Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 196 (2021) 105308

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Smooth aggregation of Bayesian experts ☆

Lorenzo Stanca

Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA

Received 10 April 2019; final version received 16 June 2021; accepted 19 June 2021
Available online 26 June 2021

Abstract

I study the ex-ante aggregation of preferences of Bayesian agents in a purely subjective framework. I relax the assumption of a Bayesian social preference while keeping the Pareto condition. Under a simple axiom that relates society's preference to those of the agents, I obtain an additively separable representation of society's preference. Adding an ambiguity aversion axiom I obtain a representation that resembles the Smooth Ambiguity Criterion of Klibanoff et al. (2005). I then briefly consider applications of this framework to inequality and treatment choice under ambiguity.

JEL classification: D71; D81

Keywords: Preference aggregation; Decision under uncertainty; Ambiguity; Precautionary principle; Treatment choice

1. Introduction

Harsanyi's (1955) Aggregation Theorem gave rise to a rich literature on the problem of aggregation of preferences under uncertainty. In such a problem one has N agents (or experts) that have preferences over uncertain prospects. Harsanyi's Aggregation Theorem considered the case of lottery choice with both agents and society's preferences being expected utility. With the further assumption that a Pareto condition is satisfied, he showed that society's preference can be represented by a weighted sum of individual utilities. A large literature subsequently has debated

☆ I am grateful to Peter Klibanoff for his guidance and support. I also thank Massimo Marinacci, Daniele Pennesi, Ernesto Savaglio and Marciano Siniscalchi for very helpful discussions.

E-mail address: lorenzo.stanca@kellogg.northwestern.edu.

https://doi.org/10.1016/j.jet.2021.105308

Table 1
Treatment diversification.

	R_1	R_2
a	2	2
b	1	4

(a)

	R_1	R_2
$\delta b + (1-\delta)a$	$2-\delta$	$2+2\delta$

(b)

both the formal structure and substantive content of this result (for a review of the literature see, among other articles, Mongin and d'Aspremont (1998)). Note that the setting of lottery choice can be thought of as a special case of choice of *acts* (state-contingent plans) in a Savage (1972) setting, with the assumption that agents share a common prior and have Subjective Expected Utility (SEU henceforth) preferences.

While Harsanyi's positive aggregation result is certainly appealing, it relies on the assumption that agents share the same belief. However, in many settings one is confronted with the opposite scenario, i.e. there is divergence on beliefs and not tastes. A typical example is when the agents are experts and society represents a public authority that seeks their advice. In any situation of *Knightian Uncertainty* or *Ambiguity* it is unlikely that experts will have the same belief. This is germane to de Finetti's interpretation of Knightian Uncertainty, where he suggested that what Knight would refer to as a situation of "uncertainty" represents a situation in which there are major discrepancies in the probabilities of different individuals.[1]

I develop axiomatically a model of aggregation of Bayesian experts, i.e. agents with a common utility function but with different beliefs. Such a model resembles the so-called smooth ambiguity model of Klibanoff et al. (2005). I aim to avoid drawbacks of existing models that have been proposed for such a purpose. In the literature, SEU and Maxmin Expected Utility (MEU) models have been proposed to aggregate preferences of experts. However, in a collective decision making setting such both criteria are known to be problematic. As noted for example by Manski (2009), they both ignore treatment diversification when comparing a treatment with unknown returns to one with known returns.

To illustrate, consider an example in which the planner is a public health agency that has to choose a medical treatment for a given population. The planner has to choose between two types of medical treatments, a and b. One treatment is an already existing one with well known effects, while the other is a newly developed treatment. There is disagreement on whether the newly developed treatment can be more effective. Diversification occurs when the individuals are not all assigned to the same treatment. A SEU or MEU criterion would never lead to treatment diversification. Table 1 above provides a numerical example. R_1 and R_2 represent two possible responses to the treatment. Payoffs represent the expected social welfare in each scenario for the two treatments. The two different responses to the treatments can be thought of as representing different medical experts' opinions about how effective the treatment is. In particular, there is agreement on the effects of the known treatment. δ represents the share of the population that will be assigned to treatment b. The table on the right describes the payoffs for the diversified treatment.[2] It is easy to see that for SEU and MEU social preferences either $\delta = 0$ or $\delta = 1$

[1] For example, as translated by Feduzi et al. (2013), De Finetti (1967) states: "[...] what Knight would refer to as 'risks' are cases in which one finds minor discrepancies in valuations made by different individuals, or by different insurers. This is what renders them insurable." For discussion and comments, see Feduzi et al. (2013).

[2] In particular, when the true treatment response is R_1 the payoff is given by $\delta + 2(1-\delta) = 2-\delta$ and analogously for R_2.

are optimal. This is a consequence of certainty independence, an implicit assumption on social preferences implied by the SEU and MEU representation.[3] However, as argued by Manski (2011) diversification is often a sensible course of action.

There are other unintended consequences from explicitly assuming a representation such as MEU. Consider the same setting as before, with the difference that both a and b are two newly developed medical treatments. Moreover, the authority also has to choose the manufacturer of the treatment. Assume there are two manufacturers, 1 and 2. Suppose there are three experts. The first expert prefers the first treatment (irrespective of the manufacturer), while the second considers the second one to be more effective, again irrespective of the manufacturer. The last expert considers the two treatments to be equally effective. However, this third expert strictly prefers the first manufacturer to the second. Formally, let the relevant acts be denoted by the following: a_i denotes vaccine a produced by manufacturer $i = 1, 2$, while b_i denotes vaccine b, produced by manufacturer $i = 1, 2$. Consider first the choice between a_1 and b_1. The first expert prefers a_1, while the second prefers b_1 and the third expert is indifferent. Similarly, when the choice is between a_2 and b_2, the first expert prefers a_2, while the second prefers b_2 and the third expert is again indifferent. Since these profiles of individual preferences are identical across the two choices, it seems inconsistent for the planner to express the preference $b_1 \succ a_1$ but not at the same time $b_2 \succ a_2$. As it will be formalized later, a MEU criterion might violate this consistency principle.

The approach that I will take will suffer from neither of the two issues above. It will allow for treatment diversification and will enforce consistency in choices like in the second example. I introduce two novel axioms. These two axioms are intimately related to the examples previously presented. The first one excludes inconsistencies described in the second example. Along with Pareto and continuity, such an axiom characterizes the criterion

$$V(f) = \sum_{i=1}^{N} \phi_i\left(\int u(f)dP_i\right). \tag{1}$$

The second axiom is related to a notion of ambiguity aversion, requiring that society prefers acts that exhibit low variability in expected utilities. Notably, this axiom implies that treatment diversification is (weakly) preferred. When added to the first one it characterizes the representation

$$V(f) = \sum_{i=1}^{N} \phi\left(\int u(f)dP_i\right)\mu_i. \tag{2}$$

The representation theorems are accomplished in a purely subjective setting; i.e., I do not assume the existence of any objective randomization. The criterion in (2) can be thought of as a social choice version of the so-called smooth ambiguity model (Klibanoff et al. (2005)).

I then briefly consider applications of this criterion. First, I apply it to the econometric literature on treatment choice under ambiguity (Manski (2009)). The application to treatment choice shows how one can use such a criterion to obtain simple comparative statics results. Finally, I consider extensions to include inequality.

[3] The certainty independence axiom was introduced by Gilboa and Schmeidler (1989).

1.1. Literature review

There are several papers that are strictly related to my work. An important paper related to my work is that of Fleming (1952). He obtains an informal derivation of additive representations for society's preferences. Crès et al. (2011) consider the problem of aggregating MEU experts when society is MEU itself. Nascimento (2012) allows for experts that can have very general perceptions of ambiguity and different attitudes toward it. Gajdos and Vergnaud (2013) axiomatically characterize preferences that independently exhibit aversion toward imprecision and conflict. Alon and Gayer (2018) consider a society of SEU agents and find conditions such that society's preferences are represented by a MEU criterion. I discuss these papers in more depth in section 5.2.

In a setting of lottery choice, Grant et al. (2010) characterize a general form of utilitarianism similar to the representation in (1). However, their primitive is much richer, being a preference relation over two types of lotteries: "identity" lotteries and standard outcome lotteries. Billot and Vergopoulos (2016) also consider an alternative model in which society is allowed to formulate probabilities on the opinion that each agent has about the actual state. An extended Pareto condition characterizes the social utility function as a convex combination of the agents' ones and the social prior as the product of individual ones. Mongin and Pivato (2015) employ a different approach to the problem of aggregation. They study a multidimensional array framework that can be thought of as modeling two sources of uncertainty: one subjective and the other objective. They derive an additive representation of preferences over a large class of domains. They then apply these results to a variety of settings, such as choice over uncertain prospects.

Gilboa et al. (2004) and Gajdos et al. (2008) provide results concerning the impossibility of aggregation of agents preferences under the Pareto condition. Following Gilboa et al. (2004), Alon and Gayer (2016) argue against the Pareto principle. They study a society of SEU agents and a social preference that admits a MEU representation. They require two Pareto-type conditions that relax the standard Pareto condition. The first is a Pareto condition for lottery acts as in Gilboa et al. (2004), while the second is a Pareto condition that applies only to acts that return agreed-upon outcomes. Under additional minor conditions, they show that this modification of the Pareto condition is equivalent to social utility being a weighted average of individual utilities, and the social set of priors containing only weighted averages of individual priors. In a related paper, Hayashi and Lombardi (2018) obtain a representation in which each agent is assigned a set of probabilities and society's preference is given. Danan et al. (2016) study aggregation under the assumption of incomplete preference. Society's preferences are represented by a unanimity rule over a set of convex combinations of individual utilities. Sprumont (2018) relaxes Bayesian aggregation by introducing belief-weighted Nash social welfare functions.

