Modeling agent's conditional preferences under objective ambiguity in Dempster-Shafer theory

Davide Petturiti[a], Barbara Vantaggi[b,*]

[a] *Dip. Economia, Università degli Studi di Perugia, 06100 Perugia, Italy*
[b] *Dip. MEMOTEF, "La Sapienza" Università di Roma, 00185 Roma, Italy*

Abstract

We manage decisions under "objective" ambiguity by considering generalized Anscombe-Aumann acts, mapping states of the world to generalized lotteries on a set of consequences. A generalized lottery is modeled through a belief function on consequences, interpreted as a partially specified randomizing device. Preference relations on these acts are given by a decision maker focusing on different scenarios (conditioning events). We provide a system of axioms which are necessary and sufficient for the representability of these "conditional preferences" through a conditional functional parametrized by a unique full conditional probability P on the algebra of events and a cardinal utility function u on consequences. The model is able to manage also "unexpected" (i.e., "null") conditioning events and distinguishes between a systematically pessimistic or optimistic behavior, either referring to "objective" belief functions or their dual plausibility functions. Finally, an elicitation procedure is provided, reducing to a Quadratically Constrained Linear Program (QCLP).

Keywords: Anscombe-Aumann acts, Conditional Choquet expected value, Belief and plausibility functions, Ambiguity, Conditional preferences

1. Introduction

In many decision problems under uncertainty in economics, we need to choose between uncertain consequences in a set X that are contingent on the states of the world in a set S. We distinguish between an "objective"

*Corresponding author.
Email addresses: davide.petturiti@unipg.it (Davide Petturiti), barbara.vantaggi@uniroma1.it (Barbara Vantaggi)

uncertainty related to X (i.e., exogenously quantified and given to the decision maker, in the spirit of [58]) and a "subjective" uncertainty related to S (i.e., encoded in the decision maker's preferences, in the spirit of [52]). This configures a two-stage process where first the state of the world is chosen by Nature, and then the consequence is chosen through "objective" uncertainty, in the spirit of [2].

Very often, due to partial knowledge, uncertainty cannot be encoded in a single probability measure, but we rather have a class of probability measures. This problem goes back to de Finetti's coherence [19]: a partially specified (conditional) probability, provided it is coherent, can be extended to any larger domain by preserving coherence. The extension of a coherent (conditional) probability is not unique in general but the set of all the possible extensions can be found through linear programming [3, 4, 8, 13, 14, 15, 18]. The class of (conditional) probabilities extending the initial assessment can be summarized by means of its lower and upper envelopes. Generally, the lower envelope of a class of probability measures is only a superadditive capacity. Nevertheless, under suitable conditions the lower envelope can be a belief function or even a necessity measure [10, 11, 16, 27, 44]. However, given a lower probability it is always possible to look for a belief function outer approximating it [45]. In this paper we restrict to the case where the lower envelope is a belief function obtained either directly as lower envelope of probability measures or as an outer approximation of it. Actually, the approach could be generalized to 2-monotone capacities, but we restrict to belief functions since we start from two motivating examples where the lower envelopes turn out to be belief functions and furthermore this restriction leads to a simpler treatment and interpretation.

We refer to situations where ambiguity is related to the "objective" probabilistic assessment, as that due to a partially known randomizing device (like an urn or a roulette wheel), that results in a class of probability measures whose lower envelope is a belief function [22, 55], like in the well-known Ellsberg's urn paradox [28].

Following [1, 48, 57], in these cases we will speak of "objective" ambiguity, which is expressed in the Dempster-Shafer theory as in [9, 25, 37]. The objects of decisions can be modeled as generalized Anscombe-Aumann acts [2], mapping S to the set $\mathbf{B}(X)$ of belief functions over X, forming the set $\mathcal{F} = \mathbf{B}(X)^S$. Then, an act associates to each state of the world a generalized lottery over the set of consequences and the generalized lottery is described by either a belief function (as in [37, 38]) or its dual plausibility function.

A crucial aspect of making decisions under uncertainty is the possibility of reasoning under hypotheses. For this reason, we consider a conditional decision model involving the above generalization of Anscombe-Aumann acts, assuming that the decision maker is able to provide a family of preference relations $\{\precsim_H\}_{H \in \wp(S)^0}$ on $\mathcal{F}$ indexed by the set $\wp(S)^0 = \wp(S) \setminus \{\emptyset\}$ of non-impossible events. Every preference relation $\precsim_H$ can be interpreted as comparing acts under the hypothesis H.

On the other hand, "subjective" uncertainty is assumed to be probabilistic, so, we model it with a full conditional probability in the sense of [19, 20, 26, 50], that allows conditioning on "null" events (i.e., events with zero probability), but possible. Notice that, as shown in [42], "null" events play a crucial role in the analysis of a game (see also [46]), thus full conditional probabilities are the most suitable "subjective" uncertainty measures for a "probabilistic" agent.

Here, the numerical model of reference is the following conditional functional $\mathbf{CEU}_{P,u}$ defined, for every $f \in \mathcal{F}$ and every $H \in \wp(S)^0$, as

$$\mathbf{CEU}_{P,u}(f|H) = \sum_{s \in S} P(\{s\}|H) \left(\oint u \mathrm{d} f(s) \right), \tag{1}$$

where $P(\cdot|\cdot)$ is a full conditional probability on $\wp(S) \times \wp(S)^0$ and $u : X \to \mathbb{R}$ is a utility function. The above conditional functional consists in a mixture with respect to a full conditional probability of Choquet expected utilities [9] contingent on the states of the world. In particular, due to the properties of the Choquet integral [53], every state-contingent Choquet expected utility is actually a lower expected utility with respect to the probabilities in $\mathbf{core}(f(s))$. The present model generalizes the conditional version of the Anscombe-Aumann model given in [46] by introducing "objective" ambiguity.

We provide a set of axioms for the family $\{\precsim_H\}_{H \in \wp(S)^0}$ that is proved to be necessary and sufficient for the existence of a unique full conditional probability $P(\cdot|\cdot)$ and a cardinal utility function u such that the corresponding $\mathbf{CEU}_{P,u}$ conditional functional *represents* the preferences, i.e., for every $f, g \in \mathcal{F}$ and every $H \in \wp(S)^0$,

$$f \precsim_H g \iff \mathbf{CEU}_{P,u}(f|H) \leq \mathbf{CEU}_{P,u}(g|H).$$

It turns out that a rational agent in this model behaves as a $\mathbf{CEU}_{P,u}$ maximizer, so, as a maximizer of a "subjective" conditional expected value of state-contingent "objective" lower expected utilities.

Hence, the present model encodes the behavior of a decision maker that is systematically pessimistic in resolving his/her uncertainty on X.

A similar model can be stated in a way to cover a systematically optimistic behavior, by considering a conditional functional $\widehat{\mathbf{CEU}}_{P,u}$ relying on a mixture, with respect to a full conditional probability, of state-contingent upper expected utilities. The state-contingent upper expectations are computed with respect to plausibility functions, that are dual to the belief functions in the range of an act. In fact, for each act and any state of the world we get a generalized lottery expressed by a plausibility function. Then, a rational agent in this case behaves as a $\widehat{\mathbf{CEU}}_{P,u}$ maximizer, so, as a maximizer of a "subjective" conditional expected value of state-contingent "objective" upper expected utilities.

A similar decision setting, limited to the unconditional case, has been considered by [57], where the author takes acts mapping states of the world to non-empty compact convex polyhedral sets of probability measures over consequences. In the same paper the author considers a representation functional different from ours, but still relying on a mixture with respect to a "subjective" probability measure.

Important efforts have been addressed in the decision theory literature to model "subjective" ambiguity, that is to ambiguity in "subjective" uncertainty evaluations (see, e.g., the survey papers [29] and [34]). For instance, in the seminal papers [54] and [35], the classical Anscombe-Aumann setting is considered, but there ambiguity is "subjective", since the mixture of state-contingent expected utilities is done through the Choquet integral with respect to a capacity over S in the first model, while a class of "subjective" probabilities is considered in the second model. Still working in the classical Anscombe-Aumann setting, we find the models [5, 6, 43]. Other lines of research take care of "subjective" ambiguity in a Savage's setting, through acts that map states of the world to non-empty sets of consequences [32, 47]. All the quoted decision models essentially focus on unconditional decisions.

Concerning the use of belief functions in decision making, we refer to [24], which provides the most up-to-date survey on the topic. The framework adopted in [24] relies on acts mapping states of the world to consequences: a belief function on the states of the world (possibly induced by a probability space and a multi-valued mapping) determines, through such an act, a generalized lottery on consequences, in the sense of our paper. Hence, the notion of act adopted in the quoted paper differs from that of generalized Anscombe-Aumann act in our sense. Nevertheless, we stress that if S is a singleton (and so mixing and conditioning are trivial), then the maximization of $\mathbf{CEU}_{P,u}$ and $\widehat{\mathbf{CEU}}_{P,u}$ reduces, respectively, to the *(generalized)* *maximin* criterion and the *(generalized) maximax* criterion of [24].

The conditional functional $\mathbf{CEU}_{P,u}$, as well as $\widehat{\mathbf{CEU}}_{P,u}$, is completely specified once the full conditional probability $P(\cdot|\cdot)$ and the utility function u have been elicited by the decision maker. In general, an agent is only able to provide few comparisons for few conditioning events. In this case, the first issue is to check the consistency of the given comparisons with the model of reference. When consistency holds, it is easily seen that an elicitation procedure relying on a finite number of arbitrary comparisons cannot guarantee any form of uniqueness for P and u in general.

We provide an elicitation procedure that reduces to a Quadratically Constrained Linear Program (QCLP). Unfortunately, the quadratic forms in the quadratic constraints of the problem are generally not positive semidefinite nor negative semidefinite, so, the problem is generally not convex: interior points algorithms are not suitable. The problem can be solved with a branch and bound algorithm coping with global optimization of non-linear problems, such as the COUENNE solver [17].

2. Motivating examples

We introduce two toy examples that help to motivate the study carried out in the rest of the paper.

2.1. An investment decision problem

Let $S = \{s_1, s_2, s_3, s_4\}$ be the set of states of the world, with

- s_1 = "North Korea and USA enter into war next year and Italian GDP increases next year";
- s_2 = "North Korea and USA enter into war next year and Italian GDP does not increase next year";
- s_3 = "North Korea and USA do not enter into war next year and Italian GDP increases next year";
- s_4 = "North Korea and USA do not enter into war next year and Italian GDP does not increase next year".

Then, $K = \{s_1, s_2\}$ = "North Korea and USA enter into war next year" and $K^c = \{s_3, s_4\}$ = "North Korea and USA do not enter into war next year".

Consider three unitary financial instruments that can result in a loss of €50, in a null gain or in a gain of €100, implying $X = \{-50, 0, 100\}$. From statistics of previous years we only have partial information on the performances of each instrument, that are listed below:

Instrument 1: It is only known that it guarantees a gain of €100 in 30% of cases;

Instrument 2: It is only known that it results in a loss of €50 in 20% of cases;

Instrument 3: No information is available.

The usual goal would be to identify each instrument with a probability distribution on X. Nevertheless, due to the lack of complete probabilistic information, the decision maker can adopt different behaviors to account for this imprecision. In particular, for each instrument, he/she can consider all the probability distributions on X compatible with the given partial assessment. This approach goes back to de Finetti's coherence [21] (see also [15]), that allows to extend any coherent (partial) assessment to a probability measure. In this light, we consider all the probability distributions on X compatible with the above partial assessments.

Hence, instrument i determines a class of probability measures $\mathbf{P}^i$ on $\wp(X)$

$$
\begin{aligned}
\mathbf{P}^1 &= \{P : \wp(X) \to [0,1] | P \text{ is a probability measure}, \gamma \in [0, 0.7], \\
&\qquad P(\{-50\}) = \gamma, P(\{0\}) = 0.7 - \gamma, P(\{100\}) = 0.3\}, \\
\mathbf{P}^2 &= \{P : \wp(X) \to [0,1] | P \text{ is a probability measure}, \gamma \in [0, 0.8] \\
&\qquad P(\{-50\}) = 0.2, P(\{0\}) = \gamma, P(\{100\}) = 0.8 - \gamma\}, \\
\mathbf{P}^3 &= \{P : \wp(X) \to [0,1] | P \text{ is a probability measure}\}.
\end{aligned}
$$

For every $i = 1, 2, 3$, if we consider the lower envelope on the elements of $\wp(X)$ defined as $Bel_i = \min \mathbf{P}^i$, we get a non-additive uncertainty measure which reveals to be a belief function (whose definition and properties are recalled in Section 3).

The use of the lower envelopes of the above sets of probabilities points out a systematically pessimistic behavior of the decision maker in dealing with "objective" uncertainty on X. On the other hand, a systematically optimistic behavior is obtained by referring to the corresponding upper envelopes that, in this example, are plausibility functions (also introduced in Section 3).

Now, consider the following investment strategies in which the adopted financial instrument is contingent on the state of the world:

	s_1	s_2	s_3	s_4
f	**Inst. 3**	**Inst. 1**	**Inst. 1**	**Inst. 2**
g	**Inst. 3**	**Inst. 3**	**Inst. 2**	**Inst. 3**

Suppose the decision maker is interested in the effect of a possible war between North Korea and USA, that is, he/she wants to considers the scenarios K and K^c. The question is: How should the decision maker decide between the two investment strategies conditionally on K and K^c? A natural answer is "according to his/her preferences". Hence, the aim is to provide a set of axioms ruling the preferences of a "rational" decision maker.

2.2. A multi-criteria decision problem with uncertain profiles

A simple situation of multi-criteria decision is shown. Let $S = \{s_1, \ldots, s_n\}$ be a finite set of *criteria* and denote by $A_i = \{a^i_1, \ldots, a^i_{n_i}\}$ the finite set of possible values for the criterion $s_i \in S$. Then, let $X = A_1 \times \cdots \times A_n = \{x_1, \ldots, x_m\}$ be the finite set of *profiles*, where each element of X is an n-tuple of evaluations in each criterion. Under certainty, each decision alternative is identified with an element of X but, due to a lack of knowledge, here we assume that the evaluation profile is uncertain and is identified with a function mapping each criterion to a belief function over X. The following example illustrates the point.

Suppose an agent needs to choose a type of restaurant in his/her town, according to the following criteria in the set $S = \{s_1, s_2, s_3\}$, with the corresponding sets of possible values A_i's:

- $s_1 =$ quality of food;
- $A_1 = \{\text{poor}, \text{medium}, \text{high}\}$;
- $s_2 =$ price;
- $A_2 = \{\text{high}, \text{medium}, \text{low}\}$;
- $s_3 =$ time to be served;
- $A_3 = \{ > 20 \text{ mins}, \text{10-20 mins}, < 10 \text{ mins}\}$.

The space of all certain profiles is then $X = A_1 \times A_2 \times A_3$.

At the moment of the choice, the agent has restricted the set of alternatives to the following types of alternatives: a Chinese restaurant or an Italian one.

Gathering all the available information resulting from evaluations of other people, possibly coming from different sources on the Internet, the agent arrives to the following partial probabilistic assessment.

Chinese Restaurant: It is only known that the quality of food is poor in 20% of cases. Furthermore, the price is high in 30% of cases, is medium in 30% of cases, and is low in 40% of cases. The time to be served is more than 20 minutes in 50% of cases and is less than 10 minutes in the other 50% of cases.

Italian Restaurant: The quality of food is low in 40% of cases, is medium in 30% of cases and is high in 30% of cases. It is only known that the price is low in 30% of cases. No information is available on the time to be served.

