# Belief-Free Preference Aggregation*

Satoshi Nakada† Shmuel Nitzan‡ Norio Takeoka§ Takashi Ui¶

August 2020

**Abstract**

A choice function, which assigns a feasible alternative to each state, is said to be belief-neutral efficient if it is ex-ante efficient under every reasonable belief such as a convex combination of individuals' beliefs (Brunnermeier et al., 2014). When every belief is considered to be reasonable, we refer to the criterion as belief-free efficiency. By definition, a choice function is belief-neutral efficient if and only if, for each reasonable belief, there exists a belief-dependent utilitarian social welfare function such that the choice function maximizes the expected social welfare under every reasonable belief. Can the utilitarian social welfare functions be taken to be common across all reasonable beliefs? We give the positive answer to this question and explore its implications on preference aggregation under belief-free settings. We show that a choice function is belief-free efficient if and only if it is rationalized by a utilitarian social preference relation. Using this result, we provide new axiomatic foundations for the aggregation of risk preferences and the separate aggregation of beliefs and risk preferences.

*Keywords: belief-neutral efficiency, belief-free efficiency, robust Pareto consistency, robust belief-wise consistency, utilitarian aggregation.*

*We are very grateful for valuable input from Susumu Cato, Larry Epstein, Marc Fleurbaey, Peter Hammond, Takashi Hayashi, Daisuke Hirata, Atsushi Kajii, So Kubota, Kaname Miyagishima, Xiangyu Qu, Naoki Yoshihara. We also thank seminar participants at Singapore Management University, SAET 2018, and EEA-ESEM 2019

†School of Management, Tokyo University of Science.

‡Department of Economics, Bar-Ilan University.

§Department of Economics, Hitotsubashi University.

¶Department of Economics, Hitotsubashi University.

# 1 Introduction

When individuals have heterogeneous subjective beliefs, the criterion of ex-ante efficiency is not so compelling due to the issue of spurious unanimity (Mongin, 2016). To illustrate it, consider a bet between Ann who believes that it will rain tomorrow and Bob who believes that it will be sunny tomorrow: Bob gives \$100 to Ann if it rains, and Ann gives \$100 to Bob otherwise. Although this bet is ex-ante efficient under the heterogeneous beliefs, it is driven by the difference in beliefs which cannot be correct simultaneously, and a win-win outcome is impossible, which questions its desirability. Thus, alternative efficiency criteria have been proposed (Gilboa et al., 2004; Gayer et al., 2014; Brunnermeier et al., 2014), incorporating common beliefs into ex-ante efficiency. One of them is belief-neutral efficiency of Brunnermeier, Simsek, and Xiong (2014) (hereafter BSX). BSX consider a social planner who has a set of reasonable beliefs over the set of states, assuming that reasonable beliefs are convex combinations of the individuals' subjective beliefs. A collective choice function, which assigns a feasible social alternative to each state (e.g., a feasible allocation of state-contingent claims), is said to be belief-neutral efficient if it is ex-ante efficient under every reasonable belief. As noted by BSX, we can incorporate an arbitrary set of reasonable beliefs into the above definition of belief-neutral efficiency.

For example, imagine that no information about states is available and the individuals' beliefs are unobservable or unreliable, so the social planner regards every belief as reasonable. We refer to this type of belief-neutral efficiency belief-free efficiency; that is, a choice function is belief-free efficient if it is ex-ante efficient under every belief. A belief-free efficient choice function avoids not only spurious unanimity but also the worst-case scenario of inefficiency under the true probability distribution, no matter what it is, implying that it is the most conservative type of belief-neutral efficiency. An example is a choice function that assigns the most preferred feasible alternative to each state with respect to an ex-post social preference relation represented by a utilitarian social welfare function, i.e., a choice function rationalized by a social preference relation with a utilitarian representation. As is well known, such a choice function is ex-ante efficient under every belief, i.e., belief-free efficient. However, it is unknown whether the converse also true. That is, some belief-free efficient choice function may not be rationalized by a social preference relation with a utilitarian representation. In that case, belief-free efficiency is a stronger requirement than utilitarianism. Otherwise, belief-free efficiency is equivalent to utilitarianism.

To study whether belief-free efficiency is equivalent to utilitarianism, we address the following more general question. BSX characterize belief-neutral efficiency in terms of a *collection* of utilitarian social welfare functions indexed by reasonable beliefs: a choice function is belief-neutral efficient if and only if, for every reasonable belief, there exists a belief-dependent utilitarian social welfare function such that the choice function maximizes the expected social welfare under the reasonable belief. This characterization is valid for an arbitrary set of reasonable beliefs. Thus, if we can take a common utilitarian social welfare function across all reasonable beliefs in BSX's characterization, then belief-free efficiency is equivalent to utilitarianism. Then, what condition guarantees the existence of such a common utilitarian social welfare function?

The purpose of this paper is not only to give an answer to the above question but also to explore their implications on preference aggregation under belief-free settings. Our first main result shows that if the set of states is finite and the set of reasonable beliefs is a convex hull of a finite set of beliefs, then we can take a common utilitarian social welfare function across all reasonable beliefs in BSX's characterization of belief-neutral efficiency; that is, a choice function is belief-neutral efficient if and only if there exists a belief-independent utilitarian social welfare function such that the choice function maximizes the expected social welfare under every belief in the finite set of beliefs. This implies that a choice function is belief-free efficient if and only if it is rationalized by a social preference relation with a utilitarian representation.

Using the notion of belief-free efficiency, we proceed to provide a new axiomatic foundation for a social preference relation with a utilitarian representation. Assume that, for every pair of social alternatives, there exists a state in which this pair is a menu. We say that a social preference relation satisfies *robust Pareto consistency* if a choice function rationalized by this social preference relation is belief-free efficient. This requirement is restated as follows: for any probability distribution over socially ranked pairs of alternatives, the corresponding lottery over the socially preferred alternatives is never Pareto-dominated by that over the socially unpreferred alternatives. Thus, if robust Pareto consistency is not satisfied, then, for some probability distribution over socially ranked pairs of alternatives, every individual prefers the lottery over the socially unpreferred alternatives to the lottery over the socially preferred alternatives, implying inconsistency between the social preference relation and the individuals' preference relations. Robust Pareto consistency requires that such inconsistency and the resulting inefficiency of the

<table>
<tr><td>(i)</td><td colspan="4">a utilitarian representation</td></tr>
<tr><td>(ii)</td><td colspan="3">a VNM representation</td><td>monotonicity</td></tr>
<tr><td>(iii)</td><td>continuity</td><td>transitivity</td><td>independence</td><td>monotonicity</td></tr>
<tr><td>(iv)</td><td>continuity</td><td colspan="3">belief-free efficiency = robust Pareto consistency</td></tr>
</table>

Table 1: Harsanyi's aggregation theorem vs. our result

choice functions should be avoided.

Our second main result establishes that a social preference relation satisfies robust Pareto consistency if and only if it has a utilitarian representation. This result holds for a social preference relation not only on a finite set of alternatives but also on an infinite set of alternatives such as the set of probability distributions (i.e. lotteries) over social outcomes under suitable topological assumptions. In the case of lotteries, robust Pareto consistency is shown to be equivalent to the set of three standard axioms, transitivity, monotonicity, and independence, which is well-defined only on the set of lotteries.

This result gives a new interpretation to Harsanyi's aggregation theorem (Harsanyi, 1955). Harsanyi's aggregation theorem states that a social preference relation on the set of lotteries is represented by a VNM utility function and satisfies monotonicity (i.e. Pareto) if and only if it has a utilitarian representation. Using VNM's axioms, we can rephrase Harsanyi's aggregation theorem as follows: a utilitarian representation is equivalent to the set of four axioms consisting of continuity, transitivity, independence, and monotonicity. Note that the set of the first three axioms is equivalent to a VNM representation, whereas the set of the last three axioms is equivalent to robust Pareto consistency (belief-free efficiency). Thus, our result rearranges the four axioms in Harsanyi's aggregation theorem and unifies the latter three into a single axiom of robust Pareto consistency. This makes the following difference between Harsanyi's aggregation theorem and our result: a VNM representation is the assumption in Harsanyi's aggregation theorem, whereas it is the consequence in our result. Table 1 summarizes the above discussion, where the equivalence of (i) and (ii) is Harsanyi's aggregation theorem, and the equivalence of (i) and (iv) is our result.

Our result also offers a useful extension of Harsanyi's aggregation theorem. In Harsanyi's theorem, society must have a complete preference relation on all lotteries that is represented by a VNM utility function. It is, however, very demanding due to the following two reasons. First, rankings of lotteries are not necessary to make a collective decision unless a randomized decision is adopted. There is also criticism on the assumption of expected social welfare maximization (Diamond, 1967). Furthermore, some pair of alternatives can be socially incomparable as emphasized by Sen (2004, 2018). By applying our results on belief-free efficiency and robust Pareto consistency, we can justify utilitarianism even when society does not have rankings of lotteries or rankings of some collections of alternatives, thus requiring less information than Harsanyi's aggregation theorem. In particular, to the best of our knowledge, no axiomatic foundation has been given to a utilitarian representation when the set of alternatives is finite.[1]

As another application of our second main result, we consider a social preference relation over Anscombe-Aumann acts (Anscombe and Aumann, 1963). When the individuals' preference relations over acts have subjective expected utility (SEU) representations, a social preference relation with an SEU representation cannot satisfy the requirement of Pareto condition unless the individuals have a common belief (Hylland and Zeckhauser, 1979; Mongin, 1995, 2016). To avoid this impossibility, follow-up papers weaken the Pareto condition and aggregate the individuals' risk preferences and beliefs (Gilboa et al., 2004; Hayashi and Lombardi, 2017; Billot and Qu, 2018). On the other hand, there is some debate about the extent to which the social planner should respect both of individuals' beliefs and tastes (e.g., Hammond, 1981) when individuals have either extreme risk attitudes or misperceived and distorted beliefs. In such a case, the social planner may want to aggregate the individuals' beliefs alone by relying on his own risk preference, or the social planner may want to aggregate the individuals' risk preferences alone by relying on his own belief.

We propose the notion of robust belief-wise consistency, which is mathematically interpreted as a special case of the notion of robust Pareto consistency. Using it, we characterize SEU representations of social preference relations aggregating beliefs alone, risk preferences alone, and both beliefs and risk preferences, respectively, where we assume monotonicity only over lotteries. In contrast to the existing literature, an SEU representation in our characterization is not an assumption but a consequence of robust belief-wise consistency, thus providing an alternative way to aggregate risk preferences and beliefs.

The organization of this paper is as follows. In Section 2, we introduce our setting. Section 3 characterizes belief-neutral and belief-free efficient choice functions. Section 4 characterizes a social preference relation satisfying robust Pareto consistency. In Section 5, robust

[1]There are papers obtaining utilitarianism assuming preferences on alternatives rather than lotteries (Blackorby et al., 1999, 2004; Fleurbaey, 2009; Mongin and Pivato, 2015), but these papers require that the set of alternatives is a non-atomic infinite set and that a social preference relation is complete. Thus, they do not obtain utilitarianism when the set of alternatives is finite.

Pareto consistency is applied to a social preference relation on lotteries, and our result is compared with Harsanyi's aggregation theorem. In Section 6, robust belief-wise consistency is applied to a social preference relation on Anscombe-Aumann acts, and our result is compared with the recent literature on utilitarian aggregation of both tastes and beliefs.

## 2 Preliminaries

A set of individuals $N = \{1, \dots, n\}$ constitutes society with a set of social alternatives $X$, which is assumed to be a finite set in Section 3 and a topological space in Section 4. The society faces uncertainty about the state of the world which determines the set of feasible social alternatives. To be more specific, let $\Omega$ be a finite set of states and assume that, when $\omega \in \Omega$ is realized, the set of feasible social alternatives is $X_\omega \subseteq X$ with $\#|X_\omega| \geq 2$, which is referred to as a menu in state $\omega$. A collective choice function is a mapping $c : \Omega \to X$ with $c(\omega) \in X_\omega$ for all $\omega \in \Omega$, which assigns a feasible social alternative to each state. The set of all collective choice functions is denoted by $\mathcal{C}$.