The techniques used in my paper are drawn extensively from the literature on additive representation of preferences, with Wakker (1989) as the main reference (see also Wakker and Zank (1999)).

1.2. Structure

Section 2 introduces the formal setting and notation. In Section 3, I present the axioms that are used in the aggregation theorems and discuss the main results. Section 4 considers applications of the model. Section 5 presents a discussion of how to extend the results to a more general setting and a more in depth discussion of strictly related papers. A supplemental appendix with additional results is included in section 8.

2. Setup and assumptions

I adopt the standard decision theoretic set-up. Uncertainty is modeled with a nonempty set S of *states of the world*. For simplicity, assume that S is finite.[4] An *event* is a subset of S. ΔS denotes the set of probability measures over S. Let $X = \mathbb{R}$ be the set of consequences. I denote with $\mathcal{F}$ the set of *acts*, i.e. functions $f : S \to X$. Therefore, the set of acts can be identified with the finite Euclidean space $\mathbb{R}^{|S|}$. Thus, I endow the space of acts with the standard Euclidean topology. A typical element of $\mathcal{F}$ is denoted with letters $f, g, h, \ell, \ldots$

Society is given by a set $I = \{1, \ldots, N\}$ of agents, or experts. Agents' preferences over the set of acts $\mathcal{F}$ are described by weak orders $\succcurlyeq_i \subset \mathcal{F} \times \mathcal{F}$, $i = 1, \ldots N$. As usual, $\succ_i$ and $\sim_i$ denote their asymmetric and symmetric components, respectively. If $f \in \mathcal{F}$, an element $x^i_f \in X$ denotes the certainty equivalent for agent i, i.e. $f \sim_i x^i_f$. A functional $V : \mathcal{F} \to \mathbb{R}$ represents a weak order $\succcurlyeq$ if $V(f) \geq V(g) \iff f \succcurlyeq g$ for every $f, g \in \mathcal{F}$.

Agents are assumed to be SEU maximizers.

Assumption 1 *(SEU agents).* For every $i \in I$, $\succcurlyeq_i$ can be represented by the functional $V_i : \mathcal{F} \to \mathbb{R}$ such that

$$V_i(f) = \int u(f(s))dP_i(s) = \sum_{s \in S} u(f(s))P_i(s),$$

where each P_i is a unique probability measure $P_i \in \Delta S$ and $u : \mathbb{R} \to \mathbb{R}$ is a cardinally unique, strictly increasing continuous utility function over real-valued outcomes. Note that agents share a common Bernoulli utility. Such an assumption is appropriate to model aggregation of experts' opinion, in which the utility function is assumed to coincide with that of a private decision maker who seeks their advice. In subsection 4.2, I present a discussion of a setting with different utilities. In subsection 5.1 I also discuss how to generalize the assumption of unbounded utility. Finally, this common Bernoulli utility function u has unbounded range, i.e.

$$u(\mathbb{R}) = \mathbb{R}.$$

More primitive axioms on preferences $\succcurlyeq_i$ can be found to insure that the above assumption is satisfied. For example, one can use the axioms in Wakker (1989) that characterize SEU preferences with a finite state space. An important consequence of the previous assumption is that for any agent i and $f \in \mathcal{F}$ the certainty equivalent $x^i_f \in X$ always exists.

Since all agents share the same utility over consequences, to aggregate preferences I require the agents' beliefs to be diverse enough.

Assumption 2 *(Linear independence).* The set $\{P_1, \ldots, P_N\}$ is a linearly independent subset of ΔS, i.e., for any given collection of scalars $\alpha_1, \ldots, \alpha_N$

$$\sum_{i=1}^{N} \alpha_i P_i(E) = 0 \quad \forall E \subseteq S \implies \alpha_1 = \ldots = \alpha_N = 0.$$

[4] More generally, what is needed here is that S is a compact metric space. By restricting the attention to finite S, I avoid having to deal with weak topologies.

Note that this assumption implies that $|S| \geq N$. Moreover, it implicitly assumes that no agents share the same belief. To overcome such a problem, one might envision grouping the agents with the same prior under a single agent and giving them more "weight" in the representation. In Section 5.1 I discuss how such an assumption can be relaxed. Interestingly, the same linear independence condition has appeared in Cerreia-Vioglio et al. (2013), with a similar purpose of identifying second order beliefs for a SEU agent. The linear independence assumption is related to the study of identifiability of mixtures of probability measures first examined by Teicher (1963). The next example shows that it is satisfied, for example, when agents have binomial priors and the state space is large enough.

Example 1. Suppose that $S = \{s_1, \ldots, s_n\}$ and $I = \{1, \ldots, N\}$. Each agent $i = 1, \ldots, N$ has a belief P_i over S defined by

$$P_i(s_k) = \binom{n}{k} \cdot p_i^k (1 - p_i)^{n-k},$$

for some $0 < p_1 < \ldots < p_N < 1$. Then as shown in Teicher (1963) (Proposition 4) the set of priors $\{P_i\}_{i=1}^N$ is linearly independent if and only if $|S| = n \geq 2N - 1$.

Another notable case where priors satisfy the independence condition is given by *orthogonal* beliefs, i.e. when experts have radical disagreement in their opinions.[5]

Society's (or the planner's) preference is modeled by another weak order over acts $\succcurlyeq_0$, with asymmetric and symmetric components $\succ_0$ and $\sim_0$, respectively. I do not assume that $\succcurlyeq_0$ has to be Bayesian. A Bayesian planner has preferences represented by

$$\int u(f) dP_0,$$

where $P_0 \in \Delta S$. The main assumption that I make on $\succcurlyeq_0$ is a standard continuity condition.

Assumption 3 *(Social preference).* $\succcurlyeq_0$ is a continuous weak order, i.e. the sets

$$\{f \in \mathcal{F} : g \succcurlyeq_0 f\} \text{ and } \{f \in \mathcal{F} : f \succcurlyeq_0 g\},$$

are closed for every $g \in \mathcal{F}$.

Continuity is required to have a utility function representing society's preferences. Notably, unlike in Mongin (1995) or Gilboa et al. (2004), society's preference is not assumed to have an expected utility representation.

These three assumptions of the utility functions will be needed in the aggregation results to have enough variation in the values of the utility vector

$$\left(\int u(f(s)) dP_1(s), \ldots, \int u(f(s)) dP_N(s) \right).$$

[5] A set of probabilities $\{P_1, \ldots, P_N\}$ is orthogonal if for every $P, P' \in \{P_1, \ldots, P_N\}$, there exists $E \subseteq S$ such that $P(E) = 0 = P'(E^c)$. Then Assumption 2 is trivially satisfied.

3. Aggregation results

3.1. Basic representation

As discussed in the Introduction, the Pareto condition is kept here under full strength.

Pareto condition: For any acts f, g, if $f \succcurlyeq_i g$ for every i then $f \succcurlyeq_0 g$. If additionally $f \succ_i g$ for some i, then $f \succ_0 g$.

The second part of the condition simply states that society cares about every agent. Denote with $\succcurlyeq_B$ any social preference that has an expected utility representation and satisfies the Pareto condition.

To introduce the next axioms, the following definition is required.

Definition 1. Given an act f and agent i denote with f_i any act that satisfies $f \sim_j f_i$ for every $j \neq i$. Formally, f_i is any act that belongs to the set

$$F_f^i = \{h \in \mathcal{F} : h \sim_j f \text{ for every } j \neq i\}.$$

The first axiom is motivated by the second example in the introduction. Consider a situation in which society is faced with two different choice problems. If all the agents regard the two problems as equivalent with the exception of one agent who is indifferent about the choice being made for both decision problems, then *a fortiori* society should have the same preferences.[6]

Agent Independence: For every agent i, acts $f, g \in \mathcal{F}$ and $f_i, g_i \in \mathcal{F}$ such that $f \sim_i g$ and $f_i \sim_i g_i$,

$$f \succcurlyeq_0 g \iff f_i \succcurlyeq_0 g_i.$$

This axiom excludes the types of inconsistencies illustrated in the second example in the introduction. Agent i is indifferent between the two alternatives and therefore should not influence social preferences over these alternatives. In other words, this axiom imposes a form of consistency across different decision problems that could be faced by society. Its conceptual relevance is similar to axioms common in social choice such as independence of irrelevant alternatives (IIA). Like for IIA, an expert who is not concerned with the decision being made (and therefore irrelevant) should not influence society's choice.

Note that when $N = 2$, the axiom is easily seen to be implied by the Pareto condition. Thus, it has strength only when $N \geq 3$. The Agent Independence axiom is equivalent to Postulate E in Fleming (1952) which was also discussed by Harsanyi (1955).[7]

In terms of attitudes toward uncertainty, this axiom has very limited implications. None of the main Savage Axioms (P2, P3, P4) are implied by this axiom. The main content of the axiom

[6] As Fleming (1952) colorfully put it,

> Again, in considering policies affecting the inhabitants of this planet we do not feel hampered by our ignorance regarding states of mind which prevail among the inhabitants of Mars.

