Desirability foundations of robust rational decision making

Marco Zaffalon[a,*], Enrique Miranda[b]

[a] *Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA). Lugano (Switzerland)*
[b] *University of Oviedo, Dep. of Statistics and Operations Research. Oviedo (Spain)*

Abstract

Recent work has formally linked the traditional axiomatisation of incomplete preferences à la Anscombe-Aumann with the theory of desirability developed in the context of imprecise probability, by showing in particular that they are the very same theory. The equivalence has been established under the constraint that the set of possible prizes is finite. In this paper, we relax such a constraint, thus *de facto* creating one of the most general theories of rationality and decision making available today. We provide the theory with a sound interpretation and with basic notions, and results, for the separation of beliefs and values, and for the case of complete preferences. Moreover, we discuss the role of conglomerability for the presented theory, arguing that it should be a rationality requirement under very broad conditions.

1. Introduction

Background

Preferences play two important roles in artificial intelligence (AI). On the one side, the axiomatisation of preferences has been a major achievement of the past century, with the pioneering work of Von Neumann and Morgestern [55], Savage [59], Anscombe and Aumann [2]. These authors have justified rational decision making through the maximisation of expected utility and at the same time have provided foundations for the Bayesian theory of probability, which is at the heart of so much work in AI. On the other side, in the recent years we have witnessed a steady increase of research in preference modelling as a ductile and practical way to assess a subject's beliefs, to learn trends from data, etc. (see, e.g., Domshlak et al [30], Pigozzi et al [58]). So preferences are very flexible objects that allow us both to work out foundations for our uncertainty models and to develop useful tools for inference.

One aspect of preferences that seems particularly promising for modern artificial intelligence is their possibility to model incompleteness: the fact that we are not always capable to compare alternatives. In fact, modelling the incompleteness in our knowledge is key to develop reliable (or *robust*) applications of AI, which is a central theme now that AI is becoming so widespread. This has eminently been pointed out by Dietterich [29], the former President of AAAI.

Research on the axiomatisation of incomplete preferences has primarily been done in communities other than AI, with the work of Bewley [8], Seidenfeld et al [60], Nau [53], Ok et al [56], Galaabaatar and Karni [34] (but see the work by Dubois et al [32]). In fact, the theme of incompleteness or robustness is not mainstream in AI yet, even though we can find well-developed algorithms and methods, like those by Corani and Zaffalon [12], Kikuti et al [41], de Bock and de Cooman [16], just to mention a few.

What these latter proposals have in common is their reference to Walley's 1991 theory of *imprecise probability* as their founding paradigm (see Augustin et al [4] for a collection of introductory papers on imprecise probability and some main applications to artificial intelligence). Such a theory is a re-elaboration of Williams' 1975 theory, which, in turn, is an extension of de Finetti's 1937 Bayesian theory (see also de Finetti [26]) made to deal with imprecision in probabilities, as originated by incompleteness or other reasons.

*Corresponding author
Email addresses: zaffalon@idsia.ch (Marco Zaffalon), mirandaenrique@uniovi.es (Enrique Miranda)

In short, Williams-Walley's theory models a subject's uncertainty by focusing on the gambles that the subject is willing to take and that for this reason are called *desirable*. The rationality of the assessments is then formalised through some axioms, which we can regard as a way to axiomatise the theory of desirability. The resulting uncertainty theory is very powerful: above all, it maintains the characteristic of logical self-consistency already present in de Finetti's theory, which it encompasses as a special case (the property of self-consistency is called *coherence* in the jargon of imprecise probability). But it is capable of modelling incompleteness in great generality too; in fact it encompasses also robust Bayesianism, and more in general the approaches that model uncertainty with sets of probabilities. And moreover it goes well beyond that, since desirability smoothly deals also with non-Archimedean problems, those that probabilities cannot capture. And it has all this done without incurring measurement problems.

Desirability and preference have essentially lived two separate lives until recently, when we have proved that they are actually the same theory (Zaffalon and Miranda [76]; an earlier report dates back to 2015). This may be surprising as preferences deal with beliefs (probabilities) and values (utilities) at the same time, while desirability was born as a theory of uncertainty alone. In fact, the equivalence has been achieved by enriching desirability models through a set of possible prizes, in addition to the usual set of states of nature.

The equivalence of the two theories thus established has proved fruitful in many respects. From the viewpoint of preference, it has allowed to show for example that the traditional notion of 'state independence', i.e., the fact that we can actually have separate models for beliefs and values, corresponds to the notion of 'strong independence' in probability [14]: this is based on the idea of using a set of stochastically independent models. It has permitted to rework the notions of Archimedeanity in preference in a new way, and to streamline the notions of precise beliefs and values, and more in general of complete preferences.

Conversely, the equivalence has allowed us to generalise desirability so as to deal with imprecise non-linear utility. This has made of desirability one the most powerful, and well founded, theories of uncertainty and decision making available today. As such, it appears to be a natural foundation for robust methods in modern AI.

Outline and contributions

With this in mind, in this paper we revisit our work from 2017, aiming to widen its scope and to discuss more broadly some important issues that we left open. After an introduction to desirability and preferences in Section 2, we then give the following contributions:

Section 3. In our 2017 formulation, we considered an unconstrained space of possibilities Ω and a finite space of prizes $\mathcal{X}$. This is particularly restrictive when we want to consider for instance continuous rewards. It is a case that has received quite some attention in the literature (see the work by del Amo and Ríos Insua [27], Dubra et al [33], Ghirardato and Siniscalchi [35], Ok et al [56], just to name a few) but it does not seem to have found a fully general formulation yet. In fact we also discussed this possibility pointing to some inherent difficulty.

The main problem in this case is that the core modelling unit in preferences, that of 'horse lottery' does not easily lend itself to work with continuous sets of prizes—on top of which, by definition, we should place a probability. We address this problem by defining the new notion of *conditional horse lottery* and reformulating the axiomatisation of preferences on that basis. Then we prove the equivalence with desirability in full generality where both the spaces involved, of possibilities and prizes, are unconstrained. We make this by showing that there is a one-to-one correspondence between the axioms of preference and those of desirability.

In the Archimedean case, this leads to represent utilities via sets of finitely additive probabilities. More generally speaking, we benefit from the representation power of desirability so as to formulate a theory that can work with any spaces, Archimedean and non-Archimedean problems, and without incurring problems of measurability. To our knowledge this type of generality is achieved here for the first time.

Section 4 gives a rigorous interpretation of the rewards that conditional horse lotteries deliver by constructing a linear utility scale based on so-called 'probability currency'.

Section 5 compares the traditional formalism based on horse lotteries and our new one based on conditional horse lotteries. We show that they are equivalent in the case where the prizes are in a finite number. When they are not, we discuss why our formalism appears to be more suited to deal with infinitely many prizes.

Section 6 focuses on the question of conglomerability. Conglomerability is a property that probabilistic, as well as desirability, modelling has to face when we deal with infinite spaces of possibilities, and in particular with infinite conditioning partitions of those spaces. This is a subject of interest to us given the unconstrained nature of the space of possibilities that we consider. The main question is whether or not conglomerability should be a rationality axiom.

Conglomerability has been first discussed by de Finetti [23] and later on by many other authors, including Hill and Lane [38], Armstrong [3], Seidenfeld et al [61], Howson [39], DiBella [28]. Berti et al [7] have recently given an overview of the subject. Despite the amount of work devoted to conglomerability, it still remains a controversial question to date.

In this panorama, we have provided an original justification of conglomerability through arguments of 'temporal coherence', that is, the idea that an uncertainty model should be dynamically coherent in time (Zaffalon and Miranda [74]). This idea holds the promise to finally resolve the controversy, given that conglomerability turns out to be a consequence of temporal coherence. But the assumptions we imposed to derive such a result make its application relatively narrow.

We reconsider the problem and show that we can weaken those assumptions in a substantial way, thus *de facto* turning conglomerability into a rationality requirement whenever a model is used to compute future beliefs and values—in the same way as probabilistic updating is usually understood.

Appendix A summarises our past results about the decomposition and completeness of preferences, and proves that those results continue to hold when the space of prizes is infinite.

Our conclusions and future perspectives are in Section 7. The proofs of the paper have been collected in Appendix B.

2. Preliminary notions

In this section, we introduce the main features of sets of desirable gambles and preference relations that shall be needed in the remainder of the paper.

2.1. Desirable gambles

We next introduce the basic theory of sets of desirable gambles; for a deeper account of the results and concepts mentioned in this section, we refer to the work of Walley [71, Section 3.7]; Couso and Moral [13], de Cooman and Miranda [21], de Cooman and Quaeghebeur [22], Miranda [45], Miranda and Zaffalon [47] and Augustin et al [4, Chapter 1]. For a philosophical defence of indeterminate probabilities, see, e.g., Hájek and Smithson [37].

In short, the idea is to consider states of nature and associated real-valued rewards called *gambles*. The focus will be on those gambles that are 'desirable' to us and that have the property of being coherent, which means that they satisfy some rationality axioms (Definition 3 as well as Axioms D1–D4 below). A coherent set of desirable gambles determines a so-called *coherent lower prevision*, which has the interpretation of a lower expectation. In turn, that determines a set of finitely additive probabilities.[1] Coherent sets of desirable gambles, as well as coherent lower previsions and sets of probabilities, can be conditioned, marginalised or extended. The former are however more general than the latter, in that they can readily deal with conditioning on events of probability zero or with choices under zero expectation, for which the traditional, Kolmogorovian, approach to probability is not applicable.

Let us see these notions in more detail now.

2.1.1. Logical foundations of desirability

Let Ω denote the set of possible outcomes of an experiment, that is, the *space of possibilities*. We let the cardinality of Ω be general, so Ω can be infinite. A *gamble* $f : \Omega \to \mathbb{R}$ is a bounded real-valued function of Ω. A gamble is interpreted as an uncertain reward in a linear utility scale. Accepting a gamble f is regarded as a commitment to receive $f(\omega)$ whatever ω occurs.

[1] Such a correspondence follows from a separating hyperplane theorem.

Denote by $\mathcal{L}(\Omega)$ the set of all the gambles on Ω and by $\mathcal{L}^+(\Omega) := \{f \in \mathcal{L}(\Omega) : f \gneq 0\}$ the subset of the *positive gambles*, where $f \gneq 0$ means that $f \geq 0$ and $f \neq 0$. We denote these sets also by $\mathcal{L}$ and $\mathcal{L}^+$, respectively, when there can be no ambiguity about the space involved. Negative gambles are defined analogously by $\mathcal{L}^- := -\mathcal{L}^+$.

We examine a set of gambles $Q \subseteq \mathcal{L}$ and come up with the subset $\mathcal{K}$ of the gambles in Q that we find desirable. How can we characterise the rationality of the assessments represented by $\mathcal{K}$?

We can follow the procedure adopted in similar cases in logic, where first of all we need to introduce a notion of deductive closure: that is, we first characterise the set of gambles that we must find desirable as a consequence of having desired $\mathcal{K}$ in the first place. This is easy to do once we assume that our utility scale is linear; those gambles are the positive linear combinations of gambles in $\mathcal{K}$:

$$\operatorname{posi}(\mathcal{K}) := \left\{ \sum_{j=1}^{r} \lambda_j f_j : f_j \in \mathcal{K}, \lambda_j > 0, r \geq 1 \right\}.$$

We must also consider that any gamble in $\mathcal{L}^+$ should be desirable as well, given that it may increase our utility without ever decreasing it. Stated differently, the set $\mathcal{L}^+$ plays the role of the tautologies in logic. This means that the actual deductive closure we are after is given by the following:

Definition 1. (**Natural extension for gambles**) Given a set $\mathcal{K}$ of desirable gambles, its *natural extension* $\mathcal{R}$ is the set of gambles given by

$$\mathcal{R} := \operatorname{posi}(\mathcal{K} \cup \mathcal{L}^+).$$

Note that $\mathcal{R}$ is the smallest convex cone that includes $\mathcal{K} \cup \mathcal{L}^+$.

The rationality of the assessments is characterised through the natural extension by the following:

Definition 2. (**Avoiding partial loss for gambles**) A set $\mathcal{K}$ of desirable gambles is said to *avoid partial loss* if $0 \notin \mathcal{R}$.

This condition is the analog of the notion of consistency in logic. The irrationality of a natural extension that incurs partial loss depends on the fact (as it is possible to show) that it must contain a negative gamble f, that is, one that cannot increase our utility and can possibly decrease it. In contradistinction, a set that avoids partial loss does not contain negative gambles.

There is a final notion that is required to make a full logical theory of desirability. This is the logical notion of a theory, that is, a set of assessments that is consistent and logically closed, in the sense that the consistent assessments coincide with their deductive closure in the examined domain Q. This means that we are fully aware of the implications of our assessments on other elements of Q. The logical notion of a theory goes in desirability under the name of coherence:

Definition 3 (**Coherence for gambles**). Say that $\mathcal{K}$ is *coherent relative to* Q if $\mathcal{K}$ avoids partial loss and $Q \cap \mathcal{R} \subseteq \mathcal{K}$ (and hence $Q \cap \mathcal{R} = \mathcal{K}$). In case $Q = \mathcal{L}$ then we simply say that $\mathcal{K}$ is *coherent*.

This definition alone, despite its conceptual simplicity, makes up all the theory of desirable gambles: in principle, every property of the theory can be derived from it. Moreover, the definition gives the theory a solid logical basis and in particular guarantees that the inferences one draws are always coherent with one another. This appears to giving a reply to Howson's 2009 question, whether probability and logic can be combined (see also Howson [40]): by reformulating 'probability' as desirability, we do not even need to combine probability with logic, as we clearly see that probability is already pure logic. At the same time, reformulating probability through desirability, we obtain a very powerful theory: as we have seen, it can be defined on any space of possibilities Ω and any domain $Q \subseteq \mathcal{L}$ (in this sense, it is not affected by measurability problems); and, as we shall make precise later on in this section, it can handle both precise and imprecise assessments, as well as model both Archimedean and non-Archimedean problems.

2.1.2. The case $Q = \mathcal{L}$

In this paper we shall primarily be concerned with the case where the set of gambles Q that we examine is the entire set $\mathcal{L}$. We shall then focus on this case to develop the theory of desirability in some detail.

The first consequence of focusing on the case $Q = \mathcal{L}$ is that coherence can equivalently be characterised by four axioms (see Walley [71, Section 3.7]; Miranda and Zaffalon [47, Proposition 2]):

Definition 4 (**Coherence for gambles when $Q = \mathcal{L}$**). In the case where $Q = \mathcal{L}$, $\mathcal{R}$ is said to be a *coherent set of desirable gambles* in $\mathcal{L}$ if and only if it satisfies the following conditions:

D1. $\mathcal{L}^+ \subseteq \mathcal{R}$ [Accepting Partial Gains];

D2. $0 \notin \mathcal{R}$ [Avoiding Null Gain];

D3. $f, g \in \mathcal{R} \Rightarrow f + g \in \mathcal{R}$ [Additivity];

D4. $f \in \mathcal{R}, \lambda > 0 \Rightarrow \lambda f \in \mathcal{R}$ [Positive Homogeneity].

A coherent set of desirable gambles is thus a convex cone (D3, D4) that excludes the origin (D2) and that contains the positive gambles (D1).

Consider an event $B \subseteq \Omega$. We denote by B both the subset of Ω and its indicator function I_B (that equals one in B and zero elsewhere). Using this convention, we can multiply B and a gamble f obtaining the gamble Bf given for all $\omega \in \Omega$ by

$$Bf(\omega) = \begin{cases} f(\omega) & \text{if } \omega \in B, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 5 (**Conditioning for gambles**). Consider a coherent set of desirable gambles $\mathcal{R}$ on $\mathcal{L}$ and let B be a non-empty subset of Ω. The set of desirable gambles conditional on B derived from $\mathcal{R}$ is defined as

$$\mathcal{R}|B := \{f \in \mathcal{R} : f = Bf\}.$$

There is a natural correspondence between Bf and the restriction of f to B, whence we can also put $\mathcal{R}|B$ in relation with

$$\mathcal{R}\rfloor B := \{f_B \in \mathcal{L}(B) : (\exists f \in \mathcal{R}|B)(\forall \omega \in B)\ f(\omega) = f_B(\omega)\},$$

which is a coherent set of desirable gambles in $\mathcal{L}(B)$. $\mathcal{R}|B$ and $\mathcal{R}\rfloor B$ are equivalent representations of the conditional set.

