PRECISION–IMPRECISION EQUIVALENCE IN A BROAD CLASS OF IMPRECISE HIERARCHICAL UNCERTAINTY MODELS

GERT DE COOMAN

ABSTRACT. Hierarchical models are rather common in uncertainty theory. They arise when there is a 'correct' or 'ideal' (so-called *first-order*) uncertainty model about a phenomenon of interest, but the modeler is uncertain about what it is. The modeler's uncertainty is then called *second-order uncertainty*. For most of the hierarchical models in the literature, both the first- and the second-order models are *precise*, i.e., they are based on classical probabilities. In the present paper, I propose a specific hierarchical model that is *imprecise* at the second level, which means that at this level, *lower* probabilities are used. No restrictions are imposed on the underlying first-order model: that is allowed to be either precise or imprecise. I argue that this type of hierarchical model generalizes and includes a number of existing uncertainty models, such as imprecise probabilities, Bayesian models, and fuzzy probabilities. The main result of the paper is what I call *Precision–Imprecision Equivalence*: the implications of the model for decision making and statistical reasoning are the same, whether the underlying first-order model is assumed to be precise or imprecise.

1. INTRODUCTION

Suppose that the information at our disposal about a certain random variable X leads us assume that it is normally distributed with mean zero: we might know, for instance, that the variable is the sum of a large number of effects that tend to cancel each other out, and that we judge to be independent. But we do not have sufficient knowledge about the sizes of the various effects leading to the result X in order to pinpoint a value for the variance σ^2. Thus we put

$$X \sim N(0, \sigma^2),$$

meaning that X is normally distributed with mean zero and variance σ^2. In this case, there is some ideal model, usually called the *first-order* model, but there still is some uncertainty about what it is exactly, as we do not know σ^2. Still, we may be able to model whatever information we have about σ^2 by another probability distribution on the possible values of σ^2, with density function p. This so-called *second-order* model represents our information about what the first-order model is. Both models considered together make up a *hierarchical model*. In the present case, it has two levels, but it is not difficult to conceive of hierarchical models with three, or four, or even more levels. In this paper however, we shall only be concerned with two-level models.

Hierarchical models are rather common in uncertainty theory. They arise when there is a 'correct' or 'ideal' (first-order) uncertainty model about a phenomenon of interest, but

---

Date: 25 February 2000.

1991 *Mathematics Subject Classification.* 60A05,62A01.

Key words and phrases. Hierarchical uncertainty model, coherence, natural extension, imprecision.

* Research Associate of the Fund for Scientific Research – Flanders (Belgium).

the modeler is uncertain about what it is. The modeler's uncertainty is then called *second-order uncertainty*. A list of examples showing that second-order uncertainty occurs in many situations can be found in De Cooman and Walley (1999).

By far the most common hierarchical model is the Bayesian one, of which the above example is a special case. Here both the first- and the second-order uncertainty are represented by probability measures. For more information about this type of model, see the work of Zellner (1971), Good (1980), Goldstein (1983; 1985), Berger (1985), von Winterfeldt and Edwards (1986), Bernardo and Smith (1994) and Robert (1994). It is an interesting property of a Bayesian hierarchical model that it can always be reduced to a first-order model, by 'integrating out the higher-order parameters'. In the example above, we can use Bayes' rule to combine the conditional normal density function $N(\cdot \mid 0, \sigma^2)$ and the marginal p into a first-order density function q, where for $x \in \mathbb{R}$:

$$q(x) = \int_0^{+\infty} N(x \mid 0, \sigma^2) p(\sigma^2) \mathrm{d}\sigma^2.$$

Whenever a Bayesian hierarchical model is used, it is assumed that the available information is at all levels detailed enough for it to be representable by a probability measure. An important lesson to be learnt from the recent work on imprecise probabilities (Smith, 1961; Levi, 1980; Walley, 1991; De Cooman et al., 1999) is that it is an exception rather than the rule that a subject has sufficient information in order to represent his beliefs in terms of a *precise* model, i.e., a probability measure. It usually makes more sense for a subject to represent his information using an *imprecise* model, in terms of so-called *upper* and *lower* probabilities, or even more expressively, in terms of upper and lower previsions. To explain this, let us consider the well-known behavioral definition of probability in terms of betting rates. A bet *on* an event A at rate r, where $0 \leq r \leq 1$, yields a reward $1 - r$ if A occurs, and a (negative) reward $-r$ if it doesn't. Clearly, the stronger a subject's belief that A will occur, the higher the rates r at which he will accept bets on A. His *lower* probability $\underline{P}(A)$ for A is defined as the highest such rate r, and it will increase with the strength of his beliefs in the occurrence of A. On the other hand, his *upper* probability $\overline{P}(A)$ for A is defined as one minus his lower probability for the opposite event $\mathrm{co}A$, and it decreases with the strength of the subject's beliefs against the occurrence of A. If the subject has little specific information regarding A, he will have no strong beliefs either in favor of or against A occurring, so his lower probability $\underline{P}(A)$ will be little higher than zero and his upper probability only slightly smaller than one. The more specific and detailed the information about A becomes, the closer $\underline{P}(A)$ and $\overline{P}(A)$ will tend to come together. Only in the limit that $\underline{P}(A) = \overline{P}(A)$ there will be a number $P(A)$, equal to this common value, such that for all $r < P(A)$, the subject accepts bets on A at rate r, and for all $s > P(A)$, he accepts bets against A at rate $1 - s$, that is, $P(A)$ is a fair betting rate, or probability, in the sense of de Finetti (1974). An uncertainty model is called imprecise if for some of the events considered, the lower probability is strictly smaller than the upper probability, and precise if upper and lower probabilities coincide everywhere. Evidently, imprecise models will be much more common than the precise ones, which are just special limit cases.