[7] In the latter paper Harsanyi claimed that his axioms were weaker than Postulate E. However, as correctly observed later by Fleming (1957), this claim was incorrect.

Table 2
Relevant payoffs.

	s_1	s_2	s_3
f	1	$\frac{1}{2}$	$\frac{4}{5}$
g	$\frac{3}{4}$	1	$\frac{4}{5}$
f_3	1	$\frac{1}{2}$	$\frac{1}{5}$
g_3	$\frac{3}{4}$	1	$\frac{1}{5}$

is consistency across decision problems. The next example shows how a MEU criterion might violate such a consistency principle.

Example 2. Let us go back to the second example in the introduction. I will now show that agent independence is violated when the social preference is MEU.

Here X represents the expected social welfare from the medical treatment. There are three states, i.e. $S = \{s_1, s_2, s_3\}$. Each state reflects a different ordering of effectiveness of the treatment: s_1 means that the first treatment is more effective, s_2 that the second is more effective and s_3 that they are equally effective. Priors are given by $P_1(s_1) = P_2(s_2) = P_3(s_3) = 1$, and the utility function is $u(x) = x$. The available acts are f, g, f_3 and g_3, whose payoffs are listed in Table 2 above.

Together with the probability assignments, such payoffs generate the preference orderings for the experts discussed in the introduction, i.e. $f \succ_1 g$, $f_3 \succ_1 g_3$, $g \succ_2 f$, $g_3 \succ_2 f_3$ and $g \sim_3 f$, $g_3 \sim_3 f_3$. Also observe that $f \sim_i f_3$ and $g \sim_i g_3$, $i = 1, 2$. Suppose that the social preference $\succcurlyeq_0$ is represented by the MEU criterion[8]

$$V(f) = \min_{i \in I} \int f dP_i \text{ for every } f \in \mathcal{F}.$$

Notice that $g \succ_0 f$, since

$$V(f) = \min_{i \in I} \int_{i \in I} f dP_i = \frac{1}{2},$$

and

$$V(g) = \min_{i \in I} \int_{i \in I} g dP_i = \frac{3}{4}.$$

However, it does not hold that $g_3 \succ_0 f_3$, since $V(f_3) = \min_{i \in I} \int_{i \in I} f_3 dP_i = \frac{1}{5}$ and $V(g_3) = \min_{i \in I} \int_{i \in I} g_3 dP_i = \frac{1}{5}$, thus violating Agent Independence.

Theorem 1 shows that when there are at least three agents and both the Pareto condition and the Agent Independence axiom hold, society's preference $\succcurlyeq_0$ admits the representation in (1).

[8] Formally, one should consider the set of convex combinations of P_i, $i = 1, \dots, N$: however, because there are only finitely many probabilities the difference is immaterial.

Theorem 1. *Suppose that $N \geq 3$. Given Assumptions 1-3, $\succcurlyeq_0$ satisfies the Pareto condition and Agent Independence if and only if there exist strictly increasing and continuous functions $\phi_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, N$ such that $\succcurlyeq_0$ is represented by the functional $V_0 : \mathcal{F} \to \mathbb{R}$*

$$V_0(f) = \sum_{i=1}^{N} \phi_i\bigg(\int u(f)dP_i\bigg).$$

Moreover, each function $\phi_i(\cdot)$ is cardinally unique, i.e., if $V_0'(f) = \sum_{i=1}^{N} \hat{\phi}_i\bigg(\int u(f)dP_i\bigg)$ also represents $\succcurlyeq_0$ then $\phi_i = a_i\hat{\phi}_i + b_i$ with $a_i \in \mathbb{R}_{++}$ and $b_i \in \mathbb{R}$.

The full proof of Theorem 1 appears in the Appendix. It is conducted in several steps. First, it is shown that $\succcurlyeq_0$ admits a utility representation $V_0 : \mathcal{F} \to \mathbb{R}$. This easily follows from Assumption 3, since the required continuity condition on $\succcurlyeq_0$ implies the desired result. Then by the Pareto condition one can write V_0 as

$$V_0(f) = \varphi\bigg(\int u(f)dP_1(s), \ldots, \int u(f)dP_N(s)\bigg).$$

Finally, by Agent Independence one can show φ satisfies a notion of "coordinate independence," which gives the representation.[9] It should be noted that Theorem 1 does not actually require that $u(X) = \mathbb{R}$. As elaborated in the supplemental appendix, one could obtain the same representation without such a richness condition, although this representation may not encompass all acts.

The functions $(\phi_i)_{i=1}^N$ in the representation measure both weights assigned to experts (analogous to Harsanyi's aggregation result) and society's attitudes toward disagreement between them. For example, suppose that $\phi_1 = \phi_2 = \ldots = \phi_N = \phi$, with ϕ concave. Then an act whose expected utility is agreed upon by all the experts will be (weakly) preferred to any other act that has the same average expected utility but with different expected utility evaluations from the experts. In other words, as much as disagreement between experts can be thought of as reflecting ambiguity, the functions $(\phi_i)_{i=1}^N$ reflect attitudes toward ambiguity. In the following section I make this connection more formal. In particular, I obtain a representation that reflects ambiguity aversion of the planner.

3.2. Ambiguity aversion

I introduce an axiom that describes when society is ambiguity averse. The idea is that society should have a preference for acts that have lower variability in expected utilities. To formalize such a notion, I introduce the set of unambiguous acts. Consider first the set of events whose probabilities are agreed upon

$$\Lambda = \{E \subseteq S : \forall i, j \in I, P_i(E) = P_j(E)\}.$$

Definition 2. An event $E \subseteq S$ is *unambiguous* if $E \in \Lambda$. It is *ambiguous* if it is not unambiguous. Λ denotes the collection of unambiguous events. An act f is unambiguous if it is measurable with respect to Λ, i.e. $f^{-1}(c) \in \Lambda$ for every $c \in \mathbb{R}$. It is *ambiguous* if it is not unambiguous.

[9] The notion of independence that is used is the same as that in Debreu (1959).

Unambiguous acts formalize the idea of an act that involves no volatility in its expected outcome, since all the agents agree on its expected value. I define ambiguity as a preference for unambiguous acts. This idea is formalized as follows. The axiom requires that there exists a Bayesian and Pareto planner such that whenever this planner is indifferent between an act and unambiguous act, then society should (weakly) prefer the unambiguous act. In this sense society is required to be "more cautious" than a Bayesian planner that satisfies the Pareto condition. An important consequence of this axiom is that when considering a treatment choice setting, diversification will always be (weakly) preferred. Indeed, treatment diversification is able to smooth out the expected social welfare. As discussed in the introduction, a preference for treatment diversification is one of the main desiderata of the theory studied in this paper.

Ambiguity aversion: There exists a Bayesian and Pareto preference $\succsim_B$ such that for every $f \in \mathcal{F}$ and unambiguous g

$$g \sim_B f \implies g \succsim_0 f.$$

The idea of using expected utility as a benchmark against which to measure ambiguity aversion comes from Ghirardato and Marinacci (2002).[10] It is important to note that Klibanoff et al. (2005) consider a weaker notion of unambiguous event inspired by the Ellsberg two-color experiment.[11]

The following result shows that, when paired with Agent Independence, this axiom characterizes the criterion (2).

Theorem 2. *Suppose that $N \geq 3$. Given Assumptions 1-3, $\succsim_0$ satisfies the Pareto condition, Agent Independence and ambiguity aversion if and only if there exist a concave and strictly increasing function $\phi : \mathbb{R} \to \mathbb{R}$ and positive weights $(\mu_i)_{i=1}^N$ with $\sum_{i=1}^N \mu_i = 1$ such that $\succsim_0$ is represented by*

$$V_0(f) = \sum_{i=1}^{N} \phi\left(\int u(f(s)) dP_i \right) \mu_i .$$

Moreover, if $V_0'(f) = \sum_{i=1}^N \hat{\phi}\left(\int u(f(s)) dP_i(s) \right) \hat{\mu}_i$ also represents $\succsim_0$ then $\phi = a\hat{\phi} + b$ with $a > 0$ and $\mu_i = \hat{\mu}_i$ for every $i = 1, \ldots, N$.

In the representation, concavity of ϕ reflects ambiguity aversion, which has great importance in applications. Unlike Theorem 1, society's attitude toward ambiguity are described by only one function. Moreover, the weights assigned to each expert are disentangled from such attitudes. Such weights $(\mu_i)_i$ can be interpreted as in Harsanyi's utilitarian aggregation setting: μ_i

10 The main difference from Ghirardato and Marinacci (2002) is that they use constant acts to make the comparison with expected utility. Since I work in a "richer" setting, I can use a richer class of acts to make such a comparison.

11 In their setting, conditions on ϕ are required for an unambiguous event to have agreement on its probability (see their Theorem 3, p. 1871). For instance, if ϕ is linear then any event is unambiguous according to their definition, while according to my definition the ambiguity of an event is independent of the shape of the function ϕ. I can take the stronger definition here because I am considering a setting of aggregation of preferences, in which the set of beliefs is taken as a primitive representing the experts' opinions.

represents the weight assigned by the planner to expert i. In the proof, the weights are derived by applying a Harsanyi-type representation theorem to the Bayesian planner and combining this result with the ambiguity aversion axiom (see Lemma 9 for details). Therefore, society weighs each expert by means of the same weights used by the Bayesian planner in the axiom.