Hence, denoting f = Chinese and g = Italian, for each type of restaurant the following probabilistic information is available, where each 3-tuple is a (possibly partially specified) probability distribution on A_i:

	s_1	s_2	s_3
f	$(0.2, ?, ?)$	$(0.3, 0.3, 0.4)$	$(0.5, 0, 0.5)$
g	$(0.4, 0.3, 0.3)$	$(?, ?, 0.3)$	$(?, ?, ?)$

Looking at alternative f, for every criterion $s_i \in S$ there is a class of probabilities $\mathbf{P}_i^f$ on $\wp(X)$ that agrees with the partial probabilistic information available on A_i. Also in this case the decision maker could consider the lower envelope $Bel_i^f = \min \mathbf{P}_i^f$ that, again, turns out to be a belief function on $\wp(X)$. The same holds for alternative g.

Hence, the two alternatives can be identified with two functions mapping criteria to belief functions over $\wp(X)$ as follows

	s_1	s_2	s_3
f	Bel_1^f	Bel_2^f	Bel_3^f
g	Bel_1^g	Bel_2^g	Bel_3^g

Suppose the agent wants to focus either on the quality of food and the price or on the time to be served, that is on $H = \{s_1, s_2\}$ or on $H^c = \{s_3\}$. The question is: How should the agent decide between the two types of restaurant focusing either on H or on H^c? Also in this case, a natural answer is "according to his/her preferences", so, the problem translates again in finding normative axioms for his/her preferences, according to which the agent is dubbed as "rational".

3. Preliminaries

In this section we recall the necessary material on belief functions and full conditional probabilities, the former used to express "objective" uncertainty on consequences, and the latter used to encode "subjective" conditional uncertainty on the states of the world. We provide a unified introduction by working on an abstract finite set.

Let $\Omega = \{\omega_1, \ldots, \omega_n\}$ be an arbitrary finite set. Denote by $\wp(\Omega)$ the power set of Ω and by $\wp(\Omega)^0 = \wp(\Omega) \setminus \{\emptyset\}$.

A *belief function* on $\wp(\Omega)$ [22, 55] is a function $Bel : \wp(\Omega) \to [0,1]$ satisfying $Bel(\emptyset) = 0$, $Bel(\Omega) = 1$ and the k-monotonicity property for every $k \geq 2$, that is, for every $A_1, \ldots, A_k \in \wp(\Omega)$,

$$Bel\left(\bigcup_{i=1}^{k} A_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} Bel\left(\bigcap_{i \in I} A_i\right).$$

The *dual* function Pl defined, for every $A \in \wp(\Omega)$, as $Pl(A) = 1 - Bel(A^c)$, is called *plausibility function* and satisfies the above inequality in the opposite direction, switching intersections and unions. Both Bel and Pl are *(normalized) capacities*, i.e., they are monotonic with respect to set inclusion. Recall that probability measures are particular belief and plausibility functions.

Both functions Bel and Pl on $\wp(\Omega)$ are completely characterized by the *Möbius inversion* of Bel [7, 36], defined for every $A \in \wp(\Omega)$ as

$$m_{Bel}(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} Bel(B),$$

also called *basic probability assignment*. The function $m_{Bel} : \wp(\Omega) \to [0,1]$ is such that $m_{Bel}(\emptyset) = 0$, $\sum_{A \in \wp(\Omega)} m_{Bel}(A) = 1$, and, for every $A \in \wp(\Omega)$,

$$Bel(A) = \sum_{B \subseteq A} m_{Bel}(B) \quad \text{and} \quad Pl(A) = \sum_{B \cap A \neq \emptyset} m_{Bel}(B). \tag{2}$$

Notice that, disregarding $m_{Bel}(\emptyset) = 0$, m_{Bel} can be formally viewed as a probability distribution over the set $\wp(\Omega)^0$, in particular, Bel (and, so, the dual Pl) is a probability measure if and only if m_{Bel} is different from 0 only on singletons.

For a function $f \in \mathbb{R}^{\Omega}$, if φ is a capacity on $\wp(\Omega)$ and σ is a permutation of $\{1, \ldots, n\}$ such that $f(\omega_{\sigma(1)}) \leq \ldots \leq f(\omega_{\sigma(n)})$ (see [23]), the *Choquet integral* of f with respect to φ is defined, denoting $E_i^{\sigma} = \{\omega_{\sigma(i)}, \ldots, \omega_{\sigma(n)}\}$ for $i = 1, \ldots, n$ and $E_{n+1}^{\sigma} = \emptyset$, as

$$\oint f \mathrm{d}\varphi = \sum_{i=1}^{n} f(\omega_{\sigma(i)})(\varphi(E_i^{\sigma}) - \varphi(E_{i+1}^{\sigma})). \tag{3}$$

In particular, if φ is a belief/plausibility function, the Choquet integral with respect to φ is the minimum/maximum of Stieltjes integrals of f with respect to the *core* of Bel (see, e.g., [23, 53]), that is the set of probability measures

$$\mathbf{core}(Bel) = \{P : P \text{ is a probability measure on } \wp(\Omega),\ Bel \leq P\},$$

for which it holds $Bel = \min \mathbf{core}(Bel)$ and $Pl = \max \mathbf{core}(Bel)$, where the minimum/maximum is pointwise on the elements of $\wp(\Omega)$.

In what follows, given a belief function Bel on $\wp(\Omega)$, to refer to the dual plausibility function we often write $\widehat{Bel}$.

In the rest of the paper we refer to the following axiomatic definition of conditional probability, essentially due to [19, 20] (see also [15]).

Definition 1. *Let $\mathcal{A}$ be an algebra of subsets of a non-empty set Ω and $\mathcal{H}$ an additive class such that $\mathcal{H} \subseteq \mathcal{A}^0$ (i.e., a class of sets closed under finite unions). A function $P : \mathcal{A} \times \mathcal{H} \to [0, 1]$ is a* **conditional probability** *if it satisfies the following conditions:*

(i) $P(E|H) = P(E \cap H|H)$, for every $E \in \mathcal{A}$ and $H \in \mathcal{H}$;

(ii) $P(\cdot|H)$ is a finitely additive probability on $\mathcal{A}$, for every $H \in \mathcal{H}$;

(iii) $P(E \cap F|H) = P(E|H) \cdot P(F|E \cap H)$, for every $H, E \cap H \in \mathcal{H}$ and $E, F \in \mathcal{A}$.

Remark 1. *Notice that, in presence of condition (i), condition (iii) is equivalent to (iii')*

$$P(A|C) = P(A|B) \cdot P(B|C),$$

for every $A \in \mathcal{A}$ and $B, C \in \mathcal{H}$, with $A \subseteq B \subseteq C$.

If $\Omega \in \mathcal{H}$, we simply write $P(E) = P(E|\Omega)$, for every $E \in \mathcal{A}$.

Following [26], we say that a conditional probability $P(\cdot|\cdot)$ is *full on $\mathcal{A}$* if $\mathcal{H} = \mathcal{A}^0$, i.e., if it is defined on $\mathcal{A} \times \mathcal{A}^0$. In particular, in [15] it is shown that every conditional probability on $\mathcal{A} \times \mathcal{H}$ can be extended to a full conditional probability on the whole $\mathcal{A} \times \mathcal{A}^0$. In what follows we will be essentially concerned with a full conditional probability defined on $\wp(\Omega) \times \wp(\Omega)^0$, for a finite set Ω. Notice that a full conditional probability allows conditioning also on "null" events, i.e., $P(\cdot|H)$ is well-defined even when $P(H) = 0$.

In this paper, full conditional probabilities are interpreted as "subjective" weights used to aggregate the "objective" information on the different states of the world, conditionally on a possibly "unexpected" (i.e., "null") event.

4. Decision-theoretic setting

The rest of the paper relies on the following decision-theoretic setting. Let $X = \{x_1, \ldots, x_m\}$ be a finite set of consequences and $\wp(X)^0 = \wp(X) \setminus \{\emptyset\}$ be the set of *opportunity sets* or *menus* [40] (i.e., non-empty sets of consequences). Denote by

$$\mathbf{B}(X) = \{Bel \,:\, Bel \text{ is a belief function on } \wp(X)\}$$

the set of all belief functions on $\wp(X)$ interpreted as generalized lotteries (according to [37]).

Let $S = \{s_1, \ldots, s_n\}$ be a finite set of states of the world endowed with its power set $\wp(S)$, forming the set of all events, and $\wp(S)^0 = \wp(S) \setminus \{\emptyset\}$, forming the set of *scenarios* (i.e., conditioning events).

An *act* $f : S \to \mathbf{B}(X)$ is a map from the states of the world to the set of belief functions on consequences, the latter expressing "objective" ambiguity. We then denote by

$$\mathcal{F} = \mathbf{B}(X)^S$$

the set of all acts which contains, in particular, the classical Anscombe-Aumann acts [2, 30, 41].

Acts in $\mathcal{F}$ can be seen as state-contingent partially known randomizing devices, thus they allow to distinguish "objective" uncertainty on consequences (which is already quantified, even though possibly partially, and provided to the agent) from "subjective" uncertainty on the states of the world (which is encoded in the agent's preferences).

A decision maker expresses his/her preferences conditionally on some scenarios $H \in \wp(S)^0$. Let $\{\precsim_H\}_{H \in \wp(S)^0}$ be a family of preference relations on $\mathcal{F}$, indexed by the set of scenarios $H \in \wp(S)^0$.

For every scenario $H \in \wp(S)^0$, we denote by $\prec_H$ and $\sim_H$ the asymmetric and symmetric parts of $\precsim_H$. Moreover, for every $f, g \in \mathcal{F}$, $f \precsim_H g$ means "f is not preferred to g under the hypothesis H", $f \prec_H g$ means "g is preferred to f under the hypothesis H", and $f \sim_H g$ means "f is indifferent to g under the hypothesis H".

Notice that the set $\mathbf{B}(X)$ contains the set

$$\mathbf{B}_0(X) = \{\delta_B \;:\; B \in \wp(X)^0\},$$

of *vacuous belief functions*, where δ_B is the belief function whose Möbius inversion is such that $m_{\delta_B}(B) = 1$ and 0 otherwise.

Remark 2. *The set* $\mathbf{B}(X)$ *is closed with respect to the convex combination operation defined, for every* $Bel_1, Bel_2 \in \mathbf{B}(X)$ *and every* $\alpha \in [0,1]$, *pointwise, for every* $A \in \wp(X)$, *as*

$$(\alpha Bel_1 + (1-\alpha)Bel_2)(A) = \alpha Bel_1(A) + (1-\alpha)Bel_2(A),$$

and it holds

$$m_{\alpha Bel_1+(1-\alpha)Bel_2}(A) = \alpha m_{Bel_1}(A) + (1-\alpha)m_{Bel_2}(A).$$

In particular, $\mathbf{B}(X)$ *turns out to be the convex closure of* $\mathbf{B}_0(X)$.

The set of acts $\mathcal{F}$ contains, in particular, the set of *constant acts* $\mathcal{F}_c$ whose elements are defined, for every $Bel \in \mathbf{B}(X)$, as

$$\overline{Bel}(s) = Bel, \text{ for all } s \in S.$$

The set $\mathcal{F}$ is closed with respect to the following operation of *convex combination*: for every $f, g \in \mathcal{F}$ and every $\alpha \in [0,1]$, $\alpha f + (1-\alpha)g$ is defined pointwise, for every $s \in S$, as

$$(\alpha f + (1-\alpha)g)(s) = \alpha f(s) + (1-\alpha)g(s).$$

For every $H \in \wp(S)^0$, the relation $\precsim_H$ determines the following other relations:

- the relation $\trianglelefteq_H$ on $\mathbf{B}(X)$ defined, for every $Bel_1, Bel_2 \in \mathbf{B}(X)$, as

$$Bel_1 \trianglelefteq_H Bel_2 \Longleftrightarrow \overline{Bel_1} \precsim_H \overline{Bel_2};$$

- the relation $\leq_H^\bullet$ on $\wp(X)^0$ defined, for every $A, B \in \wp(X)^0$, as

$$A \leq_H^\bullet B \Longleftrightarrow \delta_A \trianglelefteq_H \delta_B;$$

- the relation $\leq_H^*$ on X defined, for every $x, y \in X$, as

$$x \leq_H^* y \Longleftrightarrow \{x\} \leq_H^\bullet \{y\}.$$

Notice that if $\precsim_H$ is a *weak order* on $\mathcal{F}$, i.e., a complete and transitive binary relation, then also the induced relations $\trianglelefteq_H$, $\leq_H^\bullet$ and $\leq_H^*$ are weak orders on the corresponding domains.

We are searching for a representation of $\{\precsim_H\}_{H\in\wp(S)^0}$ in the form of a conditional mixture of Choquet integrals, i.e., for every $f \in \mathcal{F}$ and $H \in \wp(S)^0$,

$$\mathbf{CEU}_{P,u}(f|H) = \sum_{s\in S} P(\{s\}|H)\left(\oint u \mathrm{d} f(s)\right), \tag{4}$$

where $P(\cdot|\cdot)$ is a full conditional probability on $\wp(S)\times\wp(S)^0$ and $u : X \to \mathbb{R}$ is a utility function.

Notice that the $\mathbf{CEU}_{P,u}$ conditional functional expresses a pessimistic aggregation of "objective" uncertainty. A corresponding functional encoding an optimistic aggregation can be defined as

$$\widehat{\mathbf{CEU}}_{P,u}(f|H) = \sum_{s\in S} P(\{s\}|H)\left(\oint u \mathrm{d}\widehat{f(s)}\right), \tag{5}$$

where $\widehat{f(s)}$ it the dual plausibility function of $f(s)$.

Remark 3. *Let us stress that, since $P(\{s\}|H) = 0$ for every $s \in H^c$, the conditional functionals $\mathbf{CEU}_{P,u}$ and $\widehat{\mathbf{CEU}}_{P,u}$ can be rewritten restricting the external sum only to those $s \in H$.*

The family of preference relations $\{\precsim_H\}_{H\in\wp(S)^0}$ is represented by a $\mathbf{CEU}_{P,u}$ conditional functional if, for every $H \in \wp(S)^0$ and every $f, g \in \mathcal{F}$, it holds

$$f \precsim_H g \Longleftrightarrow \mathbf{CEU}_{P,u}(f|H) \leq \mathbf{CEU}_{P,u}(g|H).$$

The representation by a $\widehat{\mathbf{CEU}}_{P,u}$ conditional functional is defined analogously.

Both conditional functionals $\mathbf{CEU}_{P,u}$ and $\widehat{\mathbf{CEU}}_{P,u}$ are conditional expectations of state-contingent Choquet integrals of u. Below we provide a linear expression of the internal Choquet integral in both conditional functionals consisting in a pessimistic or optimistic aggregation of uncertainty.

Let $\leq^*$ be a weak order on X with asymmetric and symmetric parts $<^*$ and $=^*$, respectively, and assume $x_{\sigma(1)} \leq^* \ldots \leq^* x_{\sigma(m)}$, where σ is a permutation of $\{1,\ldots,m\}$. Then, denote by $X^* = X_{/=^*} = \{[x_{i_1}],\ldots,[x_{i_t}]\}$ the quotient set of X under the equivalence relation $=^*$. Thus, $<^*$ can be seen as a strict total order on X^*, and we can assume $[x_{i_1}] <^* \cdots <^* [x_{i_t}]$. The weak order $\leq^*$ will be useful in the sequel to build a utility function on consequences and to get a representation.

