Putting Context into Preference Aggregation

Philipp Peitler[*] Karl H. Schlag[§]

February 2023

[Latest version]

Abstract

The axioms underlying Arrow's impossibility theorem are very restrictive in terms of what can be used when aggregating preferences. Social preferences may not depend on the menu nor on preferences over alternatives outside the menu. But context matters. So we weaken these restrictions to allow for context to be included. The context as we define describes which alternatives in the menu and which preferences over alternatives outside the menu matter. We obtain unique representations. These are discussed in examples involving markets, bargaining and intertemporal well being of an individual. Proofs are constructive and insightful.

Keywords: social choice, preference aggregation, relative utilitarianism, menu independence, IIA, rational preferences.

1 Introduction

A central question in economics is how a social planner should compare alternatives when the preferences of the individuals involved are in conflict. To determine what is best for those individuals involved, the planer needs a rule that aggregates individual preferences into a single social preference. Different strands of the economic literature

[*]University of Vienna; philipp.peitler@univie.ac.at

[§]University of Vienna; karl.schlag@univie.ac.at

use very different rules. In the analysis of markets, the standard tool to evaluate the joint welfare of consumers is consumer surplus. In a bargaining problem, the rules for finding a fair solution can be interpreted as the choice of an implicit social planner who is trading off benefits of the different parties involved. The fair solution is determined either by the Nash product (Nash, 1950) or is taken to be the equitable outcome on the Pareto frontier (Kalai and Smorodinsky, 1975). The intertemporal welfare of a dynamically inconsistent decision maker is evaluated either by the Pareto criterion or by the long-run utility (O'Donoghue and Rabin, 1999). Even within their respective settings, these criteria can only be applied under narrow modeling assumptions. More importantly, each of them fails to conform to a minimal set of desiderata. Consumer surplus violates the Neutrality axiom, as it is sensitive to the labels of social alternatives. Neither the Nash product nor the equitable outcome is rational, in the sense of abiding by the von Neumann and Morgenstern (1944) axioms. The Pareto criterion does not generate complete preferences and long-run utility violates the Pareto principle. Some of these violations are not obvious and will be demonstrated later. In each of these settings we could search separately for other rules. However, we feel that the well-beings of individuals should be traded off based on principles that are appealing independently of the application. Therefore, we would like to have a universal rule that can be applied to every setting. In light of Arrow (1950, 1963), it is not clear whether such a rule exists. In the following we argue that some of Arrow's demands are too stringent and should be relaxed. Before we state our contentions, let us briefly revisit Arrow's theorem.

Arrow (1963) shows that there is no aggregation rule that satisfies completeness, transitivity, the Pareto principle, non-dictatorship and *independence of irrelevant alternatives* (IIA). If we want the social planner to be rational, benevolent and impartial, then the first four axioms of Arrow serve as minimal requirements. Consider now the fifth and last axiom, IIA. According to Arrow, IIA demands that when the planner chooses from a menu of alternatives, denoted by S, the choice $C(S)$ is not allowed to depend on individual preferences over alternatives outside S. In addition to the above, Arrow assumes that social choice from any menu is made according to a single social preference over all alternatives. Hence, Arrow implicitly assumes one more axiom, namely *menu independence* (MI). Together, these axioms make it impossible to aggregate preferences. Weakening either IIA or MI are potential avenues to escape the impossibility. We now argue that one can weaken either of them as neither is as desirable as the first four axioms mentioned above.

First, we asses IIA with the aid of the following example by Pearce (2021). Imagine being tasked with choosing between two social alternatives x and y for a group of five kindergarten children. Strict preferences of the children are given by the following table.

1	2	3	4	5
x	x	x	x	y
y	y	y	y	x

Figure 1: Children's ordinal preferences.

In isolation, one would be inclined to choose x over y as it is preferred by four of the five children. However, we also learn that under x each of the first four children gets 1001 toys while under y each of them only gets 1000 toys. Moreover, we are told that the fifth child has a fatal illness which would be cured under y, whereas under x the fifth child dies a long and terrifying death. Unquestionably, this additional information would change our preference to y over x. Pearce therefore concludes that information besides individual preferences between x and y must be relevant for social choice. But which information exactly? Let us embed the comparison between x and y into a social choice problem in which for each of the first four children there exists an alternative z_i where only this child i dies and every other child gets 1001 toys. While the alternatives z_1 to z_4 aren't feasible, meaning they are not in the menu, they are still possible and could be in the menu in a different situation. Assume that the following graph depicts the von Neumann-Morgenstern preferences of the children over the possible alternatives.

The inclusion of alternatives z_1 to z_4 provides a context that puts the alternatives x and y into perspective. Adding this context, it becomes clear that going from y to x represents a marginal improvement for the first four children at the expense of a substantial loss to the fifth child. IIA demands that we ignore this additional information, which seems unreasonable. In order to include the context provided by

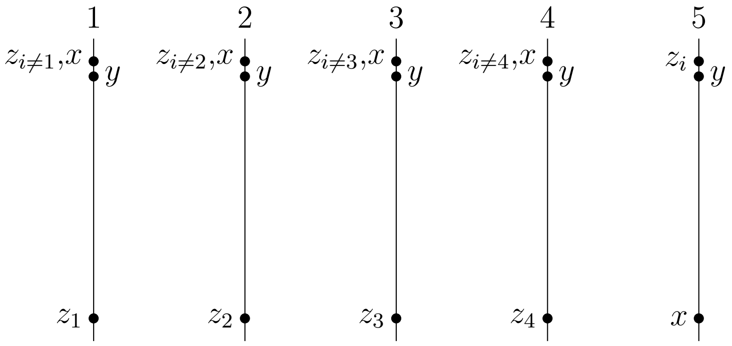

Figure 2: Children's cardinal preferences.

the infeasible alternatives, IIA has to be weakened.

Next, we asses MI. To do this, consider the following example by Sen (1993). An individual faces the menu $\{x, y\}$ where y means taking the last remaining apple from the fruit basket at the dinner table and x means taking nothing instead. Compare this to the situation where there is a second apple in the basket, such that the menu is $\{x, y, z\}$ where z means taking the other apple. One can plausibly prefer x over y under the first menu and y over x under the second. Similarly, our understanding of fairness in a social choice setting might be dependent on what is currently feasible. When comparing two social alternatives x and y, information about the feasibility of other alternatives provides a context that helps with the evaluation. In order to make use of that information, MI has to be weakened.

Above we argued that neither IIA nor MI is desirable, as they ignore additional information that can help in the assessment of social alternatives. Weakening either of these conditions opens the possibility for preference aggregation. We are looking for aggregation rules that abide by the von Neumann-Morgenstern axioms, satisfy the strong Pareto condition and are anonymous. These axioms strengthen Arrow's minimal requirements for rationality, benevolence and impartiality. We do not want to drop IIA and MI completely, as we wish to limit the information that can be used by the social planner. We therefore define two weaker axioms that identify when two social choice problems provide the same relevant information and thus yield the same social preferences. These axioms implicitly define the context as the context captures what is relevant. Our axiom *independence of irrelevant comparable alternatives* (IICA) weakens IIA by requiring that preferences over alternatives outside the menu do not influence the ranking if they are comparable for each individual to other alternatives. Specifically, an alternative is comparable to others if each individual is indifferent to some mixture of these other alternatives. *Menu independence of comparable alternatives* (MICA) weakens MI by allowing alternatives to be dropped from the menu if they are comparable to other alternatives in the menu. Each of these axioms, together with our other axioms, uniquely defines an aggregation rule that can be applied to any social choice or aggregation problem. Both of our representations are relative utilitarian, meaning the social welfare of an alternative is identified as the sum of individual von Neumann-Morgenstern utilities. When we weaken MI and maintain IIA, individual utilities are normalized relative to the menu. When we weaken IIA and maintain MI, individual utilities are normalized relative to the set of all possible alternatives. To normalize means to set the utility of the worst alternative in the respective set to 0 and of the utility of the best alternative to 1. In a later section, we then use our representations to quantify consumer welfare, fair bargaining and well-being of a dynamically inconsistent decision maker.

Relative utilitarianism has been axiomatized before by Karni (1998), Dhillon and Mertens (1999), Segal (2000), Borgers and Choo (2017), Marchant (2019), Sprumont (2019) and Brandl (2021). Most closely related is Dhillon and Mertens (1999). Dhillon and Mertens (1999) implicitly assume menu independence and characterize an aggregation rule that is the same as the one we obtain when relaxing IIA. They have a bottom-up approach, as their implicit objective is to present the weakest possible axioms that characterize relative utilitarianism. In contrast, we have a top-down approach, as we stay close to Arrow (1963) and wish to present the strongest possible axioms that don't result in an impossibility. These two different approaches lead to quite distinct axiomatic systems. Dhillon and Mertens (1999) has a higher level of technical sophistication, which is exemplified by their continuity axiom and intricate proof. In contrast, our axioms have a straightforward interpretation and our proof is more insightful, due to its simplicity. Note also that Dhillon and Mertens (1999) do not provide a representation when individuals have either identical or opposing preferences.

Another related paper is Sprumont (2019), which considers a setting where alternatives are acts (i.e. mappings from states to outcomes). The treatment is also different as axioms are formulated in terms of the set of outcomes and not in terms of sets of alternatives. Note also that the formulation of IIA therein has more the flavor of MI. In the representation, individual utilities are normalized relative to a set of alternatives the social planner deems relevant. Sprumont's axioms don't tell us what is relevant, as the relevant set is exogenous and part of the social choice problem. Without additional restrictions, this would allow the planner to justify nearly any decision, by handpicking the relevant set as needed. In our approach, the relevant set is determined endogenously from the axioms.

We mention a few of the many other papers on axiomatizations of aggregation rules that start similarly to us from Arrow (1950, 1963) and relax axioms therein. Notably, Sen (1993) drops MI and shows that impossibility still follows if IIA is replaced by *binary IIA*.[1] To our knowledge, Sen (1993) is the only previous paper that points out the implicit MI axiom in Arrow (1950, 1963). Saari (1998) and Maskin (2020) relax IIA by allowing the comparison between two alternatives to depend on how many other alternatives lie in-between them.

We proceed as follows. In Section 2 we present our framework, axioms and results. Section 3 provides a sketch of the proof. In Section 4 we apply our representation to several economic settings. Section 5 concludes.

2 Axiomatization

There is a society consisting of n individuals, where $n \in \mathbb{N}$. The set of individuals is denoted by $N := \{1, ..., n\}$. Furthermore, there is a set of possible alternatives A, where A is finite. In Appendix B we consider the case where A is infinite. Each individual in society has rational preferences over the possible alternatives. Formally, let $\triangle A$ denote the set of lotteries over A and let $\mathcal{R}$ denote the set of logically possible von Neumann-Morgenstern (vNM) preferences over $\triangle A$. Individual preferences are then captured by a preference profile $R \in \mathcal{R}^n$. For a given profile R, we sometimes

[1]Binary IIA says that the social preference between any two alternatives x and y can only depend on individual preferences between x and y. The literature often uses the notion of IIA and binary IIA interchangeably. Note however that binary IIA is only implied by IIA if MI is assumed as well.

write $\succsim_i^R$ to denote the i'th element in R. In any given situation, only a subset S of A is feasible and could be implemented for society. We call S the menu. A social planner is tasked with evaluating the alternatives in the menu, taking individual preferences into account. Formally, for any menu $S \subseteq A$ and preference profile $R \in \mathcal{R}^n$, $\succsim_*^{(S,R)}$ is a binary relation over $\triangle S$, describing the social planner's evaluation. We call (S, R) a *state* and denote by Ω the set of all states. One can think of the evaluation as a state dependent preference. We denote by $\succsim_*$ the evaluation function, which assigns to each state $(S, R) \in \Omega$ an evaluation $\succsim_*^{(S,R)}$.