Individual $i \in N$ has a von Neumann-Morgenstern (VNM) utility function $u_i : X \to \mathbb{R}$, which is given as a primitive. We write $u(a) = (u_i(a))_{i \in N} \in \mathbb{R}^N$ for each $a \in X$ and regard $u : X \to \mathbb{R}^N$ as a mapping from the set of alternatives to the utility space $u(X)$. The set of all simple lotteries over $X$ is denoted by $\Delta(X)$. Each $a \in X$ is also regarded as a degenerate lottery. For each $A \subseteq X$, the set of all simple lotteries assigning probability one to $A$ is denoted by $\Delta(A)$. For a finite index set $K$, $\rho \in \Delta(K)$, and $\{a^k\}_{k \in K} \subseteq X$, let $\sum_{k \in K} \rho(k) a^k \in \Delta(X)$ denote the lottery in which $a^k$ is drawn with probability $\rho(k)$.

For $p \in \Delta(X)$ and $i \in N$, $u_i(p)$ denotes individual $i$'s expected utility of $p$; that is, $u_i(p) = \sum_{a \in X} u_i(a) p(a)$. For $p, q \in \Delta(X)$, if $u_i(p) \geq u_i(q)$ for each $i \in N$, then we write $u(p) \geq u(q)$ and say that $p$ weakly Pareto dominates $q$. If $u(p) \geq u(q)$ and $u(p) \neq u(q)$, then we write $u(p) > u(q)$ and say that $p$ Pareto dominates $q$.

We define efficiency of social alternatives as follows.

**Definition 1.** An alternative $a \in A \subseteq X$ is efficient in $A$ if no $p \in \Delta(A)$ Pareto dominates $a$. It is strictly efficient in $A$ if no $p \in \Delta(A) \setminus \{a\}$ weakly Pareto dominates $a$.

The following characterization of an efficient alternative is well known in various fields such as statistics, game theory, economics, and control theory among others, at the latest since

Wald (1950).[2]

**Lemma 1.** *Let $A \subseteq X$ be a finite subset of $X$. Then, $a \in A$ is efficient in $A$ if and only if there exists $w \in \mathbb{R}^N_{++}$ satisfying*

$$a \in \arg\max_{x \in A} \sum_{i \in N} w_i u_i(x). \tag{1}$$

*In particular, $a \in A$ is strictly efficient in $A$ if and only if there exists $w \in \mathbb{R}^N_{++}$ satisfying*

$$\{a\} = \arg\max_{x \in A} \sum_{i \in N} w_i u_i(x). \tag{2}$$

# 3 Belief-neutral and belief-free efficiency

Imagine that a state $\omega \in \Omega$ is drawn according to a probability distribution $\pi \in \Delta(\Omega)$ and that a social planner cares about ex-ante efficiency of a collective choice function $c \in C$. In this situation, the social planner considers the lottery induced by $c$ and compares it not only with a lottery induced by another collective choice function $c'$, $\sum_{\omega \in \Omega} \pi(\omega) c'(\omega)$, but also with a lottery induced by a randomized collective choice function $P \in \Delta(C)$, $\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) c'(\omega)$. We formally define ex-ante efficiency as follows.

**Definition 2.** A collective choice function $c$ is ex-ante efficient under $\pi \in \Delta(\Omega)$ if, for any $P \in \Delta(C)$, $\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) c'(\omega)$ does not Pareto dominate $\sum_{\omega \in \Omega} \pi(\omega) c(\omega)$. A collective choice function $c$ is ex-ante strictly efficient under $\pi \in \Delta(\Omega)$ if, for any $P \in \Delta(C)$ such that $P(c')\pi(\omega) > 0$ and $c'(\omega) \neq c(\omega)$ for some $c' \in C$ and $\omega \in \Omega$, $\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) c'(\omega)$ does not weakly Pareto dominate $\sum_{\omega \in \Omega} \pi(\omega) c(\omega)$.

Definition 2 conforms to Definition 1 in the following sense. Regard $C$ as a set of social alternatives, where $c, c' \in C$ are interpreted as the same social alternative if $c(\omega) = c'(\omega)$ for all $\omega \in \Omega$ with $\pi(\omega) > 0$. Assume that individual $i \in N$ has a VNM utility function $U_i$ with $U_i(c) = \sum_{\omega \in \Omega} \pi(\omega) u_i(c(\omega))$. Then, $c$ is (strictly) efficient in $C$ if and only if it is ex-ante (strictly) efficient under $\pi$. It is also straightforward to show that ex-ante (strict) efficiency under $\pi$ implies (strict) efficiency of $c(\omega)$ in $X_\omega$ for each $\omega \in \Omega$ with $\pi(\omega) > 0$.

A social planner is assumed to have a set of reasonable beliefs over $\Omega$ denoted by $\Pi \subseteq \Delta(\Omega)$. If there is an objective belief $\pi_0 \in \Delta(S)$, then $\Pi = \{\pi_0\}$. If there is no information about the true probability distribution and any belief is considered to be reasonable, then $\Pi = \Delta(\Omega)$. If

[2]For a proof, see Theorems 5.2.5 of Blackwell and Girshick (1954) or Geoffrion (1968).

individuals have subjective beliefs $\pi_1, \dots, \pi_n \in \Delta(\Omega)$ and they are observable, then the social planner may choose $\Pi = \mathbf{co}\{\pi_1, \dots, \pi_n\}$, i.e., the set of all convex combinations of $\{\pi_1, \dots, \pi_n\}$.

We say that $c$ is belief-neutral efficient under $\Pi$ if it is ex-ante efficient under any $\pi \in \Pi$. In particular, we say that $c$ is belief-free efficient if it is belief-neutral efficient under $\Delta(\Omega)$.

**Definition 3.** A collective choice function $c$ is belief-neutral efficient (or satisfies belief-neutral efficiency) under $\Pi$ if, for any $\pi \in \Pi$ and $P \in \Delta(C)$, $\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) c'(\omega)$ does not Pareto dominate $\sum_{\omega \in \Omega} \pi(\omega) c(\omega)$. A collective choice function $c$ is belief-free efficient (or satisfies belief-free efficiency) if it is belief-neutral efficient under $\Delta(\Omega)$.

**Definition 4.** A collective choice function $c$ is strictly belief-neutral efficient (or satisfies belief-neutral strict efficiency) under $\Pi$ if, for any $\pi \in \Pi$ and $P \in \Delta(C)$ such that $P(c')\pi(\omega) > 0$ and $c(\omega) \neq c'(\omega)$ for some $c' \in C$ and $\omega \in \Omega$, $\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) c'(\omega)$ does not weakly Pareto dominate $\sum_{\omega \in \Omega} \pi(\omega) c(\omega)$. A collective choice function $c$ is strictly belief-free efficient (or satisfies strict belief-free efficiency) if it is strictly belief-neutral efficient under $\Delta(\Omega)$.

The concept of belief-neutral efficiency is introduced by BSX in their study of a welfare criterion in financial markets. BSX assume that $\Pi = \mathbf{co}\{\pi_1, \dots, \pi_n\}$, but they also emphasize that the set of reasonable beliefs can be any set of beliefs, including the set of all beliefs. Thus, Definition 3 can be understood as the suggested extension of the original definition. In contrast to BSX, however, we take account of randomization in considering Pareto-dominance. In BSX, individuals are assumed to be risk-averse, and randomization is not considered.

By Lemma 1, a collective choice function $c$ is belief-neutral efficient under $\Pi$ if and only if, for each $\pi \in \Pi$, there exists $w(\pi) \in \mathbb{R}^N_{++}$ satisfying

$$\sum_{i \in N} w_i(\pi) \sum_{\omega \in \Omega} \pi(\omega) u_i(c(\omega)) \geq \sum_{i \in N} w_i(\pi) \sum_{\omega \in \Omega} \pi(\omega) u_i(c'(\omega)) \tag{3}$$

for every $\pi \in \Pi$ and $c' \in C$, which is BSX's characterization of belief-neutral efficiency. Note that the weight vector $w(\pi)$ depends upon $\pi$.

In the following characterization of belief-neutral efficiency, we show that we can choose a common weight vector if $\Pi$ is a convex hull of a finite set of probability distributions, thus refining BSX's characterization.

**Proposition 1.** *Suppose that $X$ is a finite set and $\Pi$ is a convex hull of a finite set $\Pi^* \subsetneq \Delta(\Omega)$. Then, a collective choice function $c$ is belief-neutral efficient under $\Pi$ if and only if there exists*

*$w \in \mathbb{R}^N_{++}$ satisfying*

$$\sum_{\omega\in\Omega} \pi(\omega) \sum_{i\in N} w_i u_i(c(\omega)) \geq \sum_{\omega\in\Omega} \pi(\omega) \sum_{i\in N} w_i u_i(c'(\omega)) \tag{4}$$

*for every $\pi \in \Pi^*$ and $c' \in C$. If $c$ is strictly belief-neutral efficient under $\Pi$, then there exists $w \in \mathbb{R}^N_{++}$ satisfying*

$$\sum_{\omega\in\Omega} \pi(\omega) \sum_{i\in N} w_i u_i(c(\omega)) > \sum_{\omega\in\Omega} \pi(\omega) \sum_{i\in N} w_i u_i(c'(\omega)) \tag{5}$$

*for every $\pi \in \Pi^*$ and $c' \in C$ with $c'(\omega) \neq c(\omega)$ for some $\omega \in \Omega$ with $\pi(\omega) > 0$.*

By regarding $\sum_{i\in N} w_i u_i$ as a utilitarian social welfare function on $X$, we can interpret Proposition 1 as follows: a collective choice function $c$ is belief-neutral efficient under $\Pi = \mathbf{co}\,\Pi^*$ if and only if there exists a utilitarian social welfare function $\sum_{i\in N} w_i u_i$ over $X$ such that $c$ maximizes the expected value of the social welfare function under every $\pi \in \Pi^*$. The connection between belief-neutral efficiency of a collective choice function and a utilitarian social welfare function is more evident in the case of belief-free efficiency; that is, a choice function is belief-free efficient if and only if it assigns the most preferred feasible alternative to each state with respect to a utilitarian social welfare function.

**Corollary 2.** *Suppose that $X$ is a finite set. Then, a collective choice function $c$ is belief-free efficient if and only if there exists $w \in \mathbb{R}^N_{++}$ satisfying*

$$c(\omega) \in \arg\max_{x\in X_\omega} \sum_{i\in N} w_i u_i(x) \textit{ for all } \omega \in \Omega. \tag{6}$$

*In particular, $c$ is strictly belief-free efficient if and only if there exists $w \in \mathbb{R}^N_{++}$ satisfying*

$$\{c(\omega)\} = \arg\max_{x\in X_\omega} \sum_{i\in N} w_i u_i(x) \textit{ for all } \omega \in \Omega. \tag{7}$$

In the remainder of this section, we prove Proposition 1 and Corollary 2. Define

$$D = \{\sum_{\omega\in\Omega} \pi(\omega)(u(c'(\omega)) - u(c(\omega))) \in \mathbb{R}^N : \pi \in \Pi^*,\ c' \in C\}.$$

Note that each $d \in D$ is an expected utility vector of a collective choice function $c'$ under $\pi \in \Pi^*$ that is normalized to $\mathbf{0} = (0, \ldots, 0) \in D$ for $c' = c$. We regard $D$ as a fictitious set of social alternatives and assume that individual $i \in N$ has a VNM utility function $v_i$ with $v_i(d) = d_i$ for all $d = (d_i)_{i \in N} \in D$. In the next lemma, we show that belief-neutral efficiency of $c$ is equivalent to efficiency of $\mathbf{0}$ in $D$.

