



Harsanyi's theorem without the sure-thing principle: On the consistent aggregation of Monotonic Bernoullian and Archimedean preferences

Stéphane Zuber

► To cite this version:

Stéphane Zuber. Harsanyi's theorem without the sure-thing principle: On the consistent aggregation of Monotonic Bernoullian and Archimedean preferences. 2015. halshs-01224145

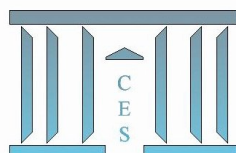
HAL Id: halshs-01224145

<https://shs.hal.science/halshs-01224145>

Submitted on 4 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Harsanyi's theorem without the sure-thing principle:
On the consistent aggregation of Monotonic Bernoullian
and Archimedean preferences**

Stéphane ZUBER

2015.69



Harsanyi's theorem without the sure-thing principle: On the consistent aggregation of Monotonic Bernoullian and Archimedean preferences

Stéphane Zuber^a

September 2015

Abstract

This paper studies the extension of Harsanyi's theorem (Harsanyi, 1955) in a framework involving uncertainty. It seeks to extend the aggregation result to a wide class of Monotonic Bernoullian and Archimedean preferences (Cerreia-Vioglio et al., 2011) that subsumes many models of choice under uncertainty proposed in the literature. An impossibility result is obtained, unless we are in the specific framework where all individuals and the decision-maker are subjective expected utility maximizers sharing the same beliefs. This implies that non-expected utility preferences cannot be aggregated consistently.

Keywords: Harsanyi's Theorem, Pareto Principle, Monotonic Bernoullian and Archimedean Preferences, Subjective Expected Utility.

JEL Classification numbers: D71, D81.

^aParis School of Economics – CNRS, Centre d'Économie de la Sorbonne (UMR 8174), Maison des Sciences Économiques, 106-112 Boulevard de l'Hôpital, 75013 Paris, France. E-mail: Stephane.Zuber@univ-paris1.fr.

This note is a deeply revised version of a previous paper entitled “Harsanyi's theorem without the sure-thing principle”, which was circulated in 2010. I am indebted to Philippe Mongin and Marcus Pivato for detailed comments on the earlier version. I am also grateful to Jean-Marc Tallon for helpful discussion on the recent literature and useful references. I remain of course responsible for any mistake or misinterpretation.

1 Introduction

Harsanyi (1955) proved the following theorem: if (i) all individuals and the social observer are expected utility maximizers on the set of simple lotteries over a set Y ; (ii) whenever all individuals prefer lottery p to lottery q , the social observer must also prefer p to q (the Pareto principle); then the social observer's utility function must be an affine combination of individuals' von Neumann and Morgenstern utility functions. This result is known as Harsanyi's aggregation theorem.

The purpose of this note is to take up the issue considered by Harsanyi, namely the aggregation of preferences, but when uncertainty rather than risk prevails. In that context, the basic model is Subjective Expected Utility (henceforth SEU, see Savage, 1954), which captures beliefs through a probability distribution. But this model has been challenged, in particular because the Sure-Thing Principle (holding that outcomes which occur regardless of which action is chosen should not affect one's preferences) is often violated by individual decision makers.

In this note, we considerably widen the investigation by considering a large class of preferences, namely Monotonic Bernoullian Archimedean preferences (henceforth MBA preferences, see Cerreia-Vioglio et al., 2011). This model permits a separation between tastes and ambiguous beliefs (represented by a set of probability distributions), in a specific sense. It also encompasses many models of decision under uncertainty that have been proposed recently, in particular the Choquet expected utility model by Schmeidler (1989), the multiple-prior preferences by Gilboa and Schmeidler (1989), the variational preferences of Maccheroni, Marinacci and Rustichini (2006) and the second order expected utility model by Klibanoff, Marinacci and, Mukerji (2005) and Ergin and Gul (2009).

Our main result is that agents' MBA preferences can be aggregated into MBA social preferences only in the very specific case where all individuals and the society have SEU preferences and share the same beliefs. In that case, it is also shown that the social aggregation rule is very specific, namely the social SEU is an affine combination of the agents SEU. Hence a consistent social aggregation of preferences is impossible as soon as at least one of the following is true: (i) agents have different beliefs; (ii) agents have non-SEU preferences.

It is already known that in models of decision under uncertainty it may not be possible to aggregate beliefs and tastes in a consistent way. Hylland and Zeckhauser

(1979) and Mongin (1995, 1998) showed that for SEU preferences: when individuals are simultaneously heterogeneous in terms of tastes and beliefs, it is impossible to aggregate individual preferences into an SEU social preference. A key issue is “spurious unanimity”: opposite beliefs and tastes may induce people to agree that a decision is better for conflicting reasons.

In the case of ambiguous beliefs, few results on preference aggregation exist in the literature. Chambers and Hayashi (2006) considered a Savage framework, where decision makers may have non-SEU preferences. They showed that, when individuals have different beliefs, eventwise monotonicity is incompatible with the Pareto axiom. Our setting further weakens eventwise dominance to statewise dominance (monotonicity). Yet, we are able to obtain an impossibility similar to the one they describe even when people have the same beliefs.

Gajdos, Tallon and Vergnaud (2008) showed that, within the class of “Rank-Dependent Additive” preferences, a consistent aggregation is impossible, even if all individuals have the same beliefs, unless the society and the individual are uncertainty neutral (which occurs in the SEU model). However, their aggregation result is obtained for a smaller class of preferences than the one studied in this note.