Definition 6 (**Marginalisation for gambles**). Let $\mathcal{R}$ be a coherent set of desirable gambles in the product space $\Omega \times \Omega'$, where Ω, Ω' are two logically independent sets. The *Ω-marginal set of desirable gambles on $\mathcal{L}(\Omega \times \Omega')$ induced by $\mathcal{R}$* is defined as

$$\mathcal{R}_\Omega := \{f \in \mathcal{R} : (\forall \omega \in \Omega)(\forall \omega_1', \omega_2' \in \Omega')\ f(\omega, \omega_1') = f(\omega, \omega_2')\}.$$

Since $\mathcal{R}_\Omega$ is made up of gambles that depend on Ω only (we call them *Ω-measurable* gambles), we can establish a bijection between $\mathcal{R}_\Omega$ and

$$\mathcal{R}'_\Omega := \{g \in \mathcal{L}(\Omega) : (\exists f \in \mathcal{R})(\forall \omega \in \Omega)(\forall \omega' \in \Omega')\ g(\omega) = f(\omega, \omega')\}.$$

$\mathcal{R}'_\Omega$ is a coherent set of gambles in $\mathcal{L}(\Omega)$.

A coherent set of desirable gambles encompasses a probabilistic model for Ω made of lower and upper expectations (also called *previsions* after de Finetti):

Definition 7 (**Coherent lower and upper previsions**). Let $\mathcal{R}$ be a coherent set of desirable gambles in $\mathcal{L}$. For all $f \in \mathcal{L}$, let

$$\underline{P}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathcal{R}\}, \tag{1}$$

where addition of a gamble with a constant is meant pointwise. It is called the *lower prevision* of f. The conjugate value given by $\overline{P}(f) := -\underline{P}(-f)$ is called the *upper prevision* of f. The functionals $\underline{P}, \overline{P} : \mathcal{L} \to \mathbb{R}$ are respectively called a *coherent lower prevision* and a *coherent upper prevision*.

Definition 8 (**Linear prevision**). Let $\underline{P}, \overline{P}$ be coherent lower and upper previsions on $\mathcal{L}$. If $\underline{P}(f) = \overline{P}(f)$ for some $f \in \mathcal{L}$, then we call the common value the *prevision* of f and we denote it by $P(f)$. If this happens for all $f \in \mathcal{L}$ then we call the functional P a *linear prevision*.

A coherent lower prevision $\underline{P}$ has a set of *dominating* linear previsions:

$$\mathcal{M}(\underline{P}) \coloneqq \{P \text{ linear prevision} : (\forall f \in \mathcal{L})\ P(f) \geq \underline{P}(f)\}, \tag{2}$$

which turns out to be closed[2] and convex. Since each linear prevision is in a one-to-one correspondence with a finitely additive probability, we can regard $\mathcal{M}(\underline{P})$ also as a set of probabilities (or a *credal set*). Moreover, $\underline{P}$ is the lower envelope of the previsions in $\mathcal{M}(\underline{P})$:

$$(\forall f \in \mathcal{L})\ \underline{P}(f) = \min\{P(f) : P \in \mathcal{M}(\underline{P})\}.$$

The coherent upper prevision $\overline{P}$ is the upper envelope of the same set; as a consequence, it follows that $\underline{P}(f) \leq \overline{P}(f)$ for all $f \in \mathcal{L}$.

Definition 9 (**Strict desirability**). A coherent set of gambles $\mathcal{R}$ is said to be *strictly desirable* if it satisfies the following condition:

D0. $f \in \mathcal{R} \setminus \mathcal{L}^+ \Rightarrow (\exists \delta > 0)\ f - \delta \in \mathcal{R}$ [Openness].

Strict desirability means that $\mathcal{R} \setminus \mathcal{L}^+$ does not include its topological border. By an abuse of terminology, $\mathcal{R}$ is said to be open too.[3]

There is a one-to-one correspondence between coherent lower previsions and strictly desirable sets: from $\underline{P}$ we can induce set

$$\mathcal{R} \coloneqq \{f \in \mathcal{L} : f \gneq 0 \text{ or } \underline{P}(f) > 0\}; \tag{3}$$

$\mathcal{R}$ is coherent and strictly desirable and moreover induces $\underline{P}$ through Eq. (1).

Remark 1 (**General vs strict and almost desirability**). Note that a general coherent set of desirable gambles is a cone that need not be open nor closed. The closure of a coherent set of desirable gambles is called a set of *almost-desirable* gambles by Walley [71, Section 3.7.3]. Strict desirability, as we have seen, is associated with open cones. Both almost and strictly desirable gambles are in a one-to-one correspondence with, and are therefore as informative as, a closed convex set of linear previsions. The extra generality of sets of desirable gambles makes them useful when dealing with the problem of conditioning on sets of probability zero, or of choosing between two options under zero expectation (Zaffalon and Miranda [74, Section 5]; see also Dubins [31], Seidenfeld et al [60], Coletti and Scozzafava [11], de Bock and de Cooman [17] for other approaches to this problem).

Remark 2 (**On the regularity assumption**). Note also that with coherent sets of desirable gambles we need not enter the controversy as to whether or not we should use the 'regularity assumption', which prescribes that probabilities of possible events should be positive (see Pedersen [57]). Such an assumption goes back to an important article by Shimony [65]. In the language of this paper, Shimony argued that de Finetti's framework could lead to the questionable undesirability of a positive gamble in case zero probabilities were present, given that the prevision (expectation) of such a gamble could be zero. In other words, Shimony argued, as we do, that D1 is an axiom of rationality. This led Shimony and a number of later authors, among whom Carnap [9] and Skyrms [66], to advocate strengthening de Finetti's theory by requiring regularity. But it has originated also much controversy given the very constraining nature of regularity on probabilistic models. Between requiring regularity and dropping the desirability of positive gambles, coherent sets of desirable gambles actually offer a third possible way that keeps all desiderata together: events can have zero probability, positive gambles are desirable, and all operations, such as conditioning or comparisons, are well defined and informative. $\Diamond$

Definition 10 (**Conditional coherent lower and upper previsions**). Let $\mathcal{R}$ be a coherent set of desirable gambles in $\mathcal{L}(\Omega)$ and B a non-empty subset of Ω. For all $f \in \mathcal{L}(\Omega)$, let

$$\underline{P}(f|B) \coloneqq \sup\{\mu \in \mathbb{R} : B(f - \mu) \in \mathcal{R}\} \tag{4}$$

[2] In the weak* topology, which is the smallest topology for which all the evaluation functionals given by $f(P) \coloneqq P(f)$, where $f \in \mathcal{L}$, are continuous.

[3] A note of caution to prevent confusion in the reader: the adjective 'strict' denotes two unrelated things in desirability and in preferences. In preferences it characterises irreflexive relations, while in desirability it formalises an Archimedean condition as it will become clear in Section 3. We are keeping the same adjective in both cases for historical reasons and given that there should be no possibility to create ambiguity by doing so.

be the *conditional lower prevision* of f given B. The conjugate value given by $\overline{P}(f|B) \coloneqq -\underline{P}(-f|B)$ is called the *conditional upper prevision* of f. The functionals $\underline{P}(\cdot|B), \overline{P}(\cdot|B) : \mathcal{L}(\Omega) \to \mathbb{R}$ are respectively called a *conditional coherent lower prevision* and a *conditional coherent upper prevision*.

Denote by $\inf_B f$ the infimum value that f takes on B. $\underline{P}(\cdot|B)$ satisfies the following conditions for all $f \in \mathcal{L}(\Omega)$ and all real $\lambda > 0$:

- $\underline{P}(f|B) \geq \inf_B f$;
- $\underline{P}(\lambda f|B) = \lambda \underline{P}(f|B)$;
- $\underline{P}(f + g|B) \geq \underline{P}(f|B) + \underline{P}(g|B)$.

These conditions are sometimes used as axioms of coherent conditional lower previsions, when not taking desirability as the primitive concept. Note that coherent lower previsions can be defined as a special case of conditional ones obtained when $B = \Omega$.

Definition 11 (**Maximal coherent set of gambles**). Let $\mathcal{R}$ be a coherent set of desirable gambles. It is called *maximal* if

$$(\forall f \in \mathcal{L} \setminus \{0\})\ f \notin \mathcal{R} \Rightarrow -f \in \mathcal{R}.$$

Requiring maximality is tantamount to assuming complete preferences.

It is important for this paper to also say something about *conglomerability*:

Definition 12 (**Conglomerability**). Consider a coherent set of desirable gambles $\mathcal{R} \subseteq \mathcal{L}(\Omega)$ and let $\mathcal{B}$ be a partition of Ω. $\mathcal{R}$ is said to be *$\mathcal{B}$-conglomerable* if it satisfies the following condition:[4]

$$(\forall f \in \mathcal{L}(\Omega))((\forall B \in \mathcal{B})\ Bf \in \mathcal{R} \cup \{0\} \Rightarrow f \in \mathcal{R} \cup \{0\}).$$

The rationale behind conglomerability is that if a gamble is desirable conditional on each event of a partition of Ω, then it should be unconditionally desirable. We refer to the work by Miranda et al [51] and Walley [71, Chapter 6] for a discussion of conglomerability in terms of sets of desirable gambles. See also Section 6.

Definition 13 (**Marginal extension for gambles**). Let $\mathcal{R}_\Omega$ be a marginal coherent set of gambles in $\mathcal{L}(\Omega \times \Omega')$ and $\mathcal{R}|\{\omega\}$, for all $\omega \in \Omega$, be conditional coherent sets. Let

$$\mathcal{R}|\Omega \coloneqq \{h \in \mathcal{L}(\Omega \times \Omega') : (\forall \omega \in \Omega)\ h(\omega, \cdot) \in \mathcal{R}|\{\omega\} \cup \{0\}\} \setminus \{0\}$$

be a set that conglomerates all the conditional information along the partition $\{\{\omega\} \times \Omega' : \omega \in \Omega\}$ of $\Omega \times \Omega'$ (by an abuse of notation, we denote this partition by Ω too). Then the following coherent and Ω-conglomerable set of gambles:

$$\hat{\mathcal{R}} \coloneqq \{g + h : g \in \mathcal{R}_\Omega \cup \{0\}, h \in \mathcal{R}|\Omega \cup \{0\}\} \setminus \{0\},$$

is called the *marginal extension* of the given marginal and conditional information.

The marginal extension is a generalisation of the law of total expectation to desirable gambles. It can be defined for lower previsions too (see for instance Walley [71, Section 6.7] and also Miranda and de Cooman [46]). To this end, we first give some preliminary notions.

Definition 14 (**Separately coherent conditional lower prevision**). Let $\mathcal{B}$ be a partition of Ω and $\underline{P}(\cdot|B)$ a coherent lower prevision conditional on B for all $B \in \mathcal{B}$. Then we call

$$\underline{P}(\cdot|\mathcal{B}) \coloneqq \sum_{B \in \mathcal{B}} B \underline{P}(\cdot|B)$$

a *separately coherent conditional lower prevision*.

For every gamble f, $\underline{P}(f|\mathcal{B})$ is the gamble on Ω that equals $\underline{P}(f|B)$ for $\omega \in B$; so it is a $\mathcal{B}$-measurable gamble.

[4] In the following we shall just call it conglomerability for short.

Definition 15 (**Marginal coherent lower prevision**). Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\Omega \times \Omega')$. Then the *$\Omega$-marginal coherent lower prevision* it induces is given by

$$\underline{P}_\Omega(f) \coloneqq \underline{P}(f)$$

for all $f \in \mathcal{L}(\Omega \times \Omega')$ that are Ω-measurable.

The Ω-marginal is simply the restriction of $\underline{P}$ to the subset of gambles in $\mathcal{L}(\Omega \times \Omega')$ that only depend on elements of Ω. For this reason, and analogously to the case of desirability, we can represent the Ω-marginal in an equivalent way also through the corresponding lower prevision $\underline{P}'_\Omega$ defined on $\mathcal{L}(\Omega)$.

Definition 16 (**Marginal extension for lower previsions**). Consider the possibility space $\Omega \times \Omega'$ and its partition $\{\{\omega\} \times \Omega' : \omega \in \Omega\}$. We shall denote this partition by Ω and its elements by $\{\omega\}$, with an abuse of notation. Let $\underline{P}'_\Omega$ be a marginal coherent lower prevision and let $\underline{P}(\cdot|\Omega)$ be a separately coherent conditional lower prevision on $\mathcal{L}(\Omega \times \Omega')$. Then the *marginal extension of $\underline{P}'_\Omega$ and $\underline{P}(\cdot|\Omega)$* is the lower prevision $\underline{P}$ given for all $f \in \mathcal{L}(\Omega \times \Omega')$ by

$$\underline{P}(f) \coloneqq \underline{P}'_\Omega(\underline{P}(f|\Omega)).$$

2.2. *Preference relations*

We turn now our attention to the second pillar of this paper: preference relations.

Let Ω denote, as before, the space of possibilities. In order to deal with preferences, we introduce now another set $\mathcal{X}$ of outcomes, or prizes. We assume that all the pairs of elements in $\Omega \times \mathcal{X}$ are possible or, which is equivalent, that Ω and $\mathcal{X}$ are logically independent.

The treatment of preferences in this paper relies on the notion of a conditional horse lottery:

Definition 17 (**Conditional horse lottery**). A *conditional horse lottery* is a function $p : \Omega \times \mathcal{X} \to [0, 1]$.

We should recall that the modelling of preferences is traditionally done by *horse lotteries* since the early work of Anscombe and Aumann [2]: those are similar to our conditional horse lotteries but actually correspond to placing a probability over prizes for each $\omega \in \Omega$ rather than an unnormalised function, as in Definition 17. Prominent examples of axiomatisations of incomplete preferences through horse lotteries are the works of Bewley [8], Seidenfeld et al [60], Nau [53], Galaabaatar and Karni [34], Zaffalon and Miranda [76]. Since it is not immediate to work with horse lotteries in the case of infinitely many prizes, in this paper we have decided to define conditional horse lotteries. The essence of the two approaches is the same but technically the latter appears more naturally suited for the extension to the general case. We shall analyse the differences of the two approaches in detail in Section 5.

Let us denote by $\mathcal{H}(\Omega \times \mathcal{X}) \coloneqq [0, 1]^{\Omega \times \mathcal{X}}$ the set of all conditional horse lotteries on $\Omega \times \mathcal{X}$. They will also be called *acts* for short. In the following we shall use the notation $\mathcal{H}$ for the set of all the acts in case there is no possibility of ambiguity.

Conditional horse lotteries reward us with so-called ‘probability currency’. In particular, if event ω occurs, conditional horse lottery p rewards us with an increase of probability to win prize x, for all $x \in \mathcal{X}$, proportional to $p(\omega, x)$. The reward process will be detailed in Section 4.[5]

Now it is convenient to give a name to the special act that provides us with no actual reward, no matter the ω that eventually occurs:

Definition 18 (**Zero act**). Let $0 \in \mathcal{H}$ denote the zero act, which is defined by $0(\omega, x) \coloneqq 0$ for all $\omega \in \Omega, x \in \mathcal{X}$.

Conditional horse lotteries are related to a behavioural interpretation through a notion of preference. The idea is that, since we aim at receiving a prize in $\mathcal{X}$, we will prefer some acts over some others; this will depend in part on our knowledge about the experiment originating an $\omega \in \Omega$, and in part on our attitude, or liking, towards the elements of $\mathcal{X}$. We are particularly interested in some rational type of preference relations that we call coherent:

Definition 19 (**Coherent preference relation**). A *preference relation* $\succ$ over conditional horse lotteries is a subset of $\mathcal{H} \times \mathcal{H}$. It is said to be *coherent* if it satisfies the next four axioms:

[5]Our conditional horse lotteries should not be confused with Luce and Krantz’s 1971 conditional acts. In such a case acts are conditional on subsets of Ω and are defined with the aim to increase their theory’s expressiveness. Our horse lotteries p are conditional in the sense that $p(\omega, \cdot)$ intuitively represents a collection of conditional probabilities. This is just a technical point that makes our theory simpler and easier to extend to the case of infinitely many prizes compared to the case of horse lotteries.

