# Putting Context into Preference Aggregation*

Philipp Peitler‡ Karl H. Schlag§

April 30, 2024

[Latest version here]

**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, the intertemporal well-being of an individual, and bargaining.

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

*We thank Hervé Moulin and Yves Sprumont for helpful comments. We gratefully acknowledge financial support from the *Jubiläumsfonds* of the Austrian National Bank (OeNB), project number 18719.

‡University of Vienna; philipp.peitler@univie.ac.at

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

# 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 planner needs a rule that aggregates individual preferences into a single social preference. Different strands of the economic literature use very different rules. In the analysis of markets, consumer surplus is used to evaluate the joint welfare of consumers. In the analysis of dynamically inconsistent decision makers, intertemporal welfare is evaluated either by the Pareto criterion or by the long-run utility (O'Donoghue & Rabin, 1999). In bargaining, the Nash bargaining solution (Nash, 1950) and the Kalai-Smorodinsky bargaining solution (Kalai & Smorodinsky, 1975) are prominent. Note here that bargaining problems can be viewed as the search for the best allocation after aggregating preferences. 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. The Pareto criterion does not generate complete preferences, and long-run utility violates the Pareto principle. Neither the Nash nor the Kalai-Smorodinsky bargaining solution can be interpreted as resulting from the most preferred allocation under a social preference relation that satisfies the von Neumann and Morgenstern (1944) axioms. Some of the above violations are not obvious and will be demonstrated. In each of these settings, we could search separately for other rules. However, we feel that the well-being 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 all settings. 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: *menu independence* (MI). Together, these axioms make it impossible to aggregate preferences. Weakening either IIA or MI is a potential avenue 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 assess 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 in Figure 1. In

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| x | x | x | x | y |
| y | y | y | y | x |

**Figure 1:** Children's ordinal preferences.

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 that 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 currently feasible (i.e., they are not in the menu), they are still possible (in the sense of conceivable) and could be in the menu in a different situation. Assume that Figure 2 depicts the von Neumann-Morgenstern preferences of the children over all these alternatives. The inclusion of alternatives $z_1$ to $z_4$ provides a context that

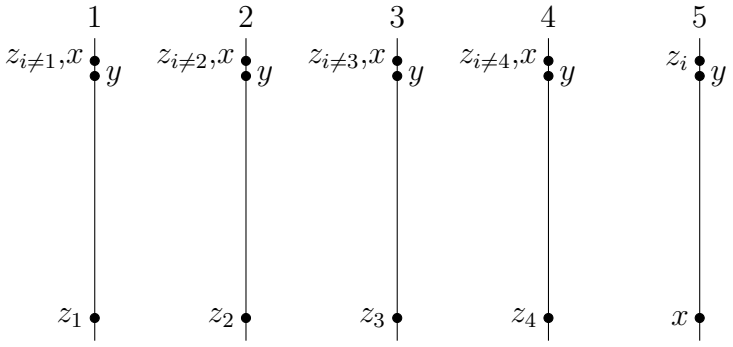


**Figure 2:** Children's cardinal utility.

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 the infeasible alternatives, IIA has to be weakened.

Next, we assess 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 depend 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. *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 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 within a given setting. Normalizing means setting the utility of the worst alternative in the respective set to 0 and the utility of the best alternative to 1. In a later section, we then use our representations to quantify consumer welfare and the well-being of a dynamically inconsistent decision maker and to derive a fair bargaining solution.

## 1.1 Literature

In the tradition of Arrow (1950, 1963), we axiomatize relative utilitarianism in a *multi-profile* setting, meaning that at least one axiom relates social preferences across different profiles of individual preference (e.g., IIA or IICA). Previous characterizations of relative utilitarianism in this setting are Dhillon and Mertens (1999), Börgers and Choo (2017) and Marchant (2019). Dhillon and Mertens (1999) implicitly assumes menu independence and characterizes 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 axioms that don't result in an impossibility. These two different

approaches result in 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 much simpler. Note also that Dhillon and Mertens (1999) does not provide a representation when some individuals have either identical or opposing preferences. Börgers and Choo (2017) takes a revealed preference approach and imposes axioms on the planner's revealed marginal rate of substitution between individuals' normalized utilities. Marchant (2019) axiomatizes a relative utilitarian choice correspondence rather than a social preference.

In a *single-profile* setting, meaning without axioms relating different profiles, relative utilitarianism has been axiomatized by Karni (1998), Segal (2000), and Karni and Weymark (2024).

We build on a large body of literature that deals with the aggregation of vNM preferences. The seminal paper is Harsanyi (1955), which shows that vNM rationality of society and a Pareto axiom implies linear aggregation of utilities. Many extensions and variations of this theorem followed (Domotor, 1979; Fishburn, 1984; Border, 1985; Selinger, 1986; Coulhon & Mongin, 1989; Hammond, 1992; Weymark, 1993; Mongin, 1994; Weymark, 1994; De Meyer & Mongin, 1995; Zhou, 1997; Blackorby et al., 1999; Mandler, 2005). While Harsanyi (1955) shows that social welfare must be a weighted sum of individual utilities, he does not identify the weights. We start from his theorem and impose additional axioms to pin down these weights. Another paper on the aggregation of vNM preferences worthy of mention is Sprumont (2013), which drops the assumption that social preferences are vNM and characterizes the relative egalitarian aggregation rule. Our axiom MICA is quite close to the "independence of inessential expansions" axiom in Sprumont (2013).

Closely related to the aggregation of risk preferences is the aggregation of subjective expected utility (SEU) preferences over uncertain acts. Since beliefs regarding the state can differ across individuals, not only conflicting tastes but also conflicting probability judgments have to be aggregated. This additional complication leads to an impossibility result, even in the single-profile setting. Mongin (1995, 1998) shows that society's preferences cannot simultaneously satisfy the Pareto condition and have

a SEU representation. To escape this impossibility, utilitarian aggregation rules either weaken the rationality postulates underlying the SEU representation (Sprumont, 2019) or the Pareto condition (Gilboa et al., 2004; Brandl, 2021; Dietrich, 2021). Most closely related is Sprumont (2019), which allows the social preferences to depend on some reference set of relevant outcomes. This is mathematically similar to our approach, where social preferences can depend on the menu. However, 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. One could not simply reinterpret the reference set as the menu, as it is a set of outcomes, not a set of acts.

Related to relative utilitarian aggregation rules is the literature on *range voting* (Smith, 2000; Gaertner & Xu, 2012; Pivato, 2014; Macé, 2018). Under range voting, a voter assigns a score between 0 and 1 to each candidate, and the candidate with the highest sum of scores is selected as the winner.

We also 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 (2022) relax IIA by allowing the comparison between two alternatives to depend on how many other alternatives lie in between them.

## 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.[2] Each individual in society has rational preferences over the

[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.

[2] In Appendix B we consider the case where $A$ is infinite.

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 write $\succcurlyeq_i^R$ to denote the $i$'th element in $R$. In any given situation, only a subset $S$ of the alternatives in $A$ is feasible and could be implemented by society. We call $S$ the *menu*. For any $S \subseteq A$ and $R \in \mathcal{R}^n$, we call $(S, R)$ the society's *state*, and we denote by $\Omega$ the set of all logically possible states. An *aggregation rule* $\succcurlyeq_*$ assigns to each state $(S, R) \in \Omega$ a binary relation $\succcurlyeq_*^{(S,R)}$ over $\triangle S$. For exposition purposes we act as if there was a social planner that employs an aggregation rule to evaluate the alternatives in the menu. Hence, for a given state $(S, R)$ we refer to $\succcurlyeq_*^{(S,R)}$ as the planner's *evaluation*.

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 $(S, R) \in \Omega$, $\succcurlyeq_*^{(S,R)}$ 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 $x \succcurlyeq_i^R y$ for all $i \in N$, then $x \succcurlyeq_*^{(S,R)} y$. 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 counterfactual 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 $\succcurlyeq_*^{(S,R)} = \succcurlyeq_*^{(S,R')}$.