We now impose axioms on how the social planner evaluates alternatives. *Rationality* says that the evaluation is rational in the sense of abiding by the vNM axioms.

Axiom RA (Rationality)**.** For each $(S, R) \in \Omega$, $\succsim_*^{(S,R)}$ satisfies the vNM axioms.

Rationality is normatively desirable and strengthens Arrow's requirement that the planner's evaluation is complete and transitive. Furthermore, we believe that, since individuals are assumed to be rational, an aggregation rule should preserve this characteristic of the individuals.

Our second axiom says that the social planner is benevolent, such that the evaluation respects the individuals' preferences whenever these are not in conflict.

Axiom SP (Strong Pareto)**.** For each $(S, R) \in \Omega$ and $x, y \in \triangle S$, if $x \succsim_i^R y$ for all $i \in N$ then $x \succsim_*^{(S,R)} y$ and if in addition $x \succ_i^R y$ for some $i \in N$ then $x \succ_*^{(S,R)} y$.

SP strengthens Arrow's Pareto condition.

Our third axiom is *anonymity*. Anonymity says that the planner's evaluation must not depend on the individual identities but only on the preferences themselves. Hence, in a counter-factual world, where preferences are interchanged across the individuals, the planner's evaluation must be the same.

Axiom AN (Anonymity)**.** For each $(S, R), (S, R') \in \Omega$, if R' is a permutation of R then $\succsim_*^{(S,R)} = \succsim_*^{(S,R')}$.

AN is an impartiality requirement and strengthens Arrow's non-dictatorship axiom.

Above we have stated our desiderata. Next, we will present the two conditions in Arrow (1950, 1963) that restrict what information can be used in the evaluation.

The first condition is the axiom of *independence of irrelevant alternatives*. We say that two preferences relations $\succcurlyeq$ and $\succcurlyeq'$ *agree* on some set of alternatives S if for any $x, y \in \triangle S$, $x \succcurlyeq y$ if and only if $x \succcurlyeq' y$. Furthermore, we say that two preference profiles $R, R' \in \mathcal{R}^n$ agree on S if for each $i \in N$, $\succcurlyeq_i^R$ and $\succcurlyeq_i^{R'}$ agree on S.

Axiom IIA (Independence of Irrelevant Alternatives)**.** Fix $(S, R) \in \Omega$. For any $R' \in \mathcal{R}^n$, if R and R' agree on S then $\succcurlyeq_*^{(S,R)} = \succcurlyeq_*^{(S,R')}$.

IIA says that the planner's evaluation of the menu cannot depend on individual preferences over alternatives outside the menu. Hence, in a counter-factual world, where individual preferences differ only on alternatives outside the menu, the planner's evaluation must be the same.

The second condition is that social preferences are menu independent. Arrow assumes this implicitly, as he writes that the social choice from a menu S is made based on a single preference relation over A, which is independent of the menu. Note that there are many equivalent ways of formalizing this condition. In anticipation of how we will relax this condition, we choose the following.

Axiom MI (Menu independence)**.** For each $(S, R) \in \Omega$ and $S' \subseteq S$, $\succcurlyeq_*^{(S,R)}$ and $\succcurlyeq_*^{(S',R)}$ agree on S'.

Menu independence says that removing alternatives from the menu does not change the planner's evaluation of the remaining alternatives.

It is well known that Arrow's axioms lead to an impossibility. Unsurprisingly, as our first three axioms strengthen Arrow's rationality, benevolence and impartiality requirements, the above axioms lead to an impossibility as well.[2]

Proposition 1. There is no evaluation function $\succcurlyeq_*$ that satisfies RA, SP, AN, IIA and MI.

As we have argued in the introduction, we believe that both IIA and MI force the social planner to ignore valuable context and should therefore be reconsidered. We will weaken each of these conditions, using the following notion of comparability.

[2]Note that formally, we weaken Arrow's *universal domain* condition, by assuming that individuals have vNM preferences. However, it has been shown that such a domain restriction is insufficient for escaping the impossibility. See Sen (1970), Kalai and Schmeidler (1977), Hylland (1980), Chichilnisky (1985) and Dhillon and Mertens (1997).

Definition 1. $a \in A$ is *comparable* relative to $B \subseteq A$ under $R \in \mathcal{R}^n$ if $a \notin B$ and for every $i \in N$ there exists $x_i \in \triangle B$ such that $x_i \sim_i^R [a]$.

An alternative is comparable to a set if for each individual there is a pay-off in the set equal to that of the alternative.

First, we weaken MI.

Axiom MICA (Menu Independence of Comparable Alternatives)**.** For each $(S, R) \in \Omega$ and $S' \subseteq S$ where every $a \in S \setminus S'$ is comparable relative to S', $\succcurlyeq_*^{(S,R)}$ and $\succcurlyeq_*^{(S',R)}$ agree on S'.

MICA says that removing comparable alternatives from the menu does not change the planner's evaluation of the remaining alternatives.

Weakening MI to MICA results in a representation we refer to as *menu contingent utilitarianism*. For any $R \in \mathcal{R}^n$ and $B \subseteq A$, let $u_i^{B,R}$ denote the representation of $\succcurlyeq_i^R$ where $\max_{a \in B} u_i^{B,R}(a) = 1$ and $\min_{a \in B} u_i^{B,R}(a) = 0$, unless $\succcurlyeq_i^R$ is indifferent on B in which case $u_i^{B,R}(a) = 0$ for all $a \in B$.[3] Furthermore, for any $B \subseteq A$ we denote by $|B|$ the number of elements in B.

Theorem 1 (Menu Contingent Utilitarianism)**.** Let $|A| \geq 2n + 4$. An evaluation function $\succcurlyeq_*$ satisfies RA, SP, AN, IIA and MICA if and only if for each $(S, R) \in \Omega$, $\succcurlyeq_*^{(S,R)}$ is represented by

$$\sum_{i \in N} u_i^{S,R}.$$

Next, we weaken IIA. For any $B \subset A$ we write B^c to denote $A \setminus B$.

Axiom IICA (Independence of Irrelevant Comparable Alternatives)**.** Fix $(S, R) \in \Omega$ and $C \subset S^c$ such that every $a \in C$ is comparable relative to C^c under R. For any $R' \in \mathcal{R}^n$, if R and R' agree on C^c and every $a \in C$ is comparable relative to C^c under R', then $\succcurlyeq_*^{(S,R')} = \succcurlyeq_*^{(S,R)}$.

IICA says that the planner's evaluation of the menu cannot depend on individual preferences over comparable alternatives outside the menu. Hence, in a counter-factual

[3]For any binary relation $\succcurlyeq$ on a set X, a utility function $u : X \to \mathbb{R}$ is said to represent $\succcurlyeq$ if for all $x, y \in X$, $u(x) \geq u(y)$ if and only if $x \succcurlyeq y$.

world, where individual preferences over these alternatives are different and these alternatives are still comparable, the planner's evaluation must be the same.

Weakening IIA to IICA results in a representation we refer to as *setting contingent utilitarianism*.

Theorem 2 (Setting Contingent Utilitarianism)**.** Let $|A| \geq 2n + 4$. An evaluation function $\succsim_*$ satisfies RA, SP, AN, IICA and MI if and only if for each $(S, R) \in \Omega$, $\succsim_*^{(S,R)}$ is represented by

$$\sum_{i \in N} u_i^{A,R}.$$

Our proofs of the theorems require a setting with many possible alternatives. However, as shown by the following proposition, we only need a few more alternatives than what is necessary for the axioms to be sufficient.

Proposition 2. Let $|A| < 2n + 1$. For both representation theorems, the axioms aren't sufficient to ensure that the planner's evaluation can be represented by the sum of normalized individual utilities across all states.

The proofs of the propositions and theorems in this section can be found in Appendix A.

Finally, we want to mention another desideratum that is satisfied by both representations, namely *neutrality*. We say that $\pi : A \mapsto A$ is a permutation of A if π is bijective and denote by Π the set of permutations of A. We abuse notation and define $\pi(S) := \{\pi(a) : a \in S\}$. We write $\pi(x) \in \triangle A$ to denote the lottery that for every $a \in A$ puts probability μ on $\pi(a)$ if and only if x puts probability μ on a. Finally, let $\pi(R) := (\succsim_i^{\pi})_{i \in N} \in \mathcal{R}^n$ denote the preference profile which has the same preferences on the permuted alternatives as R on the original alternatives. Formally, $\pi(x) \succsim_i^{\pi} \pi(y)$ if and only if $x \succsim_i y$ for all $i \in N$ and $x, y \in \triangle A$.

Axiom NE (Neutrality)**.** For each $(S, R) \in \Omega$, $x, y \in \triangle S$ and $\pi \in \Pi$, $x \succsim_*^{S,R} y$ if and only if $\pi(x) \succsim_*^{(\pi(S),\pi(R))} \pi(y)$.

Neutrality says that the labels of alternatives play no role. Hence, if the labels of a and b were interchanged and a was preferred by the planner before, then b must be preferred afterwards.

3 Sketch of Proof

To highlight some insights, this section provides a sketch of the proof. Since the proofs of both theorems are quite similar, we only sketch the proof of Theorem 1.