A1. $(\forall p \in \mathcal{H} \setminus \{0\})\ p \succ 0$ [Worst act];

A2. $(\forall p \in \mathcal{H})\ p \not\succ p$ [Irreflexivity];

A3. $(\forall p, q, r \in \mathcal{H})\ p \succ q \succ r \Rightarrow p \succ r$ [Transitivity];

A4. $(\forall p, q, r \in \mathcal{H})\ p \succ q \Leftrightarrow (\forall \alpha \in (0, 1])\ \alpha p + (1 - \alpha)r \succ \alpha q + (1 - \alpha)r$ [Mixture Independence].

If also the next axiom is satisfied, then we say that the coherent preference relation is *weakly Archimedean*.

A0. $(\forall p, q \in \mathcal{H} : \neg(p \gneq q))\ p \succ q \Rightarrow (\exists \alpha \in (0, 1))\ \alpha p \succ q$ [Weak Archimedeanity].

Let us comment on the given axioms.

- Axiom A1 is a direct consequence of the implicit assumption that the elements of $\mathcal{X}$ are prizes, that is, objects we would like to get; since the zero act surely prevents us from increasing our chances to win them—as opposed to any other act—, then it must be the worst possible act for us.
- Axioms A2 characterises the property of being 'strict' of the preferences we are defining. In particular, by writing $p \succ q$ we want to express that p is strictly more valuable than q for us. For this reason, there cannot be p such that $p \succ p$.
- As for Axiom A3, the expression $p \succ q \succ r$ means that for us p is strictly more valuable than q, which is strictly more valuable than r; if we take this to mean that we would pay a positive amount to exchange r for q, and another to exchange q for p, then we would actually pay a positive amount to exchange r for p.
- Axiom A4 can be interpreted in different ways. One way to look at it, is that what matters for establishing that $p \succ q$ is the increase, or decrease, of probabilities that comes along with dropping act q for p, and that can be expressed by the element-wise difference $p - q$. Axiom A4 will be discussed more widely in Section 5.
- Finally, Axiom A0 expresses a continuity property of the model. This is easier to see by considering a violation of the axiom, where there is a preference $p \succ q$ such that any possible worsening of p would invalidate the preference itself: that is, $\alpha p \not\succ q$ for all $\alpha \in (0, 1)$. In this case there is not continuity in the transition from $\alpha p \not\succ q$ for all $\alpha \in (0, 1)$, to $p \succ q$.

  Note that the axiom excludes from consideration the cases where $p \gneq q$. The reason is that requiring continuity in those cases would conflict with rationality. Imagine for instance the case where our preference relation is given, for all $r, s \in \mathcal{H}$, by $r \succ s$ if and only if $r \gneq s$. Consider acts p, q given by $p(\omega, x) := q(\omega, x) := \frac{1}{2}$ for all $(\omega, x) \in \Omega \times \mathcal{X}$ except for the pair (ω', x'), for which we have $p(\omega', x') > q(\omega', x')$. In this situation there is no $\alpha \in (0, 1)$ such that $\alpha p \succ q$. And yet, of course, we are fully rational in expressing $p \succ q$. Therefore continuity of preferences does not seem justified when it comes to preferences that arise from dominances like $p \gneq q$. See also Proposition 2 below.

A simple yet convenient consequence of the axioms is the following:

Proposition 1. *Suppose that for given $p, q, r, s \in \mathcal{H}$ it holds that $\alpha(p - q) = (1 - \alpha)(r - s)$ for some $\alpha \in (0, 1)$. Then for any preference relation $\succ$ that satisfies mixture independence, it holds that*

$$p \succ q \Leftrightarrow r \succ s.$$

Let us consider now a special type of preference.

Definition 20 (**Objective preference**). Given acts $p, q \in \mathcal{H}$, we say that *p is objectively preferred to q* if $p \geq q$ and $p \neq q$ (we denote objective preference also by $p \gneq q$).

The idea is that p is objectively preferred to q because the probability p assigns to outcomes is always not smaller than that of q while being strictly greater somewhere. An objective preference is indeed a preference:

Proposition 2. *Let $\succ$ be a coherent preference relation on $\mathcal{H} \times \mathcal{H}$ and $p, q \in \mathcal{H}$. Then $p \gneq q \Rightarrow p \succ q$.*

This proposition remarks in a formal way that objective preferences are those every rational subject expresses. These preferences are therefore always belonging to any coherent preference relation. If they are *the only* preferences in the relation, then we call the relation *vacuous* because we are expressing only a non-informative, or trivial, type of preferences.

3. Equivalence of desirability and preference

In the generalisation of desirability we are about to describe in this section, $\mathcal{L}$ will be redefined to be the set of gambles on $\Omega \times \mathcal{X}$. These may be considered as functional-valued gambles: for each $\omega \in \Omega$, gamble $f \in \mathcal{L}(\Omega \times \mathcal{X})$ delivers the reward function $f(\omega, \cdot)$ whose elements are measured in 'probabilistic' units associated with the respective prizes (see Section 4 for more details). $\mathcal{R}$ henceforth will denote a coherent set of desirable gambles of this form, which is a subset of the new $\mathcal{L}$. Such a generalised definition of gambles makes them very close to conditional horse lotteries.

On this basis, we shall first show that each axiom of desirability has an equivalent counterpart among the axioms of preference. From that, we will prove the equivalence of the two theories.

3.1. Axiom-to-axiom equivalence

In this section we shall start work with preferences that are not necessarily coherent, and we shall discuss then their axioms, in relations to those of desirability, one by one. To this end, we shall initially consider only a minimal requirement on preference relations:

Definition 21 (**Linear preference relation**). We say that a relation $\succ \subseteq \mathcal{H} \times \mathcal{H}$ is *linear* if $p \succ q \Rightarrow r \succ s$ for all r, s such that $p - q = r - s$.

Now we consider a special case of sets of desirable gambles that are in a bijective relation with preferences.

Lemma 3. *Let $\mathcal{L}_1(\Omega \times \mathcal{X}) := \{f \in \mathcal{L}(\Omega \times \mathcal{X}) : \sup |f| \leq 1\}$ and $\mathcal{L}_1^+ := \mathcal{L}_1 \cap \mathcal{L}^+$. Sets of desirable gambles in $\mathcal{L}_1(\Omega \times \mathcal{X})$ and linear preference relations in $\mathcal{H} \times \mathcal{H}$ are in a one-to-one correspondence, where given a set of desirable gambles $\mathcal{D}$, we consider the preference relation $\succ$ as $p \succ q \Leftrightarrow p - q \in \mathcal{D}$ and given $\succ$ we define $\mathcal{D} := \{f \in \mathcal{L}_1(\Omega \times \mathcal{X}) : (\exists p, q \in \mathcal{H})\ f = p - q, p \succ q\}$.*

Let us define coherence axioms for the special class of desirable gambles in $\mathcal{L}_1$.

Definition 22 (**Desirability axioms for linear relations**). Consider a set of gambles $\mathcal{D} \subseteq \mathcal{L}_1$. We say that:

- D0′. $\mathcal{D}$ is strictly desirable if only if for all $f \in \mathcal{D} \setminus \mathcal{L}_1^+$ and all $g \in \mathcal{L}_1$ such that $\max(0, -f) \leq g \leq \min(1, 1-f)$, there is $\alpha \in (0,1)$ such that $\alpha f - (1-\alpha) g \in \mathcal{D}$.
- D1′. $\mathcal{D}$ accepts partial gains if and only if $\mathcal{L}_1^+ \subseteq \mathcal{D}$.
- D2′. $\mathcal{D}$ avoids null gain if and only if $0 \notin \mathcal{D}$.
- D3′. $\mathcal{D}$ is additive if and only if for all $f, g \in \mathcal{D}$ such that $f = p - q$ and $g = q - r$ for some $p, q, r \in \mathcal{H}$, it holds that $f + g \in \mathcal{D}$.
- D4′. $\mathcal{D}$ satisfies positive homogeneity if and only if $f \in \mathcal{D}, \lambda > 0$ such that $\lambda f \in \mathcal{L}_1$, implies that $\lambda f \in \mathcal{D}$.

Each of these axioms is in a one-to-one correspondence with one of the axioms of preferences:

Theorem 4. *Let $\mathcal{D}$ and $\succ$ be respectively a set of gambles in $\mathcal{L}_1$ and the corresponding linear relation. Then:*

(A1⇔D1′) Relation $\succ$ has worst outcome 0 if and only if $\mathcal{D}$ accepts partial gains.

(A2⇔D2′) Relation $\succ$ is irreflexive if and only if $\mathcal{D}$ avoids null gain.

(A3⇔D3′) Relation $\succ$ is transitive if and only if $\mathcal{D}$ is additive.

(A4⇔D4′) Relation $\succ$ satisfies mixture independence if and only if $\mathcal{D}$ satisfies positive homogeneity.

(A0⇔D0′) Relation $\succ$ is weakly Archimedean if and only if $\mathcal{D}$ is strictly desirable.

3.2. *Equivalence of the theories*

At this point we start considering the case where the basic rationality axioms hold jointly.

Lemma 5. *Let $\mathcal{D}$ be a set of gambles in $\mathcal{L}_1$. If $\mathcal{D}$ satisfies* D4′ *then* D3′ *is equivalent to the following:*

D3″. *For all $f, g \in \mathcal{D}$ such that $f + g \in \mathcal{L}_1$, it holds that $f + g \in \mathcal{D}$.*

Recall that in the previous section we have shown that there is a one-to-one correspondence between linear preference relations and sets of desirable gambles in $\mathcal{L}_1$. Now we want to show what is the one-to-one correspondence between *coherent* preference relations and *coherent* sets of desirable gambles (whether in $\mathcal{L}_1$ or $\mathcal{L}$). To this end, we associate to a set of gambles $\mathcal{D} \subseteq \mathcal{L}_1$ the set

$$\mathcal{R}_{\mathcal{D}} := \{\lambda g : \lambda > 0, g \in \mathcal{D}\} \subseteq \mathcal{L}. \tag{5}$$

Vice versa, to each set of desirable gambles $\mathcal{R} \subseteq \mathcal{L}$, we associate the set

$$\mathcal{D}_{\mathcal{R}} := \left\{ \frac{f}{\sup |f|} : 0 \neq f \in \mathcal{R} \right\} \subseteq \mathcal{L}_1. \tag{6}$$

Theorem 6.

1. *Let $\mathcal{D}$ be a set of gambles in $\mathcal{L}_1$. If $\mathcal{D}$ satisfies* D1′–D4′*, then $\mathcal{R}_{\mathcal{D}}$ is coherent.*
2. *Conversely, given a coherent set of desirable gambles $\mathcal{R} \subseteq \mathcal{L}$ the set $\mathcal{D}_{\mathcal{R}}$ satisfies* D1′–D4′.
3. *Under coherence of $\mathcal{R}$ and $\mathcal{D}$, the procedures in the two statements above are each other's inverses: $\mathcal{R}_{\mathcal{D}_{\mathcal{R}}} = \mathcal{R}$ and $\mathcal{D}_{\mathcal{R}_{\mathcal{D}}} = \mathcal{D}$.*
4. *If $\mathcal{D}$ satisfies* D1′–D4′*, then $\mathcal{D}$ is coherent relative to $\mathcal{L}_1$, and $\mathcal{R}_{\mathcal{D}}$ is its natural extension.*
5. *Under coherence, $\mathcal{R}$ is a set of strictly desirable gambles if and only if $\mathcal{D}_{\mathcal{R}}$ satisfies* D0′.

Among other things, this theorem states that coherent preference relations are in a one-to-one correspondence with coherent sets of desirable gambles in $\mathcal{L}_1$, whose form is given in (6) when $\mathcal{R}$ is coherent. They are obviously also in a one-to-one correspondence with coherent sets of desirable gambles in $\mathcal{L}$ through (5).

We summarise the situation in the next theorem:

Theorem 7. *There is a one-to-one correspondence between coherent preference relations on $\mathcal{H} \times \mathcal{H}$ and coherent sets of desirable gambles in $\mathcal{L}(\Omega \times \mathcal{X})$. Moreover, strict desirability is in a one-to-one correspondence with weak Archimedeanity.*

3.3. *Other axioms considered in the literature*

Our previous results show that there is a correspondence between coherent preference relations and coherent sets of desirable gambles. In fact, this correspondence can be extended beyond the basic axioms of coherence: other interesting rationality conditions that can be imposed onto a preference relation can be formulated in terms of desirable gambles. Let us briefly discuss how this can be done.

One such axiom is called *negative transitivity*, and can be expressed as

A5. $(\forall p, q, r \in \mathcal{H})\ p \not\succ q$ and $q \not\succ r \Rightarrow p \not\succ r$ [Negative transitivity].

Negative transitivity on constant acts has been considered by Galaabaatar and Karni [34], who related it to Knightian uncertainty. In terms of sets of desirable gambles, we can prove the following:

Proposition 8. *Let $\succ$ be a coherent preference relation on $\mathcal{H} \times \mathcal{H}$, and let $\mathcal{R}$ be its associated coherent set of desirable gambles in $\mathcal{L}(\Omega \times \mathcal{X})$. Then $\succ$ is negatively transitive if and only if $\mathcal{R}$ satisfies*

$$f, g \notin \mathcal{R} \Rightarrow f + g \notin \mathcal{R}. \tag{7}$$

Interestingly, if a coherent set of desirable gambles satisfies Eq. (7) then not only $\operatorname{posi}(\mathcal{R}) = \mathcal{R}$ (due to coherence), but also $\operatorname{posi}(\mathcal{R}^c) = \mathcal{R}^c$ (due to (7)), and as a consequence it corresponds to what has been called a *lexicographic set of desirable gambles* by Van Camp et al [70]. The reason is that these sets of gambles are in a correspondence with lexicographic probabilities. See also Appendix A.2.3 for a somewhat related study of condition (7); in the framework of coherent lower previsions, we refer to the work of Miranda and Zaffalon [50].

Finally, another relevant axiom in the literature is the dominance axiom:

A6. $(\forall \omega \in \Omega)\, p \succ q^{\omega} \Rightarrow p \succ q$,

where q^{ω} is the Ω-measurable conditional horse lottery given by $q^{\omega}(\omega', x) := q(\omega, x)$ for every $(\omega', x) \in \Omega \times \mathcal{X}$. This axiom has been considered for instance by Galaabaatar and Karni [34], and it is a weaker version of Savage's postulate P7. We have shown that it is related to independent products of marginal models in the theory of coherent lower previsions, and more particularly to the *strong product* [76, Section 5.3].

Note, however, that the result does not extend immediately to the case where both $\Omega, \mathcal{X}$ are infinite; as shown by Miranda and Zaffalon [48, Example 1], independent products (and in particular the strong product) of coherent lower previsions may not even exist in that case. In fact, we find more natural to rely on a weaker, asymmetric, notion of irrelevance that always exists and that suffices to achieve the wanted decomposition of beliefs and values. We illustrate this in Appendix A following ideas from our previous work [76].

3.4. Probability and utility

We have established an equivalence result for desirability and preference at the level of the cone of desirable gambles. This allows us to focus on desirable gambles in the rest of the paper, while aiming to address questions of preference.

The result is very general since it allows us to deal also with problems that cannot be addressed with (sets of) expected utilities, that is, with all the problems that critically depend on a rigorous treatment of zero probabilities or expectations. For all the other problems, it is often convenient to work directly with a probability-utility representation. In our case, this means deducing from a coherent set of desirable gambles the corresponding lower previsions.

More precisely, given a coherent set of desirable gambles $\mathcal{R} \subseteq \mathcal{L}(\Omega \times \mathcal{X})$, we obtain its corresponding lower prevision for all $f \in \mathcal{L}(\Omega \times \mathcal{X})$ using (1):

$$\underline{P}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathcal{R}\}.$$

$\underline{P}(f)$ is interpreted as the lower expectation of f taken with respect to the probabilities and utilities that are implicit in the definition of $\mathcal{R}$. These can be made explicit very simply by applying $\underline{P}$ to some special indicator functions so as to use marginalisation (see Definition 15); we have that:

$$\underline{P}(B \times \mathcal{X}), \underline{P}(\Omega \times C)$$

are respectively the lower probability of $B \subseteq \Omega$ and the lower utility of $C \subseteq \mathcal{X}$. Note that:

- The lower probability corresponds to a set of finitely additive probabilities on Ω, as it follows from (2).
- Similarly, the lower utility corresponds to a set of finitely additive probabilities too, again using (2), which this time are interpreted as utilities. This means that a utility function in our formalism is mathematically equal to a finitely additive probability on $\mathcal{X}$. A discussion on this point can be found in our past work [76, Section 4.1].
- Separate probabilities and utilities cannot reproduce $\underline{P}$ unless this is subject to the property of 'state independence'; in the opposite condition of state dependence, the analysis needs to be done using the 'joint' model $\underline{P}$ directly. See Appendix A for details.