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 *independence of irrelevant alternatives*. We say that two binary 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 counterfactual world, where individual preferences differ only on alternatives outside the menu, the planner's evaluation must be the same.

Arrow's 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. In light of the numerous ways to formalize this notion, we have selected the following axiom, as it best aligns with our subsequent relaxation of the condition.

**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.[3]

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

[3]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).

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. Let $[a]$ denote the lottery that assigns probability 1 to the alternative $a \in A$.

**Definition 1.** $a \in A$ is *comparable* relative to $B \subseteq A$ under $R \in \mathcal{R}^n$ if (i) $a \notin B$ and (ii) 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 of the planner's evaluation 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 $a \in B$.[4] 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 aggregation rule $\succcurlyeq_*$ satisfies RA, SP, AN, IIA and MICA if and only if for each $(S, R) \in \Omega$, $\succcurlyeq_*^{(S,R)}$ is represented by

$$\sum_{i \in N} u_i^{S,R}. \tag{1}$$

We sketch the proof of Theorem 1 in Section 3. A complete proof can be found in Appendix A. Note that no axiom of Theorem 1 is implied by the other axioms of the theorem. Hence, a subset of the axioms would not suffice for the representation. We prove this in Appendix C.

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

[4]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 $x \succcurlyeq y$.

**Axiom IICA** (Independence of Irrelevant Comparable Alternatives)**.** Fix $(S, R) \in \Omega$ and $C \subseteq S^c$ 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 counterfactual world, where individual preferences over these alternatives are different in a way such that 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 aggregation rule $\succcurlyeq_*$ satisfies RA, SP, AN, IICA and MI if and only if for each $(S, R) \in \Omega$, $\succcurlyeq_*^{(S,R)}$ is represented by

$$\sum_{i \in N} u_i^{A,R}. \tag{2}$$

We prove Theorem 2 in Appendix A. No axiom of Theorem 2 is implied by the other axioms of the theorem, which we show in Appendix C.

In both theorems, we assume that there are at least $2n+4$ possible alternatives. We make this assumption, as it allows us to employ simple and insightful proofs. However, as shown by the following proposition, our proofs require only a few more alternatives than what is necessary for the axioms to be sufficient.

**Proposition 2.** Let $|A| < 2n + 1$. There exists an aggregation rule $\succcurlyeq_*$ which satisfies RA, SP, AN, IIA and MICA and which is not represented by (1). Similarly, there exists an aggregation rule $\succcurlyeq_*$ which satisfies RA, SP, AN, IICA and MI and which is not represented by (2).

We prove Proposition 2 in Appendix A.

Finally, we want to mention *neutrality*, an additional desideratum that is satisfied by both our representations. Neutrality says that the labels of alternatives play no role in the planner's evaluation. Hence, if the labels of $a$ and $b$ were interchanged and $a$ was preferred by the planner before, then $b$ must be preferred afterward. We believe this to be the most important impartiality requirement besides anonymity. If neutrality

wasn't implied by our other axioms, we would have imposed it directly. Since we will refer to neutrality in the upcoming sections, a formal definition is in order. We say that $\pi : A \mapsto A$ is a permutation of $A$ if $\pi$ is bijective and denote by $\Pi$ the set of permutations of $A$. We abuse notation and define $\pi(S) := \{\pi(a) : a \in S \subseteq A\}$. We write $\pi(x) \in \triangle A$ to denote the lottery that for every $a \in A$ assigns probability $\mu \in [0, 1]$ to $\pi(a)$ if and only if $x$ assigns probability $\mu$ to $a$. Let $\pi(R) := (\succcurlyeq_i^{\pi(R)})_{i \in N} \in \mathcal{R}^n$ denote the preference profile which has the same preferences on the permuted alternatives as $R$ on the original alternatives. Formally, $\pi(x) \succcurlyeq_i^{\pi(R)} \pi(y)$ if and only if $x \succcurlyeq_i^R y$ 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 $\pi \in \Pi$, $x \succcurlyeq_*^{(S,R)} y$ if and only if $\pi(x) \succcurlyeq_*^{(\pi(S),\pi(R))} \pi(y)$.

## 2.1 Discussion

Our objective was not to identify the weakest set of axioms leading to either of the above representations. Instead, we wanted to demonstrate that even with minor modifications to Arrow's axioms, preference aggregation becomes possible. To make this point convincingly, it is crucial that our axioms are as strong as they can be. That being said, whether weaker axioms would suffice remains an important question. This holds especially true in view of the existing literature that weakens IIA to *independence of redundant alternatives* (IRA); see Dhillon and Mertens (1999) and Brandl (2021). A redundant alternative is one for which there is a lottery over the remaining alternatives that leaves every individual indifferent between the two. While any redundant alternative is comparable, the converse does not hold. Consequently, IRA is weaker than IICA. We show in Appendix D that weakening IICA to IRA would not suffice for our representation. Similarly, weakening MICA to *menu independence of redundant alternatives* (MIRA) wouldn't suffice either. However, in another aspect IICA is indeed stronger than necessary. It would suffice if IICA only applied to cases where all alternatives outside the menu are comparable (i.e., $C = S^c$), akin to the formulation of IRA. We opted for the stronger version of the axiom as it more closely aligns with Arrow's IIA.

# 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 led 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.[5] Figure 3 shows an example for three individuals and four alternatives. By RA, the planner's

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

**Figure 3:** Pay-off matrix.

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. If individual preferences are not sufficiently diverse, we show that there exists a *dependent alternative*, whose column

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

can be expressed by a linear combination 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, one of them involving the dependent alternative, such that every individual is indifferent between them. In the 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 says that in a polar state, the planner is indifferent between all alternatives. Consider, for instance, a polar state as described by the pay-off matrix depicted on the left-hand side of Figure 4. Note

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

**Figure 4:** 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. We assume that the planner weakly prefers $e$ over $f$ and show that this implies a weak preference for $f$ over $e$. This, of course, only leaves indifference. First, notice that the pay-off matrix on the right-hand side of Figure 4 depicts the state that results from permuting the labels of Individuals 1 and

2 in the left-hand state. If $e$ is weakly preferred to $f$ in the left-hand state, by AN this must be the case in the right-hand state as well. Second, notice that the left-hand state results from permuting the labels of alternatives $e$ and $f$ in the right-hand state. If $e$ is weakly preferred to $f$ in the right-hand state, NE would demand that $f$ must be weakly preferred to $e$ in the left-hand state. With the help of MICA and PI, we show that indeed NE is satisfied in polar states. We leave the details to the formal proof in the appendix.

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

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

**Figure 5:** Resulting pay-off matrix.

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 initial menu. Note that if there are fewer than $n$ alternatives outside the initial 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, identifying the welfare of a dynamically inconsistent decision maker, and finding a fair solution to a bargaining problem. The literature has so far treated each of these problems in isolation. In the previous sections, we provided a unifying framework that aggregates individual preferences consistently across applications. With this methodology 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 $i \in N$,

$$u_i((m_1, g_1), ..., (m_n, g_n)) = \alpha_i m_i + v_i(g_i) \tag{3}$$

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 *total consumer surplus*, short TCS, given by

$$\begin{aligned} W_{\text{TCS}}((m_1, g_1), ..., (m_n, g_n)) &= \sum_{i \in N} \frac{1}{\alpha_i} u_i((m_1, g_1), ..., (m_n, g_n)) \\ &= \sum_{i \in N} m_i + \sum_{i \in N} \frac{v_i(g_i)}{\alpha_i}. \end{aligned}$$

Consider two allocations, $a$ and $b$, such that some individuals strictly prefer $a$ and some strictly prefer $b$. If $W_{\text{TCS}}(a)$ is higher than $W_{\text{TCS}}(b)$, then there exist monetary transfers after $a$, resulting in an allocation $a'$, such that $a'$ Pareto dominates $b$. This makes TCS quite appealing. However, if these transfers aren't implemented, TCS makes a judgment on how the well-being of different individuals should be traded off. We show that the way in which these trade-offs are made violates neutrality. Specifically, TCS is sensitive to which of the two consumption goods is selected as the numéraire. Let $\mathcal{R}^{QL}$ be the subset of $\mathcal{R}^n$ such that individual preferences can be represented by quasi-linear utility functions as in (3).

