathcal{M}' \subseteq \mathcal{M}. \quad \text{(symmetry)}$$

Secondly, we require $\mathcal{X}$ to be regular on A in the sense that X_A almost surely admits a unique maximin strategy. Formally,

$$P_X(\hat{\mathcal{M}}^A) = 0. \qquad \text{(regularity)}$$

The main result Theorem 6.5 states the following: if a proto game M is drawn from a distribution that is symmetric on A and regular on A, then, for every $S \subseteq A$, the probability that M_A admits a maximin strategy with support S is $2^{-(|A|-1)}$ if S has odd cardinality and 0 if S has even cardinality. So if, for example, $A = \{a, b, c\}$ and M is drawn from a distribution that is symmetric on A and regular on A, the distribution over supports of maximin strategies of M_A is as depicted below.

$\{a\}$:	$1/4$	$\{a,b\}$:	0	$\{a,b,c\}$:	$1/4$
$\{b\}$:	$1/4$	$\{a,c\}$:	0		
$\{c\}$:	$1/4$	$\{b,c\}$:	0		

We start by proving an auxiliary lemma, which shows that every strategy that is the unique maximin strategy of some symmetric game assigns positive probability to an odd number of actions. This does not hold for non-symmetric games. Consider the game known as matching pennies, where both players can choose either "heads" or "tails". If both players choose the same action, the first player wins; otherwise the second player wins. The unique maximin strategy of this game is to randomize uniformly over both actions and has thus support 2.

Lemma 6.1
Let $M \in \mathcal{M}$. If $MS(M, A) = \{p\}$, then $\text{supp}(p)$ has odd cardinality.

Proof. Assume for contradiction that $\text{supp}(p)$ has even cardinality. Let $\text{supp}(p) = S$. Since p is the unique maximin strategy of M_A, it follows from Lemma 3.2 that $(p^t M)_i < 0$ for all $i \in A \setminus S$. Now, let $i \in S$ be fixed. By definition of S, $|S \setminus \{i\}|$ is odd. Hence, $M_{S\setminus\{i\}}$ does not have full rank, i.e., there is $x \in \mathbb{R}^U \setminus \{0\}$ with $\text{supp}(x) \subseteq S \setminus \{i\}$ and $(x^t M)_{S\setminus\{i\}} = 0$. Assume without loss of generality that $(x^t M)_i \geqslant 0$ (otherwise take $-x$). Then, for $\epsilon > 0$ small enough, we have that $p^\epsilon = (1-\epsilon)p + \epsilon x \geqslant 0$ and $((p^\epsilon)^t M)_A \geqslant 0$, i.e., $p^\epsilon/|p^\epsilon| \in MS(M, A)$. This contradicts uniqueness of p. $\square$

For distributions that are regular on A, it follows quickly from Lemma 6.1 that the probability that a game has a maximin strategy with even support size is 0. If the distribution is also symmetric on A, it turns out that the probability that a game has a maximin strategy with a given support of odd size is independent of the chosen support. This is again specific to symmetric games and does not hold in general.

LEMMA 6.2

Let $\mathcal{X}$ be symmetric on A and regular on A. Then,

$$P_X(\mathcal{M}^A) = \begin{cases} 0 & \text{if } |A| \text{ is even, and} \\ 2^{-(|A|-1)} & \text{if } n \text{ is odd.} \end{cases}$$

Proof. First consider the case when $|A|$ is even. Let $M \in \mathcal{M}^A$. It follows from Lemma 6.1 that M_A admits multiple maximin strategies. Thus, $\mathcal{M}^A \subseteq \hat{\mathcal{M}}^A$, which implies that $P_X(\mathcal{M}^A) = 0$.

Now assume that $|A|$ is odd. For all $S \subseteq A$, let $\mathcal{M}^{S,=}$ be the set of symmetric proto games such that there is $x \in \mathbb{R}^U$ with $\text{supp}(x) \subseteq A$, $x_+ = A \setminus S$, and $(x^t M)_A = 0$. Note that $\mathcal{M}^{\emptyset,=} = \mathcal{M}^A$. The union of all $\mathcal{M}^{S,=}$, $S \subseteq A$, is $\mathcal{M}$, since a skew-symmetric matrix of odd size cannot have full rank. For all $S \subseteq A$, let $\mathcal{M}^{S,0} \subseteq \mathcal{M}^{S,=}$ be the set of symmetric proto games such that there is $x \in \mathbb{R}^U$ with $\text{supp}(x) \subseteq A$, $x_+ = A \setminus S$, $(x^t M)_A = 0$, and $x_i = 0$ for some $i \in A$. Let $S \subseteq A$, $M \in \mathcal{M}^{S,0}$, and $x \in \mathbb{R}^U$ be the corresponding vector with $x_i = 0$ for some fixed $i \in A$. It follows from Lemma 3.3 that $(\rho_S(x)^t \rho_S(M))_A = 0$. Since $\rho_S(x) \geqslant 0$ and $\rho_S(x)_i = 0$, it follows from Lemma 3.2 that $\rho_S(M) \in \hat{\mathcal{M}}^A$. Thus, $\rho_S(\mathcal{M}^{S,0}) \subseteq \hat{\mathcal{M}}^A$. By symmetry and regularity of $\mathcal{X}$, we then have $P_X(\mathcal{M}^{S,0}) = P_X(\rho_S(\mathcal{M}^{S,0})) \leqslant P_X(\hat{\mathcal{M}}^A) = 0$. This implies that $P_X(\mathcal{M}^{S,=}) = P_X(\rho_{S\Delta T}(\mathcal{M}^{S,=})) = P_X(\mathcal{M}^{T,=})$ for all $S, T \subseteq A$. Moreover, $\mathcal{M}^{S,=}$ and $\mathcal{M}^{A\setminus S,=}$ only differ by a null set, since $x_+ = A \setminus (-x)_+$ if x has no zero entries in A. Hence, $P_X(\mathcal{M}^{S,=}) = P_X(\mathcal{M}^{S,=} \cap \mathcal{M}^{A\setminus S,=})$ for all $S \subseteq A$. Now we show that X_A almost surely has rank $|A|-1$. Since $|A|$ is odd, X_A has rank at most $|A|-1$. If X_A has rank less than $|A|-1$, there are distinct $x, y \in \mathbb{R}^U$ such that $\text{supp}(x) \subseteq A$, $\text{supp}(y) \subseteq A$, and $(x^t X)_A = (y^t X)_A = 0$. But then $((\lambda x + (1-\lambda) y)^t X)_A = 0$ and has an entry equal to 0 for some $\lambda \in \mathbb{R}$. This is a probability zero event as shown above. Hence, X almost surely has rank $|A|-1$. This implies that $P_X(\mathcal{M}^{S,=} \cap \mathcal{M}^{T,=}) = 0$ for all $S, T \subseteq A$ with $S \neq T$ and $S \neq A \setminus T$. In summary, we get $P_X(\mathcal{M}^{S,=}) = 2^{-(|A|-1)}$ for all $S \subseteq A$. □

It was already observed by Kaplansky (1945) that a game of even size cannot have a unique, totally mixed maximin strategy, which follows from the fact that the rank of a skew-symmetric matrix is even. Moreover, Kaplansky (1995) showed that a game admits a unique, totally mixed maximin strategy if and only if the principal Pfaffians of the corresponding payoff matrix alternate in sign.[20] This result allows for a more algebraic but arguably less instructive proof of Lemma 6.2.

Now we show that if a distribution is symmetric and regular on some set of feasible actions, then it is also symmetric and regular on every subset thereof.

LEMMA 6.3

If $\mathcal{X}$ is symmetric on A and regular on A, then $\mathcal{X}$ is also symmetric on S and regular S for every $S \subseteq A$.

20 The ith principal Pfaffian is the Pfaffian of the matrix obtained by deleting the ith row and ith column.

Proof. Let $S \subseteq A$. It follows from the definition of symmetry that X is symmetric on S. Now we show by induction over $|S|$ that $\mathcal{X}$ is regular on S. If $S = A$ this is clear by the hypothesis of the lemma. For the induction step, assume that $\mathcal{X}$ is regular on T for all $T \subseteq A$ with $|T| > |S|$. Assume for contradiction that $\mathcal{X}$ is not regular on S, i.e., $P_X(\hat{\mathcal{M}}^S) > 0$. Let $i \in A \setminus S$ and $S^i = S \cup \{i\}$. We define

$$\mathcal{M}^{S^i,+} = \{M \in \hat{\mathcal{M}}^S \colon \text{there is } p \in MS(M,S) \text{ with } (p^t M)_i \geqslant 0\},$$

with $\mathcal{M}^{S^i,-}$ defined by replacing $\geqslant$ with $\leqslant$. Since $\mathcal{X}$ is symmetric on A, it follows that $P_X(\mathcal{M}^{S^i,+}) = P_X(\rho_{\{i\}}(\mathcal{M}^{S^i,+})) = P_X(\mathcal{M}^{S^i,-})$. Moreover, $\mathcal{M}^{S^i,+} \cup \mathcal{M}^{S^i,-} = \hat{\mathcal{M}}^S$ and hence, $P_X(\mathcal{M}^{S^i,+}) > 0$. Now let $M \in \mathcal{M}^{S^i,+}$. If there is $p \in MS(M,S)$ with $(p^t M)_i = 0$, then it follows from Lemma 3.2 that M_{S^i} has multiple maximin strategies. If $(p^t M)_i > 0$, let $q \in MS(M,S)$ with $q \neq p$, which exists since $M \in \hat{\mathcal{M}}^S$ by assumption. But then $(1-\lambda)p + \lambda q \in MS(M, S^i)$ for small $\lambda > 0$. In any case, M_{S^i} has two distinct maximin strategies. Thus, we have

$$P_X(\hat{\mathcal{M}}^{S^i}) \geqslant P_X(\mathcal{M}^{S^i,+}) \geqslant 1/2\, P_X(\hat{\mathcal{M}}^S) \geqslant 0,$$

which contradicts the induction hypothesis that $\mathcal{X}$ is regular on S^i. □

By combining the last two statements we get the probability that X_S admits a totally mixed maximin strategy. In the next lemma we determine the probability that X_A has a maximin strategy with support S given that X_S has a totally mixed maximin strategy.

Lemma 6.4
Let $\mathcal{X}$ be symmetric on A and regular on A and $S \subseteq A$. Then $P_X(\bar{\mathcal{M}}^S \mid \mathcal{M}^S) = 2^{-(|A|-|S|)}$.

Proof. Let $\mathcal{M}^{S,T}$ be the set of all symmetric proto games where M_S has a totally mixed maximin strategy p such that the set of actions in A yielding negative payoff for the row player against p corresponds exactly to the columns in T, i.e.,

$$\mathcal{M}^{S,T} = \{M \in \mathcal{M} \colon \text{there is } p \in MS(M,S) \text{ with } \operatorname{supp}(p) = S \\ \text{and } (p^t M)_- \cap A = T\}.$$

Note that $\mathcal{M}^{S,\emptyset} = \bar{\mathcal{M}}^S$ and $\mathcal{M}^{S,T}$ is non-empty only if $T \subseteq A \setminus S$. It follows from Lemma 3.3 that $\rho_T(\mathcal{M}^{S,T}) \subseteq \bar{\mathcal{M}}^S$ for all $T \subseteq A \setminus S$. For $M \in \bar{\mathcal{M}}^S \setminus \rho_T(\mathcal{M}^{S,T})$ we have that $(p^t M)_i = 0$ for some $p \in MS(M,S)$ and $i \in T$. Then it follows from Lemma 3.2 that M_A has multiple maximin strategies. Since $\mathcal{X}$ is symmetric on A and regular on A, we have $P_X(\mathcal{M}^{S,T}) = P_X(\rho_T(\mathcal{M}^{S,T})) = P_X(\bar{\mathcal{M}}^S)$ for all $T \subseteq A \setminus S$. For the same reason, we also have that $P_X(\mathcal{M}^{S,T} \cap \mathcal{M}^{S,T'}) = 0$ for all distinct $T, T' \subseteq A \setminus S$. Since $A \setminus S$ has $2^{|A|-|S|}$ distinct subsets, it follows that $P_X(\bar{\mathcal{M}}^S \mid \mathcal{M}^S) = 2^{-(|A|-|S|)}$. □

The main result easily follows from Lemmas 6.2, 6.3, and 6.4.

THEOREM 6.5
Let $\mathcal{X}$ be symmetric on A and regular on A. Then, for every $S \subseteq A$, the probability that X_A has a maximin strategy with support S is

$$\begin{cases} 0 & \text{if } |S| \text{ is even, and} \\ 2^{-(|A|-1)} & \text{if } |S| \text{ is odd.} \end{cases}$$

Observe that A has $2^{|A|-1}$ subsets of odd size. Hence, the probabilities above sum up to 1.

6.2 CONCLUDING REMARKS

We conclude this chapter with a number of remarks.

REMARK 6.1 *(Independence of axioms)*
Both symmetry and regularity are required to derive the conclusion of Theorem 6.5. For $A = \{r, p, s\}$, the distribution that returns the game of "rock, paper, scissors" (cf. Example 3) with probability one is not symmetric on A but regular on A. For this distribution, the uniform distribution over A is the unique maximin strategy with probability one. For $|A| > 1$, the distribution that returns the game with all payoffs equal to 0 with probability one is not regular on A but symmetric on A. For this distribution, all strategies in $\Delta(A)$ are maximin strategies with probability one.

REMARK 6.2 *(Non-symmetric games)*
Theorem 6.5 also fails for symmetric (defined analogously) and regular distributions over not necessarily symmetric games. If $|A| = |B| = 2$ and the entries in the payoff matrix follow independent standard normal random variables, the probability that a maximin strategy has full support is one third.

REMARK 6.3 *(Absolutely continuous distributions)*
Every distribution that is absolutely continuous with respect to the Lebesgue measure is regular. In particular, a distribution is absolutely continuous if all entries in the payoff matrix follow independent and absolutely continuous random variables. This implies that, if the payoffs follow independent and absolutely continuous random variables that are symmetric about 0, e.g., normal random variables or uniform random variables on intervals that are symmetric about 0, then the induced distribution is symmetric and regular.

Part II

PREFERENCE AGGREGATION AND SOCIAL CHOICE

7 SOCIAL CHOICE THEORETIC FUNDAMENTALS

7.1 SOCIAL WELFARE FUNCTIONS AND SOCIAL CHOICE FUNCTIONS

Building up on the foundations for understanding choices by single agents, we introduce a framework for studying choices by groups of agents. The set V denotes the entirety of all agents. For $i \in V$, $\succ_i$ is the preference relation of agent i. We require that every agent's preference relation is from some domain $\mathcal{D} \subseteq \mathcal{R}$ of continuous, convex, and symmetric preference relations on $\Delta(U)$. Hence, $\succ_i$ can be represented by an SSB function $\phi_i = \phi^{\succ_i} \in \Phi$. Finite subsets of agents $N \in \mathcal{F}(V)$ are called electorates. For $N \in \mathcal{F}(V)$, a preference profile on N is a function $P \colon N \to \mathcal{D}$ that assigns a preference relation to every agent, and hence an element of $\mathcal{D}^N$. The set of all preference profiles is $\mathcal{P} = \cup_{N \in \mathcal{F}(V)} \mathcal{D}^N$. The restriction of a preference profile $P \in \mathcal{D}^N$ to a set of outcomes $X \subseteq \Delta(U)$ is obtained by restriction the preference relation of every agent $i \in N$ to X. For $\succ \in \mathcal{D}$, $P(\succ)$ denotes the fraction of agents in N with preference relation $\succ$. For $A \in \mathcal{F}(U)$, $N \in \mathcal{F}(V)$, $P \in \mathcal{D}^N$, and $a \in A$, $\underline{a}$ is a Condorcet winner in $P|_{\Delta(A)}$ if for every other pure outcome in $\Delta(A)$, there is a majority of agents that prefers $\underline{a}$ to this pure outcome. Formally, for all $b \in A \setminus \{a\}$,

$$|\{i \in N \colon \underline{a} \succ_i \underline{b}\}| > |\{i \in N \colon \underline{b} \succ_i \underline{a}\}|. \qquad \text{(Condorcet winner)}$$

A *weak* Condorcet winner is a pure outcome for which the above statement holds with weak inequality instead of strict inequality.

The purpose of social choice theory is to aggregate the preferences of multiple agents into a collective preference relation. This aggregation process is typically formalized via social welfare functions (SWFs). An SWF $f \colon \mathcal{P} \to \mathcal{R}$ maps every preference profile to a collective preference relation. We require collective preferences to satisfy the axioms of SSB utility theory, which is captured by the fact that the range of SWFs is $\mathcal{R}$. We will be particularly interested in SWFs that do not discriminate among agents, i.e., SWFs that are invariant under renaming the agents. This property is called anonymity and prescribes that, for all $N, \hat{N} \in \mathcal{F}(V)$ with $|N| = |\hat{N}|$, $P \in \mathcal{D}^N$, and $\hat{P} \in \mathcal{D}^{\hat{N}}$,

$$f(P) = f(\hat{P}) \text{ whenever } P(\succ) = \hat{P}(\succ) \text{ for all } \succ \in \mathcal{D}. \qquad \text{(anonymity)}$$

The choices of society from feasible sets can be derived from the collective preferences, assuming that maximal elements are chosen. In order to reason about these choice independently from collective preferences, we consider social choice functions (SCFs), which map a preference profile and a feasible set of outcomes to a set of collectively chosen outcomes with the self-evident restriction that only feasible outcomes can be chosen. We restrict the set of feasible sets of outcomes to sets of the form $\Delta(A)$ where $A \in \mathcal{F}(U)$. This is based on the rationale that every alternative is either feasible or infeasible, and in case it is feasible, the probability assigned to it can be arbitrary. Hence, feasibility of an outcome boils down to feasibility of the alternatives in its support. An SCF is thus a function $f\colon \mathcal{P} \times \mathcal{F}(U) \to 2^{\Delta(U)} \setminus \{\emptyset\}$ with the property $f(P, A) \subseteq \Delta(A)$ for all $P \in \mathcal{P}$ and $A \in \mathcal{F}(U)$.

Anonymity can be defined for SCFs in the same way as for SWFs. A stronger invariance property, known as homogeneity, is invariance under replacing every agent by a fixed number of copies with the same preferences. Hence, homogeneity requires that the choice only depends on the fractions of agents that report a particular preference relation and not the absolute numbers. Formally, an SCF f satisfies homogeneity if, for all $P, \hat{P} \in \mathcal{P}$ and $A \in \mathcal{F}(U)$,

$$f(P, A) = f(\hat{P}, A) \text{ whenever } P(\succ) = \hat{P}(\succ) \text{ for all } \succ \in \mathcal{D}. \tag{homogeneity}$$

The fact that homogeneity implies anonymity is obvious from the definition. In the classical model of social choice, the set of outcomes is assumed to be some abstract, unstructured set of alternatives and typically, transitivity is the only restriction on individual preferences. Since the SSB axioms do not restrict the preferences over pure outcomes, this model can be embedded in ours by considering SCFs that only depend on the preferences over pure outcomes and choose all outcomes over some subset of feasible alternatives; pure outcomes in this subset correspond to chosen alternatives. An SCF f is pure if, for all $A \in \mathcal{F}(U)$, $N \in \mathcal{F}(V)$ and $P, \hat{P} \in \mathcal{D}^N$,

$$\begin{aligned} &f(P, A) = f(\hat{P}, A) \text{ whenever } P|_{\underline{A}} = \hat{P}|_{\underline{A}}, \text{ and} \\ &f(P, A) = \Delta(B) \text{ for some } B \subseteq A. \end{aligned} \tag{pure SCF}$$

Pure SCFs are sometimes assumed to be resolute, which requires that a single, pure outcome is chosen in all instances.

Particular classes of SWFs and SCFs that we are interested in are those which are affine welfare maximizing. Since the preferences of the agents can be represented by SSB functions, we can compare the collective utility of two outcomes to each other given some weighting of the agents. A preference relation $\succ \in \mathcal{R}$ is affine welfare maximizing for some profile $P \in \mathcal{D}^N$, $N \in \mathcal{F}(V)$, if there are weights

$w_i \in \mathbb{R}, i \in N$, such that $\phi^{\succ} \equiv w_i \sum_{i\in N} \phi_i$. An SWF f is affine welfare maximizing if there are weights $w_i \in \mathbb{R}$, $i \in V$, such that, for all $N \in \mathcal{F}(V)$ and $P \in \mathcal{D}^N$, $\phi^{f(P)} \equiv \sum_{i\in N} w_i\phi_i$. Note that the weights of the agents are fixed across all preference profiles. For the case when the agents' preferences satisfy the vNM axioms of linear utility theory, affine welfare maximization has been well-studied. In particular, the SWF that derives the collective preferences from adding up the vNM functions representing the agents' preferences normalized to the unit interval is known as relative utilitarianism (*RU*) (Dhillon, 1998; Dhillon and Mertens, 1999).[21] Since SSB functions in Φ are normalized such that the largest utility difference between two outcomes is 1, affine utilitarianism with weight 1 for all agents coincides with relative utilitarianism if the agents have vNM preferences. Hence, we feel justified in extending relative utilitarianism to SSB preferences in the following way. For all $N \in \mathcal{F}(V)$ and $P \in \mathcal{D}^N$,

$$\phi^{RU(P)} \equiv \sum_{i\in N} \phi_i. \qquad \text{(relative utilitarianism)}$$

By $M^P \in \mathcal{M}$ we denote the matrix representing $\phi^{RU(P)}$.

The corresponding definitions for SCFs are obtained by choosing maximal outcomes. An outcome $p \in \Delta(U)$ is affine welfare maximizing for a preference profile $P \in \mathcal{P}$ and a feasible set $A \in \mathcal{F}(U)$ if there is an affine welfare maximizing preference relation $\succ \in \mathcal{R}$ for P such that $p \in \max_{\succ} \Delta(A)$. An SCF f is affine welfare maximizing if there is an affine welfare maximizing SWF g such that, for every $A \in \mathcal{F}(U)$ and $P \in \mathcal{P}$, $f(P, A) \subseteq \max_{g(P)} \Delta(A)$. It is easy to see that every outcome that is affine welfare maximizing for positive weights is Pareto optimal with respect to the individual preferences within the respective feasible set. Conversely, it follows from Aziz et al. (2015, Theorem 1) that Pareto optimality is not only necessary but also sufficient for an outcome to be welfare maximizing (see also McLennan, 2002; Manea, 2008; Athanassoglou, 2011; Carroll, 2010; Dogan and Yildiz, 2016, for similar results for vNM preferences).

Since collective preferences are assumed to be convex, choosing maximal outcomes according to an arbitrary SWF satisfies Sen's α and Sen's γ (cf. Section 2.1). This is remarkable, since these and similar choice consistency conditions have been shown to be prohibitive for pure SCFs when combined with assumptions like non-dictatorship and Pareto optimality (cf. Chapter 8).

21 In fact relative utilitarianism also obtains for every normalization that differs from normalization to the unit interval only by an additive constant.

7.2 MAXIMAL LOTTERIES

On the domain of PC preferences, the SCF that returns the maximal outcomes according to relative utilitarianism is known as maximal lotteries (*ML*). For all $A \in \mathcal{F}(U)$, $N \in \mathcal{F}(V)$, and $P \in (\mathcal{D}^{PC})^N$,

$$ML(P, A) = \max_{RU(P)} \Delta(A). \qquad \text{(maximal lotteries)}$$

ML was first considered by Kreweras (1965) and studied in more detail by Fishburn (1984b) and Aziz et al. (2013). *ML* is Condorcet consistent as it uniquely returns a Condorcet winner whenever one exists. In general, an outcome returned by *ML* is preferred to all other outcomes by an expected majority of agents. Thus, *ML* can be seen as an extension of Condorcet's principle to all preference profiles. The outcomes of $ML(P, A)$ correspond to maximin strategies in the symmetric zero-sum game M_A^P induced by the pairwise majority margins. This implies that maximal lotteries can be computed efficiently via linear programming. Laffond et al. (1997) have shown that every symmetric zero-sum game with odd off-diagonal payoffs admits a unique maximin strategy. Thus, $ML(P, A)$ is a singleton whenever there is an odd number of agents with strict preferences over pure outcomes. Moreover, the set of symmetric zero-sum games with multiple maximin strategies $\hat{\mathcal{M}}_A^A$ is nowhere dense in $\mathcal{M}_A$ and has measure zero. Hence, $ML(P, A)$ is almost always single-valued in a well-defined sense.

8 ARROVIAN PREFERENCE AGGREGATION

> Many writers have felt that the assumption of rationality, in the sense of a one-dimensional ordering of all possible alternatives, is absolutely necessary for economic theorizing [...] There seems to be no logical necessity for this viewpoint; we could just as well build up our economic theory on other assumptions as to the structure of choice functions if the facts seemed to call for it.
>
> K. J. Arrow

Arrow's impossibility theorem (Arrow, 1951) states that every SWF that satisfies Pareto optimality and independence of irrelevant alternatives is dictatorial. Two important modeling assumptions are that individual preferences are complete and transitive but otherwise unrestricted and that collective preferences are also complete and transitive. Arrow's result has triggered a large amount of work about possible ways to circumvent its negative implications for collective decision making. With few exceptions, these attempts failed in the sense that they produced new impossibility results. They can be divided into two categories based on the aforementioned modeling assumptions.

The starting point for the first approach is the observation by Sen (1970a) that Arrow's assumption of transitivity of the collective preference relation is not necessary to guarantee the existence of maximal elements within finite feasible sets. Hence, one possible escape route from Arrow's theorem is to weaken the assumption of *transitive collective preferences*. Sen (1969) has shown that weakening transitivity to acyclicity, which is necessary and sufficient for the existence of maximal elements, allows for non-dictatorial SWFs satisfying Arrow's axioms. These SWFs do not constitute a proper resolution, however, since they are dictatorial in a weaker, but still highly undesirable, way. A number of similar results for acyclic collective preferences have been obtained for variants of Arrow's conditions (see, e.g., Mas-Colell and Sonnenschein, 1972; Brown, 1975; Blau and Deb, 1977; Blair and Pollak, 1982; Banks, 1995). (For an overview of results on weakened assumptions about collective preferences we refer to Kelly (1978), Sen (1977), Sen (1986), Schwartz (1986), and Campbell and Kelly (2002).)

The second approach weakens Arrow's assumption of *full domain of individual preferences*. Promising results have been obtained for domains of single-peaked and dichotomous preferences, which allow for attractive SWFs (see, e.g., Black, 1948; Arrow, 1951; Inada, 1969; Sen and Pattanaik, 1969; Ehlers and Storcken, 2008). They rely on the fact that the majority relation is transitive within these domains. Single-peakedness prescribes that the set of outcomes lies on a line and the individual preferences have convex upper contour sets. In a similar way, domain restrictions can be obtained for higher-dimensional convex sets of outcomes. Preferences over these sets are typically assumed to respect the structure of the outcome set by assuming that they satisfy some notion of convexity and continuity (cf. Chapter 7). Samuelson (1967) conjectured that Arrow's impossibility still holds when individual and collective preferences over lotteries on alternatives satisfy the vNM axioms, which was later proven to be the case by Kalai and Schmeidler (1977a) if there are at least four alternatives and the SWF is continuous. Hylland (1980a) showed that continuity is not needed to derive this conclusion. Similar results for restricted preferences have been obtained by Kalai et al. (1979), Border (1983), Bordes and Le Breton (1989), Bordes and Le Breton (1990a), Bordes and Le Breton (1990b), Campbell (1989), and Redekop (1995). Hence, the positive results for single-peaked preferences crucially rely on the one-dimensional structure of the outcome space.

The approach taken here is based on the observation that all of the results in the second category assume some notion of transitivity of preferences. For convex outcome sets, such an assumption is not necessary to guarantee the existence of maximal elements within convex feasible sets, however (cf. Proposition 2.1). We consider SWFs that map a profile of SSB preferences to a collective preference relation, which is also assumed to satisfy the SSB axioms. First, we show that not only does this setting allow for non-dictatorial SWFs satisfying Arrow's axioms of Pareto optimality and independence of irrelevant alternatives, even anonymous Arrovian aggregation is possible. Curiously, the unique inclusion-maximal Cartesian domain which allows for anonymous Arrovian aggregation is precisely the domain of preferences based on pairwise comparisons $\mathcal{D}^{PC}$. In contrast to single-peaked and dichotomous preferences, this domain does not restrict the preferences over pure outcomes. We go on to show that every Arrovian SWF on $\mathcal{D}^{PC}$ is affine utilitarian. This result even holds when only assuming Pareto indifference, i.e., Pareto optimality with respect to the indifference relation (cf. Harsanyi, 1955). When additionally assuming anonymity, our axioms uniquely characterize relative utilitarianism. This implies that the collective preferences over pure outcomes coincide with the majority relation.

Our second result is related to Harsanyi's *social aggregation theorem* (Harsanyi, 1955), which shows that, for individual and collective vNM preferences, every SWF satisfying Pareto indifference has to derive the collective preferences from a linear combination of the agents' vNM functions. The weights assigned to the agents may depend on the individual preferences, however. Hence, Harsanyi's theorem does not characterize affine utilitarianism as defined in Chapter 7.1. Fishburn and Gehrlein (1987) demonstrated that Harsanyi's theorem cannot be extended to SSB preferences, even when strengthening the notion of Pareto optimality (see also Turunen-Red and Weymark, 1999). The characterization of affine utilitarianism given here shows that Pareto indifference is sufficient to enforce affine utilitarianism for preferences based on pairwise comparisons when additionally assuming independence of irrelevant alternatives. A multi-profile version of Harsanyi's theorem for social welfare *functionals*, i.e., functions that map a profile of vNM functions to a vNM preference relation, was shown by Mongin (1994). He proved that every social welfare functional that satisfies Pareto optimality and IIA is affine utilitarian. When considering social welfare functionals that are invariant under positive affine transformations of the agents' vNM functions, one again obtains an SWF and the characterization of affine utilitarianism turns into an impossibility result, since affine utilitarianism does not satisfy independence of irrelevant alternatives for vNM preferences.

Related results have been obtained by Dhillon (1998), Dhillon and Mertens (1999), and Börgers and Choo (2015), who characterized relative utilitarianism for vNM preferences. The results of Dhillon (1998) and Börgers and Choo (2015) are based on a Pareto-type axiom that allows to apply Harsanyi's theorem (or similar results), while Dhillon and Mertens (1999) use a quite technical monotonicity axiom. A strengthening of their monotonicity axiom prescribes that if an agent changes his preferences between two outcomes from indifference to a preference for the collectively preferred outcome, then the collective preference between these two outcomes should not change. Compared to our result, Dhillon (1998) and Dhillon and Mertens (1999) require a weaker version of independence of irrelevant alternatives called independence of redundant alternatives, which demands the consequence of independence of irrelevant alternatives only for feasible sets that make all other outcomes redundant in that they are unanimously indifferent to some feasible outcome. Unlike independence of irrelevant alternatives, independence of redundant alternatives is satisfied by relative utilitarianism on the domain of vNM preferences. The axioms of Börgers and Choo (2015) that allow them to extend Harsanyi's single-profile utilitarianism to affine utilitarianism are formulated in terms of marginal rates of substitution based on the agents' vNM functions. All three results use anonymity to infer relative utilitarianism from affine utilitarianism.

8.1 ARROVIAN SOCIAL WELFARE FUNCTIONS

For the rest of this chapter, we assume that the set of alternatives U is finite (cf. Remark 8.7 for the case of infinitely many alternatives). Since none of the properties of SWFs considered here connects variable sets of agents to each other, we fix an electorate $N \in \mathcal{F}(V)$ with $n = |N| \geqslant 2$. Arrow (1951) showed that the only SWFs that satisfy Pareto optimality and independence of irrelevant alternatives are dictatorial functions when the domain $\mathcal{D}$ of preferences contains all transitive and complete preference relations over outcomes and collective preferences have to be transitive and complete. In contrast, we assume that individual preferences are from some domain $\mathcal{D} \subseteq \mathcal{R}$ and collective preferences are from $\mathcal{R}$. Pareto optimality prescribes that a unanimous preference of one outcome over another in the individual preferences should be reflected likewise in the collective preferences. An SWF f satisfies Pareto optimality if, for all $p, q \in \Delta(U)$, $P \in \mathcal{D}^N$, and $f(P) = \succ$,

$$p \succsim_i q \text{ for all } i \in N \text{ implies } p \succsim q, \text{ and}$$
$$\text{if additionally } p \succ_i q \text{ for some } i \in N \text{ then } p \succ q. \qquad \text{(Pareto optimality)}$$

The indifference part of Pareto optimality, which merely requires that $p \sim_i q$ for all $i \in N$ implies $p \sim q$, is usually referred to as *Pareto indifference*.

Independence of irrelevant alternatives demands that collective preferences over some feasible set of outcomes should only depend on the individual preferences over this set (and not on the preferences over outcomes outside of this set). Since outcomes are probability distributions over alternatives, our notion of feasible sets is based on the availability of alternatives. To this end, we consider the same notion of feasible sets as introduced for SCFs in Section 7.1 (see also Kalai and Schmeidler, 1977a). Hence, feasible sets are assumed to take the form $\Delta(A)$ for some $A \in \mathcal{F}(U)$. Formally, we say that an SWF f satisfies independence of irrelevant alternatives if, for all $P, \hat{P} \in \mathcal{D}^N$ and $A \in \mathcal{F}(U)$,

$$P|_{\Delta(A)} = \hat{P}|_{\Delta(A)} \text{ implies } f(P)|_{\Delta(A)} = f(\hat{P})|_{\Delta(A)}. \qquad \text{(IIA)}$$

Stronger notions of IIA for less restrictive assumptions about feasible sets are discussed in Remark 8.3.