These considerations can be extended to the conditional case in a straightforward way. For instance, we can use (4) to compute the conditional lower prevision of a gamble $f \in \mathcal{L}(\Omega \times \mathcal{X})$ given some $B \subseteq \Omega$:

$$\underline{P}(f|B \times \mathcal{X}) := \sup\{\mu \in \mathbb{R} : (B \times \mathcal{X})(f - \mu) \in \mathcal{R}\}.$$

This can be given the updating interpretation: $\underline{P}(f|B \times \mathcal{X})$ would then represent our lower expected value of f under the assumption that B occurs and that is it the only thing we get to know about Ω. Note that $\underline{P}(\cdot|B \times \mathcal{X})$ updates both our beliefs and values in the general setting of this paper.

Note also that, mathematically speaking, nothing prevents us from computing $\underline{P}(f|\Omega \times C)$: the lower prevision of f conditional on a subset C of prizes. But we cannot give it the updating interpretation as C simply does not 'occur'. See Section 6 for more details.

4. Interpreting the rewards

We have shown that there is an equivalence between coherent preferences and coherent sets of desirable gambles. We exploit this fact in order to give a clear interpretation of the process that allows us to win prizes by using the framework of gambles.

In order to avoid confusion, let us remark that the discussion in this section should better be understood as the description of a 'hypothetical' reward process: most probably we will not want to run this process in practice, but we need the intuition it conveys in order to give meaning to rewards. And we need this meaning to be able to see which gambles are desirable to us and which are not. Stated differently, the previous sections have described the theory, while here we talk about its interpretation; we should bear in mind not to mix up these two levels.

The process by which we are rewarded is based on the idea of having compound lotteries, which underlies the definition of conditional horse lotteries.

The basic framework is constituted by a collection of simple lotteries, one per pair $(\omega, x) \in \Omega \times \mathcal{X}$, which we denote by

$$L \coloneqq \{\ell_{(\omega,x)} : (\omega, x) \in \Omega \times \mathcal{X}\}.$$

In each of these we can either win prize x or nothing. Before the process starts, we have already a $\frac{1}{2}$ chance of winning prize x, for each $x \in \mathcal{X}$ under each $\omega \in \Omega$.

Then the process goes like this:

1. We are offered a finite number of gambles $f_1, \dots, f_n \in \mathcal{L}(\Omega \times \mathcal{X})$; we select a subset of them that we accept: without loss of generality, say $f_1, \dots, f_m$, with $m \leq n$.

2. As a consequence, we are given $\varepsilon \sum_{i=1}^m f_i$ in probability currency, where $0 < \varepsilon \ll 1/\sum_{i=1}^n \sup |f_i|$ is a constant. Our new probabilities to win the prizes then become

$$\rho \coloneqq \frac{1}{2} + \varepsilon \sum_{i=1}^{m} f_i. \tag{8}$$

3. Event ω occurs. We focus on the subset of simple lotteries given by

$$L_\omega \coloneqq \{\ell_{(\omega,x)} \in L : x \in \mathcal{X}\}. \tag{9}$$

   One of the simple lotteries from L_ω is chosen by an extraneous device that is unknown to us, except for the fact that it is independent of us and of our choices, and that we have no reason to believe unfair. Such a lottery is eventually run so as we can win prize x with probability $\rho(\omega, x)$.

The most important part of this process is that the rewards are given in probability currency, because this makes sure that Axioms D1–D4 are automatically satisfied. Let us show why by following the reasoning outlined by Walley [71, Section 2.2.4]:

- Axioms D1 is a trivial case: assume that among $f_1, \dots, f_n$ there is a positive gamble $f \in \mathcal{L}^+$. We are certainly willing to accept it as by doing so we increase our chance to win the prizes.

- As for Axiom D4, assume that among $f_1, \dots, f_n$ there are gambles f and λf, with $\lambda > 0$. Assume for the moment that $\lambda \geq 1$. Compare (a) the reward $\varepsilon\lambda f$ in probability currency with (b) a reward $\varepsilon\lambda f$ in probability currency if an extraneous random event C with known positive chance α occurs (C has to be unrelated to the experiments under considerations and more in general to us). Since (b) yields zero if C does not occur, then it should be desirable to us if and only if we desire it conditional on C. In this case (a) and (b) are equivalent, thus we should desire (a) if and only if we desire (b). But (b) is equivalent to a payment in probability currency of $\alpha\varepsilon\lambda f$. This shows that $\varepsilon\lambda f$ is desirable if and only if $\alpha\varepsilon\lambda f$ is desirable for all $\alpha \in (0, 1]$. Choosing $\alpha := 1/\lambda$, we obtain that λf should be desirable if and only if f is desirable. The case where $\lambda < 1$ is analogous.
- With respect to Axiom D2, we first notice that, reasoning in an analogous way to the case of D1, we are certainly not willing to accept a gamble $f \lneq 0$. In order to enforce this constraint, we require that the zero gamble not be acceptable: in this way, in case we accepted the negative gamble f, then $-f$ would be positive and hence accepted through D1, but their sum, accepted via D4, would be zero. By making zero not acceptable, we thus reject the negative gambles. Therefore Axiom D2 follows.[6]
- With respect to D3, assume that among $f_1, \dots, f_n$ there are gambles f, g that are both desirable. Consider again an extraneous event C with probability $\frac{1}{2}$ and the compound gamble h defined as follows: if C occurs we are given f and otherwise we are given g. Given that h is desirable both if C occurs and if it does not, then h should be unconditionally desirable. But the probability currency corresponding to h is equal to $\frac{1}{2}\varepsilon f + \frac{1}{2}\varepsilon g$, whence $\frac{1}{2}f + \frac{1}{2}g$ should be desirable; applying D4 we obtain that $f + g$ is desirable.

The interpretation above is based on the actual construction of a linear utility scale through the idea of probability currency and using gambles. Using such a scale allows us to have Axioms D1–D4 automatically satisfied. As a side note, observe that all this is granted also by the requirement to use a small ε in the definition of ρ that keeps it far away from the deterministic probabilities: in fact, if we allowed ρ to get to 0 or 1, then knowing that we could win or lose a prize (almost) for sure would invalidate the linear character of our wealth as a function of gamble values—given that we have non-linear utility in the prizes in general.

The same ideas and outcomes apply to preferences, given the equivalence between the formalisms of preference and desirability. This means that such a type of rewards satisfy in particular Axioms A3 and A4. What may be worth remarking, in the case of preferences, is that conditional horse lotteries return probabilities that are only proportional to the ones used in the simple lotteries that enable us to win prizes. This is a sensible point to properly interpret horse lotteries, because one could be misled to think that those are the absolute probabilities by which prizes are won. But this view would conflict in particular with the mixture independence axiom A4, that is, with the underlying requirement of linearity. Stated differently, the long-term controversy that surrounds Axiom A4 in the literature (see Allais [1], Machina [44]), which has eventually given rise to a number of alternative formalisms that drop A4, appears to be originated by using it jointly with an interpretation of horse lotteries that is, indeed, inconsistent with it. This is not to diminish the value of the criticisms of A4 in any way, it is only to remark the importance of having a correct interpretation of a mathematical theory.

Finally, we can also sketch another interpretation of the rewards that is favoured by the use of sets of desirable gambles and that may be somewhat more realistic than that based on probability currency. The idea is that we could regard reward $f(\omega, x)$ as an amount of prize x; it can also be negative, in which case it is a loss of $-f(\omega, x)$ units of x. Similarly to the case of the probability currency, at some point there would be an extraneous device deciding which $x \in \mathcal{X}$ would be selected and we would be rewarded with $f(\omega, x)$ units of x. This kind of interpretation is well posed if and only if, for all $x \in \mathcal{X}$, our wealth is a linear function of the amounts of prize x. This is in fact the rationale behind the desirability axioms D3 and D4.

As a side note, consider that the linear character of the way we desire each prize does not prevent us from desiring some prizes more than others, in general; this is the reason why we need to distinguish the prizes and the related amounts we might get. As an example, we might imagine the prizes to be different currencies (e.g., dollars, euros, francs, etc.); if we bound the amount of money we can win or lose to some small quantity, then it is known that our appreciation of the prizes is approximately linear. At the same time we will obviously desire strong currencies more than weak ones.

[6]This is not the only way to reject the negative gambles: e.g., we could accept the zero gamble and directly require in an axiom that the negative gambles not be acceptable. However, we want to deal with strict preferences in this paper and for that we need irreflexivity, that is, D2.

5. Horse lotteries vs conditional horse lotteries

The proposed formulation of rational preferences in this paper, as detailed in Section 2.2, is slightly different from the traditional one, because the main modelling unit is that of conditional horse lotteries rather than horse lotteries. This leads, in turn, to a slightly different formulation of the rationality axioms. The aim of this section is to clarify the relation between the two formulations, by distinguishing in particular the case of finite vs infinite space of prizes.

5.1. Finite case

We show that when we restrict the attention to the case of a finite set of prizes $\mathcal{X}$, the two mentioned formalisms for preference are equivalent. Let us briefly recall them:

F_1. The first is the traditional development of decision theory, à la Anscombe-Aumann.

In this formulation, the set of prizes is defined by $\mathcal{X}_z := \mathcal{X} \cup \{z\}$, where $z \notin \mathcal{X}$ is meant to represent the null prize: that is, not receiving any prize from $\mathcal{X}$.

The basic object of the formalism is that of a *horse lottery*: it is a function $p_1 : \Omega \times \mathcal{X}_z \to [0,1]$, with the additional requirement that $\sum_{x \in \mathcal{X}_z} p_1(\omega, x) = 1$ for all $\omega \in \Omega$. In other words, a horse lottery is a collection of probability mass functions $p_1(\omega, \cdot)$, one per each element of the possibility space Ω. The set of all horse lotteries is denoted by $\mathcal{H}_z$.

As usual, one defines a preference relation $\succ$, in this case between horse lotteries. This relation is assumed to have the worst act, which, without loss of generality (from Zaffalon and Miranda [76, Section 3.1]), can be represented by the act z that is identically degenerate on outcome z: $z(\omega, z) := 1$ for all $\omega \in \Omega$. The relation is subject to coherence axioms analogous to A1–A4 so that if we let ψ denote the operator that drops all the z-elements from acts, or differences of them, we get that $\mathcal{R}_1 := \{\lambda \psi(p_1 - q_1) : \lambda > 0, p_1 \succ q_1\}$ is a coherent set of desirable gambles in $\mathcal{L}(\Omega \times \mathcal{X})$ [76, Section 3.2].

The interpretation of a horse lottery is that of a pair of nested lotteries. At the outer level there is the experiment that determines the element $\omega \in \Omega$ that occurs. This part is in common with conditional horse lotteries. The inner level is different: in the case of a horse lottery, the mass function $p_1(\omega, \cdot)$ is related to a *single* simple lottery. In order to have a consistent interpretation of horse lotteries with that illustrated in Section 4 for conditional horse lotteries, we are going to assume what follows:

- Initially we have a uniform probability over winning or losing, i.e., $\frac{1}{2}$ probability to win and $\frac{1}{2}$ to lose, that is, to receive z (note that these are the same initial probabilities we have with conditional horse lotteries). In turn, we assume that the probability to win prizes is uniform as well, and so equal to $k/2 := 1/(2|\mathcal{X}|)$.
- Such a probability changes after receiving $p_1(\omega, \cdot)$, becoming equal to

$$k\left(\frac{1}{2} + \varepsilon p_1(\omega, x)\right) \tag{10}$$

for all $x \in \mathcal{X}$, and with the constant $\varepsilon > 0$ determined so as to make the resulting numbers probabilities (and far away from the deterministic cases, similarly to the discussion in Section 4). The probability of z is determined by those of the prizes.

F_2. The second formalism is precisely that defined in Section 2.2, with the given semantics.

- According to Section 4, we assume that only one lottery in L_ω, from (9), is selected with probability $k = 1/|\mathcal{X}|$, and then run; and the chance that we win prize x through a certain conditional horse lottery p_2 is equal to

$$k\left(\frac{1}{2} + \varepsilon p_2(\omega, x)\right) \tag{11}$$

for all $x \in \mathcal{X}$, as it follows from (8). (Here we are implicitly assuming, without loss of generality, that p_2 is one of the gambles $f_1, \ldots, f_m$ in Section 4.)

Note that we are using the same symbol ε for the constants in (10) and (11). The reason is that even if those two constants were different, we could always take their minimum and use it in both equations. In fact, the absolute value of the constant is not important and we could actually work out the rest of the section with two different constants, but the development would be unnecessary complicated. For this reason, in the rest of the section we take the two constants to be equal.

In the following we use subscripts 1 and 2 for acts in order to denote the specific formalism (F_1 and F_2) an act belongs to. We keep the same symbol $\succ$, though, for the preference relations defined within the different formalisms.

It turns out, perhaps not surprisingly, that formalisms F_1 and F_2 are fully equivalent ways to deal with preferences in the finite case:

Proposition 9. *Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be the two coherent sets of desirable gambles our preferences originate respectively within F_1 and F_2. Then it follows that $\mathcal{R}_1 = \mathcal{R}_2$.*

Yet, we prefer the less traditional formalism F_2 over F_1: for one thing, the axioms it is bound to seem more direct and simple than those of F_1, and the corresponding proofs seem so as well; most importantly, the infinite case appears to be much more natural, if not just possible, to formulate in full generality using conditional horse lotteries.

5.2. *Infinite case*

We move on now to consider the case where $\mathcal{X}$ can be infinite, while imposing no restrictions on its cardinality.

5.2.1. *Formalism F_2*

We start by considering formalism F_2, that is, the one used in this paper as detailed in Section 2.2.

The essential difference that arises in F_2 compared to the case of finite $\mathcal{X}$ is concerned with the question of selecting one of the lotteries in L_ω, given that this set can be infinite and even unbounded; in other words, the problem is how we should implement step 3 of the reward process (Section 4) in the case of infinite $\mathcal{X}$. Section 4 stressed that the selection is supposed to be “fair”, because otherwise we would be induced to bias our beliefs and values in order to account for the lottery-sampling process. In the finite case, the fair selection was implemented by sampling lotteries uniformly at random. How about the infinite case?

In the infinite case, uniform selection is much more of a delicate matter. The central issue is that uniform selection cannot be done in general using σ-additive probabilities; finitely additive probabilities can be used to that end, but they are not constructible in general (see, e.g., Walley [71, Sections 2.9.5–2.9.7 and 3.5.7–3.5.9]). One could wonder, in addition, what it means in practice to sample from a finitely additive probability. Stated differently, it is clear that there is no real way in which one can actually be rewarded through (conditional) horse lotteries in the general case of infinitely many prizes in $\mathcal{X}$.

However, what really matters is the viewpoint of the subject that assesses the desirability of the gambles; for this subject there is no need to know the details of the sampling process or that it can actually be implemented; it is enough that the process be regarded as fair to enable an assessment of desirability (and if the adjective ‘fair’ entails some degree of ambiguity for the subject, then this will eventually be reflected in some additional imprecision in the subject’s set of desirable gambles, which will anyway be naturally accommodated by the theory).

In other words, the present discussion serves the purpose to delineate an ideal reward process, which however allows us to have a clear interpretation behind the formalism of preferences/desirability.

To make all this clearer, we briefly rewrite the selection process that allows us to win prizes in more formal terms. We denote by $P_0 : \mathcal{L}(\mathcal{X}) \to \mathbb{R}$ the linear prevision corresponding to a finitely additive probability that selects a lottery in L_ω uniformly at random. Consider the event $\{\text{win}\}$ that a prize is won and its complementary event $\{\text{lose}\}$, and let $\mathcal{W} := \{\text{win}, \text{lose}\}$. We denote by $P_0(\cdot|\mathcal{X}) : \mathcal{L}(\mathcal{W} \times \mathcal{X}) \to \mathbb{R}$ the separately coherent conditional linear prevision that represents, for all $x \in \mathcal{X}$, our probability of winning prize x once lottery $\ell_{(\omega,x)}$ has been selected; in practice, $P_0(\{\text{win}\}|\{x\})$ is defined by (8). The joint model for the selection of lotteries and prizes is obtained by marginal extension:[7]

$$P_0' := P_0(P_0(\cdot|\mathcal{X})).$$

P_0' is a linear prevision defined on $\mathcal{L}(\mathcal{W} \times \mathcal{X})$ that corresponds to a finitely additive probability (given that so does P_0), and that represents the mechanism by which we are eventually rewarded.