**Proposition 3.** There doesn't exist an aggregation rule $\succcurlyeq_*$ on the restricted domain $\mathcal{R}^{QL}$ that is represented by $W_{\text{TCS}}$ and satisfies NE.

*Proof.* It is sufficient to show that the axiom is violated in some state, so consider a state $R \in \mathcal{R}^{QL}$ where for each $i \in N$, $v_i(g_i) = \beta_i g_i$ for some $\beta_i > 0$. Next consider the permutation $\pi$ of alternatives 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

$$W^R_{\text{TCS}}(a) \geq W^R_{\text{TCS}}(b) \text{ if and only if } W^{\pi(R)}_{\text{TCS}}(\pi(a)) \geq W^{\pi(R)}_{\text{TCS}}(\pi(b)). \tag{4}$$

Let $a = ((0,1), (0,0), ..., (0,0))$ and $b = ((0,0), (2,0), ..., (0,0))$. Then $W^R_{\text{TCS}}(a) = \frac{\beta_1}{\alpha_1}$, $W^R_{\text{TCS}}(b) = 2$, $W^{\pi(R)}_{\text{TCS}}(\pi(a)) = 1$ and $W^{\pi(R)}_{\text{TCS}}(\pi(b)) = \frac{\alpha_2}{\beta_2}$. This violates (4) for instance when $\beta_1 = 3$, $\alpha_1 = 1$, $\alpha_2 = 2$ and $\beta_2 = 1$. □

Now assume that allocations result from the following setting. A monopolist produces the good $g$ at constant marginal cost $c$. Individuals can be divided into two distinct groups: students $s$ and adults $a$. Within each group, every individual has the same preferences. The fraction of individuals belonging to Group $J \in \{s, a\}$ is denoted by $\gamma_J$. Total demand of Group $J$ is then $D_J(p) := n\gamma_J(v_i')^{-1}(p\alpha_i)$ where $i \in J$. We assume that the monopolist can identify which group an individual belongs to and hence can choose a price pair $(p_s, p_a) \in [c, \infty)^2$, where $p_J$ denotes the price charged to Group $J$. We denote by $\psi(p_s, p_a) \in \mathbb{R}^{2n}_+$ the allocation resulting from a price pair $(p_s, p_a)$. Let $p^*_J$ denote the monopoly price charged to Group $J$, let $p^*$ denote the profit-maximizing price if price discrimination was prohibited, and assume

$$p^*_a > p^* > p^*_s.$$

A policymaker can choose whether to allow or prohibit price discrimination, hence $S = \{\psi(p^*_s, p^*_a), \psi(p^*, p^*)\}$. To guide this decision, the policymaker wants to assess whether the prohibition of price discrimination benefits consumers. According to TCS, aggregate welfare is given by

$$W_{\text{TCS}}(\psi(p_s, p_a)) = CS_s(p_s) + CS_a(p_a)$$

where $CS_J(p) := \int_p^\infty D_J(p)dp$. Since total consumer surplus violates neutrality, we propose setting-contingent utilitarianism as an alternative. The setting puts natural

bounds on the possible alternatives, namely $A = \{\psi(p_s, p_a) \in \mathbb{R}_+^{2n} : (p_s, p_a) \in [c, \infty)^2\}$. Setting-contingent utilitarianism, short SCU, is then given by

$$W_{\text{SCU}}(\psi(p_s, p_a)) = \frac{n\gamma_s}{CS_s(c)} CS_s(p_s) + \frac{n\gamma_a}{CS_a(c)} CS_a(p_a).$$

We now compare these two welfare measures. To make things simple, we consider the textbook case of linear demand. It is well known that when demand is linear, total consumer surplus is higher under price discrimination if and only if one group isn't served under the uniform price. Our welfare measure, on the other hand, suggests that price discrimination can be socially beneficial, even if both groups are served under uniform pricing. We demonstrate this with the help of the following example. Figure 6 shows the demand and monopoly prices for $D_s(p) = 2000 - 50p$, $D_a(p) = 7000 - 100p$, and $c = 0$. If the firm is allowed to price discriminate, it will charge \$20 to students and

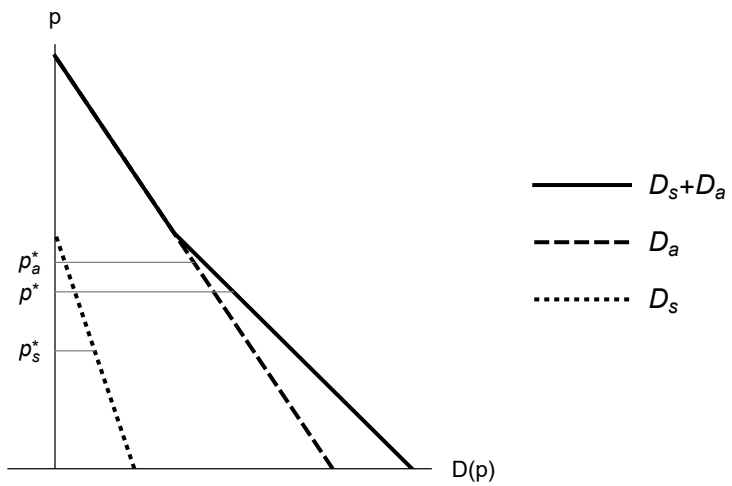


**Figure 6:** Linear demand and monopoly prices.

\$35 to adults, which yields a consumer surplus of $CS_s(\$20) = \$10,000$ and $CS_a(\$20) = \$61,250$. If price discrimination were prohibited, the firm would charge \$30 to both, which yields a consumer surplus of $CS_s(\$30) = \$2,500$ and $CS_a(\$30) = \$80,000$. Note that TCS is higher when price discrimination is prohibited. For both groups to be better off under uniform pricing, adults would have to pay between \$7,500 and \$18,750 to students. However, transfers are typically not implemented, which makes students worse off under uniform pricing. Going from price discrimination to uniform pricing, the normalized utility of students decreases from 0.25 to roughly 0.06, while it increases for adults 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 SCU. In a sense, SCU penalizes adults 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 SCU would agree with TCS since both satisfy SP. Furthermore, SCU accounts for how many individuals are positively and negatively affected by the policy, whereas TCS ignores the shares of individuals in each group. If, for instance, the share of students was below roughly 0.29, then SCU would favor the prohibition of price discrimination, as sufficiently many adults would be positively affected by the policy. SCU values individuality and has a flavor similar to the majority principle. Finally, note that SCU, unlike TCS, can be meaningfully applied even when individual utilities are not quasi-linear.

## 4.2 Intertemporal Welfare

Consider a finitely lived decision maker (DM) who consumes a good in each period up to period $n$. As in Laibson (1997), we assume the DM discounts consumption quasi-hyperbolically, such that the utility of a consumption sequence $(c_1, ..., c_n) \in \mathbb{R}^n_+$ experienced in period $t$ is given by

$$u_t(c_t, ..., c_n) = v(c_t) + \beta \sum_{i=t+1}^{n} \delta^{i-t} v(c_i) \tag{5}$$

for $\beta, \delta \in [0, 1]$ and some monotone per-period valuation $v : \mathbb{R}_+ \to \mathbb{R}$. If $\beta < 1$, the DM is dynamically inconsistent, meaning there exist sequences $(c_1, ..., c_n)$, $(c_1, ..., c_t, c'_{t+1}, ..., c'_n) \in \mathbb{R}^n_+$ 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 assess which of these 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 & Rabin, 1999) given by