Because of this result, one can formally define comparative ambiguity aversion.

Definition 3. Society $\succcurlyeq_0$ is more ambiguity averse than society $\hat{\succcurlyeq}_0$, both having the representation from Theorem 2 with functions ϕ and $\hat{\phi}$ and the same weights $(\mu_i)_{i=i}^N$, if for every act f and unambiguous act g,

$$g \hat{\succcurlyeq}_0 f \implies g \succcurlyeq_0 f.$$

Proposition 1. *$\succcurlyeq_0$ is more ambiguity averse than society $\hat{\succcurlyeq}_0$ if and only if there exists a strictly increasing, concave function $h : \mathbb{R} \to \mathbb{R}$ such that $\phi(x) = h(\hat{\phi}(x))$ for every $x \in \mathbb{R}$.*

In particular, $\succcurlyeq_0$ will be a Bayesian social preference whenever $\succcurlyeq_B$ is more ambiguity averse than $\succcurlyeq_0$.

In the next section, I discuss applications of this model.

4. Applications and extensions

4.1. Treatment diversification

In several papers (see for example Manski (2009) or Manski (2013)), Manski has studied the important problem of treatment choice under ambiguity. As discussed in the introduction, the main example is a public authority that has to decide which type of vaccination to administer to individuals who belong to a heterogeneous population. A MEU or a Bayesian criterion would almost always choose to assign all the population to only one treatment (Manski (2009)).

For this reason, Manski proposed the minimax regret, introduced by Savage (1951), as a social choice criterion. However, this criterion is known to have many flaws as recognized by Savage himself.[12] Notably, it is not possible to obtain a simple comparative statics result with a regret criterion. Indeed, in general it is not possible to find a closed form solution when such a criterion is applied (see for example Manski and Tetenov (2007), pp. 2006-2007). Because of this difficulty, it becomes unfeasible to study how the solution behaves when parameters change. On the contrary, when ϕ is strictly concave, the criterion introduced in this paper possesses a preference for treatment diversification. Manski and Tetenov (2007) consider an equivalent criterion and show that it achieves treatment diversification.[13] Here I provide an example that shows how one can obtain simple comparative statics results when applied to these type of problems.

Example 3. Let us go back to the example in the introduction on treatment choice. There are only two possible population distributions of treatment responses. Each treatment response represents an expert's opinion about the effectiveness of the medical treatment. Treatment a response is

[12] In chapter 10 of the Foundations of Statistics, he extensively discussed drawbacks of minimax regret. See Savage (1972).

[13] Also noted by Klibanoff (2013) and Marinacci (2015).

known, whereas b is a newly proposed treatment with uncertain effectiveness. Thus $S = \{P_1, P_2\}$. Let the expected social welfare for each treatment under each distribution in $P \in S$ be as follows:

	P_1	P_2
a	2	2
b	1	4

Then one can calculate that the expected social welfare for any treatment allocation $\delta \in [0, 1]$ is $W_P(\delta)$, whose values are given by:

	$W_{P_1}(\delta)$	$W_{P_2}(\delta)$
$\delta b + (1-\delta)a$	$2-\delta$	$2+2\delta$

Assume that the planner has preferences given by

$$V(\delta) = \int [\phi(W_P(\delta))]\, d\mu(P) = \mu(P_1)\phi(2-\delta) + (1-\mu(P_1))\phi(2+2\delta).$$

From this, one can calculate the optimal diversified treatment. For example, assume that $\phi(x) = \left(x^{1-\alpha}\right)/(1-\alpha)$ with $0 < \alpha \neq 1$ and $\mu(P_1) = \frac{1}{2}$. Then we get the first order condition

$$-\frac{1}{2}(2-\delta)^{-\alpha} + (2+2\delta)^{-\alpha} = 0,$$

which gives as a solution $\delta^* = 2\frac{(2^{1/\alpha}-1)}{(2+2^{1/\alpha})}$.

Observe that δ^* is decreasing in α, i.e. treatment diversification decreases the higher the ambiguity aversion, a simple comparative statics result.

4.2. Allowing for private goods

So far, I have considered a setting in which each agent $i = 1, \ldots, N$ gets the same outcome in every state of the world. However, in many settings it is relevant to allow for different outcomes for different agents. For instance, consider a situation in which looming climate change will alter the distribution of well-being on Earth. Suppose there are two states of the world. In one, the extreme latitudes gain and the low latitudes suffer, whereas the reverse occurs in the other scenario. Likewise, consider a setting in which the planner has to allocate a public good to different generations. In both settings it is not appropriate to have agents share a common outcome. In this section I extend the model to accommodate this concern.

Formally, suppose that each agent i can get a different outcome x_i, so that we have $X = \mathbb{R}^I$. Thus, an act associate maps each state $s \in S$ into a vector of outcomes $f(s) \in \mathbb{R}^I$.

Assumption 4. Each agent i has SEU preferences with utility function $u_i : \mathbb{R} \to \mathbb{R}$, assumed to be continuous, strictly increasing and satisfying $u_i(X) = \mathbb{R}$. The utilities are assumed to be fully measurable and interpersonally comparable. Moreover, the social preference $\succcurlyeq_0$ over acts is continuous.

In this section I discuss how the model introduced in this paper can be generalized to such a setting and how it can capture a concern for inequality as well. Other papers have studied similar issues, such as Epstein and Segal (1992), Adler and Sanchirico (2006), Grant et al. (2010)) and Fleurbaey (2010).

The axiomatization in this setting needs only an adjustment to the ambiguity aversion axiom. Say that $f \in \mathcal{F}$ is socially unambiguous if $\int u_i(f)dP_i = \int u_j(f)dP_j$ for every $i, j \in I$. To illustrate, when agents share the same utility an act is socially unambiguous if and only if it is unambiguous. In this more general setting $\succcurlyeq_B$ denotes a social planner with preferences represented by

$$V(f) = \sum_{i=1}^{N} \left(\int u_i(f)dP_i \right) \mu_i,$$

for some set of positive weights $(\mu_i)_{i=1}^N$ that sum to one. The ambiguity aversion axiom can now be rephrased as follows

Social ambiguity aversion: There exists a social planner $\succcurlyeq_B$ such that for every $f \in \mathcal{F}$ and socially unambiguous act g

$$g \sim_B f \implies g \succcurlyeq_0 f.$$

When agents share the same utility, the ambiguity aversion axiom captures only aversion to variability in the expected utilities. However, in this more general setting, the axiom also captures aversion to the inequality of the allocation.

Theorem 3. *Suppose that $N \geq 3$. Given Assumption 4, $\succcurlyeq_0$ satisfies the Pareto condition, Agent Independence and social ambiguity aversion if and only if there exist a concave and strictly increasing function $\phi : \mathbb{R} \to \mathbb{R}$ and positive weights $(\mu_i)_{i=1}^N$ with $\sum_{i=1}^N \mu_i = 1$ such that $\succcurlyeq_0$ is represented by*

$$V_0(f) = \sum_{i=1}^{N} \phi\left(\int u_i(f(s))dP_i \right) \mu_i.$$

Moreover, if $V_0'(f) = \sum_{i=1}^N \hat{\phi}\left(\int u_i(f(s))dP_i(s) \right) \hat{\mu}_i$ also represents $\succcurlyeq_0$ then $\phi = a\hat{\phi} + b$ with $a > 0$ and $\mu_i = \hat{\mu}_i$ for every $i = 1, \ldots, N$.

To further understand this criterion, observe that when there is no uncertainty it reduces to

$$\sum_{i=1}^{N} \mu_i \phi(u_i(x_i)),$$

for every $(x_i)_{i=1}^N \in \mathbb{R}^N$. Therefore, in this case the degree of concavity of ϕ captures aversion to inequality of the allocation. When f is such that $u_i(f(s)) = u_j(f(s))$ for every $i, j = 1, \ldots, N$ and $s \in S$, then we obtain the original criterion

$$V(f) = \sum_{i=1}^{N} \mu_i \phi\left(\int u_i(f(s))dP_i(s) \right).$$

Hence, the degree of concavity of ϕ captures attitudes toward both ambiguity and inequality. However, one might want to be able to separate these two attitudes which should not necessarily be measured by the same object. Distinguishing between these two is left to future research.

5. Discussion

I have shown that if one drops the assumption a Bayesian social preference while keeping the assumption of Bayesian agents it is still possible to obtain a rich theory of social decisions that respects the Pareto condition. In particular, in such a theory the departure from a Bayesian social preference can be understood by means of ambiguity. Moreover, this framework can be used to study important applied instances of collective decision making such as treatment choice.

5.1. Extensions and alternative representations

Assumptions 1 and 2 together imply that

$$\left\{\left(\int u(f)dP_1, \dots, \int u(f)dP_N\right) : f \in \mathcal{F}\right\} = \mathbb{R}^N. \tag{3}$$

This result allows me to adopt the techniques from the literature on additive representation of preferences. It is possible to weaken these two assumptions and generalize Theorem 1. In particular, one can allow the common utility u to be bounded and drop Assumption 2 entirely. This result could be accomplished by means of representation theorems for functions that satisfy an independence condition on non-product sets (see Segal (1992) and Chateauneuf and Wakker (1993)). However, such a representation may not apply to all acts. A precise statement is elaborated in the supplemental appendix (see Theorem 8).