Any SWF that satisfies Pareto optimality and IIA will be called an *Arrovian* SWF. Formulated in our framework, Arrow has shown that every Arrovian SWF is dictatorial on pure outcomes, i.e., there is $i \in N$ such that for all $a, b \in U$, $P \in \mathcal{D}^N$, and $f(P) = \succ$, $\underline{a} \succ_i \underline{b}$ implies $\underline{a} \succ \underline{b}$, when individual and collective preferences over pure outcomes have to be transitive and complete. Anonymity as

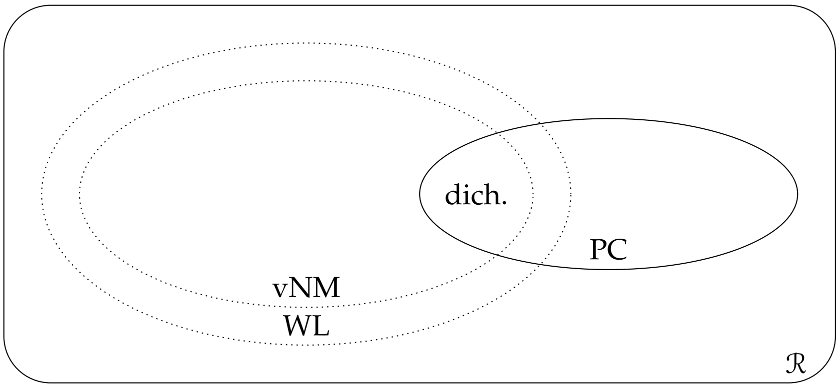


Figure 8.1: Venn diagram showing the inclusion relationships between preference domains. The intersection of the domain of vNM preferences and the domain of PC preferences exactly contains the set of dichotomous vNM preference relations. The intersection of WL preferences and PC preferences additionally contains the PC preferences based on trichotomous weak orders (see Figure 2.1 for an example). An example of PC preferences not contained in the set of WL preferences is given in Figure 2.2. Theorem 8.1 shows that the domain of PC preferences is the unique inclusion-maximal domain within $\mathcal{R}$ for which anonymous Arrovian aggregation is possible. This, for example, implies impossibilities for WL preferences and vNM preferences.

defined in Section 7.1 is obviously a stronger requirement than non-dictatorship on pure outcomes.

In order to prove our characterization, we need to assume that the domain $\mathcal{D} \subseteq \mathcal{R}$ satisfies certain richness conditions. First, we require that it is neutral in the sense that it is not biased towards certain alternatives. It is assumed that $\succ \in \mathcal{D}$ if and only if $\succ^\pi \in \mathcal{D}$ for all $\pi \in \Pi(U)$ and $\succ \in \mathcal{D}$. Secondly, it should also be possible for agents to declare completely opposed preferences, i.e., that $\mathcal{D}$ is closed under reversals. Formally, $\succ \in \mathcal{D}$ implies $\succ^{-1} \in \mathcal{D}$ for all $\succ \in \mathcal{R}$. Note that this condition is not implied by the previous neutrality condition because it allows the reversal of preferences over all outcomes, not only pure outcomes. Finally, we demand that for every transitive relation on *pure* outcomes, $\mathcal{D}$ contains at least one extension of this relation to all outcomes in $\Delta(U)$. The last assumption can be slightly weakened without affecting the correctness of our proofs (see Remark 8.6). For the rest of this chapter we assume that $\mathcal{D}$ satisfies the three conditions defined above.

8.2 CHARACTERIZATION OF THE DOMAIN

It follows from previously mentioned results that non-dictatorial Arrovian aggregation is impossible for vNM preferences. On the other hand, appealing SWFs exist in subdomains such as dichotomous vNM preferences where each agent can only assign two different utility values. In this domain, every affine utilitarian SWF with positive weights satisfies IIA and Pareto optimality. Note that vNM preferences with only two different utility values also constitute PC preferences, where every pure outcome with the higher utility value is preferred to every pure outcome with the lower utility value. This possibility banks on the fact that the majority relation on pure outcomes is transitive for dichotomous preferences. The only anonymous Arrovian SWF on this domain corresponds to approval voting and ranks pure outcomes by the number of approvals they receive from the agents. This ranking is identical to the majority relation.

Theorem 8.1 encompasses both, the negative result for vNM preferences and the positive result for dichotomous preferences, by showing that $\mathcal{D}^{PC}$ is the unique inclusion-maximal domain which allows for anonymous Arrovian aggregation.[22]

Theorem 8.1

Let $|U| \geqslant 4$ and f be an anonymous Arrovian SWF on some domain $\mathcal{D} \subseteq \mathcal{R}$. Then, $\mathcal{D} \subseteq \mathcal{D}^{PC}$.

The proof of Theorem 8.1 is given in Section 8.6. Figure 8.1 illustrates the implications of Theorem 8.1.

8.3 CHARACTERIZATION OF THE SOCIAL WELFARE FUNCTION

Theorem 8.1 has established that anonymous Arrovian aggregation is only possible if individual preferences are based on pairwise comparisons. Theorem 8.2 now shows that all Arrovian SWFs on domains of PC preferences are affine utilitarian with positive weights. Independence of irrelevant alternatives together with Pareto indifference already characterizes affine utilitarianism. Pareto optimality forces the weights to be positive, which also excludes dictatorships, i.e., affine utilitarianism with all weights except for one equal to 0.

Theorem 8.2

Let $|U| \geqslant 5$ and f be an Arrovian SWF on some domain $\mathcal{D} \subseteq \mathcal{D}^{PC}$. Then, there are $w_1, \dots, w_n \in \mathbb{R}_{>0}$ such that

$$\phi^{f(P)} \equiv \sum_{i \in N} w_i \phi_i \text{ for all } P \in \mathcal{D}^N.$$

The proof of Theorem 8.2 is given in Section 8.7. Theorem 8.2 can be seen as a multi-profile version of Harsanyi's social aggregation theorem for PC preferences, where IIA allows us to connect coefficients

22 Our domain assumptions require that $\mathcal{D}$ contains a preference relation with a strict order of all pure outcomes. Theorem 8.1 also holds for domains of dichotomous preferences, however (cf. Remark 8.6).

$$\phi = \begin{array}{c} \\ \\ \\ \\ \end{array}\!\!\begin{array}{cc} \begin{array}{ccc} a & b & c \end{array} & \\ \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} & \begin{array}{c} a \\ b \\ c \end{array} \end{array}$$

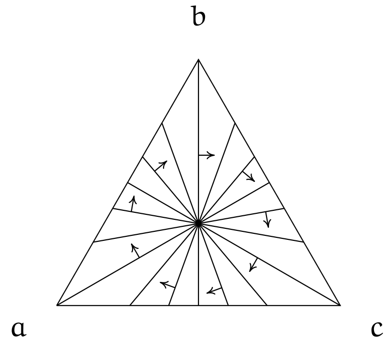


Figure 8.2: Illustration of collective preferences returned by the unique anonymous Arrovian SWF in the case of Condorcet's paradox. The left-hand side shows the collective SSB function and the right-hand side the Marschak-Machina probability triangle for the corresponding collective preferences. The arrows represent the normal vectors to the indifference curves of the outcome at the base of the arrow (pointing towards the lower contour set). Each indifference curve separates the corresponding upper and lower contour set. The unique maximal outcome is $1/3\,\underline{a} + 1/3\,\underline{b} + 1/3\,\underline{c}$.

across different profiles. When furthermore assuming anonymity, the weights of all agents have to be identical and we obtain the following characterization of relative utilitarianism.

Corollary 8.3
Let $|U| \geqslant 5$ and f be an anonymous Arrovian SWF on $\mathcal{D} \subseteq \mathcal{R}$. Then, $f = RU$.

Relative utilitarianism is computationally tractable: two outcomes can be compared by straightforward matrix-vector multiplications while maximal outcomes within feasible sets can be found using linear programming. For illustrative purposes, let $U = \{a, b, c\}$ and $N = \{1, 2, 3\}$ and consider the classic Condorcet example where the agents have the following transitive preferences over pure outcomes: $\underline{a} \succ_1 \underline{b} \succ_1 \underline{c}$, $\underline{b} \succ_2 \underline{c} \succ_2 \underline{a}$, and $\underline{c} \succ_3 \underline{a} \succ_3 \underline{b}$. The corresponding PC preferences $\succ_1, \succ_2, \succ_3$ are represented by $\phi_1, \phi_2, \phi_3 \in \Phi^{PC}$, where

$$\phi_1 = \begin{array}{cc} \begin{array}{ccc} a & b & c \end{array} & \\ \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix} & \begin{array}{c} a \\ b \\ c \end{array} \end{array}, \quad \phi_2 = \begin{array}{cc} \begin{array}{ccc} a & b & c \end{array} & \\ \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} & \begin{array}{c} a \\ b \\ c \end{array} \end{array}, \text{ and}$$

$$\phi_3 = \begin{array}{cc} \begin{array}{ccc} a & b & c \end{array} & \\ \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix} & \begin{array}{c} a \\ b \\ c \end{array} \end{array}.$$

In the preference profile $P = (\succ_1, \succ_2, \succ_3)$ the pairwise majority relation over pure outcomes is cyclic, since there are majorities for $\underline{a}$ over $\underline{b}$, $\underline{b}$ over $\underline{c}$, and $\underline{c}$ over $\underline{a}$. Relative utilitarianism aggregates preferences by adding the individual SSB representations, i.e.,

$$\phi^{RU(P)} \equiv \sum_{i\in N} \phi_i = \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ a \\ b \\ c \end{array} .$$

Figure 8.2 shows the collective preference relation represented by this matrix. The unique maximal outcome is $1/3\,\underline{a} + 1/3\,\underline{b} + 1/3\,\underline{c}$.[23]

8.4 INTERPRETATION OF RESULTS

Theorem 8.1 characterizes the domain of PC preferences as the largest domain for which anonymous Arrovian aggregation is possible. In light of many impossibility results in the context of Arrovian aggregation, the existence of such a domain is surprising. To shed some light on its characteristics, observe that it generalizes the domain of dichotomous vNM preferences in the sense that it only allows for one intensity of preference when comparing pure outcomes. In particular, the preferences over all outcomes are completely determined by the preferences over pure outcomes. This implies that, whenever the preferences over some set of pure outcomes coincide for two preference profiles, then the preferences over all outcomes in the convex hull of these pure outcomes also coincide. Moreover, the numerical SSB value for every pair of outcomes in this convex hull has to be the same for all agents in both profiles. The latter fact weakens the force of independence of irrelevant alternatives and allows it to be satisfied by affine utilitarianism. Thus, a possible interpretation of Theorem 8.1 is that Arrow's axioms deny different preference intensities among pure outcomes. They also force individual preferences to be intransitive even when preferences over pure outcomes are transitive (cf. Figure 2.2). To support the arguments we put forward in defense of intransitive preferences in Section 1.3 and Section 2.2, we refer to the following quote by Peter Fishburn:

> Transitivity is obviously a great practical convenience and a nice thing to have for mathematical purposes, but long ago this author ceased to understand why it should be a cornerstone of normative decision theory. [...] The presence of intransitive preferences complicates matters [...]

23 This outcome represents a somewhat unusual unique maximal outcome because it is not *strictly* preferred to any of the other outcomes. This is due to the contrived nature of the example and only happens if the support of a maximal outcome contains all alternatives.

> however, it is not cause enough to reject intransitivity. An analogous rejection of non-Euclidean geometry in physics would have kept the familiar and simpler Newtonian mechanics in place, but that was not to be. Indeed, intransitivity challenges us to consider more flexible models that retain as much simplicity and elegance as circumstances allow. It challenges old ways of analyzing decisions and suggests new possibilities. (Fishburn, 1991, pp. 115–117)

Theorem 8.2 is closer to Harsanyi's social aggregation theorem. It shows that Pareto optimality forces the collective preferences to be based on affine utilitarianism on the domain of PC preferences. In contrast to Harsanyi's theorem, this conclusion only holds in the presence of independence of irrelevant alternatives. It may be questioned if affine utilitarianism on the domain of PC preferences constitutes proper utilitarianism. It is in fact no more utilitarian than approval voting on the domain of dichotomous preferences. On the other hand, it is no less utilitarian than the form of utilitarianism characterized by Harsanyi. For one, even though PC preferences over pure outcomes cannot have different intensities, preferences over other outcomes *can* vary in intensity. Secondly, even in Harsanyi's case of vNM preferences, preferences are of ordinal nature despite the fact that they admit a numerical representation. The following quote of John von Neumann and Oskar Morgenstern elaborates on this point:

> It is clear that every measurement or rather every claim of measurability must ultimately be based on some immediate sensation, which possibly cannot and certainly need not be analyzed any further. In the case of utility the immediate sensation of preference of one object or aggregate of objects as against another provides this basis. But this permits us only to say when for one person one utility is greater than another. It is not in itself a basis for numerical comparison of utilities for one person nor of any comparison between different persons. Since there is no intuitively significant way to add two utilities for the same person, the assumption that utilities are of non-numerical character even seems plausible. The modern method of indifference curve analysis is a mathematical procedure to describe this situation. (von Neumann and Morgenstern, 1953, p. 16)

Finally, Corollary 8.3 implies that anonymous Arrovian preference aggregation entails that one has to be willing to accept intransitive collective preferences, even over pure outcomes. More precisely, the collective preferences over pure outcomes need to be in accordance with the majority relation. In this vein, Corollary 8.3 combines Borda's score-based and Condorcet's majoritarian conception of preference aggregation.

8.5 CONCLUDING REMARKS

We conclude this chapter with a number of technical remarks.

REMARK 8.1 *(Transitivity)*
When also requiring transitivity of individual preferences, Theorem 8.1 immediately turns into an impossibility, which follows from the fact that PC preferences are not transitive for at least four alternatives (cf. Figure 2.2). This implies the impossibility of anonymous Arrovian aggregation of vNM preferences (and thereby of WL preferences), even when collective preferences need not be transitive.[24]

REMARK 8.2 *(Anonymity)*
Theorem 8.1 does not hold without assuming anonymity. Let $U = \{a, b, c, d\}$, $N = \{1, 2, 3\}$, and $\epsilon \in (0, 1/4)$ and consider the SSB function

$$\phi = \begin{pmatrix} 0 & 1 & 1 & 1+\epsilon \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -(1+\epsilon) & -1 & -1 & 0 \end{pmatrix}.$$

Let $\mathcal{D} = \mathcal{D}^{PC} \cup \{\succ \in \mathcal{R} \colon \phi^{\succ} \equiv \phi_\pi \text{ for some } \pi \in \Pi(U)\}$. Then $\mathcal{D}$ satisfies all our domain assumptions (cf. Section 8.1). The SWF $f\colon \mathcal{D}^N \to \mathcal{R}$, $\phi^{f(P)} \equiv 2\phi_1 + 3\phi_2 + 4\phi_3$ satisfies IIA and Pareto optimality but violates anonymity. Note that f is not dictatorial (not even on pure outcomes). Hence, Theorem 8.1 does not hold when weakening anonymity to non-dictatorship.

REMARK 8.3 *(Strong IIA)*
Relative utilitarianism does not satisfy the stronger notion of IIA that considers all non-empty, compact, and convex sets feasible. To see this, let $U = \{a, b, c\}$ and $N = \{1, 2\}$ and consider the preference relations $\succ_1, \succ_2 \in \mathcal{R}$ represented by the SSB functions

$$\phi_1 = \begin{pmatrix} 0 & 3 & 4 \\ -3 & 0 & 1 \\ -4 & -1 & 0 \end{pmatrix} \quad \text{and} \quad \phi_2 = \begin{pmatrix} 0 & 1 & 3 \\ -1 & 0 & 2 \\ -3 & -2 & 0 \end{pmatrix}.$$

Then, for $p = 1/2\, \underline{a} + 1/2\, \underline{c}$ and $q = \underline{b}$, we have $p \succ_1 q$ and $q \succ_2 p$. For the profiles $P = (\succ_1, \succ_2)$ and $\hat{P} = (\succ_2^{-1}, \succ_1^{-1})$, we have $P|_{\mathrm{conv}(\{p,q\})} = \hat{P}|_{\mathrm{conv}(\{p,q\})}$ but $p\ RU(P)\ q$ and $q\ RU(\hat{P})\ p$.

REMARK 8.4 *(Symmetry)*
Theorem 8.1 also holds when collective preferences are not required to satisfy the symmetry axiom. Whether symmetry is

24 When collective preferences have to be transitive as well, this impossibility directly follows from Arrow's theorem as mentioned in Section 8.1, since IIA, Pareto optimality, and anonymity imply IIA, Pareto optimality, and non-dictatorship on pure outcomes, respectively.

required for individual preferences in Theorem 8.1 and for collective preferences in Theorem 8.2 is open.

REMARK 8.5 *(Tightness of bounds)*
Theorem 8.1 does not hold if $|U| < 4$, which is the same bound as for the result by Kalai and Schmeidler (1977b). This stems from the fact that for $U = \{a, b, c\}$, IIA only has non-trivial implications for feasible sets of the form $\Delta(\{x, y\})$ for some $x, y \in U$. For every possible preference over x and y, there is exactly one continuous and convex preference relation on $\Delta(\{x, y\})$ consistent with it. Hence, IIA only has non-trivial implications for the collective preferences over pure outcomes. However, even for three alternatives, the domains of preferences which allow for anonymous Arrovian aggregation are severely restricted. In particular, Lemmas 8.8, 8.9, 8.10, and 8.11 still hold. Any such domain contains exactly one SSB preference relation $\succ$ for every strict order over U such that

$$\phi^{\succ} \equiv \begin{pmatrix} 0 & 1 & \lambda \\ -1 & 0 & 1 \\ -\lambda & -1 & 0 \end{pmatrix}$$

for some $\lambda \in \mathbb{R}_{>0}$ that is fixed across all strict orders. For $1 < \lambda < 1 + 1/n$, relative utilitarianism constitutes an Arrovian SWF on the corresponding domain.

Theorem 8.2 does not hold if $|U| < 5$. Let $U = \{a, b, c, d\}$, $\mathcal{D} = \mathcal{D}^{PC}$, and $\hat{P} = (\hat{\succsim}_1, \hat{\succsim}_2, \hat{\succsim}_3, \hat{\succsim}_4, \dots)$ such that every SSB preference in $\mathcal{D} \setminus \{\emptyset\}$ appears exactly once in the preferences of the agents in $N \setminus \{1, 2, 3, 4\}$ and

$$\hat{\phi}_1 = \hat{\phi}_2 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix},$$

$$\hat{\phi}_3 = \begin{pmatrix} 0 & 1 & -1 & -1 \\ -1 & 0 & -1 & -1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & 0 \end{pmatrix}, \hat{\phi}_4 = \begin{pmatrix} 0 & 1 & -1 & -1 \\ -1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

Then, Pareto optimality has no implications for $\hat{P}$. Let $f\colon \mathcal{D}^N \to \mathcal{R}$, $\phi^{f(P)} \equiv \sum_{i \in N} \phi_i$ except that

$$\phi^{f(\hat{P})} \equiv \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Then, f satisfies Pareto optimality and IIA but is not affine utilitarian. The proof of Theorem 8.2 fails at Lemma 8.14.

Remark 8.6 *(Domain assumptions)*
In Section 8.1, we specified domain richness conditions that are required for our proofs. The last-named of those conditions prescribes that every transitive relation on pure outcomes is induced by at least one relation in $\mathcal{D}$. For Theorems 8.1 and 8.2, this condition is only required for every transitive relation on *four* and *five* pure outcomes, respectively. Furthermore, to derive the conclusion of Theorem 8.1, a weaker condition suffices: if $\succ \in \mathcal{D}$ with $\underline{a} \succ \underline{b} \succ \underline{c}$ and $\underline{a} \succ \underline{c}$ for some $a, b, c \in U$, then there is some $\hat{\succsim} \in \mathcal{D}$ with $\underline{a} \mathrel{\hat{\succsim}} \underline{b} \mathrel{\hat{\succsim}} \underline{c} \mathrel{\hat{\succsim}} \underline{x}$ and $\underline{a} \mathrel{\hat{\succsim}} \underline{c}$ for some $x \in U$. This condition also covers the domain of dichotomous preferences.

Remark 8.7 *(Infinite Universes)*
Fishburn (1984a) shows that under additional technical assumptions about the outcome space and the preference relations, the SSB representation holds for probability measures over arbitrary (possibly infinite) sets of alternatives. Our results extend to this framework without modifications to the proofs.

8.6 CHARACTERIZATION OF THE DOMAIN: PROOFS

We first prove a crucial lemma, which shows that continuous and convex preference relations are completely determined by their symmetric part up to orientation. This generalizes Theorem 2 by Fishburn and Gehrlein (1987), who showed the same statement for SSB preferences (i.e., they additionally assume symmetry). The weaker version by Fishburn and Gehrlein is sufficient for our main result, but we believe that Lemma 8.8 may be of independent interest, e.g., when trying to strengthen Theorems 8.1 and 8.2.

Before giving a proof of Lemma 8.8, we show four auxiliary statements about continuous and convex preference relations. Unless otherwise stated, we say that a set is open or closed if it is open or closed in $\Delta(U)$.

Lemma 8.4
Let $\succ$ be a continuous and convex preference relation. Then, $U(p)$ and $L(p)$ are open for all $p \in \Delta(U)$.

Proof. Let $p \in \Delta(U)$. We start by showing that $I(p)$ is an affine subspace of $\Delta(U)$, i.e., $I(p) = \mathrm{aff}(I(p)) \cap \Delta(U)$. To this end, let $q \in \mathrm{aff}(I(p)) \cap \Delta(U)$. Hence, there are $k \in \mathbb{N}$, $\lambda \in \mathbb{R}^k$ with $\sum_{i=1}^k \lambda_i = 1$, and $q^i \in I(p)$ such that $q = \sum_{i=1}^k \lambda_i q^i$. Equivalently,

$$r = \frac{1}{\sum_{i\in\lambda_+} \lambda_i} \left(q + \sum_{i\in\lambda_-} (-\lambda_i) q^i \right) = \frac{1}{\sum_{i\in\lambda_+} \lambda_i} \sum_{i\in\lambda_+} \lambda_i q^i. \tag{1}$$

Note that $1 + \sum_{i \in \lambda_-} -\lambda_i = \sum_{i \in \lambda_+} \lambda_i$, since $\sum_{i=1}^{k} \lambda_i = 1$. Hence, $r \in \text{conv}(I(p))$ and thus, by convexity of $\succ$, $r \in I(p)$. If $q \in U(p)$, then, by convexity of $\succ$ and (1), $r \in U(p)$, which is a contradiction. Similarly, if $q \in L(p)$. Hence, $q \in I(p)$. This proves $I(p) = \text{aff}(I(p)) \cap \Delta(U)$. Thus, as the intersection of two closed sets, $I(p)$ is closed.

Now assume for contradiction that $U(p)$ is not open, i.e., there is $q \in U(p)$ such that the ϵ-ball $B_\epsilon(q)$ around q intersects with either $I(p)$ or $L(p)$ for every $\epsilon > 0$. For $r \in B_\epsilon(q) \cap L(p)$, by continuity of $\succ$ we have that, $\text{conv}(\{q, r\}) \cap I(p) \neq \emptyset$. Hence, $B_\epsilon(q) \cap I(p) \neq \emptyset$ for every $\epsilon > 0$. This implies that q is in the closure of $I(p)$, which contradicts the fact that $I(p)$ is closed. □

Lemma 8.5

Let $\succ$ be a continuous and convex preference relation. For all $p \in \Delta(U)$, if $I(p)$ contains a non-empty open set, then $I(p) = \Delta(U)$.

Proof. Assume for contradiction that $I(p) \neq \Delta(U)$ or, equivalently, $U(p) \cup L(p) \neq \emptyset$. Without loss of generality, assume that $U(p) \neq \emptyset$. Let $q \in I(p)$ such that a neighborhood of q is contained in $I(p)$ and let $r \in U(p)$. Then convexity of $\succ$ implies that $\lambda q + (1-\lambda)r \in U(p)$ for all $\lambda \in (0, 1)$. This contradicts the assumption that a neighborhood of q is contained in $I(p)$. □

The interior of a preference relation $\text{int}(\succ) = \{p \in \Delta(U) \colon U(p) \neq \emptyset \text{ and } L(p) \neq \emptyset\}$ is the set of all outcomes with non-empty upper and lower contour sets.

Lemma 8.6

Let $\succ$ be a continuous and convex preference relation. Then, for every $p \in \text{int}(\succ)$, $I(p) = \Delta(U) \cap H$, where H is a $(|U|-1)$-dimensional hyperplane in $\mathbb{R}^U$. Moreover, $I(p)$ has dimension $|U| - 2$.

Proof. Let $p \in \text{int}(\succ)$. Then, by Lemma 8.4, $U(p)$ and $L(p)$ are non-empty and open. Since $\succ$ is convex, $U(p)$ and $L(p)$ are convex. By the separating hyperplane theorem, there are $x \in \mathbb{R}^U$ and $\lambda \in \mathbb{R}$ such that $H = \{y \in \mathbb{R}^U \colon x^T y = \lambda\}$ strictly separates $U(p)$ and $L(p)$. Thus, $\Delta(U) \cap H \subseteq I(p)$. Since $U(p)$ and $L(p)$ are non-empty and H is strictly separating, H contains an interior point of $\Delta(U)$. Hence, $\Delta(U) \cap H$ has dimension $|U| - 2$. If $I(p)$ has dimension $|U| - 1$, then, since $I(p)$ is convex, it contains an open set. Lemma 8.5 implies that $I(p) = \Delta(U)$. This contradicts $p \in \text{int}(\succ)$. □

Lemma 8.7

Let $\succ$ be a continuous and convex preference relation. If $\succ \neq \emptyset$, $\text{int}(\succ)$ is non-empty and open and $\text{cl}(\text{int}(\succ)) = \Delta(U)$.

Proof. First we show that $\text{int}(\succ) \neq \emptyset$. If $\succ \neq \emptyset$, there is $p \in \Delta(U)$ such that $L(p) \neq \emptyset$. Let $q \in L(p)$. Then, by convexity of $\succ$, $p \succ 1/2\,p + 1/2\,q \succ q$, i.e., $1/2\,p + 1/2\,q \in \text{int}(\succ)$.

To show that $\text{int}(\succ)$ is open, let $p \in \text{int}(\succ)$, $q \in U(p)$, and $r \in L(p)$. Then $p \in L(q) \cap U(r)$. Since, by Lemma 8.4, $L(q)$ and $U(r)$ are open and contain p, $L(q) \cap U(r) \subseteq \text{int}(\succ)$ contains a neighborhood of p.

To show that $\text{cl}(\text{int}(\succ)) = \Delta(U)$, let $p \in \max_{\succ} \Delta(U)$. Let $O \subseteq \Delta(U)$ be a neighborhood of p. Assume for contradiction that $O \cap \text{int}(\succ) = \emptyset$. If $O \cap (\min_{\succ} \Delta(U) \setminus \max_{\succ} \Delta(U)) \neq \emptyset$, let $q \in O \cap (\min_{\succ} \Delta(U) \setminus \max_{\succ} \Delta(U))$. Since $q \notin \max_{\succ} \Delta(U)$, it follows that $U(q) \neq \emptyset$. From Lemma 8.4 we know that $U(q)$ is open. Moreover, $q \in \text{cl}(U(q))$, since $\lambda q + (1-\lambda) r \in U(q)$ for all $\lambda > 0$ and $r \in U(q)$. Hence, $O \cap U(q) \neq \emptyset$. As the intersection of two open sets, $O \cap U(q)$ is open. Since $U(q) \cap \min_{\succ} \Delta(U) = \emptyset$, the assumption that $O \cap \text{int}(\succ) = \emptyset$ implies that $O \cap U(q) \subseteq \max_{\succ} \Delta(U)$. If $O \cap (\min_{\succ} \Delta(U) \setminus \max_{\succ} \Delta(U)) = \emptyset$, then, by assumption, $O \subseteq \max_{\succ} \Delta(U)$. In any case, $\max_{\succ} \Delta(U)$ contains an open set. Observe that, for all $p, q \in \max_{\succ} \Delta(U)$, $q \in I(p)$ and hence, by Lemma 8.5, $I(p) = \Delta(U)$. Now let $p \in O \subseteq \max_{\succ} \Delta(U)$, where O is an open set and $q, r \in \Delta(U)$ such that $r \in U(q)$. Then, by convexity, $(1-\lambda)p + \lambda r \in U((1-\lambda)p + \lambda q)$ for all $\lambda \in (0,1)$. For small $\lambda > 0$, $(1-\lambda)p + \lambda r, (1-\lambda)p + \lambda q \in O \subseteq \max_{\succ} \Delta(U)$, which contradicts $q \in I(p)$ for all $p, q \in \max_{\succ} \Delta(U)$. Hence, for every $p \in \max_{\succ} \Delta(U)$ and every neighborhood O of p, $O \cap \text{int}(\succ) \neq \emptyset$, i.e., $p \in \text{cl}(\text{int}(\succ))$. Similarly for $q \in \min_{\succ} \Delta(U)$. Hence, $\text{cl}(\text{int}(\succ)) = \Delta(U)$. $\square$

We are now ready to prove Lemma 8.8.

Lemma 8.8

Let $\succ, \hat{\succ}$ be continuous and convex preference relations. If $\sim \subseteq \hat{\sim}$, then $\hat{\succ} \in \{\succ, \succ^{-1}, \emptyset\}$.

Proof. Let $p \in \Delta(U)$. By assumption, we have $I(p) \subseteq \hat{I}(p)$. Moreover, $\Delta(U)$ is the disjoint union of $I(p)$, $U(p)$, $L(p)$ and $\hat{I}(p)$, $\hat{U}(p)$, $\hat{L}(p)$, respectively. This implies that $\hat{U}(p) \cup \hat{L}(p) \subseteq U(p) \cup L(p)$. Assume for contradiction that $\hat{U}(p) \cap U(p) \neq \emptyset$ and $\hat{U}(p) \cap L(p) \neq \emptyset$. Let $q \in \hat{U}(p) \cap U(p)$ and $r \in \hat{U}(p) \cap L(p)$. Continuity of $\succ$ implies that $\text{conv}(\{q, r\}) \cap I(p) \neq \emptyset$. Convexity of $\hat{\succ}$ implies that $\text{conv}(\{q, r\}) \subseteq \hat{U}(p)$. Hence, $\emptyset \neq \text{conv}(\{q, r\}) \cap I(p) \subseteq \hat{U}(p)$, which contradicts $I(p) \subseteq \hat{I}(p)$. Hence, $\hat{U}(p) \subseteq U(p)$ or $\hat{U}(p) \subseteq L(p)$. Similarly, $\hat{L}(p) \subseteq L(p)$ or $\hat{L}(p) \subseteq U(p)$.

Now let $p \in \text{int}(\succ) \cap \text{int}(\hat{\succ})$. From Lemma 8.6, it follows that $I(p) = \Delta(U) \cap H$ and $\hat{I}(p) = \Delta(U) \cap \hat{H}$ for $(|U| - 1)$-dimensional hyperplanes H and $\hat{H}$ through p. Moreover, $I(p)$ and $\hat{I}(p)$ have dimension $|U| - 2$. Since $I(p) \subseteq \hat{I}(p)$, it follows that $I(p) = \hat{I}(p)$. Then, either $U(p) = \hat{U}(p)$ and $L(p) = \hat{L}(p)$ or $U(p) = \hat{L}(p)$ and $L(p) = \hat{U}(p)$. Let $\succ_p$ denote the restriction of $\succ$ to those comparisons involving p, i.e., $\succ_p = \succ \cap (\{p\} \times \Delta(U) \cup \Delta(U) \times \{p\})$. Thus, either $\succ_p = \hat{\succ}_p$ or $\succ_p = \hat{\succ}_p^{-1}$.

If $\hat{\succsim} = \emptyset$, there is nothing left to show. Hence assume that $\hat{\succsim} \neq \emptyset$. By assumption, this implies that $\succ \neq \emptyset$. From Lemma 8.7, it follows that $\text{int}(\succ) \cap \text{int}(\hat{\succsim}) \neq \emptyset$. Let $p \in \text{int}(\succ) \cap \text{int}(\hat{\succsim})$ and assume without loss of generality that $\succ_p = \hat{\succsim}_p$. Let $q \in \text{int}(\succ) \cap \text{int}(\hat{\succsim})$. If $q \in U(p) = \hat{U}(p)$, then $p \in L(q) \cap \hat{L}(q)$. Hence, $\succ_q = \hat{\succsim}_q$. Similarly, if $q \in L(p)$. From Lemma 8.6, it follows that $I(p) \cup I(q) \neq \Delta(U)$. Hence, $(U(p) \cup L(p)) \cap (U(q) \cup L(q))$ is non-empty and, by Lemma 8.4, open. By Lemma 8.7, $(U(p) \cup L(p)) \cap (U(q) \cup L(q)) \cap \text{int}(\succ)\text{int}(\hat{\succsim})$ is non-empty and open. For $r \in (U(p) \cup L(p)) \cap (U(q) \cup L(q)) \cap \text{int}(\succ)\text{int}(\hat{\succsim})$, it follows from two applications of what we have shown before that $\succ_r = \hat{\succsim}_r$ and $\succ_q = \hat{\succsim}_q$.