[7]Strictly speaking, P_0 should be defined on the subset of $\mathcal{X}$-measurable gambles of $\mathcal{L}(\mathcal{W} \times \mathcal{X})$. We neglect this detail to make things simpler.

5.2.2. Formalism F_1

Formalism F_1 makes things more complicated when it comes to define the mechanism to assign prizes. The reason is that horse lotteries $p \in \mathcal{H}_z$ are traditionally defined in such a way that $p(\omega, \cdot)$ is a probability over $\mathcal{X}_z$ for all $\omega \in \Omega$. This can be made to work when $\mathcal{X}_z$ is countable by means of countable additive probabilities (see, e.g., the work by Seidenfeld et al [60]), even though there are limits to it: for instance we cannot represent the situation where for a certain ω we are rewarded with an equal amount of probability currency for all prizes—given that uniform distributions on a countable space cannot be countably additive. The situation becomes even more complex when $\mathcal{X}$ is continuous.

The straightforward idea that comes to mind is then to let $p(\omega, \cdot)$ be a finitely additive probability; this would also cover the case of continuous and unbounded $\mathcal{X}_z$. But this typically implies $p(\omega, x) = 0$ for all $x \in \mathcal{X}_z$, which makes p hard to specify in a way that is clear at all for the subject that is assumed to take gambles; moreover, finitely additive probabilities are often non-constructible, so even in principle we would be prevented from clearly seeing how a certain act is made. In addition, how would gambles look like? For in the end we need to consider scaled differences of horse lotteries, such as $f := \lambda(\psi(p) - \psi(q))$; but this would be a finitely additive set-function with all the problems mentioned above (e.g., it would be identically equal to zero on the singletons). Note that restricting the attention to a bounded space $\mathcal{X}$ would partially alleviate the problem, since if we were not to use finitely additive probabilities (one might want to use density functions for instance) we would start having measurability problems; so not all gambles (or not all spaces) would be admissible for use. We prefer a theory to be free from this type of constraints in general.

So it seems to us that traditional horse lotteries are not really suited to be extended to handle infinitely many prizes in general. We find it more natural and powerful to move to conditional horse lotteries, which solve the problem right away and in full generality, and, as a bonus, make axioms as well as proofs simpler.

As a final note, consider that all the problems of F_1 are originated by the requirement that $p(\omega, \cdot)$ be a probability. But why should it be? The requirement appears to follow from the traditional interpretation that regards $p(\omega, \cdot)$ as the probabilities by which we would win prizes running the related simple lottery. But we have already argued in Section 4 that this interpretation is unfortunate, as it is inconsistent with the linearity assumption embodied by A4 in particular. A more sensible interpretation regards $p(\omega, \cdot)$ only as a relative increase of a pre-existing probability over prizes; as such, it need not be normalised, i.e., be a probability. Once we take this step, it is immediate to consider the opportunity to work with conditional horse lotteries; so a correct interpretation favours also a more natural formulation.

6. Conglomerability

Let us start by considering an uncertainty model alone, for the moment neglecting considerations of value. We recall the formulation of conglomerability given in Definition 12:

$$(\forall f \in \mathcal{L}(\Omega))((\forall B \in \mathcal{B})\ Bf \in \mathcal{R} \cup \{0\} \Rightarrow f \in \mathcal{R} \cup \{0\}).$$

Conglomerability follows from additivity (D3) in case $\mathcal{B}$ is finite; therefore it is an issue only for infinite partitions. The debate about whether or not conglomerability should be imposed in that case has not been settled in spite of the long time passed since de Finetti's seminal work [23]. The controversy seems to be kept alive by the tension originating from two opposite tendencies: on the one hand, the fact that imposing conglomerability makes things mathematically quite harder (the corresponding models being somewhat less well behaved); on the other, that the absence of conglomerability creates a number of paradoxical situations that seem to require some fixing.

In the following we address the question from a specific point of view, that of 'temporal coherence', thus without pretending to be exhaustive on the matter. For a broader discussion about conglomerability, see, e.g., Berti et al [7]. For some logical and philosophical perspectives, see Howson [39], DiBella [28].

6.1. The scope of conglomerability in the case of beliefs

In some past work we have established a clear connection between conglomerability and the question of being coherent in time [74].

Assume that we hold beliefs at present time in the form of a coherent set of desirable gambles $\mathcal{R}$. At some later time an event $B \in \mathcal{B}$ occurs; we will then hold new beliefs in the form of a coherent set of desirable gambles $\mathcal{R}^B \subseteq \mathcal{L}(B)$. $\mathcal{R}$ and $\mathcal{R}^B \subseteq \mathcal{L}(B)$ need not entertain any relation; in particular, $\mathcal{R}^B$ need not be the set of conditional beliefs that can be derived from $\mathcal{R}$ and that are denoted by $\mathcal{R}\rfloor B \subseteq \mathcal{L}(B)$ (this point appears to have been explicitly raised first by Hacking [36]). To clarify this point, let us recall the following:

Definition 23 (**Perfect information**). If $B \in \mathcal{B}$ represents all and only the information we receive about Ω from the establishment of our current beliefs $\mathcal{R}$ up to the occurrence of B, then we say we are in a case of *perfect information* (this terminology has been introduced by Shafer et al [64]).

Thanks to this definition, we can precisely recall what updated beliefs mean by the next definition:

Definition 24 (**Updating interpretation—beliefs**). Conditional beliefs $\mathcal{R}\rfloor B$ are called *updated beliefs* when they are interpreted as beliefs under the assumption that B occurs in a case of perfect information.

Note in particular that updated beliefs are still beliefs expressed at present time, so in principle they have no relation with future beliefs, those that we hold *after* the occurrence of B.

However, there are situations where $\mathcal{R}^B$ and $\mathcal{R}\rfloor B$ are related. One of these cases arises when we decide to establish our future 'beliefs' at present time (in this case, we we will talk of future *commitments* rather than beliefs). This means that we decide in advance to stick to the future behaviour that is induced by current conditional beliefs; in other words, we decide in advance that $\mathcal{R}^B := \mathcal{R}\rfloor B$ for all $B \in \mathcal{B}$.

When may this be going to be the case?

- C1. A condition that seems necessary is that we are in a case of perfect information, otherwise conditional beliefs would not represent our actual beliefs after B occurs as these would involve considerations about evidence other than B.
- C2. What appears to be left is our capability to examine the evidence at hand to the extent of establishing our current beliefs as well as we can. If we have enough time and any other needed resource to that end, we can assume that our uncertainty model is the best, in our judgement, we can produce on the basis of the given evidence.

Given these two conditions, there seems to be no reason why we might want a model for future beliefs different from current conditional ones.

But establishing future commitments at the time of present beliefs gives us the opportunity to check, and actually impose as a rationality requirement, that the two of them cohere with each other. Note that this is an additional property, required on top of the (separate) coherence of $\mathcal{R}$ and the coherence of $\mathcal{R}^B$ (for all $B \in \mathcal{B}$), that arises out of temporal considerations. We have called it *strong temporal coherence*.[8]

The main outcome of our 2013 work, in this regard, is that strong temporal coherence coincides with the conglomerability of $\mathcal{R}$.

For us, this result essentially closes the long controversy about conglomerability: conglomerability turns out to be a form of temporal coherence, which means that it is a rationality requirement when we decide in advance to establish future commitments equal to current conditional beliefs. In case we are not interested in future beliefs, for instance because we work with unconditional models only, then conglomerability is not a rationality requirement for an uncertainty model.

However, the scope of conglomerability is still quite narrow. This happens in particular because of condition C2, as it requires the availability of resources we may as well not have, first of all the time to examine all the evidence carefully. What we are going to argue now is that there is a simple, realistic, assumption that can render the scope of conglomerability much wider, by essentially removing the need of C2. It is based on the following:

Definition 25 (**Reliable assessments**). We say that an assessment is *reliable* if it is not stronger than what is justified by the evidence and the available resources, including time. Similarly, we say that the assessment process is reliable if all assessments are.

The concept of reliable assessments appears to have been introduced first by Walley [71, Sect. 6.1.2].

It is important to realise, first of all, that the assessment process can always be reliable with imprecise probability models: the reason is that in such a setting we can weaken our assessments as much as we need, up to expressing a condition of total lack of knowledge. So what Definition 25 describes is simply a process that does not yield overconfident assessments; this can always be achieved by acting on their strength. We can even weaken our assessments some extra bit in case we prefer to be very confident not to express too strong judgements (in this case, we would

[8]For a discussion of the rationality of this axiom and the connection with other, weaker, requirements, see [74, Section 6.4].

yield an *incomplete representation* of our beliefs, rather than an *exhaustive* one; see Walley [71, Section 2.10.3] for a discussion about this point).

In any case the point is that the assumption of reliability implies that we are not taking appreciable risks in the assessment process while modelling the evidence as well as we can, given the available resources. For this reason, we should definitely be willing to set future commitments equal to current conditional beliefs. But once we do so, we can argue again, as we have done in our 2013 work, that conglomerability is a consequence of strong temporal coherence. To summarise the situation, what we argue is the following:

> *Whenever we are interested in future behaviour in a setting of perfect information, and our assessments are reliable, then conglomerability is a rationality requirement.*

And since the assumption of reliability can always be met (it is just a 'good practice' in the assessment process), it makes the scope of conglomerability very wide: it is essentially extended to all cases of perfect information—which remains then the only actual constraint.

An important remark is that all this does not prevent us from behaving in the future according to a probabilistic model $\mathcal{R}'^B$ that is more precise than that characterised by our conditional beliefs, i.e., such that $\mathcal{R}\rfloor B = \mathcal{R}^B \subseteq \mathcal{R}'^B$; there is not any form of incoherence in doing so. In fact the meaning of a probabilistic model is to prescribe the limits of what we can do, but within those limits we have ample choice; we can in fact behave as we prefer as long as the model that we use to that end contains $\mathcal{R}^B$ (this constraint has to be maintained since it represents our established future commitments). This is useful and important given that there might be a certain time between the establishment of conditional beliefs and the future occurrence of B. This may make us reflect more deeply about the evidence, thanks possibly also to the availability of resources other than time (such as computational resources), and make us see that we can be somewhat bolder in our evaluations.

Finally, let us recall that so far we have been considering the case of imprecise probability. If we instead restrict the attention to the more traditional case of precise, Bayesian, probability, it turns out that the assumption of reliability is not needed: conglomerability is a rationality requirement in a temporal setting based on perfect information alone [74, Section 6.5]. The rationale is that, for a Bayesian violating reliability amounts to violating the idea of being Bayesian itself, since in that case Bayes' rule is the only temporal-coherent way to update beliefs; whence sticking to reliability can be assumed outright for a Bayesian.

6.2. *Extension to the general case of beliefs and values*

Now we turn to the general case where $\mathcal{R} \subseteq \mathcal{L}(\Omega \times \mathcal{X})$. It is important to remember, from Section 3.4 in particular, that Ω is still the only space of possibilities: the elements of $\mathcal{X}$ do not 'occur'; as a consequence it only makes sense to update $\mathcal{R}$ on subsets of Ω alone.

Having made this point, now we proceed with our focus on updating, and in particular on the definition of conglomerability extended so as to jointly handling the case of beliefs and values.

Definition 26 (**Conglomerability—extended**). Consider a coherent set of desirable gambles $\mathcal{R} \subseteq \mathcal{L}(\Omega \times \mathcal{X})$ and let $\mathcal{B}$ be a partition of Ω. $\mathcal{R}$ is said to be *$\mathcal{B}$-conglomerable* if it satisfies the following condition:

D5. $(\forall f \in \mathcal{L}(\Omega \times \mathcal{X}))((\forall B \in \mathcal{B})\ (B \times \mathcal{X})f \in \mathcal{R} \cup \{0\} \Rightarrow f \in \mathcal{R} \cup \{0\})$ [Conglomerability].

The rationale behind this definition is exactly as before in Definition 12. Moreover, the mathematical form is also as before: we are simply focusing on a special type of partitions in the enlarged space $\Omega \times \mathcal{X}$. For this reason, the conclusions of our 2013 work apply here as well, as summarised in the previous section.

The additional discussion concerned with reliability and perfect information essentially applies too; there is however a caveat in this general case, which has to do with the different nature of beliefs and values. Once we make our joint assessments of beliefs and values in a reliable way (Definition 25), we still create the conditions, as before, to stick to the future commitments $\mathcal{R}^{B \times \mathcal{X}}$ in the form of our current conditional assessments $\mathcal{R}\rfloor(B \times \mathcal{X})$. But the time that passes from the establishment of $\mathcal{R}$ to the occurrence of B may have a different effect on our values compared to our beliefs. In fact, under perfect information, it is perfectly plausible that our beliefs can only become stronger in time. This is not the case of our values, which might change in unpredictable ways due to the occurrence of events that have nothing to do with B or $\Omega \times \mathcal{X}$: we could for instance become very risk-averse as a consequence of a financial loss. For this reason, it seems necessary to require in addition what follows:

Definition 27 (**Perfect isolation**). If $B \in \mathcal{B}$ represents the only information that can affect our values in the experiment concerned with $\Omega \times \mathcal{X}$, then we say we are in a case of *perfect isolation* of values from external factors.

Note that, as a consequence, the updating interpretation should be extended to account for isolation in the general case of beliefs and values:

Definition 28 (**Updating interpretation**). Conditional assessments $\mathcal{R}\rfloor(B \times \mathcal{X})$ are called *updated assessments* when they are interpreted as assessments under the assumption that B occurs in a case of perfect information and isolation.

Having said this, the situation can finally be summarised as follows:

> *In the general case, conglomerability is a consequence of reliability, perfect information and isolation.*

Under these assumptions (and of course in a temporal setting) D5 should be added to D1–D4 as a further rationality requirement on preferences.

Note, as discussed in the previous section, that reliability can be skipped when we are Bayesian (in this case: precise beliefs and values); whence conglomerability is a consequence of perfect information and isolation only.

6.3. A final remark

We have argued that conglomerability should be a rationality requirement under very broad conditions. This follows from the identification of conglomerability with a special type of coherence: if conglomerability is not imposed, we can find situations where we are inconsistent with ourselves at the very same point in time—in spite of the fact that some of our beliefs and values are regarded as commitments for the future. This is clearly a principled support for the adoption of conglomerability.

There is, however, also a very practical support for conglomerability that comes along with the principled justification; it is the fact that conglomerability can make our inferences very strong, and hence informative, compared to the case where we do not impose it. Consider for instance Example 4 in Miranda and Zaffalon's 2015 work. There we show that a model that is vacuous between two possibilities, that is, completely uninformative about them, becomes a precise probabilistic model after it is corrected to make it conglomerable. In this case the sole adoption of conglomerability makes such a change possible. We see that conglomerability can have a very practical role to play in our models that goes well beyond its theoretical justification.

7. Conclusions

In this paper we have reconsidered our previous work on the equivalence between (incomplete) rational preferences and desirability.

We have shown that such an equivalence continues to hold when we formulate rational preferences in full generality, with unconstrained spaces of possibilities and prizes, unlike in our original work where prizes were in a finite number. Moreover, we have discussed the role of conglomerability and deduced that it should be a rationality requirement under weak assumptions, and whence it should be widely imposed in probabilistic and decision-theoretic models.

We have also provided a rigorous interpretation of gamble values based on the idea of probability currency. This allows us to make desirability live on its own, without the need to base its development on preferences, in the sense that it becomes a self-contained theory. As a consequence, one may wonder whether the founding notion for decision-theoretic modelling should be that of desirability rather than preference. In fact we are inclined to regard desirability as the primary notion, given that both the axioms and the theory itself appear to become much more direct and easy to access.

All this provides the foundations for a very general theory of uncertainty and decision making. Where can we go from here?

There are many directions that we envisage for future research. A very natural extension of the ideas presented here involves considering the recent work by Van Camp et al [69] (see also [68]). It is based on so-called *coherent choice functions*, which generalise preference modelling to non-binary comparisons of options (the original axiomatisation for the case of incomparability is due to Seidenfeld et al [63], with slightly different axioms). This gives additional modelling power compared to incomplete preferences and desirability, while including these models as particular cases [69, Section 4]. Notably, Van Camp et al [69, Section 3] have already exploited our 2017 work so as to extend choice functions to handle values besides beliefs, albeit in the case of finitely many prizes. The present paper could allow such an approach to be eventually generalised to the case of fully unconstrained spaces.