The *pessimistic $\leq^*$-aggregated Möbius inversion* associated to $Bel \in \mathbf{B}(X)$ is the function $M_{Bel}^{\leq^*} : X^* \to [0,1]$ defined, for every $[x_{i_j}] \in X^*$, as

$$M^{\lesseqqgtr^*}_{Bel}([x_{i_j}]) = \sum_{x_{\sigma(i)} \in [x_{i_j}]} \sum_{x_{\sigma(i)} \in B \subseteq E_i^\sigma} m_{Bel}(B), \qquad (6)$$

where $E_i^\sigma = \{x_{\sigma(i)}, \ldots, x_{\sigma(m)}\}$ for $i = 1, \ldots, m$.

Note that $M^{\lesseqqgtr^*}_{Bel}([x_{i_j}]) \geq 0$ for every $[x_{i_j}] \in X^*$ and $\sum_{j=1}^t M^{\lesseqqgtr^*}_{Bel}([x_{i_j}]) = 1$, thus $M^{\lesseqqgtr^*}_{Bel}$ determines a probability distribution on X^*.

It is easily seen that, if $u : X \to \mathbb{R}$, then introducing the weak order $\leq^*$ on X such that, for every $x_i, x_j \in X$, $x_i \leq^* x_j$ if and only if $u(x_i) \leq u(x_j)$, for every $Bel \in \mathbf{B}(X)$, by equations (2), (3) and (6) it follows

$$\oint u \mathrm{d} Bel = \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M^{\lesseqqgtr^*}_{Bel}([x_{i_j}]).$$

Let us stress that $M^{\lesseqqgtr^*}_{Bel}$ encodes a pessimistic aggregation of the uncertainty expressed by m_{Bel} [9]. Indeed, it holds

$$\sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M^{\lesseqqgtr^*}_{Bel}([x_{i_j}]) = \sum_{B \in \wp(X)^0} \left(\min_{x \in B} u(x) \right) m_{Bel}(B).$$

In analogy, considering the dual plausibility function $\widehat{Bel}$, the *optimistic* $\leq^*$*-aggregated Möbius inversion* associated to $\widehat{Bel}$ is the function $M^{\lesseqqgtr^*}_{\widehat{Bel}} : X^* \to [0,1]$ defined, for every $[x_{i_j}] \in X^*$, as

$$M^{\lesseqqgtr^*}_{\widehat{Bel}}([x_{i_j}]) = \sum_{x_{\sigma(i)} \in [x_{i_j}]} \sum_{x_{\sigma(i)} \in B \subseteq (E_{i+1}^\sigma)^c} m_{Bel}(B), \qquad (7)$$

where the E_i^σ's are defined as before and $E_{m+1}^\sigma = \emptyset$.

Note that $M^{\lesseqqgtr^*}_{\widehat{Bel}}([x_{i_j}]) \geq 0$ for every $[x_{i_j}] \in X^*$ and $\sum_{j=1}^t M^{\lesseqqgtr^*}_{\widehat{Bel}}([x_{i_j}]) = 1$, thus $M^{\lesseqqgtr^*}_{\widehat{Bel}}$ determines a probability distribution on X^*.

It is easily seen that also in this case by equations (2), (3) and (7) it follows

$$\oint u \mathrm{d} \widehat{Bel} = \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M^{\lesseqqgtr^*}_{\widehat{Bel}}([x_{i_j}]).$$

Let us stress that $M^{\lesseqqgtr^*}_{\widehat{Bel}}$ encodes an optimistic aggregation of the uncertainty expressed by m_{Bel} [9]. Indeed, it holds

$$\sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M^{\lesseqqgtr^*}_{\widehat{Bel}}([x_{i_j}]) = \sum_{B \in \wp(X)^0} \left(\max_{x \in B} u(x) \right) m_{Bel}(B).$$

5. Probabilistic interpretation of numerical models

Consider the product algebra $\wp(S\times X)$ containing two isomorphic copies of $\wp(S)$ and $\wp(X)$ obtained by identifying $A \in \wp(S)$ with $A \times X$ and $B \in \wp(X)$ with $S \times B$. As usual, a Cartesian product is set to $\emptyset$ if one of the factors is equal to $\emptyset$. It is easily seen that every $A \in \wp(S\times X)$ can be written as

$$A = \bigcup_{s\in S}(\{s\} \times [A]_s),$$

with $[A]_s = \{a \in X \,:\, (s,a) \in A\}$, possibly $[A]_s = \emptyset$.

Let $P : \wp(S) \times \wp(S)^0 \to [0,1]$ be a full conditional probability and fix an act $f \in \mathbf{B}(X)^S$. For every $H \in \wp(S)^0$, the pair $(P(\cdot|H), f)$ allows to define the functions Bel^f_H and Pl^f_H on $\wp(S \times X)$ setting, for every $A \in \wp(S \times X)$,

$$Bel^f_H(A) = \sum_{s\in S} P(\{s\}|H) f(s)([A]_s), \tag{8}$$

$$Pl^f_H(A) = \sum_{s\in S} P(\{s\}|H) \widehat{f(s)}([A]_s). \tag{9}$$

The following proposition shows that Bel^f_H and Pl^f_H are a belief and a plausibility function, respectively.

Proposition 1. *The functions Bel^f_H and Pl^f_H defined as in equations* (8) *and* (9) *are a belief and a plausibility function on $\wp(S \times X)$, respectively.*

Proof. Since $[\emptyset]_s = \emptyset$ and $[S \times X]_s = X$, it holds that

$$\begin{aligned}
Bel^f_H(\emptyset) &= \sum_{s\in S} P(\{s\}|H) f(s)(\emptyset) = 0,\\
Bel^f_H(S \times X) &= \sum_{s\in S} P(\{s\}|H) f(s)(X) = 1,\\
Pl^f_H(\emptyset) &= \sum_{s\in S} P(\{s\}|H) \widehat{f(s)}(\emptyset) = 0,\\
Pl^f_H(S \times X) &= \sum_{s\in S} P(\{s\}|H) \widehat{f(s)}(X) = 1.
\end{aligned}$$

Moreover, for every $A, B \in \wp(S \times X)$ with $A \subseteq B$, we have $[A]_s \subseteq [B]_s$ for every $s \in S$, implying $f(s)([A]_s) \le f(s)([B]_s)$ and $\widehat{f(s)}([A]_s) \le \widehat{f(s)}([B]_s)$. Hence, by the monotonicity of the expectation operator with respect to $P(\cdot|H)$, it follows $Bel^f_H(A) \le Bel^f_H(B)$ and $Pl^f_H(A) \le Pl^f_H(B)$, so, both Bel^f_H and Pl^f_H are normalized capacities.

Finally, for every $E_1, \ldots, E_k \in \wp(S \times X)$ and every $\emptyset \neq I \subseteq \{1, \ldots, k\}$, since for every $s \in S$

$$\left[\bigcup_{i \in I} E_i\right]_s = \bigcup_{i \in I} [E_i]_s \quad \text{and} \quad \left[\bigcap_{i \in I} E_i\right]_s = \bigcap_{i \in I} [E_i]_s$$

we have

$$\begin{aligned} f(s)\left(\bigcup_{i=1}^{k} [E_i]_s\right) &\geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} f(s)\left(\bigcap_{i \in I} [E_i]_s\right), \\ \widehat{f(s)}\left(\bigcap_{i=1}^{k} [E_i]_s\right) &\leq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} \widehat{f(s)}\left(\bigcup_{i \in I} [E_i]_s\right). \end{aligned}$$

Hence, by the monotonicity and the linearity of the expectation operator with respect to $P(\cdot|H)$, it follows

$$\begin{aligned} Bel_H^f\left(\bigcup_{i=1}^{k} E_i\right) &\geq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} Bel_H^f\left(\bigcap_{i \in I} E_i\right), \\ Pl_H^f\left(\bigcap_{i=1}^{k} E_i\right) &\leq \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-1)^{|I|+1} Pl_H^f\left(\bigcup_{i \in I} E_i\right). \end{aligned}$$

□

In turn, given the families of belief and plausibility functions $\{Bel_H^f\}_{H \in \wp(S)^0}$ and $\{Pl_H^f\}_{H \in \wp(S)^0}$ defined on $\wp(S \times X)$ we can define two conditional measures $Bel^f(\cdot|\cdot)$ and $Pl^f(\cdot|\cdot)$ with domain $\wp(S \times X) \times \mathcal{H}$, where $\mathcal{H} = \{H \times X : H \in \wp(S)^0\}$ is an additive class isomorphic to $\wp(S)^0$. Such measures are obtained, for every $A|B \in \wp(S \times H) \times \mathcal{H}$ with $B = H \times X$ and $H \in \wp(S)^0$, as

$$Bel^f(A|B) = Bel_H^f(A) \quad \text{and} \quad Pl^f(A|B) = Pl_H^f(A). \tag{10}$$

Proposition 2. *The conditional measure $Bel^f(\cdot|\cdot)$ on $\wp(S \times X) \times \mathcal{H}$ satisfies the following properties:*

- *(i)* $Bel^f(A|B) = Bel^f(A \cap B|B)$*, for every* $A \in \wp(S \times X)$ *and* $B \in \mathcal{H}$*;*
- *(ii)* $Bel^f(\cdot|B)$ *is a belief function on* $\wp(S \times X)$*, for every* $B \in \mathcal{H}$*;*
- *(iii)* $Bel^f(A|C) = Bel^f(A|B) \cdot Bel^f(B|C)$*, for every* $B, C \in \mathcal{H}$ *and* $A \in \wp(S \times X)$ *with* $A \subseteq B \subseteq C$*.*

The conditional measure $Pl^f(\cdot|\cdot)$ on $\wp(S \times X) \times \mathcal{H}$ satisfies the following properties:

(i') $Pl^f(A|B) = Pl^f(A \cap B|B)$, for every $A \in \wp(S \times X)$ and $B \in \mathcal{H}$;

(ii') $Pl^f(\cdot|B)$ is a plausibility function on $\wp(S \times X)$, for every $B \in \mathcal{H}$;

(iii') $Pl^f(A|C) = Pl^f(A|B) \cdot Pl^f(B|C)$, for every $B, C \in \mathcal{H}$ and $A \in \wp(S \times X)$ with $A \subseteq B \subseteq C$.

Moreover, for every $A \in \wp(S \times X)$ and $B \in \mathcal{H}$ it holds

$$Bel^f(A|B) = 1 - Pl^f(A^c|B).$$

Proof. We prove that $Bel^f(\cdot|\cdot)$ satisfies properties *(i)–(iii)* as the satisfaction of properties *(i')–(iii')* by $Pl^f(\cdot|\cdot)$ is proven analogously.

Condition *(i)* follows since, taking into account Remark 3, for every $A \in \wp(S \times X)$ and $B \in \mathcal{H}$ with $B = H \times X$ and for every $s \in H$, it holds $[A]_s = [A \cap B]_s$.

Condition *(ii)* is trivially implied by Proposition 1.

Then, to prove condition *(iii)* take $A \in \wp(S \times X)$ and $B, C \in \mathcal{H}$ with $A \subseteq B \subseteq C$, where $B = H \times X$ and $C = K \times X$, where $H, K \in \wp(S)$ and $H \subseteq K$. Taking into account Remark 3, since $[B]_s$ is equal to X for $s \in H$ and to $\emptyset$ for $s \notin H$, while $[A]_s$ is equal to $\emptyset$ for $s \notin H$ we have

$$\begin{aligned}
Bel^f(A|C) &= \sum_{s \in S} P(\{s\}|K) f(s)([A]_s) = \sum_{s \in K} P(\{s\}|K) f(s)([A]_s), \\
Bel^f(A|B) &= \sum_{s \in S} P(\{s\}|H) f(s)([A]_s) = \sum_{s \in H} P(\{s\}|H) f(s)([A]_s), \\
Bel^f(B|C) &= \sum_{s \in S} P(\{s\}|K) f(s)([B]_s) = \sum_{s \in H} P(\{s\}|K) = P(H|K).
\end{aligned}$$

Hence, recalling Remark 1, for every $s \in H$, it holds that $P(\{s\}|K) = P(\{s\}|H) \cdot P(H|K)$, implying

$$Bel^f(A|C) = Bel^f(A|B) \cdot Bel^f(B|C).$$

Finally, for every $A \in \wp(S \times X)$ and $B \in \mathcal{H}$ with $B = H \times X$, note that it holds $([A]_s)^c = [A^c]_s$, where the complement in the first member is taken with respect to X, while the complement in the second member is taken with respect to $S \times X$. Hence, it follows

$$\begin{aligned} Bel^f(A|B) &= \sum_{s \in S} P(\{s\}|H) f(s)([A]_s) \\ &= \sum_{s \in S} P(\{s\}|H)(1 - \widehat{f(s)}(([A]_s)^c)) \\ &= 1 - \sum_{s \in S} P(\{s\}|H)\widehat{f(s)}(([A]_s)^c) \\ &= 1 - \sum_{s \in S} P(\{s\}|H)\widehat{f(s)}([A^c]_s) \\ &= 1 - Pl^f(A^c|B). \end{aligned}$$

□

By the previous proposition we have that the conditional measure $Pl^f(\cdot|\cdot)$ is a particular conditional submodular capacity agreeing with the axiomatic definition studied in [49], generalizing the Dempster's rule of conditioning [22]. Furthermore, $Bel^f(\cdot|\cdot)$ is the dual of $Pl^f(\cdot|\cdot)$, i.e., $Bel^f(A|B) = 1 - Pl^f(A^c|B)$ for $A \in \wp(S \times X)$ and $B \in \mathcal{H}$. It actually holds that $Bel^f(\cdot|\cdot)$ agrees with an axiomatic definition of conditional supermodular capacity that generalizes the product rule of conditioning [56]. For a discussion on the different notions of conditioning for belief functions see [10].

Notice that the utility function $u : X \to \mathbb{R}$ can be identified with a function $v : S \times X \to \mathbb{R}$ such that, for every $s, s' \in S$ and $x \in X$, it holds $v(s, x) = v(s', x) = u(x)$. Using such v, the previous proposition allows to express both conditional functionals $\mathbf{CEU}_{P,u}$ and $\widehat{\mathbf{CEU}}_{P,u}$ as conditional Choquet expected values of v with respect to Bel^f and Pl^f, respectively

$$\begin{aligned} \mathbf{CEU}_{P,u}(f|H) &= \oint v \mathrm{d} Bel^f(\cdot|H \times X), \\ \widehat{\mathbf{CEU}}_{P,u}(f|H) &= \oint v \mathrm{d} Pl^f(\cdot|H \times X), \end{aligned}$$

for every $f \in \mathcal{F}$ and $H \in \wp(S)^0$.

6. Axioms and representation theorems

The aim is to find a system of axioms for a family of preference relations assuring the representability by conditional functionals introduced in Section 4. For this aim, consider the following axioms.

(A1) Weak order: $\forall H \in \wp(S)^0$, $\precsim_H$ is a weak order on $\mathcal{F}$.

(A2) Continuity: $\forall H \in \wp(S)^0$, $\forall f, g, h \in \mathcal{F}$, if $f \prec_H g \prec_H h$, then $\exists \alpha, \beta \in (0,1)$ such that

$$\alpha f + (1-\alpha)h \prec_H g \prec_H \beta f + (1-\beta)h.$$

(A3) Independence: $\forall H \in \wp(S)^0$, $\forall f, g, h \in \mathcal{F}$ and $\forall \alpha \in (0,1)$

$$f \precsim_H g \Longleftrightarrow \alpha f + (1-\alpha)h \precsim_H \alpha g + (1-\alpha)h.$$

(A4) Monotonicity: $\forall H \in \wp(S)^0$, $\forall f, g \in \mathcal{F}$, if $f(s) \trianglelefteq_H g(s)$, $\forall s \in S$, then $f \precsim_H g$.