Now let $p \in \Delta(U) \setminus (\text{int}(\succ) \cap \text{int}(\hat{\succsim}))$. Assume for contradiction that $L(p) \setminus \hat{L}(p) \neq \emptyset$ and let $q \in L(p) \setminus \hat{L}(p)$. By Lemma 8.4, $L(p)$ is open. Hence, there is $\epsilon > 0$ such that $B_\epsilon(q) \subseteq L(p)$. If $B_\epsilon(q) \cap \hat{L}(p) = \emptyset$, then $L(p) \setminus \hat{L}(p)$ contains an open set. If $B_\epsilon(q) \cap \hat{L}(p) \neq \emptyset$, let $r \in B_\epsilon(q) \cap \hat{L}(p)$. Since $B_\epsilon(q) \cap \hat{L}(p)$ is the intersection of two open sets, it is open. Hence, there is $\epsilon' > 0$ such that $B_{\epsilon'}(r) \subseteq B_\epsilon(q) \cap \hat{L}(p)$. Let $\tau\colon B_\epsilon(q) \to B_\epsilon(q)$, $\tau(s) = q + (q - s)$ be the reflection with respect to q. Note that $q = 1/2\,(s + \tau(s)) \in \text{conv}(\{s, \tau(s)\}$ for all $s \in B_\epsilon(q)$. Hence, since convexity of $\hat{\succsim}$ implies that $\hat{L}(p)$ is convex and $q \notin \hat{L}(p)$, $\tau(s) \in L(p) \setminus \hat{L}(p)$ for all $s \in B_{\epsilon'}(r)$, i.e., $\tau(B_{\epsilon'}(r)) \subseteq L(p) \setminus \hat{L}(p)$. In any case, there is an open set $O \subseteq L(p) \setminus \hat{L}(p)$. As the intersection of two open sets, $O \cap \text{int}(\succ) \neq \emptyset$ is open. Since, by Lemma 8.7, $\text{cl}(\text{int}(\hat{\succsim})) = \Delta(U)$, it follows that $O \cap \text{int}(\succ) \cap \text{int}(\hat{\succsim}) \neq \emptyset$. Thus, there is $q \in \text{int}(\succ) \cap \text{int}(\hat{\succsim})$ such that $q \in L(p)$ but $q \notin \hat{L}(p)$. From before we know that $\succ_r = \hat{\succsim}_r$ for all $r \in \text{int}(\succ) \cap \text{int}(\hat{\succsim})$, which is a contradiction. Hence, $\hat{L}(p) = L(p)$. Similarly, we get $\hat{U}(p) = U(p)$. In summary, $\hat{L}(p) = L(p)$, $\hat{U}(p) = U(p)$, and $\hat{I}(p) \subseteq I(p)$, which implies that $\succ_p = \hat{\succsim}_p$. □

Lemma 8.8 does not hold if convexity is weakened to the assumption that $U(p)$, $L(p)$, and $I(p)$ need to be convex for all $p \in \Delta(U)$. To see this, consider the following preference relation on the closed interval $[0, 1]$. Let $\succ$ be the greater than relation and $\hat{\succsim}$ be defined such that $p \mathrel{\hat{\succsim}} q$ if $p \in (3/4, 1]$ and $q \in [0, 1/4)$ and $p \mathrel{\hat{\sim}} q$ otherwise. Both, $\succ$ and $\hat{\succsim}$ are continuous and convex according to the weaker notion of convexity defined above. For $\succ$ this is clear. To see this for $\hat{\succsim}$, observe that, for all $p \in [0, 1]$, either $\hat{I}(p) = [0, 3/4]$ and $\hat{U}(p) = (3/4, 1]$ (if $p \in [0, 1/4)$) or $\hat{I}(p) = [0, 1]$ (if $p \in [1/4, 3/4]$) or $\hat{L}(p) = [0, 1/4)$ and $\hat{I}(p) = [1/4, 1]$ (if $p \in (3/4, 1]$). In all cases, $\hat{U}(p)$ and $\hat{L}(p)$ are open and $\hat{U}(p)$, $\hat{L}(p)$, and $\hat{I}(p)$ are convex. Continuity has no consequences for $\hat{\succsim}$, since $\text{int}(\hat{\succsim}) = \emptyset$.

The next lemma is reminiscent of what is known as the *field expansion lemma* in traditional proofs of Arrow's theorem (see, e.g. Sen, 1986).[25] Let $f\colon \mathcal{D}^N \to \mathcal{R}$ be an SWF, $G, H \subseteq N$, and $a, b \in U$. We say

25 In contrast to Lemma 8.9, the consequence of the original field expansion lemma uses a stronger notion of decisiveness.

that (G,H) is decisive for $\underline{a}$ against $\underline{b}$, denoted by $\underline{a}\ D_{G,H}\ \underline{b}$, if, for all $P \in \mathcal{D}^N$, $\underline{a} \succ_i \underline{b}$ for all $i \in G$, $\underline{a} \sim_i \underline{b}$ for all $i \in H$, and $\underline{b} \succ_i \underline{a}$ for all $i \in N \setminus (G \cup H)$ implies $\underline{a} \succ \underline{b}$. Hence, $D_{G,H}$ is a relation on $\underline{U}$.

Lemma 8.9

Let f be an Arrovian SWF on some domain $\mathcal{D}$, $G, H \subseteq N$, and $a, b \in U$. Then, $\underline{a}\ D_{G,H}\ \underline{b}$ implies that $D_{G,H} = \underline{U} \times \underline{U}$.

Proof. First we show that $\underline{a}\ D_{G,H}\ \underline{x}$ and $\underline{b}\ D_{G,H}\ \underline{x}$ for all $x \in U \setminus \{a,b\}$. To this end, let $x \in U \setminus \{a,b\}$ and $\hat{\succ} \in \mathcal{D}$ such that $\underline{a} \mathrel{\hat{\succ}} \underline{b} \mathrel{\hat{\succ}} \underline{x}$ and $\underline{a} \mathrel{\hat{\succ}} \underline{x}$, which exists by our richness assumptions on $\mathcal{D}$ (cf. Section 8.1). Consider the preference profile

$$P = (\underbrace{\hat{\succ}, \dots, \hat{\succ}}_{G}, \underbrace{\emptyset, \dots, \emptyset}_{H}, \hat{\succ}^{-1}, \dots, \hat{\succ}^{-1}).$$

Since $\hat{\sim} = \hat{\sim}^{-1}$, it follows from Pareto indifference and Lemma 8.8 that $\succ\ = f(P) \in \{\hat{\succ}, \hat{\succ}^{-1}, \emptyset\}$. Since $\underline{a}\ D_{G,H}\ \underline{b}$, $\succ\ = \hat{\succ}$ remains as the only possibility. Hence, $\underline{a} \succ \underline{x}$ and $\underline{b} \succ \underline{x}$. By IIA, it follows that $\underline{a}\ D_{G,H}\ \underline{x}$ and $\underline{b}\ D_{G,H}\ \underline{x}$.

Repeated application of the second statement implies that $D_{G,H}$ is a complete relation. To show that $D_{G,H}$ is symmetric, let $x, y, z \in U$ such that $\underline{x}\ D_{G,H}\ \underline{y}$. The first part of the statement implies that $\underline{x}\ D_{G,H}\ \underline{z}$. Two applications of the second part yield $\underline{z}\ D_{G,H}\ \underline{y}$ and $\underline{y}\ D_{G,H}\ \underline{x}$. Hence, $D_{G,H} = \underline{U} \times \underline{U}$. □

Now we show that anonymous Arrovian aggregation is only possible on domains in which preferences over outcomes are completely determined by preferences over pure outcomes.

Lemma 8.10

Let f be an anonymous Arrovian SWF on some domain $\mathcal{D} \subseteq \mathcal{R}$. Then, $\succ|_{\underline{A}} = \hat{\succ}|_{\underline{A}}$ implies $\succ|_{\Delta(A)} = \hat{\succ}|_{\Delta(A)}$ for all $\succ, \hat{\succ} \in \mathcal{D}$ and $A \in \mathcal{F}(U)$.

Proof. Let $\succ_0, \hat{\succ}_0 \in \mathcal{D}$ and $A \in \mathcal{F}(U)$ such that $\succ_0|_{\underline{A}} = \hat{\succ}_0|_{\underline{A}}$. Consider the preference profile

$$P = (\succ_0, \hat{\succ}_0^{-1}, \emptyset, \dots, \emptyset).$$

Note that $P \in \mathcal{D}^N$ since $\mathcal{D}$ satisfies our richness assumptions. Assume that there are $a, b \in A$ such that $\underline{a} \succ_0 \underline{b}$ and define $\bar{P} = P_{(12)}$ to be identical to P except that the preferences of agents 1 and 2 are exchanged. Anonymity of f implies that $\bar{\succ} = f(\bar{P}) = f(P) = \succ$. Assume for contradiction that $\underline{a} \succ \underline{b}$. Then, by IIA, $(\{1\}, N \setminus \{1,2\})$ is decisive for $\underline{a}$ against $\underline{b}$. Lemma 8.9 implies that $(\{1\}, N \setminus \{1,2\})$ is also decisive for $\underline{b}$ against $\underline{a}$. Hence $\underline{b} \mathrel{\bar{\succ}} \underline{a}$, which contradicts $\bar{\succ} = \succ$. Thus, $\underline{a} \sim \underline{b}$. Hence, we get that $\underline{a} \sim \underline{b}$ for all $a, b \in A$ such that $\underline{a} \mathrel{\hat{\succ}_0} \underline{b}$. For $a, b \in A$ such that $\underline{a} \sim_0 \underline{b}$ and $\underline{a} \mathrel{\hat{\sim}_0} \underline{b}$, it follows from Pareto indifference that $\underline{a} \sim \underline{b}$. Hence, $\underline{a} \sim \underline{b}$ for all $a, b \in A$.

Since $\succ$ satisfies convexity, we get that $\succ|_{\Delta(A)} = \emptyset$. If $\succ_0|_{\Delta(A)} \neq \hat{\succ}_0|_{\Delta(A)}$, there are $p, q \in \Delta(A)$ such that $p \succ_0 q$ and not $p \hat{\succ}_0 q$, i.e., $p \hat{\succsim}_0^{-1} q$. The strict part of Pareto optimality of f implies that $p \succ q$. This contradicts $\succ|_{\Delta(A)} = \emptyset$. Hence, $\succ_0|_{\Delta(A)} = \hat{\succ}_0|_{\Delta(A)}$. □

Lemma 8.10 is the only part of the proof of Theorem 8.1 that requires anonymity. A much weaker condition would also suffice: there has to be $P \in \mathcal{D}^N$, $a, b \in U$, $i \in N$, and $f(P) = \succ$ such that $\underline{a} \succ_i \underline{b}$ and $\underline{a} \sim \underline{b}$.

Next, we show that intensities of preferences between pure outcomes have to be identical.

Lemma 8.11

Let f be an anonymous Arrovian SWF on some domain $\mathcal{D} \subseteq \mathcal{R}$ with $|U| \geqslant 4$. Then, for all $\succ_0 \in \mathcal{D}$, $\phi_0 = \phi^{\succ_0}$, and $a, b, c \in U$ with $\underline{a} \succ_0 \underline{b}$,

(i) $\underline{b} \succ_0 \underline{c}$ implies $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{b}, \underline{c})$,

(ii) $\underline{a} \succ_0 \underline{c}$ implies $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{a}, \underline{c})$,

(iii) $\underline{c} \succ_0 \underline{b}$ implies $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{c}, \underline{b})$, and

(iv) $\underline{c} \succ_0 \underline{a}$ implies $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{c}, \underline{a})$.

Proof. Ad (i): Continuity implies that $\underline{b} \sim_0 \lambda\underline{a} + (1-\lambda)\underline{c}$ for some $\lambda \in (0,1)$. Observe that $\succ_0^{(ac)}|_{\{\underline{a},\underline{b},\underline{c}\}} = \succ_0^{-1}|_{\{\underline{a},\underline{b},\underline{c}\}}$, where (ac) denotes the permutation that swaps a and c and leaves all other alternatives fixed.. Lemma 8.10 implies that $\succ_0^{(ac)}|_{\Delta(\{a,b,c\})} = \succ_0^{-1}|_{\Delta(\{a,b,c\})}$. Hence, we have $\underline{b} \sim_0 (1-\lambda)\underline{a} + \lambda\underline{c}$. Convexity of $\succ_0$ then implies that $\underline{b} \sim_0 1/2\, \underline{a} + 1/2\, \underline{c}$. This is equivalent to $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{b}, \underline{c})$.

Ad (ii): We distinguish two cases.

Case 1 ($\underline{b} \sim_0 \underline{c}$): Consider the preference profile

$$P = (\succ_0, (\succ_0^{(bc)})^{-1}, \emptyset, \dots, \emptyset).$$

Let $\succ = f(P)$. As in the proof of Lemma 8.10, we get that $\succ|_{\Delta(\{a,b,c\})} = \emptyset$. Without loss of generality, assume that $\phi_0(\underline{a}, \underline{b}) = 1$ and $\phi_0(\underline{a}, \underline{c}) = \lambda$ for some $\lambda \in (0, 1]$. Let $p = 1/2\, \underline{a} + 1/2\, \underline{c}$ and $q = 1/2\, \underline{a} + 1/2\, \underline{b}$, and denote by ϕ_1 and ϕ_2 the SSB functions representing the preference relations $\succ_0$ and $(\succ_0^{(bc)})^{-1}$, respectively. Then $\phi_1(p, q) = \phi_2(p, q) = 1/4\,(1-\lambda)$. If $\lambda < 1$, the strict part of Pareto optimality of f implies that $p \succ q$. This contradicts $\succ|_{\{\underline{a},\underline{b},\underline{c}\}} = \emptyset$. Hence, $\lambda = 1$.

Case 2 ($\underline{b} \succ_0 \underline{c}$): Our richness assumptions on the domain imply that there is $\hat{\succ}_0 \in \mathcal{D}$ with $\underline{a} \hat{\succ}_0 \underline{b} \hat{\succ}_0 \underline{c}$, $\underline{a} \hat{\succ}_0 \underline{c}$, and $\underline{c} \hat{\succ}_0 \underline{x}$ for some $x \in U$. Let $\hat{\phi}_0 = \phi^{\hat{\succ}_0}$. Lemma 8.10 implies that $\phi_{0\{\underline{a},\underline{b},\underline{c}\}} \equiv \hat{\phi}_{0\{\underline{a},\underline{b},\underline{c}\}}$. Assume without loss of generality that $\hat{\phi}_0(\underline{a}, \underline{b}) = 1$. By (i), we get $\hat{\phi}_0(\underline{a}, \underline{b}) = \hat{\phi}_0(\underline{b}, \underline{c}) = 1$. By (i), we get that $\hat{\phi}_0(\underline{a}, \underline{c}) = \hat{\phi}_0(\underline{c}, \underline{x})$ and $\hat{\phi}_0(\underline{b}, \underline{c}) = \hat{\phi}_0(\underline{c}, \underline{x}) = 1$. Hence, $\hat{\phi}_0(\underline{a}, \underline{c}) = 1$.

Ad (iii): The proof is analogous to the proof of (ii).

Ad (iv): The proof is analogous to the proof of (i). □

Theorem 8.1

Let $|U| \geqslant 4$ and f be an anonymous Arrovian SWF on some domain $\mathcal{D} \subseteq \mathcal{R}$. Then, $\mathcal{D} \subseteq \mathcal{D}^{PC}$.

Proof. Let $\succ_0 \in \mathcal{D}$ and $a, b, c, d \in U$ such that $\underline{a} \succ_0 \underline{b}$ and $\underline{c} \succ_0 \underline{d}$. We have to show that $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{c}, \underline{d})$. First assume there are $x \in \{a, b\}$ and $y \in \{c, d\}$ such that $\underline{x} \succ_0 \underline{y}$ or $\underline{y} \succ_0 \underline{x}$. Then, Lemma 8.11 implies that $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{x}, \underline{y}) = \phi_0(\underline{c}, \underline{d})$ or $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{y}, \underline{x}) = \phi_0(\underline{c}, \underline{d})$, respectively. Otherwise, $\underline{x} \sim_0 \underline{y}$ for all $x \in \{a, b\}$ and $y \in \{c, d\}$. This implies that $\succ_0|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} = {\succ_0}^{(ac)(bd)}|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$. From Lemma 8.10 we get $\succ_0|_{\Delta(\{a,b,c,d\})} = {\succ_0}^{(ac)(bd)}|_{\Delta(\{a,b,c,d\})}$. It follows that ${\phi_0}_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} = {\phi_0^{(ac)(bd)}}_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$ which in turn implies $\phi_0(\underline{a}, \underline{b}) = \phi_0(\underline{c}, \underline{d})$. □

8.7 CHARACTERIZATION OF THE SOCIAL WELFARE FUNCTION: PROOFS

In light of Theorem 8.1, we will assume throughout this section that $\mathcal{D} \subseteq \mathcal{D}^{PC}$. Except for Theorem 8.2, all results in this section only require Pareto indifference. Since for PC preferences the SSB utilities over outcomes are completely determined by the preferences over pure outcomes, we will write $\phi_{\underline{A}}$ instead of the more clumsy $\phi_{\Delta(A)}$ for any SSB function $\phi \in \Phi^{PC}$ and subset of alternatives $A \subseteq U$.

The following four lemmas show that for all preference profiles P and all alternatives a and b, $\phi(\underline{a}, \underline{b})$ only depends on the set of agents who prefer $\underline{a}$ to $\underline{b}$, whenever P is from the domain of PC preferences and ϕ represents $f(P)$. We first prove that, if a pure outcome is strictly Pareto dominated, then the intensities of collective preferences between each of the dominating pure outcomes and the dominated pure outcome are identical. (Using a symmetric argument, the same can be shown for profiles in which the Pareto dominance is reversed.)

Lemma 8.12

Let f be an Arrovian SWF on some domain $\mathcal{D}$ with $|U| \geqslant 4$. Let $a, b, c \in U$ and $P \in \mathcal{D}^N$, $\phi^P \equiv (\phi_i)_{i \in N}$, such that $\phi_i(\underline{a}, \underline{c}) = \phi_i(\underline{b}, \underline{c}) = 1$ for all $i \in N$. Then, $\phi(\underline{a}, \underline{c}) = \phi(\underline{b}, \underline{c})$, where $\phi = \phi^{f(P)}$.

Proof. The idea of the proof is to introduce a fourth alternative, which serves as a calibration device for the intensity of pairwise comparisons, and eventually disregard this alternative using IIA. To this end, let $x \in U$ and consider a preference profile $\hat{P} \in \mathcal{D}^N$ such that $P_{\{\underline{a},\underline{b},\underline{c}\}} = \hat{P}|_{\{\underline{a},\underline{b},\underline{c}\}}$ and

$$\phi^{\hat{P}}_{\{\underline{a},\underline{b},\underline{c},\underline{x}\}} = (\underbrace{\begin{pmatrix} 0 & & 1 & 1 \\ & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix}, \dots}_{N}).$$

The values of $\hat{\phi}_i(\underline{a}, \underline{b})$ for all $i \in N$ are irrelevant.[26] Let $\hat{\phi} = \phi^{f(\hat{P})}$. The Pareto indifference relation with respect to $\hat{P}|_{\{\underline{a},\underline{c},\underline{x}\}}$ is identical to $\sim_1|_{\{\underline{a},\underline{c},\underline{x}\}}$. The analogous statement holds for the Pareto indifference relation with respect to $\hat{P}|_{\{\underline{b},\underline{c},\underline{x}\}}$. Hence, Pareto indifference, Lemma 8.8, and IIA imply that there are $\alpha, \beta \in \mathbb{R}$ such that

$$\hat{\phi}_{\{\underline{a},\underline{c},\underline{x}\}} = \alpha \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix} \quad \text{and} \quad \hat{\phi}_{\{\underline{b},\underline{c},\underline{x}\}} = \beta \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix}.$$

As a consequence, $\alpha = \beta$ and $\hat{\phi}(\underline{a}, \underline{c}) = \hat{\phi}(\underline{b}, \underline{c})$. Since $P|_{\{\underline{a},\underline{b},\underline{c}\}} = \hat{P}|_{\{\underline{a},\underline{b},\underline{c}\}}$, Lemma 8.10 and IIA imply that $\phi_{\{\underline{a},\underline{b},\underline{c}\}} \equiv \hat{\phi}_{\{\underline{a},\underline{b},\underline{c}\}}$. Hence, we have that $\phi(\underline{a}, \underline{c}) = \phi(\underline{b}, \underline{c})$.[27] □

Given a preference profile P, let $N_{ab} = \{i \in N \colon \underline{a} \succ_i \underline{b}\}$ be the set of agents who strictly prefer $\underline{a}$ over $\underline{b}$ and $n_{ab} = |N_{ab}|$. Also, let $I_{ab} = N \setminus (N_{ab} \cup N_{ba})$ be the set of agents who are indifferent between $\underline{a}$ and $\underline{b}$.

Lemma 8.13 shows that for a fixed preference profile, $\phi(\underline{a}, \underline{b})$ only depends on N_{ab} and I_{ab} and not on the names of the alternatives.

Lemma 8.13

Let f be an Arrovian SWF on some domain $\mathcal{D}$ with $|U| \geqslant 5$, $a, b, c, d \in U$, and $P \in \mathcal{D}^N$ such that $N_{ab} = N_{cd}$ and $N_{ba} = N_{dc}$. Then, $\phi(\underline{a}, \underline{b}) = \phi(\underline{c}, \underline{d})$, where $\phi = \phi^{f(P)}$.

Proof. We first prove the case when all of a, b, c, d are distinct. Let $e \in U$ and consider a preference profile $\hat{P} \in \mathcal{D}^N$ such that $P|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} =$

26 Also the values $\hat{\phi}_i(\underline{x}, \underline{z})$ for all $z \in \{a, b, c\}$ are irrelevant as long as they are the same for all agents.

27 Pareto dominance also implies that $\phi(\underline{a}, \underline{c}), \phi(\underline{b}, \underline{c}) > 0$.

$\hat{P}|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$ and $\hat{\phi}_i(\underline{x},\underline{e}) = 1$ for all $x \in \{a,b,c,d\}$ and $i \in N$. Now consider a preference profile $\mathring{P} \in \mathcal{D}^N$ such that

$$\phi^{\mathring{P}}{}_{\{\underline{a},\underline{b},\underline{c},\underline{d},\underline{e}\}} = (\underbrace{\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ -1 & -1 & 0 & 1 & 1 \\ -1 & -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{pmatrix}, \dots,}_{N_{ab}}$$

$$\underbrace{\begin{pmatrix} 0 & -1 & -1 & -1 & 1 \\ 1 & 0 & -1 & -1 & 1 \\ 1 & 1 & 0 & -1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{pmatrix}}_{N_{ba}}, \dots, \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{pmatrix}, \dots).$$

Note that $\hat{P}|_{\{\underline{a},\underline{b},\underline{e}\}} = \mathring{P}|_{\{\underline{a},\underline{b},\underline{e}\}}$ and $\hat{P}|_{\{\underline{c},\underline{d},\underline{e}\}} = \mathring{P}|_{\{\underline{c},\underline{d},\underline{e}\}}$ because $N_{ab} = N_{cd}$ and $N_{ba} = N_{dc}$ by assumption. Now, let $\hat{\phi} = \phi^{f(\hat{P})}$ and $\mathring{\phi} = \phi^{f(\mathring{P})}$. Since $\hat{P}|_{\{\underline{a},\underline{b},\underline{e}\}} = \mathring{P}|_{\{\underline{a},\underline{b},\underline{e}\}}$, we have $\hat{\phi}_{\{\underline{a},\underline{b},\underline{e}\}} \equiv \mathring{\phi}_{\{\underline{a},\underline{b},\underline{e}\}}$ by IIA. Moreover, $\hat{P}|_{\{\underline{c},\underline{d},\underline{e}\}} = \mathring{P}|_{\{\underline{c},\underline{d},\underline{e}\}}$ and IIA yield $\hat{\phi}_{\{\underline{c},\underline{d},\underline{e}\}} \equiv \mathring{\phi}_{\{\underline{c},\underline{d},\underline{e}\}}$. Lemma 8.12 implies that $\mathring{\phi}(\underline{x},\underline{e}) = \lambda$ for some $\lambda \in \mathbb{R}$ for all $x \in \{a,b,c,d\}$. Thus, for some $\mu, \sigma \in \mathbb{R}$, $\mathring{\phi}$ takes the form

$$\mathring{\phi}_{\{\underline{a},\underline{b},\underline{c},\underline{d},\underline{e}\}} = \begin{pmatrix} 0 & \mu & & & \lambda \\ -\mu & 0 & & & \lambda \\ & & 0 & \sigma & \lambda \\ & & -\sigma & 0 & \lambda \\ -\lambda & -\lambda & -\lambda & -\lambda & 0 \end{pmatrix}.$$

Note that $\mathring{P}|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$ only consists of one fixed preference relation, its inverse, and complete indifference. Hence, Pareto indifference and Lemma 8.8 imply that $\mathring{\phi}_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} = \alpha\mathring{\phi}_{1\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$ for some $\alpha \in \mathbb{R}$, where $\mathring{\phi}_1$ is the SSB function that represents the preference relation of the agents in N_{ab} in $\mathring{P}$. Hence, we get that $\mu = \sigma$.

The cases when $a = c$ and $b = c$ follow from repeated application of the above case. All other cases are symmetric to one of the covered cases. □

Lemma 8.14

Let f be an Arrovian SWF, $a,b,c,d \in U$, $P, \hat{P} \in \mathcal{D}^N$, $\phi = \phi^{f(P)}$, and $\hat{\phi} = \phi^{f(\hat{P})}$. If $P|_{\{\underline{a},\underline{b}\}} = \hat{P}|_{\{\underline{a},\underline{b}\}}$ and $P|_{\{\underline{c},\underline{d}\}} = \hat{P}|_{\{\underline{c},\underline{d}\}}$, there is $\alpha > 0$ such that $\phi(\underline{a},\underline{b}) = \alpha \cdot \hat{\phi}(\underline{a},\underline{b})$ and $\phi(\underline{c},\underline{d}) = \alpha \cdot \hat{\phi}(\underline{c},\underline{d})$.

Proof. Let $e \in U \setminus \{a,b,c,d\}$ and $P', \hat{P}' \in \mathcal{D}^N$ such that $P'|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} = P|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$, $\hat{P}'|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} = \hat{P}|_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$, and $\phi'_i(\underline{x},\underline{e}) = \hat{\phi}'_i(\underline{x},\underline{e}) = 1$ for all $x \in \{a,b,c,d\}$ and $i \in N$. By $\phi' \equiv \phi^{f(P')}$ and $\hat{\phi}' \equiv \phi^{f(\hat{P}')}$ we denote the corresponding collective SSB functions. Since f satisfies

IIA, we have that $\phi_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} \equiv \phi'_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$ and $\hat{\phi}_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} \equiv \hat{\phi}'_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$. Lemma 8.12 implies that without loss of generality, ϕ' and $\hat{\phi}'$ take the following form for some $\lambda, \mu, \hat{\mu}, \sigma, \hat{\sigma} \in \mathbb{R}$ and $A = \{a, b, c, d, e\}$. Note that we can choose suitable representatives such that $\phi'(\underline{a}, \underline{e}) = \hat{\phi}'(\underline{a}, \underline{e}) = \lambda$.

$$\phi'_{\underline{A}} = \begin{pmatrix} 0 & \mu & & & \lambda \\ -\mu & 0 & & & \lambda \\ & & 0 & \sigma & \lambda \\ & & -\sigma & 0 & \lambda \\ -\lambda & -\lambda & -\lambda & -\lambda & 0 \end{pmatrix} \quad \hat{\phi}'_{\underline{A}} = \begin{pmatrix} 0 & \hat{\mu} & & & \lambda \\ -\hat{\mu} & 0 & & & \lambda \\ & & 0 & \hat{\sigma} & \lambda \\ & & -\hat{\sigma} & 0 & \lambda \\ -\lambda & -\lambda & -\lambda & -\lambda & 0 \end{pmatrix}$$

Observe that $P'|_{\{\underline{a},\underline{b},\underline{e}\}} = \hat{P}'|_{\{\underline{a},\underline{b},\underline{e}\}}$ and $P'|_{\{\underline{c},\underline{d},\underline{e}\}} = \hat{P}'|_{\{\underline{c},\underline{d},\underline{e}\}}$ by construction. Since f satisfies IIA, we get that $\phi'_{\{\underline{a},\underline{b},\underline{e}\}} = \hat{\phi}'_{\{\underline{a},\underline{b},\underline{e}\}}$ and $\phi'_{\{\underline{c},\underline{d},\underline{e}\}} = \hat{\phi}'_{\{\underline{c},\underline{d},\underline{e}\}}$. In particular, this means that $\mu = \hat{\mu}$ and $\sigma = \hat{\sigma}$. Since $\phi_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} \equiv \phi'_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$ and $\hat{\phi}_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}} \equiv \hat{\phi}'_{\{\underline{a},\underline{b},\underline{c},\underline{d}\}}$, there is $\alpha > 0$ as required. □

Lemma 8.14 shows that $\phi(\underline{a}, \underline{b})$ only depends on N_{ab} and I_{ab} and not on a, b, or P. Hence, there is a function $g\colon 2^N \times 2^N \to \mathbb{R}$ such that $g(N_{ab}, I_{ab}) = \phi^{f(P)}(\underline{a}, \underline{b})$ for all $a, b \in U$ and $P \in \mathcal{D}^N$. We now leverage Pareto indifference to show that $\phi^{f(P)}$ is a linear combination of the ϕ_i. Hence, f is affine utilitarian.

Lemma 8.15
Let f be an Arrovian SWF. Then, there are $w_1, \dots, w_n \in \mathbb{R}$ such that $\phi^{f(P)} \equiv \sum_{i \in N} w_i \phi_i$ for all $P \in \mathcal{D}^N$.

Proof. For all $G \subseteq N$, let $w_G = 1/2\,(g(N, \emptyset) + g(G, \emptyset))$. For convenience, we write w_i for $w_{\{i\}}$. Since $\phi^{f(P)}(\underline{x}, \underline{y}) = g(N_{xy}, I_{xy})$ for all $x, y \in U$ and $P \in \mathcal{D}^N$, it suffices to show that

$$g(N_{xy}, I_{xy}) = \sum_{i \in N} w_i \phi_i(\underline{x}, \underline{y}) = \sum_{i \in N_{xy}} w_i - \sum_{i \in N_{yx}} w_i, \tag{2}$$

for all $x, y \in U$. To this end, we will first show that $w_G + w_{\hat{G}} = w_{G \cup \hat{G}}$ for all $G, \hat{G} \subseteq N$ with $G \cap \hat{G} = \emptyset$. Let $G, \hat{G}$ as above, $a, b, c, x, y \in U$, and consider the following preference profile $P \in \mathcal{D}^N$ such that

$$\phi^P_{\{\underline{a},\underline{b},\underline{c},\underline{x},\underline{y}\}} = (\underbrace{\begin{pmatrix} 0 & & & -1 & 1 \\ & 0 & & -1 & 1 \\ & & 0 & -1 & 1 \\ 1 & 1 & 1 & 0 & \\ -1 & -1 & -1 & & 0 \end{pmatrix}, \ldots,}_{G}$$

$$\underbrace{\begin{pmatrix} 0 & & & -1 & -1 \\ & 0 & & 1 & 1 \\ & & 0 & 1 & -1 \\ 1 & -1 & -1 & 0 & \\ 1 & -1 & 1 & & 0 \end{pmatrix}, \ldots,}_{\hat{G}} \begin{pmatrix} 0 & & & -1 & 1 \\ & 0 & & -1 & -1 \\ & & 0 & 1 & 1 \\ 1 & 1 & -1 & 0 & \\ -1 & 1 & -1 & & 0 \end{pmatrix}, \ldots).$$

Let $\phi = f(P)$. We have that, for $p = 1/2\,\underline{x} + 1/2\,\underline{y}$ and $q = 1/3\,\underline{a} + 1/3\,\underline{b} + 1/3\,\underline{c}$, $\phi_i(p, q) = 0$ for all $i \in N$. Pareto indifference implies that $\phi(p, q) = 0$. Let $\mu = g(G, \emptyset)$, $\hat{\mu} = g(\hat{G}, \emptyset)$, and $\sigma = g(G \cup \hat{G}, \emptyset)$. By definition of w,

$$w_G + w_{\hat{G}} = w_{G \cup \hat{G}}$$

is equivalent to

$$(g(N, \emptyset) + g(G, \emptyset)) + (g(N, \emptyset) + g(\hat{G}, \emptyset)) = g(N, \emptyset) + g(G \cup \hat{G}, \emptyset).$$

Hence, we have to show that $\mu + \hat{\mu} + g(N, \emptyset) = \sigma$. By definition of g, we get that ϕ takes the following form.

$$\phi_{\{\underline{a},\underline{b},\underline{c},\underline{x},\underline{y}\}} \equiv \begin{pmatrix} 0 & & & -g(N, \emptyset) & -\hat{\mu} \\ & 0 & & \hat{\mu} & \sigma \\ & & 0 & -\mu & -\hat{\mu} \\ g(N, \emptyset) & -\hat{\mu} & \mu & 0 & \\ \hat{\mu} & -\sigma & \hat{\mu} & & 0 \end{pmatrix}$$

From $\phi(p, q) = 0$, it follows that $1/6\,(\mu + \hat{\mu} + g(N, \emptyset) - \sigma) = 0$. This proves the desired relationship.

Now we can rewrite (2) as

$$g(N_{xy}, I_{xy}) = w_{N_{xy}} - w_{N_{yx}}. \tag{3}$$

By definition of w, this is equivalent to

$$2g(N_{xy}, I_{xy}) = g(N_{xy}, \emptyset) - g(N_{yx}, \emptyset). \tag{4}$$

To prove (4), let $a, b, x, y \in U$ and consider a following preference profile $\hat{P} \in \mathcal{D}^N$ such that

$$\phi^{\hat{P}}_{\{\underline{a},\underline{b},\underline{x},\underline{y}\}} = (\underbrace{\begin{pmatrix} 0 & & 1 & 1 \\ & 0 & -1 & -1 \\ -1 & 1 & 0 & \\ -1 & 1 & & 0 \end{pmatrix}, \ldots,}_{G} \underbrace{\begin{pmatrix} 0 & & -1 & 1 \\ & 0 & -1 & 0 \\ 1 & 1 & 0 & \\ -1 & 0 & & 0 \end{pmatrix}, \ldots,}_{\hat{G}} \begin{pmatrix} 0 & & -1 & -1 \\ & 0 & 1 & 1 \\ 1 & -1 & 0 & \\ 1 & -1 & & 0 \end{pmatrix}, \ldots).$$

Let $\hat{\phi} = \phi^{f(\hat{P})}$. Observe that, for $p = 1/3\, \underline{x} + 2/3\, \underline{y}$ and $q = 1/2\, \underline{a} + 1/2\, \underline{b}$, $\hat{\phi}_i(p, q) = 0$ for all $i \in N$. Pareto indifference implies that $\hat{\phi}(p, q) = 0$. With the same definitions as before and $\epsilon = g(G, \hat{G})$, $\hat{\phi}$ takes the following form.

$$\hat{\phi}_{\{\underline{a},\underline{b},\underline{x},\underline{y}\}} \equiv \begin{pmatrix} 0 & & \mu & \sigma \\ & 0 & -\sigma & -\epsilon \\ -\mu & \sigma & 0 & \\ -\sigma & \epsilon & & 0 \end{pmatrix}$$

From $\hat{\phi}(p, q) = 0$, we get that $1/6\,(-\mu + \sigma - 2\sigma + 2\epsilon) = 0$. Hence, $2\epsilon = \mu + \sigma$. This is equivalent to

$$2g(G, \hat{G}) = g(G, \emptyset) + g(G \cup \hat{G}, \emptyset) = g(G, \emptyset) - g(N \setminus (G \cup \hat{G}), \emptyset),$$

where the last equality follows from skew-symmetry of $\hat{\phi}$ and the definition of g. This proves (4). □

Finally, the strict part of Pareto optimality implies that all weights have to be strictly positive.