A different opportunity of research lies instead in the connection established by Benavoli et al [5] between desirability and quantum mechanics: in short, the authors prove that imposing D1–D4 on gambles made of Hermitian matrices is enough to derive quantum mechanics, or, stated differently, that quantum mechanics is nothing else than desirability in an opportune domain. Subsequent work by Benavoli et al [6] has extended those ideas relying on our 2017 work so as to derive a rational theory of decision making for quantum experiments. All this could be, again, combined with the present paper to get to a higher level of generality.

A more pragmatic, though equally important, avenue of research would be to establish clear connections between the theory presented here and all the research in AI based on preferences. It seems that such developments in AI and the theory of desirability have proceeded independently so far; linking them would provide benefits in both directions and would automatically provide a formal basis for robust methods in AI.

Most of all, however, we believe that the traditional theories of preference fall short in the realism of the problems that they are able to address. On the one hand, Anscombe and Aumann's theory is well founded if we adopt the correct interpretation, but it requires paying rewards indirectly, through the idea of probability currency, which entails a certain degree of artificiality; the same happens for the theories that build on it, like the theory we present here. Savage [59] models rewards directly, but his theory suffers from an underlying assumption of linearity, much in the spirit of Axiom A4, that clashes with the non-linearity of prize values and that eventually leads to Allais' paradox. A number of attempts have been made to correct these issues in the direction of more realistic theories (e.g., see Machina [44], Tversky and Kahneman [67], Nau [54], Cerreira-Vioglio et al [10]), but the situation still appears quite open. The challenge, in our view, would be to work out a desirability-based theory that can bypass those problems. We believe this should be possible, and we leave it for future research.

Acknowledgments

The authors are grateful to the anonymous referees for a careful reading of the paper and in particular to Referee 1 for having spotted some technical problems that are now fixed. Work partially supported by the Swiss NSF grant n. 200021_146606 / 2 and by project TIN2014-59543-P.

Appendix A. Decomposition and completeness of preferences

We now turn to the formalisms of desirability and coherent lower previsions to briefly show how they can represent and generalise well-known concepts in rational preferences. The following discussion is a summary of the one we presented in [76, Section 5], with some additional remarks and proofs for the case of infinitely many values.

Appendix A.1. In terms of sets of desirable gambles

We begin by considering the case where our assessments are modelled by sets of desirable gambles.

Appendix A.1.1. State independence

State independence is the condition that allows us to have separate models for beliefs and values.

We showed in [76, Section 5.3] that the traditional notion of state independence in the literature of preferences is equivalent to the notion of 'strong independence' in imprecise probability. In the case of totally ordered Archimedean preferences (i.e., precise probability and utility models), this means that probability and utility are stochastically independent models, when utility is mathematically represented through a probability function (like we do in this paper, see Section 3.4).

Such a notion of independence extends over multiple probability-utility pairs, and even to sets of desirable gambles. However, it ceases to exist when both Ω and $\mathcal{X}$ are infinite (Miranda and Zaffalon [48, Example 1]).[9] This holds also for most other notions of independence while, interestingly, that need not be the case for notions of 'irrelevance': these

[9] This is due to the involvement of conglomerability in the definition of independence; de Bock [15] has recently shown that some other notions of independence continue to exist by dropping conglomerability.

are asymmetric notions where mutual (symmetric) independence is replaced by independence in only one direction. In the case of this paper, for instance, this means assuming that values are independent of states but not vice versa; we say that states are irrelevant to values.

This makes of irrelevance a very natural candidate for a notion of state independence that exists no matter the cardinalities of the spaces involved. But there is more to it, as it is arguable that irrelevance qualifies itself as the right notion to use in the present context, unlike independence. The reason is that while it makes sense that states may be irrelevant to values, the opposite seems to be questionable: in fact, as we have already remarked, it appears meaningless to update a model on elements of $\mathcal{X}$, which do not occur; as a consequence the irrelevance of values to states appears meaningless too.

Let us then consider the following definition of irrelevance for desirability:

Definition 29 (**Ω-$\mathcal{X}$ irrelevant product for gambles**). A coherent set of gambles $\mathcal{R}$ on $\Omega \times \mathcal{X}$ is called an Ω-$\mathcal{X}$ *irrelevant product* of its marginal sets of gambles $\mathcal{R}'_\Omega, \mathcal{R}'_\mathcal{X}$ if it includes the set

$$\mathcal{R}|\Omega \coloneqq \{h \in \mathcal{L}(\Omega \times \mathcal{X}) : (\forall \omega \in \Omega)\ h(\omega, \cdot) \in \mathcal{R}'_\mathcal{X} \cup \{0\}\} \setminus \{0\}.$$

We have already considered these products, in another context and for lower previsions, in [48] (see also the work by de Cooman and Miranda [20], de Bock and de Cooman [16] and de Bock and de Cooman [18]).

The rationale of the definition is the following. First, the requirement that $h(\omega, \cdot) \in \mathcal{R}'_\mathcal{X} \cup \{0\}$ is there so that the inferences conditional on $\{\omega\}$ encompassed by $\mathcal{R}|\Omega$ should yield the marginal set $\mathcal{R}'_\mathcal{X}$. This is just a way to formally state that ω is irrelevant to $\mathcal{X}$. The same is repeated for every ω, so that $\mathcal{R}|\Omega$ can be regarded as being born out of aggregating all the irrelevant conditional sets.[10] Finally, that $\mathcal{R}$ contains $\mathcal{R}|\Omega$ is imposed to make sure that $\mathcal{R}$ is a model coherent with the irrelevance of beliefs on values.

If $\mathcal{R}$ satisfies Definition 29 we say it models *state-independent preferences*. The least-committal among these models is given by[11]

Proposition 10. *Given two marginal coherent sets of desirable gambles $\mathcal{R}_\Omega, \mathcal{R}_\mathcal{X}$ defined on the subsets of $\mathcal{L}(\Omega \times \mathcal{X})$ given by the Ω-measurable and $\mathcal{X}$-measurable gambles, respectively, their smallest Ω-$\mathcal{X}$ irrelevant product is given by the marginal extension of $\mathcal{R}_\Omega$ and $\mathcal{R}|\{\omega\} \coloneqq \mathcal{R}_\mathcal{X}$ for all $\omega \in \Omega$. Namely,*

$$\hat{\mathcal{R}} \coloneqq \{g + h : g \in \mathcal{R}_\Omega \cup \{0\}, h \in \mathcal{R}|\Omega \cup \{0\}\} \setminus \{0\}. \tag{A.1}$$

Definition 29 and Eq. (A.1) represent our main proposal to model state-independent preferences.

Appendix A.1.2. State dependence

We say that a set of desirable gambles $\mathcal{R}$ models state-dependent preferences when it does not satisfy the conditions in Definition 29, i.e., when it does not include the set $\mathcal{R}|\Omega$ derived from its marginal $\mathcal{R}_\mathcal{X}$ and the assumption of Ω-$\mathcal{X}$ irrelevance.

Appendix A.1.3. Completeness

So far we have discussed the general case of coherent sets of gambles $\mathcal{R}$ without discussing in particular the so-called complete or maximal coherent sets for which it holds that $f \notin \mathcal{R} \Rightarrow -f \in \mathcal{R}$ for all $f \in \mathcal{L}, f \neq 0$. Given the equivalence between coherent sets of gambles in $\mathcal{L}(\Omega \times \mathcal{X})$ and coherent preferences in $\mathcal{H} \times \mathcal{H}$, discussing the case of maximal sets amounts to discussing the case of complete preferences.

Definition 30 (**Completeness of beliefs and values for gambles**). A coherent set of desirable gambles $\mathcal{R} \subseteq \mathcal{L}(\Omega \times \mathcal{X})$ is said *to represent complete beliefs* if $\mathcal{R}_\Omega$ is maximal. It is said *to represent complete values* if $\mathcal{R}_\mathcal{X}$ is maximal. Finally, if $\mathcal{R}$ is maximal, then it is said *to represent complete preferences*.

The rationale of this definition is straightforward: when we say, for instance, that $\mathcal{R}$ represents complete beliefs, we mean that there is never indecision between two options that are concerned only with Ω. The situation is analogous in the case of complete values. Finally, in the case of complete preferences, the definition is just the direct application of the definition of maximality.

[10]Given that Ω is possibly infinite, this entails an assumption of conglomerability; without such an assumption one could obtain uninformative preferences, see [48, Section 5.1] for details.

[11]We refer to the work by Moral [52, Section 2.4]; de Cooman and Miranda [21, Theorem 13] and de Bock and de Cooman [18, Theorem 8] for somewhat related results on finite referential spaces.

Appendix A.2. In terms of lower previsions

We now consider the case where preferences are modelled by means of a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega \times \mathcal{X})$.

Appendix A.2.1. State independent preferences

We proceed as in the case of desirability by focusing on irrelevance of states to values.

Definition 31 (**Ω-$\mathcal{X}$ irrelevant product for lower previsions**). Given marginal coherent lower previsions $\underline{P}'_{\Omega}, \underline{P}'_{\mathcal{X}}$, let

$$\underline{P}(f|\{\omega\}) := \underline{P}'_{\mathcal{X}}(f(\omega, \cdot)),$$

for all $f \in \mathcal{L}(\Omega \times \mathcal{X})$, and

$$\underline{P}(\cdot|\Omega) := \sum_{\omega \in \Omega} I_{\{\omega\}} \underline{P}(\cdot|\{\omega\}).$$

Then a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega \times \mathcal{X})$ is called an *Ω-$\mathcal{X}$ irrelevant product* of $\underline{P}'_{\Omega}, \underline{P}'_{\mathcal{X}}$ if

$$\underline{P} \geq \underline{P}'_{\Omega}(\underline{P}(\cdot|\Omega)).$$

In this case we also say that $\underline{P}$ models state-independent preferences.

Here $\underline{P}(\cdot|\Omega)$ plays the role that $\mathcal{R}|\Omega$ took in Section Appendix A.1, that is, conglomerating all the conditional information. The concatenation $\underline{P}'_{\Omega}(\underline{P}(\cdot|\Omega))$ is a marginal extension: in particular, it is the least-committal coherent Ω-conglomerable model built out of the given marginals and the assessment of irrelevance that defines $\underline{P}(\cdot|\Omega)$ through $\underline{P}'_{\mathcal{X}}$. Every coherent lower prevision that dominates $\underline{P}'_{\Omega}(\underline{P}(\cdot|\Omega))$ is compatible with the irrelevance assessment but is also more informative than that:

Proposition 11. *Given two marginal coherent lower previsions $\underline{P}'_{\Omega}, \underline{P}'_{\mathcal{X}}$, their smallest Ω-$\mathcal{X}$ irrelevant product is given by the marginal extension*

$$\underline{P}'_{\Omega}(\underline{P}(\cdot|\Omega)).$$

We refer to the work of Walley [71, Section 6.7], Miranda and de Cooman [46], de Cooman and Miranda [20] and de Cooman and Hermans [19] for additional information on the marginal extension and its use in a number of different contexts.

Appendix A.2.2. State dependent preferences

As in the case of desirable gambles, we define state dependence as the lack of state independence: this means that if we consider a coherent lower prevision $\underline{P}$ with marginals $\underline{P}'_{\Omega}, \underline{P}'_{\mathcal{X}}$, it is said to model *state-dependent preferences* when it does not dominate the concatenation $\underline{P}'_{\Omega}(\underline{P}(\cdot|\Omega))$, where $\underline{P}(\cdot|\Omega)$ is derived from $\underline{P}'_{\mathcal{X}}$ by an assumption of epistemic irrelevance.

Appendix A.2.3. Completeness

The definition of completeness for lower previsions is a rephrasing of that for desirable gambles:

Definition 32 (**Completeness of beliefs and values for lower previsions**). A coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega \times \mathcal{X})$ is said *to represent complete beliefs* if its marginal $\underline{P}_{\Omega}$ is linear. It is said *to represent complete values* if its marginal $\underline{P}_{\mathcal{X}}$ is linear. Finally, if $\underline{P}$ is linear, then it is said *to represent complete preferences*.

There are a number of equivalent ways in which we can characterise the linearity of previsions in terms of the corresponding set of desirable gambles:

Proposition 12. *Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}$ and $\mathcal{R}$ its corresponding coherent set of strictly desirable gambles. The following are equivalent:*

(i) *$\underline{P}$ is a linear prevision.*

(ii) *If $f \notin \mathcal{R}$ then $\varepsilon - f \in \mathcal{R}$ for all $\varepsilon > 0$.*

(iii) *$\mathcal{R}$ is negatively additive, meaning that $f, g \notin \mathcal{R} \Rightarrow (\forall \varepsilon > 0)\ f + g - \varepsilon \notin \mathcal{R}$.*

We can then focus on any of those formulations to immediately deduce characterisations for all the cases of completeness in preferences:

Proposition 13. *Consider a coherent preference relation $\succ$ on gambles represented by a coherent lower prevision $\underline{P}$ on $\mathcal{L}(\Omega \times \mathcal{X})$, whose corresponding coherent set of gambles is denoted by $\mathcal{R}$. Let $\mathcal{R}_\Omega, \mathcal{R}_\mathcal{X}$ denote its marginals. Then:*

(i) *$\underline{P}$ represents complete beliefs $\Leftrightarrow$ $\mathcal{R}_\Omega$ is negatively additive.*

(ii) *$\underline{P}$ represents complete values $\Leftrightarrow$ $\mathcal{R}_\mathcal{X}$ is negatively additive.*

(iii) *$\underline{P}$ represents complete preferences $\Leftrightarrow$ $\mathcal{R}$ is negatively additive.*

Appendix B. Proofs

Proof of Proposition 1. Using the mixture-independence axiom A4 together with the assumption,

$$p \succ q \Leftrightarrow \alpha p + (1-\alpha) r \succ \alpha q + (1-\alpha) r = \alpha p + (1-\alpha) s \Leftrightarrow r \succ s.$$

□

Proof of Proposition 2. By $p \gneq q$ we deduce that $0 \lneq p - q \leq 1$. Then $(p-q) \in \mathcal{H}$ and $(p-q) \succ 0$ by A1. And since

$$\frac{1}{2}(p-q) = \frac{1}{2}[(p-q) - 0],$$

we deduce by Proposition 1 that $p \succ q$. □

Proof of Lemma 3. Consider a set of desirable gambles $\mathcal{D} \subseteq \mathcal{L}_1(\Omega \times \mathcal{X})$ and define the linear relation $\succ \subseteq \mathcal{H} \times \mathcal{H}$ by $p \succ q$ if and only if $p - q \in \mathcal{D}$.

Observe also that for any $f \in \mathcal{D}$ we can always find $p, q \in \mathcal{H}$ such that $f = p - q$: it is sufficient to let $p := \max(f, 0)$ and $q := \min(0, f)$.

Let us now prove injectivity. Assume by contradiction that there are sets $\mathcal{D}_1 \neq \mathcal{D}_2$ in $\mathcal{L}_1$ such that the respective linear relations are equal: $\succ_1 = \succ_2$. Since the sets are different, then we can assume without loss of generality that there is $f_1 \in \mathcal{D}_1 \setminus \mathcal{D}_2$. At the same time for all $p, q \in \mathcal{H}$ with $p - q = f_1$, we have that $p \succ_1 q$ and $p \succ_2 q$. This implies that there is $f_2 \in \mathcal{D}_2$ such that $f_2 = p - q$. But then $f_1 = f_2$, a contradiction.