(A5) Non-triviality: $\forall H \in \wp(S)^0$, $\exists f, g \in \mathcal{F}$ such that $f \prec_H g$.

(A6) Relevance: $\forall H \in \wp(S)^0$, $\forall f, g \in \mathcal{F}$ with $f(s) = g(s)$, $\forall s \in H$, then $f \sim_H g$.

(A7) Weak dynamic consistency: $\forall f, g \in \mathcal{F}$ and $\forall H, K \in \wp(S)^0$, if $f \precsim_H g$, $f \precsim_K g$, and $H \cap K = \emptyset$, then $f \precsim_{H \cup K} g$.

(A8) State neutrality: $\forall s, t \in S$, if $f(s) = f(t)$, $g(s) = g(t)$, and $f \precsim_{\{s\}} g$, then $f \precsim_{\{t\}} g$.

(A9) Pessimistic aggregated indifference: $\forall H \in \wp(S)^0$, if $M_{f(s)}^{\leq_H^*} = M_{g(s)}^{\leq_H^*}$, $\forall s \in S$, then $f \sim_H g$.

Axioms **(A1)**–**(A5)** are the usual Anscombe-Aumann axioms in the formulation of [33], stated for generalized Anscombe-Aumann acts and every preference relation in $\{\precsim_H\}_{H \in \wp(S)^0}$. Axioms **(A6)**–**(A8)** cope with conditioning. In particular, axiom **(A6)** expresses a focusing conditioning rule, i.e., it states that in conditioning on H, only the part of acts inside of H counts. Axiom **(A7)** copes with relating different conditioning events, while axiom **(A8)** encodes a form of consistency between different states. Finally, axiom **(A9)** is responsible for the $\mathbf{CEU}_{P,u}$ representation: it says that if two possibly distinct acts have the same pessimistic $\leq_H^*$-aggregated Möbius inversion (i.e., the same pessimistic aggregation of "objective" uncertainty) in every state then, they should be judged indifferent given H.

Let us stress that none of axioms **(A6)** and **(A9)** implies the other. Indeed, axiom **(A9)** looks at acts f, g whose pessimistic $\leq_H^*$-aggregated Möbius inversion coincide over the entire S, but it can be $f(s) \neq g(s)$ for some $s \in S$. On the other hand, axiom **(A6)** considers acts f, g which coincide on H but may differ on H^c, so they could have different pessimistic $\leq^*_H$-aggregated Möbius inversions over H^c. The following example illustrates the point.

Example 1. *Take $S = \{s_1, s_2, s_3\}$, $X = \{x_1, x_2, x_3\}$ and $H = \{s_1, s_2\}$, and let $\leq^*_H$ be the weak order on X such that $x_1 <^*_H x_2 <^*_H x_3$. Consider the acts*

S	s_1	s_2	s_3
f	$\delta_{\{x_3\}}$	$\delta_{\{x_2\}}$	$\delta_{\{x_1\}}$
g	$\delta_{\{x_3\}}$	$\delta_{\{x_2\}}$	$\delta_{\{x_3\}}$
h	$\delta_{\{x_1\}}$	$\delta_{\{x_1\}}$	$\delta_{\{x_1\}}$
l	$\delta_{\{x_1,x_2\}}$	$\delta_{\{x_1\}}$	$\delta_{\{x_1,x_2\}}$

Since $\leq^*_H$ is a total order, we can identify $X_{/=^*_H}$ with X, thus we have

X	x_1	x_2	x_3
$M^{\leq^*_H}_{\delta_{\{x_1\}}}$	1	0	0
$M^{\leq^*_H}_{\delta_{\{x_2\}}}$	0	1	0
$M^{\leq^*_H}_{\delta_{\{x_3\}}}$	0	0	1
$M^{\leq^*_H}_{\delta_{\{x_1,x_2\}}}$	1	0	0

Since $f(s) = g(s)$, for every $s \in H$, then axiom **(A6)** *implies $f \sim_H g$, but since $M^{\leq^*_H}_{f(s_3)} \neq M^{\leq^*_H}_{g(s_3)}$ axiom* **(A9)** *does not apply. On the other hand, since $M^{\leq^*_H}_{h(s)} = M^{\leq^*_H}_{l(s)}$, for every $s \in S$, then axiom* **(A9)** *implies $h \sim_H l$, but since $f(s_1) \neq g(s_1)$, axiom* **(A6)** *does not apply.*

The following theorem (already stated in [12] without the proof) shows that axioms **(A1)**–**(A9)** are necessary and sufficient to get a $\mathbf{CEU}_{P,u}$ representation.

Theorem 1. *The following statements are equivalent:*

(i) the family of relations $\{\precsim_H\}_{H \in \wp(S)^0}$ satisfies **(A1)**–**(A9)***;*

(ii) there exist a full conditional probability $P : \wp(S) \times \wp(S)^0 \to [0, 1]$ and a non-constant utility function $u : X \to \mathbb{R}$ such that, for every $f, g \in \mathcal{F}$ and every $H \in \wp(S)^0$,

$$f \precsim_H g \Longleftrightarrow \mathbf{CEU}_{P,u}(f|H) \leq \mathbf{CEU}_{P,u}(g|H).$$

Moreover, P is unique and u is unique up to positive linear transformations.

Proof. Necessity of axioms **(A1)**–**(A9)** is trivially proven, so, we prove only sufficiency.

Identify every belief function $Bel \in \mathbf{B}(X)$ with the corresponding Möbius inversion m_{Bel}. In turn, this allows to identify $\mathbf{B}(X)$ with the set $\Delta(\wp(X)^0)$ of probability distributions over $\wp(X)^0$, i.e., assignments of non-negative weights over $\wp(X)^0$ summing up to 1. So, $\mathcal{F}$ can be identified with $\Delta(\wp(X)^0)^S$, that is a mixture space with respect to pointwise convex combinations.

By the Anscombe-Aumann representation theorem (see [33, 41, 54]), for every $H \in \wp(S)^0$, axioms **(A1)**–**(A5)** imply the existence of a probability measure $P_H : \wp(S) \to [0,1]$ and a non-constant function $v_H : \wp(X)^0 \to \mathbb{R}$ such that the functional $\mathbf{V}_H : \mathcal{F} \to \mathbb{R}$ defined, for every $f \in \mathcal{F}$, as

$$\mathbf{V}_H(f) = \sum_{s \in S} P_H(\{s\}) \left(\sum_{B \in \wp(X)^0} v_H(B) m_{f(s)}(B) \right),$$

represents the relation $\precsim_H$, i.e., for every $f, g \in \mathcal{F}$, $f \precsim_H g \Longleftrightarrow \mathbf{V}_H(f) \leq \mathbf{V}_H(g)$. Notice that, since $\mathbf{V}_H(\overline{\delta_B}) = v_H(B)$, for every $B \in \wp(X)^0$, then v_H represents $\leq^\bullet_H$, i.e., $A \leq^\bullet_H B \Longleftrightarrow v_H(A) \leq v_H(B)$, for every $A, B \in \wp(X)^0$. Moreover, P_H is unique, v_H is unique up to positive linear transformations, and the functional $\mathbf{V}_H$ is linear, i.e., for every $f, g \in \mathcal{F}$ and $\alpha \in [0,1]$, it holds (see Remark 2)

$$\mathbf{V}_H(\alpha f + (1-\alpha) g) = \alpha \mathbf{V}_H(f) + (1-\alpha)\mathbf{V}_H(g).$$

We first show that axioms **(A1)**–**(A8)** imply the existence of a full conditional probability $P : \wp(S) \times \wp(S)^0 \to [0,1]$ and a non-constant function $v : \wp(X)^0 \to \mathbb{R}$ such that the conditional functional defined, for every $f \in \mathcal{F}$ and every $H \in \wp(S)^0$, as

$$\mathbf{V}(f|H) = \sum_{s \in S} P(\{s\}|H) \left(\sum_{B \in \wp(X)^0} v(B) m_{f(s)}(B) \right),$$

represents $\{\precsim_H\}_{H \in \wp(S)^0}$, i.e., for every $f, g \in \mathcal{F}$ and every $H \in \wp(S)^0$,

$$f \precsim_H g \Longleftrightarrow \mathbf{V}(f|H) \leq \mathbf{V}(g|H).$$

Fix an arbitrary $s \in S$ and let $\underline{A}$ and $\overline{A}$ be elements of $\wp(X)^0$ such that, for every $B \in \wp(X)^0$, it holds

$$\underline{A} \leq^\bullet_{\{s\}} B \leq^\bullet_{\{s\}} \overline{A},$$

which is equivalent to $\overline{\delta_{\underline{A}}} \precsim_{\{s\}} \overline{\delta_B} \precsim_{\{s\}} \overline{\delta_{\overline{A}}}$, whose existence follows by axiom **(A1)** and the finiteness of $\wp(X)^0$, due to the finiteness of X. For $t \in S$ with $t \neq s$, since $\overline{\delta_{\underline{A}}} \precsim_{\{s\}} \overline{\delta_B} \precsim_{\{s\}} \overline{\delta_{\overline{A}}}$, axiom **(A7)** implies

$$\overline{\delta_{\underline{A}}} \precsim_{\{t\}} \overline{\delta_B} \precsim_{\{t\}} \overline{\delta_{\overline{A}}}.$$

Furthermore, axiom **(A8)** implies, for every $B \in \wp(X)^0$,

$$\overline{\delta_{\underline{A}}} \precsim_{\{s,t\}} \overline{\delta_B} \precsim_{\{s,t\}} \overline{\delta_{\overline{A}}},$$

and proceeding analogously by adding progressively a state each time, for every $H \in \wp(S)^0$, it holds, for every $B \in \wp(X)^0$,

$$\overline{\delta_{\underline{A}}} \precsim_H \overline{\delta_B} \precsim_H \overline{\delta_{\overline{A}}},$$

which is equivalent to $\underline{A} \leq^{\bullet}_H B \leq^{\bullet}_H \overline{A}$.

Hence, for every $H \in \wp(S)^0$ and every $B \in \wp(X)^0$, since $\mathbf{V}_H$ represents $\precsim_H$ and $\mathbf{V}_H(\overline{\delta_B}) = v_H(B)$, it follows

$$v_H(\underline{A}) \leq v_H(B) \leq v_H(\overline{A}).$$

By the uniqueness of each v_H up to positive linear transformations we can assume

$$v_H(\underline{A}) = 0 \quad \text{and} \quad v_H(\overline{A}) = 1.$$

Denote $\mathbf{1} = \overline{\delta_{\overline{A}}}$ and $\mathbf{0} = \overline{\delta_{\underline{A}}}$, for which we have $\mathbf{V}_H(\mathbf{1}) = 1$ and $\mathbf{V}_H(\mathbf{0}) = 0$.

Fix an arbitrary $s \in S$. For every $B \in \wp(X)^0$ we have $\mathbf{V}_{\{s\}}(\overline{\delta_B}) = v_{\{s\}}(B) = \beta = \mathbf{V}_{\{s\}}(\beta\mathbf{1} + (1-\beta)\mathbf{0})$ with $\beta \in [0,1]$, and since $\mathbf{V}_{\{s\}}$ represents $\precsim_{\{s\}}$ it follows

$$\overline{\delta_B} \sim_{\{s\}} \beta\mathbf{1} + (1-\beta)\mathbf{0}.$$

For $t \in S$ with $t \neq s$, since $\overline{\delta_B} \sim_{\{s\}} \beta\mathbf{1} + (1-\beta)\mathbf{0}$, axiom **(A7)** implies

$$\overline{\delta_B} \sim_{\{t\}} \beta\mathbf{1} + (1-\beta)\mathbf{0}.$$

Furthermore, axiom **(A8)** implies,

$$\overline{\delta_B} \sim_{\{s,t\}} \beta\mathbf{1} + (1-\beta)\mathbf{0},$$

and proceeding analogously by adding progressively a state each time, for every $H \in \wp(S)^0$, it holds

$$\overline{\delta_B} \sim_H \beta\mathbf{1} + (1-\beta)\mathbf{0},$$

which implies

$$v_H(B) = \beta,$$

i.e., for every $H, K \in \wp(S)^0$, we have $v_H = v_K$. Hence, we can define $v : \wp(X)^0 \to \mathbb{R}$ by setting $v = v_H$, where $H \in \wp(S)^0$ is arbitrary.

Now, for every $E \in \wp(S)$, define the act

$$\mathbf{1}_E(s) = \begin{cases} \delta_{\overline{A}} & \text{if } s \in E, \\ \delta_{\underline{A}} & \text{if } s \notin E, \end{cases}$$

for which it holds, for every $H \in \wp(S)^0$,

$$\mathbf{V}_H(\mathbf{1}_E) = P_H(E) = \mathbf{V}_H(P_H(E)\mathbf{1} + (1 - P_H(E))\mathbf{0}),$$

and since $\mathbf{V}_H$ represents $\precsim_H$ it follows

$$\mathbf{1}_E \sim_H P_H(E)\mathbf{1} + (1 - P_H(E))\mathbf{0}. \tag{11}$$

Define $P : \wp(S) \times \wp(S)^0 \to [0, 1]$ setting, for every $E|H \in \wp(S) \times \wp(S)^0$,

$$P(E|H) = P_H(E).$$

We show that $P(\cdot|\cdot)$ is a full conditional probability.

First, since $\mathbf{1}_E(s) = \mathbf{1}_{E\cap H}(s)$ for every $s \in H$, axiom **(A6)** implies $\mathbf{1}_E \sim_H \mathbf{1}_{E\cap H}$ and since $\mathbf{V}_H$ represents $\precsim_H$ it follows

$$P(E|H) = \mathbf{V}_H(\mathbf{1}_E) = \mathbf{V}_H(\mathbf{1}_{E\cap H}) = P(E \cap H|H).$$

Second, for every $H \in \wp(S)^0$, $P(\cdot|H)$ is a probability measure since $P_H(\cdot)$ is.

Finally, take $A \in \wp(S)$ and $B, C \in \wp(S)^0$ with $A \subseteq B \subseteq C$. Since $\mathbf{1}_B(s) = \mathbf{1}(s)$ for every $s \in B$, axiom **(A6)** implies $\mathbf{1}_B \sim_B \mathbf{1}$, moreover, since by (11) it holds $\mathbf{1}_A \sim_B P_B(A)\mathbf{1} + (1 - P_B(A))\mathbf{0}$, applying axiom **(A3)** we derive

$$P_B(A)\mathbf{1}_B + (1 - P_B(A))\mathbf{0} \sim_B P_B(A)\mathbf{1} + (1 - P_B(A))\mathbf{0},$$

and by axiom **(A1)**

$$\mathbf{1}_A \sim_B P_B(A)\mathbf{1}_B + (1 - P_B(A))\mathbf{0}.$$

Since for every $s \in C \setminus B$ we have $\mathbf{1}_A(s) = \mathbf{1}_B(s) = \mathbf{0}(s) = \delta_{\underline{A}}$, axiom **(A6)** implies $\mathbf{1}_B \sim_{C\setminus B} \mathbf{0}$, and applying axioms **(A1)** and **(A3)** we derive

$$\mathbf{1}_A \sim_{C\setminus B} P_B(A)\mathbf{1}_B + (1 - P_B(A))\mathbf{0}.$$

Applying axiom **(A7)** we have

$$\mathbf{1}_A \sim_C P_B(A)\mathbf{1}_B + (1 - P_B(A))\mathbf{0},$$

moreover, by (11) we know that

$$\mathbf{1}_B \sim_C P_C(B)\mathbf{1} + (1 - P_C(B))\mathbf{0}$$

and applying axiom **(A3)** we have

$$\begin{aligned} P_B(A)\mathbf{1}_B + (1 - P_B(A))\mathbf{0} &\sim_C P_B(A)[P_C(B)\mathbf{1} + (1 - P_C(B))\mathbf{0}] + (1 - P_B(A))\mathbf{0} \\ &= P_B(A)P_C(B)\mathbf{1} + (1 - P_B(A)P_C(B))\mathbf{0}, \end{aligned}$$

and by axiom **(A1)** we get

$$\mathbf{1}_A \sim_C P_B(A)P_C(B)\mathbf{1} + (1 - P_B(A)P_C(B))\mathbf{0}.$$

Since $\mathbf{V}_C$ represents $\precsim_C$, and $\mathbf{V}_C(\mathbf{1}_A) = P_C(A)$ and $\mathbf{V}_C(P_B(A)P_C(B)\mathbf{1} + (1 - P_B(A)P_C(B))\mathbf{0}) = P_B(A)P_C(B)$ it follows $P(A|C) = P(A|B)P(B|C)$.