The paper closest to ours is Mongin and Pivato (2015), which was written independently of the present note. Their conclusion, like ours, is that consistent aggregations of preferences can be achieved only if (i) the individuals and the observer are all SEU maximizers, and (ii) they use the same subjective probabilities. Technically, their method of proof is similar to ours. It is based on the key insight provided by Blackorby, Donaldson and Mongin (2004) that the Pareto principle and monotonicity (or statewise dominance) impose separability conditions. Combining these separability conditions yield an additive separable representation. Their paper has a more general framework than the one in this note: it encompasses other frameworks than choice under uncertainty. But they also assume technical structural conditions (that the set of feasible alternative is connected, sectionally connected, and indifference connected for the social ordering). In the Anscombe-Aumann framework (Anscombe and Aumann, 1963), we do not need such structural assumptions, but only standard and easily interpretable choice-theoretic premisses.¹

¹In different but related settings, Blackorby, Donaldson and Mongin (2004) and Chambers and Hayashi (2014) use the same ideas to derive additive representations, but they need to assume that the

Recently, Herzberg (2013) also obtained an impossibility result for the class of MBA preferences we consider in this note. But he does so within an Arrovian aggregation framework, using an aggregator desideratum of systematicity. In this note, we do not impose such regularity constraints on the social aggregator.

Other papers are related to this note. Fleurbaey (2009) studied the aggregation of expected utility individual preferences into a non-expected utility social preference ordering, which satisfies statewise dominance. He obtained that the social aggregation must be affine so that society must use an expected utility to evaluate prospects. We also find this affine aggregation when individuals' preferences are SEU. This note broadens the investigation to non SEU individual preferences and does not assume objective probabilities. Our proof is nonetheless related to his, for it relies on functional equations.

More generally, there is a strand of the literature studying the consistency of ex ante and ex post social judgements. Most of the literature has focused on the expected utility case, to show that the Utilitarian rule is the only one that can be used consistently ex ante and ex post (see in particular Hammond, 1981; Myerson, 1981; Blackorby, Donaldson and Weymark, 1999). The present note also seeks consistent social judgements, but enlarges the study to non-expected utility preferences. It shows that not only the social aggregation rule matters, but also the kind of preferences that are aggregated. Specifically, we show that only SEU preferences with common beliefs can be aggregated consistently.

The remainder of the note is divided into three sections and one appendix. In Section 2, we lay down the model and define MBA preferences. In Section 3, we present the result concerning the possibility of consistent aggregation of preferences (Theorem 1). In Section 4, we discuss this result and several directions that can be explored to go beyond it. An appendix provides the proof of Theorem 1.

2 The framework

Consider a society made of a finite number of agents $\mathcal{N} = \{1, \dots, N\}$. Let $\mathcal{N}_0 = \{0, 1, \dots, N\}$ where 0 refers to society.

set of alternatives is a product space. This is less general than Mongin and Pivato (2015) and the present note.

Uncertainty is modeled using the Anscombe-Aumann framework (Anscombe and Aumann, 1963). There is a finite set $\mathcal{S} = \{1, \dots, S\}$ of states of the world. Let C be a non-empty set of consequences and $X = \Delta(C)$ be the set of simple lotteries over C (that is lotteries with finite support). Let \mathcal{A} be the set of acts, that is, mappings $f : \mathcal{S} \rightarrow X$. Since X is a mixture space, the mixture operation $\alpha x + (1 - \alpha)y$ is well-defined for all $\alpha \in [0, 1]$ and $x, y \in X$, and $\alpha x + (1 - \alpha)y \in X$. We will identify X with the subset of sure acts, that is acts $x \in \mathcal{A}$ such that $x(s) = x$ for all $s \in \mathcal{S}$, where $x \in X$. One can also define a mixture operation on \mathcal{A} in the following way: for any $f, g \in \mathcal{A}$ and $\alpha \in [0, 1]$, the act $\alpha f + (1 - \alpha)g$ in \mathcal{A} is such that $(\alpha f + (1 - \alpha)g)(s) = \alpha f(s) + (1 - \alpha)g(s) \in X$ for every state $s \in \mathcal{S}$.

We model the preferences of an agent $i \in \mathcal{N}$ on \mathcal{A} by a binary relation \succsim_i , and denote by \succ_i and \sim_i its symmetric and asymmetric components. Society's preferences are denoted \succsim_0 , with similar definitions.

Let us now define some properties of binary relations \succsim over \mathcal{A} . The first property guarantees that preferences are complete orderings that can be represented by continuous utility functions.

Property 1 (Regularity). *The binary relation \succsim satisfies:*

- a. Completeness. *For all $f, g \in \mathcal{A}$ either $f \succsim g$ or $g \succsim f$ (or both).*
- b. Transitivity. *For all $f, g, h \in \mathcal{A}$, if $f \succsim g$ and $g \succsim h$ then $f \succsim h$.*
- c. Archimedean. *For all $f, g, h \in \mathcal{A}$, if $f \succ g$ and $g \succ h$ then there exist $\alpha, \beta \in (0, 1)$ such that $\alpha f + (1 - \alpha)h \succ g$ and $g \succ \beta f + (1 - \beta)h$.*

The second property is a dominance property that guarantees the consistency between preferences over acts and preferences over lotteries in each state of the world.

Property 2 (Monotonicity). *For all $f, g \in \mathcal{A}$, if $f(s) \succsim g(s)$ for all $s \in \mathcal{S}$, then $f \succsim g$. If furthermore there exists $t \in \mathcal{S}$ such that $f(t) \succ g(t)$, then $f \succ g$.*

Note that we use a strong version of Monotonicity, which implies the sensitivity to each state of the world. This also means that we rule out the existence of null states. This is done for the sake of simplicity, to avoid some complications with dealing with null states that would not change the message of the present note.

The last property is the VNM independence axiom restricted to risk situations where probabilities are known.

Property 3 (Risk independence). For all $x, y, z \in X$, if $x \succsim y$, then for all $\alpha \in [0, 1]$, $x \succsim y \iff \alpha x + (1 - \alpha)z \succsim \alpha y + (1 - \alpha)z$.