Theorem 8.2

Let $|U| \geqslant 5$ and f be an Arrovian SWF on some domain $\mathcal{D} \subseteq \mathcal{D}^{PC}$. Then, there are $w_1, \ldots, w_n \in \mathbb{R}_{>0}$ such that

$$\phi^{f(P)} \equiv \sum_{i \in N} w_i \phi_i \text{ for all } P \in \mathcal{D}^N.$$

Proof. From Lemma 8.15 we know that there are $w_1, \ldots, w_n \in \mathbb{R}$ such that, for all $P \in \mathcal{D}^N$, $\phi^{f(P)} \equiv \sum_{i \in N} w_i \phi_i$. Assume for contradiction that $w_i \leqslant 0$ for some $i \in N$. Let G be the set of agents such that $w_i \leqslant 0$ and consider a preference profile $P \in \mathcal{D}^N$ with $a, b, c \in U$ such that

$$\phi^{P}_{\{\underline{a},\underline{b},\underline{c}\}} \equiv (\underbrace{\begin{pmatrix} 0 & & 1 \\ & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix}, \ldots,}_{G} \begin{pmatrix} 0 & & 1 \\ & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}, \ldots).$$

Let $\phi \equiv \phi^{f(P)}$. Then, for $p = 1/2\, \underline{a} + 1/2\, \underline{b}$, we have that $\phi_i(p, \underline{c}) > 0$ for all $i \in G$ and $\phi_i(p, \underline{c}) = 0$ for all $i \in N \setminus G$. Pareto optimality of f implies that $\phi(p, \underline{c}) > 0$. However, we have

$$\begin{aligned}\phi(p, \underline{c}) &= \alpha \left(\sum_{i \in G} w_i \phi_i(p, \underline{c}) + \sum_{i \in N \setminus G} w_i \underbrace{\phi_i(p, \underline{c})}_{=0} \right) \\ &= \alpha \sum_{i \in G} \underbrace{w_i \phi_i(p, \underline{c})}_{\leqslant 0} \leqslant 0\end{aligned}$$

for some $\alpha > 0$. This is a contradiction. □

9 RELATIVE UTILITARIAN SOCIAL CHOICE

> It is common practice, on the eve of Election Day, to call upon the public to exercise their right and vote. The argument is that by voting one can sometimes influence the outcome and secure the election of a preferred candidate. It turns out, however, that some popular voting rules may give rise to situations where one's vote results in the election of a less preferred candidate [...]
>
> — R. Holzman

It has been established in Chapter 8 that anonymous Arrovian preference aggregation necessitates relative utilitarianism. In this chapter, we consider the SCF that chooses maximal elements according to relative utilitarianism, called relative utilitarian outcomes, in a framework where agents may choose to opt-out from the aggregation process by not reporting their preferences. Fishburn and Brams (1983) observed that some SCFs may incentivize agents to abstain since this yields a more preferred outcome. Moulin (1988) showed that this phenomenon, called the *no-show paradox*, pertains to all resolute and Condorcet consistent *pure* SCFs. A number of authors provided strengthenings of this result (Holzman, 1988; Sanver and Zwicker, 2009; Brandt et al., 2017), extensions to not necessarily strict preferences (Duddy, 2014), extensions to non-resolute SCFs (Pérez, 2001; Jimeno et al., 2009; Brandl et al., 2015a), and extensions to not necessarily pure SCFs (Brandl et al., 2015b). An SCF that is not susceptible to the no-show paradox entices participation. Moulin's result even holds when only considering abstention by single agents. We study SCFs that entice participation for groups of agents in the sense that no group of agents can obtain a more preferred outcome (in terms of accumulated utility) by abstaining. In case an SCF returns multiple outcomes, the above condition has to hold for *any* pair of chosen outcomes. This property will be called utilitarian participation.

Our first result shows that choosing relative utilitarian outcomes entices utilitarian participation. This is obvious when the agents' preferences admit representations through vNM functions and is the reason why scoring rules (such as Borda's rule or plurality rule) entice participation. Choosing relative utilitarian outcomes is not the only way to entice utilitarian participation, since, e.g., every constant SCF trivially entices utilitarian participation. It can however be singled-out under additional assumptions. In certain cases choosing relative utilitarian outcomes is particularly natural, e.g., if there exists a pure relative utilitarian outcome, i.e., a pure outcome that yields positive accumulated utility compared to every other pure outcome (and therefore, by convexity, every outcome). An SCF that uniquely chooses such a pure outcome whenever it exists is called weakly utilitarian. Our second result shows that every homogeneous and weakly utilitarian SCF that entices utilitarian participation has to choose relative utilitarian outcomes. This result requires the domain of preferences to be sufficiently rich.

The first result implies that on the domain of preferences based on pairwise comparisons, choosing maximal lotteries entices utilitarian participation. As a consequence, maximal lotteries satisfies *SD*-group-participation as introduced by Brandl et al. (2015b), which prescribes that no group of agents can abstain and thereby obtain an outcome that stochastically dominates the outcome obtained by participating with respect to the preferences of all its members. Since maximal lotteries is Condorcet consistent, this can be seen as a possible resolution of Moulin's no-show paradox for non-pure SCFs.

9.1 RELATIVE UTILITARIAN OUTCOMES AND UTILITARIAN PARTICIPATION

None of the conditions considered here connects choices from different feasible sets to each other nor do the results require any assumptions about the feasible set. Hence, the feasible set will be some fixed $A \in \mathcal{F}(U)$ for the rest of this chapter. Since we wish to consider arbitrarily large electorates, the set of agents V is assumed to be infinite. Moulin's notion of participation requires that a single agent can never be better off by abstaining. In a framework where the agents' preferences admit representations through SSB functions, this notion can be extended to groups of agents by accumulating the utility comparisons between outcomes. A group of agents prefers one outcome to another if the former yields positive accumulated utility when compared to the latter. An SCF entices utilitarian participation if no group of agents can obtain a preferred outcome by abstaining. For $N, G \in \mathcal{F}(V)$ with $G \subsetneq N$ and $P \in \mathcal{D}^N$, let $P_{-G} = (\succ_i)_{i \in N \setminus G}$ be the preference profile that is obtained from P by removing the preference relations of agents in G. An SCF f entices utilitarian participation if, for all $N, G \in \mathcal{F}(V)$ with $G \subsetneq N$, $P \in \mathcal{D}^N$, $p \in f(P, A)$, and $q \in f(P_{-G}, A)$,

$$\sum_{i \in G} \phi_i(p, q) \geqslant 0. \qquad \text{(utilitarian participation)}$$

In particular, if an SCF entices utilitarian participation, no group of agents can abstain while all its members prefer the newly obtained outcome over the originally obtained outcome. This notion of participation for groups of agents was considered by Brandl et al. (2015b) in a framework where the agents' preferences can be represented by vNM functions that are unknown to the SCF except for their ranking over pure outcomes.

The following two theorems show that choosing relative utilitarian outcomes is closely connected to utilitarian participation. First, it is shown that choosing relative utilitarian outcomes entices utilitarian participation.

THEOREM 9.1
Let $\mathcal{D} \subseteq \mathcal{R}$. Then, choosing relative utilitarian outcomes entices utilitarian participation.

Proof. Let $N, G \in \mathcal{F}(V)$ with $G \subsetneq N$, $P \in \mathcal{D}^N$, $p \in \max_{RU(P)} \Delta(A)$, and $p' \in \max_{RU(P_{-G})} \Delta(A)$. Then, we have that

$$\begin{aligned} \sum_{i \in N} \phi_i(p, q) \geqslant 0 \text{ for all } q \in \Delta(A), \text{ and} \\ \sum_{i \in N \setminus G} \phi_i(p', q) \geqslant 0 \text{ for all } q \in \Delta(A), \end{aligned} \tag{5}$$

by definition of *RU*. It follows that

$$\sum_{i \in G} \phi_i(p, p') = \underbrace{\sum_{i \in N} \phi_i(p, p')}_{\geqslant 0} - \underbrace{\sum_{i \in N \setminus G} \phi_i(p, p')}_{\leqslant 0} \geqslant 0.$$

The subscripted inequalities follow from (5) and the fact that SSB functions are skew-symmetric. Hence, choosing relative utilitarian outcomes entices utilitarian participation. □

Clearly, Theorem 9.1 does not require that the agents' preferences are from a common domain. Despite the fact that Theorem 9.1 is seemingly trivial and admits a very simple proof, it has important consequences, whose correctness is far less obvious at first sight (cf. Section 9.2).

Our second result shows that, on sufficiently rich domains, utilitarian participation in combination with homogeneity and weak utilitarianism necessitates choosing relative utilitarian outcomes. For some profiles, there is a *pure* outcome that is preferred to every other outcome in terms of accumulated utility. Such pure outcomes should arguably be chosen whenever they exist, since they are preferred to all other pure outcomes in terms of accumulated utility even *ex post*, i.e., after the randomization is executed. An SCF f is weakly utilitarian if, for all $N \in \mathcal{F}(V)$ and $P \in \mathcal{D}^N$,[28]

$$f(P, A) = \{\underline{a}\} \text{ whenever } \sum_{i \in N} \phi_i(\underline{a}, \underline{b}) > 0 \text{ for all } b \in A \setminus \{a\}. \tag{weak utilitarianism}$$

Note that for vNM preferences pure relative utilitarian outcomes always exist.

We make two assumptions about the domain $\mathcal{D}$ that are required for the proof of Theorem 9.2. The first assumption is that it should be possible for two agents to completely disagree with each other by reporting completely reversed preferences. A domain $\mathcal{D} \subseteq \mathcal{R}$ is closed under reversals if, for every $\succ \in \mathcal{D}$, $\succ^{-1} \in \mathcal{D}$ (cf. Chapter 8). Secondly, a domain should not be heavily biased towards certain alternatives. For every pure outcome, it has to be possible to find a preference profile for which this pure outcome is preferred to every other pure outcome in terms of accumulated utility. Formally, a domain $\mathcal{D}$ is non-imposing if, for every $a \in A$, there are $N \in \mathcal{F}(V)$ and $P \in \mathcal{D}^N$ such that $\sum_{i \in N} \phi_i(\underline{a}, \underline{b}) > 0$ for all $b \in A \setminus \{a\}$.

Theorem 9.2

Let $\mathcal{D} \subseteq \mathcal{R}$ be non-imposing and closed under reversals. Then, every homogeneous and weakly utilitarian SCF on $\mathcal{D}$ that entices utilitarian participation only chooses relative utilitarian outcomes.

Proof. Let f be a homogeneous and weakly utilitarian SCF that satisfies utilitarian participation. Assume for contradiction that f does not only choose relative utilitarian outcomes, i.e., there are $N \in \mathcal{F}(V)$ and $P \in \mathcal{D}^N$ such that $f(P, A) \not\subseteq \max_{RU(p)} \Delta(A)$. Hence, there are $p \in f(P, A)$ and $q \in \Delta(A)$ such that $\sum_{i \in N} \phi_i(p, q) < 0$. By linearity of the ϕ_i, there is an alternative $a \in A$ such that $\sum_{i \in N} \phi_i(p, \underline{a}) = \alpha < 0$. Since V is assumed to be infinite and $\mathcal{D}$ is closed under reversals, there are $\hat{N} \in \mathcal{F}(V)$, $\hat{N} \cap N = \emptyset$, and $\hat{P} \in \mathcal{D}^{\hat{N}}$ such that $\hat{P}(\succ) = P(\succ^{-1})$ for all $\succ \in \mathcal{D}$. Since $\mathcal{D}$ is non-imposing, there are $G \in \mathcal{F}(V)$, $G \cap (N \cup \hat{N}) = \emptyset$, and $P^G \in \mathcal{D}^G$ such that $\sum_{i \in G} \phi_i(\underline{a}, \underline{b}) > 0$ for all $b \in A \setminus \{a\}$. Now let $\beta = \sum_{i \in G} \phi_i(p, \underline{a}) < 0$ and $k \in \mathbb{N}$ such that $k\alpha - \beta < 0$. Let $N^k, \hat{N}^k \in \mathcal{F}(V)$, $|N^k| = |\hat{N}^k| = k|N|$, $G \cap N^k = \emptyset$ and $(G \cup N^k) \cap \hat{N}^k = \emptyset$, and $P^k \in \mathcal{D}^{N^k}$, $\hat{P}^k \in \mathcal{D}^{\hat{N}^k}$ such that, for all $\succ \in \mathcal{D}$, $P(\succ) = P^k(\succ)$ and $\hat{P}(\succ) = \hat{P}^k(\succ)$. It follows from homogeneity of f that $p \in f(P, A) = f(P^k, A)$. By definition of $\hat{P}$ it follows that $\sum_{i \in N^k \cup \hat{N}^k \cup G} \phi_i = \sum_{i \in G} \phi_i$. By the choice of P^G and since f is

28 The weakening of weak utilitarianism that only requires that $\underline{a} \in f(P, A)$ also suffices to derive the conclusion of Theorem 9.2.

weakly utilitarian, it follows that $f(P^k \cup \hat{P}^k \cup P^G, A) = \{\underline{a}\}$. Moreover, it holds that

$$\sum_{i \in \hat{N}^k \cup G} \phi_i(p, \underline{a}) = \underbrace{\sum_{i \in \hat{N}^k} \phi_i(p, \underline{a})}_{-k\alpha} + \underbrace{\sum_{i \in \cup G} \phi_i(p, \underline{a})}_{=\beta} = -(k\alpha - \beta) > 0.$$

Hence, the group of agents $\hat{N}^k \cup G$ prefers abstaining (which yields p) to not abstaining (which yields $\underline{a}$) in terms of accumulated utility. This contradicts the assumption that f entices utilitarian participation. □

9.2 PREFERENCES BASED ON PAIRWISE COMPARISONS

The results in Section 9.1 are particularly relevant when preferences are based on pairwise comparisons, i.e., $\mathcal{D} \subseteq \mathcal{D}^{PC}$. In this case, choosing relative utilitarian outcomes coincides with maximal lotteries and the agents' preferences over all outcomes are completely determined by their preferences over pure outcomes. Hence, an SCF can be seen as a function that maps preferences over pure outcomes to sets of outcomes. Optional participation for this class of SCFs was studied by Brandl et al. (2015b). Their notion of participation for groups of agents with respect to stochastic dominance (*SD*-group-participation) prescribes that no group of agents can, by abstaining, obtain an outcome that stochastically dominates the original outcome according to the preferences of all agents in the group. Since preferences based on pairwise comparisons are a refinement of preferences based on stochastic dominance (cf. Fishburn, 1984a; Aziz et al., 2015), it follows from Theorem 9.1 that maximal lotteries entices *SD*-group-participation.

Corollary 9.3
Let $\mathcal{D} \subseteq \mathcal{D}^{PC}$. Then, *ML* entices *SD*-group-participation.

Proof. Let $N, G \in \mathcal{F}(V)$, $G \subsetneq N$, and $P \in \mathcal{D}^N$. Let $p \in ML(P, A)$ and $q \in ML(P_{-G}, A)$. Utilitarian participation of *ML* implies that $\sum_{i \in G} \phi_i(p, q) \geqslant 0$. In particular, there is $i \in G$ such that $\phi_i(p, q) \geqslant 0$. This implies that q does not stochastically dominate p according to $\succ_i|_{\underline{A}}$. □

SD-group-participation is in fact quite demanding. E.g., Brandl et al. (2015b) showed that no majoritarian and *ex post* efficient SCF entices *SD*-participation even for single agents.[29]

29 An SCF is majoritarian if it only depends on the majority relation on pairs of pure outcomes. An SCF is *ex post* efficient if it never returns an outcome with positive probability on a Pareto dominated alternative.

Recall that an SCF satisfies Condorcet consistency if it uniquely returns pure outcomes that are preferred to all other pure outcomes by a majority of agents, so-called Condorcet winners, whenever they exist. For preferences based on pairwise comparisons, Condorcet consistency is equivalent to weak utilitarianism. We thus obtain the following corollary of Theorem 9.2.

Corollary 9.4
Let $\mathcal{D} \subseteq \mathcal{D}^{PC}$ be non-imposing and closed under reversals. Then, every homogeneous and Condorcet consistent SCF that entices utilitarian participation chooses a subset of maximal lotteries.

Corollary 9.4 is in contrast to a result by Moulin (1988), who showed that no resolute and Condorcet consistent *pure* SCF entices participation.

9.3 CONCLUDING REMARKS

Remark 9.1 *(One-way monotonicity)*
The proof of Theorem 9.1 can be adapted to show that choosing relative utilitarian outcomes satisfies *one-way monotonicity* (Sanver and Zwicker, 2009). As a consequence, maximal lotteries satisfies one-way monotonicity on the domain of PC preferences. This is in contrast to Sanver and Zwicker (2009) and Peters (2017) who showed that no Condorcet consistent pure SCF satisfies half-way monotonicity, a weakening of both one-way monotonicity and participation.

Remark 9.2 *(Domain assumptions)*
The assumption that the domain $\mathcal{D}$ is non-imposing is indispensable to derive the conclusion of Theorem 9.2. To see this, let $U = A = \{a, b, c\}$ and consider the following SSB function.

$$\begin{array}{cc} & \begin{array}{ccc} a & b & c \end{array} \\ \phi = & \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{array}{c} a \\ b \\ c \end{array} \end{array}$$

The domain $\mathcal{D} = \{\succ \in \mathcal{R} \colon \phi^{\succ} = \pm\phi\}$ is closed under reversals but imposing. Weak utilitarianism has no implications on $\mathcal{D}$. Hence, every constant function on $\mathcal{D}$ is homogeneous and weakly utilitarian and entices utilitarian participation. It is unknown, whether the assumption that $\mathcal{D}$ is closed under reversals is necessary.

Remark 9.3 *(Non-imposition and cancellation)*
Using similar arguments as in the proof of Theorem 9.2, it can be shown that every weakly utilitarian SCF that satisfies non-imposition and cancellation has to choose relative utilitarian outcomes. Non-imposition requires that every *pure* outcome is chosen for at least one preference profile. Cancellation prescribes that an SCF ignores agents with completely opposed preferences (cf. Young, 1974b).

Remark 9.4 *(SD-participation)*
Corollary 9.4 does not hold if utilitarian participation is weakened to *SD*-group-participation. For example, the SCF that uniquely chooses the Condorcet winner if one exists and the uniform distribution over A otherwise is homogeneous and Condorcet consistent and satisfies *SD*-group-participation.[30]

Remark 9.5 *(Strong SD-participation)*
A stronger notion of *SD*-participation prescribes that the outcome obtained by participating stochastically dominates the outcome obtained by abstaining. On the domain of PC preferences, this notion of *SD*-participation is incompatible with Condorcet consistency (Brandt et al., 2017, Theorem 9). In this sense, the impossibility result by Moulin (1988) also holds for non-pure SCFs.

30 However, this SCF violates *ex post* efficiency, which is satisfied by maximal lotteries.

10 CONSISTENT SOCIAL CHOICE

> Consistency can be viewed as a condition of social stability. For suppose that society has adopted a concept of equity that is not consistent. Then in some situation there will exist a subgroup of individuals who find that the way they divide the amount of property allotted to them [...] is unfair. In other words, it does not accord with the normative concept that everyone in this society subscribes to.
>
> — H. P. Young

Consistency conditions capture the rationality of the choices made by a choice function. They feature in many results in the literature on social choice theory. A very early example is Condorcet consistency, which goes back to Condorcet (1785) and prescribes that a Condorcet winner should be chosen whenever one exists. Hence it requires an SCF to be consistent with majority rule in profiles where majority rule is unambiguously defined. Arrow's theorem for SWFs can be turned into a result for SCFs, which states that every SCF satisfying Pareto optimality and independence of infeasible alternatives is dictatorial, if one requires its choices to be consistent with the weak axiom of revealed preference (Samuelson, 1938). The weak axiom of revealed preference is equivalent to the conjunction of Sen's α and a strong expansion condition called β (Arrow, 1948).

Chapters 8 and 9 have established that considering SCFs that need not be pure allows to circumvent Arrow's impossibility theorem and Moulin's no-show paradox on the domain of preferences based on pairwise comparisons. We shall see in this chapter that it also yields a way around impossibility results based on consistency conditions. Two well-known consistency conditions for SCFs are consistency with respect to variable electorates and consistency with respect to components of similar pure outcomes. Population consistency prescribes that every outcome that is chosen by two disjoint electorates should also be chosen by the union of both electorates. The second consistency condition takes into account the structure of preference profiles that it is given by the preferences of the agents. A component is a set of pure outcomes that is an interval in every agent's preference relation over pure outcomes. Components can be thought of as variants or clones of one representative. Cloning consistency prescribes that the probabilities assigned to alternatives outside the component must not depend on the presence of clones of the representative or the preferences over the clones. Composition consistency additionally requires that the probabilities for alternatives inside the component have to be directly proportional to their probabilities when the SCF is applied to the component alone. First, we show that no pure SCF satisfies population consistency and cloning consistency (Theorem 10.1). For non-pure SCFs, population consistency and composition consistency are not only compatible, but even characterize maximal lotteries (Theorem 10.2). Both axioms are required for the characterization. Alternatively, *ML* can be characterized as the unique SCF satisfying population consistency, cloning consistency, and Condorcet consistency (cf. Remark 10.4). All three results assume that SCFs satisfy a number of weak properties some which have normative appeal, while others preclude some sort of irregularity.

10.1 PRELIMINARIES

For the rest of this chapter, we assume that both, the set of alternatives U and the set of agents V are infinite. We consider SCFs on the domain $\mathcal{D} \subset \mathcal{D}^{PC}$ that contains all PC preferences based on a complete, transitive, and asymmetric order over pure outcomes.[31] In the following, a number of basic properties of SCFs are stated that will be used for the characterization of maximal lotteries.

If an SCF is homogeneous, it only depends on the *fraction* of agents reporting a particular preference relation (cf. Section 7.1). Hence, a homogeneous SCF f can be viewed as a function with domain $\Delta^{\mathbb{Q}}(\mathcal{D})$, where the restriction to rational-valued distributions follows from the assumption that electorates have to be finite. Since only homogeneous SCFs will be considered, we assume that SCFs operate on the domain $\mathcal{P}^{\Delta} = \Delta^{\mathbb{Q}}(\mathcal{D})$ from now on. Elements of $\mathcal{P}^{\Delta}$ will be called fractional preference profiles. The specification "fractional" will be omitted whenever it is clear from the context. This representation of preference profiles abstracts away from electorates. Similar models (sometimes even assuming a continuum of agents) have been considered by Young (1974a), Young (1975), Young and Levenglick (1978), Saari (1995), Dasgupta and Maskin (2008), Che and Kojima (2010), and Budish and Cantillion (2012), for example.

Since preferences are assumed to be based on pairwise comparisons, every preference profile is completely determined by its restriction to pure outcomes. The restriction of a preference profile $P \in \mathcal{P}^{\Delta}$

31 All axioms considered in this chapter only reference to the agents' preferences over pure outcomes. Hence, any domain that contains one preference relation for every strict order of pure outcomes would work to characterize *ML* as a function of preferences over pure outcomes. Preferences based on pairwise comparisons are the natural domain of *ML*, however, since it is relative utilitarian (and thus Pareto optimal) on this domain.

to a set of pure outcomes $\underline{A}$, $P|_{\underline{A}}$, is an element of $\Delta^{\mathbb{Q}}(\mathcal{D}|_{\underline{A}})$. For $\hat{\succ} \in \mathcal{D}|_{\underline{A}}$, $P(\hat{\succ})$ denotes the fraction of agents whose preferences over pure outcomes in $\underline{A}$ coincide with $\hat{\succ}$, i.e.,

$$P(\hat{\succ}) = \sum_{\succ \in \mathcal{D}\colon \hat{\succ} \subseteq \succ} P(\succ).$$

Hence, $P|_{\underline{A}}$ can be depicted as a table that gives $P(\hat{\succ})$ for every $\hat{\succ} \in \mathcal{D}|_{\underline{A}}$ with $P(\hat{\succ}) > 0$. The table below shows an example with $A = \{a, b, c\}$.[32]

$1/2$	$1/3$	$1/6$
a	a	b
b	c	c
c	b	a

(6)

For all $x, y \in A$, $P(\underline{x}, \underline{y}) = P(\{(\underline{x}, \underline{y})\})$ is the fraction of agents who prefer $\underline{x}$ to $\underline{y}$ (the set $\{(\underline{x}, \underline{y})\}$ represents the relation on $\{\underline{x}, \underline{y}\}$ with $\underline{x} \succ \underline{y}$). In Example 6, $P(\underline{a}, \underline{b}) = 5/6$.

Independence of *infeasible* alternatives is the choice theoretic analog of independence of irrelevant alternatives (cf. Chapter 8). It requires that an SCF only depends on the preferences over feasible outcomes; the preferences over infeasible outcomes are irrelevant. An SCF f satisfies independence of infeasible alternatives if, for all $P, \hat{P} \in \mathcal{P}^{\Delta}$ and $A \in \mathcal{F}(U)$,

$$f(P, A) = f(\hat{P}, A) \text{ whenever } P|_{\underline{A}} = \hat{P}|_{\underline{A}}. \tag{IIA}$$

Every SCF can be represented by an SWF, which returns a relation over outcomes whose maximal elements are exactly the outcomes returned by the SCF. If one requires that this SWF returns preference relations with convex weak upper contour sets, then sets of maximal elements and hence, the sets of outcomes chosen by the SCF are convex. In particular, this is the case if the SWF returns SSB preferences. Whenever these SSB preferences can be represented by a rational-valued SSB matrix, the set of maximal elements for a finite feasible set is a polytope whose vertices lie in $\mathbb{Q}^U$. Based on these considerations, we require that f chooses convex sets with rational-valued extreme points.[33]

$$f(P, A) \text{ is convex with rational-valued extreme points for all } P \in \mathcal{P}^{\Delta} \text{ and } A \in \mathcal{F}(U). \tag{convexity}$$

32 For $x \in A$, we write x within preference profiles instead of the more clumsy $\underline{x}$ to increase readability.

33 It may well be required that *all* chosen outcomes have to be rational-valued due to the conceptual difficulty with choosing non-rational-valued outcomes that cannot be carried out exactly in practice.

Fishburn (1973, pp. 248–249) argued that the set of outcomes returned by an SCF should be convex because it would be unnatural if two outcomes were socially acceptable while a randomization between them was not (see also Fishburn, 1972, p. 201).

The next condition targets robustness of SCFs with respect to small changes in the preference profile. An SCF f is continuous if it is upper hemi-continuous in the first argument. This prevents that small groups of agents have too much influence on the societal choice, i.e., the images of two preference profiles that are close to each other should be close to each other. This definition of continuity relies on the usage of fractional preference profiles.

$$f(\cdot, A) \text{ is upper hemi-continuous for all } A \in \mathcal{F}(U). \quad \text{(continuity)}$$

It is not clear how to interpret situations in which SCFs return multiple outcomes. One might assume that eventually a single outcome is chosen using some tie-breaking scheme or that choosing multiple outcomes is acceptable as a final result. In order to avoid leaving too much to this issue, we require that non-unique choices constitute an exceptional case. This is captured by the requirement that the set of preference profiles for which a unique outcome is returned is dense in the set of all preference profiles.

$$\{P \in \mathcal{P}^\Delta \colon |f(P, A)| = 1\} \text{ is dense in } \mathcal{P}^\Delta \text{ for all } A \in \mathcal{F}(U). \quad \text{(decisiveness)}$$

None of the conditions introduced above interprets the preference relations in that the preferences of the agents should be correlated with the choices of the SCF. Unanimity states that in the case of one agent and a feasible set containing only two pure outcomes, the less preferred pure outcome should not be chosen uniquely. This condition is weaker than *ex post* efficiency for agendas of size two, which in turn is weaker than Young's faithfulness (Young, 1974b). Formally, f satisfies faithfulness if, for all $P \in \mathcal{P}^\Delta$ and $x, y \in U$,

$$f(P, \{x, y\}) \neq \{\underline{y}\} \text{ whenever } P(\underline{x}, \underline{y}) = 1. \quad \text{(unanimity)}$$

An SCF that satisfies homogeneity, independence of infeasible alternatives, convexity, continuity, decisiveness, and unanimity is called a proper SCF.

Most SCFs considered in the literature are proper SCFs. The most well-known example is *random dictatorship (RD)*, which chooses the outcome that assigns to every alternative the probability of it being ranked first by an agent chosen uniformly at random. Formally, for all $P \in \mathcal{P}^\Delta$ and $A \in \mathcal{F}(U)$,

$$RD(P, A) = \left\{ \sum_{\succ \in \mathcal{D}|_{\underline{A}}} P(\succ) \cdot \max_{\succ} \underline{A} \right\}, \qquad \text{(random dictatorship)}$$

where $\max_\succ \underline{A}$ denotes the unique pure outcome $\underline{x}$ such that $\underline{x} \succ p$ for all $p \in \Delta(A) \setminus \{\underline{x}\}$. For the preference profile P given in Example 6 and $A = \{a, b, c\}$,

$$RD(P, A) = \{5/6\, \underline{a} + 1/6\, \underline{b}\}.$$

It is clear from the definition that *RD* satisfies homogeneity, independence of infeasible alternatives, and unanimity. Since *RD* is single-valued, it is trivially decisive and convex-valued. It is also easily verified that *RD* satisfies continuity.

10.2 POPULATION CONSISTENCY AND COMPOSITION CONSISTENCY

We require collective choices to satisfy two choice consistency conditions called population consistency and composition consistency. Population consistency relates choices from varying electorates to each other. Given some fixed agenda, it requires that every outcome that is chosen by two disjoint electorates is also chosen by the union of both electorates. When considering fractional preference profiles, the union of two preference profiles on disjoint electorates amounts to a convex combination of both profiles. For example, consider the two preference profiles $\hat{P}$ and $\bar{P}$ whose restriction to $A = \{a, b, c\}$ is given below.

$$\begin{array}{cc} 1/2 & 1/2 \\ \hline a & b \\ b & c \\ c & a \\ \\ \multicolumn{2}{c}{\hat{P}|_{\underline{A}}} \end{array} \qquad \begin{array}{cc} 1/2 & 1/2 \\ \hline a & b \\ c & c \\ b & a \\ \\ \multicolumn{2}{c}{\bar{P}|_{\underline{A}}} \end{array} \qquad \begin{array}{ccc} 1/4 & 1/4 & 1/2 \\ \hline a & a & b \\ b & c & c \\ c & b & a \\ \\ \multicolumn{3}{c}{(1/2\,\hat{P} + 1/2\,\bar{P})|_{\underline{A}}} \end{array} \qquad (7)$$

Population consistency prescribes that every outcome that is chosen by both $\hat{P}$ and $\bar{P}$ given the agenda A (say $1/2\, \underline{a} + 1/2\, \underline{b}$) is also chosen when $\hat{P}$ and $\bar{P}$ are merged. Formally, an SCF f satisfies population consistency if for all $\hat{P}, \bar{P} \in \mathcal{P}^\Delta$, $A \in \mathcal{F}(U)$, and $\lambda \in [0,1] \cap \mathbb{Q}$,

$$f(\hat{P}, A) \cap f(\bar{P}, A) \quad \subseteq \quad f(P, A), \qquad \text{(population consistency)}$$

where $P = \lambda\hat{P} + (1-\lambda)\bar{P}$. Hence, the set of profiles for which a given outcome is returned (possibly among other outcomes) has to be convex. Observe that population consistency is agnostic to the type of output of f. Its definition does not need to be adjusted for other types of aggregation functions. Population consistency and variants thereof have been considered plenty in the literature. Reinforcement requires the above set inclusion to hold with equality whenever the left-hand side is non-empty. It is thus stronger than population consistency. Reinforcement was introduced by Young and is the driving force in his characterizations of Borda's rule (Young, 1974b) and scoring rules (Young, 1975).[34] Variants of reinforcement have been used by Smith (1973) to characterize SWFs based on scoring rules and Fine and Fine (1974) to characterize positional rules. The frequent occurrence of population consistency and its variants in different contexts in the social choice literature highlights its compelling nature (see also Young, 1974a; Fishburn, 1978; Young and Levenglick, 1978; Saari, 1990; Saari, 1995; Myerson, 1995; Congar and Merlin, 2012).