Moreover, $P(\cdot|\cdot)$ is unique since every P_H is, and v is unique up to positive linear transformations since every v_H is.

We finally show that the conditional functional $\mathbf{V}$ has a $\mathbf{CEU}_{P,u}$ expression. Since, for every $H, K \in \wp(S)^0$, $v_H = v_K$, it follows that $\leq^*_H$ and $\leq^*_K$ are the same weak order on X, so, we simply denote $\leq^* = \leq^*_H$ for an arbitrary $H \in \wp(S)^0$ and set $X^* = X_{/=^*} = \{[x_{i_1}], \ldots, [x_{i_t}]\}$ assuming $[x_{i_1}] <^* \ldots <^* [x_{i_t}]$.

For every $Bel \in \mathbf{B}(X)$ and every $H \in \wp(S)^0$, the constant act $\overline{Bel}$ is such that

$$\mathbf{V}(\overline{Bel}|H) = \sum_{B \in \wp(X)^0} v(B) m_{Bel}(B),$$

moreover, for $Bel_1, Bel_2 \in \mathbf{B}(X)$ with $M^{\leq^*}_{Bel_1} = M^{\leq^*}_{Bel_2}$ axiom **(A9)** implies $\overline{Bel_1} \sim_H \overline{Bel_2}$ that, in turn, implies $\mathbf{V}(\overline{Bel_1}|H) = \mathbf{V}(\overline{Bel_2}|H)$, that is

$$\sum_{B \in \wp(X)^0} v(B) m_{Bel_1}(B) = \sum_{B \in \wp(X)^0} v(B) m_{Bel_2}(B).$$

Since for every $Bel \in \mathbf{B}(X)$ there is a probability measure $\mu \in \mathbf{B}(X)$ such that $M^{\leq^*}_{Bel} = M^{\leq^*}_{\mu}$, axiom **(A9)** implies $\overline{Bel} \sim_H \overline{\mu}$, from which we have

$$\sum_{B \in \wp(X)^0} v(B) m_{Bel}(B) = \sum_{B \in \wp(X)^0} v(B) m_{\mu}(B) = \sum_{x \in X} v(\{x\}) \mu(\{x\}).$$

Define $u : X \to \mathbb{R}$ setting, for every $x \in X$,

$$u(x) = v(\{x\}),$$

which is easily seen to represent $\leq^*$, i.e., $x_i \leq^* x_j \iff u(x_i) \leq u(x_j)$, for every $x_i, x_j \in X$. In particular, this implies that u is constant on the equivalence classes in X^*.

This allows to write

$$\sum_{x \in X} v(\{x\})\mu(\{x\}) = \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) \sum_{x \in [x_{i_j}]} \mu(\{x\}),$$

where $\sum_{x \in [x_{i_j}]} \mu(\{x\}) = M_{\mu}^{\leq^*}([x_{i_j}])$, and recalling that $M_{Bel}^{\leq^*} = M_{\mu}^{\leq^*}$ we have

$$\begin{aligned} \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) \sum_{x \in [x_{i_j}]} \mu(\{x\}) &= \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M_{\mu}^{\leq^*}([x_{i_j}]) \\ &= \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M_{Bel}^{\leq^*}([x_{i_j}]). \end{aligned}$$

Hence, we have

$$\sum_{B \in \wp(X)^0} v(B) m_{Bel}(B) = \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M_{Bel}^{\leq^*}([x_{i_j}]) = \oint u \mathrm{d} Bel.$$

Now, we can conclude that, for every $f \in \mathcal{F}$ and every $H \in \wp(S)^0$,

$$\begin{aligned} \mathbf{V}(f|H) &= \sum_{s \in S} P(\{s\}|H) \left(\sum_{B \in \wp(X)^0} v(B) m_{f(s)}(B) \right) \\ &= \sum_{s \in S} P(\{s\}|H) \left(\sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M_{f(s)}^{\leq^*}([x_{i_j}]) \right) \\ &= \sum_{s \in S} P(\{s\}|H) \left(\oint u \mathrm{d} f(s) \right) = \mathbf{CEU}_{P,u}(f|H). \end{aligned}$$

Notice that u cannot be constant otherwise $\mathbf{CEU}_{P,u}$ would be constant too and we would have a contradiction with axiom **(A5)**, since $\mathbf{CEU}_{P,u}$ represents $\{\precsim_H\}_{H \in \wp(S)^0}$. Moreover, independently of $H \in \wp(S)^0$, since

$$\mathbf{CEU}_{P,u}(\overline{\delta_{\underline{A}}}|H) = \min_{x \in \underline{A}} u(x) = 0 \quad \text{and} \quad \mathbf{CEU}_{P,u}(\overline{\delta_{\overline{A}}}|H) = \min_{x \in \overline{A}} u(x) = 1,$$

there are $\underline{x}, \overline{x} \in X$ such that

$$\mathbf{CEU}_{P,u}(\overline{\delta_{\{\underline{x}\}}}|H) = u(\underline{x}) = 0 \quad \text{and} \quad \mathbf{CEU}_{P,u}(\overline{\delta_{\{\overline{x}\}}}|H) = u(\overline{x}) = 1,$$

implying that $\overline{\delta_{\{\underline{x}\}}} \sim_H \overline{\delta_{\underline{A}}}$ and $\overline{\delta_{\{\overline{x}\}}} \sim_H \overline{\delta_{\overline{A}}}$, thus we can take $\{\underline{x}\}, \{\overline{x}\}$ for $\underline{A}, \overline{A}$.

Finally, u is unique up to positive linear transformations since v is. □

In the following we prove that to get a $\widehat{\mathbf{CEU}}_{P,u}$ representation it is sufficient to replace axiom **(A9)** with the following axiom

(A9') Optimistic aggregated indifference: $\forall H \in \wp(S)^0$, if $M_{\widehat{f(s)}}^{\leq^*_H} = M_{\widehat{g(s)}}^{\leq^*_H}$, $\forall s \in S$, then $f \sim_H g$.

Theorem 2. *The following statements are equivalent:*

(i) the family of relations $\{\precsim_H\}_{H \in \wp(S)^0}$ *satisfies* **(A1)**–**(A8)** *and* **(A9')***;*

(ii) there exist a full conditional probability $P : \wp(S) \times \wp(S)^0 \to [0,1]$ *and a non-constant utility function* $u : X \to \mathbb{R}$ *such that, for every* $f, g \in \mathcal{F}$ *and every* $H \in \wp(S)^0$,

$$f \precsim_H g \Longleftrightarrow \widehat{\mathbf{CEU}}_{P,u}(f|H) \leq \widehat{\mathbf{CEU}}_{P,u}(g|H).$$

Moreover, P *is unique and* u *is unique up to positive linear transformations.*

Proof. Necessity of axioms **(A1)**–**(A8)** and **(A9')** is trivially proven, so, we prove only sufficiency.

The proof of sufficiency is completely analogous to that of Theorem 1 up to the derivation of the conditional functional $\mathbf{V}$. We show that **(A9')** implies that the conditional functional $\mathbf{V}$ has a $\widehat{\mathbf{CEU}}_{P,u}$ expression. As in the proof of Theorem 1, since, for every $H, K \in \wp(S)^0$, $v_H = v_K$, it follows that $\leq^*_H$ and $\leq^*_K$ are the same weak order on X, so, we simply denote $\leq^* = \leq^*_H$ for an arbitrary $H \in \wp(S)^0$ and set $X^* = X_{/=^*} = \{[x_{i_1}], \ldots, [x_{i_t}]\}$ assuming $[x_{i_1}] <^* \ldots <^* [x_{i_t}]$.

For every $Bel \in \mathbf{B}(X)$, consider its dual plausibility function $\widehat{Bel}$ and its Möbius inversion m_{Bel}. The constant act $\overline{Bel}$ is such that

$$\mathbf{V}(\overline{Bel}|H) = \sum_{B \in \wp(X)^0} v(B) m_{Bel}(B),$$

moreover, for $Bel_1, Bel_2 \in \mathbf{B}(X)$ with $M^{\leq^*}_{\widehat{Bel_1}} = M^{\leq^*}_{\widehat{Bel_2}}$ axiom **(A9')** implies $\overline{Bel_1} \sim_H \overline{Bel_2}$ that, in turn, implies $\mathbf{V}(\overline{Bel_1}|H) = \mathbf{V}(\overline{Bel_2}|H)$, that is

$$\sum_{B \in \wp(X)^0} v(B) m_{Bel_1}(B) = \sum_{B \in \wp(X)^0} v(B) m_{Bel_2}(B).$$

Since for every $Bel \in \mathbf{B}(X)$ with dual $\widehat{Bel}$ there is a probability measure $\mu \in \mathbf{B}(X)$ such that $M^{\leq^*}_{\widehat{Bel}} = M^{\leq^*}_{\mu}$, axiom **(A9')** implies $\overline{Bel} \sim_H \overline{\mu}$, from which we have

$$\sum_{B \in \wp(X)^0} v(B) m_{Bel}(B) = \sum_{B \in \wp(X)^0} v(B) m_{\mu}(B) = \sum_{x \in X} v(\{x\}) \mu(\{x\}).$$

Define $u : X \to \mathbb{R}$ setting, for every $x \in X$,

$$u(x) = v(\{x\}),$$

which is easily seen to represent $\leq^*$, i.e., $x_i \leq^* x_j \iff u(x_i) \leq u(x_j)$, for every $x_i, x_j \in X$. In particular, this implies that u is constant on the equivalence classes in X^*.

This allows to write

$$\sum_{x \in X} v(\{x\}) \mu(\{x\}) = \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) \sum_{x \in [x_{i_j}]} \mu(\{x\}),$$

where $\sum_{x \in [x_{i_j}]} \mu(\{x\}) = M^{\leq^*}_{\mu}([x_{i_j}])$, and recalling that $M^{\leq^*}_{\widehat{Bel}} = M^{\leq^*}_{\mu}$ we have

$$\begin{aligned} \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) \sum_{x \in [x_{i_j}]} \mu(\{x\}) &= \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M^{\leq^*}_{\mu}([x_{i_j}]) \\ &= \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M^{\leq^*}_{\widehat{Bel}}([x_{i_j}]). \end{aligned}$$

Hence, we have

$$\sum_{B \in \wp(X)^0} v(B) m_{Bel}(B) = \sum_{[x_{i_j}] \in X^*} u(x_{i_j}) M^{\leq^*}_{\widehat{Bel}}([x_{i_j}]) = \oint u \mathrm{d}\widehat{Bel}.$$

Now, we can conclude that, for every $f \in \mathcal{F}$ and every $H \in \wp(S)^0$,

$$\begin{aligned}
\mathbf{V}(f|H) &= \sum_{s\in S} P(\{s\}|H)\left(\sum_{B\in\wp(X)^0} v(B) m_{f(s)}(B)\right) \\
&= \sum_{s\in S} P(\{s\}|H)\left(\sum_{[x_{i_j}]\in X^*} u(x_{i_j}) M^{\leq^*}_{\widehat{f(s)}}([x_{i_j}])\right) \\
&= \sum_{s\in S} P(\{s\}|H)\left(\oint u \mathrm{d}\widehat{f(s)}\right) = \widehat{\mathbf{CEU}}_{P,u}(f|H).
\end{aligned}$$

Notice that u cannot be constant otherwise $\widehat{\mathbf{CEU}}_{P,u}$ would be constant too and we would have a contradiction with axiom **(A5)**, since $\widehat{\mathbf{CEU}}_{P,u}$ represents $\{\precsim_H\}_{H\in\wp(S)^0}$. Moreover, independently of $H \in \wp(S)^0$, since

$$\widehat{\mathbf{CEU}}_{P,u}(\overline{\delta_{\underline{A}}}|H) = \max_{x\in\underline{A}} u(x) = 0 \quad \text{and} \quad \widehat{\mathbf{CEU}}_{P,u}(\overline{\delta_{\overline{A}}}|H) = \max_{x\in\overline{A}} u(x) = 1,$$

there are $\underline{x}, \overline{x} \in X$ such that

$$\widehat{\mathbf{CEU}}_{P,u}(\overline{\delta_{\{\underline{x}\}}}|H) = u(\underline{x}) = 0 \quad \text{and} \quad \widehat{\mathbf{CEU}}_{P,u}(\overline{\delta_{\{\overline{x}\}}}|H) = u(\overline{x}) = 1,$$

implying that $\overline{\delta_{\{\underline{x}\}}} \sim_H \overline{\delta_{\underline{A}}}$ and $\overline{\delta_{\{\overline{x}\}}} \sim_H \overline{\delta_{\overline{A}}}$, thus we can take $\{\underline{x}\}, \{\overline{x}\}$ for $\underline{A}, \overline{A}$.

Finally, u is unique up to positive linear transformations since v is.

□

Let us stress that both $\mathbf{CEU}_{P,u}$ and $\widehat{\mathbf{CEU}}_{P,u}$ conditional functionals allow to take "null" (possible) conditioning events as hypotheses and, even more, they allow to order events in $\wp(S)^0$ according to their "unexpectation".

For that, we define, for every $H, K \in \wp(S)^0$,

$$H \sqsubseteq K \Longleftrightarrow \mathbf{1}_\emptyset \prec_{H\cup K} \mathbf{1}_H,$$

with the meaning "H is no more unexpected than K", where the act $\mathbf{1}_E$, for $E \in \wp(S)$, is defined as in the proof of Theorem 1. We then denote by $\sqsubset$ and $=^\square$, respectively, the asymmetric and symmetric parts of $\sqsubseteq$, where $H \sqsubset K$ means "K is more unexpected than H" and $H =^\square K$ means "none between H and K is more unexpected". It turns out that the relation $\sqsubseteq$ considers the probability of the event H under the hypothesis that either H or K occurs.

Proposition 3. *For every $H, K \in \wp(S)^0$ it holds*

$$\begin{aligned} H \sqsubseteq K &\iff P(H|H \cup K) > 0, \\ H \sqsubset K &\iff P(H|H \cup K) > 0 \text{ and } P(K|H \cup K) = 0, \\ H =^{\square} K &\iff P(H|H \cup K) > 0 \text{ and } P(K|H \cup K) > 0. \end{aligned}$$

Proof. We prove only the case of a $\mathbf{CEU}_{P,u}$ representation, since the case of a $\widehat{\mathbf{CEU}}_{P,u}$ representation is analogous. Suppose the family $\{\precsim_H\}_{H \in \wp(S)^0}$ satisfies axioms **(A1)**–**(A9)**, then by Theorem 1 it has a $\mathbf{CEU}_{P,u}$ representation, where P is unique and u is unique up to positive linear transformations. Thus, we can assume $\min_{x \in X} u(x) = 0$ and $\max_{x \in X} u(x) = 1$.