Before we prove the theorem, we derive two interim results. Note that SP implies *Pareto indifference* (PI), which says that if every individual in society is indifferent between two lotteries, then so is the social planner. The first interim result states that if both RA and PI are satisfied, then the utility function of the social planner can be expressed as a weighted sum of the individual utility functions. This result is well known and has first been postulated by Harsanyi (1955). However, Harsanyi's original proof contains a mistake, which lead to a variety of proofs by the subsequent literature. We in turn provide our own proof of Harsanyi's theorem, similar to those by Border (1985), Selinger (1986) and Hammond (1992), albeit ours is self contained as we do not refer to mathematical theorems. Our proof makes use of the *pay-off matrix*, which indicates for each individual the vNM utility of every alternative in the menu.[4] See Figure 3 for an example with three individuals and four alternatives. By RA,

$$\begin{array}{c c} & \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{array}{c} \\ i{=}1 \\ i{=}2 \\ i{=}3 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 \\ .5 & .5 & 0 & 1 \\ 1 & 0 & .7 & .3 \\ 0 & 1 & .2 & .8 \end{pmatrix} \end{array}$$

Figure 3: Example of a pay-off matrix.

the planner's evaluation can be represented by a row vector of vNM utilities as well. What needs to be shown is that whenever PI is satisfied, the planner's utility vector is equal to some linear combination of the rows in the pay-off matrix. If individual preferences are sufficiently diverse, such that the rows of the pay-off matrix span the entire vector space, then any logically possible vNM preference over the menu can be expressed by a linear combination of the individual utility functions. On the other hand, if individual preferences are not sufficiently diverse, we show that there exists a *dependent alternative*, who's column can be expressed by a linear combination

[4] A row of 1's is included in the pay-off matrix to allow for a constant in the representation.

of the other columns. For instance, in our example Column d is equal to a plus b minus c. We then sequentially drop dependent alternatives until the columns of the remaining alternatives are linearly independent. Then individual preferences over the remaining alternatives are sufficiently diverse such that any vNM preference over these alternatives can be expressed by a linear combination of rows. For each of the dependent alternatives we can identify two lotteries such that every individual is indifferent between them. In case of alternative d, these lotteries would be $\frac{1}{2}[a] + \frac{1}{2}[b]$ and $\frac{1}{2}[c] + \frac{1}{2}[d]$. By PI, also the planner must be indifferent between these lotteries and hence the planner's utility of the dependent alternative is pinned down by the same linear combination as for the individuals.

For our second interim result, we introduce the concept of *polar* alternatives. An alternative is polar if it is best among the menu for one individual and worst among the menu for everyone else. If the menu consists only of polar alternatives, we call such a state a *polar state*. Our second interim result then states that in a polar state, the planner is indifferent between all alternatives. Consider for instance a polar state that is described by the pay-off matrix depicted on the left-hand side of Figure 4. Note

$$
\begin{array}{c}
 \\ \\ i{=}1 \\ i{=}2 \\ i{=}3
\end{array}
\begin{array}{c}
\begin{array}{ccc} e & f & g \end{array} \\
\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{array}
\quad \xrightleftharpoons[\text{neutrality}]{\text{anonymity}} \quad
\begin{array}{c}
 \\ \\ i{=}1 \\ i{=}2 \\ i{=}3
\end{array}
\begin{array}{c}
\begin{array}{ccc} e & f & g \end{array} \\
\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{array}
$$

Figure 4: The polar state.

that the pay-off matrix doesn't describe the state completely, as it does not specify individual preferences over infeasible alternatives. However, by IIA, preferences over infeasible alternatives can be ignored. Assume the planner weakly prefers e to f. We begin by permuting the preferences of Individual's 1 and 2, which corresponds to swapping these individuals' rows in the pay-off matrix. By AN, the planner must weakly prefer e over f in the resulting state as well. The pay-off matrix of the resulting state is depicted on the right hand side of Figure 4. In the right-hand state we then interchange the labels of alternatives e and f, which corresponds to swapping the first two columns. If the planner satisfies NE, then if e was preferred over f before, then f must be preferred over e after their labels have been swapped. With the help of MICA we can show that indeed NE is satisfied in polar states. Note that swapping the first two columns in the right-hand pay-off matrix, brings us back to the initial state. Hence, we have shown that if e is weakly preferred to f, then f must also be weakly preferred to e, implying indifference.

Finally, we prove the theorem. Consider individual preferences over the menu $\{a, b, c, d\}$ as depicted in Figure 3 and assume that for each individual there is a comparable polar alternative outside the menu, denoted by e, f and g. In the menu $\{e, f, g\}$, the planner must be totally indifferent as we have shown previously. Now we add the original menu, resulting in Figure 5. Since we have only added comparable

$$\begin{array}{c c}
 & \begin{array}{ccccccc} a & b & c & d & e & f & g \end{array} \\
\begin{array}{c} \\ i{=}1 \\ i{=}2 \\ i{=}3 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ .5 & .5 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & .7 & .3 & 0 & 1 & 0 \\ 0 & 1 & .2 & .8 & 0 & 0 & 1 \end{pmatrix}
\end{array}$$

Figure 5: Adding the original menu to the polar state.

alternatives, by MICA the planner must still be indifferent between e, f and g. This indifference, together with our first interim result, then implies that equal weights on the individual utility functions represent the planner's evaluation. SP ensures that these common weights are positive and can be normalized to 1. Finally, we remove the polar alternatives and by MICA the same linear combination must still represent the planner's evaluation. By IIA, this must hold even if there are no comparable polar alternatives outside the original menu. Note that if there are less than n alternatives outside the original menu, the proof is more involved.

4 Applications

We apply our aggregation rules to three classic economic situations: aggregating the welfare of different consumers in a market, finding a fair solution to a bargaining problem and identifying the total welfare of a dynamically inconsistent decision maker. The literature has so far treated each of these problems in isolation. We provide a unifying framework, which aggregates individual preferences consistently across applications. In particular, we uncover policy recommendations contrary to those by the established welfare criteria.

4.1 Consumer Welfare

Consider a society with n individuals and two consumption goods m and g where m is the numéraire. A social alternative is an allocation of these goods to the individuals in society, hence an element of $\mathbb{R}_+^{2n}$. Individuals are self interested and their utility is quasi-linear, formally for each $i \in N$,

$$u_i((m_1, g_1), ..., (m_n, g_n)) = \alpha_i m_i + v_i(g_i) \tag{1}$$

for some $v_i : \mathbb{R}_+ \to \mathbb{R}$ and $\alpha_i \in \mathbb{R}_+$. The standard measure of aggregate welfare in this setting is *aggregate consumer surplus*

$$ACS((m_1, g_1), ..., (m_n, g_n)) = \sum_{i \in N} \frac{1}{\alpha_i} u_i((m_1, g_1), ..., (m_n, g_n)) = \sum_{i \in N} m_i + \sum_{i \in N} \frac{v_i(g_i)}{\alpha_i}.$$

Consider two allocations a and b such that some individuals strictly prefer a and some strictly prefer b. If $ACS(a)$ is higher than $ACS(b)$, then there exists monetary transfers after a, resulting in an allocation a', such that a' Pareto dominates b. This makes consumer surplus quite appealing. However, if these transfers aren't implemented, then ACS makes a judgment on how the well-being of different individuals should be traded off. We can show that the way in which ACS makes these trade-offs violates neutrality. Specifically, consumer surplus is sensitive to which of the two consumption goods is selected as the numéraire. Let $\mathcal{R}^{QL} \subset \mathcal{R}^n$ denote the set of preference profiles that can be represented by (1).

Proposition 3. There doesn't exist an evaluation function $\succcurlyeq$ on the restricted domain $\mathcal{R}^{QL}$ that satisfies NE.

Proof. It is sufficient to show that the axiom is violated in some state, so consider the state where for each $i \in N$, $v_i(g_i) = \beta_i g_i$ for some $\beta_i > 0$. Next consider the permutation of alternatives π such that $\pi((m_1, g_1), ..., (m_n, g_n)) = ((g_1, m_1), ..., (g_n, m_n))$

and note that $\pi(R) \in \mathcal{R}^{QL}$. For NE to be satisfied, it must hold that $ACS^R(a) \geq ACS^R(b)$ if and only if $ACS^{\pi(R)}(\pi(a)) \geq ACS^{\pi(R)}(\pi(b))$. We will show that for the following alternatives this is violated. Let $a = ((0,1),(0,0),...,(0,0))$ and $b = ((0,0),(2,0),...,(0,0))$. Then $ACS^R(a) = \frac{\beta_1}{\alpha_1}$, $ACS^R(b) = 2$, $ACS^{\pi(R)}(\pi(a)) = 1$ and $ACS^{\pi(R)}(\pi(b)) = \frac{\alpha_2}{\beta_2}$. Neutrality is violated for instance when $\beta_1 = 3$, $\alpha_1 = 1$, $\alpha_2 = 2$ and $\beta_2 = 1$. $\square$

Next consider a setting where a monopolist produces the good g at constant marginal cost c. We assume that individuals can be divided into two groups, low elasticity consumers L and high elasticity consumers H. Within each group every individual has the same preferences. The fraction of individuals belonging to Group $J \in \{L, H\}$ is denoted by γ_J. Total demand of Group J is then $D_J(p) := n\gamma_J(v_i')^{-1}(p\alpha_i)$ where $i \in J$. Assume the monopolist can identify which group an individual belongs to and therefore engage in third-degree price discrimination. Let p_J^* denote the monopoly price charged to Group J and let p^* denote the optimal price if price discrimination was prohibited. Assume

$$p_L^* > p^* > p_H^*.$$

A social planner has to decide whether to allow or prohibit price discrimination. Aggregate consumer surplus as a function of prices is given by

$$ACS(p_l, p_H) = CS_L(p_L) + CS_H(p_H)$$

where $CS_J(p) := \int_p^\infty D_J(p)dp$ and would recommend prohibition whenever $CS_L(p^*) + CS_H(p^*) > CS_L(p_L^*) + CS_H(p_H^*)$. As an alternative measure of aggregate welfare we propose setting contingent utilitarianism. We assume that the planner cannot directly implement an allocation, but instead can only affect the pricing scheme in the market, such that set of possible alternatives is the set of allocations that can arise from any price tuple in $[c, \infty)^2$. Then aggregate welfare under setting contingent utilitarianism is given by

$$u_*(p_l, p_H) = n\gamma_L \frac{1}{CS_L(c)} CS_L(p_L) + n\gamma_H \frac{1}{CS_H(c)} CS_H(p_H).$$

We find that for the text-book case of linear demand, price discrimination can be socially beneficial, even if low demand consumers would be served under uniform pricing.

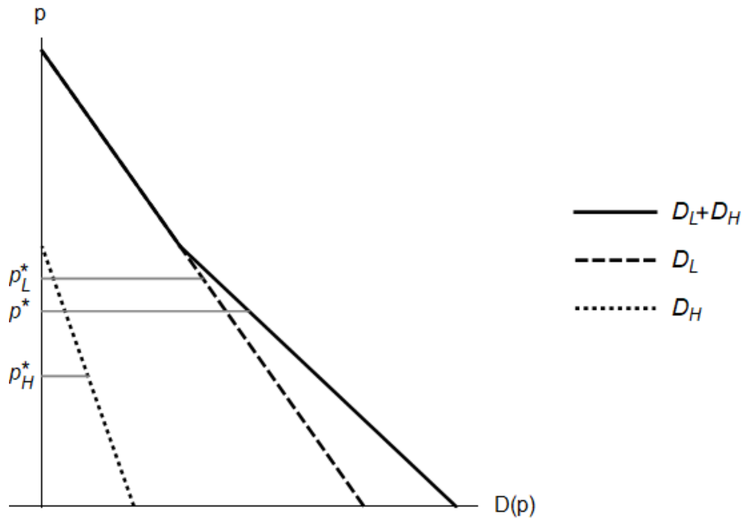

Figure 6: Linear demand and monopoly prices.