A major assumption of the paper is that all agents have the same utility. Observe that if agents have different utilities, (3) is not guaranteed to hold. To preserve Theorem 1, one would have to directly make an assumption analogous to (3).

Authors in the literature have questioned the validity of the Pareto condition. For example, Mongin (1997) (pp. 524-525) presents an example in which two experts might agree that an event is more likely than its complement but society need not agree with this judgment. If one wanted to accommodate such a concern, one could adopt the following weakening of Pareto

For any acts f, g, if $f \sim_i g$ for every $i = 1, \dots, N$ then $f \sim_0 g$.

Under such an assumption, it is possible to obtain a representation analogous to that of Theorem 2 but with weights that are not necessarily all positive.

5.2. Discussion of related papers

Gajdos et al. (2008) obtain a negative aggregation result that generalizes Mongin (1995). They consider a planner and agents whose preferences belong to a large class that includes most known models of decision under uncertainty. They show that aggregating agents' preferences is possible and necessarily linear if and only if (society's) preferences are uncertainty neutral. More precisely, they show that if an agent has a non neutral attitude towards uncertainty, then he must be either a dictator or he is assigned zero weight by society.

Crès et al. (2011) consider a setting of aggregation of experts' opinion. Society and the experts have MEU preferences with different sets of priors. They also introduce a notion of ambiguity aversion for society, which they call Expert Uncertainty Aversion (EUA). Similarly to the ambiguity aversion axiom, their axiom also requires society to be averse to disagreement between experts. Differently from my paper, however, their axiom is stated in terms of mixture of acts.

Formally, it requires that if each expert evaluates an act f above the weighted average of the evaluations of other acts $f_1, \ldots, f_m$, so should society. They show that EUA is equivalent to the existence of a set of probability vectors over the experts (interpreted as possible allocations of weights to the experts) such that the decision maker's set of priors is given by all the weighted-averages of priors.

Nascimento (2012) studies aggregation of preferences of experts that agree on the ranking of risky prospects, but have different perceptions and attitudes towards ambiguity. He considers the original Anscombe-Aumann framework (the primitive being a preference over lotteries over acts) used by Seo (2009). He obtains an aggregation rule that, when restricted to MEU preferences, is a generalization of the decision criterion from Crès et al. (2011). His main axiom, convexity, also conveys the idea that society should hedge against the uncertainty about which expert is right. Differently from my axiom, it is stated in terms of mixture of acts.

Gajdos and Vergnaud (2013) provide an axiomatic foundation for a decision criterion that allows one to distinguish on a behavioral basis the decision maker's attitude towards imprecision and disagreement. Experts are modeled as MEU maximizers. They assume that society satisfies a Pareto-like axiom and a notion of aversion to disagreement between experts. They show that such assumptions are equivalent to a decision criterion in which society aggregates the experts' preferences with a two step procedure: first, society transforms experts' information through a function that represents society's attitude toward conflicting information, and uses the resulting sets of probability to evaluate the act under consideration. Second, society aggregates linearly the experts' evaluations, using the worst weight vector in a set, which models society's attitude toward disagreement between experts.

Alon and Gayer (2018) consider the aggregation of SEU agents, allowing for different opinions and tastes. Their main axiom, called ambiguity aversion, requires society to have a preference for hedging. However, differently from the literature on ambiguity aversion, a mixture between acts is defined by means of unambiguous partitions of the state space. They provide an axiomatic characterization of a planner who follows a MEU decision rule. Moreover, the derived set of priors is linked to the agents' prior in that it never exceeds the convex hull of the underlying set of prior probabilities. Notably, the utility function need not be related to the agents' utilities in any way.

6. Appendix

6.1. *Proof of Theorem 1*

I divide the proof into several steps. The first standard Lemma asserts the existence of a certainty equivalent of any act for society's preference.

Lemma 1. *For any $f \in \mathcal{F}$, there exists $x_f \in X$ such that $x_f \sim_0 f$.*

Proof. The proof is standard and therefore omitted. □

Lemma 2. *There exists a continuous $V_0 : \mathcal{F} \to \mathbb{R}$ that represents $\succsim_0$.*

Proof. The proof is standard and therefore omitted. □

Now let $\mathcal{U} = \{u(x) : x \in X\}$. By Assumption 1, it holds $\mathcal{U} = \mathbb{R}$.

Let R be the set defined by

$$R = \left\{ \left(\int u(f(s))dP_1(s), \dots, \int u(f(s))dP_N(s) \right) : f \in \mathcal{F} \right\}.$$

The following Lemma gives an alternative characterization of R.

Lemma 3. *It holds*

$$R = \left\{ \left(\int g(s)dP_1(s), \dots, \int g(s)dP_N(s) \right) : g \in \mathbb{R}^S \right\}.$$

Proof. It is easy to see that

$$R \subseteq \left\{ \left(\int g(s)dP_1(s), \dots, \int g(s)dP_N(s) \right) : g \in \mathbb{R}^S \right\}.$$

Indeed, if $\mathbf{x} \in R$ then

$$\mathbf{x} = \left(\int u(f(s))dP_1(s), \dots, \int u(f(s))dP_N(s) \right),$$

for some $f \in \mathcal{F}$. But then letting $g : S \to \mathbb{R}$ be defined by $g \equiv u \circ f$, we obtain

$$\mathbf{x} = \left(\int g(s)dP_1(s), \dots, \int g(s)dP_N(s) \right),$$

so that

$$\mathbf{x} \in \left\{ \left(\int g(s)dP_1(s), \dots, \int g(s)dP_N(s) \right) : g \in \mathbb{R}^S \right\},$$

as desired.

Conversely, if

$$\mathbf{x} \in \left\{ \left(\int g(s)dP_1(s), \dots, \int g(s)dP_N(s) \right) : g \in \mathbb{R}^S \right\},$$

then

$$\mathbf{x} = \left(\int g(s)dP_1(s), \dots, \int g(s)dP_N(s) \right),$$

for some $g : S \to \mathbb{R}$. Define $f : S \to \mathbb{R}$ by

$$f(s) = u^{-1}(g(s)) \quad \forall s \in S.$$

Note that f is well defined since $u : X \to \mathbb{R}$ is strictly increasing. Then clearly $g = u \circ f$, so that

$$\mathbf{x} = \left(\int u(f(s))dP_1(s), \dots, \int u(f(s))dP_N(s) \right).$$

Hence $\mathbf{x} \in R$. This shows that

$$R = \left\{ \left(\int g(s)dP_1(s), \dots, \int g(s)dP_N(s) \right) : g \in \mathbb{R}^S \right\},$$

as desired. □

Lemma 4. *Under the independence assumption, it holds*

$$R = \prod_i \mathcal{U}_i = \mathcal{U}^N = \mathbb{R}^N. \tag{4}$$

The proof of the above results requires the following preliminary result, which essentially is a reformulation of Theorem 1 in Teicher (1963). Let $s_1, \dots, s_{|S|}$ denote an enumeration of S.

Lemma 5. *Under Assumption 2 (Linear Independence) the matrix*

$$\begin{bmatrix} P_1(s_1) & P_2(s_1) & \dots & P_N(s_1) \\ P_1(s_2) & P_2(s_2) & \dots & \\ \vdots & & \ddots & \\ P_1(s_{|S|}) & P_2(s_{|S|}) & \dots & P_N(s_{|S|}) \end{bmatrix} \tag{5}$$

has full rank.

Proof. Suppose that the matrix

$$\begin{bmatrix} P_1(s_1) & P_2(s_1) & \dots & P_N(s_1) \\ P_1(s_2) & P_2(s_2) & \dots & \\ \vdots & & \ddots & \\ P_1(s_{|S|}) & P_2(s_{|S|}) & \dots & P_N(s_{|S|}) \end{bmatrix}$$

does not have full rank. But this implies that there are coefficients $\alpha_1, \dots, \alpha_N$, with some different from zero, such that

$$\sum_{i=1}^N \alpha_i P_i(s) = 0 \quad \forall s \in S.$$

But then it is easy to see that

$$\sum_{i=1}^N \alpha_i P_i(E) = 0 \quad \forall E \subset S,$$

which contradicts Assumption 2 (Linear Independence) since there exists $\alpha_i \neq 0$. Indeed, any E can be written as a disjoint union $E = \cup_{s \in E}\{s\}$ so that

$$\sum_{i=1}^N \alpha_i P_i(E) = \sum_{i=1}^N \alpha_i P_i(\cup_{s\in E}\{s\}) = \sum_{i=1}^N \alpha_i \sum_{s\in E} P_i(s) = \sum_{s\in E}\sum_{i=1}^N \alpha_i P_i(s) = \sum_{s\in E} 0 = 0,$$

which gives the desired contradiction. □

We are ready for the proof of Lemma 4. As standard, I identify the set $\mathbb{R}^S$ of all functions from S to $\mathbb{R}$ with the vector space $\mathbb{R}^{|S|}$.