Since

$$\mathbf{1}_\emptyset \prec_{H \cup K} \mathbf{1}_H \iff \mathbf{CEU}_{P,u}(\mathbf{1}_\emptyset | H \cup K) < \mathbf{CEU}_{P,u}(\mathbf{1}_H | H \cup K),$$

and since $\mathbf{CEU}_{P,u}(\mathbf{1}_\emptyset | H \cup K) = P(\emptyset | H \cup K) = 0$ and $\mathbf{CEU}_{P,u}(\mathbf{1}_H | H \cup K) = P(H|H \cup K)$, the first claim follows.

Then, the other two claims follow since

$$\begin{aligned} H \sqsubset K &\iff H \sqsubseteq K \text{ and } \neg(K \sqsubseteq H), \\ H =^{\square} K &\iff H \sqsubseteq K \text{ and } K \sqsubseteq H. \end{aligned}$$

□

The relation $\sqsubseteq$ reveals to be a weak order on $\wp(S)^0$ and has been originally introduced by [21, 39, 51].

Every full conditional probability $P(\cdot|\cdot)$ on $\wp(S)$ gives rise to a linearly ordered class of probability measures $\{P_0, \ldots, P_k\}$ on $\wp(S)$, said *complete agreeing class*, whose supports form a partition of S [13, 15]. Vice versa, every complete agreeing class $\{P_0, \ldots, P_k\}$ on $\wp(S)$ generates a full conditional probability $P(\cdot|\cdot)$ on $\wp(S)$.

Events with probability 0 essentially determine the structure of a full conditional probability $P(\cdot|\cdot)$ on $\wp(S)$ and actually the relation $\sqsubseteq$ is intimately related to the corresponding $\{P_0, \ldots, P_k\}$.

Remark 4. *Given $P(\cdot|\cdot)$, the corresponding complete agreeing class $\{P_0, \ldots, P_k\}$ can be built through the events (see [15])*

$$H_0^\alpha = \{s \in H_0^{\alpha-1} : P(\{s\}|H_0^{\alpha-1}) = 0\} \quad \text{for } \alpha = 1, \ldots, k,$$

with $H_0^0 = S$, by setting $P_\alpha(\cdot) = P(\cdot|H_0^\alpha)$ where $H_0^\alpha \neq \emptyset$ for $\alpha = 0, \ldots, k$ and $H_0^{k+1} = \emptyset$. On the other hand, given $\{P_0, \ldots, P_k\}$, for every $E|H \in \wp(S) \times$

$\wp(S)^0$ there is a minimum index $\alpha_H \in \{0, \ldots, k\}$ such that $P_{\alpha_H}(H) > 0$ and it holds

$$P(E|H) = \frac{P_{\alpha_H}(E \cap H)}{P_{\alpha_H}(H)}.$$

The class of events $\{H_0^0, \ldots, H_0^k\}$ determines a decreasing class $\{\mathcal{I}_0, \ldots, \mathcal{I}_k\}$ of ideals of $\wp(S)$, singled out by the relation $\sqsubseteq$, defined as

$$\begin{aligned} \mathcal{I}_\alpha &= \{A \in \wp(S)^0 \,:\, H_0^\alpha \sqsubseteq A\} \cup \{\emptyset\} \\ &= \{A \in \wp(S) \,:\, A \subseteq H_0^\alpha\}. \end{aligned}$$

The class of events $\{H_0^0, \ldots, H_0^k\}$ also gives rise to a partition $\mathcal{E} = \{E_0^0, \ldots, E_0^k\}$ of S obtained by setting

$$E_0^\alpha = H_0^\alpha \setminus H_0^{\alpha-1} \quad \textit{for } \alpha = 0, \ldots, k-1,$$

with $E_0^k = H_0^k$, where $E_0^\alpha = \mathbf{supp}(P_\alpha) = \{s \in S \;:\; P_\alpha(\{s\}) > 0\}$ in the complete agreeing class representing $P(\cdot|\cdot)$.

7. Model elicitation

The conditional functionals $\mathbf{CEU}_{P,u}$ and $\widehat{\mathbf{CEU}}_{P,u}$ are completely specified once the full conditional probability $P(\cdot|\cdot)$ and the utility function u have been elicited by the decision maker. In the following, we will focus on $\mathbf{CEU}_{P,u}$ as for every act f and every state $s \in S$ we consider $M_{f(s)}^{\leq^*}$. For the elicitation of $\widehat{\mathbf{CEU}}_{P,u}$ it is sufficient to take $M_{\widehat{f(s)}}^{\leq^*}$ in place of $M_{f(s)}^{\leq^*}$.

In general, the decision maker is only able to provide few comparisons for few conditioning events. In this case, the first issue is to check the consistency of the given comparisons with the model of reference. When consistency holds, it is easily seen that an elicitation procedure relying on a finite number of arbitrary comparisons cannot guarantee any form of uniqueness of P and u in general.

Fixed X and S, we propose an elicitation procedure based on three different cognitive tasks:

1. We ask the decision maker to single out a subset $\mathcal{L} = \{H_1, \ldots, H_N\} \subseteq \wp(S)^0$ that correspond to those events considered as "scenarios of interest" and then to order them according to their unexpectation, by providing a weak order $\sqsubseteq$ on $\mathcal{L}$. The weak order $\sqsubseteq$ is allowed to be trivial, i.e., its asymmetric part can be empty.

2. We ask the decision maker to provide a non-trivial weak order $\leq^*$ on X, i.e., on consequences obtained with certainty: denote $X^* = X_{/=^*} = \{[x_{i_1}], \ldots, [x_{i_t}]\}$ for which $<^*$ is a strict total order, and we can assume $[x_{i_1}] <^* \cdots <^* [x_{i_t}]$.
3. For every $H \in \mathcal{L}$, we ask the decision maker to provide a finite number of strict $\{f_l \prec_H g_l\}_{l \in L_H}$ and weak comparisons $\{f_w \precsim_H g_w\}_{w \in W_H}$, with $L_H \neq \emptyset$ while W_H is allowed to be empty. This assures non-triviality.

Given the weak order $\sqsubseteq$ on $\mathcal{L}$, a full conditional probability $P(\cdot|\cdot)$ is said to be *compatible* with $\sqsubseteq$ if the relation $\sqsubseteq$ is the restriction to $\mathcal{L}$ of the unexpectation relation induced by $P(\cdot|\cdot)$ on the entire $\wp(S)^0$, that is, if for every $H_i, H_j \in \mathcal{L}$ it holds

$$H_i \sqsubseteq H_j \Longleftrightarrow P(H_i | H_i \cup H_j) > 0.$$

On the other hand, a utility function $u : X \to \mathbb{R}$ is said to *represent* the weak order $\leq^*$ on X, if $x_i \leq^* x_j \Longleftrightarrow u(x_i) \leq u(x_j)$, for every $x_i, x_j \in X$. In particular, this implies that u is constant on the equivalence classes in X^*.

The issue is to find:

- a full conditional probability $P(\cdot|\cdot)$ on $\wp(S)$ (and, so, a complete agreeing class $\{P_0, \ldots, P_k\}$ on $\wp(S)$ by Remark 4) *compatible* with the relation $\sqsubseteq$ on $\mathcal{L}$;
- a utility function $u : X \to \mathbb{R}$ *representing* $\leq^*$ on X;

such that the corresponding $\mathbf{CEU}_{P,u}$ conditional functional preserves all the strict and weak preference comparisons.

Let us stress that if we can find $P(\cdot|\cdot)$ and u as above then taking $u' = au + b$ with $a > 0$ we obtain an equivalent representation. For this, we can assume without loss of generality that $\min_{x \in X} u(x) = 0$ and $\max_{x \in X} u(x) = 1$.

At this aim, let $\mathcal{L}_{/=^\square} = \{[H_{i_0}], \ldots, [H_{i_M}]\}$ and assume $[H_{i_0}] \sqsubset \ldots \sqsubset [H_{i_M}]$, where $M = 0$ in case $\sqsubseteq$ is trivial. Now, define $B_0^{M+1} = \emptyset$ and for $\alpha = 0, \ldots, M$

$$B_0^\alpha = \bigcup_{\beta=\alpha}^{M} \bigcup_{H \in [H_{i_\beta}]} H \quad \text{and} \quad E_0^\alpha = B_0^\alpha \setminus B_0^{\alpha+1}.$$

Remark 5. *A necessary condition for the relation $\sqsubseteq$ to admit a compatible full conditional probability $P(\cdot|\cdot)$ is that $E_0^\alpha \neq \emptyset$ and $H \cap E_0^\alpha \neq \emptyset$ for all*

$H \in [H_{i_\alpha}]$, for $\alpha = 0, \ldots, M$. Hence, in what follows we will tacitly assume that this is the case, otherwise the elicitation process must be stopped as the relation $\sqsubseteq$ must be revised by the decision maker.

We call *minimal agreeing class* compatible with $\sqsubseteq$ every linearly ordered class of probability measures $\{P_0^*, \ldots, P_M^*\}$ on $\wp(S)$ such that $\mathbf{supp}(P_\alpha^*) \subseteq E_0^\alpha$ and $P_\alpha^*(H \cap E_0^\alpha) > 0$ for all $H \in [H_{i_\alpha}]$, for $\alpha = 0, \ldots, M$. As shown in the following Remark 6 recalling results already known in the literature [15, 49], the search of a full conditional probability $P(\cdot|\cdot)$ on $\wp(S)$ compatible with $\sqsubseteq$ is equivalent to the search of a minimal agreeing class $\{P_0^*, \ldots, P_M^*\}$ on $\wp(S)$ compatible with $\sqsubseteq$.

Remark 6. *Every minimal agreeing class $\{P_0^*, \ldots, P_M^*\}$ on $\wp(S)$ compatible with $\sqsubseteq$ induces a conditional probability $P^*(\cdot|\cdot)$ on $\wp(S) \times \mathcal{H}$ compatible with $\sqsubseteq$, where $\mathcal{H} = \mathbf{additive}(\mathcal{L})$ is the set of events obtained closing $\mathcal{L}$ with respect to unions. Indeed, see [49], for every $E|H \in \wp(S) \times \mathcal{H}$, there exists a minimum index $\alpha_H \in \{0, \ldots, M\}$ such that $P_{\alpha_H}^*(H) > 0$ and the function defined as*

$$P^*(E|H) = \frac{P_{\alpha_H}^*(E \cap H)}{P_{\alpha_H}^*(H)},$$

is shown to be a conditional probability on $\wp(S) \times \mathcal{H}$ for which it holds, for every $H_i, H_j \in \mathcal{L}$, $H_i \sqsubseteq H_j \iff P^*(H_i|H_i \cup H_j) > 0$. Vice versa, every conditional probability $P^*(\cdot|\cdot)$ on $\wp(S) \times \mathcal{H}$ compatible with $\sqsubseteq$ induces a minimal agreeing class $\{P_0^*, \ldots, P_M^*\}$ on $\wp(S)$ compatible with $\sqsubseteq$, by setting $P_\alpha^*(\cdot) = P^*(\cdot|B_0^\alpha)$, for $\alpha = 0, \ldots, M$.

The conditional probability $P^*(\cdot|\cdot)$ can be further extended (generally not in a unique way) to a full conditional probability $P(\cdot|\cdot)$ on $\wp(S)$ compatible with $\sqsubseteq$ on $\mathcal{L}$. One of the possible extensions is given by the complete agreeing class $\{P_0^*, \ldots, P_M^*, P_{M+1}^*\}$ where P_{M+1}^* is an arbitrary probability measure on $\wp(S)$ whose support is such that $\mathbf{supp}(P_{M+1}^*) = S \setminus \bigcup_{\alpha=0}^M \mathbf{supp}(P_\alpha^*)$. The adjunct of P_{M+1}^* is necessary only if $S \setminus \bigcup_{\alpha=0}^M \mathbf{supp}(P_\alpha^*) \neq \emptyset$, otherwise, $\{P_0^*, \ldots, P_M^*\}$ is already a complete agreeing class on $\wp(S)$ that can be used to generate a full conditional probability on $\wp(S)$ as in Remark 4.

With such an input, the elicitation procedure consists in solving the following system with unknowns the minimal agreeing class $\{P_0^*, \ldots, P_M^*\}$, the utility function u and the control variable δ:

$$\begin{cases} \text{For } \alpha = 0, \ldots, M \text{ and all } H \in [H_{i_\alpha}], \\ \text{For all } l \in L_H, \\ \displaystyle\sum_{s \in H \cap E_0^\alpha} P_\alpha^*(\{s\}) \left(\sum_{[x_{i_j}] \in X^*} u(x_{i_j}) \left(M^{\leq^*}_{f_l(s)}([x_{i_j}]) - M^{\leq^*}_{g_l(s)}([x_{i_j}]) \right) \right) + \delta \leq 0, \\ \text{For all } w \in W_H, \\ \displaystyle\sum_{s \in H \cap E_0^\alpha} P_\alpha^*(\{s\}) \left(\sum_{[x_{i_j}] \in X^*} u(x_{i_j}) \left(M^{\leq^*}_{f_w(s)}([x_{i_j}]) - M^{\leq^*}_{g_w(s)}([x_{i_j}]) \right) \right) \leq 0, \\ \sum_{s \in H \cap E_0^\alpha} P_\alpha^*(\{s\}) - \delta \geq 0, \\ \sum_{s \in E_0^\alpha} P_\alpha^*(\{s\}) = 1, \quad P_\alpha^*(\{s\}) \geq 0, \ \forall s \in E_0^\alpha, \\ u(x_{i_1}) = 0, \quad u(x_{i_t}) = 1, \\ u(x_{i_j}) - u(x_{i_{j+1}}) + \delta \leq 0, \text{ for } j = 1, \ldots, t-1, \\ -1 \leq \delta \leq 1. \end{cases} \tag{12}$$

Solving the above system allows to check both the consistency of the given preference statements and, if consistency holds, to find a full conditional probability $P(\cdot|\cdot)$ and a utility function u giving rise to the conditional functional $\mathbf{CEU}_{P,u}$. Indeed, the preference statements are consistent with the model if and only if we can find a solution $\{P_0^*, \ldots, P_M^*\}, u, \delta$ with $\delta > 0$ and in this case the solution of the system determines $P(\cdot|\cdot)$ and u, up to the possible arbitrary choice of the probability measure P_{M+1}^* (see Remark 6).