In contrast, aggregate consumer surplus is higher under price discrimination if and only if low demand consumers aren't served under the uniform price. Figure 6 shows the demand and monopoly prices for $D_H(p) = 2000 - 50p$, $D_L(p) = 7000 - 100p$ and $c = 0$. If the firm is allowed to price discriminate it will charge \$20 to high elasticity and \$35 to low elasticity consumers, which yields a consumer surplus of $CS_H(\$20) = \$10,000$ and $CS_L(\$20) = \$61,250$. If price discrimination was prohibited the firm would charge \$30 to both, which yields a consumer surplus of $CS_H(\$30) = \$2,500$ and $CS_L(\$30) = \$80,000$. In order for both groups to be better off under uniform pricing, low elasticity consumers would have to pay between $\$7,500$ and $\$18,750$ to high elasticity consumers. However, if these transfers are not implemented, and they usually aren't, then high elasticity consumers are worse off under uniform pricing. Going from price discrimination to uniform pricing, normalized utility of high elasticity consumers drops from 0.25 to roughly 0.06, while it increases for low elasticity consumers from 0.25 to roughly 0.33. If both groups consist of an equal number of individuals, then total utility is higher under price discrimination and is therefore favored by u_*. In a sense u_* penalizes low elasticity consumers through a lower weight on their surplus, in order to account for the missing transfers. Note that if transfers would in fact be implemented, then u_* would agree with ACS, since both satisfy SP. Furthermore, u_* accounts for how many individuals are positively and negatively affected by the policy, whereas ACS ignores the shares of individuals in each group. If for instance the share of consumers in Group H is below roughly 0.29 then u_* favors the prohibition of price discrimination, as sufficiently many individuals in Group L are positively affected by the policy. u_* values individuality and has a flavor similar to the majority principle. Finally, note that u_*, unlike ACS, can be meaningfully applied even when individual utility is not quasi-linear.

4.2 Bargaining

Consider a worker W and an employer E bargaining over the worker's wage. If no agreement can be reached, then both take their respective outside option, which we call the *disagreement point* and denote by d. For each $i \in \{E, W\}$ let $u_i : \mathbb{R}_+ \cup \{d\} \to \mathbb{R}$ denote the individual's utility over the possible alternatives. The set of feasible alternatives S is given by disagreement point and any wage that is weakly better than d for both individuals, formally $S = \{d\} \cup \{w \in \mathbb{R}_+ : u_i(w) \geq u_i(d) \text{ for } i \in \{W, E\}\}$. We assume that for both the worker and the employer there exists a wage in S that is strictly better than the disagreement point, which allows us to normalize u_i without loss of generality such that $u_i(d) = 0$ and $\max_{w \in S} u_i(w) = 1$.

In this bargaining situation, what is the fair wage? Nash (1950) proposes

$$Nash := \arg\max_{x \in \triangle S} \{u_W(x) u_E(x)\},$$

whereas Kalai and Smorodinsky (1975) recommend

$$KS := \arg\max_{x \in \triangle S} \{u_W(x)\} \text{ s.t. } u_W(x) = u_E(x).$$

Unlike our aggregation rules, which produce a complete ordering over all wages in S, the solutions offered by Nash (1950) and Kalai and Smorodinsky (1975) are choice rules, as they only identify the planner's most preferred alternative. To make the different frameworks comparable, we assume that the lotteries over wages selected by the choice rules are strictly preferred to all other lotteries by the planner. Given this assumption, we find that neither of these rules is rational.

Proposition 4. There doesn't exist an evaluation function $\succcurlyeq$ that is consistent with either $Nash$ or KS and that satisfies RA.

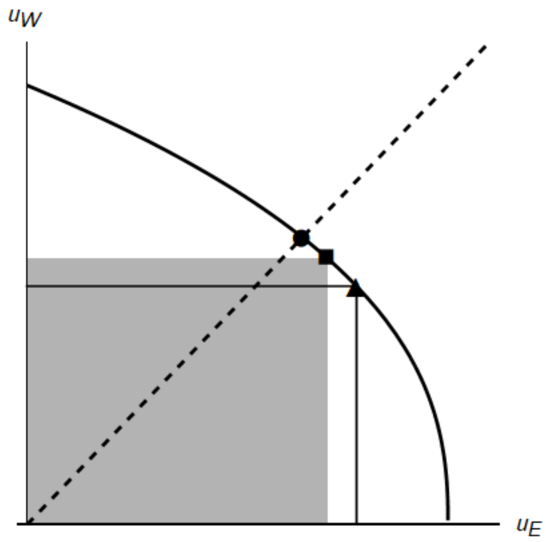

Figure 7: Comparing the different bargaining solutions.

Proof. It is sufficient to show that the axiom is violated in some state, so consider the state where both the worker and employer are risk neutral. Note that in this state $Nash$ and KS is the same, namely $\{x \in \triangle S : u_W(x) = u_E(x) = \frac{1}{2}\}$. Define $\underline{w}$, $\overline{w}$ such that $u_E(\overline{w}) = u_W(\underline{w}) = 0$ and $u_W(\overline{w}) = u_E(\underline{w}) = 1$. The lottery $\frac{1}{2}[\overline{w}] + \frac{1}{2}[\underline{w}]$ is part of the solution, but neither $\overline{w}$ nor $\underline{w}$ is. Hence, under both rules the planner strictly prefers a mixture over every pure alternative in the mixture, which violates RA. $\square$

We believe that a solution to the bargaining problem should be rational and offer menu contingent utilitarianism as an alternative. Since individual utilities are already normalized to $[0, 1]$ on the menu S, the socially optimal wage simply maximizes the sum of the workers and employers utility, formally

$$u_* := \arg\max_{x \in \triangle S} \{u_W(x) + u_E(x)\}.$$

Like $Nash$ and KS, u_* has an elegant geometric interpretation, which we demonstrate with the help of Figure 7. The curve shows the Pareto frontier of bargaining set for $u_E(w) = 1 - w$ and $u_W(w) = w^{\frac{2}{5}}$. KS (circle) is the intersection of the Pareto frontier and the dashed 45° line, $Nash$ (square) maximizes the area of the rectangle spanned by d and the Pareto frontier and u_* (triangle) maximizes the circumference of the rectangle spanned by d and the Pareto frontier. If the Pareto frontier is smooth, then the slope of the Pareto frontier at u_* is -1, as it would otherwise be possible to increase the utility of one individual by decreasing the utility of the other individual by a lesser

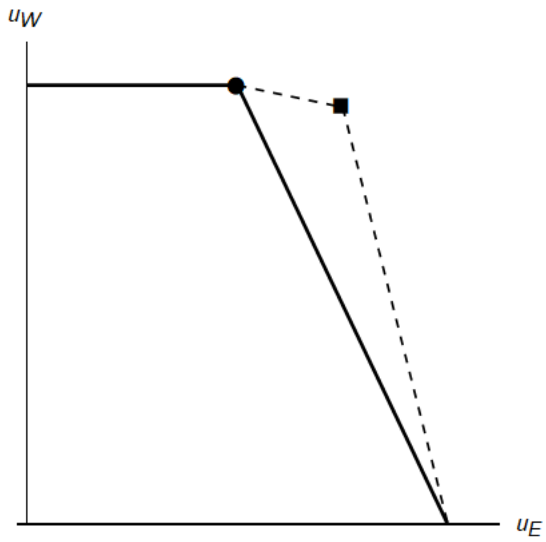

Figure 8: Violation of Monotonicity by u_*.

amount. Hence, in this example, u_* favors the employer relative to KS because the slope at KS is flatter than -1 and the employer can be made significantly better off at the cost of making the worker slightly worse off. If the slope at KS were steeper than -1, u_* would favor the worker relative to KS.

We now consider the framework of Nash (1950) and Kalai and Smorodinsky (1975) to see which of their axioms is violated by our bargaining solution. It is easy to see that u_* respects Nash's and Kalai & Smorodinsky's *Pareto optimality* and *invariance to affine transformations*. To a lesser extend u_* respects their *symmetry* axiom, as the symmetric outcome is optimal in a symmetric environment, but not uniquely optimal if the Pareto frontier is flat. Finally, u_* violates Nash's IIA axiom and Kalai & Smorodinsky's *monotonicity* axiom. First consider Nash's IIA axiom and note that despite the name, it is in fact a menu independence condition equivalent to MI. Hence, the violation is unsurprising and of no concern, since what is fair can depend on what is on the table. The violation of monotonicity is demonstrated by Figure 8. In the initial bargaining set, indicated by the solid line, u_* (circle) lies on the kink as the slope of the Pareto frontier is flatter than -1 to the left of the kink and steeper than -1 to the right. Increasing the bargaining set by the dashed line gives the worker a weakly larger utility for each level of u_E. Monotonicity would then require the solution to give weakly higher utility to the worker, which in this example would mean that the solution stays at the circle. However, the new bargaining set provides an opportunity for the social planner to increase the employer's utility substantially at the cost of a slight decrease in utility for the worker. Under menu contingent utilitarianism, the planner takes this opportunity, resulting in a new u_* (square).

4.3 Inter-temporal Welfare

Consider the consumption-savings problem of an inter-temporal decision maker (DM) with finite time horizon $n \geq 3$. For ease of exposition, we assume that there is no interest on savings and no per period income. Hence, the problem reduces to allocating the initial endowment, normalized to size 1, across time periods. The set of feasible consumption sequences is then $S = \{(c_1, ..., c_n) \in [0,1]^n : \sum_1^n c_t \leq 1\}$. As an example, one could think of a retired worker, who did not pay into a pension fund and who has to decide on how to consume the wealth that she saved up prior to retirement. As in Laibson (1997), we assume the DM discounts consumption quasi-hyperbolically, such that the utility in period t is given by

$$u_t(c_t, ..., c_n) = v(c_t) + \beta \sum_{i=t+1}^{n} \delta^{i-t} v(c_i) \tag{2}$$

with $\beta, \delta \in [0,1]$ and v monotonically increasing and bounded on $[0,1]$. If $\beta < 1$, the DM is dynamically inconsistent, meaning there exist consumption sequences $(c_1, ..., c_n), (c_1, ..., c_t, c'_{t+1}, ..., c'_n) \in S$ such that the DM in Period t strictly prefers $(c_1, ..., c_n)$ over $(c_1, ..., c_t, c'_{t+1}, ..., c'_n)$, while in Period $t+1$ her preference is reversed. It is as if the DM consists of different *selves*, Self 1 to Self n, with conflicting interests. If we want to asses which of the two sequences is better for the DM overall, we need a welfare criterion that incorporates the perspective of each self. Two criteria are common in the literature, the Pareto criterion and *long-run utility* (O'Donoghue and Rabin, 1999)

$$LR(c_1, ..., c_n) := \sum_{t=1}^{n} \delta^{t-1} v(c_i).$$

Before we discuss these criteria, let us first consider the case where $\beta = 1$, such that the DM is dynamically consistent. It is conventional wisdom among economists that in this case there is no conflict between the selves and the appropriate welfare measure is the utility of Self 1. This view however ignores the possibility that, even though there is no conflict looking forward, there could be a conflict looking backward.