For any set Z , and function $v : Z \rightarrow \mathbb{R}$, we define $Rg(v) = \{a \in \mathbb{R} : \exists z \in Z, a = v(z)\}$. Let B be an interval in \mathbb{R} , a monotonic function $I : B^S \rightarrow \mathbb{R}$ is normalized if $I(b, \dots, b) = b$ for all $b \in B$.

Definition 1. The binary relation \succsim on \mathcal{A} is a Monotonic Bernoullian Archimedean (MBA) preference ordering if there exist an affine function $u : X \rightarrow \mathbb{R}$ and a continuous increasing and normalized function $I : Rg(u) \rightarrow \mathbb{R}$ such that, for all $f, g \in \mathcal{A}$,

$$f \succsim g \iff I(u(f(1)), \dots, u(f(S))) \geq I(u(g(1)), \dots, u(g(S))). \quad (1)$$

Moreover, if $v : X \rightarrow \mathbb{R}$ and $I_v : Rg(v) \rightarrow \mathbb{R}$ also satisfy (1) and I_v is normalized, then there exist $\lambda \in \mathbb{R}_{++}$ and $\mu \in \mathbb{R}$ such that $v(x) = \lambda u(x) + \mu$ for all $x \in X$ and $I_v(b_1, \dots, b_S) = \lambda I(\frac{b_1 - \mu}{\lambda}, \dots, \frac{b_S - \mu}{\lambda})$ for all $(b_1, \dots, b_S) \in Rg(v)^S$.

MBA preferences are characterized by the three properties introduced above.

Proposition 1 (Cerreia-Vioglio et al., 2011). A binary relation \succsim on \mathcal{A} is an MBA preference ordering if and only if it satisfies Regularity, Monotonicity and Risk Independence.

Proof. Similar to the proof of Proposition 1 in Cerreia-Vioglio et al. (2011), the only difference being that we only consider a finite number of states of the world, which are all essential by Monotonicity, so that I is increasing. \square

MBA preferences encompass many decision models proposed in the literature on choice under uncertainty. In particular it encompasses the SEU model.

Definition 2. A binary relation \succsim on \mathcal{A} is a subjective expected utility (SEU) preference ordering if there exists a vector $(p_1, \dots, p_S) \in (0, 1)^S$ with $\sum_{s \in S} p_s = 1$ and an affine function $u : X \rightarrow \mathbb{R}$ such that, for all $f, g \in \mathcal{A}$,

$$f \succsim g \iff \sum_{s \in S} p_s u(f(s)) \geq \sum_{s \in S} p_s u(g(s)). \quad (2)$$

In the above definition, the vector (p_1, \dots, p_S) is the subjective probability distribution used by the decision maker. When two decision makers have preferences that

are SEU preference orderings, and when they have the same subjective probability distribution, we say that they share the same beliefs.

3 The (im)possibility of consistent aggregations of MBA preferences

Let us now turn to the aggregation of preferences. We assume that all decision makers (including society) have MBA preferences orderings \succsim_i , $i \in \mathcal{N}_0$. We also assume that there exists some diversity among the individual preferences that we seek to aggregate. It is represented by the following assumption that is known in the literature as the ‘independent prospects’ property (Weymark, 1991; Mongin, 1998).

Assumption 1. *For any $i \in \mathcal{N}$, there exists $x_i, y_i \in X$ such that $x_i \succ_i y_i$ and $x_i \sim_j y_i$ for all $j \in \mathcal{N} \setminus \{i\}$.*

The aim of a social aggregation is to derive the social preference ordering from individual preference orderings. Many ethical principles can be invoked to determine how this aggregation should be made. But the most widely accepted principle is the Pareto principle, whose strong version is as follows.

Axiom 1 (Strong Pareto). *For all $f, g \in \mathcal{A}$, if $f \succsim_i g$ for all $i \in \mathcal{N}$, then $f \succsim_0 g$. If moreover there exists $j \in \mathcal{N}$ such that $f \succ_j g$, then $f \succ_0 g$.*

The following (im)possibility theorem, whose proof is in the Appendix, delineates the scope of consistent preference aggregations.

Theorem 1. *Assume that all individuals have MBA preferences satisfying Assumption 1. Social preferences can be MBA preferences and satisfy the Strong Pareto principle only if all individuals have SEU preferences and share the same beliefs.*

In that case, there exist weights $(a_1, \dots, a_N) \in \mathbb{R}_+^N$ such that, for all $f, g \in \mathcal{A}$:

$$f \succsim_0 g \iff \sum_{s \in \mathcal{S}} p_s \left(\sum_{i \in \mathcal{N}} a_i u_i(f(s)) \right) \geq \sum_{s \in \mathcal{S}} p_s \left(\sum_{i \in \mathcal{N}} a_i u_i(g(s)) \right), \quad (3)$$

where $(p_1, \dots, p_S) \in (0, 1)^S$ is the individuals’ subjective probability distribution. Furthermore the weight vector $(a_1, \dots, a_N) \in \mathbb{R}_+^N$ is unique, and the affine functions u_i are unique up to positive affine transformations with a common multiplier.

The first part of Theorem 1 is an impossibility result. It states that it is impossible to consistently aggregate preferences under uncertainty, unless individuals and the society have SEU preferences and share the same beliefs. In particular, SEU preferences cannot be aggregated when individuals have different beliefs (Mongin, 1995). And non-SEU preferences cannot be aggregated even when people share the same beliefs (Gajdos, Tallon and Vergnaud, 2008). Theorem 1 thus generalizes existing impossibility results.

The second part is the extension of Harsanyi's results to the Anscombe-Aumann framework when people are SEU and share the same beliefs. The social welfare function must be an affine aggregation of individuals' SEU. This specific aggregation has been best described as generalized utilitarian (for clear and comprehensive reviews of the ethical significance of Harsanyi's theorem, see Weymark, 1991; Mongin and d'Aspremont, 1998). It may be problematic for societies wishing to incorporate fairness into their decision making process.