Composition consistency relates choices from different agendas for a fixed preference profile to each other. It only restricts the choices for preference profiles that are decomposable, however. An agenda $B \in \mathcal{F}(U)$ is a component in $P \in \mathcal{P}^\Delta$ if $\underline{B}$ constitutes an interval in every agent's preference relation over pure outcomes, i.e., in every $\succ \in \mathcal{D}|_{\underline{U}}$ with $P(\succ) > 0$. The set $\underline{B}$ is an interval in $\succ$ if pure outcomes in $\underline{B}$ cannot be distinguished by their relationship to pure outcomes outside of $\underline{B}$ or, formally, if for all $x, y \in B$ and $z \in U \setminus B$, $\underline{x} \succ \underline{z}$ if and only if $\underline{y} \succ \underline{z}$. Alternatives within a component are called clones. For example, consider the preference profile P with component $B = \{b, b'\}$ and let $A = \{a, b, b'\}$ and $\hat{A} = \{a, b\}$.

$$
\begin{array}{ccc}
1/3 & 1/6 & 1/2 \\
\hline
a & a & b \\
b' & b & b' \\
b & b' & a
\end{array}
\qquad
\begin{array}{cc}
1/2 & 1/2 \\
\hline
a & b \\
b & a
\end{array}
\qquad
\begin{array}{cc}
1/3 & 2/3 \\
\hline
b' & b \\
b & b'
\end{array}
\qquad (8)
$$

$$P|_{\underline{A}} \qquad P|_{\underline{\hat{A}}} \qquad P|_{\underline{B}}$$

Composition consistency states that the choice from P for A can be decomposed into two choices, one for $\hat{A}$ and one for B. The choice for A is obtained by first making the choice for $\hat{A}$ and then substituting the choice for B therein. Let $\hat{A}, B \in \mathcal{F}(U)$ such that $\hat{A} \cap B = \{b\}$ and $A = \hat{A} \cup B$. Then, an SCF f satisfies composition consistency if, for all $P \in \mathcal{P}^\Delta$ such that B is a component in P,

$$f(P, \hat{A}) \times_b f(P, B) = f(P, A). \qquad \text{(composition consistency)}$$

34 Reinforcement is called "consistency" by Young (1974b) and Young (1975).

In Example 8 above, composition consistency implies that $1/2\,\underline{a} + 1/2\,\underline{b} \in f(P, \hat{A})$ and $2/3\,\underline{b} + 1/3\,\underline{b}' \in f(P, B)$ if and only if $1/2\,\underline{a} + 1/3\,\underline{b} + 1/6\,\underline{b}' \in f(P, A)$. Note that by independence of infeasible alternatives, the implications of composition consistency hold whenever the pure outcomes in $\underline{B}$ are indistinguishable by pure outcomes in $\underline{A} \setminus \underline{B}$ even when B is not a component in P. Composition consistency was introduced by Laffond et al. (1996), who examined various tournament solutions and other well-known SCFs with regards to composition consistency. It was further studied by Laslier (1996), Laslier (1997), Brandt (2011), Brandt et al. (2011), and Horan (2013). Example 8 shows that *RD* violates composition consistency, since $RD(P, \hat{A}) \times_b RD(P, B) = (1/2\,\underline{a} + 1/2\,\underline{b}) \times_b (2/3\,\underline{b} + 1/3\,\underline{b}') \neq 1/2\,\underline{a} + 1/2\,\underline{b} = RD(P, A)$.

Composition consistency implies that the probabilities assigned to non-clones (alternatives in $A \setminus \{b\}$) must not change by cloning b. This weakening of composition consistency is called cloning consistency. As before, let $\hat{A}, B \in \mathcal{F}(U)$ such that $\hat{A} \cap B = \{b\}$ and $A = \hat{A} \cup B$. An SCF f satisfies cloning consistency if, for all $P \in \mathcal{P}^\Delta$ such that B is a component in P,

$$f(P, \hat{A})_{\hat{A}\setminus\{b\}} = f(P, A)_{\hat{A}\setminus\{b\}}. \qquad \text{(cloning consistency)}$$

Cloning consistency as defined here was proposed by Tideman (1987) and further studied by Zavist and Tideman (1989). Similar conditions have already been considered by Chernoff (1954), Arrow and Hurwicz (1972), and Maskin (1979) in the decision theory literature, where it is called deletion of repetitious states. Moulin (1986) considered cloning consistency for choice functions that are based on binary trees. Since cloning an alternative b has no effect on first rank nominations of pure outcomes in $\underline{A} \setminus \{\underline{b}\}$, we can infer that *RD* satisfies cloning consistency. In Example 8, $RD(P, \hat{A})_{\hat{A}\setminus\{b\}} = \{1/2\,\underline{a}\} = RD(P, A)_{\hat{A}\setminus\{b\}}$.

10.3 PURE SOCIAL CHOICE FUNCTIONS

The social choice literature displays two streams of research, whose origins can be traced back to Borda (1784) and Condorcet (1785): scoring rules, of which Borda's rule is one representative, and Condorcet extensions, i.e., Condorcet consistent SCFs. A number of results have shown that population consistency is essentially the characterizing property of scoring rules (see, e.g., Smith, 1973; Young, 1974a; Young, 1975). Condorcet observed that Borda's rule may fail to select a Condorcet winner, and hence, violates Condorcet consistency. Young and Levenglick (1978) have shown that this shortcoming is shared with all pure SCFs that satisfy population consistency.[35] Condorcet consistency seems to be more compatible with composition consistency, since a number of SCFs that are known to satisfy composition consistency are Condorcet extensions (cf. Laffond et al., 1996). Laslier (1996) has shown that no Pareto optimal rank-based pure SCF—a generalization of scoring rules—can satisfy composition consistency. Hence, the ideas of Borda and Condorcet are largely incompatible for pure SCFs. One exception is the Pareto rule, which returns all *ex post* efficient outcomes. It is not a proper SCF in the sense defined here, however, since it violates decisiveness. The following theorem shows that the conflict between the ideas of Borda and Condorcet prevails, even when weakening composition consistency to cloning consistency.

THEOREM 10.1

> No proper pure SCF satisfies cloning consistency and population consistency.

For non-pure SCFs, population consistency and cloning consistency are compatible with each other as witnessed by *RD*, for example. If cloning consistency is strengthened to composition consistency, these properties uniquely characterize *ML*.

10.4 CHARACTERIZATION OF MAXIMAL LOTTERIES

We start our characterization of *ML* by considering the case of two-element feasible sets, e.g., $A = \{a, b\}$. For pure SCFs, majority rule is the only reasonable SCF for this case (cf. May, 1952; Dasgupta and Maskin, 2008). For possibly non-pure SCFs, there are a number of interesting SCFs, even on two-element feasible sets (see, e.g., Saunders, 2010; Fishburn and Gehrlein, 1977). By independence of infeasible alternatives, the choice of a proper SCF f for the feasible set A can only depend on the fraction of agents who prefer $\underline{a}$ to $\underline{b}$. Hence, $f(\cdot, A)$ can be seen as a correspondence from the unit interval to the unit interval. By convexity, continuity, and decisiveness, this correspondence has to be convex-valued with rational-valued extreme points, upper hemi-continuous, and single-valued on a dense subset. Unanimity prohibits that 0 gets mapped to $\{1\}$ and that 1 gets mapped to $\{0\}$. When additionally requiring population consistency, it follows that the function has to be monotonically increasing. Composition consistency has no implications when only considering two-element feasible sets.

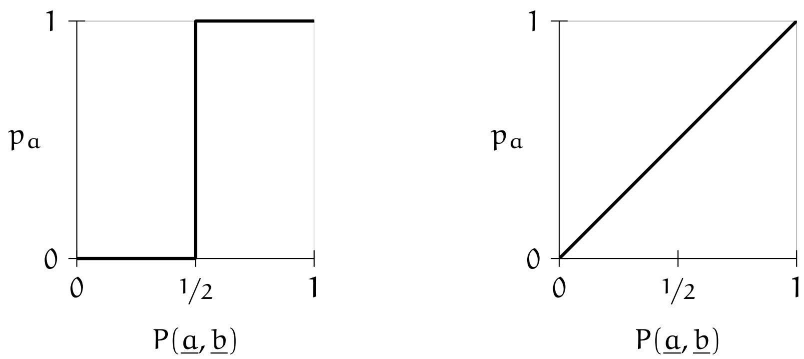
Maximal lotteries can be seen as the natural extension of majority rule, since it uniquely chooses the pure outcome that is preferred

35 Theorem 2 by Young and Levenglick (1978) actually assumes reinforcement, but its proof can be made work for population consistency as defined here with minor adjustments.

by a majority. Hence, it completely suppresses minorities. Random dictatorship on the other hand is perfectly proportional. For all $P \in \mathcal{P}^\Delta$,

$$ML(P,\{a,b\}) = \begin{cases} \{\underline{a}\} & \text{if } P(\underline{a},\underline{b}) > 1/2, \\ \{\underline{b}\} & \text{if } P(\underline{a},\underline{b}) < 1/2, \\ \Delta(\{a,b\}) & \text{otherwise,} \end{cases} \qquad \text{and}$$

$$RD(P,\{a,b\}) = \{P(\underline{a},\underline{b})\,\underline{a} + P(\underline{b},\underline{a})\,\underline{b}\}.$$



(a) Maximal lotteries **(b)** Random dictatorship

Figure 10.1: Maximal lotteries and random dictatorship on two-element feasible sets. Here, p_a denotes the probability assigned to a by the corresponding outcome in $ML(P,\{a,b\})$ and $RD(P,\{a,b\})$, respectively.

Fishburn and Gehrlein (1977) compared these two SCFs on two-element feasible sets on the basis of expected agent satisfaction and found that the simple majority rule outperforms the proportional rule. Curiously, when allowing for three-element feasible sets, population consistency and composition consistency characterize majority rule and thus, maximal lotteries on two-element feasible sets. If arbitrary feasible sets are allowed, the following characterization of maximal lotteries is obtained.

THEOREM 10.2

A proper SCF f satisfies population consistency and composition consistency if and only if $f = ML$.

As a lighthouse to the reader, we give a short outline of the proof of Theorem 10.2. We start by showing that maximal lotteries is a proper SCF that satisfies population consistency and composition consistency. This follows from properties of maximin strategies in symmetric zero-sum games in a relatively straightforward way.

The converse direction is divided into two statements. The key part is to show that every proper SCF f that satisfies population consistency and composition consistency has to choose a subset of maximal lotteries. First, it is shown that f has to be equal to maximal lotteries on two-element feasible sets, which requires applications of composition consistency to three-element feasible sets. Then, it is assumed for contradiction that f returns an outcome that is not maximal for some preference profile and some feasible set. This yields a preference profile that admits a Condorcet winner on a possibly larger feasible set, say A, and for which f returns the uniform distribution over a non-singleton set of alternatives, say A'. The existence thereof is critical to construct a set of preference profiles whose affine hull contains $\mathcal{P}^\Delta$ and for each of which f returns the uniform distribution over A'. Population consistency allows to choose this set arbitrarily close to the uniform distribution on $\mathcal{D}|_{\underline{A}}$. Along the way we show that f has to be Condorcet consistent for all preference profiles that are close to this uniform profile. Hence, there is a profile with a strict Condorcet winner (close to the uniform profile) such that f returns the uniform distribution over A' as well as the Condorcet winner for every profile in a neighborhood of this Condorcet profile. This contradicts decisiveness. Lastly, we show that f has to return all maximal lotteries. To this end, we show that for every preference profile and feasible set, every vertex of the set of maximal lotteries can be approached by a sequence of maximal lotteries for a sequence of preference profiles that approaches the original profile. From $f \subseteq \mathit{ML}$ and continuity, we obtain that f has to select all these vertices in the original preference profile. Convexity implies that $f = \mathit{ML}$.

10.5 CONCLUDING REMARKS

We conclude this chapter with a number of remarks.

Remark 10.1 *(Independence of axioms)*

Population consistency and composition consistency are both required for the characterization of *ML* in Theorem 10.2. Random dictatorship satisfies population consistency, but violates composition consistency. The same is true for Borda's rule. When defining ML^3 as choosing the set of outcomes that correspond to maximin strategies in $(M_A^P)^3$ (where the power is taken for each entry separately) it is a proper SCF that satisfies composition consistency but violates population consistency.[36]

Also, continuity, decisiveness, and unanimity, which are part of the definition of proper SCFs, are required. Continuity is needed because the relative interior of *ML*, known as strict maximal lotteries (Aziz et al., 2018), satisfies all remaining axioms. When not requiring decisiveness, the Pareto rule, which returns all *ex post* efficient outcomes, is consistent with the remaining axioms. The SCF that returns all *minimal* lotteries violates una-

36 Such variants of maximal lotteries have also been considered by Fishburn (1984b).

nimity but none of the other axioms. Homogeneity is essential to define continuity and decisiveness. Whether independence of infeasible alternatives and convexity are needed is open.

REMARK 10.2 *(Size of the Universe)*
The proof of Theorem 10.2 exploits the infinity of the universe U. The SCF ML^3 as defined in Remark 10.1 satisfies population consistency when there are at most three alternatives, but violates population consistency if there are more alternatives. This implies that the statement of Theorem 10.2 requires the universe to contain at least four alternatives.

REMARK 10.3 *(Strong population consistency)*
Maximal lotteries does not satisfy reinforcement, a strengthening of population consistency in which the set inclusion is replaced with equality whenever the left-hand side is non-empty (see Section 10.2). Consider the following two preference profiles $\hat{P}$ and $\bar{P}$ and $A = \{a, b, c\}$.

1/3	1/3	1/3
a	b	c
b	c	a
c	a	b

$\hat{P}|_{\underline{A}}$

1/3	1/3	1/3
c	a	b
b	c	a
a	b	c

$\bar{P}|_{\underline{A}}$

It can be checked that $ML(\hat{P}, A) = ML(\bar{P}, A) = \{1/3\, \underline{a} + 1/3\, \underline{b} + 1/3\, \underline{c}\}$. Hence, $ML(\hat{P}, A) \cap ML(\bar{P}, A)$ is non-empty. Reinforcement implies that for $P = 1/2\, \hat{P} + 1/2\, \bar{P}$, $ML(P, A) = \{1/3\, \underline{a} + 1/3\, \underline{b} + 1/3\, \underline{c}\}$. However, $ML(P, A) = \Delta(A)$, since $M^P_A = 0$. To demonstrate the strength of reinforcement, observe that it is even violated by the Pareto rule (cf. Remark 10.1).

REMARK 10.4 *(Cloning consistency and Condorcet consistency)*
Theorem 10.2 does not hold if composition consistency is weakened to cloning consistency, since then for example random dictatorship also qualifies. If however Condorcet consistency is assumed in addition, the axioms single-out maximal lotteries again.

THEOREM 10.3
A proper SCF f satisfies population consistency, cloning consistency, and Condorcet consistency if and only if $f = ML$.

At the end of Section 10.8 we sketch how the proof of Theorem 10.2 can be adjusted to prove Theorem 10.3. As above, all three axioms are required for the characterization.

Remark 10.5 *(Relationship to the characterization of maximin strategies)*
The results in this chapter are related to the characterization of maximin strategies in Chapter 5. Every solution concept g induces an SCF f by defining $f(P, A) = g(M^P, A)$ for all $P \in \mathcal{P}^\Delta$ and $A \in \mathcal{F}(U)$. With this definition consistency of g is equivalent to population consistency of f, since M_A^P is a symmetric zero-sum game (cf. Remark 5.3). Consequentialism is a weakening of composition consistency and very similar to cloning consistency. Rationality of a solution concept is implied by Condorcet consistency of the corresponding SCF. Hence, Theorems 10.2 and 10.3 can be seen as the equivalents of Theorem 5.3 for SCFs. Stronger axioms are required for the characterization of maximal lotteries, since the choices of an SCF need not be solely based on M^P. Readers may find the proof of Theorem 5.3 helpful to understand the structure of the proof of Theorem 10.2.

10.6 Cloning consistency implies neutrality

Recall that U is an infinite set of alternatives. For convenience we will assume that $\mathbb{N} \subseteq U$. As a tool for the upcoming proofs, we show that cloning consistency implies neutrality, a well-known symmetry condition, for proper SCFs. Neutrality requires that all alternatives are treated equally in the sense that renaming alternatives is reflected by the same renaming in the set of outcomes. Formally, an SCF f satisfies neutrality if

$$(f(P, A))_\pi = f(P^\pi, \pi(A)) \text{ for all } \pi \in \Pi(U),\, A \in \mathcal{F}(U), \text{ and } P \in \mathcal{P}^\Delta. \qquad \text{(neutrality)}$$

The proof of Lemma 10.4 exploits the fact that the probability assigned to an alternative stays fixed when replacing another alternative by a component of size 2 for cloning consistent SCFs.

Lemma 10.4
Every proper SCF that satisfies cloning consistency satisfies neutrality.

Proof. Let f be a proper SCF satisfying cloning consistency. Let $\pi \in \Pi(U)$, $A = \{a_1, \dots, a_m\} \in \mathcal{F}(U)$, and $P \in \mathcal{P}^\Delta$. We have to show that $(f(P, A))_\pi = f(P^\pi, \pi(A))$. To this end, let $p^A \in f(P, A)$. Since U is infinite, there is $B = \{b_1, \dots, b_m\} \in \mathcal{F}(U)$ such that $B \cap A = \emptyset$ and $B \cap \pi(A) = \emptyset$. Now let $\hat{P} \in \mathcal{P}^\Delta$ such that $\hat{P}|_{\underline{A}} = P|_{\underline{A}}$ and $\{a_i, c_i\}$ is a component in $\hat{P}$ for all $i \in [m]$. By IIA, we have that $p^A \in f(\hat{P}, A)$. We now repeatedly apply cloning consistency to a_i and the component $\{a_i, b_i\}$ for all $i \in [m]$. For $i = 1$, this implies that

$$f(\hat{P}, A)_{A\setminus\{a_1\}} = f(\hat{P}, A \cup \{b_1\})_{A\setminus\{a_1\}}.$$

Hence, there is $\hat{p} \in f(P, A \cup \{b_1\})$ such that $\hat{p}_{a_i} = p^A_{a_i}$ for all $i \in \{2, \ldots, m\}$. This implies that $\hat{p}_{a_1} + \hat{p}_{b_1} = p^A_{a_1}$. Repeated application of cloning consistency to a_i and $\{a_i, b_i\}$ for all $i \in \{2, \ldots, m\}$, implies that there is $p^{AB} \in f(\hat{P}, A \cup B)$ such that $p^{AB}_{a_i} + p^{AB}_{b_i} = p^A_{a_i}$ for all $i \in [m]$. Applying cloning consistency analogously to b_i and $\{a_i, b_i\}$ for all $i \in [m]$ yields that there is $p^B \in f(\hat{P}, B)$ such that $p^B_{b_i} = p^{AB}_{a_i} + p^{AB}_{b_i} = p^A_{a_i}$ for all $i \in [m]$. Finally, let $\bar{P} \in \mathcal{P}^\Delta$ such that $\bar{P}|_{\underline{B}} = \hat{P}|_{\underline{B}}$ and $\{\pi(a_i), b_i\}$ is a component in $\bar{P}$ for all $i \in [m]$. By IIA, we have that $p^B \in f(\bar{P}, B)$. As before, it follows from cloning consistency that there is $p^{\pi(A)} \in f(\bar{P}, \pi(A))$ with $p^{\pi(A)}_{\pi(a_i)} = p^B_{b_i} = p^A_{a_i}$ for all $i \in [m]$. Hence, $p^{\pi(A)} = p^A_\pi$ by construction. Since $\bar{P}|_{\underline{\pi(A)}} = P^\pi|_{\underline{\pi(A)}}$ by construction of $\bar{P}$, we have $p^{\pi(A)} \in f(P^\pi, \pi(A))$ by IIA. Hence, $(f(P, A))_\pi \subseteq f(P^\pi, \pi(A))$. The fact that $f(P^\pi, \pi(A)) \subseteq (f(P, A))_\pi$ follows from application of the above to P^π and π^{-1}. □

10.7 PURE SOCIAL CHOICE FUNCTIONS: PROOFS

We now prove Theorem 10.1

THEOREM 10.1
No proper pure SCF satisfies cloning consistency and population consistency.

Proof. Assume for contradiction that f is a proper pure SCF that satisfies population consistency and cloning consistency. By cloning consistency and Lemma 10.4, f satisfies neutrality. Observe that if for $a, b \in A$ and $\hat{P}, \bar{P} \in \mathcal{P}^\Delta$ with $\hat{P}(\underline{a}, \underline{b}) = 1$ and $\bar{P}(\underline{b}, \underline{a}) = 1$, we have $f(\hat{P}, \{a, b\}) = \Delta(\{a, b\})$, then by neutrality, $f(\bar{P}, \{a, b\}) = \Delta(\{a, b\})$. Population consistency implies that $f(P, \{a, b\}) = \Delta(\{a, b\})$ for all $P \in \mathcal{P}^\Delta$, which contradicts decisiveness. Hence, by unanimity, $f(\hat{P}, \{a, b\}) = \{\underline{a}\}$.

Now let $A = \{a, b, c\}$ and consider the profiles $P^1, \ldots, P^6 \in \mathcal{P}^\Delta$ as depicted below. We will construct a full-dimensional subset of $\mathcal{P}^\Delta$ for which f chooses $\Delta(\{a, b\})$.

1/3	1/3	1/3
a	b	c
b	c	a
c	a	b

$P^1|_{\underline{A}}$

1/2	1/2
a	b
c	c
b	a

$P^2|_{\underline{A}}$

1/2	1/2
a	b
b	c
c	a

$P^3|_{\underline{A}}$

1/2	1/2
b	a
a	c
c	b

$P^4|_{\underline{A}}$

It follows from neutrality that $f(P^1, A) = \Delta(A)$. Again, by neutrality, $f(P^2, \{a,b\}) = \Delta(\{a,b\})$. Notice that $\{b,c\}$ is a component in P^2. Hence, by cloning consistency, $\underline{a} \in f(P^2, A)$. Neutrality implies that $\underline{b} \in f(P^2, A)$ and thus, by convexity, $\Delta(\{a,b\}) \subseteq f(P^2, A)$.

By neutrality, $f(P^3, \{a,b\}) = \Delta(\{a,b\})$. Notice that $\{b,c\}$ is a component in P^3. Hence, by cloning consistency, $\underline{a} \in f(P^3, A)$. Assume for contradiction that $\underline{b} \notin f(P^3, A)$. From before, we know that $\underline{b} \in f(P^3, \{a,b\})$. Then cloning consistency implies that $\underline{c} \in f(P^3, A)$. Using this, neutrality implies that $\underline{c} \in f(P^4, A)$. Let $P^{34}, P^c \in \mathcal{P}^\Delta$ as depicted below.

$1/4$	$1/4$	$1/4$	$1/4$		$1/2$	$1/2$
a	b	b	a		c	c
b	c	a	c		a	b
c	a	c	b		b	a

$P^{34}|_{\underline{A}}$ $\qquad$ $P^c|_{\underline{A}}$

Population consistency applied to P^3 and P^4 yields that $\underline{c} \in f(P^{34}, A)$. By unanimity, $f(P^c, \{a,c\}) = \{\underline{c}\}$ and $\{a,b\}$ is a component. Cloning consistency implies that $f(P^c, A) = \{\underline{c}\}$. If, for every $\epsilon > 0$, there is $P \in B_\epsilon(P^c)$ such that $\underline{a} \in f(P, A)$, then, by continuity, $\underline{a} \in f(P^c, A)$. Similarly for b. Hence, there is $\epsilon > 0$ such that $f(P, A) = \{\underline{c}\}$ for all $P \in B_\epsilon(P^c)$. Now let $P \in \mathcal{P}^\Delta$ such that $P|_{\underline{A}} \in B_{\epsilon/3}(\mathrm{uni}(\mathcal{D}|_{\underline{A}}))$. Then, there are $\hat{P} \in B_\epsilon(P^c)$ and $\lambda \in [0,1]$ such that $P = \lambda P^{34} + (1-\lambda)\hat{P}$. Population consistency implies that $\underline{c} \in f(P, A)$. Since the choice of P was arbitrary, neutrality implies that $\{a,b,c\} \subseteq f(P, A)$ for all $P \in \mathcal{P}^\Delta$ with $P|_{\underline{A}} \in B_{\epsilon/3}(\mathrm{uni}(\mathcal{D}|_{\underline{A}}))$. This contradicts decisiveness and we have that $\underline{b} \in f(P^3, A)$. Convexity implies that $\Delta(\{a,b\}) \subseteq f(P^3, A)$. By neutrality, $\Delta(\{a,b\}) \subseteq f(P^4, A)$.

Now, for $\lambda \in [0, 1/2]$, consider $P^{5,\lambda}$ and $P^{6,\lambda}$ depicted below.

$1/2-\lambda$	$1/2-\lambda$	2λ		$1/2-\lambda$	$1/2-\lambda$	2λ
a	c	b		b	c	a
b	b	a		a	a	b
c	a	c		c	b	c

$P^{5,\lambda}|_{\underline{A}}$ $\qquad$ $P^{6,\lambda}|_{\underline{A}}$

By neutrality and convexity, we have that $f(P^{5,0}, \{a,b\}) = \Delta(\{a,b\})$. Notice that $\{b,c\}$ is a component is $P^{5,0}$. Hence, by cloning consistency, $\underline{a} \in f(P^{5,0}, A)$. By neutrality, $\{\underline{a}, \underline{c}\} \subseteq f(P^{5,0}, A)$. By unanimity, we have that $f(P^{5,1/2}, A) = \{\underline{b}\}$. Again by unanimity, $f(P^{5,1/2}, \{a,c\}) = \{\underline{a}\}$. Thus, population consistency implies that $\underline{a} \in f(P^{5,\lambda}, \{a,c\})$ for all $\lambda \in [0, 1/2]$. If $\underline{c} \in f(P^{5,\lambda^*}, \{a,c\})$ for some $\lambda^* \in (0, 1/2]$, then, by population consistency, $\{\underline{a}, \underline{c}\} \subseteq f(P^{5,\lambda}, \{a,c\})$ for all $\lambda \in [0, \lambda^*]$, which contradicts decisiveness. Hence, $f(P^{5,\lambda}, \{a,c\}) = \{\underline{a}\}$ for all $\lambda \in (0, 1/2]$. Since $\{a,b\}$ is a component in $P^{5,\lambda}$, cloning consistency implies that

$\underline{c} \notin f(P^{5,\lambda}, A)$ for all $\lambda \in (0, 1/2]$. If $\underline{a} \in f(P^{5,\lambda}, A)$ for λ arbitrarily close to $1/2$, then continuity implies that $\underline{a} \in f(P^{5,1/2}, A)$. This contradicts $f(P^{5,1/2}, A) = \{\underline{b}\}$. Hence, by convexity, there is $\lambda^* \in [0, 1/2)$ such that $\Delta(\{a, b\}) \subseteq f(P^{5,\lambda^*}, A)$. By neutrality, $\Delta(\{a, b\}) \subseteq f(P^{6,\lambda^*}, A)$. Let $P^i = P^{i,\lambda^*}$ for $i \in \{5, 6\}$.

Every $P^i|_{\underline{A}}$ is a vector in the five-dimensional unit simplex $\mathcal{P}^\Delta|_{\underline{A}}$ in $\mathbb{Q}^6$. The corresponding vectors are depicted below.

$$\begin{pmatrix} P^1|_{\underline{A}} \\ P^2|_{\underline{A}} \\ P^3|_{\underline{A}} \\ P^4|_{\underline{A}} \\ P^5|_{\underline{A}} \\ P^6|_{\underline{A}} \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 - \lambda & 0 & 0 & 0 & 1/2 - \lambda & 2\lambda \\ 2\lambda & 0 & 1/2 - \lambda & 0 & 0 & 1/2 - \lambda \end{pmatrix}$$

It can be checked that $P^1|_{\underline{A}}, \ldots, P^6|_{\underline{A}}$ are affinely independent for all $\lambda^* \in [0, 1/2)$, i.e., $\dim(\{P^1|_{\underline{A}}, \ldots, P^6|_{\underline{A}}\}) = 5$. Moreover, $\Delta(\{a, b\}) \subseteq f(P^i, A)$ for all $i \in [6]$.

Hence, by IIA, $\{P \in \mathcal{P}^\Delta : |f(P, A)| = 1\}$ is not dense in $\mathcal{P}^\Delta$ at $1/6\, P^1 + \cdots + 1/6\, P^6$, which contradicts decisiveness of f. $\square$

10.8 CHARACTERIZATION OF MAXIMAL LOTTERIES: PROOFS

In this section we prove Theorem 10.2. The high-level structure of the proof is described after Theorem 10.2 in Section 10.8.

10.8.1 *ML* Satisfies Population consistency and Composition consistency

We first show that *ML* satisfies all axioms required in Theorem 10.2. This statement is split into two lemmas.

Lemma 10.5

ML is a proper SCF.

Proof. The fact that *ML* satisfies homogeneity, IIA, and unanimity is clear by definition.

The fact that $f(P, A)$ is convex for every $P \in \mathcal{P}^\Delta$ and $A \in \mathcal{F}(U)$ follows from convexity of the set of maximin strategies for all (symmetric) zero-sum games.

ML is continuous, since the correspondence that maps a (symmetric) zero-sum game to the set of maximin strategies is (upper hemi-) continuous.

ML satisfies decisiveness. Let $P \in \mathcal{P}^\Delta$ and $A \in \mathcal{F}(U)$. It is easy to see that, for every $\epsilon > 0$, we can find $\hat{P} \in B_\epsilon(P) \cap \mathcal{P}^\Delta$ and $k \in \mathbb{N}$ such that $k\hat{P}(\underline{x}, \underline{y})$ is an odd integer for all $x, y \in A$ with $x \neq y$. Laffond et al. (1997) have shown that every symmetric zero-sum game whose off-diagonal entries are odd integers admits a unique maximin strategy. Hence, $|f(\hat{P}, A)| = 1$ and f is decisive.

ML obviously satisfies unanimity by definition. □

LEMMA 10.6
ML satisfies population consistency and composition consistency.

Proof. ML satisfies population consistency. Let $\hat{P}, \bar{P} \in \mathcal{P}^{\Delta}$, $A \in \mathcal{F}(U)$, and $p \in ML(\hat{P}, A) \cap ML(\bar{P}, A)$. Then, by definition of *ML*, $p^t M^{\hat{P}} q \geqslant 0$ and $p^t M^{\bar{P}} q \geqslant 0$ for all $q \in \Delta(A)$. Hence, for all $\lambda \in [0,1]$ and $q \in \Delta(A)$,

$$p^t \left(\lambda M^{\hat{P}} + (1-\lambda) M^{\bar{P}}\right) q = \lambda \underbrace{p^t M^{\hat{P}} q}_{\geqslant 0} + (1-\lambda) \underbrace{p^t M^{\bar{P}} q}_{\geqslant 0} \geqslant 0.$$

This implies that $p \in ML(\lambda \hat{P} + (1-\lambda)\bar{P}, A)$.

ML satisfies composition consistency. Let $P \in \mathcal{P}^{\Delta}$, $\hat{A}, B \in \mathcal{F}(U)$ such that $\hat{A} \cap B = \{b\}$, $A = \hat{A} \cup B$ such that B is a component in P. To simplify notation, let $C = A \setminus B$ and $M = M^P$. Notice first that M takes the following form for some $v \in \mathbb{Q}^{A \setminus B}$:

$$M_A = \left(\begin{array}{c|c} \begin{matrix} \\ M_C \\ \\ \end{matrix} & \begin{matrix} | & & | \\ v & \dots & v \\ | & & | \end{matrix} \\ \hline \begin{matrix} — & (-v^t) & — \\ & \vdots & \\ — & (-v^t) & — \end{matrix} & M_B \end{array}\right).$$

Let $p \in ML(P, \hat{A}) \times_b ML(P, B)$. Then, there are $p^{\hat{A}} \in ML(P, \hat{A})$ and $p^B \in ML(P, B)$ such that $p = p^{\hat{A}} \times_b p^B$. Then, for all $q \in \Delta(A)$,

$$\begin{aligned} p^t M q &= p_C^t M_C q_C + \|p_B\| (-v)^t q_C + p_C^t v \|q_B\| + p_B^t M_B q_B \\ &= (p_C, \|p_B\|)^t M_{\hat{A}} (q_C, \|q_B\|)^t + p_B^t M_B q_B \\ &= \underbrace{(p_{\hat{A}}^{\hat{A}})^t M_{\hat{A}} (q_C, \|q_B\|)^t}_{\geqslant 0} + \|p_B\| \underbrace{(p_B^B)^t M_B q_B}_{\geqslant 0} \geqslant 0, \end{aligned}$$

since $p^{\hat{A}} \in ML(P, \hat{A})$ and $p^B \in ML(P, B)$. Hence $p \in ML(P, A)$.

For the other direction, let $p \in ML(P, A)$. We have to show that there are $p^{\hat{A}} \in ML(P, \hat{A})$ and $p^B \in ML(P, B)$ such that $p = p^{\hat{A}} \times_b p^B$. First, if $\|p_B\| = 0$ let $p^{\hat{A}} = p$ and $p^B \in ML(P, B)$ be arbitrary. Then, $p = p^{\hat{A}} \times_b p^B$ and $p^{\hat{A}} \in ML(P, \hat{A})$. Otherwise, let $p^{\hat{A}} \in \Delta(\hat{A})$ and

$p^B \in \Delta(B)$ such that $p^{\hat{A}}_{\hat{A}} = (p_C, \|p_B\|)$ and $p^B_B = p_B/\|p_B\|$. Then, $p = p^{\hat{A}} \times_b p^B$ and, for all $q \in \Delta(\hat{A})$,

$$\begin{aligned}(p^{\hat{A}})^t M q &= p_C^t M_C q_C + \|p_B\|(-\nu)^t q_C + p_C^t \nu q_b \\ &= p_C^t M_C q_C + \|p_B\|(-\nu)^t q_C + p_C^t \nu q_b \\ &+ \frac{q_b}{\|p_B\|} \underbrace{p_B^t M_B p_B}_{=0} \\ &= p_A^t M_A (q_C, \frac{q_b}{\|p_B\|} p_B)^t \geqslant 0,\end{aligned}$$

since $p \in \mathit{ML}(P, A)$. Hence, $p^{\hat{A}} \in \mathit{ML}(P, \hat{A})$. For all $q \in \Delta(B)$,

$$\begin{aligned}\|p_B\|^2 (p^B)^t M q &= \|p_B\| p_B^t M_B q_B \\ &= \|p_B\| p_B^t M_B q_B + \underbrace{p_C^t M_C p_C}_{=0} \\ &+ \underbrace{\|p_B\|(-\nu)^t p_C + \|p_B\| p_C^t \nu}_{=0} \\ &= (p_C, p_B)^t M_A (p_C, \|p_B\| q_B) \\ &= p_A^t M_A (p_C, \|p_B\| q_B) \geqslant 0.\end{aligned}$$

Hence, $p^B \in \mathit{ML}(P, B)$. □

10.8.2 Binary Choice

The basis of our characterization of *ML* is the special case for agendas of size 2. The following lemma states that, on two-element agendas, whenever a composition consistent proper SCF returns a non-pure outcome, it has to return all feasible outcomes. Interestingly, the proof uses composition consistency on three-element agendas, even though the statement itself only concerns agendas of size 2. In order to simplify notation, let $A = \{a, b\}$ and

$$p^\lambda = \lambda a + (1 - \lambda) b.$$

Lemma 10.7
: Let f be a proper SCF that composition consistency. Then, for all $P \in \mathcal{P}^\Delta$ and $\lambda \in (0, 1)$, $p^\lambda \in f(P, A)$ implies $f(P, A) = \Delta(A)$.