For instance, if $\beta = 1$ and δ was close to 0, the DM would consume nearly all of the endowment in Period 1 and leave nearly nothing for future periods. At a future period, the DM could regret her earlier decision and prefer she had saved more in the past. So even though the DM is dynamically consistent, there is disagreement among the selves. To formalize this idea, let each self t have a vNM preference $\succcurlyeq_t$ over $\triangle S$. This means that Self t can compare sequences that differ in the consumption levels before Period t. Note that, since one cannot change the past, $\succcurlyeq_t$ is only partially revealed and the revealed part is represented by u_t. Now that all selves have preferences over the same domain, we can view these conflicting interested from the perspective of preference aggregation. This allows us to show that the common view, that for $\beta = 1$ total welfare is given Self's 1 utility, violates the Pareto criterion. Let $\mathcal{R}^{QH} \subset \mathcal{R}^n$ denote the set of preference profiles, where the revealed part can be represented by (2).

Proposition 5. There doesn't exist an evaluation function $\succcurlyeq$ on the restricted domain $\mathcal{R}^{QH}$ that (i) is represented by u_1 whenever $R \in \mathcal{R}^{QH}$ is represented by (2) with $\beta = 1$ and (ii) satisfies SP.

Proof. Note that it is sufficient to show that the axiom is violated in some state. Consider the state where each self is *past indifferent*, meaning for any $t \in \{2, ..., n\}$, $(c_1, ..., c_n), (c'_1, ..., c'_{t-1}, c_t, ..., c_n) \in S$,

$$(c_1, ..., c_n) \sim_t (c'_1, ..., c'_{t-1}, c_t, ..., c_n).$$

Now consider two sequences $s = (c_1, ..., c_n)$ and $s' = (c'_1, c'_2, c_3..., c_n)$ such that $c_1 > c'_1$ and $v(c_1) + \delta v(c_2) = v(c'_1) + \delta v(c'_2)$. Then $s \sim_1 s'$, $s' \succ_2 s$ and $s \sim_t s'$ for all $t \geq 3$ and hence by SP s' should give strictly higher total welfare than s. If however u_1 measure total welfare of the DM, then s and s' are equally desirable. $\square$

Let us now return to the aforementioned welfare criteria. The Pareto principle cannot compare every sequence in S, hence is incomplete and violates RA. Long-run utility is equal to u_1 for $\beta = 1$ and therefore, as shown by Proposition 5, violates SP. As an alternative to the established criteria, we propose menu contingent utilitarianism. In order to apply our criterion, we have to make assumptions on individual preferences over histories. Note that the same is true for the Pareto principle (see Goldman (1979)).

As a benchmark, we assume past indifference. Then preferences over entire sequences are represented by (2). Next we normalize each self's utility with respect to the best and worst alternative in S. Without loss of generality assume that $v(0) = 0$ and let $\bar{u}_t = \max_{s \in S} u_t(s)$ denote the optimal allocation from the perspective of Self t. Then total welfare of the DM according to menu contingent utilitarianism is given by

$$u_*(c_1, ..., c_n) := \sum_{t=1}^{n} \frac{1}{\bar{u}_t} u_t(c_t, ..., c_n) = \sum_{t=1}^{n} \underbrace{\left(\frac{1}{\bar{u}_t} + \beta \sum_{i=1}^{t-1} \frac{\delta^i}{\bar{u}_{t-i}} \right)}_{\gamma_t} v(c_t).$$

In contrast to long-run utility, the weights γ_t on the per period valuation are increasing in t. Hence, our criterion recommends an increasing consumption profile. This is because future consumption has a positive externality on earlier selves in the form of anticipatory utility, while later selves do not benefit from past consumption. Of course, this is driven by our assumption of past indifference. We leave it to future research to determine whether this assumption is plausible.

5 Conclusion

In economics we are used to modeling rational agents in the sense that these maximize expected utility. This is the canonical model that is applied independently of the setting. This motivates us to consider such individuals when aggregating preferences. It also motivates us to look for a social preference relation that is rational in this sense. Thereafter one can then treat the group of rational individuals as a single rational individual. In particular, it allows to evaluate uncertainty as a group in the traditional way by reducing uncertainty to risk and computing expected utility.

This paper implicitly launches a call to treat aggregated preferences as we treat individual preferences, and thus to apply a method that is motivated universally and not by the specific application. In particular, aggregation is more than just social choice. We present in detail examples of bargaining, consumer welfare and individual well being.

We pay tribute to Arrow (1963) by relaxing the axioms and assumptions therein as little as possible. Consequently we obtain two different aggregation rules, one that relaxes MI and one that relaxes IIA. The rule to be applied is determined by the aspects of the comparison. MI should be relaxed whenever outside options play a role as in bargaining. Relaxing IIA gives insights when the specific comparison has to be put into context. In fact, one objective of this paper is to formally introduce the notion of a context. In our understanding, there are no small or large changes in contexts. Instead, the context is what matters. To change the context means that something different can matter.

The key concept in our paper is that of comparable alternatives. This concept can also be found in the axiom of *independence of inessential expansions* in Sprumont (2013, 2019). However, unlike Sprumont we do not regard this axiom as "practical convenience" (Sprumont, 2013, p. 1021). We find that our formulation (equivalent to that of Sprumont (2013, 2019)) is as close as possible to MI. Alternatives that do not change the attainable payoffs of each of the individuals are irrelevant. Note that this allows each individual to have their own reason why the alternative dropped does not change the possible payoffs of this individual.

Our axioms lead to relative utilitarianism. The additivity comes from the requirement that social preferences are rational (Harsanyi, 1955). Here the linearity of von Neumann-Morgenstern preferences comes into play. Using the additivity of appropriately scaled preferences our comparability axioms allow to connect any social choice problem to one with a polar state. Anonymity and neutrality imply that all alternatives in a polar state are equally good in aggregation. This insight has implications on the original social choice problem, namely that normalized utilities have to be summed. The extreme payoffs that pin down the normalization is the part that is left over from the polar state.

Aggregating preferences when these differ among individuals implicitly requires that individual preferences can be compared. This is because the aggregate preference relation will implicitly contain a judgment about how the well being of one individual is traded off against that of another. Note that we do not make explicit assumptions on how preferences can be compared apart from assuming anonymity. It is our axioms that reveal that preferences can be compared once they are appropriately normalized.

We see our paper as identifying a method for applications. Thus we have included three. Many interesting questions arise, such as the role of the individual as opposed to the role of the group.

Appendix A

This section contains the proofs of Theorem 1, Theorem 2, Proposition 2 and some additional propositions and lemmas required for the proofs of the theorems.

First off, note that SP implies *Pareto indifference*.

Axiom PI (Pareto Indifference)**.** For any $(S,R) \in \Omega$ and $x, y \in \triangle S$, if $x \sim_i^R y$ for all $i \in N$ then $x \sim_*^{(S,R)} y$.

We use this axiom for our first interim result.

Proposition 6. Let RA and PI be satisfied. Then for any $(S,R) \in \Omega$ and for any representations $(u_i)_{i\in N}$ of R and u_* of $\succsim_*^{(S,R)}$ there exists $(\lambda_i)_{i=0}^n \in \mathbb{R}^{n+1}$ such that for all $a \in S$,

$$u_*(a) = \lambda_0 + \sum_{i\in N} \lambda_i u_i(a).$$

Proof. Fix $(S,R) \in \Omega$. We denote the alternatives in S by a_1 to a_m where $m = |S|$. The pay-off vector $\vec{u}_i := (u_i(a_1), ..., u_i(a_m))$ describes individual $i's$ vNM preferences over S. By RA, the representation u_* of $\succsim_*^{(S,R)}$ must be an expected utility representation and is therefore fully described by a pay-off vector as well. We denote this vector by $\vec{u}_* := (u_*(a_1), ..., u_*(a_m))$. Let

$$M := \begin{bmatrix} (1, ..., 1) \\ \vec{u}_1 \\ \vdots \\ \vec{u}_n \end{bmatrix}$$

be the *pay-off matrix*, which describes individual preferences over S. Note that the proposition states that $\vec{u}_*$ is equal to some linear combination of the rows in M. We distinguish two cases. If individual preferences are sufficiently diverse, then there are m linearly independent rows in M, which span the entire m-fold vector space. Hence, a linear combination of rows in M equal to $\vec{u}_*$ exists trivially, even without invoking PI. Next consider the case where individual preferences are not sufficiently diverse, such that not every vector of length m can be expressed as a linear combination of rows. In this case the maximal number of linearly independent rows is strictly below m. Hence, the maximal number of linearly independent columns must be strictly below m as well. As M has m columns, linearly dependent columns must exist. Sequentially drop linearly dependent columns from M until all remaining columns are linearly independent. Denote the set of alternatives associated with the remaining columns by B. We call the alternatives in B the *independent alternatives* and the alternatives in $S \setminus B$ the *dependent alternatives*. Note that in the resulting matrix, the rows span the entirety of the reduced vector space. Hence, by dropping the dependent alternatives, preferences are again sufficiently diverse to express any utility vector for the independent alternatives. This means that we can find a linear combination of rows of M that matches $\vec{u}_*$ in the utilities for the independent alternatives. Now consider the dependent alternatives. For each $a \in S \setminus B$ there must exist a linear combination of columns $\gamma^a : B \to \mathbb{R}$ such that

$$\sum_{b \in B} \gamma^a(b)\vec{u}(b) = \vec{u}(a). \tag{3}$$

We will now show that each γ^a can be decomposed into two lotteries $\gamma^a_+, \gamma^a_- \in \triangle S$ such that every individual is indifferent between these lotteries. Let $B^a_+ := \{b \in B : \gamma^a(b) \geq 0\}$ and $B^a_- := \{b \in B : \gamma^a(b) < 0\}$. Since by definition the first row of M consists only of 1's, $\sum_{b \in B} \gamma^a(b) = 1$ and furthermore $k := \sum_{b \in B^a_+} \gamma^a(b) = 1 - \sum_{b \in B^a_-} \gamma^a(b)$. Now define two lotteries

$$\gamma^a_+(b) := \begin{cases} \frac{1}{k}\gamma^a(b) & b \in B^a_+ \\ 0 & b \notin B^a_+ \end{cases} \qquad \text{and} \qquad \gamma^a_-(b) := \begin{cases} -\frac{1}{k}\gamma^a(b) & b \in B^a_- \\ \frac{1}{k} & b = a \\ 0 & b \notin B^a_- \cup \{a\} \end{cases}$$

Note that for each $i \in N$, $\sum_{b \in S} \gamma^a_+(b)u_i(b) = \sum_{b \in S} \gamma^a_-(b)u_i(b)$, meaning each individual is indifferent between the two lotteries. By PI, the planner must be indifferent as well. Therefore, $\sum_{b \in S} \gamma^a_+(b)u_*(b) = \sum_{b \in S} \gamma^a_-(b)u_*(b)$ and furthermore

$$\sum_{b \in B} \gamma^a(b)u_*(b) = u_*(a). \tag{4}$$

Compare this to Equation (3). The planner's utility for any dependent alternatives is determined from the independent alternatives in the same way as for every individual. Therefore, if a linear combination of rows of M matches the planner's utilities for the independent alternatives, it must also match the utilities for the dependent alternatives. $\square$

We say that a state $(S, R) \in \Omega$ is *polar* if (i) S has exactly n elements, which we denote by p_1 to p_n, (ii) for every $i \in N$, $p_j \sim_i^R p_k$ for all $j, k \in N \setminus \{i\}$ and (iii) either $p_i \succ_i^R p_j$ for all $j \in N \setminus \{i\}$ or $p_i \sim_i^R p_j$ for all $j \in N \setminus \{i\}$. We say that p_i is i's *polar alternative*.