As already noted by Diamond (1967), Harsanyi's aggregation procedure prevents ex ante equity considerations requiring progressive transfers of expected utility to be welfare improving. Myerson (1981) also noted the difficulty of defining egalitarian principles in risky situations. In the more general framework of dynamic choice under uncertainty, Bommier and Zuber (2012) have already proved that rational Paretian social observers must have the same aversion to inequality. Theorem 1 further specifies this result by implying that inequality aversion is nil (with respect to specific normalizations of the expected utilities). Another issue is that of ex post equity. Fleurbaey (2010) has proposed a specific weakening of the Pareto principle to deal with it.

4 Discussion

Theorem 1 states that it is generally impossible to consistently aggregate MBA individual preferences into MBA social preferences. To escape the impossibility, several directions can be explored.

An important condition for Theorem 1 to hold is Assumption 1. Of course, if the assumption is true for some individuals in \mathcal{N} only, the impossibility would still hold (we only lose the uniqueness of the weights a_i in Equation 3). To recover a possibility result, we need to assume that all individuals have the same preferences on X . This assumption was made by Crès, Gilboa and Vieille (2011). They showed that it is then possible to aggregate agents beliefs in a consistent way within the multiple prior

preference model. This finding is interesting but it seems restrictive to assume that people have exactly the same risk preferences. The result is however useful if we focus on the aggregation of preferences restricted to common-taste uncertain options (see Danan et al., 2015).

Another option would be to widen the scope of admissible social preferences beyond MBA preferences. An option that has widely been explored is to allow social preferences to be state-dependent. Within the SEU framework, Mongin (1998) obtained a possibility result: although agents have different beliefs, it is possible to obtain a state dependent SEU social ordering, which is an affine combination of agents ex ante expected utilities. Similar results have been presented by Chambers and Hayashi (2006). However, allowing for state dependence at the social level cannot restore the possibility for non-SEU preferences (see Gajdos, Tallon and Vergnaud, 2008; Mongin and Pivato, 2015). The results of the present note could be extended to allow state dependence, without modifying the result. Another possibility would be to completely drop the requirement of Monotonicity (or statewise dominance). This is actually what was suggested by the literature promoting ex ante equity in the context of risk and uncertainty (see for instance Diamond, 1967; Epstein and Segal, 1992), but one then has to deal with the issue of time consistency (Machina, 1989).

A last option to restore possibilities would be to weaken the Pareto principle. A first direction would be to only use the Weak Pareto principle. This would make possible dictatorial choices, where the society always makes decision using the preference ordering of a specific individual. Positional dictatorships (such as Maxmin) are however not possible in general (see Chambers and Hayashi, 2014). Other weakened versions of the Pareto principle have been proposed. To avoid spurious unanimities Gilboa, Samet and Schmeidler (2004) proposed to restrict the use of the Pareto principle to cases where SEU agents have the same beliefs. Danan et al. (2015) have recently shown that this not enough if we want to study MBA preferences: even with identical sets of belief, agents may agree to prefer situations which are socially dominated in each state of the world. Danan et al. (2015) hence further restricted the Pareto principle to apply to common-taste outcomes (they call this principle the common-taste robust Pareto principle). Brunnermeier et al. (2014) proposed a similar version of the Pareto principle. Gilboa, Samuelson and Schmeidler (2014) proposed a stronger notion of no-betting Pareto dominance. This route is thus a promising avenue for research. One

limitation, though, is that affine aggregations of utilities are in general still obtained (see for instance Gilboa, Samet and Schmeidler, 2004; Danan et al., 2015). Fairness issues, such as those discussed Section 3, remain unresolved by the use of such weakened Pareto principles.

Appendix

Assume that all agents have MBA preferences on \mathcal{A} , and that social preferences are also MBA. By Proposition 1, for any $i \in \mathcal{N}_0$ there exist an affine function $u_i : X \rightarrow \mathbb{R}$ and a continuous increasing and normalized function $I_i : Rg(u_i) \rightarrow \mathbb{R}$ such that, for all $f, g \in \mathcal{A}$,

$$f \succsim_i g \iff I_i(u_i(f(1)), \dots, u_i(f(S))) \geq I_i(u_i(g(1)), \dots, u_i(g(S))). \quad (4)$$

The following Lemma concerns the aggregation in risky situations.

Lemma 1. *There exist unique weights $(a_1, \dots, a_N) \in \mathbb{R}_{++}^N$ and a constant $b \in \mathbb{R}$ such that, for all $x \in X$:*

$$u_0(x) = \sum_{i \in \mathcal{N}} a_i u_i(x) + b, \quad (5)$$

Proof. Applying the Strong Pareto principle to preferences restricted to X , this follows from the results by Fishburn (1984, Theorem 2 and Corollary 1). \square

Define $\mathcal{U} = \{(v_1, \dots, v_N) \in \mathbb{R}^N \mid \exists x \in X : v_i = u_i(x), \forall i \in \mathcal{N}\}$. The set \mathcal{U} is a nonempty convex subspace of \mathbb{R}^N of dimension N .² Hence \mathcal{U}^S is also a nonempty convex subspace of $\mathbb{R}^{N \times S}$ of dimension NS . A vector $v \in \mathcal{U}^S$ has coordinates v_i^s , meaning that there exists $f \in \mathcal{A}$ such that, for any $i \in \mathcal{N}$ and any $s \in \mathcal{S}$, $v_i^s = u_i(f(s))$.