Proof. Let $P \in \mathcal{P}^\Delta$ and assume that $p^\lambda \in f(P, A)$ for some $\lambda \in (0, 1)$. Choose $c \in U \setminus A$ and $\hat{P} \in \mathcal{P}^\Delta$ as depicted below.

$P(\underline{a}, \underline{b})$	$P(\underline{b}, \underline{a})$
a	c
b	b
c	a

$\hat{P}|_{\{\underline{a},\underline{b},\underline{c}\}}$

Notice that $\hat{P}|_A = P|_A$ and thus, by IIA, $p^\lambda \in f(\hat{P}, A)$. Neutrality implies that $\lambda\underline{a} + (1-\lambda)\underline{c} \in f(\hat{P}, \{a, c\})$. Since A is a component in $\hat{P}$, we have $\lambda p^\lambda + (1-\lambda)\underline{c} \in f(\hat{P}, \{a, c\}) \times_a f(\hat{P}, A) = f(\hat{P}, \{a, b, c\})$. Since $\{b, c\}$ is also a component in $\hat{P}$, composition consistency implies that $\lambda p^\lambda + (1-\lambda)\underline{c} \in f(\hat{P}, \{a, b, c\}) = f(\hat{P}, A) \times_b f(\hat{P}, \{b, c\})$. Observe that $\lambda p^\lambda_a = p^{\lambda^2}_a$ and hence, $p^{\lambda^2} \in f(\hat{P}, A) = f(P, A)$.

Applying this argument repeatedly yields $p^{\lambda^{2^k}} \in f(P, A)$ for all $k \in \mathbb{N}$. Since $\lambda^{2^k} \to 0$ for $k \to \infty$ and f is continuous, we get $p^0 = \underline{b} \in f(P, A)$. Similarly, it follows that $p^1 = \underline{a} \in f(P, A)$. The fact that f is convex-valued implies that $f(P, A) = \Delta(A)$. □

The characterization of *ML* for agendas of size 2 proceeds along the following lines. By unanimity, neutrality, and Lemma 10.7, we know which outcomes have to be returned by every composition consistent SCF for three particular profiles. Then population consistency implies that every such SCF has to return all maximal lotteries. Last, we again use population consistency to show that the function is not decisive if it additionally returns outcomes that are not chosen by *ML*.

Lemma 10.8

> Let f be a proper SCF that satisfies population consistency and composition consistency. Then $f(P, A) = ML(P, A)$ for every $P \in \mathcal{P}^\Delta$.

Proof. First, note that for all $P \in \mathcal{P}^\Delta$, $f(P, A)$ only depends on $P(\underline{a}, \underline{b})$ by IIA. Let $P \in \mathcal{P}^\Delta$ be a profile such that $P(\underline{a}, \underline{b}) = 1/2$. Since $f(P, A) \neq \emptyset$, there is $\lambda \in [0, 1]$ such that $p^\lambda \in f(P, A)$. Neutrality implies that $p^{1-\lambda} \in f(P, A)$ and hence, by convexity of $f(P, A)$, $p^{1/2} = 1/2\,(p^\lambda + p^{1-\lambda}) \in f(P, A)$. It follows from Lemma 10.7 that $f(P, A) = \Delta(A)$.

Now, let $\hat{P} \in \mathcal{P}^\Delta$ be a profile such that $\hat{P}(\underline{a}, \underline{b}) = 1$. Unanimity implies that $f(\hat{P}, A) \neq \{\underline{b}\}$. Hence, by Lemma 10.7, $\underline{a} \in f(\hat{P}, A)$. By population consistency and the first part of the proof, we get $\underline{a} \in f(P)$ for all $P \in \mathcal{P}^\Delta$ with $P(\underline{a}, \underline{b}) \in [1/2, 1]$. Similarly, $\underline{b} \in f(P, A)$ for all $P \in \mathcal{P}^\Delta$ with $P(\underline{a}, \underline{b}) \in [0, 1/2]$. This already shows that $ML(P, A) \subseteq f(P, A)$ for every $P \in \mathcal{P}^\Delta$.

Finally, let $\hat{P} \in \mathcal{P}^\Delta$ be a profile such that $\hat{P}(\underline{a}, \underline{b}) = \mu > 1/2$. If $f(\hat{P}, A) \neq \{\underline{a}\}$, there is $\lambda \in [0, 1)$ such that $p^\lambda \in f(\hat{P}, A)$. Recall that $f(P, A) = \Delta(A)$ for all $P \in \mathcal{P}^\Delta$ such that $P(\underline{a}, \underline{b}) = 1/2$. Hence, it follows from population consistency that $p^\lambda \in f(P, A)$ for all $P \in \mathcal{P}^\Delta$ with $P(\underline{a}, \underline{b}) \in [1/2, \mu]$. But then $\{P \in \mathcal{P}^\Delta \colon P(\underline{a}, \underline{b}) \in [1/2, \mu]\} \subseteq \{P \in \mathcal{P}^\Delta \colon |f(P, A)| > 1\}$ and hence, $\{P \in \mathcal{P}^\Delta \colon |f(P, A)| = 1\}$ is not dense in $\mathcal{P}^\Delta$. This contradicts decisiveness of f. Thus, $f(P, A) = \{\underline{a}\}$ whenever $P(\underline{a}, \underline{b}) > 1/2$. An analogous argument shows that $f(P, A) = \{\underline{b}\}$ whenever $P(\underline{a}, \underline{b}) < 1/2$.

In summary, we have that $f(P, A) = \{\underline{a}\}$ if $P(\underline{a}, \underline{b}) \in (1/2, 1]$, $f(P, A) = \{\underline{b}\}$ if $P(\underline{a}, \underline{b}) \in [0, 1/2)$, and $f(P, A) = \Delta(A)$ if $P(\underline{a}, \underline{b}) = 1/2$. Thus, $f = ML$ (as depicted in Figure 10.1(a)). □

10.8.3 $f \subseteq ML$

The first lemma in this section shows that every proper SCF that satisfies population consistency and composition consistency returns weak Condorcet winner whenever they exist for profiles whose restriction to an agenda A is close to the uniform distribution on $\mathcal{D}|_{\underline{A}}$, i.e., profiles which in which every relation in $\mathcal{D}|_{\underline{A}}$ is assigned approximately the same fraction of agents. We prove this statement by induction on the size of the agenda. Every such profile that admits a weak Condorcet winner in A can be written as a convex combination of profiles that have a component and admit the same weak Condorcet winner. For the latter profiles we know from the induction hypothesis that the weak Condorcet winner has to be chosen.

Lemma 10.9

Let f be a proper SCF that satisfies population consistency and composition consistency. Then, for all $A \in \mathcal{F}(U)$, there is $\epsilon > 0$ such that $\underline{a} \in f(P, A)$ for every profile $P \in \mathcal{P}^{\Delta}$ such that $P|_{\underline{A}} \in B_{\epsilon}(\text{uni}(\mathcal{D}|_{\underline{A}}))$ and $\underline{a}$ is a weak Condorcet winner in $P|_{\underline{A}}$.

Proof. Let $A \in \mathcal{F}(U)$, $|A| = m$, and $P \in \mathcal{P}^{\Delta}$ be such that $\underline{a} \in \underline{A}$ is a weak Condorcet winner in $P|_{\underline{A}}$ and $\|P|_{\underline{A}} - \text{uni}(\mathcal{D}|_{\underline{A}})\| \leqslant \epsilon_m = (4^m \prod_{k=1}^{m} k!)^{-1}$. We show that $\underline{a} \in f(P, A)$ by induction over m. An example for $m = 3$ illustrating the idea is given after the proof. For $m = 2$, the claim follows directly from Lemma 10.8.

For $m > 2$, fix $b \in A \setminus \{a\}$. First, we introduce some notation. For $\succ \in \mathcal{D}|_{\underline{A}}$, we denote by $\succ^{b \to a} \in \mathcal{D}|_{\underline{A}}$ the relation that is identical to $\succ$ except that $\underline{b}$ is moved upwards or downwards (depending on whether $\underline{a} \succ \underline{b}$ or $\underline{b} \succ \underline{a}$) until no more pure outcome is "in between" $\underline{a}$ and $\underline{b}$ (without switching the order of $\underline{a}$ and $\underline{b}$. Formally, $\succ^{b \to a}|_{\underline{A \setminus \{b\}}} = \succ_{\underline{A \setminus \{b\}}}$, $\underline{a} \succ^{b \to a} \underline{b}$ if and only if $\underline{a} \succ \underline{b}$, and $\{a, b\}$ is a component in $\succ^{b \to a}$, i.e., there is no $x \in A$ such that $\underline{a} \succ \underline{x} \succ \underline{b}$ or $\underline{b} \succ \underline{x} \succ \underline{a}$. Notice that for every $\hat{\succ} \in \mathcal{D}|_{\underline{A}}$, there are at most $m - 1$ distinct relations $\succ \in \mathcal{D}|_{\underline{A}}$ such that $\hat{\succ} = \succ^{b \to a}$.

We first show that, by composition consistency, weak Condorcet winners have to be chosen whenever they exist for a particular type of profiles. For $\succ \in \mathcal{D}|_{\underline{A}}$, let $S \in \mathcal{P}^{\Delta}$ such that A is a component in S and $S(\succ) + S(\succ)^{b \to a} = S(\succ^{-1}) = 1/2$. We have that $S(\underline{a}, \underline{x}) = 1/2$ for all $x \in A \setminus \{a\}$ and hence, $\underline{a}$ is a weak Condorcet winner in $S|_{\underline{A}}$. We prove that $\underline{a} \in f(S, A)$ by induction over m. For $m = 2$, this follows from Lemma 10.8. For $m > 2$, let $x \in A \setminus \{b\}$ such that $\underline{x} \succ \underline{y}$ for all $y \in A$ or $\underline{y} \succ \underline{x}$ for all $y \in A$. Such an x exists, since $m > 2$. Notice that $A \setminus \{x\}$ is a component in $S|_{\underline{A}}$ and $S(\underline{x}, \underline{y}) = 1/2$ for all $y \in A \setminus \{x\}$. If $x = a$, it follows from composition consistency and Lemma 10.8 that $\underline{a} \in f(S, A)$. If $x \neq a$, it follows from the induction hypothesis that $\underline{a} \in f(S, A \setminus \{x\})$. Lemma 10.8 implies that $\underline{a} \in f(S, \{a, x\})$ as $S(\underline{a}, \underline{x}) = 1/2$. Then, it follows from composition consistency that $\underline{a} \in f(S, \{a, x\}) \times_a f(S, A \setminus \{x\}) = f(S, A)$.

Now, for every $\succ \in \mathcal{D}|_{\underline{A}}$ such that $\{a,b\}$ is *not* a component in $\succ$ and $0 < P(\succ) \leqslant P(\succ^{-1})$, let $S^{\succ} \in \mathcal{P}^{\Delta}$ such that

$$S^{\succ}(\succ) + S^{\succ}(\succ^{b \to a}) = S^{\succ}(\succ^{-1}) = 1/2, \text{ and}$$
$$S^{\succ}(\succ)/S^{\succ}(\succ^{-1}) = P(\succ)/P(\succ^{-1}).$$

From what we have shown before, it follows that $\underline{a} \in f(S^{\succ}, A)$ for all $\succ \in \mathcal{D}|_{\underline{A}}$.

The rest of the proof proceeds as follows. We show that P can be written as a convex combination of profiles of the type $S^{\succ}$ and a profile $\hat{P}$ in which $\{a,b\}$ is a component and $\underline{a}$ is a weak Condorcet winner in $P|_{\underline{A}}$. Since $P|_{\underline{A}}$ is close to the uniform distribution on $\mathcal{D}|_{\underline{A}}$, $P(\succ)$ is almost identical for all relations $\succ \in \mathcal{D}|_{\underline{A}}$. Hence, $S^{\succ}(\succ)$ is close to 0 for all relations $\succ$ in which $\{a,b\}$ is a component. As a consequence, $\hat{P}(\succ)$ is almost identical for all preference relations $\succ$ in which $\{a,b\}$ is a component and $\hat{P}|_{\underline{A}\setminus\{\underline{b}\}}$ is close to the uniform distribution on $\mathcal{D}|_{\underline{A}\setminus\{\underline{b}\}}$. By the induction hypothesis, $\underline{a} \in f(\hat{P}, A \setminus \{b\})$. Since $\{a,b\}$ is a component in $\hat{P}$ and $\hat{P}(\underline{a},\underline{b}) \geqslant 1/2$, it follows from composition consistency that $\underline{a} \in f(\hat{P}, A)$.

We define $S \in \mathbb{Q}^{\mathcal{D}}_{\geqslant 0}$ such that

$$S = 2\sum_{\succ} P(\succ^{-1}) S^{\succ},$$

where the sum is taken over all $\succ \in \mathcal{D}|_{\underline{A}}$ such that $\{a,b\}$ is *not* a component in $\succ$ and $0 < P(\succ) \leqslant P(\succ^{-1})$ (in case $P(\succ) = P(\succ^{-1})$ we pick one of $\succ$ and $\succ^{-1}$ arbitrarily). Now, let $\hat{P} \in \mathcal{P}^{\Delta}$ such that

$$P = (1 - \|S\|)\hat{P} + S.$$

Note that, by definition of S, $\hat{P}(\succ) = 0$ for all $\succ \in \mathcal{D}|_{\underline{A}}$ such that $\{a,b\}$ is not a component in $\succ$. Hence, $\{a,b\}$ is a component in $\hat{P}$. By the choice of P, we have that

$$\|S\| = \sum_{\succ \in \mathcal{D}|_{\underline{A}}} S(\succ) \leqslant \frac{m! - 2(m-1)!}{m!} + \epsilon_m = 1 - \frac{2}{m} + \epsilon_m.$$

Using this fact, a simple calculation shows that

$$\hat{P}(\succ) \leqslant \frac{P(\succ) - S(\succ)}{\frac{2}{m} - \epsilon_m} \leqslant \frac{\frac{1}{m!} + \epsilon_m}{\frac{2}{m} - \epsilon_m} \leqslant \frac{1}{2(m-1)!} + \frac{\epsilon_{m-1}}{4(m-1)!},$$

for every $\succ \in \mathcal{D}|_{\underline{A}}$ in which $\{a,b\}$ is a component. Since for every such relation, there is exactly one other relation in $\mathcal{D}|_{\underline{A}}$ that is identical to $\succ$ except that $\underline{a}$ and $\underline{b}$ are swapped, we have that

$$\hat{P}(\succ) \leqslant \frac{1}{(m-1)!} + \frac{\epsilon_{m-1}}{2(m-1)!},$$

for every $\succ \in \mathcal{D}|_{\underline{A}\setminus\{\underline{b}\}}$. By the above calculation, we have that

$$\left\| \hat{P}|_{\underline{A}\setminus\{\underline{b}\}} - \mathrm{uni}(\mathcal{D}|_{\underline{A}\setminus\{\underline{b}\}}) \right\| \leqslant \epsilon_{m-1}.$$

Since $S^{\succ}(\underline{a},\underline{x}) = 1/2$ for all $x \in A \setminus \{a\}$ and $\succ \in \mathcal{D}|_{\underline{A}}$, we have that $\hat{P}(\underline{a},\underline{x}) \geqslant 1/2$ for all $x \in A \setminus \{a\}$. Thus, $\underline{a}$ is a weak Condorcet winner in $\hat{P}|_{\underline{A} \setminus \{\underline{b}\}}$. From the induction hypothesis it follows that $\underline{a} \in f(\hat{P}, A \setminus \{b\})$. Using the fact that $\hat{P}(\underline{a},\underline{b}) \geqslant 1/2$, Lemma 10.8 implies that $\underline{a} \in f(\hat{P}, \{a,b\})$. Finally, composition consistency entails $\underline{a} \in f(\hat{P}, A \setminus \{b\}) \times_a f(\hat{P}, \{a,b\}) = f(\hat{P}, A)$.

In summary, $\underline{a} \in f(S^{\succ}, A)$ for all $\succ \in \mathcal{D}|_{\underline{A}}$ and $\underline{a} \in f(\hat{P}, A)$. Since P is a convex combination of profiles of the type $S^{\succ}$ and $\hat{P}$, it follows from population consistency that $\underline{a} \in f(P, A)$. $\square$

We now give an example for $A = \{a, b, c\}$ that illustrates the proof of Lemma 10.9. Let $0 \leqslant \epsilon \leqslant \epsilon_3$ and consider a preference profile $P \in \mathcal{P}^{\Delta}$ of the following form.

$(1+2\epsilon)/6$	$1/6$	$1/6$	$(1-\epsilon)/6$	$(1-\epsilon)/6$	$1/6$
a	a	b	b	c	c
b	c	a	c	a	b
c	b	c	a	b	a

$P|_{\underline{A}}$

Then, we have that $\|P|_{\underline{A}} - \text{uni}(\mathcal{D}|_{\underline{A}})\| \leqslant \epsilon_3$. Now consider $\succ \in \mathcal{D}|_{\underline{A}}$ with $\underline{b} \succ \underline{c} \succ \underline{a}$, which yields $S^{\succ} \in \mathcal{P}^{\Delta}$ as depicted below.

$1/2$	$(1-\epsilon)/2$	$\epsilon/2$
a	b	c
c	c	b
b	a	a

$S^{\succ}|_{\underline{A}}$

Here, $\underline{y} \succ \underline{a}$ for all $y \in A$. Hence, it follows from what we have shown before that $\underline{a} \in f(S^{\succ}, A)$. No other profiles of this type need to be considered, as $\succ$ and $\succ^{-1}$ are the only relations in $\mathcal{D}|_{\underline{A}}$ in which $\{a,b\}$ is not a component. Thus $S = 1/3\, S^{\succ}$.

Then, we get $\hat{P}$ as follows.

$(1+2\epsilon)/4$	$1/4$	$(1-\epsilon)/4$	$(1-\epsilon)/4$
a	b	c	c
b	a	a	b
c	c	b	a

$\hat{P}|_{\underline{A}}$

$(1+\epsilon)/2$	$(1-\epsilon)/2$
a	c
c	a

$\hat{P}|_{\{\underline{a},\underline{c}\}}$

$(2+\epsilon)/4$	$(2-\epsilon)/4$
a	b
b	a

$\hat{P}|_{\{\underline{a},\underline{b}\}}$

It follows from Lemma 10.8 that $\underline{a} \in f(\hat{P}, \{a, c\})$ and $\underline{a} \in f(\hat{P}, \{a, b\})$. Then, composition consistency implies that

$$\underline{a} \in f(\hat{P}, A) = f(\hat{P}, \{a, c\}) \times_a f(\hat{P}, \{a, b\}).$$

In summary, we have that

$$P = 2/3\, \hat{P} + 1/3\, S^{\succ},$$

$\underline{a} \in f(\hat{P}, A)$, and $\underline{a} \in f(S^{\succ}, A)$. Thus, population consistency implies that $\underline{a} \in f(P, A)$.

Lemma 10.10

Let f be a proper SCF that satisfies population consistency and composition consistency. Then, for all $A \in \mathcal{F}(U)$, there is $\epsilon > 0$ such that f returns the uniform distribution over all weak Condorcet winners for all profiles $P \in \mathcal{P}^{\Delta}$ such that $P|_{\underline{A}} \in B_{\epsilon}(\mathcal{D}|_{\underline{A}})$.

Proof. Let $A \in \mathcal{F}(U)$, $|A| = m$, and $P \in \mathcal{P}^{\Delta}$ with $\|P|_{\underline{A}} - \text{uni}(\mathcal{D}|_{\underline{A}})\| \leqslant \epsilon_m$. Let $\hat{A} \subseteq A$ be the set of weak Condorcet winners in $P|_{\underline{A}}$. We actually prove a stronger statement, namely that $\Delta(\hat{A}) \subseteq f(P, A)$. Every pure outcome $\underline{x} \in \underline{\hat{A}}$ is a weak Condorcet winner in $P|_{\underline{A}}$. Thus, it follows from Lemma 10.9 that $\underline{\hat{A}} \subseteq f(P)$. Since $f(P, A)$ is convex, $\Delta(\hat{A}) \subseteq f(P, A)$ follows. □

For the remainder of the proof, we need to define two classes of profiles that are based on regularity conditions imposed on the corresponding majority margins. Let $A \in \mathcal{F}(U)$. A profile $P \in \mathcal{P}^{\Delta}$ is

$$\textit{regular on } A \text{ if } \sum_{y \in A} M^P_{xy} = 0 \text{ for all } x \in A, \text{ and}$$
$$\textit{strongly regular on } A \text{ if } M^P_A = 0.$$

By $\mathcal{P}^A$ and $\mathcal{S}^A$ we denote the set of all profiles in $\mathcal{P}^{\Delta}$ that are regular or strongly regular on A, respectively.

For the following five lemmas, fix $A \in \mathcal{F}(U)$ and $|A| = m$. We show that, for every $\hat{A} \subseteq A$, every profile can be affinely decomposed into profiles of three different types: profiles that are strongly regular on $\hat{A}$, certain profiles that are regular on A, and profiles that admit a strict Condorcet winner in $\hat{A}$.[37] Lemmas 10.11, 10.12, and 10.13 do not make any reference to population consistency, composition consistency, or maximal lotteries and may be of independent interest. First, we determine the dimension of the space of strongly regular profiles on $\hat{A}$ restricted to A.

Lemma 10.11

Let $\hat{A} \subseteq A \in \mathcal{F}(U)$. Then, $\dim(\mathcal{S}^{\hat{A}}|_{\underline{A}}) = m! - \binom{|\hat{A}|}{2} - 1$.

37 Similar decompositions of majority margin matrices have been explored by Zwicker (1991) and Saari (1995).

Proof. We will characterize $\mathcal{S}^{\hat{A}}|_{\underline{A}}$ using a set of linear constraints. By definition, $\mathcal{S}^{\hat{A}}|_{\underline{A}} = \{S|_{\underline{A}} \colon S \in \mathcal{P}^{\Delta}$ such that $M^S_{\hat{A}} = 0\}$. Recall that $M^S_{xy} = \sum_{\succ \in \mathcal{D} \colon \underline{x} \succ \underline{y}} S(\succ) - \sum_{\succ \in \mathcal{D} \colon \underline{y} \succ \underline{x}} S(\succ)$ for all $x, y \in A$. Since $M^P_{xx} = 0$ for all $P \in \mathcal{P}^{\Delta}$ and $x \in A$, $\mathcal{S}^{\hat{A}}|_{\underline{A}}$ can be characterized by $\binom{|\hat{A}|}{2}$ homogeneous linear constraints in the $(m! - 1)$-dimensional space $\mathcal{D}|_{\underline{A}}$, which implies that $\dim(\mathcal{S}^{\hat{A}}|_{\underline{A}}) \geqslant m! - \binom{|\hat{A}|}{2} - 1$. Equality holds but is not required for the following arguments. We therefore omit the proof. □

Second, we determine the dimension of the set of skew-symmetric $m \times m$ matrices that correspond to profiles that are regular on A and vanish outside their upper left $\hat{m} \times \hat{m}$ submatrix, i.e.,

$$\mathcal{M}^{\hat{m}} = \Big\{M \in \mathcal{M}_{[m]} \cap \mathbb{Q}^{m \times m} \colon \sum_{j=1}^{m} M_{ij} = 0 \text{ if } i \in [m] \text{ and}$$
$$M_{ij} = 0 \text{ if } \{i, j\} \not\subseteq [\hat{m}]\Big\}.$$

In Lemma 10.13, we then proceed to show that every matrix of this type can be decomposed into matrices induced by a subset of profiles that are regular on A and for which we know that every SCF has to return the uniform distribution over the first $\hat{m}$ alternatives (possibly among other outcomes) for the agenda A.

LEMMA 10.12

$\dim(\mathcal{M}^{\hat{m}}) = \binom{\hat{m}}{2} - (\hat{m} - 1)$.

Proof. First note that the space of all $m \times m$ matrices has dimension m^2. We show that $\mathcal{M}^{\hat{m}}$ can be characterized by a set of $(m^2 - \hat{m}^2) + (\binom{\hat{m}}{2} + \hat{m}) + (\hat{m} - 1)$ homogeneous linear constraints. Let $M \in \mathbb{Q}^{m \times m}$ and observe that $(m^2 - \hat{m}^2)$ constraints are needed to ensure that M vanishes outside of $[\hat{m}] \times [\hat{m}]$, $\binom{\hat{m}}{2} + \hat{m}$ constraints are needed to ensure skew-symmetry of $M_{[\hat{m}]}$, and $(\hat{m} - 1)$ constraints are needed to ensure that the first $\hat{m}$ rows (and hence also the columns) of M sum up to 0, i.e., $\sum_{j=1}^{m} M_{ij} = 0$ for all $i \in [\hat{m} - 1]$. It follows from skew-symmetry and the latter $\hat{m} - 1$ constraints that the $\hat{m}$th row of M sums up to 0, since

$$\sum_{j=1}^{m} M_{\hat{m}j} = \sum_{i,j=1}^{m} M_{ij} - \sum_{i=1}^{\hat{m}-1} \sum_{j=1}^{m} M_{ij} = 0.$$

The last $m - \hat{m}$ rows of M trivially sum up to 0. Hence, $\dim(\mathcal{M}^{\hat{m}}) \geqslant \hat{m}^2 - (\binom{\hat{m}}{2} + \hat{m}) - (\hat{m} - 1) = \binom{\hat{m}}{2} - (\hat{m} - 1)$. Equality holds but is not required for the following arguments. We therefore omit the proof. □

Let $\Pi^{\circ}_{[m]}(B)$ be the set of all permutations on $[m]$ that are cyclic on $B \subseteq [m]$ and coincide with the identity permutation outside of B.[38]

38 A permutation π of $[m]$ is cyclic on $B \subseteq [m]$ if $\pi^{|B|}$ is the smallest positive power of π that is the identity function on B.

We denote by $\mathcal{M}^{\hat{m},\circ}$ the space of all matrices in $\mathcal{M}^{\hat{m}}$ induced by a permutation in $\Pi^{\circ}_{[m]}(B)$ for some $B \subseteq [\hat{m}]$, i.e.,

$$\mathcal{M}^{\hat{m},\circ} = \Big\{ M \in \mathcal{M}^{\hat{m}} \colon \text{there are } \pi \in \Pi^{\circ}_{[m]}(B) \text{ and } B \subseteq [\hat{m}] \text{ with } M_{ij} = \begin{cases} 1 & \text{if } j = \pi(i),\, i \in B, \\ -1 & \text{if } i = \pi(j),\, j \in B, \\ 0 & \text{otherwise,} \end{cases} \Big\},$$

with the convention that $\mathcal{M}^{2,\circ} = \{0\}$. We now show that the linear hull of $\mathcal{M}^{\hat{m},\circ}$ is $\mathcal{M}^{\hat{m}}$.

Lemma 10.13
$\operatorname{lin}(\mathcal{M}^{\hat{m},\circ}) = \mathcal{M}^{\hat{m}}$.

Proof. The idea underlying the proof is as follows: every matrix $M \in \mathcal{M}^{\hat{m}}$ corresponds to a weighted directed graph with vertex set $[m]$ where the weight of the edge from i to j is M_{ij}. If $M \neq 0$, there exists a cycle along edges with positive weight of length at least 3 in the subgraph induced by $[\hat{m}]$. We obtain a matrix $\hat{M}$ with smaller norm than M by subtracting the matrix in $\mathcal{M}^{\hat{m},\circ}$ from M that corresponds to the cycle identified before.

Let $M \in \mathcal{M}^{\hat{m}}$ and $k \in \mathbb{N} \setminus \{0\}$ such that $kM \in \mathbb{N}^{m \times m}$. We show, by induction over $k\|M\|$, that $M = \sum_{i=1}^{\ell} \lambda_i M^i$ for some $\lambda \in \mathbb{Q}^{\ell}$ and $M^i \in \mathcal{M}^{\hat{m},\circ}$ for all $i \in [\ell]$ for some $\ell \in \mathbb{N}$. If $k\|M\| = 0$ then $M = 0$. Hence, the induction hypothesis is trivial.

If $k\|M\| \neq 0$, i.e., $M \neq 0$, we can find $B \subseteq [\hat{m}]$ with $|B| \geqslant 3$ and $\pi \in \Pi^{\circ}_{[m]}(B)$ such that $M_{ij} > 0$ if $\pi(i) = j$ and $i \in B$. Note that π defines a cycle of length at least 3 in the graph that corresponds to M. We define $M^1 \in \mathcal{M}^{\hat{m},\circ}$ by letting

$$M^1_{ij} = \begin{cases} 1 & \text{if } \pi(i) = j \text{ and } i \in B, \\ -1 & \text{if } \pi(j) = i \text{ and } j \in B, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Let $\lambda_1 = \min\{M_{ij} \colon i, j \in [m] \text{ and } M^1_{ij} > 0\}$ and $\hat{M} = M - \lambda_1 M^1$. By construction, we have that $\hat{M}_{ij} = M_{ij} - \lambda_1$ if $\pi(i) = j$ and $i \in B$, $\hat{M}_{ij} = M_{ij} + \lambda_1$ if $\pi(j) = i$ and $j \in B$, and $\hat{M}_{ij} = M_{ij}$ otherwise. Note that $M_{ij} \geqslant \lambda_1$ if $\pi(i) = j$ and $i \in B$ and $M_{ij} \leqslant -\lambda_1$ if $\pi(j) = i$ and $j \in B$ by definition of λ_1. Recall that $kM \in \mathbb{N}^{m \times m}$ and, in particular, $k\lambda_1 \in \mathbb{N}$. Hence, $k\hat{M} \in \mathbb{N}^{m \times m}$. Moreover, $k\|\hat{M}\| = k\|M\| - 2k\lambda_1|B| \leqslant k\|M\| - 1$. From the induction hypothesis we know that $\hat{M} = \sum_{i=2}^{\ell} \lambda_i M^i$ with $\lambda_i \in \mathbb{Q}$ and $M^i \in \mathcal{M}^{\hat{m},\circ}$ for all $i \in [\ell] \setminus \{1\}$ for some $\ell \in \mathbb{N}$. By construction of $\hat{M}$, we have that $M = \sum_{i=1}^{\ell} \lambda_i M^i$. □

Lemma 10.14 leverages Lemmas 10.10, 10.11, 10.12, and 10.13 to show two statements. First, it determines the dimension of the space of profiles that are regular on $\hat{A} \subseteq A$ restricted to A. Second, it proves that there is a full-dimensional subset of this space for which every proper SCF that satisfies population consistency and composition consistency returns the uniform distribution over $\hat{A}$.

Lemma 10.14

> Let f be a proper SCF that satisfies population consistency and composition consistency. Then, for all $\hat{A} \subseteq A$, there is $\mathcal{X} \subseteq \mathcal{P}^{\Delta}$ such that $\mathcal{X}|_{\underline{A}}$ has dimension $|A|! - |\hat{A}|$ and $\text{uni}(\hat{A}) \in f(P, A)$ for every $P \in \mathcal{X}$.

Proof. To simplify notation, we assume without loss of generality that $A = [m]$ and $\hat{A} = [\hat{m}]$. By Lemma 10.11 we can find a set $\mathcal{S} = \{S^1, \dots, S^{m! - \binom{\hat{m}}{2}}\} \subseteq \mathcal{S}^{[\hat{m}]}$ of profiles such that $\mathcal{S}|_{\underline{A}}$ has dimension $m! - \binom{\hat{m}}{2} - 1$. Since $\mathcal{S}$ can be chosen such that every $S|_{\underline{A}}$ is close to $\text{uni}(\mathcal{D}|_{\underline{[m]}})$ for every $S \in \mathcal{S}$, it follows from Lemma 10.10 that $\text{uni}([\hat{m}]) \in f(S, A)$ for all $S \in \mathcal{S}$. Therefore, it suffices to find a set of profiles $\mathcal{T} = \{P^1, \dots, P^{\binom{\hat{m}}{2} - (\hat{m}-1)}\} \subseteq \mathcal{P}^{[\hat{m}]}$ such that $\text{uni}([\hat{m}]) \in f(P, A)$ for every $P \in \mathcal{T}$ and $\mathcal{S}|_{\underline{A}} \cup \mathcal{T}|_{\underline{A}}$ is a set of affinely independent vectors. If $\hat{m} = 2$, we can choose $\mathcal{T} = \emptyset$. For $\hat{m} \geqslant 3$ we construct a suitable set of profiles as follows.