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

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. However, this view ignores the possibility that, even though there is no conflict looking forward, there could be a conflict looking backward. For instance, the DM might regret an earlier decision and prefer they had saved more in the past while simultaneously agreeing with the earlier self on how much to save today. To formalize this idea, let each self $t$ have a vNM preference $\succcurlyeq_t$ over $\triangle(\mathbb{R}^n_+)$. This means that Self $t$ can compare sequences that differ in the consumption levels before Period $t$. However, note that since one cannot affect one's past, $\succcurlyeq_t$ is only partially revealed. Let $\mathcal{R}^{QH}$ be the subset of $\mathcal{R}^n$ such that the revealed part of the individual preference can be represented by quasi-hyperbolically discounted utility as in (5). Formally, for any $(\succcurlyeq_1, ..., \succcurlyeq_n) \in \mathcal{R}^{QH}$, $(c_1, ..., c_n) \in \mathbb{R}^n_+$ and $t \in \{1, ..., n\}$, preferences of Self $t$ over lotteries that assign probability 1 on sequences starting with $(c_1, ..., c_{t-1})$ can be represented by (5). Note that the selves are in conflict unless $\succcurlyeq_i = \succcurlyeq_j$ for all $i, j \in \{1, ..., n\}$. Hence, for all selves to agree, the DM would have to value past consumption more than current consumption and value it more the further it lies in the past. Since this doesn't seem very plausible, we would expect backward-looking disagreements to be common. Now that all selves have preferences over the same domain, we can view these conflicting interests from the perspective of preference aggregation. This enables us to demonstrate that taking Self 1's utility as a measure of total welfare whenever $\beta = 1$ violates our Pareto condition.

**Proposition 4.** Let $\succcurlyeq_*$ be an aggregation rule on the restricted domain $\mathcal{R}^{QH}$, such that $\succcurlyeq_*^{(S,R)} = \succcurlyeq_1^R$ whenever $\beta = 1$. Then $\succcurlyeq_*$ violates 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 $t \in \{2, ..., n\}$, $(c_1, ..., c_n), (c'_1, ..., c'_{t-1}, c_t, ..., c_n) \in \mathbb{R}^n_+$,

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

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, $u_1$ measures total welfare of the DM, then $s$ and $s'$ are equally desirable. □

Let us now return to the aforementioned welfare criteria. The Pareto principle cannot compare every sequence in $\mathbb{R}^n_+$; hence, is incomplete and violates RA. Long-run utility is equal to $u_1$ for $\beta = 1$, and therefore, as shown by Proposition 4, violates SP. As an alternative to the established criteria, we propose menu-contingent utilitarianism. We assume that the social planner can distribute an initial endowment of size 1 across periods, hence $S = \{(c_1, ..., c_n) \in [0,1]^n : \sum_1^n c_t \leq 1\}$. Furthermore, 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 (5). Next, we have to 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

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

In contrast to long-run utility, the weights $\gamma_t$ 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. Finally, note that MCU can be applied even when the DM is not a quasi-hyperbolic discounter.

## 4.3 Bargaining

Consider $n$ individuals who have the possibility to cooperate and create a surplus. In order to generate this surplus, they must agree on an alternative to be implemented. For instance, one could think of a buyer and seller bargaining over the price or a worker and an employer negotiating a wage. If the group cannot reach an agreement, each individual resorts to their respective outside options, which we call the *disagreement alternative* and denote by $a_0$. In the following exposition, we consider bargaining

through the eyes of an arbitrator who, to facilitate cooperation, has to choose an alternative for the group. Equivalently, one could phrase the selection as a consequence of fairness postulates.

The literature on axiomatic bargaining suggests that the arbitrator adheres to the following approach. The problem is first reformulated in terms of vectors of utilities, and then a bargaining solution is applied. This reformulation is discussed in more detail below. Here, we recall the bargaining approach. A *bargaining problem* $(d, U)$ specifies a set of utility vectors $U \subset \mathbb{R}^n$, called the *bargaining set*, and a *disagreement point* $d = (d_1, ..., d_n) \in U$. We assume that $U$ is compact and denote by $\mathcal{B}$ the set of all such bargaining problems. A bargaining solution $f$ assigns to each $(d, U)$ a subset of $U$, hence $f(d, U) \subseteq U$. Two bargaining solutions are prominent in the literature: the *Nash bargaining solution* (Nash, 1950), short Nash, and the *Kalai-Smorodinsky bargaining solution* (Kalai & Smorodinsky, 1975), short KS. Note that both solutions require the bargaining set to be convex. Denote by $\mathcal{B}_{\text{con}}$ the subset of $\mathcal{B}$ for which $U$ is convex. Then Nash is given by

$$f_{\text{Nash}}(d, U) := \underset{(v_1,...,v_n)\in U}{\arg\max} \prod_{i\in N} (v_i - d_i)$$

for all $(d, U) \in \mathcal{B}_{\text{con}}$. Next, we define KS, which requires some additional notation. For any $(d, U) \in \mathcal{B}$, let $\Lambda(d, U) := \{(v_1, ..., v_n) \in U : v_i \geq d_i \text{ for all } i \in N\}$. Hence, $\Lambda(d, U)$ is the set of utility vectors in $U$ that weakly Pareto-dominate $d$. Furthermore, let $\alpha_i(d, U)$ denote $i$'s maximal utility among all points in $\Lambda(d, U)$. Note that KS only applies for $n = 2$. So $f_{\text{KS}}(d, U)$ is the intersection of the Pareto frontier of $U$ and the line connecting $d$ and $(\alpha_1(d, U), \alpha_2(d, U))$ for all $(d, U) \in \mathcal{B}_{\text{con}}$.

We now investigate whether these bargaining solutions could be used by an arbitrator who first evaluates the different alternatives and then selects the best according to this evaluation. Let $A$ denote the set of alternatives and note that $a_0 \in A$. Let $\mathcal{U}$ denote the set of possible vNM utility functions over $\triangle A$. A utility profile $u \in \mathcal{U}^n$ specifies for each individual a utility function over the possible alternatives. We say that a bargaining problem $(d, U) \in \mathcal{B}$ is *associated* with a utility profile $u \in \mathcal{U}^n$ (and vice versa) if $d = (u_1(a_0), ..., u_n(a_0))$ and $U = \{(u_1(a), ..., u_n(a)) : a \in A\}$. To ensure that any bargaining problem in $\mathcal{B}$ can be generated by some utility profile in $\mathcal{U}^n$, we require $A$ to be sufficiently rich. In line with our motivating examples, the reader might

think of $A$ as the set of prices or wages, i.e., $A = \mathbb{R}_+$. We denote by $\succcurlyeq_+$ the arbitrator's aggregation rule, such that for every $u \in \mathcal{U}^n$, $\succcurlyeq_+^u$ is a binary relation over $\triangle A$. Note that we deviate from the aggregation rule as defined in Section 2 in two ways. We have indicated this by the use of a different subscript. First, for notational convenience, $\succcurlyeq_+$ is a function of utility profiles rather than preference profiles. Second, $\succcurlyeq_+$ does not depend on the menu because, as we will argue later, in a bargaining setting, the menu is determined endogenously. Neither of these deviations is responsible for our subsequent finding. Furthermore, in line with Section 2 we require the arbitrator to evaluate lotteries over $A$. We don't insist that randomization is feasible, but if it were feasible, the arbitrator should be able to evaluate the resulting lotteries. Finally, we connect the aggregation rule to a bargaining solution through the following definition.

**Definition 2.** $\succcurlyeq_+$ is *consistent* with $f$ if for any $(d, U) \in \mathcal{B}_{\text{con}}$ and associated $u \in \mathcal{U}^n$, the following holds. For all $x \in \triangle A$,

$$(u_1(x), ..., u_n(x)) \in f(d, U) \text{ if and only if } x \in \{y \in \triangle A : y \succcurlyeq_+^u z \text{ for all } z \in \triangle A\}.$$

The definition captures an implicit assumption of the bargaining approach, namely that alternatives are selected solely based on their utility vectors. Hence, if some lottery's utility vector is chosen by the bargaining solution, this lottery must be among the most preferred. Conversely, if a lottery is among the most preferred, its utility vector must be chosen by the bargaining solution. We have restricted the definition to convex bargaining sets as this is the domain of the prominent bargaining solutions.