**Lemma 2.** *A collective choice function $c$ is belief-neutral efficient if and only if $\mathbf{0}$ is efficient in $D$. If $c$ is strictly belief-neutral efficient, then $\mathbf{0}$ is strictly efficient in $D$.*

*Proof.* Let $\mathbf{co}\, D \subset \mathbb{R}^N$ be the convex hull of $D$, which should be distinguished from the set of all lotteries $\Delta(D)$. We first show that

$$\mathbf{co}\, D = \left\{ \sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega)(u(c'(\omega)) - u(c(\omega))) : P \in \Delta(C),\ \pi \in \Pi \right\}.$$

The set in the right-hand side is clearly a subset of $\mathbf{co}\, D$. Thus, it is enough to show that, for each $\sum_{k \in K} \lambda(k) d^k \in \mathbf{co}\, D$, where $d^k = \sum_{\omega \in \Omega} \pi^k(\omega)(u(c^k(\omega)) - u(c(\omega)))$, there exist $P \in \Delta(C)$ and $\pi \in \Pi$ such that

$$\sum_{k \in K} \lambda(k) d^k = \sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega)(u(c'(\omega)) - u(c(\omega))). \tag{8}$$

In fact, we can easily verify that $P$ and $\pi$ given below satisfy (8):

$$\pi(\omega) = \sum_{k \in K} \lambda(k) \pi^k(\omega),$$

$$P(c') = \begin{cases} \displaystyle\prod_{\omega \in \Omega,\ \pi(\omega) \neq 0} \frac{\sum_{k \in K,\ c^k(\omega) = c'(\omega)} \lambda(k) \pi^k(\omega)}{\pi(\omega)} & \text{if } c'(\omega) \in \{c^k(\omega)\}_{k \in K} \text{ for some } \omega \in \Omega, \\ 0 & \text{otherwise.} \end{cases}$$

Assume that $\mathbf{0}$ is not efficient in $D$. Then there exists a lottery $\sum_{k \in K} \rho(k) d^k \in \Delta(D)$ such that $(v_i(\sum_{k \in K} \rho(k) d^k))_{i \in N} = (\sum_{k \in K} \rho(k) d_i^k)_{i \in N} > 0$. Thus, for $P \in \Delta(C)$ and $\pi \in \Pi$ given above, we have

$$\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) u(c'(\omega)) > \sum_{\omega \in \Omega} \pi(\omega) u(c(\omega))), \tag{9}$$

which implies that $c$ is not belief-neutral efficient under $\Pi$. Conversely, assume that $c$ is not belief-neutral efficient under $\Pi$. Then, there exist $P \in \Delta(C)$ and $\pi \in \Pi$ satisfying (9), and thus $\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega)(u(c'(\omega)) - u(c(\omega))) > 0$, which implies that $\mathbf{0}$ is not efficient in $D$.

Assume that $\mathbf{0}$ is not strictly efficient in $D$. Then, there exists a lottery $\sum_{k \in K} \rho(k) d^k \in \Delta(D) \setminus \{\mathbf{0}\}$ such that $(v_i(\sum_{k \in K} \rho(k) d^k))_{i \in N} = (\sum_{k \in K} \rho(k) d_i^k)_{i \in N} \geq 0$. Thus, for $P \in \Delta(C)$ and $\pi \in \Pi$

given above, we have

$$\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) u(c'(\omega)) \geq \sum_{\omega \in \Omega} \pi(\omega) u(c(\omega))), \tag{10}$$

which implies that $c$ is not strictly belief-neutral efficient under $\Pi$. $\square$

We are ready to prove Proposition 1 and Corollary 2 using Lemmas 1 and 2.

*Proof of Proposition 1.* By Lemma 2, $c$ is belief-neutral efficient under $\Pi$ if and only if $\mathbf{0}$ is efficient in $D$. By Lemma 1, this is true if and only if there exists $w \in \mathbb{R}^N_{++}$ satisfying $\sum_{i \in N} w_i d_i \leq 0$ for all $d \in D$, which is rewritten as (4).

By Lemma 2, if $c$ is strictly belief-neutral efficient under $\Pi$, then $\mathbf{0}$ is strictly efficient in $D$. By Lemma 1, there exists $w \in \mathbb{R}^N_{++}$ satisfying $\sum_{i \in N} w_i d_i < 0$ for all $d \in D \setminus \{\mathbf{0}\}$. Thus, (5) holds for every $\pi \in \Pi^*$ and $c' \in C$ with $\sum_{\omega \in \Omega} \pi(\omega) u(c(\omega)) \neq \sum_{\omega \in \Omega} \pi(\omega) u(c'(\omega))$. Because $c$ is strictly belief-neutral efficient under $\Pi$, $\sum_{\omega \in \Omega} \pi(\omega) u(c(\omega)) = \sum_{\omega \in \Omega} \pi(\omega) u(c'(\omega))$ if and only if $c(\omega) = c'(\omega)$ for all $\omega \in \Omega$ with $\pi(\omega) > 0$. This implies that (5) holds for every $\pi \in \Pi^*$ and $c' \in C$ with $c'(\omega) \neq c(\omega)$ for some $\omega \in \Omega$ with $\pi(\omega) > 0$. $\square$

*Proof of Corollary 2.* When $\Pi = \Delta(\Omega)$, (4) is reduced to (6), and (5) is reduced to (7). To complete the proof, we show that (7) implies strict belief-free efficiency. By (7), $\sum_{i \in N} w_i u_i(c(\omega)) \geq \sum_{i \in N} w_i u_i(c'(\omega))$ for all $c' \in C$ and $\omega \in \Omega$, and $\sum_{i \in N} w_i u_i(c(\omega)) > \sum_{i \in N} w_i u_i(c'(\omega))$ if $c'(\omega) \neq c(\omega)$. Thus, for any $\pi \in \Pi$ and $P \in \Delta(C)$ such that $P(c')\pi(\omega) > 0$ and $c(\omega) \neq c'(\omega)$ for some $c' \in C$ and $\omega \in \Omega$,

$$\sum_{i \in N} w_i \sum_{\omega \in \Omega} \pi(\omega) u_i(c(\omega)) > \sum_{i \in N} w_i \sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) u_i(c'(\omega)),$$

which implies that $\sum_{c' \in C} P(c') \sum_{\omega \in \Omega} \pi(\omega) c'(\omega)$ cannot weakly Pareto dominate $\sum_{\omega \in \Omega} \pi(\omega) c(\omega)$. $\square$

# 4 Belief-free efficiency and robust Pareto consistency

In this section, we consider a social preference relation $\succsim$ over $X$ and explore the connection between strict belief-free efficiency and a utilitarian representation of $\succsim$. We say that $\succsim$ has a utilitarian representation, or $\succsim$ is utilitarian for short, if it is represented by a utilitarian social welfare function; that is, there exists $w \in \mathbb{R}^N_+$ satisfying

$$a \succsim b \iff \sum_{i \in N} w_i u_i(a) \geq \sum_{i \in N} w_i u_i(b) \text{ for all } a, b \in X. \tag{11}$$

Applying the notion of strict belief-free efficiency to a social preference relation $\succsim$ over $X$, we provide a new axiomatic foundation for a utilitarian representation.

To begin with, we demonstrate the equivalence of a utilitarian representation of an anti-symmetric social preference relation and strict belief-free efficiency of the collective choice function rationalized by the social preference relation. Let $X$ be a finite set and assume that (i) $\#|X_\omega| = 2$ for all $\omega \in \Omega$, (ii) for every pair $\{a, b\} \subset X$, there exists $\omega$ such that $X_\omega = \{a, b\}$, and (iii) $X_\omega = X_{\omega'}$ if and only if $\omega = \omega'$. Then, we can identify each state $\omega \in \Omega$ with a pair of social alternatives, which leads us to write

$$\Omega = \{\{a, b\} : a, b \in X\}. \tag{12}$$

For a collective choice function $c \in C$, let $\succsim$ be a complete anti-symmetric (i.e. strict)[3] binary relation on $X$ satisfying

$$c(\omega) = a \iff a \succsim b \text{ with } \omega = \{a, b\}. \tag{13}$$

Then, by Corollary 2, $c$ is strictly belief-free efficient if and only if $\succsim$ has a utilitarian representation (11) with a strictly positive weight $w \in \mathbb{R}^N_{++}$.

Motivated by the above observation, we redefine strict belief-free efficiency as a requirement for a complete anti-symmetric binary relation. In the following redefinition of strict belief-free efficiency, we do not have to consider randomized collective choice functions because each menu contains exactly two elements by (12).

**Lemma 3.** *Let $c \in C$ and $\succsim$ be a collective choice function and a complete anti-symmetric binary relation on $X$ satisfying* (13)*. Then, $c$ is strictly belief-free efficient if and only if, for any collection of socially ranked pairs of alternatives $(a^k, b^k)_{k \in K}$ such that $a^k \succsim b^k$ (i.e., $a^k \succ b^k$ or $a^k = b^k$) for all $k \in K$ and $a^k \succ b^k$ for at least one $k \in K$ and any probability distribution $\rho \in \Delta(K)$ with full support, the lottery $\sum_{k \in K} \rho(k) b^k$ over the socially unpreferred alternatives does not weakly Pareto dominate the lottery $\sum_{k \in K} \rho(k) a^k$ over the socially preferred alternatives.*

[3]A binary relation $\succsim$ is anti-symmetric if $a \succsim b$ and $b \succsim a$ imply $a = b$.

*Proof.* It is enough to show that $c$ is strictly belief-free efficient if and only if, for any $\pi \in \Delta(\Omega)$ and $c' \in C$ such that $c(\omega) \neq c'(\omega)$ for at least one $\omega \in \Omega$ with $\pi(\omega) > 0$, $\sum_{\omega\in\Omega} \pi(\omega)c'(\omega)$ does not weakly Pareto dominate $\sum_{\omega\in\Omega} \pi(\omega)c(\omega)$. By Definition 4, it is also enough to show the "if" part of this statement. Suppose that $c$ is not strictly belief-free efficient. By definition, there exist $\pi \in \Delta(\Omega)$ and $P \in \Delta(C)$ such that $P(c')\pi(\omega) > 0$ and $c(\omega) \succ c'(\omega)$ for some $c' \in C$ and $\omega \in \Omega$, and $\sum_{c'\in C} P(c') \sum_{\omega\in\Omega} \pi(\omega)c'(\omega)$ weakly Pareto dominates $\sum_{\omega\in\Omega} \pi(\omega)c(\omega)$. Let $c^* \in C$ be such that $c^*(\omega) \neq c(\omega)$ for all $\omega$, which is uniquely determined because each menu contains exactly two social alternatives. Let $\pi^* \in \Delta(\Omega)$ be given by

$$\pi^*(\omega) = \frac{\pi(\omega) \sum_{c'(\omega)\neq c(\omega)} P(c')}{\sum_{\omega'\in\Omega} \pi(\omega') \sum_{c'(\omega')\neq c(\omega')} P(c')}.$$

Then,

$$\sum_{c'\in C} P(c') \sum_{\omega\in\Omega} \pi(\omega)(u_i(c'(\omega)) - u_i(c(\omega))) = \frac{\sum_{\omega\in\Omega} \pi^*(\omega)(u_i(c^*(\omega)) - u_i(c(\omega)))}{\sum_{\omega'\in\Omega} \pi(\omega') \sum_{c'(\omega')\neq c(\omega')} P(c')}.$$

This implies that $\sum_{\omega\in\Omega} \pi^*(\omega)c^*(\omega)$ weakly Pareto dominates $\sum_{\omega\in\Omega} \pi^*(\omega)c(\omega)$. □

Now let $\succsim$ be a complete binary relation on a topological space $X$ and call it a social preference relation on $X$. We consider a social preference relation satisfying the condition in Lemma 3, which is referred to as robust Pareto consistency.

**Definition 5.** A social preference relation $\succsim$ on $X$ satisfies robust Pareto consistency if, for any finite collection of socially ranked pairs of alternatives $(a^k, b^k)_{k\in K}$ such that $a^k \succsim b^k$ for all $k \in K$ and $a^k \succ b^k$ for at least one $k \in K$ and any probability distribution $\rho \in \Delta(K)$ with full support, the lottery $\sum_{k\in K} \rho(k)b^k$ over the socially unpreferred alternatives does not weakly Pareto dominate the lottery $\sum_{k\in K} \rho(k)a^k$ over the socially preferred alternatives.

Recall that if $X$ is finite and $\succsim$ is anti-symmetric, then robust Pareto consistency is equivalent to a utilitarian representation by Corollary 2 and Lemma 3. We will show that a similar claim is true even if $X$ is not finite and $\succsim$ is not anti-symmetric.

To understand the normative implication of robust Pareto consistency, suppose that $\succsim$ does not satisfy robust Pareto consistency. Then, for some collection of socially ranked pairs of alternatives and some probability distribution over the collection, every individual prefers the lottery over the socially unpreferred alternatives to the lottery over the socially preferred alternatives, implying inconsistency between the social preference relation and the individuals' preference relations. This inconsistency implies inefficiency of a collective choice function which selects a socially preferred alternative from every pair of alternatives. Robust Pareto consistency requires that such inconsistency and inefficiency are avoided.