Let us focus on surjectivity. Consider a linear relation $\succ \subseteq \mathcal{H} \times \mathcal{H}$ and let $\mathcal{D}' := \{f \in \mathcal{L}_1(\Omega \times \mathcal{X}) : (\exists p, q \in \mathcal{H})\ f = p - q, p \succ q\}$. The relation it induces is given by $p \succ' q \Leftrightarrow p - q \in \mathcal{D}'$. Now, if $p \succ q$, then $p - q \in \mathcal{D}'$, whence $p \succ' q$. Conversely, if $p \succ' q$, then $p - q \in \mathcal{D}'$; as a consequence, there are $p', q' \in \mathcal{H}$, $p' \succ q'$ such that $p - q = p' - q'$. But relation $\succ$ is linear, whence $p \succ q$. It follows that $\succ = \succ'$. □

Proof of Theorem 4. Consider a set of desirable gambles $\mathcal{D} \subseteq \mathcal{L}_1$ and its corresponding linear relation $\succ$. They are related through

$$\mathcal{D} = \{f \in \mathcal{L}_1(\Omega \times \mathcal{X}) : (\exists p, q \in \mathcal{H})\ f = p - q, p \succ q\},$$
$$p \succ q \Leftrightarrow p - q \in \mathcal{D}.$$

(A1$\Leftrightarrow$D1′) If $p \succ 0$ for all $p \in \mathcal{H}, p \neq 0$, then $p \in \mathcal{D}$; and since $\mathcal{L}_1^+ = \mathcal{H} \setminus \{0\}$, then we have the thesis. Conversely, if $\mathcal{L}_1^+ \subseteq \mathcal{D}$, then $p \in \mathcal{D}$ for all $p \in \mathcal{H} \setminus \{0\}$; as a consequence $p \succ 0$ for all $p \in \mathcal{H}, p \neq 0$.

(A2$\Leftrightarrow$D2′) By contradiction, if $p \succ p$ for some $p \in \mathcal{H}$, then $0 \in \mathcal{D}$; and if $0 \in \mathcal{D}$, then for some $p \in \mathcal{D}$, $p \succ p$.

(A3$\Leftrightarrow$D3′) Assume that the relation is transitive and take $f, g \in \mathcal{D}$ such that $f = p - q, g = q - r$ for some $p, q, r \in \mathcal{H}$. Then we have that $p \succ q$ and $q \succ r$; by transitivity we conclude that $p \succ r$, whence $f + g \in \mathcal{D}$.

Conversely, assume that $\mathcal{D}$ is additive and take $p \succ q, q \succ r$. Letting $f := p - q, g := q - r$, we get by additivity that $f + g = p - r \in \mathcal{D}$, so that $p \succ r$.

(A4⇔D4′) Assume that $\succ$ satisfies mixture independence and take $f \in \mathcal{D}$ and $\lambda > 0$ such that $\lambda f \in \mathcal{L}_1$. Then there are $p \succ q$ such that $f = p - q$. Moreover, there must be $r, s \in \mathcal{H}$ such that $\lambda f = r - s$ given that $\lambda f \leq 1$ (it is enough as usual to take the positive and negative parts of λf, respectively). Using Proposition 1, we get that $r \succ s$, whence $\lambda f \in \mathcal{D}$.

Conversely, assume that $f \in \mathcal{D}$ implies that $\lambda f \in \mathcal{D}$ for all $\lambda > 0$ such that $\lambda f \in \mathcal{L}_1$. Taking $p \succ q$, we have then that $\lambda p + (1-\lambda)r \succ \lambda q + (1-\lambda)r$ for all $p, q, r \in \mathcal{H}$ and $\lambda \in (0, 1]$. On the other hand, if $\lambda p + (1-\lambda)r \succ \lambda q + (1-\lambda)r$ for all $p, q, r \in \mathcal{H}$ and $\lambda \in (0, 1]$, then trivially $p \succ q$.

(A0⇔D0′) Assume that for all $p \succ q$, with $p \gneq q$ not holding, there is $\alpha \in (0, 1)$ such that $\alpha p \succ q$. We want to show that for all $f \in \mathcal{D} \setminus \mathcal{L}_1^+$ and all $g \in \mathcal{L}_1$ such that $\max(0, -f) \leq g \leq \min(1, 1-f)$, there is $\alpha \in (0, 1)$ such that $\alpha f - (1-\alpha)g \in \mathcal{D}$.

Let $q := g$ and $p := f + q$, so that $f = p - q$. Let us show that $q \in \mathcal{H}$; it is enough to prove that $\max(0, -f) \leq \min(1, 1-f)$. In case $f(\omega, x) \geq 0$, the inequality reduces itself to $0 \leq 1 - f(\omega, x)$, which is true since $f(\omega, x) \leq 1$. In case $f(\omega, x) < 0$, the inequality becomes $-f(\omega, x) \leq 1$, which is again true since $f(\omega, x) \geq -1$.

Let us similarly show that $p = f + q \in \mathcal{H}$ too. We know that $f + \max(0, -f) \leq p \leq f + \min(1, 1-f)$. Again, if $f(\omega, x) \geq 0$, we see that $0 \leq f(\omega, x) \leq p(\omega, x) \leq 1$; if $f(\omega, x) < 0$, we have that $0 \leq p(\omega, x) \leq f + 1 \leq 1$.

At this point we can apply the hypothesis, recalling that $f \notin \mathcal{L}_1^+$ and hence that $p \gneq q$ does not hold, and so we deduce that there is $\alpha \in (0, 1)$ such that $\alpha p - q = \alpha(p - q) - (1-\alpha)q = \alpha f - (1-\alpha)q = \alpha f - (1-\alpha)g \in \mathcal{D}$.

Conversely, suppose that for all $f \in \mathcal{D} \setminus \mathcal{L}_1^+$ and all $g \in \mathcal{L}_1$ such that $\max(0, -f) \leq g \leq \min(1, 1-f)$, there is $\alpha \in (0, 1)$ such that $\alpha f - (1-\alpha)g \in \mathcal{D}$. Let us show that $p \succ q$, with $p \gneq q$ not holding, implies that there is $\alpha \in (0, 1)$ such that $\alpha p \succ q$.

Let $f := p - q$, from which it follows that $f \notin \mathcal{L}_1^+$, and $g := q$. Observe that if $\max(0, -f) \leq g \leq \min(1, 1-f)$ holds, then from the assumption the thesis follows immediately. In other words, we have to show that $\max(0, q - p) \leq q \leq \min(1, q + (1-p))$. This is immediate since on the left-hand side we decrease q by some non-negative amount while on the right-hand side we are increasing it by a non-negative amount.

□

Proof of Lemma 5. That D3″ implies D3′ is trivial. Therefore let us assume that $\mathcal{D}$ satisfies D3′ and prove that D3″ holds.

Consider $f, g \in \mathcal{D}$ such that $f + g \in \mathcal{L}_1$. By D4′ we have that $\frac{1}{2}f, \frac{1}{2}g \in \mathcal{D}$. We can always write $f = p - q, g = r - s$ for some $p, q, r, s \in \mathcal{H}$. It follows that

$$\left(\frac{1}{2}p + \frac{1}{2}r\right) - \left(\frac{1}{2}q + \frac{1}{2}r\right) \in \mathcal{D},$$
$$\left(\frac{1}{2}q + \frac{1}{2}r\right) - \left(\frac{1}{2}q + \frac{1}{2}s\right) \in \mathcal{D}.$$

Using D3′ we obtain that $(\frac{1}{2}p + \frac{1}{2}r) - (\frac{1}{2}q + \frac{1}{2}s) = \frac{1}{2}f + \frac{1}{2}g \in \mathcal{D}$. Using D4′ again, we have the thesis. □

Proof of Theorem 6. Let us prove the five points of the theorem.

1. Assume that $\mathcal{D}$ satisfies D1′–D4′ and show that $\mathcal{R}_{\mathcal{D}}$ is coherent.

   D1. Consider $f \in \mathcal{L}^+$ and let $g := f / \sup|f|$, so that $g \in \mathcal{L}_1^+$. As a consequence of D1′, $g \in \mathcal{D}$, whence $f \in \mathcal{R}_{\mathcal{D}}$.

   D2. Trivial.

   D3. Take $f, g \in \mathcal{R}_{\mathcal{D}}$, so that there are $f', g' \in \mathcal{D}, \lambda', \lambda'' > 0$ with $f = \lambda' f', g = \lambda'' g'$. If we let $\lambda := \lambda' + \lambda''$, we have that $f/(2\lambda), g/(2\lambda) \in \mathcal{D}$ by D4′ and at the same time $(f+g)/(2\lambda) \in \mathcal{L}_1$, so that $(f+g)/(2\lambda) \in \mathcal{D}$ by D3″. As a consequence, $f + g \in \mathcal{R}_{\mathcal{D}}$.

   D4. Trivial.

2. That $\mathcal{D}_{\mathcal{R}}$ satisfies D1′–D4′ once $\mathcal{R}$ is coherent is straightforward.

3. We need to show that, under coherence of $\mathcal{R}, \mathcal{D}$, it holds that $\mathcal{R}_{\mathcal{D}_{\mathcal{R}}} = \mathcal{R}$ and $\mathcal{D}_{\mathcal{R}_{\mathcal{D}}} = \mathcal{D}$. For the first equality, that $\mathcal{R}_{\mathcal{D}_{\mathcal{R}}} \supseteq \mathcal{R}$ follows from

$$\mathcal{R}_{\mathcal{D}_{\mathcal{R}}} = \{\lambda g : \lambda > 0, g \in \mathcal{D}_{\mathcal{R}}\} = \left\{\lambda \frac{g}{\sup|g|} : \lambda > 0, g \in \mathcal{R}\right\} \supseteq \mathcal{R}$$

   by taking $\lambda := \sup|g|$.

   To show that $\mathcal{R}_{\mathcal{D}_{\mathcal{R}}} \subseteq \mathcal{R}$, consider any h in $\mathcal{R}$, implying that $\frac{h}{\sup|h|} \in \mathcal{D}_{\mathcal{R}}$. Note that $\sup|h| > 0$; indeed, if $\sup|h|$ would be 0 then $h = 0$. For any given $\lambda > 0$, this implies that $\lambda\frac{h}{\sup|h|} = \lambda' h$ for $\lambda' := \frac{\lambda}{\sup|h|} > 0$ indeed belongs to $\mathcal{R}$, because this set satisfies D4.

   Let us prove now the equality $\mathcal{D}_{\mathcal{R}_{\mathcal{D}}} = \mathcal{D}$. Take $\mathcal{D}$ satisfying D1′–D4′. To see that $\mathcal{D}_{\mathcal{R}_{\mathcal{D}}} \subseteq \mathcal{D}$, take $f \in \mathcal{D}_{\mathcal{R}_{\mathcal{D}}}$. Then there is some $g \in \mathcal{R}_{\mathcal{D}}$ such that $f = \frac{g}{\sup|g|}$. Since $g \in \mathcal{R}_{\mathcal{D}}$, there is some $h \in \mathcal{D}$ and some $\lambda > 0$ such that $g = \lambda h$, whence $f = \frac{\lambda h}{\sup|\lambda h|} = \frac{h}{\sup|h|}$. Since $f \in \mathcal{L}_1$, we deduce from D4′ that $f \in \mathcal{D}$.

   Conversely, take $f \in \mathcal{D}$. Then $f \in \mathcal{R}_{\mathcal{D}}$ and as a consequence $\frac{f}{\sup|f|} \in \mathcal{D}_{\mathcal{R}_{\mathcal{D}}}$. Since $\mathcal{D}_{\mathcal{R}_{\mathcal{D}}}$ satisfies D4′ and $f \in \mathcal{L}_1$, we deduce that $f \in \mathcal{D}_{\mathcal{R}_{\mathcal{D}}}$.

4. Since $\mathcal{D} \subseteq \mathcal{R}_{\mathcal{D}}$ and $\mathcal{R}_{\mathcal{D}}$ is coherent (apply 1), then $\mathcal{D}$ avoids partial loss by [47, Proposition 3(e)]. This implies that the minimal coherent extension of $\mathcal{D}$ exists. Let us call it $\mathcal{E}$. That $\mathcal{R}_{\mathcal{D}} \subseteq \mathcal{E}$ is trivial. Since $\mathcal{R}_{\mathcal{D}}$ is coherent, then by definition of $\mathcal{E}$, it must be that $\mathcal{R}_{\mathcal{D}} \supseteq \mathcal{E}$. As a consequence $\mathcal{R}_{\mathcal{D}}$ is the natural extension.

   Let us show that $\mathcal{R}_{\mathcal{D}} \cap \mathcal{L}_1 \subseteq \mathcal{D}$ (the converse inclusion is trivial). Consider $f \in \mathcal{R}_{\mathcal{D}} \cap \mathcal{L}_1$. Then $f = \lambda g$ for some $\lambda > 0, g \in \mathcal{D}$. And since $\lambda g \in \mathcal{L}_1$, we get by D4′ that $f = \lambda g \in \mathcal{D}$. This shows that $\mathcal{D}$ is coherent relative to $\mathcal{L}_1$.

5. ($\Rightarrow$) Showing that $\mathcal{D}_{\mathcal{R}}$ satisfies D0′ is equivalent to showing that the preference relation it induces is weakly Archimedean, thanks to Theorem 4. Therefore consider $g \in \mathcal{D}_{\mathcal{R}} \setminus \mathcal{L}_1^+$ such that $g = p - q$ for some $p, q \in \mathcal{H}$. We want to show that there is $\alpha \in (0,1)$ such that $\alpha p - q \in \mathcal{D}_{\mathcal{R}}$.

   Thanks to point 4, we have that $g \in \mathcal{R} \setminus \mathcal{L}^+$ and hence by strict desirability that there is $\delta > 0$ such that $g - \delta \in \mathcal{R}$. Choose $\alpha \in (0,1)$ so that $(1-\alpha)p \leq \delta$. Then $\alpha p - q \geq p - q - \delta = g - \delta$, whence $\alpha p - q \in \mathcal{D}_{\mathcal{R}}$.

   ($\Leftarrow$) Finally, let us assume that $\mathcal{D}_{\mathcal{R}}$ satisfies D0′–D4′ and show that for all $h \in \mathcal{R} \setminus \mathcal{L}^+$ and $k \in \mathcal{L}$, there is $\alpha \in (0,1)$ such that $\alpha h + (1-\alpha)k \in \mathcal{R}$. If we prove this, then it is enough to take a constant $k < 0$ to get that $h - \delta \in \mathcal{R}$, with $\delta := -\frac{1-\alpha}{\alpha}k > 0$. This means that $\mathcal{R}$ would be strictly desirable.

   Now then consider the positive and negative parts of k, namely, k^+, k^-, such that $k = k^+ - k^-$. Note that for the thesis it is enough to show that there is $\alpha \in (0,1)$ such that $\alpha\frac{h}{2} - (1-\alpha)k^- \in \mathcal{R}$, given that $\alpha\frac{h}{2} + (1-\alpha)k^+$ belongs to $\mathcal{R}$ already, whence by additivity the result follows.

   To that end, choose $\lambda > 0$ such that $f := \frac{h}{2\lambda}$ and $g := \frac{k^-}{\lambda}$ satisfy the inequalities in D0′. Then there is $\alpha \in (0,1)$ such that $\alpha\frac{h}{2\lambda} - (1-\alpha)\frac{k^-}{\lambda} \in \mathcal{D}_{\mathcal{R}}$; and as a consequence, $\alpha\frac{h}{2} - (1-\alpha)k^- \in \mathcal{R}$.

□

Proof of Theorem 7. Thanks to Theorems 4 and 6, we have an equivalence between coherent preference relations $\succ$ and sets of desirable gambles coherent relative to $\mathcal{L}_1$; the latter are in a one-to-one correspondence with their natural extensions, from which we deduce the first part of the theorem. The second part is granted again by Theorems 4 and 6. □

Proof of Proposition 8. Assume first of all that $\succ$ is negatively transitive, and let $f, g \notin \mathcal{R}$. Then $f \not\succ 0$ and $g \not\succ 0$, from which it follows that $0 \not\succ -g$. Applying the negative transitivity of $\succ$, we deduce that $f \not\succ -g$, or, equivalently, that $f + g \not\succ 0$, whence $f + g \notin \mathcal{R}$.

Conversely, assume that Eq. (7) holds and let $p, q, r \in \mathcal{H}$ satisfy $p \not\succ q$ and $q \not\succ r$. Then $p - q \notin \mathcal{R}$ and $q - r \notin \mathcal{R}$, whence $p - r \notin \mathcal{R}$, and using the correspondence between $\succ$ and $\mathcal{R}$ we conclude that $p \not\succ r$. □

Proof of Proposition 9. From (10) and (11), we see that simply defining

$$p_2 := \psi(p_1) \tag{B.1}$$

for all horse lotteries $p_1 \in \mathcal{H}_z$, we obtain a correspondence between $\mathcal{H}$ and $\mathcal{H}_z$ made of elements that are equally valuable for us: their respective probability currencies are just the same. Moreover, the correspondence is one-to-one as it follows from [76, Remark 5 in p. 1098] (note also that ψ is a linear operator).