Proof of Lemma 4. Let $T^* : \mathbb{R}^S \to \mathbb{R}^N$ be the map defined by

$$T^*(f) \equiv \left(\int f(s) dP_1(s), \dots, \int f(s) dP_N(s) \right).$$

I claim that $T^*(\mathbb{R}^S) = \mathbb{R}^N$. To see this, note that by Lemma 5 the matrix

$$M = \begin{bmatrix} P_1(s_1) & P_2(s_1) & \dots & P_N(s_1) \\ P_1(s_2) & P_2(s_2) & \dots & \\ \vdots & & \ddots & \\ P_1(s_{|S|}) & P_2(s_{|S|}) & \dots & P_N(s_{|S|}) \end{bmatrix}$$

has full rank and recall that $|S| \geq N$. Now note that T^* can be written as

$$T^*(\mathbf{x}) = M^T \mathbf{x}.$$

Since M^T has full rank it follows

$$T^*(\mathbb{R}^S) = \mathbb{R}^N.$$

Thus by Lemma 3 it follows that

$$T(\mathbb{R}^S) = T^*(\mathbb{R}^S) = \mathbb{R}^N,$$

as desired. □

The next simple results states that the map T is continuous.

Lemma 6. *The map $T : \mathcal{F} \to R$ defined by*

$$T(f) = \left(\int u(f(s))dP_1(s), \dots, \int u(f(s))dP_N(s) \right), \tag{6}$$

is continuous.

Proof. The proof is routine and thus omitted. □

Below we will need the following independence notion well known in the literature (see for example Wakker (1989).

Definition 4. A function $f : E \subset R^N \to \mathbb{R}$ is called *completely separable* if for every $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N)$, $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_i, \dots, \mathbf{y}_N)$, $\mathbf{y}' = (\mathbf{x}_1, \dots, \mathbf{y}_i, \dots, \mathbf{x}_N)$ and $\mathbf{x}' = (\mathbf{y}_1, \dots, \mathbf{x}_i, \dots, \mathbf{y}_N)$ in E,

$$\varphi(\mathbf{x}) \geq \varphi(\mathbf{y}) \\ \iff \varphi(\mathbf{x}_{-i}, \mathbf{y}_i) \geq \varphi(\mathbf{y}).$$

Compared to usual results on additive preferences (Gorman (1968), Lemma 1), proving that φ is continuous requires a different argument. The complication is that in the standard additively separable setting each coordinate is independent from every other one, i.e. φ can be written as $\varphi(u_1(x_1), \dots, u(x_N)))$, whereas here an act f determines the value for all coordinates. The next lemma delivers this result.

Lemma 7. *There exists $\varphi : R \subset \mathbb{R}^N \to \mathbb{R}$ such that $V_0(f) = \varphi(V_1(f), \dots, V_N(f))$ for every $f \in \mathcal{F}$. Moreover, φ satisfies the following properties*

(1) *φ is continuous.*
(2) *For any $\mathbf{x}, \mathbf{y} \in R$ if $\mathbf{x}_i \geq \mathbf{y}_i$ for every $i = 1, \ldots N$ and $\mathbf{x}_j > \mathbf{y}_j$ for some j then $\varphi(\mathbf{x}) > \varphi(\mathbf{y})$.*
(3) *φ is completely separable.*

Proof. As standard, for any $\mathbf{x}' \in R$ define $\varphi : R \to \mathbb{R}$ by

$$\varphi(\mathbf{x}') = V(f),$$

where f is any act that satisfies $T(f) = \mathbf{x}'$. To prove that φ is well defined, I show that $\varphi(T(f))$ is independent of the choice of f. If f, h are such that $T(f) = T(h)$ this means that

$$f \sim_i h \quad \text{for all } i \in I.$$

Hence since this implies that $f \succcurlyeq_i h$ and $h \succcurlyeq_i f$ for every $i \in I$ by the Pareto condition we obtain $f \sim_0 h$ so that $\varphi(T(f)) = V_0(f) = V_0(h) = \varphi(T(h))$, as desired.

To prove (1), let $(\mathbf{x}_n)_{n=1}^{\infty}$ be a sequence that converges to $\mathbf{x} \in \mathbb{R}^N$. I construct a sequence $(f_n)_{n\geq 1}$ such that $T(f_n) = \mathbf{x}_n$ for every $n \geq 1$ and $f_n \to f$ with $T(f) = \mathbf{x}$. This would imply that $\varphi(\mathbf{x_n}) \to \varphi(\mathbf{x})$ and thus that φ is continuous. Indeed, because V_0 and T are continuous then $V_0(f_n) \to V_0(f)$ and $T(f_n) \to T(f) = \mathbf{x}$ so that

$$\varphi(\mathbf{x}_n) = V_0(f_n) \to V_0(f) = \varphi(T(f)) = \varphi(\mathbf{x}),$$

as desired. To construct the sequence f_n, consider the sequence of sets given by

$$I_n = \{f \in \mathcal{F} : T(f) = \mathbf{x}_n\}.$$

By continuity of T, each I_n is a closed set. This implies that we can find a closest element to f in every I_n. Let this element be f_n. To see why f_n converges to f, consider the ball

$$N = \{f' : ||f - f'||_1 \leq K, K > 0\},$$

around f ($|| \cdot ||_1$ denotes the ℓ_1 norm in $\mathbb{R}^N$). Then $T(N)$ is a neighborhood for $\mathbf{x}$. To see this, note that $T(N) = T^*(N^*)$ where

$$N^* = \{u \circ f' : f' \in N\}.$$

Clearly N^* is also a neighborhood of $u \circ f$ (recall that u is continuous and strictly increasing; thus, it is a homeomorphism). Since $T^*(u \circ f) = \mathbf{x}$ and by Assumption 2 (Linear Independence) T^* is a surjective linear operator, and thus by the Open Mapping Theorem (e.g. see Rudin (1973), Theorem 2.11), $T^*(\text{int}N^*)$ (int denotes the topological interior) is an open set so that

$$\mathbf{x} = T^*(u \circ f) \in T^*(\text{int}N^*).$$

In other words, $T^*(N^*) = T(N)$ is a neighborhood for $\mathbf{x}$. Hence there exists k such that

$$\mathbf{x}_n \in T(N) \text{ for every } n \geq k.$$

Let $f_n \in \mathcal{F}$ be such that $T(f_n) = \mathbf{x}_n$ (these exist by Lemma 4). Therefore, $f_n \in N$ for every $n \geq k$. Hence $f_n \to f$ as desired. By the previous reasoning, this concludes the proof.

As for (2), this is a straightforward application of the Pareto condition. Indeed, suppose that $\mathbf{x}$ and $\mathbf{y}$ satisfy

$$\mathbf{x_i} \geq \mathbf{y_i}, \tag{7}$$

for every $i \in I$. Let $T(f) = \mathbf{x}$ and $T(g) = \mathbf{y}$. Now (7) implies that

$$f \succsim_i g \text{ for every } i \in I.$$

Hence by the Pareto principle $f \succsim_0 g$, i.e. $g(\mathbf{x}) = V_0(f) \geq V_0(g) = g(\mathbf{y})$ as desired. If $\mathbf{x}_i > \mathbf{y}_i$ for some i by the same reasoning as above $f \succ_i g$ and by the Pareto condition $f \succ_0 g$ so that $g(\mathbf{x}) = V_0(f) > V_0(g) = g(\mathbf{y})$ as desired.

Finally, to verify (4), let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N)$, $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_i, \ldots, \mathbf{y}_N)$, $\mathbf{y}' = (\mathbf{x}_1, \ldots, \mathbf{y}_i, \ldots, \mathbf{x}_N)$ and $\mathbf{x}' = (\mathbf{y}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{y}_N)$ be elements of R. By definition, there exist $f, g, f', g' \in \mathcal{F}$ such that $T(f) = \mathbf{x}$, $T(g) = \mathbf{x}'$, $T(f') = \mathbf{y}$ and $T(g') = \mathbf{y}'$. Moreover, by the definition of $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}'$, we have that

$$f \sim_i g \text{ and } f' \sim_i g'$$

and for all $j \in I \setminus \{i\}$,

$$f \sim_j f' \text{ and } g \sim_j g'.$$

By Agent Independence,

$$f \succsim g \iff f' \succsim g',$$

so that by Lemma 7, and which gives

$$\varphi(\mathbf{x}) \geq \varphi(\mathbf{y})$$
$$\iff \varphi(\mathbf{x}_{-i}, \mathbf{y}_i) \geq \varphi(\mathbf{y}),$$

as desired. □

The next key lemma needs the following important result proved in Wakker (1989).

Definition 5. Given a weak order $\succsim$ on a product set $\Gamma^n = \prod_{i=1}^n \Gamma$, a coordinate $i \in \{1, \ldots, n\}$ is said *essential* if there exist $\gamma_i, \lambda_i \in \Gamma$ and γ_{-i} such that

$$\gamma_i \gamma_{-i} \succ \lambda_i \gamma_{-i}.$$

Definition 6. A weak order $\succsim$ on a product set Γ^n satisfies coordinate independence (CI) if, for all $i \in \{1, \ldots, n\}$ and $x, y \in \Gamma^n$ and $\gamma, \lambda \in \Gamma$ it holds

$$\gamma x_{-i} \succsim \lambda x_{-i} \iff \gamma y_{-i} \succsim \lambda y_{-i}.$$

Theorem 4 *(Wakker (1989), Theorem III.6.6.). Let $\succsim$ be a weak order on defined Γ^n, with Γ being a connected topological space and $n \geq 3$. Consider a weak order $\succsim$ such that every coordinate is essential. Then $\succsim$ satisfies continuity and CI if and only if it is represented by $U : \Gamma^n \to \mathbb{R}$*

$$U(\gamma) = \sum_{i=1}^n u_i(\gamma_i),$$

for every $\gamma \in \Gamma^n$.

Moreover, each u_i is cardinally unique.