Definition 5 immediately implies monotonicity and transitivity of $\succsim$.

**Lemma 4.** *Suppose that $\succsim$ satisfies robust Pareto consistency. Then, the following holds.*

(i) *$\succsim$ satisfies monotonicity; that is, $a \succsim b$ whenever $a$ weakly Pareto dominates $b$.*

(ii) *$\succsim$ satisfies transitivity.*

*Proof.* As a direct consequence of the requirement when $\#|K| = 1$, if $a \succ b$, then $b$ does not weakly Pareto dominate $a$, which implies monotonicity. To prove transitivity, suppose otherwise that $\succsim$ does not satisfy transitivity. Then, there exist $x, y, z \in X$ such that $x \succsim y$, $y \succsim z$, and $z \succ x$. Let $a^1 = x$, $a^2 = y$, $a^3 = z$, $b^1 = y$, $b^2 = z$, and $b^3 = x$. By robust Pareto consistency, $\sum_{k=1}^{3} 1/3 \cdot b^k$ does not weakly Pareto dominate $\sum_{k=1}^{3} 1/3 \cdot a^k$. However,

$$u_i(\sum_{k=1}^{3} 1/3 \cdot a^k) = 1/3 \cdot (u_i(x) + u_i(y) + u_i(z)) = u_i(\sum_{k=1}^{3} 1/3 \cdot b^k),$$

which is a contradiction. Therefore, $\succsim$ satisfies transitivity. □

We first establish the necessity of robust Pareto consistency for a utilitarian representation.

**Proposition 3.** *If $\succsim$ is utilitarian, then it satisfies robust Pareto consistency.*

*Proof.* Suppose that (11) holds. For $\sum_{k\in K} \rho(k)b^k$ and $\sum_{k\in K} \rho(k)a^k$ given in Definition 5,

$$\sum_{i\in N} w_i u_i(p) = \sum_{k\in K} \rho(k) \sum_{i\in N} w_i u_i(a^k) > \sum_{k\in K} \rho(k) \sum_{i\in N} w_i u_i(b^k) = \sum_{i\in N} w_i u_i(q)$$

because there exists $k \in K$ such that $a^k \succ b^k$, i.e., $\rho(k) \sum_{i\in N} w_i u_i(a^k) > \rho(k) \sum_{i\in N} w_i u_i(b^k)$. Thus, $\sum_{k\in K} \rho(k)b^k$ cannot weakly Pareto dominate $\sum_{k\in K} \rho(k)a^k$. □

If $X$ is a finite set, then robust Pareto consistency is not only necessary but also sufficient for a utilitarian representation, which generalizes the above discussion in the case of an anti-symmetric social preference relation. To the best of our knowledge, no axiomatic foundation has been given to a utilitarian representation with a finite set of social alternatives

**Proposition 4.** *Suppose that $X$ is a finite set. Then, $\succsim$ satisfies robust Pareto consistency if and only if it is utilitarian.*

*Proof.* See Appendix A. □

If $X$ is an infinite set, robust Pareto consistency alone does not imply a utilitarian representation. In fact, if $\succsim$ is utilitarian, then $\succsim$ must satisfy the standard continuity axiom, and moreover, it must satisfy an analogous axiom concerning the utility space $u(X)$.

**Definition 6.** A social preference relation $\succsim$ satisfies continuity if, for each $x \in X$, the upper contour set $\{y \in X : y \succsim x\}$ and the lower contour set $\{y \in X : x \succsim y\}$ are closed.

**Definition 7.** A social preference relation $\succsim$ satisfies utility space continuity if, for each $x \in X$, the upper contour utility set $\{u(y) \in u(X) : y \succsim x\}$ and the lower contour utility set $\{u(y) \in u(X) : x \succsim y\}$ are closed with respect to the relative topology of $u(X) \subset \mathbb{R}^N$.

We use the latter axiom to obtain a utilitarian representation. Yet, if $u$ is a closed or open mapping, then the standard continuity axiom implies the utility space continuity axiom. For example, if $X$ is a compact space, then $u$ is a closed mapping by the closed map lemma.

**Lemma 5.** *Suppose that $X$ is a compact space and utility functions are continuous. If $\succsim$ satisfies continuity, then it also satisfies utility space continuity.*

*Proof.* According to the closed map lemma, every continuous function from a compact space to a Hausdorff space is closed. Thus, $u$ is a closed mapping. This implies that $\{u(y) : y \succsim x\}$ and $\{u(y) : x \succsim y\}$ are closed because they are the images of $\{y \in X : y \succsim x\}$ and $\{y \in X : x \succsim y\}$, respectively, and these sets are closed by continuity. □

We also use the following topological property of $u(X)$ possessed by a convex or open set. It is well-known that a convex set $C \subset \mathbb{R}^N$ is a subset of the closure of the relative interior of $C$; that is, $C \subseteq \mathbf{cl}\,\mathbf{ri}\,C$, where $\mathbf{cl}$ is the closure operator and $\mathbf{ri}$ is the relative interior operator.[4] An open set $C \subseteq \mathbb{R}^N$ also has this property because $C = \mathbf{int}\,C \subset \mathbf{cl}\,\mathbf{int}\,C$ and $\mathbf{int}\,C = \mathbf{ri}\,C$, where $\mathbf{int}$ is the interior operator. The following proposition states that if $u(X)$ has this property, then utility space continuity and robust Pareto consistency implies a utilitarian representation.

**Proposition 5.** *Suppose that $u(X)$ is a connected set satisfying $u(X) \subseteq \mathbf{cl}\,\mathbf{ri}\,u(X)$. Then, a social preference relation $\succsim$ satisfies utility space continuity and robust Pareto consistency if and only if it is utilitarian.*

*Proof.* See Appendix A. □

[4]For example, see Rockafellar (1996, Theorem 6.3).

In particular, Proposition 5 together with Lemma 5 implies the following corollary.

**Corollary 6.** *Suppose that $X$ is a compact space and that $u(X)$ is a convex set. Then, a social preference relation $\succsim$ satisfies continuity and robust Pareto consistency if and only if it is utilitarian.*

# 5 Social preferences over lotteries

As an application of the results in Section 4, we consider a social preference relation over lotteries and compare our results with Harsanyi's aggregation theorem (Harsanyi, 1955).

Let $Z$ be a finite set of social outcomes. We consider a social preference relation $\succsim$ on $X \equiv \Delta(Z)$, where generic elements of $X$ are denoted by $p$, $q$, and $r$. As in the previous sections, individual $i \in N$ has a continuous VNM utility function $u_i : X \to \mathbb{R}$. A special case is a VNM utility function defined on $Z$, where $\Delta(X) = \Delta(\Delta(Z))$ is identified with $\Delta(Z)$ through the reduction of compound lotteries. In this case, $u_i$ is linear, i.e., $\lambda u_i(p) + (1 - \lambda)u(q) = u_i(\lambda p + (1 - \lambda)q)$ for all $p, q \in X$ and $\lambda \in [0, 1]$.

To emphasize the difference between nonlinear and linear VNM utility functions, let us restate the definition of robust Pareto consistency as follows.

**Definition 8.** A social preference relation $\succsim$ on $\Delta(Z)$ satisfies robust Pareto consistency if, for any finite collection of socially ranked pairs of lotteries $(p^k, q^k)_{k\in K}$ such that $p^k \succsim q^k$ for all $k \in K$ and $p^k \succ b^k$ for at least one $k \in K$ and any probability distribution $\rho \in \Delta(K)$ with full support, $\sum_{k\in K} \rho(k)q^k$ does not weakly Pareto dominate $\sum_{k\in K} \rho(k)p^k$.

It should be noted that $\sum_{k\in K} \rho(k)p^k$ and $\sum_{k\in K} \rho(k)q^k$ are treated as compound lotteries rather than reduced lotteries unless $u_i$ is linear. That is, $\sum_{k\in K} \rho(k)q^k$ does not weakly Pareto dominate $\sum_{k\in K} \rho(k)p^k$ if and only if $\sum_{k\in K} \rho(k)u(q^k) \not\gtrsim \sum_{k\in K} \rho(k)u(p^k)$, which may be different from $u(\sum_{k\in K} \rho(k)q^k) \not\gtrsim u(\sum_{k\in K} \rho(k)p^k)$. This is because individuals may have different risk attitudes toward different sources of uncertainty, i.e., uncertainty about menus $\rho \in \Delta(K)$ and uncertainty about objective lotteries $p^k, q^k \in \Delta(Z)$ (Segal, 1990).

Because $\Delta(Z)$ is a connected and compact set and $u$ is continuous, $u(\Delta(Z))$ is connected, and continuity implies utility space continuity by Lemma 5. Thus, we can apply Proposition 5 to obtain the following characterization of a utilitarian representation.

**Corollary 7.** *Suppose that $u(X) \subseteq \mathbf{cl}\,\mathbf{ri}\,u(X)$. Then, a social preference relation $\succsim$ on $\Delta(Z)$ satisfies continuity and robust Pareto consistency if and only if it is utilitarian.*

A social welfare function characterized by this corollary can exhibit inequality aversion as illustrated by the following example.

**Example 1.** Two individuals 1 and 2 have VNM utility functions on $X$ given by $u_1(p) = \phi(v_1(p))$ and $u_2(p) = \phi(v_2(p))$, respectively, where $v_1$ and $v_2$ are VNM utility functions on $Z$, and $\phi$ is a strictly increasing function. The set of social alternatives is $Z = \{z_1, z_2\}$ with $(v_1(z_1), v_2(z_1)) = (1, 0)$ and $(v_1(z_2), v_2(z_2)) = (0, 1)$. The social welfare function on $\Delta(Z)$ is

$$W(p) = \phi(v_1(p)) + \phi(v_2(p)),$$

which is characterized by Corollary 7. Compare $p = z_1$ and $q = 1/2 \cdot z_1 + 1/2 \cdot z_2$. Then, $W(p) = \phi(1) + \phi(0)$ and $W(q) = \phi(1/2) + \phi(1/2)$. If $\phi$ is linear, then $W(p) = W(q)$ and thus $p$ and $q$ are socially indifferent, which is criticized by Diamond (1967) as the limitation of utilitarianism characterized by Harsanyi (1955). In contrast, if $\phi$ is strictly concave, then $W(q) > W(p)$ and thus $q$ is socially preferred to $q$. Note that inequality aversion here comes from the individuals' risk attitudes rather than the social planner's fairness concerns. Concavity of $\phi$ means that an individual prefers a later resolution of compound lotteries; that is, he or she is more risk averse about the first stage lotteries over $\Delta(Z)$ than the second stage lotteries over $Z$. Thus, we do not explicitly consider the social planner's fairness concerns, which is in contrast to the models of Epstein and Segal (1992) and Grant et al. (2010).

In the remainder of this section, we focus on a special case of linear VNM utility functions, i.e., VNM utility functions on $Z$. In this case, the two compound lotteries in Definition 8 are treated as reduced lotteries. Because $X$ is a compact set and $u(X)$ is a convex set, we can apply Corollary 6 to obtain the following characterization of a utilitarian representation.

**Corollary 8.** *Suppose that each individual has a VNM utility function on $Z$. Then, a social preference relation $\succsim$ on $\Delta(Z)$ satisfies continuity and robust Pareto consistency if and only if it is utilitarian.*

Note that a utilitarian representation characterized by Corollary 8 is an expected utility (or

VNM) representation because

$$\sum_{i\in N} w_i u_i(p) = \sum_{i\in N} w_i \left( \sum_{x\in X} p(x) u_i(x) \right) = \sum_{x\in X} p(x) \left( \sum_{i\in N} w_i u_i(x) \right),$$

where $\sum_{i\in N} w_i u_i(x)$ is a VNM index for the social expected utility. That is, a VNM representation of $\succsim$ is a consequence of Corollary 8.

Contrastingly, a VNM representation of $\succsim$ is an assumption of Harsanyi's aggregation theorem as stated below.

**Proposition 9.** *Suppose that a social preference relation $\succsim$ on $\Delta(Z)$ has a VNM representation. Then, $\succsim$ satisfies monotonicity if and only if it is utilitarian.*

It is well-known that a VNM representation is equivalent to VNM's set of axioms consisting of transitivity, continuity, and independence given below.