Theorem 3. *The following statements are equivalent:*

- *(i) the system* (12) *admits a solution* $\{P_0^*, \ldots, P_M^*\}, u, \delta$ *with* $\delta > 0$*;*
- *(ii) there exist a full conditional probability* $P : \wp(S) \times \wp(S)^0 \to [0,1]$ *compatible with* $\sqsubseteq$ *and a non-constant utility function* $u : X \to \mathbb{R}$ *representing* $\leq^*$ *and with* $\min_{x \in X} u(x) = 0$ *and* $\max_{x \in X} u(x) = 1$ *such that,*

for every $H \in \mathcal{L}$,

$$\begin{cases} f_l \prec_H g_l & \Longrightarrow \quad \mathbf{CEU}_{P,u}(f_l|H) < \mathbf{CEU}_{P,u}(g_l|H) \quad \textit{for all } l \in L_H, \\ f_w \precsim_H g_w & \Longrightarrow \quad \mathbf{CEU}_{P,u}(f_w|H) \leq \mathbf{CEU}_{P,u}(g_w|H) \quad \textit{for all } w \in W_H. \end{cases}$$

Proof. Let $\mathcal{H} = \mathbf{additive}(\mathcal{L})$ be the additive class generated by $\mathcal{L}$ obtained by closing $\mathcal{L}$ with respect to unions. Let $\mathcal{L}_{/=^\square} = \{[H_{i_0}], \ldots, [H_{i_M}]\}$ and assume $[H_{i_0}] \sqsubset \ldots \sqsubset [H_{i_M}]$. Define $B_0^{M+1} = \emptyset$ and for $\alpha = 0, \ldots, M$

$$B_0^\alpha = \bigcup_{\beta=\alpha}^{M} \bigcup_{H \in [H_{i_\beta}]} H \quad \text{and} \quad E_0^\alpha = B_0^\alpha \setminus B_0^{\alpha+1}.$$

Finally, denote $X^* = X_{/=^*} = \{[x_{i_1}], \ldots, [x_{i_t}]\}$ for which $<^*$ is a strict total order with $[x_{i_1}] <^* \cdots <^* [x_{i_t}]$.

(ii) $\Longrightarrow$ *(i)*. Suppose *(ii)* holds. Let $P^* = P_{|\wp(S)\times\mathcal{H}}$ be the restriction of P to $\wp(S) \times \mathcal{H}$, which is a conditional probability. We have that P^* is compatible with $\sqsubseteq$ since P is, so, taking $P^*_\alpha(\cdot) = P^*(\cdot|B_0^\alpha)$ for $\alpha = 0, \ldots, M$, we get a minimal agreeing class $\{P^*_0, \ldots, P^*_M\}$ (see Remark 6) of probability measures on $\wp(S)$ such that $\mathbf{supp}(P^*_\alpha) \subseteq E_0^\alpha$ and $P^*_\alpha(H \cap E_0^\alpha) > 0$ for all $H \in [H_{i_\alpha}]$.

We also have that $0 = u(x_{i_1}) < \cdots < u(x_{i_t}) = 1$ since u represents $\leq^*$ and is such that $\min_{x\in X} u(x) = 0$ and $\max_{x\in X} u(x) = 1$.

For $\alpha = 0, \ldots, M$ and all $H \in [H_{i_\alpha}]$, fixing two acts $f, g \in \mathbf{B}(X)^S$, the quantity $P^*_\alpha(H) \cdot (\mathbf{CEU}_{P,u}(f|H) - \mathbf{CEU}_{P,u}(g|H))$ reduces to

$$\sum_{s\in H\cap E_0^\alpha} P^*_\alpha(\{s\}) \left(\sum_{[x_{i_j}]\in X^*} u(x_{i_j}) \left(M^{\leq^*}_{f(s)}([x_{i_j}]) - M^{\leq^*}_{g(s)}([x_{i_j}]) \right) \right).$$

Defining

$$\begin{aligned} \delta_1 &= \min\{P^*_\alpha(H \cap E_0^\alpha) \,:\, H \in [H_{i_\alpha}], \alpha = 0, \ldots, M\}, \\ \delta_2 &= \min\{u(x_{i_{j+1}}) - u(x_{i_j}) \,:\, j = 1, \ldots, t-1\}, \\ \delta_3 &= \min\{P^*_\alpha(H) \cdot (\mathbf{CEU}_{P,u}(g_l|H) - \mathbf{CEU}_{P,u}(f_l|H)) \,:\, l \in L_H, H \in [H_{i_\alpha}], \alpha = 0, \ldots, M\}, \\ \delta &= \min\{\delta_1, \delta_2, \delta_3\}, \end{aligned}$$

we have that $\{P^*_0, \ldots, P^*_M\}, u, \delta$ is a solution of (12) with $\delta > 0$.

(i) $\Longrightarrow$ *(ii)*. Let $\{P^*_0, \ldots, P^*_M\}, u, \delta$ be a solution of (12) with $\delta > 0$. By construction (see Remark 6) we have that $\{P^*_0, \ldots, P^*_M\}$ is a minimal

agreeing class of probability measures on $\wp(S)$ such that $\mathbf{supp}(P^*_\alpha) \subseteq E^\alpha_0$ and $P^*_\alpha(H \cap E^\alpha_0) > 0$ for all $H \in [H_{i_\alpha}]$. For every $E|H \in \wp(S) \times \mathcal{H}$, let $\alpha_H \in \{0, \dots, M\}$ be the minimum index such that $P^*_{\alpha_H}(H) > 0$ and define

$$P^*(E|H) = \frac{P^*_{\alpha_H}(E \cap H)}{P^*_{\alpha_H}(H)}.$$

The function $P^*(\cdot|\cdot)$ is a conditional probability that can be extended (generally not in a unique way) to a full conditional probability $P(\cdot|\cdot)$ on $\wp(S)$.

We show that $P^*(\cdot|\cdot)$ is compatible with the relation $\sqsubseteq$ on $\mathcal{L}$ as this implies that also every full conditional probability extending it is. Let $H_i, H_j \in \mathcal{L}$ with $H_i \sqsubseteq H_j$, which implies that there are indices $\alpha \leq \beta$ such that $H_i \in [H_{i_\alpha}]$ and $H_j \in [H_{i_\beta}]$. In turn, the construction implies $\alpha_{H_i} = \alpha_{H_i \cup H_j} = \alpha$, from which it follows

$$P^*(H_i|H_i \cup H_j) = \frac{P^*_\alpha(H_i)}{P^*_\alpha(H_i \cup H_j)} > 0.$$

Vice versa, let $H_i, H_j \in \mathcal{L}$ with $P^*(H_i|H_i \cup H_j) > 0$. By the monotonicity with respect to set inclusion of the P^*_α's it must be $\alpha_{H_i} = \alpha_{H_i \cup H_j} = \alpha \leq \beta = \alpha_{H_j}$, so, we have $H_i \in [H_{i_\alpha}]$ and $H_j \in [H_{i_\beta}]$ that implies $H_i \sqsubseteq H_j$.

Furthermore, setting $u(x) = u(x_{i_j})$ for every $x \in [x_{i_j}]$, with $j = 1, \dots, t$, we get a non-constant utility function $u : X \to \mathbb{R}$ representing $\leq^*$ and with $\min_{x \in X} u(x) = 0$ and $\max_{x \in X} u(x) = 1$.

Finally, since for $\alpha = 0, \dots, M$ and all $H \in [H_{i_\alpha}]$, fixing two acts $f, g \in \mathbf{B}(X)^S$, the expression

$$\sum_{s \in H \cap E^\alpha_0} P^*_\alpha(\{s\}) \left(\sum_{[x_{i_j}] \in X^*} u(x_{i_j}) \left(M^{\leq^*}_{f(s)}([x_{i_j}]) - M^{\leq^*}_{g(s)}([x_{i_j}]) \right) \right)$$

is equal to the quantity $P^*_\alpha(H) \cdot (\mathbf{CEU}_{P,u}(f|H) - \mathbf{CEU}_{P,u}(g|H))$, we have that all strict and weak conditional preference comparisons are preserved by the resulting $\mathbf{CEU}_{P,u}$ functional. □

In order to search for a solution $\{P^*_0, \dots, P^*_M\}, u, \delta$ of system (12) with $\delta > 0$ we can introduce the following optimization problem:

$$\begin{gathered}\text{maximize } \delta \\ \text{subject to: system (12).}\end{gathered}$$

In case the above problem is not feasible or the optimal solution is such that $\delta \leq 0$ we can conclude that the information elicited by the decision maker is not consistent with the model $\mathbf{CEU}_{P,u}$.

The above optimization problem is a Quadratically Constrained Linear Program (QCLP). Unfortunately, as shown in Example 2, the quadratic forms in the quadratic constraints of the problem are generally not positive semidefinite nor negative semidefinite, so, the problem is generally not convex: interior points algorithms are not suitable. The problem can be solved with a branch and bound algorithm coping with global optimization of non-linear problems, such as the COUENNE solver [17].

Example 2. *Let $X = \{x_1, x_2\}$ be such that $x_1 <^* x_2$, and $S = \{s_1, s_2\}$ with $\mathcal{L} = \{S\}$ where $\sqsubseteq$ is the trivial weak order $S \sqsubseteq S$, implying that $E_0^0 = S$. Consider the vacuous belief functions $\delta_{\{x_1\}}$ and $\delta_{\{x_2\}}$, whose pessimistic $\leq^*$-aggregated Möbius inversions are (we can identify X^* with X, since $\leq^*$ is a total order)*

	x_1	x_2
$M^{\leq^*}_{\delta_{\{x_1\}}}$	1	0
$M^{\leq^*}_{\delta_{\{x_2\}}}$	0	1

Take the acts

	s_1	s_2
f	$\delta_{\{x_1\}}$	$\delta_{\{x_2\}}$
g	$\delta_{\{x_2\}}$	$\delta_{\{x_1\}}$

and consider the preference statement $f \prec_S g$ that corresponds to the constraint

$$p_1^0 u_1 + p_2^0 u_2 - p_1^0 u_2 - p_2^0 u_1 + \delta \leq 0,$$

where $p_i^0 = P_0^*(\{s_i\})$, for $i = 1, 2$, and $u_j = u(x_j)$, for $j = 1, 2$. The above constraint can be rewritten in matrix notation as

$$\frac{1}{2}\mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x} \leq 0,$$

where

$$\mathbf{x} = \begin{bmatrix} p_1^0 \\ p_2^0 \\ u_1 \\ u_2 \\ \delta \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The eigenvalues of matrix A are $-2, 0, 2$, so, it is neither positive semidefinite nor negative semidefinite.

The following example illustrates the elicitation procedure.

Example 3. *Let $S = \{s_1, s_2, s_3\}$ and $X = \{x_1, x_2, x_3\}$ with $x_1 <^* x_2 <^* x_3$ and take Bel_1, Bel_2 and Bel_3 whose corresponding Möbius inversions are:*

	$\emptyset$	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	X
m_{Bel_1}	0	$\frac{1}{3}$	0	$\frac{1}{3}$	0	$\frac{1}{3}$	0	0
m_{Bel_2}	0	$\frac{1}{5}$	$\frac{1}{5}$	0	$\frac{1}{5}$	0	$\frac{2}{5}$	0
m_{Bel_3}	0	$\frac{1}{4}$	0	$\frac{1}{4}$	0	0	0	$\frac{2}{4}$

The pessimistic $\leq^*$-aggregated Möbius inversions are (we can identify X^* with X, since $\leq^*$ is a total order)

	x_1	x_2	x_3
$M^{\leq^*}_{Bel_1}$	$\frac{2}{3}$	0	$\frac{1}{3}$
$M^{\leq^*}_{Bel_2}$	$\frac{2}{5}$	$\frac{3}{5}$	0
$M^{\leq^*}_{Bel_3}$	$\frac{3}{4}$	0	$\frac{1}{4}$

Take the acts

	s_1	s_2	s_3
f	Bel_1	Bel_2	Bel_2
g	Bel_3	Bel_3	Bel_2
h	Bel_2	Bel_1	Bel_1

Take $\mathcal{L} = \{H_1 = \{s_1, s_2\}, H_2 = \{s_3\}\}$ with $H_1 \sqsubset H_2$ and consider the preferences

$$f \prec_{H_1} g, \quad g \prec_{H_1} h, \quad g \prec_{H_2} h.$$

In this case we have that $E_0^0 = H_1$ and $E_0^1 = H_2$. To avoid cumbersome notation, denote $p_i^\alpha = P_\alpha^*(\{s_i\})$ and $u_j = u(x_j)$. We need to solve the *following optimization problem*

$$\text{maximize } \delta$$

$$\textit{subject to:}$$

$$\begin{cases}
p_1^0 u_1\left(-\frac{1}{12}\right) + p_1^0 u_3\left(\frac{1}{12}\right) + p_2^0 u_1\left(-\frac{7}{20}\right) + p_2^0 u_2\left(\frac{3}{5}\right) + p_2^0 u_3\left(-\frac{1}{4}\right) + \delta \leq 0, \\
p_1^0 u_1\left(\frac{7}{20}\right) + p_1^0 u_2\left(-\frac{3}{5}\right) + p_1^0 u_3\left(\frac{1}{4}\right) + p_2^0 u_1\left(\frac{1}{12}\right) + p_2^0 u_3\left(-\frac{1}{12}\right) + \delta \leq 0, \\
p_3^1 u_1\left(-\frac{4}{15}\right) + p_3^1 u_2\left(\frac{3}{5}\right) + p_3^1 u_3\left(-\frac{1}{3}\right) + \delta \leq 0, \\
p_1^0 + p_2^0 - \delta \geq 0, \\
p_1^0 + p_2^0 = 1, \quad p_1^0, p_2^0 \geq 0, \\
p_3^1 - \delta \geq 0, \\
p_3^1 = 1, \quad p_3^1 \geq 0, \\
u_1 = 0, \quad u_3 = 1, \\
u_1 - u_2 + \delta \leq 0, \\
u_2 - u_3 + \delta \leq 0, \\
-1 \leq \delta \leq 1,
\end{cases}$$

for which the COUENNE *solver finds the solution* $p_1^0 = 0$, $p_2^0 = 1$, $p_3^1 = 1$, $u_1 = 0$, $u_2 = 0.0850159$, $u_3 = 1$, *and* $\delta = 0.0833333$. *Since* $\delta > 0$ *the preference statements are consistent with the model and a full conditional probability* $P(\cdot|\cdot)$ *on* $\wp(S)$ *is that represented by the complete agreeing class* $\{P_0^*, P_1^*, P_2^*\}$ *(to which the probability measure* P_2^* *has been added) whose distributions are*

	$\{s_1\}$	$\{s_2\}$	$\{s_3\}$
P_0^*	0	1	0
P_1^*	0	0	1
P_2^*	1	0	0

It actually holds that

$$\mathbf{CEU}_{P,u}(f|H_1) = 0.0510095 < \mathbf{CEU}_{P,u}(g|H_1) = 0.25 < \mathbf{CEU}_{P,u}(h|H_1) = 0.\overline{3},$$

$$\mathbf{CEU}_{P,u}(g|H_2) = 0.0510095 < \mathbf{CEU}_{P,u}(h|H_2) = 0.\overline{3}.$$

8. Motivating examples: continuation

8.1. An investment decision problem (continued)

Consider the investment decision problem described in Subsection 2.1.

Suppose that our decision maker is not able to express directly his/her preference between f and g, conditionally on K and K^c. Nevertheless, our decision maker is a profit maximizer and believes that a war between North Korea and USA next year is unexpected, while it is more likely a decrease of Italian GDP next year.

The fact that event K is unexpected, i.e., it is judged as "null" by our decision maker, does not rule out its possible realization. In particular, if event K were true then our decision maker believes that it would be more likely an increase of Italian GDP, due to a profit of Italian weapons manufacturers.