For all $i \in \mathcal{N}$, we now define the relations \succsim_i^* over \mathcal{U}^S in the following way. For any $v, \tilde{v} \in \mathcal{U}^S$ and any $f, g \in \mathcal{A}$ such that $v_i^s = u_i(f(s))$ and $\tilde{v}_i^s = u_i(g(s))$ for all $i \in \mathcal{N}$ and $s \in \mathcal{S}$:

$$v \succsim_i^* \tilde{v} \iff f \succsim_i g. \quad (6)$$

²Indeed, if $v = (v_1, \dots, v_N) \in \mathcal{U}$ and $\tilde{v} = (\tilde{v}_1, \dots, \tilde{v}_N) \in \mathcal{U}$, this implies that there exist $x, y \in X$ such that $v_i = u_i(x)$ and $\tilde{v}_i = u_i(y)$, $\forall i \in \mathcal{N}$. For any $\lambda \in (0, 1)$, let $z = \lambda x + (1 - \lambda)y$. Then, for all $i \in \mathcal{N}$, $u_i(z) = \lambda u_i(x) + (1 - \lambda)u_i(y) = \lambda v_i + (1 - \lambda)\tilde{v}_i$. Hence, $\lambda v + (1 - \lambda)\tilde{v} \in \mathcal{U}$. Assumption 1 implies that the dimension of \mathcal{U} is N , because all components are independent.

Hence, using Equation (4), for all $i \in \mathcal{N}$ and all $v, \tilde{v} \in \mathcal{U}^S$, $v \succsim_i^* \tilde{v} \iff I_i(v_i^1, \dots, v_i^S) \geq I_i(\tilde{v}_i^1, \dots, \tilde{v}_i^S)$. By Lemma 1, it must also be the case that, for all $v, \tilde{v} \in \mathcal{U}^S$,

$$v \succsim_0^* \tilde{v} \iff I_0\left(\sum_{i \in \mathcal{N}} a_i v_i^1 + b, \dots, \sum_{i \in \mathcal{N}} a_i v_i^S + b\right) \geq I_0\left(\sum_{i \in \mathcal{N}} a_i \tilde{v}_i^1 + b, \dots, \sum_{i \in \mathcal{N}} a_i \tilde{v}_i^S + b\right) \quad (7)$$

The following Lemma lists some properties of relation \succsim_0^* .

Lemma 2. *The relation \succsim_0^* on \mathcal{U}^S is complete, transitive, continuous and all indifference sets of \succsim_0^* are connected subsets of \mathcal{U}^S .*

Proof. That \succsim_0^* is complete, transitive and continuous is proved by the fact that it is represented on \mathcal{U}^S by a continuous real-valued function (Equation 7).

Consider any $v, \tilde{v} \in \mathcal{U}^S$ such that $v \sim_0^* \tilde{v}$. Denote $\bar{S} = \{s \in \mathcal{S} \mid \sum_{i \in \mathcal{N}} a_i v_i^s > \sum_{i \in \mathcal{N}} a_i \tilde{v}_i^s\}$ and $\underline{S} = \{s \in \mathcal{S} \mid \sum_{i \in \mathcal{N}} a_i v_i^s < \sum_{i \in \mathcal{N}} a_i \tilde{v}_i^s\}$. For any $\lambda, \mu \in [0, 1]$, also denote ${}^\mu_\lambda v$ the vector such that

(a) ${}^\mu_\lambda v_i^s = \lambda v_i^s + (1 - \lambda) \tilde{v}_i^s$ for all $i \in \mathcal{N}$ and $s \in (\mathcal{S} \setminus \underline{S})$,

(b) and ${}^\mu_\lambda v_i^s = \mu \tilde{v}_i^s + (1 - \mu) v_i^s$ for all $i \in \mathcal{N}$ and $s \in \underline{S}$.

By convexity of \mathcal{U} , we know that ${}^\mu_\lambda v \in \mathcal{U}^S$ for any $\lambda, \mu \in [0, 1]$.

Case 1: $\bar{S} = \underline{S} = \emptyset$. In that case, $v \sim_0^* {}^\lambda_\lambda v \sim_0^* \tilde{v}$ for any $\lambda \in [0, 1]$. Hence the map $\psi : [0, 1] \rightarrow \mathcal{U}^S$ such that $\psi(\lambda) = {}^\lambda_\lambda v$ for all $\lambda \in [0, 1]$ defines a path between v and \tilde{v} whose elements are on the same indifference curve as v and \tilde{v} . The segment of the indifference curve between v and \tilde{v} is path connected.

Case 2: $\bar{S} = \emptyset$ and $\underline{S} \neq \emptyset$ or $\bar{S} \neq \emptyset$ and $\underline{S} = \emptyset$. This is impossible: because I_0 is increasing, this would imply either $v \prec \tilde{v}$ or $v \succ \tilde{v}$, which is a contradiction.

Case 3: $\bar{S} \neq \emptyset$ and $\underline{S} \neq \emptyset$. For $\lambda \in (0, 1)$, we have ${}^0_\lambda v \prec_0^* v \sim_0^* \tilde{v} \prec_0^* {}^1_\lambda v$.³ the map $\Psi : [0, 1] \rightarrow \mathcal{U}^S$ such that $\Psi(\mu) = {}^\mu_\lambda v$ for all $\mu \in [0, 1]$ defines a path between ${}^0_\lambda v$ and ${}^1_\lambda v$. Given that \succsim_0^* is continuous, there exist $\mu \in (0, 1)$ such that ${}^\mu_\lambda v \sim_0^* v \sim_0^* \tilde{v}$. We denote $\mu(\lambda)$ this number, which is unique.⁴ We also take $\mu(0) = 1$ and $\mu(1) = 0$. Hence the map $\psi : [0, 1] \rightarrow \mathcal{U}^S$ such that $\psi(\lambda) = {}^{\mu(\lambda)}_\lambda v$ for all $\lambda \in [0, 1]$ defines a path between v