Proposition 7. Let RA, SP, AN, IIA and MICA be satisfied. Then for any polar state $(S, R) \in \Omega$, if $i, j \in N$ have a strict preference on S then $p_i \sim_*^{(S,R)} p_j$.

Proof. Assume $(S, R) \in \Omega$ is polar and $i, j \in N$ have strict preferences on S. If $i = j$ the proposition is trivially satisfied so assume $i \neq j$. We assume $p_i \succcurlyeq_*^{(S,R)} p_j$ and show that $p_j \succcurlyeq_*^{(S,R)} p_i$ is implied, which only leaves $p_i \sim_*^{(S,R)} p_j$. We prove the proposition by going through a sequence of preference profiles and menus. We use roman numerals as subscripts to keep track of the different preference profiles and menus. Let $R_{\mathrm{I}} \in \mathcal{R}^n$ denote the permutation of R where only preferences of i and j are permuted. By AN, $p_i \succcurlyeq_*^{(S,R_{\mathrm{I}})} p_j$. Let $R_{\mathrm{II}} \in \mathcal{R}^n$ denote a preference profile which agrees with R_{I} on S and where there is $q_i, q_j \in S^c$ such that $q_i \sim_k^{R_{\mathrm{II}}} p_i$ and $q_j \sim_k^{R_{\mathrm{II}}} p_j$ for all $k \in N$. By our assumption that A has at least $2n + 4$ elements, such alternatives must exist. By IIA, $p_i \succcurlyeq_*^{(S,R_{\mathrm{II}})} p_j$. Let $S_{\mathrm{I}} := S \cup \{q_i, q_j\}$. Then $p_i \succcurlyeq_*^{(S_{\mathrm{I}},R_{\mathrm{II}})} p_j$ by MICA. Furthermore, $p_i \sim_*^{(S_{\mathrm{I}},R_{\mathrm{II}})} q_i$ and $p_j \sim_*^{(S_{\mathrm{I}},R_{\mathrm{II}})} q_j$ by SP and hence $q_i \succcurlyeq_*^{(S_{\mathrm{I}},R_{\mathrm{II}})} q_j$ by RA. Let $S_{\mathrm{II}} := S_{\mathrm{I}} \setminus \{p_i, p_j\}$. Then $q_i \succcurlyeq_*^{(S_{\mathrm{II}},R_{\mathrm{II}})} q_j$ by MICA. Let $R_{\mathrm{III}} \in \mathcal{R}^n$ denote a preference profile which agrees with R_{II} on S_{II} and where $q_i \sim_k^{R_{\mathrm{III}}} p_j$ and $q_j \sim_k^{R_{\mathrm{III}}} p_i$ for all $k \in N$. Then $q_i \succcurlyeq_*^{(S_{\mathrm{II}},R_{\mathrm{III}})} q_j$ by IIA, $q_i \succcurlyeq_*^{(S_{\mathrm{I}},R_{\mathrm{III}})} q_j$ by MICA and $p_j \succcurlyeq_*^{(S_{\mathrm{I}},R_{\mathrm{III}})} p_i$ by SP and RA. Furthermore, $p_j \succcurlyeq_*^{(S,R_{\mathrm{III}})} p_i$ by MICA. Note that R_{III} and R agree on S. Therefore, $p_j \succcurlyeq_*^{(S,R)} p_i$ by IIA, which implies $p_i \sim_*^{(S,R)} p_j$. $\square$

We now show two properties of binary relations that we will need for the proofs of the theorems.

Lemma 1. Let $\succcurlyeq$, $\succcurlyeq'$ and $\succcurlyeq''$ be binary relations over $\triangle S$. If $\succcurlyeq$ and $\succcurlyeq'$ agree on $B \subseteq S$ and $\succcurlyeq'$ and $\succcurlyeq''$ agree on $C \subseteq S$ then $\succcurlyeq$ and $\succcurlyeq''$ agree on $B \cap C$.

Proof. Consider any $B, C \subseteq S$ s.t. $B \cap C \neq \emptyset$ and any $x, y \in \triangle(B \cap C)$. If $\succcurlyeq$ and $\succcurlyeq'$ agree on B then $x \succcurlyeq y$ if and only if $x \succcurlyeq' y$ and if $\succcurlyeq'$ and $\succcurlyeq''$ agree on C then $x \succcurlyeq' y$

if and only if $x \succcurlyeq'' y$. Hence, for any $x, y \in \triangle(B \cap C)$, $x \succcurlyeq y$ if and only if $x \succcurlyeq'' y$, meaning $\succcurlyeq$ and $\succcurlyeq''$ agree on $B \cap C$. $\square$

Lemma 2. Let $\succcurlyeq$ be a binary relation over $\triangle S$ satisfying RA. For any $u : S \to \mathbb{R}$, if for each distinct $b, c, d \in S$ the part of $\succcurlyeq$ on $\triangle\{b, c, d\}$ is represented by u, then $\succcurlyeq$ is represented by u.

Proof. Let u satisfy the premise above. Select $b, c, d \in S$ such that $b \succ c$. If this is not possible, then $\succcurlyeq$ must be totally indifferent on S, in which case the proof is trivial. By assumption u represents $\succcurlyeq$ on $\triangle\{b, c, d\}$. Let $\hat{u} : S \to \mathbb{R}$ denote a representation of $\succcurlyeq$ where $\hat{u}(a) = u(a)$ for all $a \in \{b, c, d\}$. Now consider any $e \in S \setminus \{b, c, d\}$. By assumption, both u and $\hat{u}$ represent $\succcurlyeq$ on $\triangle\{b, c, e\}$. Since vNM representations are unique up to a positive affine transformation, there must exist $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}$ such that $u(a) = \alpha\hat{u}(a) + \beta$ for all $a \in \{b, c, e\}$. As $\hat{u}(b) = u(b)$ and $\hat{u}(c) = u(c)$, we find that $(1 - \alpha)u(b) = (1 - \alpha)u(c)$ which further implies $\alpha = 1$ and $\beta = 0$. Hence, $\hat{u}(e) = u(e)$. As this holds true for any $e \in S \setminus \{b, c, d\}$, we find that $\hat{u} = u$ and hence u represents $\succcurlyeq$. $\square$

Proof of Theorem 1

Let RA, SP, AN, IIA and MICA be satisfied and assume that $|A| \geq 2n + 4$. We will prove that for each $(S, R) \in \Omega$, $\succcurlyeq_*^{(S,R)}$ is represented by

$$\sum_{i \in N} u_i^{S,R}.$$

Fix $(S, R) \in \Omega$ where $|S| \geq 3$. Later we consider the case $|S| = 2$. We will show that for any distinct $b, c, d \in S$, the the part of $\succcurlyeq_*^{(S,R)}$ on $\triangle\{b, c, d\}$ is represented by $\sum_{i \in N} u_i^{S,R}$. It then follows from Lemma 2 that $\succcurlyeq_*^{(S,R)}$ is represented by $\sum_{i \in N} u_i^{S,R}$. The proof is done in three steps. First, we define a sequence of states, connecting (S, R) to a polar state. Second, we show that in the final state of that sequence, social preferences on $\triangle\{b, c, d\}$ are represented by $\sum_{i \in N} u_i^{S,R}$. Third, we show that social preferences in the final state agree with the social preferences in the initial state on $\{b, c, d\}$.

We begin by defining the aforementioned sequence of states. We use roman numerals as subscripts to keep track of the different preference profiles and menus. If $S \neq A$

then define $S_{\mathrm{I}} := S$. If on the other hand $S = A$, we construct S_{I} in the following way. $S \setminus \{b, c, d\}$ must have at least $2n + 1$ alternatives and hence at least one alternative $a \in S \setminus \{b, c, d\}$ must be comparable relative to $S \setminus \{a\}$ under R. In this case, let $S_{\mathrm{I}} := S \setminus \{a\}$. Either way, there is at least one alternative in S_{I}^c. Let $R_{\mathrm{I}} \in \mathcal{R}^n$ denote a preference profile that agrees with R on S_{I} and where there is an $e \in S_{\mathrm{I}}^c$ such that for each $i \in N$, $e \sim_i^{R_{\mathrm{I}}} a$ for all $a \in \{a \in S_{\mathrm{I}} : a \succcurlyeq_i^{R_{\mathrm{I}}} a' \text{ for all } a' \in S\}$. Let $S_{\mathrm{II}} := S_{\mathrm{I}} \cup \{e\}$. Next identify the smallest subset $S_{\mathrm{III}} \subseteq S_{\mathrm{II}}$ with the property that $\{b, c, d\} \subseteq S_{\mathrm{III}}$ and every $a \in S_{\mathrm{II}} \setminus S_{\mathrm{III}}$ is comparable relative to S_{III} under R_{I}. Note that there are at most $n + 4$ alternatives in S_{III}, namely b, c, d, e and n alternatives that are each worst for exactly one individual. Hence, there are at least n alternatives in S_{III}^c. For any state $\omega \in \Omega$, we denote the set of individuals that are not totally indifferent on the menu by $N_{\succ}^{\omega}$. Let $R_{\mathrm{II}} \in \mathcal{R}^n$ denote a preference profile that agrees with R_{I} on S_{III} and where there exists $P \subseteq S_{\mathrm{III}}^c$ such that (i) (P, R_{II}) is polar, where $p_i \in P$ denotes i's polar alternative, (ii) for each $i \in N$, $i \in N_{\succ}^{(P,R_{\mathrm{II}})}$ if and only if $i \in N_{\succ}^{(S_{\mathrm{III}},R_{\mathrm{II}})}$ and (iii) for each $i \in N$, $p_i \sim_i^{R_{\mathrm{II}}} a$ for all $a \in \{a \in S_{\mathrm{III}} : a \succcurlyeq_i^{R_{\mathrm{II}}} a' \text{ for all } a' \in S_{\mathrm{III}}\}$ and $p_j \sim_i^{R_{\mathrm{II}}} a$ for all $a \in \{a \in S_{\mathrm{III}} : a' \succcurlyeq_i^{R_{\mathrm{II}}} a \text{ for all } a' \in S_{\mathrm{III}}\}$ and $j \in N \setminus \{i\}$. This concludes the construction of the sequence of states.