For every $B \subseteq [\hat{m}]$ with $|B| = k \geqslant 3$ and $\pi \in \Pi^{\circ}_{[m]}(B)$, let $[m] \setminus B = \{a_1, \dots, a_{m-k}\}$ and $P^{B,\pi} \in \mathcal{P}^{\Delta}$ be defined as follows: for $\succ \in \mathcal{D}|_{\underline{A}}$, $P^{B,\pi}(\succ) = 1/(2k)$ if

$$\underline{\pi^0(i)} \succ \underline{\pi^1(i)} \succ \underline{\pi^2(i)} \succ \dots \succ \underline{\pi^{k-1}(i)} \succ \underline{a_1} \succ \dots \succ \underline{a_{m-k}} \quad \text{or}$$
$$\underline{a_{m-k}} \succ \dots \succ \underline{a_1} \succ \underline{\pi^{k-1}(i)} \succ \dots \succ \underline{\pi^2(i)} \succ \underline{\pi^0(i)} \succ \underline{\pi^1(i)},$$

for some $i \in B$. Note that $P^{B,\pi}$ is regular on $[\hat{m}]$, since

$$P^{B,\pi}(\underline{i,j}) = \begin{cases} \lambda & \text{if } \pi(i) = j \text{ and } i \in B, \\ -\lambda & \text{if } \pi(j) = i \text{ and } j \in B, \text{ and} \\ 0 & \text{otherwise,} \end{cases}$$

where $\lambda = 1/k > 0$. Hence, for every $M \in \mathcal{M}^{\hat{m},\circ}$, there are $B \subseteq [\hat{m}]$ and $\pi \in \Pi^{\circ}_{[m]}(B)$ such that $\lambda M = M_A^{P^{B,\pi}}$. Notice that B and $[m] \setminus B$ are components in $P^{B,\pi}$. For $j \in B$, we have by construction that $P^{B,\pi}(\underline{j}, \underline{a_1}) = 0$. Hence, it follows from Lemma 10.8 that $\{\underline{j}, \underline{a_1}\} \subseteq f(P^{B,\pi}, \{j, a_1\})$. Moreover, neutrality, convexity, and composition consistency imply that $\text{uni}(B) \in f(P^{B,\pi}, A)$ by the symmetry of $P^{B,\pi}$ with respect to B. Now let $a_i \in \{a_1, \dots, a_{m-k}\}$. Observe that $\{a_1, \dots, a_{i-1}\}$ is a component in $P^{B,\pi}$ and $P^{B,\pi}(\underline{a_1}, \underline{a_i}) = 0$. Thus, composition consistency and Lemma 10.8 imply that

$$\underline{a_i} \in f(P^{B,\pi}, \{a_1, a_i\}) \times_{a_1} f(P^{B,\pi}, \{a_1, \dots, a_{i-1}\}) = f(P^{B,\pi}, \{a_1, \dots, a_i\}).$$

Furthermore, $P^{B,\pi}(\underline{a_i}, \underline{a_{m-k}}) = 0$ and $\{a_{i+1}, \dots, a_{m-k}\}$ is a component in $P^{B,\pi}$. As before, we get

$$\begin{aligned}\underline{a_i} \in f(P^{B,\pi}, \{a_i, a_{m-k}\}) \times_{a_{m-k}} f(P^{B,\pi}, \{a_{i+1}, \dots, a_{m-k}\}) \\ = f(P^{B,\pi}, \{a_i, \dots, a_{m-k}\}).\end{aligned}$$

Also $\{a_i, \dots, a_{m-k}\}$ is a component in $P^{B,\pi}$ and thus,

$$\begin{aligned}\underline{a_i} \in f(P^{B,\pi}, \{a_1, \dots, a_i\}) \times_{a_i} f(P^{B,\pi}, \{a_i, \dots, a_{m-k}\}) \\ = f(P^{B,\pi}, [m] \setminus B).\end{aligned}$$

Since $P^{B,\pi}(\underline{j}, \underline{a_1}) = 0$, we get

$$\underline{a_i} \in f(P^{B,\pi}, \{j, a_1, \dots, a_{m-k}\}) \times_j f(P^{B,\pi}, B) = f(P^{B,\pi}, A).$$

Then, it follows from convexity of $f(P^{B,\pi}, A)$ that

$$\text{uni}([\hat{m}]) = \frac{k}{\hat{m}}\text{uni}(B) + \frac{1}{\hat{m}} \sum_{a_i \in [\hat{m}] \setminus B} \underline{a_i} \in f(P^{B,\pi}, A),$$

since $\text{uni}(B) \in f(P^{B,\pi}, A)$ and $\underline{a_i} \in f(P^{B,\pi}, A)$ for every $i \in [m-k]$.

We know from Lemma 10.12 that $\dim(\mathcal{M}^{\hat{m},\circ}) \geqslant \binom{\hat{m}}{2} - (\hat{m}-1)$ and, from Lemma 10.13, that $\text{lin}(\mathcal{M}^{\hat{m},\circ}) = \mathcal{M}^{\hat{m}}$. Thus, we can find a basis $\{M^1, \dots, M^{\binom{\hat{m}}{2}-(\hat{m}-1)}\}$ of $\mathcal{M}^{\hat{m},\circ}$ and a set of corresponding profiles

$$\mathcal{T} = \{P^1, \dots, P^{\binom{\hat{m}}{2}-(\hat{m}-1)}\} \subseteq \{P^{B,\pi} \colon B \subseteq [\hat{m}] \text{ and } \pi \in \Pi^\circ_{[m]}(B)\}.$$

We claim that $\mathcal{S}|_{\underline{A}} \cup \mathcal{T}|_{\underline{A}}$ is a set of affinely independent vectors in $\mathcal{P}^\Delta|_{\underline{A}}$. Let $S^1, \dots, S^k \in \mathcal{S}$ and $P^1, \dots, P^\ell \in \mathcal{T}$ be pairwise disjoint. Assume that $\sum_i \lambda_i S^i|_{\underline{A}} + \sum_j \mu_j P^j|_{\underline{A}} = 0$ for some $\lambda \in \mathbb{Q}^k$ and $\mu \in \mathbb{Q}^\ell$ such that $\sum_i \lambda_i + \sum_j \mu_j = 0$. This implies that $\sum_j \mu_j M^j = 0$, which in turn implies $\mu = 0$, since the M^j are linearly independent. Hence, $\sum_i \lambda_i S^i|_{\underline{A}} = 0$ and $\sum_i \lambda_i = 0$, which implies that $\lambda = 0$, since $S^1|_{\underline{A}}, \dots, S^{m!-\binom{\hat{m}}{2}}|_{\underline{A}}$ are affinely independent. Thus, $\mathcal{S}|_{\underline{A}} \cup \mathcal{T}|_{\underline{A}}$ is a set of affinely independent vectors and $\dim(\mathcal{S}|_{\underline{A}} \cup \mathcal{T}|_{\underline{A}}) = |\mathcal{S} \cup \mathcal{T}| - 1 = m! - \hat{m}$. The above stated fact that $\text{uni}([\hat{m}]) \in f(P^{B,\pi}, A)$ for every $B \subseteq [\hat{m}]$ and $\pi \in \Pi^\circ_{[m]}(B)$ finishes the proof. □

We now consider proper SCFs that returns an outcome that is *not* maximal. The following lemma shows that for every such SCF, there is a set of profiles with a strict Condorcet winner on some agenda for which it returns the uniform distribution over a fixed subset of the agenda if we additionally require population consistency and composition consistency. Furthermore, this set of profiles has only one regular profile in its linear hull. Later this statement is leveraged to show that every population consistent and composition consistent proper SCF returns a subset of maximal lotteries.

LEMMA 10.15

Let f be a proper SCF that satisfies population consistency and composition consistency. If $f \not\subseteq ML$, there are $\mathcal{Y} \subseteq \mathcal{P}^\Delta$ and $\hat{A} \subseteq A \in \mathcal{F}(U)$ such that

(i) $\mathcal{Y}|_{\underline{A}}$ has dimension $|\hat{A}| - 1$,

(ii) $\mathrm{uni}(\hat{A}) \in f(P, A)$ for every $P \in \mathcal{Y}$, and

(iii) $\dim(\mathrm{lin}(\mathcal{Y}|_{\underline{A}}) \cap \mathrm{lin}(\mathcal{P}^{\hat{A}}|_{\underline{A}})) = 1$.

Proof. If $f \not\subseteq ML$, there are $P \in \mathcal{P}^\Delta$, $A \in \mathcal{F}(U)$, and $p \in f(P, A)$ such that $p \notin ML(P, A)$. Since $ML(P, A)$ is closed and $f(P, A)$ is convex with rational-valued extreme points, we can assume without loss of generality that $p \in \Delta^{\mathbb{Q}}(A)$. Since p is not a maximal lottery, by definition, there is $q \in \Delta(A)$ such that $q^t M^P p > 0$. Linearity of matrix multiplication implies that there is $x \in A$ such that $(M^P p)_x > 0$. We first use composition consistency to "blow up" alternatives such that the resulting outcome is the uniform distribution over a subset of alternatives. Let κ be the greatest common divisor of $\{p_y : y \in A\}$, i.e., $\kappa = \max\{s \in \mathbb{Q} : p_y/s \in \mathbb{N} \text{ for all } y \in A\}$. For every $y \in A$, let $A^y \in \mathcal{F}(U)$ such that $|A^y| = \max\{1, p_y/\kappa\}$, $A^y \cap A = \{y\}$, and all A^y are pairwise disjoint. The A^y exist, since U is assumed to be infinite. Moreover, let $A^u = \bigcup_{y \in A} A^y$. Now, choose $P^u \in \mathcal{P}^\Delta$ such that $P^u|_{\underline{A}} = P|_{\underline{A}}$, A^y is a component in P^u for every $y \in A$, and $P^u|_{\underline{A^y}} = \mathrm{uni}(\mathcal{D}|_{\underline{A^y}})$ for every $y \in A^y$. Hence, $\mathrm{uni}(A^y) \in f(P^u, A^y)$ for all $y \in A$ as f is neutral and $f(P^u, A^y)$ is convex. To simplify notation, let $A^p = \bigcup_{y \in \mathrm{supp}(p)} A^y$. By composition consistency, it follows that $\hat{p} = \mathrm{uni}(A^p) \in f(P^u, A^u)$. Observe that

$$(M^{P^u}\hat{p})_x = \sum_{y \in \mathrm{supp}(p) \setminus \{x\}} \frac{|A^y|}{|A^p|} M^{P^u}_{xy} = \sum_{y \in A \setminus \{x\}} p_y M^P_{xy} > 0.$$

We now construct a profile $\hat{P} \in \mathcal{P}^\Delta$ such that $\underline{x}$ is a strict Condorcet winner in $\hat{P}|_{\underline{A^u}}$ and $\mathrm{uni}(A^p) \in f(\hat{P}, A^u)$. To this end, let $\hat{P} \in \mathcal{P}^\Delta$ be the uniform mixture of all profiles that arise from P^u by permuting all alternatives in $A^p \setminus \{x\}$, i.e.,

$$\hat{P} = \frac{1}{|A^p \setminus \{x\}|!} \sum_{\substack{\pi \in \Pi(U) \colon \pi(y) = y \\ \text{for all } y \in U \setminus A^p \cup \{x\}}} (P^u)^\pi.$$

Then, $M^{\hat{P}}_{xy} = M^{\hat{P}}_{xz} > 0$ for all $y, z \in A^p \setminus \{x\}$. Neutrality and population consistency imply that $\hat{p} \in f(\hat{P}, A^u)$.

Let $P^{\mathrm{uni}} \in \mathcal{P}^\Delta$ such that $P^{\mathrm{uni}}|_{\underline{A^u}} = \mathrm{uni}(\mathcal{D}|_{\underline{A^u}})$ and, for $\lambda \in [0, 1]$, define

$$P^\lambda = \lambda \hat{P} + (1 - \lambda) P^{\mathrm{uni}}.$$

It follows from Lemma 10.9 that $\underline{y} \in f(P^{\mathrm{uni}}, A^u)$ for all $y \in A^u$. Convexity of $f(P^{\mathrm{uni}}, A^u)$ implies that $f(P^{\mathrm{uni}}, A^u) = \Delta(A^u)$. Hence, by population consistency, $\hat{p} \in f(P^\lambda, A^u)$ for all $\lambda \in [0, 1]$.

Now, let $S \in \mathcal{P}^\Delta$ such that $M^S_{A^p \cup \{x\}} = 0$ and $M^S_{yz} = 1$ for all $y \in A^p \cup \{x\}$, $z \in A^u \setminus (A^p \cup \{x\})$. For $\lambda \in [0,1]$, let

$$S^\lambda = \lambda S + (1-\lambda)P^{uni}.$$

Note that every $y \in A^p \cup \{x\}$ is a weak Condorcet winner in S^λ. It follows from population consistency and Lemma 10.9 that, for small $\lambda > 0$, $y \in f(S^\lambda, A^u)$ for all $y \in A^p \cup \{x\}$ and, from convexity of $f(S^\lambda, A^u)$, that $\Delta(A^p \cup \{x\}) \subseteq f(S^\lambda, A^u)$. In particular, $\hat{p} \in f(S^\lambda, A^u)$ for small $\lambda > 0$.

Finally, let

$$P^x = {}^1\!/_3\, P^\lambda + {}^2\!/_3\, S^\lambda,$$

for some small $\lambda > 0$. Population consistency implies that $\hat{p} \in f(P^x, A^u)$. Moreover, $M^{P^x}_{xy} > 0$ for all $y \in A^u \setminus \{x\}$, i.e., $\underline{x}$ is a strict Condorcet winner in $P^x|_{\underline{A^u}}$ Hence, it follows from Lemma 10.9 that $\underline{x} \in f(P^x, A^u)$.

If $p_x > 0$ then, by construction, $\hat{p} = \text{uni}(A^p \cup \{x\}) \in f(P^x, A^u)$. If $p_x = 0$ then $\hat{p} = \text{uni}(A^p) \in f(P^x, A^u)$. In this case it follows from convexity of $f(P^x, A^u)$ that

$$\text{uni}(A^p \cup \{x\}) = {}^1\!/_{(|A^p|+1)}\, \underline{x} + {}^{|A^p|}\!/_{(|A^p|+1)}\, \text{uni}(A^p) \in f(P^x, A^u).$$

Hence, in either case, we get a profile P^x such that $\text{uni}(A^p \cup \{x\}) \in f(P^x, A^u)$ and $M^x = M^{P^x}_{A^p \cup \{x\}}$ takes the form

$$M^x = \lambda \cdot \begin{pmatrix} 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ 1 & \dots & 1 & 0 & 1 & \dots & 1 \\ 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & 0 & \dots & 0 \end{pmatrix}$$

for some $\lambda > 0$ where all entries except the xth row and column are zero. Let $\hat{m} = |A^p \cup \{x\}|$. Let $\pi^y \in \Pi(U)$ such that $\pi^y(x) = y$ and $\pi^y(z) = z$ for all $z \in U \setminus \{x, y\})$ and $P^y = (P^x)^{\pi^y}$. Then, for every $y \in A^p \cup \{x\}$, $M^{P^y}_{A^p \cup \{x\}} = M^x_{\pi^y} = M^y$ and, by neutrality, $\text{uni}(A^p \cup \{x\}) \in f(P^y, A^u)$.

Let $\mathcal{Y} = \{P^y \colon y \in A^p \cup \{x\}\}$. We have that $\dim(\mathcal{Y}|_{\underline{A^u}}) = \hat{m} - 1$ since $\mathcal{Y}|_{\underline{A^u}}$ is a set of affinely independent vectors. Now we determine $\dim(\text{lin}(\mathcal{Y}|_{\underline{A^u}}) \cap \text{lin}(\mathcal{P}^{A^p \cup \{x\}}|_{\underline{A^u}}))$. To this end, let $\lambda P \in \text{lin}(\mathcal{Y}) \cap \text{lin}(\mathcal{P}^{A^p \cup \{x\}})$ with $\lambda \in \mathbb{Q}$ and $P \in \mathcal{P}^\Delta$. Then, there are $\lambda^z \in \mathbb{Q}$ such that $\lambda M^P_{A^p \cup \{x\}} = \sum_{z \in A^p \cup \{x\}} \lambda^z M^z$ and, for all $y \in A^p \cup \{x\}$, $\lambda \sum_{z \in A^p \cup \{x\}} M^P_{yz} = 0$. It follows that $(\hat{m}-1)\lambda^y = \sum_{z \in A^p \cup \{x\} \setminus \{y\}} \lambda^z$ for all $y \in A^p \cup \{x\}$. Hence, $\lambda^y = \lambda^z$ for all $y, z \in A^p \cup \{x\}$ and $\text{lin}(\mathcal{Y}|_{\underline{A^u}}) \cap \text{lin}(\mathcal{P}^{A^p \cup \{x\}}|_{\underline{A^u}}) = \{\lambda \sum_{y \in A^p \cup \{x\}} P^y|_{\underline{A^u}} \colon \lambda \in \mathbb{Q}\}$. □

In Lemma 10.16, we finally show that every proper SCF that satisfies population consistency and composition consistency has to yield maximal lotteries. The structure of the proof is as follows. We assume for contradiction that a proper SCF satisfies population consistency and composition consistency, but returns an outcome that is not maximal. Then we can find a set of profiles whose restriction to the corresponding agenda has full dimension and the uniform profile is in its interior and for which the uniform distribution over a fixed subset of at least two alternatives from the agenda is returned. Thus, this set contains a profile with a strict Condorcet winner whose restriction to the agenda is close to the uniform distribution. For every profile in an ϵ-ball around this strict Condorcet profile, the function has to return the uniform distribution over a non-singleton subset as well as the Condorcet winner, which contradicts decisiveness.

Lemma 10.16

> Every proper SCF f that satisfies population consistency and composition consistency has to yield maximal lotteries, i.e., $f \subseteq$ *ML*.

Proof. Let $A \in \mathcal{F}(U)$. For $|A| = 2$ the statement follows directly from Lemma 10.8. For $|A| > 2$, assume for contradiction that $f \not\subseteq$ *ML*. By Lemma 10.15, there is $\hat{A} \subseteq A$ and $\mathcal{Y} \subseteq \mathcal{P}^\Delta$ such that $\mathcal{Y}|_{\underline{A}}$ has dimension $|\hat{A}| - 1$, $\text{uni}(\hat{A}) \in f(P, A)$ for every $P \in \mathcal{Y}$ and $\text{lin}(\mathcal{Y}|_{\underline{A}}) \cap \text{lin}(\mathcal{P}^{\hat{A}}|_{\underline{A}})$ has dimension 1. By Lemma 10.14, there is $\mathcal{X} \subseteq \mathcal{P}^{\hat{A}}$ such that $\mathcal{X}|_{\underline{A}}$ has dimension $|A|! - |\hat{A}|$ and $\text{uni}(\hat{A}) \in f(P, A)$ for every $P \in \mathcal{X}$. Since $0 \notin \mathcal{X}|_{\underline{A}}$ and $0 \notin \mathcal{Y}|_{\underline{A}}$, $\text{lin}(\mathcal{X}|_{\underline{A}})$ has dimension $|A|! - |\hat{A}| + 1$ and $\text{lin}(\mathcal{Y}|_{\underline{A}})$ has dimension $|\hat{A}|$. Thus, $\text{lin}(\mathcal{X}|_{\underline{A}} \cup \mathcal{Y}|_{\underline{A}})$ has dimension $|A|!$. This implies that $\mathcal{X} \cup \mathcal{Y}$ has dimension $|A|! - 1$.

Furthermore, it follows from population consistency that $\text{uni}(\hat{A}) \in f(P, A)$ for every $P \in \text{conv}(\mathcal{X} \cup \mathcal{Y})$. Since $\text{uni}(\mathcal{D}|_{\underline{A}})$ is in the interior of $\text{conv}(\mathcal{X}|_{\underline{A}} \cup \mathcal{Y}|_{\underline{A}})$, there are $x \in \hat{A}$ and $P^x \in \mathcal{P}^\Delta$ such that $P^x|_{\underline{A}} \in \text{int}_{\mathcal{P}^\Delta|_{\underline{A}}}(\mathcal{X}|_{\underline{A}} \cup \mathcal{Y}|_{\underline{A}})$ and $\underline{x}$ is a strict Condorcet winner in $P^x|_{\underline{A}}$. Hence, there is $\epsilon > 0$ such that, for every $\hat{P} \in B_\epsilon(P^x|_{\underline{A}})$, $\hat{P}|_{\underline{A}} \in \text{conv}(\mathcal{X}|_{\underline{A}} \cup \mathcal{Y}|_{\underline{A}})$ and $\underline{x}$ is a strict Condorcet winner in $\hat{P}|_{\underline{A}}$. Then, we get that $\underline{x} \in f(\hat{P}, A)$ and $\text{uni}(\hat{A}) \in f(\hat{P}, A)$ for every $\hat{P} \in B_\epsilon(P^x)$. Thus, $\{\hat{P} \in \mathcal{P}^\Delta \colon |f(\hat{P}, A)| = 1\}$ is not dense in $\mathcal{P}^\Delta$ at P^x. This contradicts decisiveness of f. □

10.8.4 *ML* $\subseteq f$

In this section we show that every proper SCF f that satisfies population consistency and composition consistency has to yield *all* maximal lotteries. To this end, we first prove an auxiliary lemma. It was shown by McGarvey (1953) that every complete and anti-symmetric relation is the majority relation of some profile with a bounded number of agents. We show an analogous statement for skew-symmetric matrices and fractional preference profiles.

Lemma 10.17

Let $m \in \mathbb{N}$ and $M \in \mathcal{M}$. Then, there are $P \in \mathcal{P}^\Delta$ and $\kappa \in \mathbb{Q}_{>0}$ such that $\kappa M_{[m]} = M^P_{[m]}$. Furthermore, if there is $\pi \in \Pi([m])$ such that $M_{[m]} = (M_\pi)_{[m]}$, then $P|_{\underline{[m]}} = P^\pi|_{\underline{[m]}}$.

Proof. For all $i, j \in [m]$ with $i \neq j$, let $P^{ij} \in \mathcal{P}^\Delta$ be a profile such that, for all $\succ \in \mathcal{D}|_{\underline{[m]}}$, $P^{ij}(\succ) = 1/(m-1)!$ if $\underline{i} \succ \underline{j}$ and $\{i, j\}$ is a component in P^{ij} and $P^{ij}(\succ) = 0$ otherwise. By construction, we have that $P^{ij}(\underline{i}, \underline{j}) = 1$ and $P^{ij}(\underline{x}, \underline{y}) = 0$ for all $x, y \in [m]$ with $\{x, y\} \neq \{i, j\}$. Let $\kappa = 1/\sum_{i,j \in [m] \colon M_{ij} > 0} M_{ij}$ and $P = \kappa \sum_{i,j \colon M_{ij} > 0} M_{ij} P^{ij}$. Then, we have that $\kappa M_{[m]} = M^P_{[m]}$. The second part of the lemma follows from the symmetry of the construction. □

Fix some agenda A. For a profile P which admits a unique maximal lottery on A, it follows from Lemma 10.16 that $f(P, A) = \mathit{ML}(P, A)$. In Lemma 10.18, we show that for every remaining profile P and every vertex of $\mathit{ML}(P, A)$, there is a sequence of profiles converging to P such that every sequence of maximal lotteries for this sequence of profiles converges to this vertex. Since $f \subseteq \mathit{ML}$ by Lemma 10.16 and f is continuous, it follows that this vertex is in $f(P, A)$. Convexity of $f(P, A)$ then implies that $f(P, A) = \mathit{ML}(P, A)$.

Lemma 10.18

Let f be a proper SCF that satisfies population consistency and composition consistency. Then, $\mathit{ML} \subseteq f$.

Proof. Let $P \in \mathcal{P}^\Delta$ and $A \in \mathcal{F}(U)$. We want to show that $f(P, A) = \mathit{ML}(P, A)$. It follows from Lemma 10.16 that $f \subseteq \mathit{ML}$. If $\mathit{ML}(P, A)$ is a singleton, it follows from $f \subseteq \mathit{ML}$ that $f(P, A) = \mathit{ML}(P, A)$. Hence, consider the case where $\mathit{ML}(P, A)$ is not a singleton. By neutrality, we can assume without loss of generality that $A = [m]$ and for simplicity $M = M^P$. Let $p \in \Delta^{\mathbb{Q}}(U)$ be an extreme point of $\mathit{ML}(P, A)$ and assume without loss of generality that $\mathrm{supp}(p) = [k]$.

We first consider the case where k is odd. By Lemma 10.17, there are $S \in \mathcal{P}^\Delta$ and $\kappa \in \mathbb{Q}_{>0}$ such that

$$M^S_A = \kappa \left(\begin{array}{cccccc|ccc}
0 & -\frac{1}{p_1 p_2} & 0 & \cdots & 0 & \frac{1}{p_k p_1} & 1 & \cdots & 1 \\
\frac{1}{p_1 p_2} & & & \ddots & & 0 & & & \\
0 & & \ddots & & & \vdots & \vdots & \ddots & \vdots \\
\vdots & & & & & 0 & & & \\
0 & \ddots & & & & -\frac{1}{p_{k-1} p_k} & & & \\
-\frac{1}{p_k p_1} & 0 & \cdots & 0 & \frac{1}{p_{k-1} p_k} & 0 & 1 & \cdots & 1 \\
\hline
-1 & & \cdots & & & -1 & 0 & \cdots & 0 \\
\vdots & & \ddots & & & \vdots & \vdots & \ddots & \vdots \\
-1 & & \cdots & & & -1 & 0 & \cdots & 0
\end{array}\right)$$

Intuitively, M^S_A corresponds to a weighted cycle on $[k]$. Note that $(p^tM^S)_i = 0$ for all $i \in [k]$ and $(p^tM^S)_i > 0$ for all $i \in A \setminus \text{supp}(p)$, i.e., p is a quasi-strict maximin strategy in M^S_A (cf. Chapter 3). Since p is a maximin strategy in M^S_A, it follows that $p \in ML(S, A)$. For $\epsilon \in [0, 1]$, we define $P^\epsilon = (1-\epsilon)P + \epsilon S$ and $M^\epsilon = M^{P^\epsilon}$. Population consistency implies that $p \in ML(P^\epsilon, A)$ for all $\epsilon \in [0, 1]$. Observe that p is a quasi-strict maximin strategy in M^ϵ_A for every $\epsilon \in (0, 1]$. Hence, for every maximin strategy q in M^ϵ_A, it follows that $(q^tM^\epsilon)_i = 0$ for every $i \in [k]$ and $\text{supp}(q) \subseteq [k]$. It follows from basic linear algebra that

$$\det\left(M^S_{[k-1]}\right) = \kappa^{k-1} \prod_{i=1}^{k-2} \left(\frac{1}{p_i p_{i+1}}\right)^2 \neq 0,$$

and hence, $M^S_{[k]}$ has rank at least $k-1$. In fact, $M^S_{[k]}$ has rank $k-1$, since skew-symmetric matrices of odd size cannot have full rank (cf. Chapter 3). Furthermore, $\det(M^\epsilon_{[k-1]})$ is a polynomial in ϵ of order at most $k-1$ and hence, has at most $k-1$ zeros. Thus, we can find a sequence $(\epsilon^\ell)_{\ell\in\mathbb{N}}$ which converges to zero such that $M^{\epsilon^\ell}_{[k]}$ has rank $k-1$ for all $\ell \in \mathbb{N}$. In particular, if $(q^tM^\epsilon)_i = 0$ for all $i \in [k]$, then $q = p$. This implies that p is the unique maximin strategy in $M^{\epsilon^\ell}_A$ for all $\ell \in \mathbb{N}$. By Lemma 10.16, we know that $f(P^{\epsilon^\ell}, A) \subseteq ML(P^{\epsilon^\ell}, A)$ for all $\ell \in \mathbb{N}$. Hence, $\{p\} = ML(P^{\epsilon^\ell}, A) \subseteq f(P^{\epsilon^\ell}, A)$ for all $\ell \in \mathbb{N}$. It follows from continuity of f that $p \in f(P, A)$.

Now consider the case where k is even. $ML(P, A)$ is a polytope because it is the solution space of a linear feasibility program. Assume that p is a vertex of $ML(P, A)$. Lemma 3.2 implies that p is not a quasi-strict maximin strategy of M^P_A. Hence, we may assume without loss of generality that $(p^tM)_{k+1} = 0$. Let $e_1 = M_{k+1,1}/p_2$ and $e_i = (M_{k+1,i}+p_{i-1}e_{i-1})/p_{i+1}$ for $i \in \{2, \ldots, k-1\}$. By Lemma 10.17, there are $S \in \mathcal{P}^\Delta$ and $\kappa \in \mathbb{Q}_{>0}$ such that

$$M^S_A = \kappa \left(\begin{array}{cccccc|c|ccc}
0 & e_1 & 0 & \ldots & & 0 & 0 & 1 & \ldots & 1 \\
-e_1 & & & & \ddots & \vdots & & & & \\
0 & & \ddots & & & 0 & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & & & & e_{k-1} & & & & \\
0 & \ldots & 0 & & -e_{k-1} & 0 & 0 & & & \\
\hline
0 & & \ldots & & & 0 & 0 & 1 & \ldots & 1 \\
\hline
-1 & & & \ldots & & & -1 & 0 & \ldots & 0 \\
\vdots & & & \ddots & & & \vdots & \vdots & \ddots & \vdots \\
-1 & & & \ldots & & & -1 & 0 & \ldots & 0
\end{array}\right)$$

Note that $M^S_{1,k} = M^S_{k,1} = 0$. For $\epsilon > 0$, let $P^\epsilon = (1-\epsilon)P + \epsilon S$ and $M^\epsilon = M^{P^\epsilon}$. We claim that p^ϵ defined as follows is a maximin strategy in M^ϵ_A. To this end, let $s_\epsilon = \frac{\epsilon\kappa}{1-\epsilon+\epsilon\kappa}$ and

$$p^\epsilon_i = \begin{cases} (1-s_\epsilon)p_i & \text{if } i \in [k], \\ s_\epsilon & \text{if } i = k+1, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Note that $1/\kappa\,(p^t M^S)_1 = -p_2 e_1 = -M_{k+1,1}$ and, for $i \in \{2, \dots, k-1\}$,

$$\begin{aligned} \frac{1}{\kappa}(p^t M^S)_i &= p_{i-1}e_{i-1} - p_{i+1}e_i \\ &= p_{i-1}e_{i-1} - (M_{k+1,i} + p_{i-1}e_{i-1}) = -M_{k+1,i}. \end{aligned}$$

To determine $(p^t M^S)_k$, we first prove inductively that, for all $i \in [k-1]$,

$$p_i e_i = \frac{1}{p_{i+1}} \sum_{j=1}^{i} M_{k+1,j} p_j.$$

For $i = 1$, this follows from the definition of e_1. Now, let $i \in \{2, \dots, k-1\}$. Then,

$$\begin{aligned} p_i e_i &= \frac{p_i}{p_{i+1}}(M_{k+1,i} + p_{i-1}e_{i-1}) \\ &= \frac{p_i}{p_{i+1}}(M_{k+1,i} + \frac{1}{p_i}\sum_{j=1}^{i-1} M_{k+1,j}p_j) \\ &= \frac{1}{p_{i+1}}\sum_{j=1}^{i} M_{k+1,j}p_j, \end{aligned}$$

where the second equality follows from the induction hypothesis. Now,

$$\begin{aligned} \frac{1}{\kappa}(p^t M^S)_k = p_{k-1}e_{k-1} &= \frac{1}{p_k}\sum_{j=1}^{k-1} M_{k+1,j}p_j \\ &= -\frac{1}{p_k}M_{k+1,k}p_k = -M_{k+1,k}, \end{aligned}$$

where the third equality follows from the fact that $(p^t M)_{k+1} = 0$. For $i \in [k]$, it follows from $(p^t M)_i = 0$ that $((p^\epsilon)^t M)_i = s_\epsilon M_{k+1,i}$. Then, for $i \in [k]$,

$$((p^\epsilon)^t M^\epsilon)_i = (1-\epsilon)s_\epsilon M_{k+1,i} + \epsilon\kappa(1-s_\epsilon)(-M_{k+1,i}) = 0.$$

Furthermore, it follows from $(p^t M)_{k+1} = 0$ that $((p^\epsilon)^t M^\epsilon)_{k+1} = 0$ as $M_{k+1,k+1} = 0$, and, for $i \in A \setminus [k+1]$,

$$((p^\epsilon)^t M^\epsilon)_i \geqslant (1-\epsilon)s_\epsilon M_{k+1,i} + \epsilon\kappa \geqslant -(1-\epsilon)s_\epsilon + \epsilon\kappa \geqslant 0.$$

This shows that p^ϵ is a maximin strategy in M_A^ϵ and hence, $p^\epsilon \in \mathit{ML}(P^\epsilon, A)$. Since $|\text{supp}(p^\epsilon)|$ is odd, it follows from the first case that $p^\epsilon \in f(P^\epsilon, A)$. Note that s_ϵ goes to 0 as ϵ goes to 0. Hence, p^ϵ goes to p as ϵ goes to 0. It now follows from continuity of f that $p \in f(P, A)$.

Together, we have that $p \in f(P, A)$ for every vertex p of $\mathit{ML}(P, A)$. Since every outcome in $\mathit{ML}(P, A)$ can be written as a convex combination of vertices, convexity of $f(P, A)$ implies that $f(P, A) = \mathit{ML}(P, A)$.□

Theorem 10.2 then follows directly from Lemmas 10.16 and 10.18.

Theorem 10.2

A proper SCF f satisfies population consistency and composition consistency if and only if $f = \mathit{ML}$.

10.8.5 Proof of Theorem 10.3

The proof of Theorem 10.2 can be adjusted to prove Theorem 10.3 as follows. Lemma 10.8 is not required in the remainder of the proof when assuming Condorcet consistency. It can however be shown easily by observing that Condorcet consistency, continuity, and convexity imply that $\mathit{ML} \subseteq f$ for agendas of size two. The rest of its proof does not make use of composition consistency. Condorcet consistency and continuity imply that weak Condorcet winners have to be chosen whenever they exist. To see this, observe that whenever a pure outcome is a weak Condorcet winner in some profile, there is another profile arbitrarily close to it where this pure outcome is a strict Condorcet winner. Hence, Lemma 10.9 and Lemma 10.10 follow directly from Condorcet consistency and continuity, even when not restricting to profiles that are close to the uniform profile. All implications of composition consistency in the proof of Lemma 10.14 can be derived either from cloning consistency or from the observation that weak Condorcet winners have to be chosen whenever they exist. For the only time composition consistency is applied in the proof of Lemma 10.15, cloning consistency suffices as well. The proofs of all other lemmas do not make use of composition consistency apart from references to earlier lemmas.