We can now investigate whether a given bargain solution is consistent with certain desirable features of our aggregation rule. We find that neither Nash nor KS is consistent with rationality. More generally, any bargaining rule that imposes ex-ante symmetry is inconsistent with the vNM postulates. This is well known; see for instance Diamond (1965), Harsanyi (1975), Broome (1984) or Machina (1989). Nevertheless, we provide the following proposition for the sake of completeness.

**Proposition 5.** There doesn't exist an aggregation rule $\succcurlyeq_+$ that satisfies RA and is consistent with either $f_{\text{Nash}}$ or $f_{\text{KS}}$.

*Proof.* We prove the proposition for $n = 2$ as this is required for KS. The proof can be easily extended to a general $n$. The proof strategy goes as follows. We assume

that $\succcurlyeq_+$ is consistent with a bargaining solution $f$ that satisfies Nash's *symmetry* and Pareto axiom. Symmetry says that in a symmetric bargaining problem, any element of the bargaining solution must assign equal utility to all individuals. Note that both Nash and KS satisfy symmetry. We then show that $\succcurlyeq_+$ must violate RA.

Consider the bargaining problem $((0,0),\mathcal{S})$ where $\mathcal{S}$ is the unit simplex. As $((0,0),\mathcal{S})$ is symmetric, $f$ selects $\{(0.5,0.5)\}$. See Figure 7 for an illustration. For any

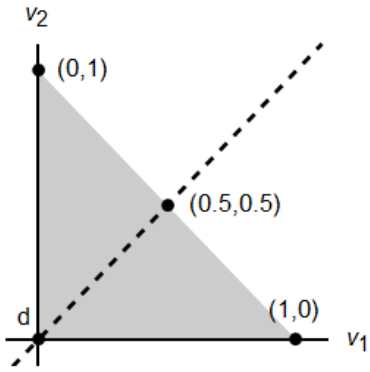


**Figure 7:** Symmetric bargaining solution for $((0,0),\mathcal{S})$.

$u \in \mathcal{U}^2$ associated with $((0,0),\mathcal{S})$, there must exist two alternatives $a_1, a_2 \in A$ that generate the extreme points of $\mathcal{S}$, i.e., $(u_1(a_1), u_2(a_1)) = (1,0)$ and $(u_1(a_2), u_2(a_2)) = (0,1)$. Now consider the coin-flip between $a_1$ and $a_2$, denoted by $x := \frac{1}{2}[a_1] + \frac{1}{2}[a_2]$. As $(u_1(x), u_2(x)) = (0.5, 0.5) \in f((0,0),\mathcal{S})$, $x$ must be among the most preferred elements according to $\succcurlyeq^u_+$. Formally $x \in \{y \in \triangle A : y \succcurlyeq^u_+ z \text{ for all } z \in \triangle A\}$. But neither $a_1$ nor $a_2$ are among the most preferred elements, as neither $(1,0)$ nor $(0,1)$ is selected by $f$. Therefore,

$$\frac{1}{2}[a_1] + \frac{1}{2}[a_2] \succ^u_+ [a_1], \tag{6}$$

$$\frac{1}{2}[a_1] + \frac{1}{2}[a_2] \succ^u_+ [a_2]. \tag{7}$$

By the vNM independence axiom, (6) would imply $[a_2] \succ^u_+ [a_1]$ and (7) would imply $[a_1] \succ^u_+ [a_2]$, a contradiction. Hence, the vNM independence axiom and consequently RA must be violated. □

We believe that the arbitrator should make the decision based on a rational evaluation of the alternatives. As this is not reconcilable with the prominent bargaining solutions, we offer menu-contingent utilitarianism as an alternative. The menu arises naturally in this setting. Because an individual would not agree to an alternative

worse than $a_0$, it is the set of alternatives that weakly Pareto-dominate $a_0$. Formally, given $u \in \mathcal{U}^n$ the menu is $S := \{a \in A : u_i(a) \geq u_i(a_0) \text{ for all } i \in N\}$. Note that a bargaining problem $(d, U)$ contains all relevant information for our rule to evaluate the utility vectors in $\Lambda(d, U)$. Hence, menu-contingent utilitarianism leads to a well-defined bargaining solution, which we abbreviated by MCU and which is given by

$$f_{\text{MCU}}(d, U) := \underset{(v_1,\ldots,v_n)\in\Lambda(d,U)}{\arg\max} \sum_{i\in N} \frac{v_i - d_i}{\alpha_i(d, U) - d_i}$$

for all $(d, U) \in \mathcal{B}$. MCU selects the utility vectors that are most preferred under a menu-contingent utilitarian evaluation. Note that MCU does not require convexity of the bargaining set. The same bargaining solution has been axiomatized before by Pivato (2009) under the assumption of convexity of the bargaining set. A similar bargaining solution has been axiomatized by Cao (1982) and Baris (2018), again assuming convexity of the bargaining set. Their solution differs from ours, as they normalize individual utilities with respect to each individual's highest utility in $U$, whereas we normalize with respect to each individual's highest utility in $\Lambda(d, U)$.

Like Nash and KS, MCU has a neat geometric interpretation, which we demonstrate with the help of Figure 8. The disagreement point is the origin and the curve shows the

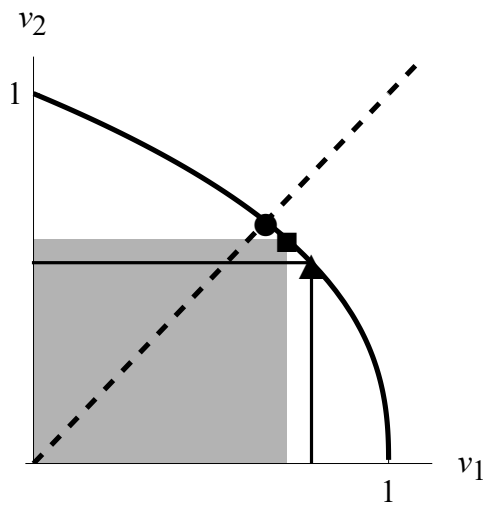


**Figure 8:** Nash, KS and MCU.

Pareto frontier of bargaining set for $u_1(a) = 1 - a$ and $u_2(a) = a^{\frac{2}{5}}$. 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 MCU (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 MCU 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 amount. Hence, in this example, MCU favors Individual 1 relative to KS because the slope at KS is flatter than -1 such that Individual 1 can be made significantly better off at the cost of making Individual 2 slightly worse off.

# 5 Conclusion

We initiated this project in pursuit of a rule that is rational, benevolent, and impartial, as these desiderata are not satisfied by the rules that are typically used in applications. We learned from Arrow (1950, 1963) that for a rule to satisfy these desiderata, it must be sensitive to context. We offer two rules, each motivated by adhering closely to Arrow. Each rule results from slightly weakening one of Arrow's two context-restricting axioms, MI and IIA. Each rule makes an explicit normative judgment about what is allowed to matter. We believe that in different situations, different considerations are allowed to matter. Consequently, we find that there is no single unifying rule. This revelation could pave the way for numerous rules derived from various modifications of MI and IIA. However, if we permit a planner to tailor the context to a specific application, almost any social choice can be justified. To maintain discipline, we recommend selecting from our two rules, as each represents a normative judgment on one side of the spectrum.

The menu-contingent utilitarian rule satisfies IIA and thus reflects the judgment that nothing outside the menu is allowed to matter. However, the menu itself does matter. We believe this normative consideration to be relevant for bargaining situations, as bargaining involves finding a compromise that provides each individual with a fair share of the surplus. Our understanding of fairness is relative to what can be achieved and must disregard alternatives that individuals would never agree upon. We also propose this rule for practical convenience in situations where there is no information about individual preferences outside the menu or when the setting does not impose natural bounds on individual utilities.

On the other side of the spectrum is the setting-contingent utilitarian rule. This rule satisfies MI, hence reflects the judgment that the menu itself is not allowed to matter. Individual preferences over all possible alternatives are taken into account, independently of whether an alternative is currently feasible. We believe this normative consideration to be relevant for policymaking, where the policymaker should take a birds-eye view and assess welfare on the basis of the entire setting. In addition, policymakers often face binary decisions sequentially, namely whether to adopt a specific policy or not. Menu independence ensures that these decisions are consistent across time.