Lemma 8. *There exist continuous and strictly increasing functions $\phi_i : \mathcal{U} \to \mathbb{R}$ and an increasing function h such that $\varphi(\mathbf{x}) = h(\sum_{i=1}^N \phi_i(\mathbf{x}_i))$.*

Proof. Define the weak order $\succsim^*$ on $\mathcal{U}^N$ by

$$\mathbf{x} \succsim^* \mathbf{y} \iff \varphi(\mathbf{x}) \geq \varphi(\mathbf{y}),$$

for every $\mathbf{x}, \mathbf{y} \in \mathcal{U}^N$. By Lemma 7, $\succsim_0^*$ satisfies coordinate independence and continuity with every coordinate being essential. Therefore by Theorem 4 there exists $\phi_i : \mathcal{U} \to \mathbb{R}$ such that

$$\mathbf{x} \succsim^* \mathbf{y} \iff \sum_{i=1}^{N} \phi_i(\mathbf{x}_i) \geq \sum_{i=1}^{N} \phi_i(\mathbf{y}_i), \tag{8}$$

for every $\mathbf{x}, \mathbf{y} \in \mathcal{U}^N$; moreover, by the Pareto condition each ϕ_i must be strictly increasing. Thus,

$$\varphi(\mathbf{x}) \succsim^* \varphi(\mathbf{y}) \iff \sum_{i=1}^{N} \phi_i(\mathbf{x}_i) \geq \sum_{i=1}^{N} \phi_i(\mathbf{y}_i),$$

for every $\mathbf{x}$, $y \in \mathcal{U}^N$. This means that there must be a strictly increasing function h such that

$$\varphi(\mathbf{x}) = h\left(\sum_{i=1}^{N} \phi_i(\mathbf{x}_i)\right),$$

as desired. Moreover, if there exist other functions ϕ_i' such that (8) holds then it must be that $\phi_i' = a\phi_i + b$ for $a, b \in \mathbb{R}$ with $a > 0$. □

We are now ready to show Theorem 1.

Proof of Theorem 1. Sufficiency follows by a straightforward application of Lemma 8. As for necessity, consider any acts f, g and every f_i, g_i such that $f \sim_i g$, $f_i \sim_i g_i$. Using the representation we get

$$\sum_{j=1}^{N} \phi_j\left(\int u(f(s))dP_j(s)\right) = \phi_i\left(\int u(f(s))dP_i(s)\right) + \sum_{j \neq i}^{N} \phi_j\left(\int u(f(s))dP_j(s)\right),$$
$$\sum_{j=1}^{N} \phi_j\left(\int u(g(s))dP_j(s)\right) = \phi_i\left(\int u(g(s))dP_i(s)\right) + \sum_{j \neq i}^{N} \phi_j\left(\int u(g(s))dP_j(s)\right).$$

Now note that

$$\phi_i\left(\int u(f(s))dP_i(s)\right) = \phi_i\left(\int u(g(s))dP_i(s)\right),$$

so that $f \succsim_0 g$ if and only if

$$\sum_{j \neq i}^{N} \phi_j\left(\int u(f(s))dP_j(s)\right) \geq \sum_{j \neq i}^{N} \phi_j\left(\int u(g(s))dP_j(s)\right).$$

Likewise, $f_i \succsim_0 g_i$ if and only if

$$\sum_{j \neq i}^{N} \phi_j\left(\int u(f_i(s))dP_j(s)\right) \geq \sum_{j \neq i}^{N} \phi_j\left(\int u(g_i(s))dP_j(s)\right).$$

However, because by assumption for every $j \neq i$ it holds $f_i \sim_j f$, we obtain

$$f \succcurlyeq_0 g \iff f_i \succcurlyeq_0 g_i,$$

which proves Agent Independence. The Pareto condition simply follows by the fact that each ϕ_i is strictly increasing. Finally, if $\succcurlyeq_0$ admits such a representation then it is clearly a continuous weak order. □

6.2. Proof of Theorem 2

The next lemma is key for the proof of the theorem.

Lemma 9. *A Bayesian planner $\succcurlyeq_B$ satisfies the Pareto condition if and only if there exists a set of positive weights $(\mu_i)_{i=1}^N$ with $\sum_{i=1}^N \mu_i = 1$ such that $\succcurlyeq_B$ is represented by*

$$V(f) = \int u(f(s))d\bar{P},$$

where $\bar{P} = \sum_{i=1}^N \mu_i P_i$.

Proof. Let P_0 be the planner's belief. Define the mapping $F : \mathcal{F} \to \mathbb{R}^{n+1}$ by $f \mapsto (\int u(f)dP_0, \int u(f)dP_1, \ldots, \int u(f)dP_N)$. Observe that $F(\mathcal{F})$ is convex. To see this, as in Lemma 3 one can show that

$$F(\mathcal{F}) = \left\{\left(\int gdP_0, \int gdP_1, \ldots, \int gdP_N\right) : g \in \mathbb{R}^S\right\}.$$

Because the mapping $g \mapsto \left(\int gdP_0, \int gdP_1, \ldots, \int gdP_N\right)$ is linear, it follows that $F(\mathcal{F})$ is convex as desired. Therefore, by the Pareto condition we can apply Proposition 2 in De Meyer and Mongin (1995) so that there exist weights $(a_i)_{i=0}^N$ with $a_i > 0$ for $i = 1, \ldots, N$ such that

$$\int u(f)dP_0 = \sum_{i=1}^N \int u(f)dP_i a_i + a_0.$$

The result then follows by appropriately rescaling u. The converse is straightforward. □

Given this result, the proof now follows closely the proof of Theorem 2 in Grant et al. (2009).

Proof of Theorem 2. Suppose that $\succcurlyeq_0$ is represented by

$$f \mapsto \sum_{i=1}^N \phi\left(\int u(f(s))dP_i(s)\right)\mu_i.$$

Take $f, g \in \mathcal{F}$ with g unambiguous such that $g \sim_B f$ where $\succcurlyeq_B$ is represented by

$$V(f) = \int u(f(s))d\bar{P},$$

where $\bar{P} = \sum_{i=1}^{N} \mu_i P_i$. This means that

$$\int u(g) d\bar{P} = \int u(f) d\bar{P}.$$

Observe that $g \sim_0 u^{-1}(\int u(g) d\bar{P}) \sim_0 u^{-1}(\int u(f) d\bar{P})$. By applying Jensen's inequality it follows that

$$\phi\Big(\sum_{i=1}^{n} \int u(g(s)) dP_i(s) \mu_i\Big) = \phi\Big(\int u(f) d\bar{P}\Big) \geq \sum_{i=1}^{N} \phi\Big(\int u(f(s)) dP_i(s)\Big)\mu_i,$$

so that $g \succcurlyeq_0 f$ as desired.

Conversely, note that by Theorem 1 $\succcurlyeq_0$ is represented by $V_0(f) = \varphi(T(f))$ such that φ is continuous and completely separable. Moreover, by the ambiguity aversion axiom there exists a set of positive weights $(\mu_i)_{i=1}^{N}$ summing to 1 such that for every $\mathbf{x} = (x_i)_{i=1}^{N} \in \mathbb{R}^N$ it holds that $\varphi((\sum_{i=1}^{N} x_i \mu_i, \ldots, \sum_{i=1}^{N} x_i \mu_i)) \geq \varphi(\mathbf{x})$. Indeed, let $\mathbf{x} = T(f)$. By Lemma 9, $u^{-1}(\int u(f) d\bar{P}) \sim_B f$. Therefore, ambiguity aversion implies that $u^{-1}(\int u(f) d\bar{P}) \succcurlyeq_0 f$ which means that

$$\varphi\Big(\Big(\sum_{i=1}^{N} x_i \mu_i, \ldots, \sum_{i=1}^{N} x_i \mu_i\Big)\Big) \geq \varphi(\mathbf{x}),$$

as desired.

Thus, the result follows by applying Theorem 1 in Werner (2005). It is routine to check that the representation is unique. □

6.3. *Proof of Proposition 1*

I prove that $\succcurlyeq_0$ is more ambiguity averse than $\hat{\succcurlyeq}_0$ if and only if there exists a strictly increasing, concave function $h : \mathbb{R} \to \mathbb{R}$ such that $\phi(x) = h(\hat{\phi}(x))$ for every $x \in \mathbb{R}$.

Let $h = \phi \circ \hat{\phi}^{-1}(x)$ for every $x \in \mathbb{R}$. It is obvious that h is strictly increasing. I claim h is concave. To see this, note that for any $f \in \mathcal{F}$, there exists $x_f \in \mathbb{R}$ such that $x_f \hat{\sim}_0 f$ (Lemma 1). Hence $x_f \succcurlyeq_0 f$, which implies

$$\hat{\phi}^{-1}\Big(\sum_{i=1}^{N} \hat{\phi}\Big(\int u(f) dP_i\Big)\mu_i\Big) \geq \phi^{-1}\Big(\sum_{i=1}^{N} \phi\Big(\int u(f) dP_i\Big)\mu_i\Big),$$

so that

$$h\Big(\sum_{i=1}^{N} \mu_i \hat{\phi}\Big(\int u(f) dP_i\Big)\Big) \geq \sum_{i=1}^{N} \phi\Big(\int u(f) dP_i\Big) = \sum_{i=1}^{N} \mu_i h \circ \hat{\phi}\Big(\int u(f) dP_i\Big).$$

By Lemma 6 in Klibanoff et al. (2005), it follows that h is concave as wanted.