**Definition 9.** A social preference relation $\succsim$ on $\Delta(Z)$ satisfies independence if for $p, q, r \in \Delta(Z)$ and $\lambda \in [0, 1]$,

$$p \succsim q \iff \lambda p + (1 - \lambda) r \succsim \lambda q + (1 - \lambda) r.$$

Because Corollary 8 and Proposition 9 characterize the same utilitarian representation, the condition in Corollary 8 is equivalent to that in Proposition 9. Thus, we obtain the following summary of Corollary 8 and Proposition 9.

**Corollary 10.** *The following four statements are equivalent.*

(i) $\succsim$ *on* $\Delta(Z)$ *is utilitarian.*

(ii) $\succsim$ *on* $\Delta(Z)$ *satisfies continuity, transitivity, independence, and monotonicity.*

(iii) $\succsim$ *on* $\Delta(Z)$ *has a VNM representation and satisfies monotonicity.*

(iv) $\succsim$ *on* $\Delta(Z)$ *satisfies continuity and robust Pareto consistency.*

As suggested by Corollary 10, robust Pareto consistency plays the role of transitivity, independence, and monotonicity when $\succsim$ satisfies continuity. This is true even without continuity.

**Lemma 6.** *A social preference relation $\succsim$ on $\Delta(Z)$ satisfies robust Pareto consistency if and only if it satisfies transitivity, independence, and monotonicity.*

*Proof.* See Appendix B. □

Therefore, Corollary 8 provides a new interpretation of Harsanyi's aggregation theorem by unifying the axioms of transitivity, independence, and monotonicity into the single axiom of robust Pareto consistency, i.e., belief-free efficiency. Table 1 summarizes the relationship between the axioms in the four statements of Corollary 10.

Before closing this section, we also consider a social preference relation on $X \subsetneq \Delta(Z)$, of which a special case is $X = Z$ consisting of degenerate lotteries. Because $\Delta(X) \subseteq \Delta(Z)$ by the reduction of compound lotteries, we can regard the restriction of $u_i$ to $X$ as a VNM utility function on $X$. Thus, we can apply the results in the previous sections to obtain the following characterization of utilitarianism in terms of robust Pareto consistency.

- Suppose that $X$ is a finite set. Then, $\succsim$ satisfies robust Pareto consistency if and only if it is utilitarian by Corollary 4.
- Suppose that $X$ is a convex set. Then, $\succsim$ satisfies continuity and robust Pareto consistency if and only if it is utilitarian by Proposition 5 because $u$ is a closed mapping.
- Suppose that $u(X)$ is a connected open set. Then, $\succsim$ satisfies continuity and robust Pareto consistency if and only if it is utilitarian by Proposition 5 because $u$ is a closed mapping.
- Suppose that $X$ is a finite set and that $\succsim$ is an incomplete strict social preference relation on $X$. Let $c$ be a collective choice function such that $c(\omega) \succ x$ for all $x \in X_\omega \setminus \{c(\omega)\}$. Then, $c$ is strictly belief-free efficient if and only if $c$ has a utilitarian representation (7) by Corollary 2.

In Harsanyi's aggregation theorem, society must have a complete social preference relation that is represented by a VNM social welfare function, which requires $X = \Delta(Z)$. Contrastingly, our results allow society to have not only a social preference relation over the set of all lotteries but also a (possibly) incomplete social preference relation on a finite set of lotteries or alternatives. In this sense, our characterization of utilitarianism in terms of robust Pareto consistency is more general and versatile than Harsanyi's aggregation theorem.

# 6 Social preferences over acts

As an application of the results in Section 4, we consider a social preference relation over Anscombe-Aumann acts (Anscombe and Aumann, 1963). When the individuals' preference relations over acts have subjective expected utility (SEU) representations, a social preference relation with an SEU representation cannot satisfy the Pareto condition unless the individuals have a common belief (Hylland and Zeckhauser, 1979; Mongin, 1995, 2016). To avoid this impossibility, follow-up papers weaken the Pareto condition and aggregate the individuals' risk preferences and beliefs (Gilboa et al., 2004; Hayashi and Lombardi, 2017; Billot and Qu, 2018) by assuming an SEU representation of a social preference relation.

In this section, we introduce the notion of robust belief-wise consistency, which is mathematically interpreted as a special case of the notion of robust Pareto consistency. Using it, we characterize SEU representations of social preference relations aggregating beliefs alone, risk preferences alone, and both beliefs and risk preferences, respectively, where monotonicity is assumed only over lotteries. In contrast to the existing literature, an SEU representation in our characterization is not an assumption but a consequence of robust belief-wise consistency, thus providing an alternative way to aggregate risk preferences and beliefs.

## 6.1 Aggregation of beliefs

Let $S$ and $Z$ be, respectively, a finite set of states and a finite set of social outcomes. An act $f$ assigns a lottery $f(s) \in \Delta(Z)$ to each state $s \in S$. We consider a social preference relation $\succsim$ on the set of acts, which is denoted by

$$X \equiv \Delta(Z)^S.$$

Each lottery $p \in \Delta(Z)$ is regarded as a constant act which assigns the same lottery $p$ to every state. As discussed in Section 4, $\succsim$ can be associated with a collective choice function on $\{\{f, g\} : f, g \in X\}$, but we do not explicitly consider such a collective choice function. It should be noted that an act on $S$ and a collective choice function on $\{\{f, g\} : f, g \in X\}$ are totally different objects.

To evaluate acts, a social planner has a set of reasonable beliefs over $S$, which is denoted by $\Gamma \subseteq \Delta(S)$. If there is an objective belief $\gamma_0 \in \Delta(S)$, then $\Gamma = \{\gamma_0\}$. If every belief is considered to be possible, then $\Gamma = \Delta(S)$. If all individuals' subjective beliefs $\gamma_1, \dots, \gamma_n \in \Delta(S)$ are observable, then the social planner may choose $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ or $\Gamma = \mathbf{co}\{\gamma_1, \dots, \gamma_n\}$, as assumed by BSX.

Consider a randomization over acts, i.e., $P \in \Delta(X)$, and take any reasonable belief $\gamma \in \Gamma$. Then, we have a three-stage compound lottery in which (i) an act $f \in X$ is drawn according to

$P$, (ii) a lottery $f(s) \in \{f(s')\}_{s' \in S}$ is drawn according to $\gamma$, and (iii) a social outcome $z \in Z$ is drawn according to the lottery $f(s)$. Let

$$\sum_{f \in X} P(f) \sum_{s \in S} \gamma(s) f(s) \in \Delta(Z)$$

be the lottery reduced from this compound lottery.

For $P, Q \in \Delta(X)$, we say that $P$ weakly belief-wise dominates $Q$ under $\Gamma$ if, for all $\gamma \in \Gamma$, the lottery reduced from the compound lottery consisting of $P$ and $\gamma$ is socially preferred to that consisting of $Q$ and $\gamma$; that is, for all $\gamma \in \Gamma$,

$$\sum_{f \in X} P(f) \sum_{s \in S} \gamma(s) f(s) \succsim \sum_{f \in X} Q(f) \sum_{s \in S} \gamma(s) f(s). \tag{14}$$

As a special case, let $\Gamma$ be the set of all degenerate probability distributions. Then, $P$ weakly belief-wise dominates $Q$ under $\Gamma$ if and only if, for all $s \in S$,

$$\sum_{f \in X} P(f) f(s) \succsim \sum_{f \in X} Q(f) f(s).$$

In this case, we say that $P$ weakly state-wise dominates $Q$. Because (14) is a linear inequality, $P$ weakly state-wise dominates $Q$ if and only if $P$ weakly belief-wise dominates $Q$ under $\Delta(S) = \mathbf{co}\,\Gamma$. More generally, the following lemma holds.

**Lemma 7.** *Suppose that $\Gamma = \mathbf{co}\,\Gamma^* \subseteq \Delta(S)$. Then, for $P, Q \in \Delta(X)$, $P$ weakly belief-wise dominates $Q$ under $\Gamma$ if and only if $P$ weakly belief-wise dominates $Q$ under $\Gamma^*$.*

To introduce the notion of robust belief-wise consistency, we start with a special case. Consider two acts $f, g \in X$ such that $g$ weakly belief-wise dominates $f$ under $\Gamma$. Note that $g$ always gives a socially preferred reduced lottery over outcomes than $f$ no matter what reasonable belief over $S$ is taken into account. In this case, we require that $g$ should be socially preferred to $f$. Equivalently, if $f$ is strictly socially preferred to $g$, then we require that $g$ should not weakly belief-wise dominate $f$.

In the following definition of robust belief-wise consistency, we extend the above requirement to randomizations over acts. Let $(f^k, g^k)_{k \in K}$ be a collection of socially ranked pairs of acts such that $f^k \succsim g^k$ for all $k \in K$ and $f^k \succ g^k$ for at least one $k \in K$. Consider randomizations over $\{f^k\}_{k \in K}$ and $\{g^k\}_{k \in K}$, respectively, in terms of a common probability distribution $\rho \in \Delta(K)$

with full support: $\sum_{k\in K}\rho(k)f^k \in \Delta(X)$ draws a socially preferred act $f^k$ with probability $\rho(k)$ and $\sum_{k\in K}\rho(k)g^k \in \Delta(X)$ draws a socially unpreferred act $g^k$ with the same probability.

**Definition 10.** A social preference relation $\succsim$ on $X$ satisfies robust belief-wise consistency under $\Gamma$ if, for any finite collection of socially ranked pairs of acts $(f^k, g^k)_{k\in K}$ such that $f^k \succsim g^k$ for all $k \in K$ and $f^k \succ g^k$ for at least one $k \in K$ and any probability distribution $\rho \in \Delta(K)$ with full support, $\sum_{k\in K}\rho(k)g^k$ does not weakly belief-wise dominate $\sum_{k\in K}\rho(k)f^k$ under $\Gamma$.

To get an intuition, suppose that $\sum_{k\in K}\rho(k)g^k$ weakly belief-wise dominates $\sum_{k\in K}\rho(k)f^k$ under $\Gamma$. That is, $\sum_{k\in K}\rho(k)g^k$ is socially preferred to $\sum_{k\in K}\rho(k)f^k$ under every $\gamma \in \Gamma$, which is understood as ex-ante social preferences before $k$ realizes. For each $k \in K$, however, $f^k$ is socially preferred to $g^k$ for every $k \in K$, which is understood as ex-post social preferences after $k$ realizes. Therefore, ex-ante social preferences are inconsistent with ex-post social preferences, i.e., the social preference relation is not dynamically consistent in terms of every belief in $\Gamma$. Robust belief-wise consistency is a normative axiom requiring dynamic consistency in this sense.

To verify robust belief-wise consistency under $\Gamma$, we have to take account of every belief in $\Gamma$. However, if $\Gamma = \mathbf{co}\,\Gamma^*$, then it is enough to consider beliefs in $\Gamma^*$ by Lemma 7.

**Lemma 8.** *Suppose that $\Gamma = \mathbf{co}\,\Gamma^* \subseteq \Delta(S)$. Then, a social preference relation $\succsim$ on $X$ satisfies robust belief-wise consistency under $\Gamma$ if and only if it satisfies robust belief-wise consistency under $\Gamma^*$.*

Robust belief-wise consistency and robust Pareto consistency are inspired by the same idea. To see this, assume the following.

- When $\succsim$ is restricted to $\Delta(Z)$, it has a VNM representation; that is, there exists a VNM utility function $u_0$ on $Z$ such that $p \succsim q$ if and only if $u_0(p) \geq u_0(q)$ for all $p, q \in \Delta(Z)$.
- Individual $i \in N$ has the same VNM utility function and a subjective belief $\gamma_i \in \Delta(S)$ such that he or she prefers $f$ to $g$ if and only if $\sum_{s\in S}\gamma_i(s)u_0(f(s)) \geq \sum_{s\in S}\gamma_i(s)u_0(g(s))$.

Then, it is easy to see that robust Pareto consistency under these assumptions is equivalent to belief-wise consistency under $\{\gamma_1, \ldots, \gamma_n\}$.

Using the above observation and Corollary 6, we show that if $\Gamma$ is a convex hull of a finite set, then robust belief-wise consistency under $\Gamma$ together with continuity is equivalent to an SEU representation with some belief in $\Gamma$.