Allais, M. (1953). "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine". In: *Econometrica* 21.4, pp. 503–546 [p. 2].

Allesina, S. and J. M. Levine (2011). "A competitive network theory of species diversity". In: *Proceedings of the National Academy of Sciences (PNAS)* 108.14, pp. 5638–5642 [p. 9].

Anand, P. (1993). "The Philosophy of Intransitive Preference". In: *The Economic Journal* 103.417, pp. 337–346 [p. 2].

Anand, P. (2009). "Rationality and intransitive preference: Foundations for the modern view". In: *The Handbook of Rational and Social Choice*. Ed. by P. Anand, P. K. Pattanaik, and C. Puppe. Oxford University Press. Chap. 6 [p. 2].

Arrow, K. J. (1948). *The Possibility of a Universal Social Welfare Function*. Technical Report RAD(L)-289. Project RAND [p. 95].

Arrow, K. J. (1951). *Social Choice and Individual Values*. 1st. 2nd edition 1963. New Haven: Cowles Foundation [pp. 7, 11, 63, 64, 66].

Arrow, K. J. and L. Hurwicz (1972). "An optimality criterion of decision-making under ignorance". In: *Uncertainty and Expectations in Economics*. Ed. by C. F. Carter and J. L. Ford. Basil Blackwell [p. 101].

Athanassoglou, S. (2011). "Efficiency under a combination of ordinal and cardinal information on preferences". In: *Journal of Mathematical Economics* 47.2, pp. 180–185 [p. 61].

Aumann, R. J. (1959). "Acceptable Points in General Cooperative n-Person Games". In: *Contributions to the Theory of Games IV*. Ed. by A. W. Tucker and R. D. Luce. Vol. 40. Annals of Mathematics Studies. Princeton University Press, pp. 287–324 [p. 30].

Aumann, R. J. (1974). "Subjectivity and Correlation in Randomized Strategies". In: *Journal of Mathematical Economics* 1, pp. 67–96 [p. 37].

Aumann, R. J. and A. Brandenburger (1995). "Epistemic Conditions for Nash Equilibrium". In: *Econometrica* 63.5, pp. 1161–1180 [pp. 7, 38, 39].

Aumann, R. J. and J. H. Drèze (2008). "Rational Expections in Games". In: *American Economic Review* 98.1, pp. 72–86 [pp. 7, 38–40].

Aziz, H., F. Brandl, and F. Brandt (2015). "Universal Pareto Dominance and Welfare for Plausible Utility Functions". In: *Journal of Mathematical Economics* 60, pp. 123–133 [pp. 26, 61, 91].

Aziz, H., F. Brandl, F. Brandt, and M. Brill (2018). "On the Tradeoff between Efficiency and Strategyproofness". In: *Games and Economic Behavior* 110, pp. 1–18 [pp. 26, 104].

Aziz, H., F. Brandt, and M. Brill (2013). "On the Tradeoff between Economic Efficiency and Strategyproofness in Randomized Social Choice". In: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IFAAMAS, pp. 455–462 [p. 62].

Bach, C. W. and E. Tsakas (2014). "Pairwise epistemic conditions for Nash equilibrium". In: *Games and Economic Behavior* 85, pp. 48–59 [p. 39].

Banks, J. S. (1995). "Acyclic Social Choice From Finite Sets". In: *Social Choice and Welfare* 12, pp. 293–310 [p. 63].

Bar-Hillel, M. and A. Margalit (1988). "How vicious are cycles of intransitive choice?" In: *Theory and Decision* 24.2, pp. 119–145 [p. 2].

Barelli, P. (2009). "Consistency of beliefs and epistemic conditions for Nash and correlated equilibria". In: *Games and Economic Behavior* 67.2, pp. 363–375 [pp. 7, 39].

Bergson, A. (1938). "A Reformulation of Certain Aspects of Welfare Economics". In: *The Quarterly Journal of Economics* 52.2, pp. 310–334 [p. 11].

Bergstrom, T. C. (1992). "When Non-transitive Relations Take Maxima and Competitive Equilibrium Can't be Beat". In: *Economic Theory and International Trade (Essays in Memoriam of J. Trout Rader)*. Ed. by W. Neuefeind and R. G. Riezmann. Springer-Verlag, pp. 29–52 [p. 22].

Bernheim, B. D. (1984). "Rationalizable Strategic Behavior". In: *Econometrica* 52.4, pp. 1007–1028 [p. 39].

Black, D. (1948). "On the Rationale of Group Decision-making". In: *Journal of Political Economy* 56.1, pp. 23–34 [p. 64].

Blair, D. H. and R. A. Pollak (1982). "Acyclic Collective Choice Rules". In: *Econometrica* 50.4, pp. 931–943 [p. 63].

Blau, J. H. and R. Deb (1977). "Social Decision Functions and the Veto". In: *Econometrica* 45.4, pp. 871–879 [p. 63].

Blavatskyy, P. R. (2006). "Axiomatization of a preference for most probable winner". In: *Theory and Decision* 60.1, pp. 17–33 [pp. 3, 13, 24–26].

Blyth, C. R. (1972). "Some Probability Paradoxes in Choice from Among Random Alternatives". In: *Journal of the American Statistical Association* 67.338, pp. 366–373 [pp. 25, 26].

Bohnenblust, H. F., S. Karlin, and L. S. Shapley (1950). "Solutions of discrete, two-person games". In: *Contributions to the Theory of Games*. Ed. by H. W. Kuhn and A. W. Tucker. Vol. 1. Princeton University Press, pp. 51–72 [p. 5].

Borda, Chevalier de (1784). *Memoire sur les Elections au Scrutin*. Histoire de l'Academie Royale des Sciences [pp. 10, 101].

Border, K. C. (1983). "Social welfare functions for economic environments with and without the Pareto principle". In: *Journal of Economic Theory* 29.2, pp. 205–216 [p. 64].

Bordes, G. and M. Le Breton (1989). "Arrovian theorems with private alternatives domains and selfish individuals". In: *Journal of Economic Theory* 47.2, pp. 257–281 [p. 64].

Bordes, G. and M. Le Breton (1990a). "Arrovian theorems for economic domains: Assignments, matchings and pairings". In: *Social Choice and Welfare* 7.3, pp. 193–208 [p. 64].

Bordes, G. and M. Le Breton (1990b). "Arrovian theorems for economic domains. The case where there are simultaneously private and public goods". In: *Social Choice and Welfare* 7.1, pp. 1–17 [p. 64].

Börgers, T. and Y.-M. Choo (2015). "Revealed Relative Utilitarianism". Working paper [p. 65].

Brandl, F., F. Brandt, C. Geist, and J. Hofbauer (2015a). "Strategic Abstention based on Preference Extensions: Positive Results and Computer-Generated Impossibilities". In: *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*. AAAI Press, pp. 18–24 [p. 87].

Brandl, F., F. Brandt, and J. Hofbauer (2015b). "Incentives for Participation and Abstention in Probabilistic Social Choice". In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. IFAAMAS, pp. 1411–1419 [pp. 87–89, 91].

Brandl, F., F. Brandt, and J. Hofbauer (2018). "Welfare Maximization Entices Participation". In: *Games and Economic Behavior*. Forthcoming [p. 26].

Brandt, F. (2011). "Minimal Stable Sets in Tournaments". In: *Journal of Economic Theory* 146.4, pp. 1481–1499 [p. 101].

Brandt, F., M. Brill, and H. G. Seedig (2011). "On the Fixed-Parameter Tractability of Composition-Consistent Tournament Solutions". In: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*. AAAI Press, pp. 85–90 [p. 101].

Brandt, F., C. Geist, and D. Peters (2017). "Optimal Bounds for the No-Show Paradox via SAT Solving". In: *Mathematical Social Sciences* 90. Special Issue in Honor of Hervé Moulin, pp. 18–27 [pp. 87, 93].

Brown, D. J. (1975). "Aggregation of Preferences". In: *Quarterly Journal of Economics* 89.3, pp. 456–469 [p. 63].

Budish, E. and E. Cantillion (2012). "The Multi-Unit Assignment Problem: Theory and Evidence from Course Allocation at Harvard". In: *American Economic Review* 102.5, pp. 2237–2271 [p. 96].

Butler, D., G. Pogrebna, and P. R. Blavatskyy (2016). "Predictably Intransitive Preferences". In: Working paper [pp. 13, 26].

Campbell, D. E. (1989). "Arrow's Theorem for Economic Environments and Effective Social Preferences". In: *Social Choice and Welfare* 6.4, pp. 325–329 [p. 64].

Campbell, D. E. and J. S. Kelly (2002). "Impossibility Theorems in the Arrovian Framework". In: *Handbook of Social Choice and Welfare*.

Ed. by K. J. Arrow, A. K. Sen, and K. Suzumura. Vol. 1. Elsevier. Chap. 1 [p. 63].

Carroll, G. (2010). "An efficiency theorem for incompletely known preferences". In: *Journal of Economic Theory* 145.6, pp. 2463–2470 [p. 61].

Casti, J. L. (1996). *Five golden rules : great theories of 20th-century mathematics, and why they matter*. John Wiley & Sons [p. 6].

Che, Y.-K. and F. Kojima (2010). "Asymptotic Equivalence of Probabilistic Serial and Random Priority Mechanisms". In: *Econometrica* 78.5, pp. 1625–1672 [p. 96].

Chen, X., X. Deng, and S.-H. Teng (2009). "Settling the Complexity of Computing Two-Player Nash Equilibria". In: *Journal of the ACM* 56.3 [p. 37].

Chernoff, H. (1954). "Rational Selection of Decision Functions". In: *Econometrica* 22.4, pp. 422–443 [pp. 22, 101].

Chew, S. H. (1983). "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income and Decision Theory Resolving the Allais Paradox". In: *Econometrica* 51.4, pp. 1065–1092 [p. 24].

Condorcet, Marquis de (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Facsimile published in 1972 by Chelsea Publishing Company, New York. Imprimerie Royale [pp. 10, 95, 101].

Congar, R. and V. Merlin (2012). "A characterization of the maximin rule in the context of voting". In: *Theory and Decision* 72.1, pp. 131–147 [p. 100].

Dasgupta, P. and E. Maskin (2008). "On the robustness of majority rule". In: *Journal of the European Economic Association* 6.5, pp. 949–973 [pp. 96, 102].

Daskalakis, C., P. Goldberg, and C. Papadimitriou (2009). "The complexity of computing a Nash equilibrium". In: *SIAM Journal on Computing* 39.1, pp. 195–259 [p. 37].

Dhillon, A. (1998). "Extended Pareto rules and relative utilitarianism". In: *Social Choice and Welfare* 15.4, pp. 521–542 [pp. 61, 65].

Dhillon, A. and J.-F. Mertens (1997). "An impossibility theorem with von Neumann-Morgenstern preferences". In: *Economic Letters* 56.3, pp. 305–309 [p. 12].

Dhillon, A. and J.-F. Mertens (1999). "Relative Utilitarianism". In: *Econometrica* 67.3, pp. 471–498 [pp. 12, 61, 65].

Dogan, B. and K. Yildiz (2016). "Efficiency and stability of probabilistic assignments in marriage problems". In: *Games and Economic Behavior* 95, pp. 47–58 [p. 61].

Duddy, C. (2014). "Electing a representative committee by approval ballot: An impossibility result". In: *Economic Letters* 124.1, pp. 14–16 [p. 87].

Ehlers, L. and T. Storcken (2008). "Arrow's Possibility Theorem for One-Dimensional Single-Peaked Preferences". In: *Games and Economic Behavior* 64.2, pp. 533–547 [p. 64].

Fan, K. (1952). "Fixed-Point and Minimax Theorems in Locally Convex Topological Linear Spaces". In: *Proceedings of the National Academy of Sciences of the United States of America* 38.2, pp. 121–126 [p. 6].

Fan, K. (1953). "Minimax Theorems". In: *Proceedings of the National Academy of Sciences (PNAS)* 39.1, pp. 42–47 [p. 6].

Faris, W. G. and R. S. Maier (1987). "The Value of a Random Game: The Advantage of Rationality". In: *Complex Systems* 1, pp. 235–244 [pp. 8, 50].

Fine, B. and K. Fine (1974). "Social Choice and Individual Ranking I". In: *Review of Economic Studies* 41.127, pp. 303–323 [pp. 15, 100].

Fishburn, P. C. (1970). "The Irrationality of Transitity in Social Choice". In: *Behavioral Science* 15, pp. 119–123 [p. 2].

Fishburn, P. C. (1972). "Lotteries and Social Choices". In: *Journal of Economic Theory* 5.2, pp. 189–207 [p. 98].

Fishburn, P. C. (1973). *The Theory of Social Choice*. Princeton University Press [pp. 10, 98].

Fishburn, P. C. (1978). "Axioms for approval voting: Direct proof". In: *Journal of Economic Theory* 19.1, pp. 180–185 [p. 100].

Fishburn, P. C. (1982). "Nontransitive measurable utility". In: *Journal of Mathematical Psychology* 26.1, pp. 31–67 [pp. 3, 21, 23, 24].

Fishburn, P. C. (1983). "Transitive measurable utility". In: *Journal of Economic Theory* 31.2, pp. 293–317 [p. 24].

Fishburn, P. C. (1984a). "Dominance in SSB utility theory". In: *Journal of Economic Theory* 34.1, pp. 130–148 [pp. 74, 91].

Fishburn, P. C. (1984b). "Probabilistic Social Choice Based on Simple Voting Comparisons". In: *Review of Economic Studies* 51.4, pp. 683–692 [pp. 14, 62, 104].

Fishburn, P. C. (1984c). "SSB utility theory: An economic perspective". In: *Mathematical Social Sciences* 8.1, pp. 63–94 [pp. 3, 23–25].

Fishburn, P. C. (1988). *Nonlinear preference and utility theory*. The Johns Hopkins University Press [pp. 21, 23].

Fishburn, P. C. (1991). "Nontransitive preferences in decision theory". In: *Journal of Risk and Uncertainty* 4.2, pp. 113–134 [pp. 2, 3, 71].

Fishburn, P. C. and S. J. Brams (1983). "Paradoxes of Preferential Voting". In: *Mathematics Magazine* 56.4, pp. 207–214 [pp. 14, 87].

Fishburn, P. C. and W. V. Gehrlein (1977). "Towards a Theory of Elections with Probabilistic Preferences". In: *Econometrica* 45.8, pp. 1907–1924 [pp. 102, 103].

Fishburn, P. C. and W. V. Gehrlein (1987). "Aggregation theory for SSB utility functionals". In: *Journal of Economic Theory* 42.2, pp. 352–369 [pp. 13, 65, 74].

Fisher, D. C. and R. B. Reeves (1995). "Optimal Strategies for Random Tournament Games". In: *Linear Algebra and its Applications* 217, pp. 83–85 [p. 51].

Fisher, D. C. and J. Ryan (1992). "Optimal Strategies for a Generalized "Scissors, Paper, and Stone" Game". In: *American Mathematical Monthly* 99.10, pp. 935–942 [p. 50].

Fisher, D. C. and J. Ryan (1995). "Tournament Games and Positive Tournaments". In: *Journal of Graph Theory* 19.2, pp. 217–236 [p. 32].

Gale, D., H. W. Kuhn, and A. W. Tucker (1950). "On Symmetric Games". In: *Contributions to the Theory of Games*. Ed. by H. W. Kuhn and A. W. Tucker. Vol. 1. Princeton University Press, pp. 81–87 [pp. 33–35].

Gintis, H. (2009). *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press [p. 39].

Grether, D. M. and C. R. Plott (1979). "Economic theory of choice and the preference reversal phenomenon". In: *American Economic Review* 69.4, pp. 623–638 [p. 2].

Grilli, J., G. Barabás, M. J. Michalska-Smith, and S. Allesina (2017). "Higher-order interactions stabilize dynamics in competitive network models". In: *Nature* 548, pp. 210–213 [p. 9].

Harsanyi, J. C. (1955). "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility". In: *Journal of Political Economy* 63.4, pp. 309–321 [pp. 12, 13, 64].

Harsanyi, J. C. (1967). "Games with Incomplete Information Played by "Bayesian" Players, Part I". In: *Management Science* 50.12, pp. 1804–1817 [pp. 7, 38].

Harsanyi, J. C. (1973a). "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-strategy Equilibrium Points". In: *International Journal of Game Theory* 2.1, pp. 1–23 [p. 30].

Harsanyi, J. C. (1973b). "Oddness of the Number of Equilibrium Points: A New Proof". In: *International Journal of Game Theory* 2.1, pp. 235–250 [p. 49].

Heims, S. J. (1980). *John Von Neumann and Norbert Wiener: from mathematics to the technologies of life and death*. MIT Press [p. 6].

Hellman, Z. (2013). "Weakly rational expectations". In: *Journal of Mathematical Economics* 49.6, pp. 496–500 [p. 39].

Holzman, R. (1988). "To vote or not to vote: What is the quota?" In: *Discrete Applied Mathematics* 22.2, pp. 133–141 [p. 87].

Horan, S. (2013). "Implementation of Majority Voting Rules". In: Working paper [p. 101].

Hylland, A. (1980a). "Aggregation procedure for cardinal preferences: A comment". In: *Econometrica* 48.2, pp. 539–542 [pp. 12, 64].

Hylland, A. (1980b). "Strategyproofness of Voting Procedures with Lotteries as Outcomes and Infinite Sets of Strategies". Mimeo [p. 12].

Inada, K. (1969). "The simple majority decision rule". In: *Econometrica* 37, pp. 490–506 [p. 64].

Jimeno, J. L., J. Pérez, and E. García (2009). "An extension of the Moulin No Show Paradox for voting correspondences". In: *Social Choice and Welfare* 33.3, pp. 343–459 [p. 87].

Jonasson, J. (2004). "On the Optimal Strategy in a Random Game". In: *Electronic Communications in Probability* 9, pp. 132–139 [pp. 8, 50].

Kahnemann, D. and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk". In: *Econometrica* 47.2, pp. 263–292 [p. 2].

Kalai, E., E. Muller, and M. A. Satterthwaite (1979). "Social Welfare Functions When Preferences Are Convex, Strictly Monotonic, and Continuous". In: *Public Choice* 34.1, pp. 87–97 [p. 64].

Kalai, E. and D. Schmeidler (1977a). "Aggregation Procedure for Cardinal Preferences: A Formulation and Proof of Samuelson's Impossibility Conjecture". In: *Econometrica* 45.6, pp. 1431–1438 [pp. 12, 64, 66].

Kalai, E. and D. Schmeidler (1977b). "An admissible set occurring in various bargaining situations". In: *Journal of Economic Theory* 14, pp. 402–411 [pp. 12, 73].

Kaplansky, I. (1945). "A Contribution to von Neumann's Theory of Games". In: *Annals of Mathematics* 46.3, pp. 474–479 [pp. 5, 53].

Kaplansky, I. (1995). "A contribution to von Neumann's theory of games. II". In: *Linear Algebra and its Applications* 226-228, pp. 371–373 [p. 53].

Karni, E. (2014). "Axiomatic Foundations of Expected Utility and Subjective Probability". In: *Handbook of the Economics of Risk and Uncertainty*. Ed. by M. Machina and K. Viscusi. Vol. 1. Elsevier. Chap. 1 [p. 21].

Kelly, J. S. (1978). *Arrow Impossibility Theorems*. Economic Theory and Mathematical Economics. Academic Press [p. 63].

Kjeldsen, T. H. (2001). "John von Neumann's Conception of the Minimax Theorem: A journey through different mathematical contexts". In: *Archive for History of Exact Solutions* 56.1, pp. 39–68 [p. 6].

Knebel, J., M. F. Weber, T. Krüger, and E. Frey (2015). "Evolutionary games of condensates in coupled birth–death processes". In: *Nature Communications* 6.6977 [p. 9].

Kreweras, G. (1965). "Aggregation of Preference Orderings". In: *Mathematics and Social Sciences I: Proceedings of the seminars of Menthon-Saint-Bernard, France (1–27 July 1960) and of Gösing, Austria (3–27 July 1962)*, pp. 73–79 [pp. 14, 62].

Kuhn, H. W. and A. W. Tucker, eds. (1950). *Contributions to the Theory of Games I*. Annals of Mathematics Studies 24. Princeton University Press [p. 5].

Laffond, G., J. Lainé, and J.-F. Laslier (1996). "Composition-Consistent Tournament Solutions and Social Choice Functions". In: *Social Choice and Welfare* 13.1, pp. 75–93 [pp. 16, 101, 102].

Laffond, G., J.-F. Laslier, and M. Le Breton (1997). “A Theorem on Symmetric Two-Player Zero-Sum Games”. In: *Journal of Economic Theory* 72.2, pp. 426–431 [pp. 50, 62, 109].

Laslier, J.-F. (1996). “Rank-Based Choice Correspondences”. In: *Economics Letters* 52.3, pp. 279–286 [pp. 101, 102].

Laslier, J.-F. (1997). *Tournament Solutions and Majority Voting*. Springer-Verlag [p. 101].

Le Breton, M. (1986). “Essais sur les fondements de l’analyse économique de l’inégalité”. Thèse pour le Doctorat d’État en Sciences Économiques. Université de Rennes 1 [p. 12].

Le Breton, M. (2005). “On the Uniqueness of Equilibrium in Symmetric Two-Player Zero-Sum Games with Integer Payoffs”. In: *Économie publique* 17.2, pp. 187–195 [p. 50].

Levine, J. M., J. Bascompte, P. B. Adler, and S. Allesina (2017). “Beyond pairwise mechanisms of species coexistence in complex communities”. In: *Nature* 546, pp. 56–64 [p. 9].

Llinares, J.-V. (1998). “Unified treatment of the problem of existence of maximal elements in binary relations: A characterization”. In: *Journal of Mathematical Economics* 29.3, pp. 285–302 [p. 22].

Loomes, G. and R. Sugden (1982). “Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty”. In: *The Economic Journal* 92.368, pp. 805–824 [p. 24].

Loomes, G. and R. Sugden (1987). “Some implications of a more general form of regret theory”. In: *Journal of Economic Theory* 41.2, pp. 270–287 [p. 24].

Loomis, L. H. (1947). “On A Theorem of von Neumann”. In: *Proceedings of the National Academy of Sciences (PNAS)* 33.11, pp. 329–331 [p. 6].

Machina, M. J. (1983). “Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom”. In: *Foundations of Utility and Risk Theory with Applications*. Ed. by B. Stigum and F. Wenstop. Springer. Chap. 5 [p. 2].

Machina, M. J. (1989). “Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty”. In: *Journal of Economic Literature* 27.4, pp. 1622–1668 [p. 2].

Manea, M. (2008). “A constructive proof of the ordinal efficiency welfare theorem”. In: *Journal of Economic Theory* 141.1, pp. 276–281 [p. 61].

Mas-Colell, A. and H. Sonnenschein (1972). “General Possibility Theorems for Group Decisions”. In: *Review of Economic Studies* 39.2, pp. 185–192 [p. 63].

Maskin, E. (1979). “Decision-making under ignorance with implications for social choice”. In: *Theory and Decision* 11.3, pp. 319–337 [p. 101].

May, K. (1952). “A Set of Independent, Necessary and Sufficient Conditions for Simple Majority Decisions”. In: *Econometrica* 20.4, pp. 680–684 [p. 102].

May, K. (1954). “Intransitivity, Utility, and the Aggregation of Preference Patters”. In: *Econometrica* 22.1, pp. 1–13 [p. 2].

McClennen, E. F. (1988). “Sure-Thing Doubts”. In: *Decision, Probability and Utility*. Ed. by P. Gärdenfors and N.-E. Sahlin. Cambridge University Press. Chap. 10 [p. 2].

McGarvey, D. C. (1953). “A Theorem on the Construction of Voting Paradoxes”. In: *Econometrica* 21.4, pp. 608–610 [p. 123].

McLennan, A. (2002). “Ordinal Efficiency and the Polyhedral Separating Hyperplane Theorem”. In: *Journal of Economic Theory* 105.2, pp. 435–449 [p. 61].

McLennan, A. (2005). “The Expected Number of Nash Equilibria of a Normal Form Game”. In: *Econometrica* 73.1, pp. 141–174 [pp. 8, 49–51].

McLennan, A. and J. Berg (2005). “Asymptotic expected number of Nash equilibria of two-player normal form games”. In: *Games and Economic Behavior* 51.2, pp. 264–295 [pp. 8, 50].

Mongin, P. (1994). “Harsanyi’s Aggregation Theorem: multi-profile version and unsettled questions”. In: *Social Choice and Welfare* 11.4, pp. 331–354 [pp. 12, 65].

Monjardet, B. (2008). “Statement of precedence and a comment on IIA terminology”. In: *Games and Economic Behavior* 62, pp. 736–738 [p. 22].

Moulin, H. (1986). “Choosing from a Tournament”. In: *Social Choice and Welfare* 3.4, pp. 271–291 [p. 101].

Moulin, H. (1988). “Condorcet’s Principle implies the No Show Paradox”. In: *Journal of Economic Theory* 45.1, pp. 53–64 [pp. 14, 87, 88, 92, 93].

Myerson, R. B. (1995). “Axiomatic derivation of scoring rules without the ordering assumption”. In: *Social Choice and Welfare* 12.1, pp. 59–74 [p. 100].

Nash, J. F. (1950a). “Equilibrium Points in n-Person Games”. In: *Proceedings of the National Academy of Sciences (PNAS)* 36, pp. 48–49 [pp. 9, 37].

Nash, J. F. (1950b). “The Bargaining Problem”. In: *Econometrica* 18.2, pp. 155–162 [p. 22].

Nash, J. F. (1951). “Non-cooperative games”. In: *Annals of Mathematics* 54.2, pp. 286–295 [p. 10].

Packard, D. J. (1982). “Cyclical Preference Logic”. In: *Theory and Decision* 14.4, pp. 415–426 [pp. 25, 26].

Pearce, D.G. (1984). “Rationalizable Strategic Behavior and the Problem of Perfection”. In: *Econometrica* 52.4, pp. 1029–1050 [p. 39].

Perea, A. (2007). “A One-Person Doxastic Characterization of Nash strategies”. In: *Synthese* 158, pp. 251–271 [p. 38].

Perea, A. (2012). *Epistemic Game Theory: Reasoning and Choice*. Cambridge University Press [pp. 7, 38].

Pérez, J. (2001). "The Strong No Show Paradoxes are a common flaw in Condorcet voting correspondences". In: *Social Choice and Welfare* 18.3, pp. 601–616 [p. 87].

Peters, D. (2017). "Condorcet's Principle and the Preference Reversal Paradox". In: *Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, pp. 455–469 [p. 92].

Puppe, C. (2016). "The Single-Peaked Domain Revisited: A Simple Global Characterization". In: Working paper [p. 15].

Raghavan, T. E. S. (1994). "Zero-Sum Two Person Games". In: *Handbook of Game Theory with Economic Applications*. Ed. by R. J. Aumann and S. Hart. North-Holland. Chap. 20 [p. 30].

Redekop, J. (1995). "Arrow theorems in econonmic environments". In: *Social choice, welfare, and ethics*. Ed. by W. A. Barnett, H. Moulin, M. Salles, and N. J. Schofield. Cambridge University Press, pp. 163–185 [p. 64].

Roberts, D. P. (2004). "Kernel Sizes for Random Matrix Games". Unpublished manuscript [p. 50].

Roberts, D. P. (2006). "Nash equilibria of Cauchy-random zero-sum and coordination matrix games". In: *International Journal of Game Theory* 34.2, pp. 167–184 [p. 50].

Rubinstein, A. and U. Segal (2012). "On the likelihood of cyclic comparisons". In: *Journal of Economic Theory* 147.6, pp. 2483–2491 [p. 26].

Saari, D. G. (1990). "Consistency of decision processes". In: *Annals of Operations Research* 23.1, pp. 103–137 [p. 100].

Saari, D. G. (1995). *Basic Geometry of Voting*. Springer [pp. 96, 100, 116].

Samuelson, P. A. (1938). "A Note on the Pure Theory of Consumers' Behaviour". In: *Econometrica* 5, pp. 61–71 [p. 95].

Samuelson, P. A. (1967). "Arrow's Mathematical Politics". In: *Human Values and Economic Policy*. Ed. by S. Hook. New York University Press, pp. 41–51 [p. 64].

Sanver, M. R. and W. S. Zwicker (2009). "One-way monotonicity as a form of strategy-proofness". In: *International Journal of Game Theory* 38.4, pp. 553–574 [pp. 87, 92].

Saunders, B. (2010). "Democracy, Political Equality, and Majority Rule". In: *Ethics* 121.1, pp. 148–177 [p. 102].

Schwartz, T. (1970). "On the Possibility of Rational Policy Evaluation". In: *Theory and Decision* 1.1, pp. 89–106 [p. 12].

Schwartz, T. (1986). *The Logic of Collective Choice*. Columbia University Press [p. 63].

Sen, A. K. (1969). "Quasi-Transitivity, Rational Choice and Collective Decision". In: *Review of Economic Studies* 36.3, pp. 381–393 [pp. 14, 22, 63].

Sen, A. K. (1970a). *Collective Choice and Social Welfare*. North-Holland [p. 63].

Sen, A. K. (1970b). "Interpersonal Aggregation and Partial Comparability". In: *Econometrica* 38.3, pp. 393–409 [p. 12].

Sen, A. K. (1971). "Choice Functions and Revealed Preference". In: *Review of Economic Studies* 38.3, pp. 307–317 [pp. 14, 22].

Sen, A. K. (1977). "Social Choice Theory: A Re-Examination". In: *Econometrica* 45.1, pp. 53–89 [p. 63].

Sen, A. K. (1986). "Social Choice Theory". In: *Handbook of Mathematical Economics*. Ed. by K. J. Arrow and M. D. Intriligator. Vol. 3. Elsevier. Chap. 22, pp. 1073–1181 [pp. 63, 77].

Sen, A. K. and P. K. Pattanaik (1969). "Necessary and Sufficient Conditions for Rational Choice under Majority Decision". In: *Journal of Economic Theory* 1, pp. 178–202 [p. 64].

Smith, J. H. (1973). "Aggregation of Preferences with Variable Electorate". In: *Econometrica* 41.6, pp. 1027–1041 [pp. 15, 100, 101].

Sonnenschein, H. (1971). "Demand Theory Without Transitive Preference with Applications to the Theory of Competitive Equilibrium". In: *Preferences, Utility and Demand*. Ed. by J. Chipman, L. Hurwicz, M. Richter, and H. Sonnenschein. Houghton Mifflin Harcourt [pp. 3, 22].

Steinhaus, H. and S. Trybula (1959). "On a paradox in applied probabilities". In: *Bulletin of the Polish Academy of Sciences* 7, pp. 67–69 [p. 26].

Stiemke, E. (1915). "Über positive Lösungen homogener linearer Gleichungen". In: *Mathematische Annalen* 76, pp. 340–342 [p. 33].

Tan, T. C.-C. and S. R. da Costa Werlang (1988). "The Bayesian foundations of solution concepts of games". In: *Journal of Economic Theory* 45.2, pp. 370–391 [p. 39].

Tideman, T. N. (1987). "Independence of Clones as a Criterion for Voting Rules". In: *Social Choice and Welfare* 4.3, pp. 185–206 [pp. 15, 101].

Turunen-Red, A. H. and J. A. Weymark (1999). "Linear Aggregation of SSB Utility Functionals". In: *Theory and Decision* 46.3, pp. 281–294 [pp. 13, 65].

Von Neumann, J. (1928). "Zur Theorie der Gesellschaftspiele". In: *Mathematische Annalen* 100.1, pp. 295–320 [pp. 5, 6, 30, 33].

Von Neumann, J. (1937). "Über ein ökonomisches Gleichgewichtssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes". In: *Ereignisse des Mathematischen Kolloquiums*, pp. 73–83 [p. 6].

Von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton University Press [pp. 2, 5, 6].

Von Neumann, J. and O. Morgenstern (1953). *Theory of Games and Economic Behavior*. 3rd. Princeton University Press [pp. 21, 71].

Wald, A. (1945a). "Generalization of a Theorem By v. Neumann Concerning Zero Sum Two Person Games". In: *Annals of Mathematics* 46.2, pp. 281–286 [p. 6].

Wald, A. (1945b). "Statistical Decision Functions Which Minimize the Maximum Risk". In: *Annals of Mathematics* 46.2, pp. 265–280 [p. 5].

Wilson, R. (1971). "Computing Equilibria of N-Person Games". In: *SIAM Journal on Applied Mathematics* 21.1, pp. 80–87 [pp. 49, 50].

Wilson, R. (1987). "Game Theoretic Analysis of Trading Processes". In: *Advances in Economic Theory*. Ed. by T. Bewley. Cambridge University Press [p. 38].

Young, H. P. (1974a). "A Note on Preference Aggregation". In: *Econometrica* 42.6, pp. 1129–1131 [pp. 96, 100, 101].

Young, H. P. (1974b). "An axiomatization of Borda's rule". In: *Journal of Economic Theory* 9.1, pp. 43–52 [pp. 15, 93, 98, 100].

Young, H. P. (1975). "Social Choice Scoring Functions". In: *SIAM Journal on Applied Mathematics* 28.4, pp. 824–838 [pp. 15, 96, 100, 101].

Young, H. P. and A. Levenglick (1978). "A Consistent Extension of Condorcet's Election Principle". In: *SIAM Journal on Applied Mathematics* 35.2, pp. 285–300 [pp. 10, 96, 100–102].

Zavist, T. M. and T. N. Tideman (1989). "Complete independence of clones in the ranked pairs rule". In: *Social Choice and Welfare* 6.2, pp. 167–173 [p. 101].

Zwicker, W. S. (1991). "The voter's paradox, spin, and the Borda count". In: *Mathematical Social Sciences* 22.3, pp. 187–227 [p. 116].