# 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 $\succcurlyeq_*^{(S,R)}$, 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 $\succcurlyeq_*^{(S,R)}$ 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. In the first case, individual preferences are sufficiently diverse, such that 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

$$\sum_{b \in B} \gamma^a(b)\vec{u}(b) = \vec{u}(a). \tag{8}$$

We will now show that each $\gamma^a$ 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

$$\sum_{b\in B} \gamma^a(b)u_*(b) = u_*(a). \tag{9}$$

Compare this to Equation (8). 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. □

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 $R_{\mathrm{I}}$ 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 $R_{\mathrm{II}}$ on $S_{\mathrm{II}}$ 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 $R_{\mathrm{III}}$ and $R$ agree on $S$. Therefore, $p_j \succcurlyeq_*^{(S,R)} p_i$ by IIA, which implies $p_i \sim_*^{(S,R)} p_j$. □

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 $B \cap C$. □

**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 $\succcurlyeq$. □

*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 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 $S_{\mathrm{I}} := S$ and denote an arbitrary alternative in $S^c$ by $e$. If, on the other hand, $S = A$, we construct $S_{\mathrm{I}}$ in the following way. As $|A| \geq 2n+4$, $S \setminus \{b,c,d\}$ must have at least $2n+1$ alternatives. Consequently, at least one alternative $e \in S \setminus \{b,c,d\}$ must be comparable relative to $S \setminus \{e\}$ under $R$. So let $S_{\mathrm{I}} := S \setminus \{e\}$. Note that either way, there is an alternative $e$ in $S_{\mathrm{I}}^c$. Let $R_{\mathrm{I}} \in \mathcal{R}^n$ denote a preference profile that agrees with $R$ on $S_{\mathrm{I}}$ and where for each $i \in N$, $e \sim_i^{R_{\mathrm{I}}} a$ for all $a \in \{a \in S_{\mathrm{I}} : a \succcurlyeq_i^{R_{\mathrm{I}}} a' \text{ for all } a' \in S_{\mathrm{I}}\}$. 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 $\{b,c,d,e\} \subseteq S_{\mathrm{III}}$ and every $a \in S_{\mathrm{II}} \setminus S_{\mathrm{III}}$ is comparable relative to $S_{\mathrm{III}}$ under $R_{\mathrm{I}}$. Note that there are at most $n+4$ alternatives in $S_{\mathrm{III}}$, namely $b,c,d,e$, and $n$ alternatives that are each worst for exactly one individual. Hence, as $|A| \geq 2n+4$, there are at least $n$ alternatives in $S_{\mathrm{III}}^c$. 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 $R_{\mathrm{I}}$ on $S_{\mathrm{III}}$ and where there exists $P \subseteq S_{\mathrm{III}}^c$ such that (i) $(P, R_{\mathrm{II}})$ 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 $\succcurlyeq_*^{(P\cup S_{\text{III}},R_{\text{II}})}$ 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 $\succcurlyeq_*^{(P\cup S_{\text{III}},R_{\text{II}})}$ is represented by

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

By Proposition 7, $p_i \sim_*^{(P,R_{\text{II}})} p_j$ for all $i,j \in N_{\succ}^{(P,R_{\text{II}})}$. Because every alternative in $S_{\text{III}}$ is comparable relative to $P$ under $R_{\text{II}}$, by MICA $p_i \sim_*^{(P\cup S_{\text{III}},R_{\text{II}})} p_j$ for all $i,j \in N_{\succ}^{(P\cup S_{\text{III}},R_{\text{II}})}$ as well. This implies $\lambda_i = \lambda_j =: \lambda$ for all $i,j \in N_{\succ}^{(P\cup S_{\text{III}},R_{\text{II}})}$. Note that if $i \notin N_{\succ}^{(P\cup S_{\text{III}},R_{\text{II}})}$, then $u_i^{P\cup S_{\text{III}},R_{\text{II}}}$ is constant on $P\cup S_{\text{III}}$, and $\lambda_i$ can be normalized to $\lambda$ as the planner's vNM representation is unique up to positive affine transformations. Similarly, $\lambda_0$ can be normalized to 0. We now show that $\lambda > 0$ and, therefore, $\lambda$ can be normalized to 1. We distinguish three cases. First, consider the case where $N_{\succ}^{(P\cup S_{\text{III}},R_{\text{II}})} = N$ and for any alternative $a \in S_{\text{III}}$ there exists a lottery $x_a \in \triangle P$ such that $a \sim_i^{R_{\text{II}}} x_a$ for all $i \in N$. Then $\succcurlyeq_*^{(P\cup S_{\text{III}},R_{\text{II}})}$ is totally indifferent on $P\cup S_{\text{III}}$ and any $\lambda$ represents the same preferences. Second, consider the case where $N_{\succ}^{(P\cup S_{\text{III}},R_{\text{II}})} = N$ and for some alternative $a \in S$ such a lottery does not exist. Depending on whether $\sum_{i=1}^n u_i^{S_{\text{III}},R_{\text{II}}}(a)$ 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, $\lambda$ has to be strictly positive. Third, if $N_{\succ}^{(P\cup S_{\text{III}},R_{\text{II}})} \neq N$ then for SP to be satisfied, $p_i \succ_*^{(P\cup S_{\text{III}},R_{\text{II}})} p_j$ for any $i \in N_{\succ}^{(P\cup S_{\text{III}},R_{\text{II}})}$ and $j \notin N_{\succ}^{(P\cup S_{\text{III}},R_{\text{II}})}$. This requires $\lambda$ to be strictly positive as well. In any of the three cases, $\lambda$ can be normalized to 1. Hence, we have shown that $\succcurlyeq_*^{(P\cup S_{\text{III}},R_{\text{II}})}$ is represented by

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

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

$$\begin{aligned}
\succcurlyeq_*^{(S,R)} \text{ and } \succcurlyeq_*^{(S_{\mathrm{I}},R)} &\text{ agree on } \{b,c,d\}. \text{ (MICA)}\\
\succcurlyeq_*^{(S_{\mathrm{I}},R)} \text{ and } \succcurlyeq_*^{(S_{\mathrm{I}},R_{\mathrm{I}})} &\text{ agree on } \{b,c,d\}. \text{ (IIA)}\\
\succcurlyeq_*^{(S_{\mathrm{I}},R_{\mathrm{I}})} \text{ and } \succcurlyeq_*^{(S_{\mathrm{II}},R_{\mathrm{I}})} &\text{ agree on } \{b,c,d\}. \text{ (MICA)}\\
\succcurlyeq_*^{(S_{\mathrm{II}},R_{\mathrm{I}})} \text{ and } \succcurlyeq_*^{(S_{\mathrm{III}},R_{\mathrm{I}})} &\text{ agree on } \{b,c,d\}. \text{ (MICA)}\\
\succcurlyeq_*^{(S_{\mathrm{III}},R_{\mathrm{I}})} \text{ and } \succcurlyeq_*^{(S_{\mathrm{III}},R_{\mathrm{II}})} &\text{ agree on } \{b,c,d\}. \text{ (IIA)}\\
\succcurlyeq_*^{(S_{\mathrm{III}},R_{\mathrm{II}})} \text{ and } \succcurlyeq_*^{(P\cup S_{\mathrm{III}},R_{\mathrm{II}})} &\text{ agree on } \{b,c,d\}. \text{ (MICA)}
\end{aligned}$$

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 $R_{\mathrm{I}}$ 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 $(S,R)\in\Omega$, $\succcurlyeq_*^{(S,R)}$ 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 $R_{\text{II}}$ is selected such that every alternative in $(P \cup S_{\text{III}})^c$ is comparable relative to $(P \cup S_{\text{III}})$ under $R_{\text{II}}$. Then one can simply replace the use of Proposition 7 with Proposition 8 and the use of IIA with IICA for the proof to go through. This concludes the proof of Theorem 2.