³Indeed, by definition, for all $s \in \bar{S}$, $\sum_{i \in \mathcal{N}} a_i v_i^s > \lambda (\sum_{i \in \mathcal{N}} a_i v_i^s) + (1 - \lambda) (\sum_{i \in \mathcal{N}} a_i \tilde{v}_i^s) = \sum_{i \in \mathcal{N}} a_i {}^0_\lambda v_i^s$; in the meantime, $\sum_{i \in \mathcal{N}} a_i v_i^s = \sum_{i \in \mathcal{N}} a_i {}^0_\lambda v_i^s$ for all $s \in (\mathcal{S} \setminus \bar{S})$. Similarly, for all $s \in \bar{S}$, $\sum_{i \in \mathcal{N}} a_i \tilde{v}_i^s < \lambda (\sum_{i \in \mathcal{N}} a_i v_i^s) + (1 - \lambda) (\sum_{i \in \mathcal{N}} a_i \tilde{v}_i^s) = \sum_{i \in \mathcal{N}} a_i {}^1_\lambda v_i^s$; in the meantime, $\sum_{i \in \mathcal{N}} a_i \tilde{v}_i^s = \sum_{i \in \mathcal{N}} a_i {}^1_\lambda v_i^s$ for all $s \in (\mathcal{S} \setminus \bar{S})$. The results thus hold because I_0 is increasing.

⁴Indeed, we have $\sum_{i \in \mathcal{N}} a_i {}^\mu_\lambda v_i^s = \lambda (\sum_{i \in \mathcal{N}} a_i v_i^s) + (1 - \lambda) (\sum_{i \in \mathcal{N}} a_i \tilde{v}_i^s)$ for all $\mu \in [0, 1]$ and $s \in (\mathcal{S} \setminus \underline{S})$.

and \tilde{v} whose elements are on the same indifference curve as v and \tilde{v} . The segment of the indifference curve between v and \tilde{v} is path connected.

All indifference sets of \succsim_0^* are path-connected and thus connected. \square

The function $V_i : \mathcal{U}^S \rightarrow \mathbb{R}$ defined by $V_i(v) = I_i(v_i^1, \dots, v_i^S)$ for all $v \in \mathcal{U}^S$ is a representation of \succsim_i^* for each $i \in \mathcal{N}$. The function $V_0 : \mathcal{U}^S \rightarrow \mathbb{R}$ defined by $V_0(v) = I_0\left(\sum_{i \in \mathcal{N}} a_i u_i^1 + b, \dots, \sum_{i \in \mathcal{N}} a_i u_i^S + b\right)$ for all $v \in \mathcal{U}^S$ is a representation of \succsim_0^* . The Strong Pareto principle and the definition of \succsim_i^* also imply that:

(P*) For all $v, \tilde{v} \in \mathcal{U}^S$, if $v \succsim_i^* \tilde{v}$ for all $i \in \mathcal{N}$, then $v \succsim_0^* \tilde{v}$. If moreover there exists $j \in \mathcal{N}$ such that $v \succ_j^* \tilde{v}$, then $v \succ_0^* \tilde{v}$.

We have the following Lemma:

Lemma 3. Let $B = \{(b_1, \dots, b_N) \in \mathbb{R}^N \mid \exists v \in \mathcal{U}^S : b_i = V_i(v) \text{ for all } i \in \mathcal{N}\}$. There exists a continuous and increasing function $W : K \rightarrow \mathbb{R}$ such that for all $f \in \mathcal{A}$:

$$V_0(f) = W\left(V_1(f), \dots, V_N(f)\right). \quad (8)$$

Proof. Under Pareto Indifference (which is implied for preferences over \mathcal{U}^S by property P*), Blackorby, Donaldson and Weymark (1999, Lemma 1, p. 369; Lemma 2, p.370) showed that there must exist a continuous social aggregator function W . Condition P* further entails that W must be increasing in each of its components. \square

By definition of the V_i functions, using Equation (8), and denoting, for each $s \in \mathcal{S}$, $w_s : \mathcal{U} \rightarrow \mathbb{R}$ the function such that $w_s(v_1^s, \dots, v_N^s) = \sum_{i \in \mathcal{N}} a_i v_i^s + b$ for any $(v_1^s, \dots, v_N^s) \in \mathcal{U}$, the following equation holds for any $v \in \mathcal{U}^S$:

$$W\left(I_1(v_1^1, \dots, v_1^S), \dots, I_N(v_N^1, \dots, v_N^S)\right) = I_0\left(w_1(v_1^1, \dots, v_N^1), \dots, w_S(v_1^S, \dots, v_N^S)\right). \quad (9)$$

Consider any open box K in $\mathbb{R}^{N \times S}$ which is a subspace of \mathcal{U}^S . The functional equation (9) holds for any vector in K . Given that the functions I_0 , W , w and I_i are

And, for all $s \in \underline{S}$:

$$\sum_{i \in \mathcal{N}} a_i \overset{\mu}{\lambda} v_i^s > \sum_{i \in \mathcal{N}} a_i \overset{\hat{\mu}}{\lambda} v_i^s \iff \mu < \hat{\mu}.$$

The only possibility to have $\overset{\mu}{\lambda} v \sim_0^* \overset{\hat{\mu}}{\lambda} v$ is that $\mu = \hat{\mu}$.

all continuous and increasing, the solution of the functional equation is (see Aczél and Maksa, 1996, Theorem 5):

$$I_0(w_1, \dots, w_S) = \Phi \left(\sum_{s \in \mathcal{S}} \psi_s(w_s) \right) \quad (10)$$

$$W(I_1, \dots, I_N) = \Phi \left(\sum_{i \in \mathcal{N}} \phi_i(I_i) \right) \quad (11)$$

$$\forall s \in \mathcal{S}, \quad w_s(v_1^s, \dots, v_N^s) = \psi_s^{-1} \left(\sum_{i \in \mathcal{N}} \varphi_i^s(v_i^s) \right) \quad (12)$$

$$\forall i \in \mathcal{N}, \quad I_i(v_i^1, \dots, v_i^S) = \phi_i^{-1} \left(\sum_{s \in \mathcal{S}} \varphi_i^s(v_i^s) \right) \quad (13)$$

where the real functions $\Phi, \psi_s, \phi_i, \varphi_i^s$ are continuous and increasing.