Next, we show that the part of $\succcurlyeq_*^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})}$ on $\triangle\{b, c, d\}$ is represented by $\sum_{i \in N} u_i^{S,R}$. By Proposition 6, there exists weights $(\lambda_i)_{i=0}^n \in \mathbb{R}^{n+1}$ such that $\succcurlyeq_*^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})}$ is represented by

$$\lambda_0 + \sum_{i \in N} \lambda_i u_i^{P \cup S_{\mathrm{III}}, R_{\mathrm{II}}}.$$

By Proposition 7, $p_i \sim_*^{(P,R_{\mathrm{II}})} p_j$ for all $i, j \in N_{\succ}^{(P,R_{\mathrm{II}})}$. Because every alternative in S_{III} is comparable relative to P under R_{II}, by MICA $p_i \sim_*^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})} p_j$ for all $i, j \in N_{\succ}^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})}$ as well. This implies $\lambda_i = \lambda_j =: \lambda$ for all $i, j \in N_{\succ}^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})}$. Note that if $i \notin N_{\succ}^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})}$, then $u_i^{P \cup S_{\mathrm{III}}, R_{\mathrm{II}}}$ is constant on $P \cup S_{\mathrm{III}}$ and λ_i can be normalized to λ as the planner's vNM representation is unique up to positive affine transformations. Similarly, λ_0 can be normalized to 0. We now show that $\lambda > 0$ and therefore λ can be normalized to 1. We distinguish three cases. First, consider the case where $N_{\succ}^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})} = N$ and for any alternative $a \in S_{\mathrm{III}}$ there exists a lottery $x_a \in \triangle P$ such that $a \sim_i^{R_{\mathrm{II}}} x_a$ for all $i \in N$. Then $\succcurlyeq_*^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})}$ is totally indifferent on $P \cup S_{\mathrm{III}}$ and any λ represents the same preferences. Second, consider the case where $N_{\succ}^{(P \cup S_{\mathrm{III}}, R_{\mathrm{II}})} = N$

and for some alternative $a \in S$ such a lottery does not exist. Depending on whether $\sum_{i=1}^{n} u_i^{S_{\mathrm{III}},R_{\mathrm{II}}}(a)$ is greater or smaller than 1, one can construct a lottery over P that either Pareto dominates a or is Pareto dominated by a. Then for SP to be satisfied, λ has to be strictly positive. Third, if $N_{\succ}^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})} \neq N$ then for SP to be satisfied, $p_i \succ_*^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})} p_j$ for any $i \in N_{\succ}^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})}$ and $j \notin N_{\succ}^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})}$. This requires λ to be strictly positive as well. In any of the three cases, λ can be normalized to 1. Hence, we have shown that $\succcurlyeq_*^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})}$ is represented by

$$\sum_{i\in N} u_i^{P\cup S_{\mathrm{III}},R_{\mathrm{II}}}.$$

Finally, note that we have constructed $(P\cup S_{\mathrm{III}}, R_{\mathrm{II}})$ in such a way that for each $i \in N$, $u_i^{P\cup S_{\mathrm{III}},R_{\mathrm{II}}}(a) = u_i^{S,R}(a)$ for all $a \in \{b,c,d\}$. Hence, we have shown that the part of $\succcurlyeq_*^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})}$ on $\triangle\{b,c,d\}$ is indeed represented by $\sum_{i\in N} u_i^{S,R}$.

In the third and final step, we show that $\succcurlyeq_*^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})}$ and $\succcurlyeq_*^{(S,R)}$ agree on $\{b,c,d\}$.

$\succcurlyeq_*^{(S,R)}$ and $\succcurlyeq_*^{(S_{\mathrm{I}},R)}$ agree on $\{b,c,d\}$. (MICA)

$\succcurlyeq_*^{(S_{\mathrm{I}},R)}$ and $\succcurlyeq_*^{(S_{\mathrm{I}},R_{\mathrm{I}})}$ agree on $\{b,c,d\}$. (IIA)

$\succcurlyeq_*^{(S_{\mathrm{I}},R_{\mathrm{I}})}$ and $\succcurlyeq_*^{(S_{\mathrm{II}},R_{\mathrm{I}})}$ agree on $\{b,c,d\}$. (MICA)

$\succcurlyeq_*^{(S_{\mathrm{II}},R_{\mathrm{I}})}$ and $\succcurlyeq_*^{(S_{\mathrm{III}},R_{\mathrm{I}})}$ agree on $\{b,c,d\}$. (MICA)

$\succcurlyeq_*^{(S_{\mathrm{III}},R_{\mathrm{I}})}$ and $\succcurlyeq_*^{(S_{\mathrm{III}},R_{\mathrm{II}})}$ agree on $\{b,c,d\}$. (IIA)

$\succcurlyeq_*^{(S_{\mathrm{III}},R_{\mathrm{II}})}$ and $\succcurlyeq_*^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})}$ agree on $\{b,c,d\}$. (MICA)

Hence, by Lemma 1 $\succcurlyeq_*^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})}$ and $\succcurlyeq_*^{(S,R)}$ agree on $\{b,c,d\}$. This concludes the proof for $|S| \geq 3$.

Now consider $|S| = 2$. There must at least be $2n+2$ alternatives in S^c. We consider a different profile R_{I} where there is a set of polar alternatives P outside of S. Then $\succcurlyeq_*^{(P\cup S,R_{\mathrm{I}})}$ must be represented by equal weights by the above argument. By MICA, $\succcurlyeq_*^{(S,R_{\mathrm{I}})}$ must be represented by equal weights and by IIA $\succcurlyeq_*^{(S,R)}$ must be represented by equal weights. This concludes the proof of Theorem 1.

Next we derive an interim result required for the proof of Theorem 2.

Proposition 8. Let RA, SP, AN, MI and IICA be satisfied. Then for any polar $(S,R) \in \Omega$, if every $a \in S^c$ is comparable relative to S under R and $i,j \in N$ have a strict preference on S then $p_i \sim_*^{(S,R)} p_j$.

The proof is nearly identical to the proof of Proposition 7, with the only difference that IICA is used instead of IIA. This is possible as Proposition 8 is restricted to polar states where all alternatives outside are comparable.

Proof of Theorem 2

Let RA, SP, AN, IICA and MI be satisfied and assume that $|A| \geq 2n+4$. We show that for each $(S,R) \in \Omega$, $\succsim_*^{(S,R)}$ is represented by

$$\sum_{i \in N} u_i^{A,R}.$$

Note that it suffices to show that this holds for $S = A$, as by MI the same representation must hold for any $S \subset A$. Furthermore, note that for $S = A$, the representations of both theorems coincide. Hence, we follow the proof of Theorem 1, with the caveat that R_{II} is selected such that every alternative in $(P \cup S_{\mathrm{III}})^c$ is comparable relative to $(P \cup S_{\mathrm{III}})$ under R_{II}. Then one can simply replace the use of Proposition 7 with Proposition 8 and the use of IIA with IICA for the proof to follow through. This concludes the proof of Theorem 2.

Proof of Proposition 2

We prove the proposition by providing a counterexample. Specifically, we identify an evaluation function that satisfies the axioms but is not represented by the normalized sum of individual utilities across all states. We begin with a counterexample for the representation of Theorem 1. Assume that RA, SP, AN, IIA and MICA are satisfied. Fix a state $(A, \hat{R}) \in \Omega$ where no alternative is comparable relative to the remaining alternatives under $\hat{R}$. As $|A| < 2n+1$, such a state must exist. Let $\pi(\hat{R})$ denote the set containing all permutations of $\hat{R}$ and $\hat{R}$ itself and let $\hat{\Omega} := \{(S,R) \in \Omega : S = A, R \in \pi(\hat{R})\}$. Now consider an evaluation function where $\succsim_*^{(S,R)}$ is represented by

$$\sum_{i \in N} \left(\sum_{a \in S} u_i^{S,R}(a) \right) u_i^{S,R} \tag{5}$$

whenever $(S,R) \in \hat{\Omega}$ and by

$$\sum_{i \in N} u_i^{S,R} \tag{6}$$

whenever $(S,R) \notin \hat{\Omega}$. Note that it could be that for all $(S,R) \in \hat{\Omega}$, (5) and (6) are positive affine transformations of each other. So assume that $(A,\hat{R})$ has been selected such that this is not the case, which is possible by the richness of Ω. If this evaluation function indeed satisfies our axioms, we have produced a counterexample. Note that AN, MICA and IIA *connect* states, meaning they impose restrictions between states, whereas RA and SP impose restrictions on each state separately. So to prove that no axiom is violated, we will show that (i) no axiom connects a state in $\hat{\Omega}$ to a state outside of $\hat{\Omega}$, (ii) (5) satisfies the restrictions imposed between any two states in $\hat{\Omega}$, and (iii) (5) satisfies RA and SP for each $(S,R) \in \hat{\Omega}$. First, $\hat{\Omega}$ has been constructed such that AN doesn't connect any state in $\hat{\Omega}$ to a state outside of $\hat{\Omega}$. IIA doesn't connect any state in $\hat{\Omega}$ to another state, as there are no alternatives outside the menu. MICA doesn't connect any state in $\hat{\Omega}$ to another state, as no alternative is comparable relative to the other alternatives in A under any $R \in \pi(\hat{R})$. Second, (5) satisfies AN as the weight on each utility function only depends on the utility function itself but not on the index. Third, (5) assigns positive weights to all individual utility functions, unless an individual is indifferent on A, in which case the weight can be chosen arbitrarily. Hence, RA and SP are satisfied for each $(S,R) \in \hat{\Omega}$. This concludes the proof of Proposition 2 in case of Theorem 1.

Next we provide a counterexample for the representation of Theorem 2. Assume that RA, SP, AN, IICA and MI are satisfied. Fix a state $(A,\hat{R}) \in \Omega$ where no alternative is comparable relative to the remaining alternatives under $\hat{R}$. As $|A| < 2n+1$, such a state must exist. Let $\pi(\hat{R})$ denote the set containing all permutations of $\hat{R}$ and $\hat{R}$ itself and let $\hat{\Omega} := \{(S,R) \in \Omega : R \in \pi(\hat{R})\}$. Now consider an evaluation function where $\succsim_*^{(S,R)}$ is represented by

$$\sum_{i\in N}\left(\sum_{a\in S} u_i^{A,R}(a)\right) u_i^{A,R} \tag{7}$$

whenever $(S,R) \in \hat{\Omega}$ and by

$$\sum_{i\in N} u_i^{A,R} \tag{8}$$

whenever $(S,R) \notin \hat{\Omega}$. As before, assume that $(A,\hat{R})$ has been selected such that (7) and (8) are not positive affine transformations of each other. Note that both MI and IICA connect states. As before, we show that (i) no axiom connects a state in $\hat{\Omega}$ to a state outside of $\hat{\Omega}$, (ii) (7) satisfies the restrictions imposed between any two states in $\hat{\Omega}$, and (iii) (7) satisfies RA and SP for each $(S, R) \in \hat{\Omega}$. First, $\hat{\Omega}$ has been constructed such that both MI and AN doesn't connect any state in $\hat{\Omega}$ to a state outside of $\hat{\Omega}$. IICA doesn't connect any state in $\hat{\Omega}$ to another state, as no alternative is comparable relative to the other alternatives in A under any $R \in \pi(\hat{R})$. Second, (5) satisfies AN as the weight on each utility function only depends on the utility function itself but not on the index. (7) satisfies MI as the weights are independent of the menu. Third, (7) assigns positive weights to all individual utility functions, unless an individual is indifferent on A, in which case the weight can be chosen arbitrarily. Hence, RA and SP are satisfied for each $(S, R) \in \hat{\Omega}$. This concludes the proof of Proposition 2 in case of Theorem 2.