**Proposition 11.** *Suppose that $\Gamma = \mathbf{co}\,\Gamma^*$, where $\Gamma^* \subset \Delta(S)$ is a finite set. Then, a social preference relation $\succsim$ on $X$ satisfies continuity and robust belief-wise consistency under $\Gamma^*$ if and only if it has an SEU utilitarian representation with a subjective belief $\gamma \in \Gamma$; that is, there exist a VNM utility function $u_0$ on $Z$ and $\gamma \in \Gamma$ satisfying*

$$f \succsim g \iff \sum_{s\in S} \gamma(s)u_0(f(s)) \geq \sum_{s\in S} \gamma(s)u_0(g(s)) \text{ for all } f, g \in X. \tag{15}$$

*In particular, when $\Gamma^* = \{\gamma_1, \ldots, \gamma_n\}$ and $\gamma_i \in \Delta(S)$ is a subjective belief of individual $i \in N$, $\gamma$ in* (15) *is a convex combination of the individuals' beliefs.*

*Proof.* See Appendix C. □

To explain Proposition 11, consider the special case with $\Gamma^* = \{\gamma_1, \ldots, \gamma_n\}$. Then, $\gamma$ in (15) is understood as aggregating the individuals beliefs. On the other hand, $u_0$ in (15) is interpreted as the social planner's VNM utility function. Thus, Proposition 11 allows the social planner to use his own risk attitude, while aggregating the individuals' subjective beliefs into the social belief $\gamma \in \Gamma$.

Proposition 11 also admits another interpretation. Consider a decision maker with a preference relation $\succsim$ over acts, who is not necessarily a social planner. The decision maker has objective but imprecise information about states, which is represented by the set of probability distributions $\Gamma$. A typical example is an Ellsberg's urn, where the decision maker is partially informed about the composition of colors in the urn. Proposition 11 states that robust belief-wise consistency is essential to ensure an SEU representation with a subjective belief that is consistent with the objective but imprecise information.

## 6.2 Aggregation of beliefs and tastes

Proposition 11 is also useful to study aggregation of both beliefs and risk preferences. Once we admit belief aggregation, we can obtain preference aggregation with a fairly mild assumption. Suppose that every individual $i \in N$ has a VNM utility function $u_i$ on $Z$. We say that a social preference relation on $X$ satisfies risk preference monotonicity if $\succsim$ restricted to the set of lotteries $\Delta(Z)$ satisfies monotonicity.

**Definition 11.** A social preference relation $\succsim$ on $X$ satisfies risk preference monotonicity if, for all $p, q \in \Delta(Z)$, $p \succsim q$ whenever $p$ weakly Pareto dominates $q$.

We show that risk preference monotonicity together with the axioms in Proposition 11 is equivalent to an SEU utilitarian representation.

**Proposition 12.** *Suppose that $\Gamma = \mathbf{co}\,\Gamma^*$, where $\Gamma^* \subset \Delta(S)$ is a finite set. Then, a social preference relation $\succsim$ on $X$ satisfies continuity, robust belief-wise consistency under $\Gamma^*$, and risk preference monotonicity if and only if it has an SEU utilitarian representation with a subjective belief $\gamma \in \Gamma$; that is, there exist $w \in \mathbb{R}^N_+$ and $\gamma \in \Gamma$ satisfying*

$$f \succsim g \iff \sum_{s\in S} \gamma(s) \sum_{i\in N} w_i u_i(f(s)) \geq \sum_{s\in S} \gamma(s) \sum_{i\in N} w_i u_i(g(s)) \text{ for all } f, g \in X. \tag{16}$$

*In particular, when $\Gamma^* = \{\gamma_1, \ldots, \gamma_n\}$ and $\gamma_i \in \Delta(S)$ is a subjective belief of individual $i \in N$, $\gamma$ in* (16) *is a convex combination of the individuals' beliefs.*

*Proof.* See Appendix C. □

To explain Proposition 12, consider the special case with $\Gamma^* = \{\gamma_1, \ldots, \gamma_n\}$ again. Then, the social belief $\gamma$ in (16) aggregates the individuals beliefs. Moreover, the social VNM utility function in (16) is a linear combination of the individuals' VNM utility functions, which is understood as aggregating the individuals' risk preferences. Thus, a social preference relation with the representation (16) aggregates both beliefs and risk preferences.

Gilboa et al. (2004) were the first to provide an axiomatic foundation for utilitarian aggregation of tastes and beliefs, followed by many papers including Hayashi and Lombardi (2017) and Billot and Qu (2018). To aggregate tastes and beliefs, these papers adopt weaker criteria of ex-ante Pareto efficiency because the standard criterion of ex-ante Pareto efficiency is inconsistent with utilitarian preference aggregation and it causes a problem of spurious unanimity (Mongin, 2016). Contrastingly, we introduce the axiom of robust belief-wise consistency, which requires dynamic consistency of social preferences under heterogeneous beliefs, thus providing an alternative axiomatic foundation for utilitarian aggregation of preferences and beliefs.

To elaborate on the difference between the approach in the above literature and our approach, let us discuss the key axiom in Hayashi and Lombardi (2017) and Billot and Qu (2018). They adopt the following weaker criterion of ex-ante Pareto dominance proposed by Gayer et al. (2014). For $f, g \in X$, $f$ is said to weakly unanimity Pareto dominates $g$ if

$$\sum_{s\in S} \gamma_j(s) u_i(f(s)) \geq \sum_{s\in S} \gamma_j(s) u_i(g(s))$$

for all $i, j \in N$, where $\gamma_j \in \Delta(S)$ is individual $j$'s subjective belief, respectively. That is, whenever the individuals share some individual's belief, they unanimously prefer $f$ to $g$. Using this criterion, Hayashi and Lombardi (2017) and Billot and Qu (2018) introduce the following axiom.[5]

**Definition 12.** A social preference relation $\succsim$ on $X$ satisfies collective ex-ante monotonicity if $f \succsim g$ whenever $f$ weakly unanimity Pareto dominates $g$.

Using collective ex-ante monotonicity, Hayashi and Lombardi (2017) characterize a maxmin representation of a social preference relation on a domain which is different from ours. On the other hand, Billot and Qu (2018) characterize the same SEU utilitarian representation as that in the special case of Proposition 12 with $\Gamma^* = \{\gamma^1, \ldots, \gamma^n\}$, as stated in the following proposition.

**Proposition 13.** *Suppose that a social preference relation $\succsim$ on $X$ has an SEU representation. Then, $\succsim$ satisfies collective ex-ante monotonicity if and only if it has an SEU utilitarian representation* (16)*, where $\gamma$ is a convex combination of the individuals' beliefs.*

Billot and Qu (2018) assume an SEU representation and ask what additional axiom implies the representation (16) with $\gamma \in \mathbf{co}\{\gamma_1, \ldots, \gamma_n\}$, where beliefs and tastes are aggregated simultaneously. This is a standard approach in the literature, and their answer is collective ex-ante monotonicity. In contrast, we disentangle the aggregation process; that is, we first establish aggregation of beliefs, and then proceed to aggregation of both beliefs and tastes by assuming risk preference monotonicity. Put it differently, we assume risk preference monotonicity as a minimum requirement and ask what additional axiom implies the representation (16) with $\gamma \in \mathbf{co}\{\gamma_1, \ldots, \gamma_n\}$. Our answer is robust belief-wise consistency, and an SEU representation is its consequence rather than the assumption, which is in clear contrast to Billot and Qu (2018).

## 6.3 Robust state-wise consistency

As a final special case, let $\Gamma$ be the set of all beliefs. Because $\Gamma$ is the convex hull of all degenerate beliefs, robust belief-wise consistency under $\Gamma$ is equivalent to robust state-wise consistency defined below, which is an immediate consequence of Lemma 8.

**Definition 13.** A social preference relation $\succsim$ on $X$ satisfies robust state-wise consistency under $\Gamma$ if, for any finite collection of socially ranked pairs of acts $(f^k, g^k)_{k \in K}$ such that $f^k \succsim g^k$ for all

[5]Hayashi and Lombardi (2017) call this condition collective ex-ante Pareto. Billot and Qu (2018) call this condition belief-proof Pareto.

$k \in K$ and $f^k \succ g^k$ for at least one $k \in K$ and any probability distribution $\rho \in \Delta(K)$ with full support, $\sum_{k \in K} \rho(k) g^k$ does not weakly state-wise dominate $\sum_{k \in K} \rho(k) f^k$ under $\Gamma$.

The following corollary of Propositions 11 and 12 characterizes a social preference relation satisfying robust state-wise consistency.

**Corollary 14.** *A social preference relation $\succsim$ on $X$ satisfies continuity and robust state-wise consistency if and only if it has an SEU representation* (15) *with $\gamma \in \Delta(S)$. Moreover, $\succsim$ on $X$ satisfies continuity, robust state-wise consistency, and risk preference monotonicity if and only if it has an SEU utilitarian representation* (16) *with $w \in \mathbb{R}^N_+$ and $\gamma \in \Delta(S)$.*

The second statement of Corollary 14 characterizes a social preference relation aggregating risk preferences alone. In fact, the social VNM utility function in (16) aggregates the individuals' risk preferences, whereas the social belief in (16) is independent of the individuals' beliefs.

The first statement of Corollary 14 provides an alternative axiomatic foundation for an SEU representation of a preference relation on Anscombe-Aumann acts. In the following corollary, we combine it with the result of Anscombe and Aumann (1963).

**Corollary 15.** *Let $\succsim$ be a complete binary relation on $X$. Then, the following three statements are equivalent.*

(i) *$\succsim$ has an SEU representation* (15) *with $\gamma \in \Delta(S)$.*

(ii) *$\succsim$ satisfies continuity and robust state-wise consistency.*

(iii) *$\succsim$ satisfies continuity, transitivity, independence, and state-wise monotonicity in the following sense: $f(s) \succsim g(s)$ for all $s \in S$ implies $f \succsim g$.*

The equivalence of (i) and (iii) is the result of Anscombe and Aumann (1963), and that of (i) and (ii) is Corollary 14. By the equivalence of (ii) and (iii), we find that robust state-wise consistency plays the role of transitivity, independence, and state-wise monotonicity when $\succsim$ satisfies continuity. This is true even without continuity.

**Lemma 9.** *A social preference relation $\succsim$ on $X$ satisfies robust state-wise consistency if and only if it satisfies transitivity, independence, and state-wise monotonicity.*

*Proof.* See Appendix C. □

| (i) | an SEU representation | | | |
|---|---|---|---|---|
| (ii) | continuity | transitivity | independence | state-wise monotonicity |
| (iii) | continuity | robust belief-wise consistency | | |

Table 2: Ansombe and Aumanns's theorem vs. our result

The above discussion is summarized in Table 2. It is now clear that robust state-wise consistency in Corollary 15 corresponds to robust Pareto consistency in Corollary 10. In fact, robust state-wise consistency implies robust Pareto consistency when we regard $S$ and $u_0(f(s))$ as a fictitious set of individuals and a VNM utility function of individual $s \in S$ defined on $X$, respectively.

# Appendix A Proofs for the results in Section 4

In Section 4, we study a weak social preference relation, whereas, in Sections 2 and 3, we study a collective choice function, which corresponds to a strict social preference relation. Thus, we cannot directly use the results in Sections 2 and 3 to prove the results in Section 4.

Let

$$V_I \equiv \{(u_i(a) - u_i(b))_{i\in N} \in \mathbb{R}^N : a \sim b\},\ V_P \equiv \{(u_i(a) - u_i(b))_{i\in N} \in \mathbb{R}^N : a \succ b\}$$

denote the set of utility difference vectors for indifferent pairs of alternatives and that for strictly ranked pairs of alternatives, respectively. Using $V_I$ and $V_P$, we can rephrase the requirement of robust Pareto consistency as follows.

**Lemma A.** *If $\succsim$ satisfies robust Pareto consistency, then*

$$(\mathbf{aff}\, V_I + \mathbb{R}^N_-) \cap \mathbf{co}\, V_P = \emptyset, \tag{17}$$

*where* $\mathbf{aff}\, V_I$ *is the affine hull of* $V_I$ *and* $\mathbf{aff}\, V_I + \mathbb{R}^N_- = \{v + t : v \in \mathbf{aff}\, V_I,\ t \in \mathbb{R}^N_-\}$.