But we also know that, for any $s \in \mathcal{S}$ and $(v_1^s, \dots, v_N^s) \in \mathcal{U}$, $w_s(v_1^s, \dots, v_N^s) = \sum_{i \in \mathcal{N}} a_i v_i^s + b$. Combining with Equation (12), we obtain the following Pexider equation on a rectangular non-empty subset of \mathcal{U} :

$$\sum_{i \in \mathcal{N}} \varphi_i^s(v_i^s) = \psi_s \left(\sum_{i \in \mathcal{N}} a_i v_i^s + b \right).$$

Thus there must exist a positive number π_s and real numbers κ_i^s such that $\psi_s(z) = \pi_s z + \left(\sum_{i \in \mathcal{N}} \kappa_i^s - \pi_s b \right)$ and $\varphi_i^s(v_i^s) = \pi_s a_i v_i^s + \kappa_i^s$, for all $i \in \mathcal{N}$, on the relevant domain (see Rádo and Baker, 1987, Theorem 1 and Corollary 3).

Let $\text{int}(\mathcal{U}^S)$ be the interior of the set \mathcal{U}^S . For any $v \in \text{int}(\mathcal{U}^S)$, it is possible to construct an open box $K(v)$ in $\mathbb{R}^{N \times S}$ around v , which is a subspace of \mathcal{U}^S . On $K(v)$, \succsim_0^* is locally represented by the fully additive representation:

$$W(v) = \Phi \left(\sum_{i \in \mathcal{N}} \sum_{s \in \mathcal{S}} \pi_s a_i v_s^i + \sum_{i \in \mathcal{N}} \sum_{s \in \mathcal{S}} \kappa_i^s \right).$$

But, since $\text{int}(\mathcal{U}^S)$ is convex and \succsim_0^* -indifference connected (Lemma 2), we know by Theorem 2.2 of Chateauneuf and Wakker (1993) that this local fully additive representation extends globally to the whole set $\text{int}(\mathcal{U}^S)$. Furthermore, this global additive representation is unique up to a positive affine transformation. This implies that the weight vectors $(p_1, \dots, p_S) \in (0, 1)^S$ and $(a_1, \dots, a_N) \in \mathbb{R}_+^N$ are uniquely defined. Last,

since \mathcal{U}^S is convex and \succsim_0^* is continuous, the representation also extends to the boundaries of \mathcal{U}^S .

Combining Equations (13), that $\varphi_i^s(v_i^s) = \pi_s a_i v_i^s + \kappa_i^s$ on the relevant domain, and the fact that I_i is normalized, we obtained that, for any $(v_i^1, \dots, v_i^S) \in \mathbb{R}^S$ such that there exists $f \in \mathcal{A}$ for which $u_i(f(s)) = v_i^s$:

$$I_i(v_i^1, \dots, v_i^S) = \sum_{s \in \mathcal{S}} \frac{\pi_s}{\sum_{s \in \mathcal{S}} \pi_s} v_i^s.$$

Denote $p_s = \frac{\pi_s}{\sum_{s \in \mathcal{S}} \pi_s}$ for all $s \in \mathcal{S}$. The vector $(p_1, \dots, p_S) \in (0, 1)^S$ uniquely defines a probability distribution over states of the world. And, because individuals have MBA preferences, we obtain that for all $i \in \mathcal{N}$ and $f, g \in \mathcal{A}$:

$$\begin{aligned} f \succsim_i g &\iff I_i(u_i(f(1)), \dots, u_i(f(S))) \geq I_i(u_i(g(1)), \dots, u_i(g(S))) \\ &\iff \sum_{s \in \mathcal{S}} p_s u_i(f(s)) \geq \sum_{s \in \mathcal{S}} p_s u_i(g(s)). \end{aligned}$$

By Definition 2 the individuals must therefore have SEU preferences and use the same subjective probability distribution (p_1, \dots, p_S) .

Combining Equations (10) and (12), and using $\varphi_i^s(v_i^s) = \pi_s a_i v_i^s + \kappa_i^s$ on the relevant domain and that I_0 is normalized, we also get that, for any $v \in \mathcal{U}^S$:

$$I_0(w_1(v_1^1, \dots, v_1^S), \dots, w_S(v_1^S, \dots, v_N^S)) = \sum_{s \in \mathcal{S}} p_s \left(\sum_{i \in \mathcal{N}} a_i v_i^s + b \right).$$

Using the definition of \succsim_0^* , we have that for any $f, g \in \mathcal{A}$:

$$\begin{aligned} f \succsim_0 g &\iff I_0(w_1(u_1(f(1)), \dots, u_N(f(1))), \dots, w_S(u_1(f(S)), \dots, u_N(f(S)))) \\ &\geq I_0(w_1(u_1(g(1)), \dots, u_N(g(1))), \dots, w_S(u_1(g(S)), \dots, u_N(g(S)))) \\ &\iff \sum_{s \in \mathcal{S}} p_s \left(\sum_{i \in \mathcal{N}} a_i u_i(f(s)) \right) \geq \sum_{s \in \mathcal{S}} p_s \left(\sum_{i \in \mathcal{N}} a_i u_i(g(s)) \right). \end{aligned}$$

The uniqueness of the global additive representation of \succsim up to a positive affine transformation implies that the affine functions u_i must be unique up to positive affine transformations with a common multiplier.