It is now a simple step to prove the two inclusions.

$\mathcal{R}_2 \subseteq \mathcal{R}_1$ If $\lambda(p_2 - q_2) \in \mathcal{R}_2$ for some $p_2, q_2 \in \mathcal{H}$ and $\lambda > 0$, then also $\lambda\psi(p_1 - q_1) \in \mathcal{R}_1$, where p_1, q_1 are obtained through (B.1). This follows from the identity $p_2 - q_2 = \psi(p_1 - q_1)$, as obtained from linearity of ψ, and the fact that the lotteries defined through (B.1) are equally valuable.

$\mathcal{R}_1 \subseteq \mathcal{R}_2$ Similarly, consider any $\lambda\psi(p_1 - q_1) \in \mathcal{R}_1$ and let $p_2 := \psi(p_1), q_2 := \psi(q_1)$, so that $p_2, q_2 \in \mathcal{H}$. It follows that $\lambda(p_2 - q_2) \in \mathcal{R}_2$ for the same reasons given in the proof of the converse inclusion.

□

Proof of Proposition 10. It follows from [51, Proposition 29] that $\hat{\mathcal{R}}$ is the smallest conglomerable and coherent set of gambles that includes $\mathcal{R}_\Omega$ and $\mathcal{R}|\{\omega\}$ (for all $\omega \in \Omega$). It includes $\mathcal{R}|\Omega$ by construction. We are left to show that $\hat{\mathcal{R}}$ induces both $\mathcal{R}_\Omega$ and $\mathcal{R}_\mathcal{X}$.

We begin by proving that the Ω-marginal of $\hat{\mathcal{R}}$ is $\mathcal{R}_\Omega$. Consider an Ω-measurable gamble $f \in \hat{\mathcal{R}}$. Then there are $g \in \mathcal{R}_\Omega$ and $h \in \mathcal{R}|\Omega$ such that $f \geq g + h$. For any $\omega \in \Omega$, it holds that

$$0 \leq \sup_{x \in \mathcal{X}} h(\omega, x) \leq \sup_{x \in \mathcal{X}} (f(\omega, x) - g(\omega, x)) = f'(\omega) - g'(\omega),$$

where in last equality we are using that both f, g are Ω-measurable, and are denoting by f', g' their equivalent representations as gambles on Ω. Thus, $f \geq g$, and since $\mathcal{R}_\Omega$ is a coherent set of gambles we conclude that also $f \in \mathcal{R}_\Omega$. The converse inclusion follows from Eq. (A.1).

Next, consider an $\mathcal{X}$-measurable gamble $f \in \hat{\mathcal{R}}$. Then there are $g \in \mathcal{R}_\Omega$ and $h \in \mathcal{R}|\Omega$ such that $f \geq g + h$. If $g \neq 0$, then there exists $\omega \in \Omega$ such that $g'(\omega) > 0$. Then,

$$fI_{\{\omega\}} \geq g'(\omega) + hI_{\{\omega\}} \geq hI_{\{\omega\}},$$

and since $h(\omega, \cdot) \in \mathcal{R}_\mathcal{X} \cup \{0\}$, either $h(\omega, \cdot) \in \mathcal{R}_\mathcal{X}$, in which case $f(\omega, \cdot) \in \mathcal{R}_\mathcal{X}$, or $h(\omega, \cdot) = 0$, whence $f(\omega, \cdot) \geq g'(\omega) > 0$, meaning that also $f(\omega, \cdot) \in \mathcal{R}_\mathcal{X}$. Finally, the proof when $g = 0$ is similar (in that case we can pick any $\omega \in \Omega$ such that $hI_{\{\omega\}} \neq 0$, since one is bound to exist). Again, the converse inclusion follows from Eq. (A.1). □

Proof of Proposition 11. The result follows immediately from Definition 31, taking into account that $\underline{P}'_\Omega(\underline{P}(\cdot|\Omega))$ is a coherent lower prevision by [71, Section 6.7.2]. □

Proof of Proposition 12. We make a circular proof.

$(i) \Rightarrow (ii)$ It $f \notin \mathcal{R}$, it follows from Eq. (3) that $\underline{P}(f) = \overline{P}(f) \leq 0$, whence $\underline{P}(-f) = -\overline{P}(f) \geq 0$, and therefore $\underline{P}(\varepsilon - f) > 0$ for every $\varepsilon > 0$. Thus, Eq. (3) implies that $\varepsilon - f \in \mathcal{R}$.

$(ii) \Rightarrow (iii)$ Assume ex-absurdo the existence of $f, g \notin \mathcal{R}$ such that $f + g - \varepsilon \in \mathcal{R}$. It follows from (ii) that $\frac{\varepsilon}{4} - f, \frac{\varepsilon}{4} - g \in \mathcal{R}$, and applying D3 we conclude that $-\frac{\varepsilon}{2} \in \mathcal{R}$, a contradiction with D2.

$(iii) \Rightarrow (i)$ If $\underline{P}$ is not linear, we can find a gamble f such that $\underline{P}(f) < \overline{P}(f)$. The conjugacy of the lower and upper previsions and Eq. (1) implies that for every $\varepsilon > 0$ the gambles $f - \underline{P}(f) - \varepsilon$ and $\overline{P}(f) - f - \varepsilon$ do not belong to $\mathcal{R}$, and (iii) implies then that for every $\delta > 0$ the gamble $(f - \underline{P}(f) - \varepsilon) + (\overline{P}(f) - f - \varepsilon) - \delta$ does not belong to $\mathcal{R}$. But for ε, δ satisfying $2\varepsilon + \delta < \overline{P}(f) - \underline{P}(f)$ this sum is a non-negative gamble. This is a contradiction with D1.

□

Proof of Proposition 13. This is an immediate consequence of Proposition 12. □

References

[1] Allais M (1953) Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. Econometrica 21(4):503–546
[2] Anscombe FJ, Aumann RJ (1963) A definition of subjective probability. The Annals of Mathematical Statistics 34:199–2005
[3] Armstrong T (1990) Conglomerability of probability measures on Boolean algebras. Journal of Mathematical Analysis and Applications 150(2):335–358
[4] Augustin T, Coolen F, de Cooman G, Troffaes M (eds) (2014) Introduction to Imprecise Probabilities. Wiley
[5] Benavoli A, Facchini A, Zaffalon M (2016) Quantum mechanics: the Bayesian theory generalized to the space of Hermitian matrices. Physical Review A 94(042106)
[6] Benavoli A, Facchini A, Zaffalon M (2016) Quantum rational preferences and desirability. In: Proceedings of The 1st International Workshop on "Imperfect Decision Makers: Admitting Real-World Rationality" (NIPS 2016), JMLR.org, PMLR, vol 58, pp 87–96
[7] Berti P, Miranda E, Rigo P (2017) Basic ideas underlying conglomerability and disintegrability. International Journal of Approximate Reasoning 88:387–400
[8] Bewley T (2002) Knightian decision theory: part I. Decisions in economics and finance 25:79–110, reprint of Discussion Paper 807, Cowles Foundation, 1986
[9] Carnap R (1980) The problem of a more general concept of regularity, vol 2, University of Califormia Press, pp 145–155
[10] Cerreira-Vioglio S, Maccheroni F, Marinacci M (2016) Stochastic dominance analysis without the independence axiom. Management Science 63(4):1097–1109
[11] Coletti G, Scozzafava R (2002) Probabilistic logic in a coherent setting. Kluwer
[12] Corani G, Zaffalon M (2008) Learning reliable classifiers from small or incomplete data sets: the naive credal classifier 2. Journal of Machine Learning Research 9:581–621
[13] Couso I, Moral S (2011) Sets of desirable gambles: Conditioning, representation, and precise probabilities. International Journal of Approximate Reasoning 52:1034–1055
[14] Cozman FG (2012) Sets of probability distributions, independence, and convexity. Synthese 186(2):577–600
[15] De Bock J (2017) Independent natural extension for sets of desirable gambles: Williams-coherence to the rescue. In: Antonucci A, Corani G, Couso I, Destercke S (eds) ISIPTA'17: Proceedings of the 10th International Symposium on Imprecise Probability: Theories and Applications. PMLR, vol 62, pp 121–132
[16] De Bock J, de Cooman G (2014) An efficient algorithm for estimating state sequences in imprecise hidden Markov models. Journal of Artificial Intelligence Research 50:189–233
[17] De Bock J, de Cooman G (2015) Conditioning, updating and lower probability zero. International Journal of Approximate Reasoning 67:1–36
[18] De Bock J, de Cooman G (2015) Credal networks under epistemic irrelevance: The sets of desirable gambles approach. International Journal of Approximate Reasoning 56:178–207
[19] De Cooman G, Hermans F (2008) Imprecise probability trees: bridging two theories of imprecise probability. Artificial Intelligence 172(11):1400–1427
[20] De Cooman G, Miranda E (2009) Forward irrelevance. Journal of Statistical Planning and Inference 139(2):256–276
[21] De Cooman G, Miranda E (2012) Irrelevance and independence for sets of desirable gambles. Journal of Artificial Intelligence Research 45:601–640
[22] De Cooman G, Quaeghebeur E (2012) Exchangeability and sets of desirable gambles. International Journal of Approximate Reasoning 53:363–395, special issue in honour of Henry E. Kyburg, Jr.
[23] De Finetti B (1931) Sul significato soggettivo della probabilità. Fundamenta Mathematicae 17:298–329
[24] De Finetti B (1937) La prévision: ses lois logiques, ses sources subjectives. Annales de l'Institut Henri Poincaré 7:1–68, English translation in [42]
[25] De Finetti B (1970) Teoria delle Probabilità. Einaudi, Turin
[26] De Finetti B (1974) Theory of Probability: A Critical Introductory Treatment, vol 1. John Wiley & Sons, Chichester, English translation of [25]
[27] Del Amo AG, Ríos Insua D (2002) A note on an open problem in the foundations of statistics. RACSAM 96(1):55–61
[28] DiBella N (2018) The qualitative paradox of non-conglomerability. Synthese 195(3):1181–1210
[29] Dietterich TG (2017) Steps toward robust artificial intelligence. AI Magazine 38(3):3–24
[30] Domshlak C, Hüllermeier E, Kaci S, Prade H (2011) Preferences in AI: an overview. Artificial Intelligence 175(7-8):1037–1052
[31] Dubins LE (1975) Finitely additive conditional probabilities, conglomerability and disintegrations. The Annals of Probability 3:88–99
[32] Dubois D, Fargier H, Perny P (2003) Qualitative decision theory with preference relations and comparative uncertainty: an axiomatic approach. Artificial Intelligence 148(1-2):219–260
[33] Dubra J, Maccheroni F, Ok E (2004) Expected utility theory without the completeness axiom. Journal of Economic Theory 115:118–133
[34] Galaabaatar T, Karni E (2013) Subjective expected utility with incomplete preferences. Econometrica 81(1):255–284
[35] Ghirardato P, Siniscalchi M (2012) Ambiguity in the small and in the large. Econometrica 80(6):2827–2847
[36] Hacking I (1967) Slightly more realistic personal probability. Philosophy of Science 34(4):311–325
[37] Hájek A, Smithson M (2012) Rationality and indeterminate probabilities. Synthese 187(1):33–48
[38] Hill B, Lane D (1985) Conglomerability and countable additivity. Sankhya 47:366–379
[39] Howson C (2009) Can logic be combined with probability? Probably. Journal of Applied Logic 7:177–187
[40] Howson C (2011) Bayesianism as a pure logic of inference. In: Bandyopadhyay PS, Forster MR (eds) Handbook of the Philosophy of Science, vol 7: Philosophy of Statistics, Elsevier, pp 441–471
[41] Kikuti D, Cozman FG, Filho RS (2011) Sequential decision making with partially ordered preferences. Artificial Intelligence 175(7-8):1346–1365
[42] Kyburg Jr HE, Smokler HE (eds) (1964) Studies in Subjective Probability. Wiley, New York, second edition (with new material) 1980
[43] Luce RD, Krantz DH (1971) Conditional expected utility. Econometrica 39(2):253–271

[44] Machina M (1982) “Expected utility” analysis without the independence axiom. Econometrica 50(2):277–323
[45] Miranda E (2008) A survey of the theory of coherent lower previsions. International Journal of Approximate Reasoning 48(2):628–658
[46] Miranda E, de Cooman G (2007) Marginal extension in the theory of coherent lower previsions. International Journal of Approximate Reasoning 46(1):188–225
[47] Miranda E, Zaffalon M (2010) Notes on desirability and conditional lower previsions. Annals of Mathematics and Artificial Intelligence 60(3–4):251–309
[48] Miranda E, Zaffalon M (2015) Independent products in infinite spaces. Journal of Mathematical Analysis and Applications 425(1):460–488
[49] Miranda E, Zaffalon M (2015) On the problem of computing the conglomerable natual extension. International Journal of Approximate Reasoning 56:1–27
[50] Miranda E, Zaffalon M (2016) Conformity and independence with coherent lower previsions. International Journal of Approximate Reasoning 78:125–137
[51] Miranda E, Zaffalon M, de Cooman G (2012) Conglomerable natural extension. International Journal of Approximate Reasoning 53(8):1200–1227
[52] Moral S (2005) Epistemic irrelevance on sets of desirable gambles. Annals of Mathematics and Artificial Intelligence 45:197–214
[53] Nau R (2006) The shape of incomplete preferences. The Annals of Statistics 34(5):2430–2448
[54] Nau R (2011) Risk, ambiguity, and state-preference theory. Economic Theory 48(2–3):437–467
[55] Von Neumann J, Morgestern O (1947) Theory of games and economic behaviour. Princeton University Press
[56] Ok E, Ortoleva P, Riella G (2012) Incomplete preferences under uncertainty: indecisiveness in beliefs versus tastes. Econometrica 80(4):1791–1808
[57] Pedersen AP (2014) Comparative expectations. Studia Logica 102(4):811–848
[58] Pigozzi G, Tsoukiàs A, Viappiani P (2016) Preferences in artificial intelligence. Annals of Mathematics and Artificial Intelligence 77(3-4):361–401
[59] Savage LJ (1954) The Foundations of Statistics. Wiley, New York
[60] Seidenfeld T, Schervish MJ, Kadane JB (1995) A representation of partially ordered preferences. The Annals of Statistics 23:2168–2217, reprinted in [62], pp. 69–129
[61] Seidenfeld T, Schervisch MJ, Kadane JB (1998) Non-conglomerability for finite-valued finitely additive probability. Sankhya 60(3):476–491
[62] Seidenfeld T, Schervish MJ, Kadane JB (1999) Rethinking the Foundations of Statistics. Cambridge University Press, Cambridge
[63] Seidenfeld T, Schervisch MJ, Kadane JB (2010) Coherent choice functions under uncertainty. Synthese 172(1):157–176
[64] Shafer G, Gillett PR, Scherl RB (2003) A new understanding of subjective probability and its generalization to lower and upper prevision. International Journal of Approximate Reasoning 33:1–49
[65] Shimony A (1955) Coherence and the axioms of confirmation. The Journal of Symbolic Logic 20(1):1–28
[66] Skyrms B (1980) Causal Necessity. Yale University Press
[67] Tversky A, Kahneman D (1992) Advances in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty 5(4):297–323
[68] Van Camp A (2018) Choice functions as a tool to model uncertainty. PhD thesis, University of Ghent
[69] Van Camp A, de Cooman G, Miranda E, Quaeghebeur E (2018) Coherent choice functions, desirability and indifference. Fuzzy Sets and Systems 341(C):1–36
[70] Van Camp A, Miranda E, de Cooman G (2018) Lexicographic choice functions. International Journal of Approximate Reasoning 92:97–119
[71] Walley P (1991) Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London
[72] Williams PM (1975) Notes on conditional previsions. Tech. rep., School of Mathematical and Physical Science, University of Sussex, UK, reprinted in [73]
[73] Williams PM (2007) Notes on conditional previsions. International Journal of Approximate Reasoning 44:366–383, revised journal version of [72]
[74] Zaffalon M, Miranda E (2013) Probability and time. Artificial Intelligence 198(1):1–51
[75] Zaffalon M, Miranda E (2015) Desirability and the birth of incomplete preferences. CoRR abs/1506.00529, `http://arxiv.org/abs/1506.00529`
[76] Zaffalon M, Miranda E (2017) Axiomatising incomplete preferences through sets of desirable gambles. Journal of Artificial Intelligence Research 60:1057–1126