*Proof.* Suppose that $(\mathbf{aff}\, V^I + \mathbb{R}^N_-) \cap \mathbf{co}\, V^P \neq \emptyset$. Then, there exists $v_I \in \mathbf{aff}\, V^I$, $v_P \in \mathbf{co}\, V_P$, and $t \in \mathbb{R}^N_-$ such that $v_I + t = v_P$, i.e., $v_P - v_I \leq 0$. Note that $v_P$ and $v_I$ are represented as follows:

$$v_P = \sum_{k=1}^{M} \lambda^k v_P^k,\ v_I = \sum_{k=M+1}^{K} \lambda^k v_I^k,$$

where $v_P^1, \ldots, v_P^M \in V_P$, $v_I^{M+1}, \ldots, v_I^K \in V_I$, $\lambda^1, \ldots, \lambda^M > 0$ with $\sum_{k=1}^M \lambda^k = 1$, and $\lambda^{M+1}, \ldots, \lambda^K \in \mathbb{R} \setminus \{0\}$ with $\sum_{k=M+1}^K \lambda^k = 1$. Without loss of generality, assume that $\lambda^k < 0$ if and only if $M + 1 \leq k \leq M'$. Then, we can rewrite $v_I$ as follows:

$$v_I = -\sum_{k=M+1}^{M'} |\lambda^k| v_I^k - \sum_{k=M'+1}^{K} |\lambda^k|(-v_I^k).$$

Plugging the above into $v_P - v_I \leq 0$, we obtain

$$v_P - v_I = \sum_{k=1}^{M} \lambda^k v_P^k + \sum_{k=M+1}^{M'} |\lambda^k| v_I^k + \sum_{k=M'+1}^{K} |\lambda^k|(-v_I^k) \leq 0.$$

Define $v^1, \ldots, v^K \in V_I \cup V_P$ and $\rho \in \Delta(K)$ as follows:

$$v^k = \begin{cases} v_P^k & \text{if } 1 \leq k \leq M, \\ v_I^k & \text{if } M + 1 \leq k \leq M', \\ -v_I^k & \text{if } M' + 1 \leq k \leq K, \end{cases}$$

$$\rho(k) = |\lambda^k| / \sum_{k'=1}^{K} |\lambda^{k'}| \text{ for all } k.$$

Then, $\sum_{k=1}^K \rho(k) v^k = (v_P - v_I)/\sum_{k'=1}^K |\lambda^{k'}| \leq 0$. Consider $\sum_{k=2}^K \rho(k)/(1-\rho(1)) v^k \in \mathbf{co}\,(V_I \cup V_P)$. By Carathéodory's theorem, there exists at most $n + 1$ points $w^2, \ldots, w^{n+2} \in V_I \cup V_P$ such that $\sum_{k=2}^{n+2} \pi(k) w^k = \sum_{k=2}^K \rho(k)/(1-\rho(1)) v^k$. Thus, $\rho(1) v^1 + (1 - \rho(1)) \sum_{k=2}^{n+2} \pi(k) w^k \leq 0$, which implies that $\succsim$ does not satisfy robust Pareto consistency. □

Using Lemma A, we provide a partial characterization of a utilitarian representation in terms of robust Pareto consistency.

**Proposition A.** *If $\succsim$ satisfies robust Pareto consistency, then there exists $w \in \mathbb{R}_+^N$ satisfying*

$$a \succsim b \;\Rightarrow\; \sum_{i \in N} w_i u_i(a) \geq \sum_{i \in N} w_i u_i(b) \;\; \textit{for all } a, b \in X. \tag{18}$$

*Proof.* Suppose that $\succsim$ satisfies robust Pareto consistency. By Lemma A and the separating hyperplane theorem, there exists $w \in \mathbb{R}^N$ and $\lambda \in \mathbb{R}$ such that, for all $v_I \equiv (v_{I,i})_{i \in N} \in \mathbf{aff}\, V_I + \mathbb{R}_-^N$ and $v_P \equiv (v_{P,i})_{i \in N} \in \mathbf{co}\, V_P$,

$$\sum_{i \in N} w_i v_{I,i} \leq \lambda \leq \sum_{i \in N} w_i v_{P,i}. \tag{19}$$

Because $v_{I,i}$ is unbounded from below for each $i$ and $\sum_{i \in N} w_i v_{I,i}$ is bounded from above, we must have $w \in \mathbb{R}_+^N$. This implies that

$$\sum_{i \in N} w_i v_i \leq 0 \text{ for all } v \in \mathbb{R}_-^N. \tag{20}$$

Moreover, because $\mathbf{aff}\, V_I$ is a linear subspace and $\sum_{i \in N} w_i(\alpha v_i) = \alpha \sum_{i \in N} w_i v_i$ is bounded from above for all $v \in \mathbf{aff}\, V_I$ and $\alpha \in \mathbb{R}$, we must have

$$\sum_{i \in N} w_i v_i = 0 \text{ for all } v \in \mathbf{aff}\, V_I. \tag{21}$$

By (20) and (21), we can choose $\lambda = 0$ in (19), which together with (21) implies that

$$\sum_{i \in N} w_i v_i \geq 0 \text{ for all } v \in V_I \cup V_P \subset \mathbf{aff}\, V_I \cup \mathbf{co}\, V_P.$$

Therefore, (18) holds. □

To obtain a utilitarian representation, we must also have

$$a \succ b \;\Rightarrow\; \sum_{i \in N} w_i u_i(a) > \sum_{i \in N} w_i u_i(b) \;\text{ for all } a, b \in X, \tag{22}$$

but robust Pareto consistency alone does not suffice. Propositions 4 and 5 give sufficient conditions for (22), of which proofs are given below.

*Proof of Proposition 4.* By Propositions 3 and A, it is enough to prove (22). Assume that $\succsim$ satisfies robust Pareto consistency. Because $X$ is a finite set, $\mathbf{co}\, V_P$ is a compact set. Thus, $\mathbf{aff}\, V_I + \mathbb{R}_-^N$ and $\mathbf{co}\, V_P$ are strictly separated by the separating hyperplane theorem. That is, there exists $w \in \mathbb{R}^N$ and $\lambda \in \mathbb{R}$ such that, for all $v_I \equiv (v_{I,i})_{i \in N} \in \mathbf{aff}\, V_I + \mathbb{R}_-^N$ and $v_P \equiv (v_{P,i})_{i \in N} \in \mathbf{co}\, V_P$,

$$\sum_{i \in N} w_i v_{I,i} \leq \lambda < \sum_{i \in N} w_i v_{P,i}. \tag{23}$$

By applying the same argument as that in the proof of Proposition A, we obtain $w \in \mathbb{R}_+^N$ and $\lambda = 0$, so

$$\sum_{i \in N} w_i v_i > 0 \text{ for all } v \in \mathbf{co}\, V_P,$$

which implies (22). □

*Proof of Proposition 5.* By Propositions 3 and A, it is enough to prove (22). Assume that $\succsim$ satisfies utility space continuity and robust Pareto consistency.

Let $S = \textbf{aff}\, u(X) - u^0$ with $u^0 \in u(X)$, which is a $k$-dimensional linear subspace (i.e., $k$ is the affine dimension of $u(X)$). Note that $\textbf{aff}\,(V_I \cup V_P) = S$ and thus $\textbf{aff}\, V_I$ is a linear subspace of $S$ and $\textbf{co}\, V_P$ is a convex subset of $S$. Let $\Psi : S \to \mathbb{R}^k$ be a linear bijection from $S$ to $\mathbb{R}^k$. For each $v \in S$, $\Psi_i(v)$ denotes the $i$-th element of $\Psi(v) = (\Psi_i(v))_{i=1}^k \in \mathbb{R}^k$.

By Proposition A and its proof, there exists $w \in \mathbb{R}^N_+$ such that $\sum_{i \in N} w_i v_i \geq 0$ for all $v \in \textbf{co}\, V_P$. Thus, there exists $\bar{w} \in \mathbb{R}^k$ such that $\sum_{i \in N} w_i v_i = \sum_{i=1}^k \bar{w}_i \Psi_i(v) \geq 0$ for all $v \in \textbf{co}\, V_P$ because $\Psi$ is a linear bijection. This implies that

$$\sum_{i=1}^k \bar{w}_i \bar{v}_i \geq 0 \text{ for all } \bar{v} \in \Psi(\textbf{co}\, V_P). \tag{24}$$

If the equality never holds in (24), i.e.,

$$\sum_{i=1}^k \bar{w}_i \bar{v}_i > 0 \text{ for all } \bar{v} \in \Psi(\textbf{co}\, V_P), \tag{25}$$

then $\sum_{i \in N} w_i v_i = \sum_{i=1}^k \bar{w}_i \Psi_i(v) > 0$ for all $v \in \textbf{co}\, V_P$, which implies (22) and completes the proof.

It is enough to establish (25) for $\bar{v} = \Psi(v)$ with $v = u(a) - u(b) \in V_P$ because $\Psi(\textbf{co}\, V_P) = \textbf{co}\, \Psi(V_P)$. First, suppose that $\{u(a), u(b)\} \cap \textbf{ri}\, u(X) \neq \emptyset$. Then, $v \in \textbf{ri}\,\textbf{co}\, V_P$.[6] This implies that $\bar{v} \in \textbf{int}\,\textbf{co}\, \Psi(V_P) \subset S$, and the interior point of a convex set cannot be on the separating hyperplane, so (24) implies (25). Next, suppose that $\{u(a), u(b)\} \cap \textbf{ri}\, u(X) = \emptyset$. Note that $\{u(y) : a \succ y \succ b\}$ is nonempty and open with respect to the relative topology of $u(X)$ because it is the complement of the union of two closed sets with no intersection, $\{u(y) : a \succsim y\}$ and $\{u(y) : y \succsim b\}$, and $u(X)$ is connected. Thus, there exists an open set $O \subset \mathbb{R}^N$ such that $O \cap u(X) = \{u(y) : a \succ y \succ b\}$. Because $u(X) \subseteq \textbf{cl}\,\textbf{ri}\, u(X)$, there exists $c \in X$ such that $u(c) \in (O \cap u(X)) \cap \textbf{ri}\, u(X) = \{u(y) : a \succ y \succ b\} \cap \textbf{ri}\, u(X)$. Then, $u(a) - u(c), u(c) - u(b) \in \textbf{ri}\,\textbf{co}\, V_P$, and thus $\sum_{i=1}^k \bar{w}_i \bar{v}_i = \sum_{i=1}^k \bar{w}_i \Psi_i(u(a) - u(c)) + \sum_{i=1}^k \bar{w}_i \Psi_i(u(c) - u(b)) > 0$. □

[6] To see this, suppose that $u(a) \in \textbf{ri}\, u(X)$. Then, there exists $\varepsilon > 0$ such that $N_\varepsilon(u(a)) \cap \textbf{aff}\, u(X) \subset u(X)$ where $N_\varepsilon(u(a))$ is a $\varepsilon$-neighborhood of $u(a)$. By utility space continuity, the set $SU_\succ(u(b)) = \{u(y) \in U(X) | y \succ b\}$ is open with respect to the relative topology of $u(X)$. Since $u(a) \in SU_\succ(u(b))$, there exists $\varepsilon' \in (0, \varepsilon)$ such that $N_{\varepsilon'}(u(a)) \cap u(X) \subset SU_\succ(u(b))$. Hence, $N_{\varepsilon'}(u(a)) \cap u(X) - u(b) \subset \textbf{co}\, V_p$. Moreover, for any $v' \in N_{\varepsilon'}(v) \cap \textbf{aff}\,\textbf{co}\, V_p$, by construction, $v' \in N_{\varepsilon'}(u(a)) \cap u(X) - u(b)$. Therefore, $N_{\varepsilon'}(v) \cap \textbf{aff}\,\textbf{co}\, V_p \subset \textbf{co}\, V_p$, which implies that $v \in \textbf{ri}\,\textbf{co}\, V_P$. A similar argument can be applied for $u(b) \in \textbf{ri}\, u(X)$.