*Proof of Proposition 2*

We prove the proposition by providing a counterexample. Specifically, we identify an aggregation rule 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 aggregation rule where $\succcurlyeq_*^{(S,R)}$ is represented by

$$\sum_{i \in N} \left( \sum_{a \in S} u_i^{S,R}(a) \right) u_i^{S,R} \tag{10}$$

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

$$\sum_{i \in N} u_i^{S,R} \tag{11}$$

whenever $(S, R) \notin \hat{\Omega}$. Note that it could be that for all $(S, R) \in \hat{\Omega}$, (10) and (11) 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 $\Omega$. If this aggregation rule 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) (10) satisfies the restrictions imposed between any two states in $\hat{\Omega}$, and (iii) (10) 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, (10) satisfies AN as the weight on each utility function only depends on the utility function itself but not on the index. Third, (10) assigns positive weights to all individual utility functions that are not indifferent on $A$. 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. Let $\hat{\Omega}$ be defined as above. Now consider an aggregation rule where $\succcurlyeq_*^{(S,R)}$ is represented by

$$\sum_{i \in N} \left( \sum_{a \in S} u_i^{A,R}(a) \right) u_i^{A,R} \tag{12}$$

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

$$\sum_{i \in N} u_i^{A,R} \tag{13}$$

whenever $(S, R) \notin \hat{\Omega}$. As before, assume that $(A, \hat{R})$ has been selected such that (12) and (13) are not positive affine transformations of each other. Note that both MI and IICA connect states. As before, it holds that (i) no axiom connects a state in $\hat{\Omega}$ to a state outside of $\hat{\Omega}$, (ii) (12) satisfies the restrictions imposed between any two states in $\hat{\Omega}$, and (iii) (12) satisfies RA and SP for each $(S, R) \in \hat{\Omega}$. This can be shown, similarly to how it was shown for Theorem 1 above. Just note that 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 $R \in \pi(\hat{R})$. This concludes the proof of Proposition 2 in the 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 $\Omega$ to be the set of states, such that for each $(S, R) \in \Omega$, both $\left\{a \in A : a \succcurlyeq_i^R b \text{ for all } b \in A\right\}$ 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 3.** $a \in A$ 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.

**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

$$(1-\varepsilon)u_i^R(a) + \varepsilon u_i^R(x_{i,\varepsilon}) = u_i^R(y_{i,\varepsilon}). \tag{14}$$

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 (14) is satisfied for all $\varepsilon$. So assume $u_i^R(a) = \sup_{b \in B} u_i^R(b)$ and fix $\varepsilon$. Choose $x_{i,\varepsilon}$ arbitrarily. The left-hand side of (14) 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 (14). 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 $S'$, $\succcurlyeq_*^{(S,R)}$ and $\succcurlyeq_*^{(S',R)}$ 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 $\succcurlyeq_i^R$ where $\sup_{a \in B} \hat{u}_i^{B,R}(a) = 1$ and $\inf_{a \in B} u_i^{B,R}(a) = 0$, unless $\succcurlyeq_i^R$ 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 aggregation rule $\succcurlyeq_*$ satisfies RA, SP, AN, IIA and MICA* if and only if for each $(S, R) \in \Omega$, $\succcurlyeq_*^{(S,R)}$ is represented by

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

*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 $R_{\mathrm{I}}$ and $R_{\mathrm{II}}$, 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 $S_{\mathrm{I}}$ 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 $\succcurlyeq_i^{R_{\mathrm{I}}}$, $u_i^{R_{\mathrm{I}}}(e) = \sup_{a \in S_{\mathrm{I}}} u_i^{R_{\mathrm{I}}}(a)$. Let $R_{\mathrm{II}} \in \mathcal{R}^n$ denote a preference profile that agrees with $R_{\mathrm{I}}$ on $S_{\mathrm{III}}$ and where there exists $P \subseteq S_{\mathrm{III}}^c$ such that (i) $(P, R_{\mathrm{II}})$ 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 $\succcurlyeq_i^{R_{\mathrm{II}}}$, $u_i^{R_{\mathrm{II}}}(p_i) = \sup_{a \in S_{\mathrm{III}}} u_i^{R_{\mathrm{II}}}(a)$ 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 $S_{\mathrm{I}}$ under $R_{\mathrm{I}}$ and every alternative in $P$ is approximately comparable relative to $S_{\mathrm{III}}$ under $R_{\mathrm{II}}$. Then the proof of Theorem 1 goes through. □

Theorem 2 holds without adjustment to the axioms.

# Appendix C

In this section, we show that none of the axioms can be dispensed with. We do so by separately dropping each axiom and then identifying a representation that satisfies

the remaining axioms, different from the representation of the respective theorem.

When SP is dropped in either theorem, total indifference of the planner satisfies the remaining axioms. Formally, for each $(S, R) \in \Omega$ and $x \in \triangle S$, $u_*^{(S,R)}(x) = 0$. When AN is dropped in Theorem 1, any weighted sum of individual utilities satisfies the remaining axioms as long as the weights are strictly positive. Formally, there exists $(\lambda_i)_{i \in N} \in \mathbb{R}^n_+$ such that for each $(S, R) \in \Omega$, $u_*^{(S,R)}(x) = \sum_{i=1}^n \lambda_i u_i^{S,R}(x)$. Similarly, when AN is dropped in Theorem 2, $u_*^{(S,R)}(x) = \sum_{i=1}^n \lambda_i u_i^{A,R}(x)$ satisfies the remaining axioms. If RA is dropped in either theorem, then the relative egalitarian rule by Sprumont (2013) satisfies the remaining axioms. The rule ranks alternatives according to the lowest normalized individual utility it generates. If two alternatives generate the same lowest utility, they are further ranked by the second lowest utility, and so on. Depending on the theorem for which we want to provide the counterexample, individual utilities are normalized either with respect to $S$ or $A$. When IIA is dropped in Theorem 1, then the representation of Theorem 2 satisfies the remaining axioms. Similarly, when MI is dropped in Theorem 2, the then representation of Theorem 1 satisfies the remaining axioms. For a representation when either MICA is dropped in Theorem 1 or IICA is dropped in Theorem 2 we refer to Propositions 9 and 10 in Appendix D.

# Appendix D

In this section we consider weaker versions of our context-defining axioms MICA and IICA. The following notion of redundant alternatives has been considered by the literature.

**Definition 4.** $a \in A$ is *redundant* relative to $B \subseteq A$ under $R \in \mathcal{R}^n$ if $a \notin B$ and there exists $x \in \triangle B$ such that $x \sim_i^R [a]$ for all $i \in N$.

Given this definition, we can state the central axiom of Dhillon and Mertens (1999). Note that we adapted the axiom to our setting where the menu is explicit.

**Axiom IRA** (Independence of Redundant Alternatives)**.** Fix $(S, R) \in \Omega$ such that every $a \in S^c$ is redundant relative to $S$ under $R$. For any $R' \in \mathcal{R}^n$, if $R$ and $R'$ agree on $S$ and every $a \in S^c$ is redundant relative to $S$ under $R'$, then $\succsim_*^{(S,R')} = \succsim_*^{(S,R)}$.

Note that IRA differs from IICA in two ways. First, the independence applies to redundant rather than comparable alternatives. Second, IRA applies only if all alternatives outside the menu are redundant, whereas IICA applies when only some alternatives outside the menu are comparable as well. In order to disentangle these difference, we define two intermediate axioms.

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

**Axiom IIRA** (Independence of Irrelevant Redundant Alternatives)**.** Fix $(S, R) \in \Omega$ and $C \subseteq S^c$ such that every $a \in C$ is redundant 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 redundant relative to $C^c$ under $R'$, then $\succcurlyeq_*^{(S,R')} = \succcurlyeq_*^{(S,R)}$.

It is easy to see that IICA* would suffice in Theorem 2, since whenever IICA is used in the proof, all alternatives outside the menu are comparable. However, replacing IICA with IIRA would not suffice, as demonstrated by the following proposition. Consequently, replacing IICA with IRA wouldn't suffice either, as IRA is weaker than IIRA.