Appendix B

In this section we consider the case where A is either countably or uncountable infinite. This requires us to make some adjustments to the framework and the axioms. First, individual utilities might not be bounded, in which case they cannot be normalized as in the representations of Theorems 1 and 2. We deal with this by introducing a domain restriction, namely we only impose axioms on states where each individual has a best and worst alternative in A. Formally, we define Ω to be the set of states, such that for each $(S, R) \in \Omega$, both $\{a \in A : a \succcurlyeq_i^R b \text{ for all } b \in A\}$ and $\{a \in A : b \succcurlyeq_i^R a \text{ for all } b \in A\}$ are non-empty for all $i \in N$. Second, even if individual preferences have a best and worst alternative in A, they might not have one for $S \subset A$. For example, if $A = [0, 1]$ and $u_i^R(a) = a$ for some $R \in \mathcal{R}^n$ and $i \in N$, then $\succcurlyeq_i^R$ does not have a best alternative in $S = [0, 1)$. The non-existence of a best or worst alternative for an individual in a given menu means that we cannot construct the specific polar state required for the proof of Theorem 1. Therein, every alternative in the menu must be comparable relative to the set of polar alternatives and every polar alternative must be comparable relative to the menu. For the proof to go through, we introduce a weaker notation of comparability.

Definition 2. $a \in A$ is *approximately comparable* relative to $B \subseteq A$ under $R \in \mathcal{R}^n$ if

$a \notin B$ and for every $i \in N$ and $\varepsilon \in (0,1)$ there exists $x_{i,\varepsilon}, y_{i,\varepsilon} \in \triangle B$ such that

$$(1-\varepsilon)[a] + \varepsilon x_{i,\varepsilon} \sim_i^R y_{i,\varepsilon}.$$

Note that if a is comparable relative to B under R, then a is approximately comparable relative to B under R. If A is finite, the two concepts coincide. Furthermore, if a is approximately comparable, its utility must lie between the supremum and infimum utility in the set for every individual, as shown by the following lemma. We will later use this result for the representation theorem.

Lemma 3. For any $R \in \mathcal{R}^n$ and $B \subseteq A$, if $a \notin B$ and for any utility profile $(u_i^R)_{i \in N}$ of R,

$$\sup_{b \in B} u_i^R(b) \geq u_i^R(a) \geq \inf_{b \in B} u_i^R(b)$$

for all $i \in N$ then a is approximately comparable relative to B under R.

Proof. Fix $(u_i^R)_{i \in N}$. Note that $a \in A$ is approximately comparable relative to $B \subseteq A$ under $R \in \mathcal{R}^n$ if and only if for every $i \in N$ and $\varepsilon \in (0,1)$ there exists $x_{i,\varepsilon}, y_{i,\varepsilon} \in \triangle B$ such that

$$(1-\varepsilon) u_i^R(a) + \varepsilon u_i^R(x_{i,\varepsilon}) = u_i^R(y_{i,\varepsilon}). \tag{9}$$

If $u_i^R(a)$ is strictly between $\sup_{b \in B} u_i^R(b)$ and $\inf_{b \in B} u_i^R(b)$ one can simply select a $z_i \in \triangle B$ such that $u_i^R(z_i) = u_i^R(a)$ and then set $x_{i,\varepsilon} = y_{i,\varepsilon} = z_i$. Then (9) is satisfied for all ε. So assume $u_i^R(a) = \sup_{b \in B} u_i^R(b)$ and fix ε. Choose $x_{i,\varepsilon}$ arbitrarily. The left hand side of (9) is strictly between $\sup_{b \in B} u_i^R(b)$ and $\inf_{b \in B} u_i^R(b)$ and hence there must be a $y_{i,\varepsilon} \in \triangle B$ to satisfy (9). The same argument applies when $u_i^R(a) = \inf_{b \in B} u_i^R(b)$. $\square$

In MICA, comparability is then replaced by approximate comparability.

Axiom MICA*. For each $(S,R) \in \Omega$ and $S' \subseteq S$ where every $a \in S \setminus S'$ is approximately comparable relative to S', $\succsim_*^{(S,R)}$ and $\succsim_*^{(S',R)}$ agree on S'.

With these adjustments, we can now state the equivalent of Theorem 1 when A is infinite. For any $R \in \mathcal{R}^n$ and $B \subseteq A$, let $\hat{u}_i^{B,R}$ denote the representation of $\succsim_i^R$ where $\sup_{a \in B} \hat{u}_i^{B,R}(a) = 1$ and $\inf_{a \in B} u_i^{B,R}(a) = 0$, unless $\succsim_i^R$ is indifferent on B in which case $\hat{u}_i^{B,R}(a) = 0$ for all $a \in B$.

Theorem 3. Let A be infinite. An evaluation function $\succsim_*$ satisfies RA, SP, AN, IIA and MICA* if and only if for each $(S, R) \in \Omega$, $\succsim_*^{(S,R)}$ is represented by

$$\sum_{i \in N} \hat{u}_i^{S,R}.$$

Proof. First off, note that the proofs of Proposition 7, Lemma 1 and Lemma 2 go through when there are infinitely many possible alternatives. For a proof of Proposition 6 under infinite A we refer to Mandler (2005). Now consider the proof of Theorem 1. When constructing the sequence of states, specifically R_{I} and R_{II}, there might not be a best or worst alternative in the menu for some individuals. We make the following adjustments to the proof. Let $R_{\mathrm{I}} \in \mathcal{R}^n$ denote a preference profile that agrees with R on S_{I} and where there is an $e \in S_{\mathrm{I}}^c$ such that for each $i \in N$ and any $u_i^{R_{\mathrm{I}}}$ representing $\succsim_i^{R_{\mathrm{I}}}$, $u_i^{R_{\mathrm{I}}}(e) = \sup_{a \in S_{\mathrm{I}}} u_i^{R_{\mathrm{I}}}(a)$. Let $R_{\mathrm{II}} \in \mathcal{R}^n$ denote a preference profile that agrees with R_{I} on S_{III} and where there exists $P \subseteq S_{\mathrm{III}}^c$ such that (i) (P, R_{II}) is polar, where $p_i \in P$ denotes i's polar alternative, (ii) for each $i \in N$, $i \in N_{\succ}^{(P,R_{\mathrm{II}})}$ if and only if $i \in N_{\succ}^{(S_{\mathrm{III}},R_{\mathrm{II}})}$ and (iii) for each $i \in N$ and any $u_i^{R_{\mathrm{II}}}$ representing $\succsim_i^{R_{\mathrm{II}}}$, $u_i^{R_{\mathrm{II}}}(p_i) = \sup_{a \in S_{\mathrm{III}}} u_i^{R_{\mathrm{II}}}(a)$ and $u_i^{R_{\mathrm{II}}}(p_j) = \inf_{a \in S_{\mathrm{III}}} u_i^{R_{\mathrm{II}}}(a)$ for all $j \in N \setminus \{i\}$. By Lemma 3, e is approximately comparable to S_{I} under R_{I} and every alternative in P is approximately comparable relative to S_{III} under R_{II}. Then the proof of Theorem 1 goes through. $\square$

Note that Theorem 2 goes through without adjustment to the axioms.

References

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346.

Arrow, K. J. (1963). *Social Choice and Individual Values*, volume 12. Yale University Press.

Border, K. C. (1985). More on harsanyi's utilitarian cardinal welfare theorem. *Social Choice and Welfare*, 1(4):279–281.

Borgers, T. and Choo, Y. M. (2017). Revealed relative utilitarianism. CESIfo Working Paper Nr. 6613.

Brandl, F. (2021). Belief-averaging and relative utilitarianism. *Journal of Economic Theory*, 198:105368.

Chichilnisky, G. (1985). Von neumann-morgenstern utilities and cardinal preferences. *Mathematics of Operations Research*, 10(4):633–641.

Dhillon, A. and Mertens, J.-F. (1997). An impossibility theorem with von neumann-morgenstern preferences. *Economics Letters*, 56(3):305–309.

Dhillon, A. and Mertens, J.-F. (1999). Relative utilitarianism. *Econometrica: Journal of the Econometric Society*, 67(3):471–498.

Goldman, S. M. (1979). Intertemporally inconsistent preferences and the rate of consumption. *Econometrica: Journal of the Econometric Society*, 47(3):621–626.

Hammond, P. J. (1992). Harsanyi's utilitarian theorem: A simpler proof and some ethical connotations. In *Rational Interaction*, pages 305–319. Springer.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63(4):309–321.

Hylland, A. (1980). Aggregation procedure for cardinal preferences: a comment. *Econometrica: Journal of the Econometric Society*, 48(2):539–542.

Kalai, E. and Schmeidler, D. (1977). Aggregation procedure for cardinal preferences: A formulation and proof of samuelson's impossibility conjecture. *Econometrica: Journal of the Econometric Society*, 45(6):1431–1438.

Kalai, E. and Smorodinsky, M. (1975). Other solutions to nash's bargaining problem. *Econometrica: Journal of the Econometric Society*, 43(3):513–518.

Karni, E. (1998). Impartiality: definition and representation. *Econometrica: Journal of the Econometric Society*, 66(6):1405–1415.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–478.

Mandler, M. (2005). Harsanyi's utilitarianism via linear programming. *Economics Letters*, 88(1):85–90.

Marchant, T. (2019). Utilitarianism without individual utilities. *Social Choice and Welfare*, 53(1):1–19.

Maskin, E. (2020). Arrow's theorem, may's axioms, and borda's rule. https://scholar.harvard.edu/maskin/publications/arrows-theorem-mays-axiom-bordas-rule. Accessed on 2023-01-16.

Nash, J. F. (1950). The bargaining problem. *Econometrica: Journal of the Econometric Society*, 18(2):155–162.

O'Donoghue, T. and Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89(1):103–124.

Pearce, D. (2021). Individual and social welfare: a bayesian perspective. In *Frisch Lecture delivered to the World Congress of the Econometric Society.*

Saari, D. G. (1998). Connecting and resolving sen's and arrow's theorems. *Social Choice and Welfare*, 15(2):239–261.

Segal, U. (2000). Let's agree that all dictatorships are equally bad. *Journal of Political Economy*, 108(3):569–589.

Selinger, S. (1986). Harsanyi's aggregation theorem without selfish preferences. *Theory and decision*, 20(1):53—-62.

Sen, A. (1970). Interpersonal aggregation and partial comparability. *Econometrica: Journal of the Econometric Society*, 38(3):393–409.

Sen, A. (1993). Internal consistency of choice. *Econometrica: Journal of the Econometric Society*, 61(3):495–521.

Sprumont, Y. (2013). On relative egalitarianism. *Social Choice and Welfare*, 40(4):1015–1032.

Sprumont, Y. (2019). Relative utilitarianism under uncertainty. *Social Choice and Welfare*, 53(4):621–639.

von Neumann, J. and Morgenstern, O. (1944). *Theory of games and economic behavior.* Princeton University Press.