# Appendix B Proofs for the results in Section 5

*Proof of Lemma 6.* Suppose that $\succsim$ satisfies transitivity, monotonicity, and independence. Let $(p^k, q^k)_{k\in K}$ be a collection of socially ranked pairs of lotteries such that $p^k \succsim q^k$ for all $k \in K$ and $p^k \succ q^k$ for at least one $k \in K$. For any $\rho \in \Delta(k)$ with full support, $\sum_{k\in K} \rho(k)p^k \succ \sum_{k\in K} \rho(k)q^k$ by transitivity and independence. By monotonicity, $\sum_{k\in K} \rho(k)q^k$ does not weakly Pareto dominate $\sum_{k\in K} \rho(k)p^k$. Thus, $\succsim$ satisfies robust Pareto consistency.

Conversely, assume that $\succsim$ satisfies robust Pareto consistency. By Lemma 4, $\succsim$ satisfies transitivity and monotonicity. To show independence of $\succsim$, it is enough to show that

$$p \succsim q \;\Rightarrow\; \lambda p + (1-\lambda)r \succsim \lambda q + (1-\lambda)r, \tag{26}$$

$$p \succ q \;\Leftarrow\; \lambda p + (1-\lambda)r \succ \lambda q + (1-\lambda)r. \tag{27}$$

Suppose that (26) does not hold. Then, there exist $p, q, r \in \Delta(Z)$ and $\lambda \in [0, 1]$ such that $p \succsim q$ and $\lambda q + (1-\lambda)r \succ \lambda p + (1-\lambda)r$. Define

$$\bar{p} = \frac{\lambda}{1+\lambda}p + \frac{1}{1+\lambda}(\lambda q + (1-\lambda)r) \text{ and } \bar{q} = \frac{\lambda}{1+\lambda}q + \frac{1}{1+\lambda}(\lambda p + (1-\lambda)r). \tag{28}$$

Then, $\bar{p} = \bar{q}$, and thus $\bar{q}$ weakly Pareto dominates $\bar{p}$, which implies that $\succsim$ does not satisfy robust Pareto consistency.

Suppose that (27) does not hold. Then, there exist $p, q, r \in \Delta(Z)$ and $\lambda \in [0, 1]$ such that $q \succsim p$ and $\lambda p + (1-\lambda)r \succ \lambda q + (1-\lambda)r$. For $\bar{p}$ and $\bar{q}$ defined in (28), $\bar{p} = \bar{q}$, and thus $\bar{p}$ weakly Pareto dominates $\bar{q}$, which implies that $\succsim$ does not satisfy robust Pareto consistency. □

# Appendix C Proofs for the results in Section 6

*Proof of Proposition 11.* Because it is straightforward to check the "if" part, we prove the "only if" part. We can easily verify that robust belief-wise consistency implies transitivity and independence of $\succsim$ restricted on $\Delta(Z)$, which is also discussed in Lemma 9. Thus, $\succsim$ has a VNM representation with a VNM utility function $u_0$ on $Z$. Define $U(f, \gamma) = u_0(\sum_{s\in S} \gamma(s)f(s))$ for $f \in X$ and $\gamma \in \Gamma^*$. Then,

$$\sum_{k\in K} \rho(k) \sum_{s\in S} \gamma(s) f^k(s) \succsim \sum_{k\in K} \rho(k) \sum_{s\in S} \gamma(s) g^k(s) \iff \sum_{k\in K} \rho(k) U(f^k, \gamma) \geq \sum_{k\in K} \rho(k) U(g^k, \gamma).$$

Regard $\Gamma^*$ as a fictitious group of individuals, where individual $\gamma \in \Gamma^*$ has a VNM utility function on $X$ given by $U(f, \gamma)$ for all $f \in X$. Then, robust belief-wise consistency implies robust Pareto consistency in terms of the fictitious group of individuals $\Gamma^*$. Thus, by Corollary 6, there exists $\lambda \in \Delta(\Gamma^*)$ satisfying

$$f \succsim g \iff \sum_{\gamma \in \Gamma^*} \lambda(\gamma) U(f, \gamma) \geq \sum_{\gamma \in \Gamma^*} \lambda(\gamma) U(g, \gamma)$$

because $X$ is compact and $\{(U(f, \gamma))_{\gamma \in \Gamma^*} : f \in X\} \subset \mathbb{R}^{\Gamma^*}$ is convex. Therefore, we obtain (15) with $\gamma = \sum_{\gamma' \in \Gamma^*} \lambda(\gamma')\gamma' \in \mathbf{co}\,\Gamma^*$. □

*Proof of Proposition 12.* Because it is straightforward to check the "if" part, we prove the "only if" part. By Proposition 11, there exist a VNM utility function $u_0$ on $Z$ and $\gamma \in \Gamma$ satisfying (15). Because $\succsim$ has a VNM representation on $\Delta(Z)$ and satisfies risk preference monotonicity, $\succsim$ has a utilitarian representation on $\Delta(Z)$ by Harsanyi's aggregation theorem; that is, there exists $w \in \mathbb{R}^N_+$ such that $\sum_{i \in N} w_i u_i$ is a VNM utility function of $\succsim$. By replacing $u_0$ with $\sum_{i \in N} w_i u_i$ in (15), we obtain (16). □

*Proof of Lemma 9.* Regard $S$ as a fictitious set of individuals. Individual $s \in S$ has the following preference relation on $\Delta(X)$: for $P, Q \in \Delta(X)$, individual $s$ prefers $P$ to $Q$ if and only if $\sum_{f \in X} P(f) f(s) \succsim \sum_{f \in X} Q(f) f(s)$. Then, $P$ weakly state-wise dominates $Q$ if and only if $P$ weakly Pareto dominates $Q$. Thus, $\succsim$ satisfies state-wise monotonicity if and only if it satisfies monotonicity in terms of the fictitious set of individuals. Moreover, $\succsim$ satisfies robust state-wise consistency if and only if it satisfies robust Pareto consistency in terms of the fictitious set of individuals. Therefore, Lemma 9 is implied by the proof of Lemma 6, which does not use the individuals' VNM utility functions.

For completeness, we provide a proof. Suppose that $\succsim$ satisfies transitivity, independence, and state-wise monotonicity. Let $(f^k, g^k)_{k \in K}$ be a collection of socially ranked pairs of acts such that $f^k \succsim g^k$ for all $k \in K$ and $f^k \succ g^k$ for at least one $k \in K$. For any $\rho \in \Delta(K)$ with full support, $\sum_{k \in K} \rho(k) f^k \succ \sum_{k \in K} \rho(k) g^k$ by transitivity and independence. By state-wise monotonicity, $\sum_{k \in K} \rho(k) g^k$ does not weakly state-wise dominate $\sum_{k \in K} \rho(k) f^k$. Thus, $\succsim$ satisfies robust state-wise consistency.

Suppose that $\succsim$ satisfies state-wise consistency. If $f \succ g$, then $g$ does not weakly state-wise dominate $f$. Thus, if $g$ weakly state-wise dominates $f$, i.e., $g(s) \succsim f(s)$ for all $s \in S$, then we must have $g \succsim f$, which implies state-wise monotonicity.

To prove transitivity, suppose that $\succsim$ does not satisfy transitivity. Then, there exist $f, g, h \in X$ such that $f \succsim g$, $g \succsim h$, and $h \succ f$. Let $f^1 = f$, $f^2 = g$, $f^3 = h$, $g^1 = g$, $g^2 = h$, and $g^3 = f$. Note that, for each $s \in S$, $\sum_{k=1}^{3} 1/3 \cdot f^k(s) = \sum_{k=1}^{3} 1/3 \cdot g^k(s)$. Therefore, $\sum_{k=1}^{3} 1/3 \cdot g^k$ weakly state-wise dominates $\sum_{k=1}^{3} 1/3 \cdot f^k$, which implies that $\succsim$ does not satisfy robust belief-wise consistency.

To prove independence, we show that

$$f \succsim g \;\Rightarrow\; \lambda f + (1-\lambda)h \succsim \lambda g + (1-\lambda)h, \tag{29}$$

$$f \succ g \;\Leftarrow\; \lambda f + (1-\lambda)h \succ \lambda g + (1-\lambda)h. \tag{30}$$

Suppose that (29) does not hold. Then, there exist $f, g, h \in X$ and $\lambda \in [0, 1]$ such that $f \succsim g$ and $\lambda g + (1-\lambda)h \succ \lambda f + (1-\lambda)h$. Define

$$P = \frac{\lambda}{1+\lambda} f + \frac{1}{1+\lambda}(\lambda g + (1-\lambda)h) \text{ and } Q = \frac{\lambda}{1+\lambda} g + \frac{1}{1+\lambda}(\lambda f + (1-\lambda)h). \tag{31}$$

Then, $P = Q$, and thus $Q$ weakly belief-wise dominates $P$, which implies that $\succsim$ does not satisfy robust belief-wise consistency.

Suppose that (30) does not hold. Then, there exist $f, g, h \in X$ and $\lambda \in [0, 1]$ such that $g \succsim f$ and $\lambda f + (1-\lambda)g \succ \lambda g + (1-\lambda)h$. For $P$ and $Q$ defined in (31), $P = Q$, and thus $P$ weakly Pareto dominates $Q$, which implies that $\succsim$ does not satisfy robust belief-wise consistency. □

# References


Anscombe, F. J,. Aumann, R. J., 1963. A definition of subjective probability. Ann. Math. Stat. 34, 199–205.

Billot, A., Qu, X., 2018. Utilitarian aggregation with reasonably heterogeneous beliefs. Working paper.

Blackorby, C., Donaldson, D., Weymark, J. A., 1999. Harsanyi's social aggregation theorem for state-contingent alternatives. J. Math. Econ. 32, 365–387.

Blackorby, C., Donaldson, D., Mongin, P., 2004. Social aggregation without the expected utility hypothesis. Cahier no 2004-020 du Laboratoire d'Econométrie, Ecole Polytechnique.

Blackwell, D., Girshick, M. A., 1954. Theory of Games and Statistical Decisions. John Wiley & Sons.

Brunnermeier, M., Simsek, A., Xiong, W., 2014. A welfare criterion for models with distorted beliefs. Q. J. Econ. 129, 1753–1797.

Diamond, P., 1967. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility: comment. J. Polit. Econ. 67, 765–766.

Epstein, L. G., Segal, U., 1992. Quadratic social welfare functions. J. Polit. Econ. 100, 691–712.

Fleurbaey, M., 2009. Two variants of Harsanyi's aggregation theorem. Econ. Letters 105, 300–302.

Gayer, G., Gilboa, I., Samuelson, L., Schmeidler, D., 2014. Pareto efficiency with different beliefs. J. Legal Stud. 43, 151–171.

Geoffrion, A. M., 1968. Proper efficiency and the theory of vector maximization. J. Math. Anal. Appl. 22, 618–630.

Gilboa, I., Samet, D., Schmeidler, D., 2004. Utilitarian aggregation of beliefs and tastes. J. Polit. Econ. 112, 932–938.

Gilboa, I., Samuelson, L. Schmeidler, D., 2014. No-betting-Pareto dominance. Econometrica 82, 1405–1442.

Grant, S., Kajii, A., Polak, B., Safra, Z., 2010. Generalized utilitarianism and Harsanyi's impartial observer theorem. Econometrica 78, 1939–1971.

Hammond, P. J., 1981. Ex-ante and ex-post welfare optimality under uncertainty. Economica 48, 235–250.

Harsanyi, J., 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. J. Polit. Econ. 63, 309–321.

Hayashi, T., Lombardi, M., 2017. Fair social decision under uncertainty and belief disagreements. Econ. Theory, 1–42.

Hylland, A., Zeckhauser, R., 1979. The impossibility of Bayesian group decision making with separate aggregation of beliefs and values. Econometrica 47, 1321–1336.

Mongin, P., 1995. Consistent Bayesian aggregation. J. Econ. Theory 66, 313–351.

Mongin, P., 2016. Spurious unanimity and the Pareto principle. Econ.Philos. 32, .511–532.

Mongin, P., Pivato, M., 2015. Ranking multidimensional alternatives and uncertain prospects. J. Econ. Theory 157, 146–171.

Rockafellar, R. T., 1996. Convex Analysis. Princeton University Press.

Segal, U., 1990. Two-Stage lotteries without the reduction axiom. Econometrica 58, 349–377.

Sen, A., 2004. Incompleteness and reasoned choice. Synthese 140, 43–59.

Sen, A., 2018. The importance of incompleteness. Int. J. Econ. Theory 14, 9–20.

Wald, A., 1950. Statistical Decision Functions. Chapman & Hall.