Hence, our decision maker is able to provide the following information:

- $\mathcal{L} = \{K, K^c\}$ with $K^c \sqsubset K$,
- $-50 <^* 0 <^* 100$,

that allow to set, according to the proof of Theorem 1, $\underline{A} = \{-50\}$, $\overline{A} = \{100\}$ and to define, for every $E \in \wp(S)$, the act

$$\mathbf{1}_E(s) = \begin{cases} \delta_{\overline{A}}, & \text{if } s \in E, \\ \delta_{\underline{A}}, & \text{if } s \notin E. \end{cases}$$

In turn, the beliefs of our decision maker can be translated as follows:

$$\mathbf{1}_{\{s_3\}} \prec_{K^c} \mathbf{1}_{\{s_4\}} \quad \text{and} \quad \mathbf{1}_{\{s_2\}} \prec_K \mathbf{1}_{\{s_1\}}.$$

In this case we have that $E_0^0 = K^c$ and $E_0^1 = K$. To avoid cumbersome notation, denote $p_i^\alpha = P_\alpha^*(\{s_i\})$ and $u_1 = u(-50), u_2 = u(0), u_3 = u(100)$. We need to solve the following optimization problem

$$\text{maximize } \delta$$
$$\text{subject to:}$$
$$\begin{cases} -p_3^0 u_1 + p_3^0 u_3 + p_4^0 u_1 - p_4^0 u_3 + \delta \leq 0, \\ p_1^1 u_1 - p_1^1 u_3 - p_2^1 u_1 + p_2^1 u_3 + \delta \leq 0, \\ p_3^0 + p_4^0 - \delta \geq 0, \\ p_3^0 + p_4^0 = 1, \quad p_3^0, p_4^0 \geq 0, \\ p_1^1 + p_2^1 - \delta \geq 0, \\ p_1^1 + p_2^1 = 1, \quad p_1^1, p_2^1 \geq 0, \\ u_1 = 0, \quad u_3 = 1, \\ u_1 - u_2 + \delta \leq 0, \\ u_2 - u_3 + \delta \leq 0, \\ -1 \leq \delta \leq 1, \end{cases}$$

for which the COUENNE solver finds the solution $p_3^0 = 0.19795$, $p_4^0 = 0.80205$, $p_1^1 = 0.80205$, $p_2^1 = 0.19795$, $u_1 = 0$, $u_2 = 0.5$, $u_3 = 1$, and $\delta = 0.5$. Since $\delta > 0$ the preference statements are consistent with the model and a full conditional probability $P(\cdot|\cdot)$ on $\wp(S)$ is that represented by the complete agreeing class $\{P_0^*, P_1^*\}$ whose distributions are

	$\{s_1\}$	$\{s_2\}$	$\{s_3\}$	$\{s_4\}$
P_0^*	0	0	0.19795	0.80205
P_1^*	0.80205	0.19795	0	0

With such $P(\cdot|\cdot)$ and u we have

$$\begin{aligned} \mathbf{CEU}_{P,u}(g|K) = 0 &< 0.059385 = \mathbf{CEU}_{P,u}(f|K), \\ \mathbf{CEU}_{P,u}(g|K^c) = 0.07918 &< 0.380205 = \mathbf{CEU}_{P,u}(f|K^c), \end{aligned}$$

that imply, $g \prec_K f$ and $g \prec_{K^c} f$, i.e., under both hypotheses the decision maker should choose f.

8.2. A multi-criteria decision problem with uncertain profiles (continued)

Consider the multi-criteria decision problem of Subsection 2.2 and denote $A_1 = \{a_1^1, a_2^1, a_3^1\}$, $A_2 = \{a_1^2, a_2^2, a_3^2\}$ and $A_3 = \{a_1^3, a_2^3, a_3^3\}$, where the values a_j^i's have the meaning discussed before. Suppose that every A_i is totally ordered according to the indices of its elements and endow $X = A_1 \times A_2 \times A_3 = \{x_1, \ldots, x_{27}\}$ with the lexicographic order $\leq^*$ induced by the total orders of the A_i's. We can assume $x_1 <^* \cdots <^* x_{27}$, where $x_1 = (a_1^1, a_1^2, a_1^3)$ and $x_{27} = (a_3^1, a_3^2, a_3^3)$.

If we interpret S as a set of criteria, then the relation $\sqsubseteq$ on $\wp(S)^0$ can be given an "irrelevance" interpretation, that is, for every $A, B \in \wp(S)^0$, $A \sqsubseteq B$ means "A is no more irrelevant than B". Under this interpretation, take $\mathcal{L} = \{H = \{s_1, s_2\}, H^c = \{s_3\}\}$ with $H =^\square H^c$, meaning that none of the two groups of criteria is irrelevant compared to the other. In this case we have $E_0^0 = S$.

If we consider the preference statements $f \prec_H g$ and $f \prec_{H^c} g$, then we need to solve the problem (we can identify X^* with X, since $\leq^*$ is a total order)

$$\text{maximize } \delta$$
$$\text{subject to:}$$
$$\begin{cases} \displaystyle\sum_{s \in \{s_1,s_2\}} P_0^*(\{s\}) \left(\sum_{x_i \in X} u(x_i) \left(M_{f(s)}^{\leq^*}(x_i) - M_{g(s)}^{\leq^*}(x_i) \right) \right) + \delta \leq 0, \\ P_0^*(\{s_3\}) \left(\displaystyle\sum_{x_i \in X} u(x_i) \left(M_{f(s)}^{\leq^*}(x_i) - M_{g(s)}^{\leq^*}(x_i) \right) \right) + \delta \leq 0, \\ P_0^*(\{s_1\}) + P_\alpha^*(\{s_2\}) - \delta \geq 0, \\ P_0^*(\{s_3\}) - \delta \geq 0, \\ \displaystyle\sum_{s \in S} P_0^*(\{s\}) = 1, \quad P_0^*(\{s\}) \geq 0, \ \forall s \in S, \\ u(x_1) = 0, \quad u(x_{27}) = 1, \\ u(x_i) - u(x_{i+1}) + \delta \leq 0, \text{ for } i = 1, \ldots, 26, \\ -1 \leq \delta \leq 1. \end{cases}$$

The above optimization problem does not admit a solution $\{P_0^*\}, u, \delta$ such that $\delta > 0$. Indeed, every full conditional probability $P(\cdot|\cdot)$ on $\wp(S) \times \wp(S)^0$ is such that $P(\{s_3\}|\{s_3\}) = 1$, so, taking a utility function u representing $\leq^*$ with $u(x_1) = 0$ and $u(x_{27}) = 1$, we have

$$\mathbf{CEU}_{P,u}(f|H^c) = 0.5u(x_3) > 0 = \mathbf{CEU}_{P,u}(g|H^c),$$

where $x_3 = (a_1^1, a_1^2, a_3^3)$ with $u(x_3) > 0$, that contradicts $f \prec_{H^c} g$.

9. Conclusions

The presented models generalize the conditional version of the Anscombe-Aumann model given in [46] by introducing "objective" ambiguity on consequences, modeled in the Dempster-Shafer theory. The main reason for restricting to acts mapping states of the world to belief functions on consequence, is their interpretation as state-contingent partially known randomizing devices as in Ellsberg's paradox [28]. Nevertheless, a possible natural generalization is to consider acts mapping states of the world to supermodular/submdoular capacities or, even, sets of probability measures on consequences, as done in [57].

We provide rationality axioms for conditional preferences encoding either a systematically pessimistic or optimistic behavior of an agent in resolving uncertainty on consequences. These axioms assure the representability of the conditional preference relations by means of a conditional functional parametrized by a full conditional probability on the states of the world and a utility function on consequences. For each scenario, this conditional functional consists in a conditional expectation of state-contingent Choquet expected utility models.

Our models allow conditioning on "null" events and, even more, they allow to order (possible) events according to their "unexpectation". This feature comes from the use of full conditional probabilities.

Here we consider comparisons of acts only conditionally on the same event H, i.e., $f \precsim_H g$: a possible extension of our models could be to consider comparisons under different conditioning events $(f, H) \precsim (f, K)$ as done in [31].

Our models do not allow for ambiguity on the states of the world, as "subjective" uncertainty is assumed to be probabilistic. Nevertheless, the presented models could be extended in a way to take into account also "subjective" ambiguity on the states of the world. To achieve this, we need to relax some axioms of the presented axiomatization, in particular axiom **(A3)** which is responsible for the additivity of the conditional measure. In doing so, we could achieve a representation where "subjective" uncertainty is expressed by a conditional submodular/supermodular capacity [49], or other types of conditional uncertainty measures. This will be the subject of future research.

Acknowledgements

The authors are members of the GNAMPA-INdAM research group. The first author was supported by Università degli Studi di Perugia, Fondo Ricerca di Base 2019, project "Modelli per le decisioni economiche e finanziarie in condizioni di ambiguità ed imprecisione".

References

References

[1] D.S. Ahn. Ambiguity without a state space. *The Review of Economic Studies*, 75(1):3–28, 2008.

[2] F.J. Anscombe and R.J. Aumann. A definition of subjective probability. *The Annals of Mathematical Statistics*, 34(1):199–205, 1963.

[3] V. Biazzo, A. Gilio, T. Lukasiewicz, and G. Sanfilippo. Probabilistic logic under coherence: complexity and algorithms. *Annals of Mathematics and Artificial Intelligence*, 45:35–81, 2005.

[4] A. Capotorti and B. Vantaggi. Locally strong coherence in inference processes. *Annals of Mathematics and Artificial Intelligence*, 35:125–149, 2002.

[5] P. Casaca, A. Chateauneuf, and J.H. Faro. Ignorance and competence in choices under uncertainty. *Journal of Mathematical Economics*, 54:143–150, 2014.

[6] A. Chateauneuf and J.H. Faro. Ambiguity through confidence functions. *Journal of Mathematical Economics*, 45(9):535–558, 2009.

[7] A. Chateauneuf and J.-Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17(3):263–283, 1989.

[8] G. Coletti. Coherent numerical and ordinal probabilistic assessments. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(12):1747–1754, 1994.

[9] G. Coletti, D. Petturiti, and B. Vantaggi. Rationality principles for preferences on belief functions. *Kybernetika*, 51(3):486–507, 2015.

[10] G. Coletti, D. Petturiti, and B. Vantaggi. Conditional belief functions as lower envelopes of conditional probabilities in a finite setting. *Information Sciences*, 339:64–84, 2016.

[11] G. Coletti, D. Petturiti, and B. Vantaggi. When upper conditional probabilities are conditional possibility measures. *Fuzzy Sets and Systems*, 304:45–64, 2016.

[12] G. Coletti, D. Petturiti, and B. Vantaggi. Decisions on generalized Anscombe-Aumann acts under possibly "unexpected" scenarios. *Proc. WUPES 2018*, pages 49–60, 2018.

[13] G. Coletti and R. Scozzafava. Characterization of coherent conditional probabilities as a tool for their assessment and extension. *Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 44:101–132, 1996.

[14] G. Coletti and R. Scozzafava. The role of coherence in eliciting and handling imprecise probabilities and its application to medical diagnosis. *Information Sciences*, 130:41–65, 2000.

[15] G. Coletti and R. Scozzafava. *Probabilistic Logic in a Coherent Setting*, volume 15 of *Trends in Logic*. Kluwer Academic Publisher, Dordrecht/Boston/London, 2002.

[16] G. Coletti, R. Scozzafava, and B. Vantaggi. Inferential processes leading to possibility and necessity. *Information Sciences*, 245:132–145, 2013.

[17] Couenne. https://projects.coin-or.org/Couenne/.

[18] F.G. Cozman and L. F. di Ianni. Probabilistic satisfiability and coherence checking through integer programming. *International Journal of Approximate Reasoning*, 58:57–70, 2015.

[19] B. de Finetti. Sull'impostazione assiomatica del calcolo delle probabilità. *Annali Triestini*, 19:29–81, 1949.

[20] B. de Finetti. Aggiunta alla nota sull'assiomatica della probabilità. *Annali Triestini*, 20:3–20, 1950.

[21] B. de Finetti. *Theory of Probability 1-2*. John Wiley & Sons, London, New York, Sydney, Toronto, 1975.

[22] A.P. Dempster. Upper and Lower Probabilities Induced by a Multivalued Mapping. *Annals of Mathematical Statistics*, 38(2):325–339, 1967.

[23] D. Denneberg. *Non-Additive Measure and Integral*. Kluwer Academic Publisher, 1994.

[24] T. Denoeux. Decision-making with belief functions: A review. *International Journal of Approximate Reasoning*, 109:87–110, 2019.

[25] T. Denoeux and P.P. Shenoy. An Axiomatic Utility Theory for Dempster-Shafer Belief Functions. In *Proceedings of Machine Learning Research*, volume 103, pages 145–155. 2019.

[26] L.E. Dubins. Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability*, 3(1):89–99, 1975.

[27] D. Dubois and H. Prade. When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49(1):65–74, 1992.

[28] D. Ellsberg. Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643–669, 1961.

[29] J. Etner, M. Jeleva, and J.-M. Tallon. Decision Theory under Ambiguity. *Journal of Economic Surveys*, 26(2):234–270.

[30] P.C. Fishburn. *Utility Theory for Decision Making.* John Wiley and Sons, New York, 1970.

[31] P.C. Fishburn. *The foundations of expected utility.* Springer Science, Dordrecht, 1982.

[32] P. Ghirardato. Coping with ignorance: unforeseen contingencies and non-additive uncertainty. *Economic Theory*, 17:247–276, 2001.

[33] I. Gilboa. *Theory of Decision under Uncertainty.* Cambridge University Press, 2009.

[34] I. Gilboa and M. Marinacci. *Ambiguity and the Bayesian Paradigm*, chapter 21. Springer International Publishing Switzerland 2016.

[35] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

[36] M. Grabisch. *Set Functions, Games and Capacities in Decision Making.* Springer, 2016.

[37] J.-Y. Jaffray. Linear utility theory for belief functions. *Operations Research Letters*, 8(2):107–112, 1989.

[38] J.-Y. Jaffray and P. Wakker. Decision making with belief functions: Compatibility and incompatibility with the sure-thing principle. *Journal of Risk and Uncertainty*, 7(3):255–271, 1993.

[39] P.H. Krauss. Representation of conditional probability measures on Boolean algebras. *Acta Mathematica Hungarica*, 19(3):229–241, 1968.

[40] D.M. Kreps. A representation theorem for "preference for flexibility". *Econometrica*, 47(3):56–577, 1979.

[41] D.M. Kreps. *Notes On The Theory Of Choice.* Westview Press, 1988.

[42] D.M. Kreps and R. Wilson. Sequential equilibria. *Econometrica*, 50(4):863–894, 1982.

[43] F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica*, 74(6):1447–1498.

[44] E. Miranda, I. Couso, and P. Gil. A random set characterization of possibility measures. *Information Sciences*, 168(1):51–75, 2004.

[45] I. Montes, E. Miranda, and P. Vicig. Outer approximating coherent lower probabilities with belief functions. *International Journal of Approximate Reasoning*, 110:1–30, 2019.

[46] R.B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.

[47] K. Nehring. Preference for flexibility in a Savage framework. *Econometrica*, 67:101–119, 1999.

[48] W. Olszewski. Preferences over sets of lotteries. *The Review of Economic Studies*, 74(2):567–595, 2007.

[49] D. Petturiti and B. Vantaggi. Conditional submodular Choquet expected values and conditional coherent risk measures. *International Journal of Approximate Reasoning*, 113:14–38, 2019.

[50] A. Rényi. On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungarica*, 6(3):285–335, 1955.

[51] A. Rényi. On conditional probability spaces generated by a dimensionally ordered set of measures. *Theory of Probability & Its Applications*, 1(1):55–64, 1956.

[52] L.J. Savage. *The foundations of statistics*. Wiley, New York, 1954.

[53] D. Schmeidler. Integral representation without additivity. *Proceedings of the American Mathematical Society*, 97(2):255–261, 1986.

[54] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

[55] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.

[56] G. Shafer. A theory of statistical evidence. In W.L. Harper and C.A. Hooker, editors, *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, volume 6b of *The University*

of Western Ontario Series in Philosophy of Science, pages 365–436. Springer Netherlands, 1976.

[57] M.-L. Vierø. Exactly what happens after the Anscombe–Aumann race? *Economic Theory*, 41(2):175–212, 2009.

[58] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.