**Proposition 9.** Let $|A| \geq 2n + 4$. There exists an aggregation rule $\succcurlyeq_*$ that satisfies RA, SP, AN, IIRA and MI and $\succcurlyeq_*^{(S,R)}$ can not be represented by $\sum_{i \in N} u_i^{A,R}$ for all $(S, R) \in \Omega$.

*Proof.* The following proof is adapted from Sprumont (2013). We prove the proposition by identifying a representation that satisfies the axioms but is not a positive affine transformation of the normalized sum of individual utility functions. For every $(S, R) \in \Omega$, define

$$\Psi(S, R) := \arg\max_{x \in \triangle S} \prod_{i \in N} (u_i^{S,R}(x) + 1)$$

and note that $\Psi(S, R)$ is non-empty and for every $x, y \in \Psi(S, R)$, $u_i^{S,R}(x) = u_i^{S,R}(y)$ for all $i \in N$. Hence, $\psi_i^{(S,R)} := u_i^{S,R}(x) + 1$ for some $x \in \Psi(S, R)$ is well defined. Now consider the following representation of $\succcurlyeq_*^{(S,R)}$,

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

As $\psi_i^{(A,R)}$ is not necessarily the same for all $i \in N$, the representation is not necessarily a positive affine transformation of $\sum_{i \in N} u_i^{A,R}$. What remains to show is that the representation satisfies the axioms stated in the proposition. RA and SP are satisfied as the representation is a weighted sum of individual utility functions with strictly positive weights. AN is satisfied as any individual's weight does not depend on the individual's index. MI is satisfied as the weights do not depend on the menu. Finally, to see that IIRA is satisfied, assume that for $(S, R) \in \Omega$ there exists a set of redundant alternatives $C \subseteq S^c$. Now consider $R' \in \mathcal{R}^n$ where $R' \neq R$, $R$ and $R'$ agree on $C^c$ and $C$ is still redundant under $R'$. As any redundant alternative has a lottery that yields the same product, redundant alternatives can be ignored in the maximization problem involved in $\Psi(S, R)$. Hence, $(\psi_i^{(A,R)})_{i \in N} = (\psi_i^{(A,R')})_{i \in N}$. Note that IICA is violated as changing individual preferences over comparable alternatives outside the menu can affect the maximization problem and, consequently, the weights. □

Similarly, we can replace the notion of comparable alternatives in MICA with redundant alternatives, resulting in the following axiom.

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

Analogously, MIRA does not suffice in Theorem 1.

**Proposition 10.** Let $|A| \geq 2n + 4$. There exists an aggregation rule $\succcurlyeq_*$ that satisfies RA, SP, AN, IIA and MIRA and $\succcurlyeq_*^{(S,R)}$ can not be represented by $\sum_{i \in N} u_i^{S,R}$ for all $(S, R) \in \Omega$.

A representation of $\succcurlyeq_*^{(S,R)}$ that satisfies the stated axioms is $\sum_{i \in N} \psi_i^{(S,R)} u_i^{S,R}$. The proof is analogous to that of Proposition 9.

## 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* (2nd ed.). Yale University Press.

Baris, O. F. (2018). Timing effect in bargaining and ex ante efficiency of the relative utilitarian solution. *Theory and Decision*, *84*(4), 547–556.

Blackorby, C., Donaldson, D., & Weymark, J. A. (1999). Harsanyi's social aggregation theorem for state-contingent alternatives. *Journal of Mathematical Economics*, *32*(3), 365–387.

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

Börgers, T., & Choo, Y.-M. (2017). *Revealed relative utilitarianism* (CESifo Working Paper Series No. 6613).

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

Broome, J. (1984). Uncertainty and fairness. *The Economic Journal*, *94*(375), 624–632.

Cao, X. (1982). Preference functions and bargaining solutions. *1982 21st IEEE Conference on Decision and Control*, 164–171.

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

Coulhon, T., & Mongin, P. (1989). Social choice theory in the case of von Neumann-Morgenstern utilities. *Social Choice and Welfare*, *6*(3), 175–187.

De Meyer, B., & Mongin, P. (1995). A note on affine aggregation. *Economics Letters*, *47*(2), 177–183.

Dhillon, A., & Mertens, J.-F. (1997). An impossibility theorem with von Neumann-Morgenstern preferences. *Economics Letters*, *56*(3), 305–309.

Dhillon, A., & Mertens, J.-F. (1999). Relative utilitarianism. *Econometrica*, *67*(3), 471–498.

Diamond, P. A. (1965). The evaluation of infinite utility streams. *Econometrica*, *33*(1), 170–177.

Dietrich, F. (2021). Fully Bayesian aggregation. *Journal of Economic Theory*, *194*, 105255.

Domotor, Z. (1979). Ordered sum and tensor product of linear utility structures. *Theory and Decision*, *11*(4), 375–399.

Fishburn, P. C. (1984). On Harsanyi's utilitarian cardinal welfare theorem. *Theory and Decision*, *17*(1), 21–28.

Gaertner, W., & Xu, Y. (2012). A general scoring rule. *Mathematical Social Sciences*, *63*(3), 193–196.

Gilboa, I., Samet, D., & Schmeidler, D. (2004). Utilitarian aggregation of beliefs and tastes. *Journal of Political Economy*, *112*(4), 932–938.

Goldman, S. M. (1979). Intertemporally inconsistent preferences and the rate of consumption. *Econometrica*, *47*(3), 621.

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

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

Harsanyi, J. C. (1975). Nonlinear social welfare functions. *Theory and Decision*, *6*(3), 311–332.

Hylland, A. (1980). Aggregation procedure for cardinal preferences: A comment. *Econometrica*, *48*(2), 539.

Kalai, E., & Schmeidler, D. (1977). Aggregation procedure for cardinal preferences: A formulation and proof of Samuelson's impossibility conjecture. *Econometrica*, *45*(6), 1431.

Kalai, E., & Smorodinsky, M. (1975). Other solutions to Nash's bargaining problem. *Econometrica*, *43*(3), 513–518.

Karni, E. (1998). Impartiality: Definition and representation. *Econometrica*, *66*(6), 1405.

Karni, E., & Weymark, J. A. (2024). Impartiality and relative utilitarianism. *Social Choice and Welfare*.

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

Macé, A. (2018). Voting with evaluations: Characterizations of evaluative voting and range voting. *Journal of Mathematical Economics*, *79*, 10–17.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, *27*(4), 1622–1668.

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. (2022). *Arrow's theorem, may's axioms, and borda's rule* (Working Paper).

Mongin, P. (1994). Harsanyi's Aggregation Theorem: Multi-profile version and unsettled questions. *Social Choice and Welfare*, *11*(4), 331–354.

Mongin, P. (1995). Consistent Bayesian aggregation. *Journal of Economic Theory*, *66*(2), 313–351.

Mongin, P. (1998). The paradox of the Bayesian experts and state-dependent utility theory. *Journal of Mathematical Economics*, *29*(3), 331–361.

Nash, J. (1950). The bargaining problem. *Econometrica*, *18*(2), 155–162.

O'Donoghue, T., & 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* (Working Paper).

Pivato, M. (2009). Twofold optimality of the relative utilitarian bargaining solution. *Social Choice and Welfare*, *32*(1), 79–92.

Pivato, M. (2014). Formal utilitarianism and range voting. *Mathematical Social Sciences*, *67*, 50–56.

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*, *38*(3), 393–409.

Sen, A. (1993). Internal consistency of choice. *Econometrica*, *61*(3), 495–521.

Smith, W. D. (2000). *Range voting* (Working Paper).

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., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ, USA: Princeton University Press.

Weymark, J. A. (1993). Harsanyi's social aggregation theorem and the weak Pareto principle. *Social Choice and Welfare*, *10*(3), 209–221.

Weymark, J. A. (1994). Harsanyi's social aggregation theorem with alternative Pareto principles. In W. Eichhorn (Ed.), *Models and Measurement of Welfare and Inequality* (pp. 869–887). Springer Berlin Heidelberg.

Zhou, L. (1997). Harsanyi's utilitarianism theorems: General societies. *Journal of Economic Theory*, *72*(1), 198–207.