The other direction of the proof follows from Jensen's inequality.

6.4. *Proof of Theorem 3*

The equivalent of Theorem 1 can be proved in the same way, except Lemma 4 follows immediately by using the fact that $u_i(X) = \mathbb{R}$ for every $i = 1, \ldots, N$. The proof then works as the proof of Theorem 2.

Suppose that $\succcurlyeq_0$ is represented by

$$f \mapsto \sum_{i=1}^{N} \phi\Big(\int u_i(f(s))dP_i(s)\Big)\mu_i .$$

Take $f, g \in \mathcal{F}$ with g socially unambiguous such that $g \sim_B f$. This means that

$$\int \bar{u}(g)d\bar{P} \equiv \sum_{i=1}^{n} \int u_i(g(s))dP_i(s)\mu_i = \sum_{i=1}^{n} \int u_i(f(s))dP_i(s)\mu_i .$$

By applying Jensen's inequality it follows that

$$\begin{aligned} V(g) = \phi\Big(\int \bar{u}(g)d\bar{P}\Big) &= \phi\Big(\sum_{i=1}^{n} \int u_i(g(s))dP_i(s)\mu_i\Big) \\ &= \phi\Big(\sum_{i=1}^{n} \int u_i(f(s))dP_i(s)\mu_i\Big) \\ &\geq \sum_{i=1}^{N} \phi\Big(\int u_i(f(s))dP_i(s)\Big)\mu_i = V(f), \end{aligned}$$

so that $g \succcurlyeq_0 f$ as desired.

Conversely, note that by the equivalent version of Theorem 1, $\succcurlyeq_0$ is represented by $V_0(f) = \varphi(T(f))$ such that φ is continuous and completely separable. Moreover by the social ambiguity aversion axiom, for every $\mathbf{x} = (x_i)_{i=1}^N \in \mathbb{R}^N$ it holds that $\varphi((\sum_{i=1}^N x_i\mu_i, \ldots, \sum_{i=1}^N x_i\mu_i)) \geq \varphi(\mathbf{x})$. Indeed, let f satisfy $\mathbf{x}=T(f)$. Find y_i, $i=1, \ldots, N$ such that $u_i(y_i)=\sum_{i=1}^N \mu_i(\int u_i(f)dP_i)$. Observe that $(y_1, \ldots, y_N) \sim_B f$. Therefore, the social ambiguity aversion implies that $(y_1, \ldots, y_N) \succcurlyeq_0 f$ which means that

$$\varphi\Big(\Big(\sum_{i=1}^{N} x_i\mu_i, \ldots, \sum_{i=1}^{N} x_i\mu_i\Big)\Big) \geq \varphi(\mathbf{x}),$$

as desired.

Thus the result follows by applying Theorem 1 in Werner (2005). It is routine to check that the representation is unique.

Appendix. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jet.2021.105308.

References

Adler, M.D., Sanchirico, C.W., 2006. Inequality and uncertainty: theory and legal applications. Univ. Pa. Law Rev. 155, 279–377.

Alon, S., Gayer, G., 2016. Utilitarian preferences with multiple priors. Econometrica 84, 1181–1201.

Alon, S., Gayer, G., 2018. Maxmin expected utility as resulting from information on others' preferences. Working Paper.

Billot, A., Vergopoulos, V., 2016. Aggregation of Paretian preferences for independent individual uncertainties. Soc. Choice Welf. 47, 973–984.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., Montrucchio, L., 2013. Classical subjective expected utility. Proc. Natl. Acad. Sci. 110, 6754–6759.

Chateauneuf, A., Wakker, P., 1993. From local to global additive representation. J. Math. Econ. 22 (6), 523–545.
Crès, H., Gilboa, I., Vieille, N., 2011. Aggregation of multiple prior opinions. J. Econ. Theory 146 (6), 2563–2582.
Danan, E., Gajdos, T., Hill, B., Tallon, J.-M., 2016. Robust social decisions. Am. Econ. Rev. 106 (9), 2407–2425.
De Finetti, B., 1967. L'incertezza nell'economia, Part One of B. de Finetti and F. Emanuelli, Economia delle Assicurazioni, Unione Tipografico-Editrice Torinese, Torino.
De Meyer, B., Mongin, P., 1995. A note on affine aggregation. Econ. Lett. 47 (2), 177–183.
Debreu, G., 1959. Topological methods in cardinal utility theory. Technical report. Cowles Foundation for Research in Economics, Yale University.
Epstein, L.G., Segal, U., 1992. Quadratic social welfare functions. J. Polit. Econ. 100 (4), 691–712.
Feduzi, A., Runde, J., Zappia, C., 2013. De Finetti on uncertainty. Camb. J. Econ. 38 (1), 1–21.
Fleming, J.M., 1957. Cardinal welfare and individualistic ethics: a comment. J. Polit. Econ. 65 (4), 355–357.
Fleming, M., 1952. A cardinal concept of welfare. Q. J. Econ. 66 (3), 366–384.
Fleurbaey, M., 2010. Assessing risky social situations. J. Polit. Econ. 118 (4), 649–680.
Gajdos, T., Tallon, J.-M., Vergnaud, J.-C., 2008. Representation and aggregation of preferences under uncertainty. J. Econ. Theory 141 (1), 68–99.
Gajdos, T., Vergnaud, J.-C., 2013. Decisions with conflicting and imprecise information. Soc. Choice Welf. 41 (2), 427–452.
Ghirardato, P., Marinacci, M., 2002. Ambiguity made precise: a comparative foundation. J. Econ. Theory 102 (2), 251–289.
Gilboa, I., Samet, D., Schmeidler, D., 2004. Utilitarian aggregation of beliefs and tastes. J. Polit. Econ. 112 (4), 932–938.
Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with non-unique prior. J. Math. Econ. 18 (2), 141–153.
Gorman, W.M., 1968. The structure of utility functions. Rev. Econ. Stud. 35 (4), 367–390.
Grant, S., Kajii, A., Polak, B., Safra, Z., 2010. Generalized utilitarianism and Harsanyi's impartial observer theorem. Econometrica 78 (6), 1939–1971.
Grant, S., Polak, B., Strzalecki, T., 2009. Second-Order Expected Utility, Available at SSRN.
Harsanyi, J.C., 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. J. Polit. Econ. 63 (4), 309–321.
Hayashi, T., Lombardi, M., 2018. Fair social decision under uncertainty and belief disagreements. Econ. Theory, 1–42.
Klibanoff, P., 2013. Thoughts on policy choice under ambiguity. J. Inst. Theor. Econ. 169, 134–138.
Klibanoff, P., Marinacci, M., Mukerji, S., 2005. A smooth model of decision making under ambiguity. Econometrica 73 (6), 1849–1892.
Manski, C.F., 2009. Diversified treatment under ambiguity. Int. Econ. Rev. 50 (4), 1013–1041.
Manski, C.F., 2011. Choosing treatment policies under ambiguity. Annu. Rev. Econ. 3 (1), 25–49.
Manski, C.F., 2013. Status quo deference and policy choice under ambiguity. J. Inst. Theor. Econ. 169 (1), 116–128.
Manski, C.F., Tetenov, A., 2007. Admissible treatment rules for a risk-averse planner with experimental data on an innovation. J. Stat. Plan. Inference 137 (6), 1998–2010.
Marinacci, M., 2015. Model uncertainty. J. Eur. Econ. Assoc. 13 (6), 1022–1100.
Mongin, P., 1995. Consistent Bayesian aggregation. J. Econ. Theory 66 (2), 313–351.
Mongin, P., 1997. Spurious unanimity and the Pareto principle. In: Conférence de l'International Society for Utilitarian Studies, Nouvelle Orléans.
Mongin, P., d'Aspremont, C., 1998. Utility theory and ethics. In: Handbook of Utility Theory, vol. 1, pp. 371–481.
Mongin, P., Pivato, M., 2015. Ranking multidimensional alternatives and uncertain prospects. J. Econ. Theory 157, 146–171.
Nascimento, L., 2012. The ex ante aggregation of opinions under uncertainty. Theor. Econ. 7 (3), 535–570.
Rudin, W., 1973. Functional Analysis. Mcgraw-Hill Series in Higher Mathematics.
Savage, L.J., 1951. The theory of statistical decision. J. Am. Stat. Assoc. 46 (253), 55–67.
Savage, L.J., 1972. The Foundations of Statistics. Courier Corporation.
Segal, U., 1992. Additively separable representations on non-convex sets. J. Econ. Theory 56 (1), 89–99.
Seo, K., 2009. Ambiguity and second-order belief. Econometrica 77 (5), 1575–1605.
Sprumont, Y., 2018. Belief-weighted Nash aggregation of Savage preferences. J. Econ. Theory 178, 222–245.
Teicher, H., 1963. Identifiability of finite mixtures. Ann. Math. Stat. 34, 1265–1269.
Wakker, P.P., 1989. Additive Representations of Preferences: A New Foundation of Decision Analysis, vol. 4. Springer Science & Business Media.
Wakker, P.P., Zank, H., 1999. A unified derivation of classical subjective expected utility models through cardinal utility. J. Math. Econ. 32 (1), 1–19.
Werner, J., 2005. A simple axiomatization of risk-averse expected utility. Econ. Lett. 88 (1), 73–77.