References

- Aczél, J., Maksa, G. (1996). “Solution of the rectangular $m \times n$ generalized bisymmetry equation and of the problem of consistent aggregation”, *Journal of Mathematical Analysis and Applications*, 203, 104–126.
- Anscombe, F.J., Aumann, R.J. (1963). “A definition of subjective probability”, *Annals of Mathematical Statistics*, 34, 199–205.
- Blackorby, C., Donaldson, D., Mongin, P. (2004). “Social aggregation without the expected utility hypothesis”, mimeo.
- Blackorby, C., Donaldson, D., Weymark, J.A. (1999). “Harsanyi’s social aggregation theorem for state-contingent alternatives”, *Journal of Mathematical Economics*, 32, 365–387.
- Bommier, A., Zuber, S. (2012). “The Pareto principle of optimal inequality”, *International Economic Review*, 53, 593–607.
- Brunnermeier, M. K., Simsek, A., Xiong, W. (2014). “A welfare criterion for models with distorted beliefs”, *Quarterly Journal of Economics*, 129, 1753–1797.
- Cerreia-Vioglio, S., Ghirardato, P., Maccheroni, F., Marinacci, M., Siniscalchi, M. (2011). “Rational preferences under ambiguity”, *Economic Theory*, 48, 341–375.
- Chambers, C.P., Hayashi, T. (2006). “Preference aggregation under uncertainty: Savage vs. Pareto”, *Games and Economic Behavior*, 54, 430–440.
- Chambers, C.P., Hayashi, T. (2014). “Preference aggregation with incomplete information”, *Econometrica*, 82, 589–599.
- Chateauneuf, A., Wakker, P. (1993). “From local to global additive representation”, *Journal of Mathematical Economics*, 22, 523–545.
- Crès, A., Gilboa, I., Vieille, N. (2011). “Aggregation of multiple prior opinions”, *Journal of Economic Theory*, 146, 2563–2582.
- Danan, E., Gajdos, T., Hill, B., Tallon, J.-M. (2015). “Robust social decisions”, mimeo.
- Diamond, P.A. (1967). “Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility: Comment”, *Journal of Political Economics*, 75, 765–766.
- Epstein, L.G., Segal, U. (1992). “Quadratic social welfare functions”, *Journal of Political Economy*, 100, 691–712.
- Ergin, H., Gul, F. (2009). “A subjective theory of compound lotteries”, *Journal of Economic Theory*, 144, 899–929.
- Fishburn, P.C. (1984). “On Harsanyi’s utilitarian cardinal welfare theorem”, *Theory and Decision*, 17, 21–28.
- Fleurbaey, M. (2009). “Two variants of Harsanyi’s aggregation theorem”, *Economics Letters*, 105, 300–302.
- Fleurbaey, M. (2010). “Assessing risky social situations”, *Journal of Political Economy*, 118, 649–680.
- Gajdos, T., Tallon, J.-M., Vergnaud, J.-C. (2008). “Representation and aggregation of preferences under uncertainty”, *Journal of Economic Theory*, 141, 68–99.
- Gilboa, I., Samet, D., Schmeidler, D. (2004). “Utilitarian aggregation of beliefs and tastes”, *Journal of Political Economy*, 112, 932–938.
- Gilboa, I., Schmeidler, D. (1989). “Maxmin expected utility with a non-unique prior”, *Journal of Mathematical Economics*, 18, 141–153.
- Gilboa, I., Samuelson, L., Schmeidler, D. (2014). “No betting Pareto dominance”, *Econometrica*, 82, 1405–1442.

- Hammond, P.J. (1981). "Ex-ante and ex-post welfare optimality under uncertainty", *Economica*, 48, 235–250.
- Harsanyi, J.C. (1955). "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility", *Journal of Political Economics*, 63, 309–321.
- Herzberg, F. (2013). "Aggregation of Monotonic Bernoullian Archimedean preferences: Arrowian impossibility results", Discussion Paper W.P. 488, Center for Mathematical Economics, Bielefeld.
- Hylland, A. , Zeckhauser, R. (1979). "The impossibility of Bayesian group decision making with separate aggregation of beliefs and values", *Econometrica*, 47, 1321–1336.
- Klibanoff, P., Marinacci, M., Mukerji, S. (2005). "A smooth model of decision making under ambiguity", *Econometrica*, 73, 1849–1892.
- Maccheroni, F., Marinacci, M., Rustichini, A. (2006). "Ambiguity aversion, robustness and the variational representation of preferences", *Econometrica*, 74, 1447–1498.
- Machina, M.J. (1989). "Dynamic consistency and non-expected utility models of choice under uncertainty", *Journal of Economic Literature*, 27, 1622–1688.
- Mongin, P. (1995). "Consistent Bayesian aggregation", *Journal of Economic Theory*, 66, 313–351.
- Mongin, P. (1998). "The paradox of the Bayesian expert and state-dependent utility theory", *Journal of Mathematical Economics*, 29, 331–361.
- Mongin, P., d'Aspremont, C. (1998). "Utility theory and ethics", in Barberà, S., Hammond, P.J., Seidl, C. (Eds.), *Handbook of Utility Theory - Volume 1: Principles*, Dordrecht: Kluwer Academic Publishers, 371–481.
- Mongin, P., Pivato, M. (2015). "Ranking multidimensional alternatives and uncertain prospects", *Journal of Economic Theory*, 157, 146–171.
- Myerson, R.B. (1981). "Utilitarianism, egalitarianism, and the timing effect in social choice problems", *Econometrica*, 77, 883–897.
- Rádo, F., Baker, J.A. (1987). "Pexider's equation and aggregation of allocations", *Aequationes Mathematicae*, 32, 277–239.
- Savage, L.J. (1954). *The Foundations of Statistics*, New York: Wiley.
- Schmeidler, D. (1989). "Subjective probability and expected utility without additivity", *Econometrica*, 57, 571–587.
- Weymark, J. A. (1991). "A reconsideration of the Harsanyi-Sen debate on Utilitarianism", in Elster, J., Roemer, J.E. (eds.), *Interpersonal Comparisons of Well-Being*, Cambridge: Cambridge University Press, 255–320.