In this light, it seems strange that while there is not enough information available to identify a precise first-order model, it would nevertheless allow us to specify a precise model at the more abstract second level. This has been recognized as an important conceptual problem for Bayesian hierarchical models, and it should therefore not be surprising that there are a number of models in the literature that allow imprecision at the second level, while still assuming that the first-order model is a (precise) probability measure or prevision; examples are the robust Bayesian models (Berger, 1994), models involving second-order possibility distributions (Walley, 1997; De Cooman, 1998; De Cooman and Walley, 1999), fuzzy probabilities (Zadeh, 1976; Watson et al., 1979; Zadeh, 1984) and Gärdenfors and Sahlin's (1982) epistemic reliability model. Nau's (1992) approach involves an imprecise first-order model bearing a formal relationship to fuzzy probabilities. I know of no detailed analysis where imprecision is explicitly allowed at both levels, but see Walley (1991, Section 5.10.5) for a brief discussion.

In the present paper, I introduce and study a behavioral hierarchical model involving lower probabilities at the second level, while the underlying first-order model is allowed to be either precise or imprecise. This model is closely related to, but in a sense more general than the buying functions introduced in De Cooman and Walley (1999).[1] It also generalizes many of the hierarchical models in the literature. I argue that it can be reduced to an imprecise first-order model, using similar principles to the case of Bayesian hierarchical models. This first-order model can then be used in decision making and statistical reasoning. Interestingly, I show that as far as the induced first-order model and a number of other interesting behavioral implications are concerned, it does not matter whether we assume that the underlying ideal first-order model is precise or imprecise: they are identical under either assumption. I call this result *Precision–Imprecision Equivalence*. It generalizes a number of results known in the literature: the close *formal* analogy between Walley's behavioral theory of imprecise probabilities (Walley, 1991) and Bayesian sensitivity analysis (Berger, 1994), the results concerning second-order possibility distributions in Walley (1997) and the representation theorems in De Cooman and Walley (1999).

The paper is organized as follows. Section 2 gives an overview of the basic notions of the behavioral theory of imprecise probabilities, which are needed for the development of the hierarchical model. This model, in terms of so-called lower desirability functions, is introduced in Section 3. It can be made more explicit mathematically by providing more details about the underlying ideal first-order model. In Section 4, I assume that the first-order model is imprecise. Based on this assumption, I introduce a number of rationality criteria which can be imposed on the second-order model, together with a notion of natural extension, which can be used to explore the behavioral implications of given lower desirability assessments. A similar treatment is given in Section 5, based on the assumption that the underlying ideal first-order model is precise. The main results of the paper are gathered in Section 6, where I show that the behavioral implications of lower desirability functions do not depend on whether the underlying first-order model is assumed to be precise or not. I also prove a number of properties of natural extension that clarify the analogy with natural extension in the theory of first-order imprecise probabilities. In Section 7 the present model is related to Walley's theory of first-order imprecise probabilities (Walley, 1991), Bayesian first- and second-order models (de Finetti, 1974; Good, 1980), and the theory of fuzzy probability and buying functions explored in De Cooman (1998), De Cooman and Walley (1999) and De Cooman (2000). Section 8 concludes the paper.

2. Precise and Imprecise First-Order Models

In this section, I discuss a number of aspects of the precise and imprecise uncertainty models that will serve as a basis for the development of the more complex second-order models in the later sections. This necessarily brief exposition is based on Walley's (1991) behavioral account of imprecise probabilities, which should be consulted for more details and further discussion.

---

[1] Indeed, many of the ideas discussed here go back to that paper, which also contains a much more detailed discussion of hierarchical uncertainty models, their history, and the motivations for studying them.

Call a *possibility space* the set Ω of possible states ω of the world—mutually exclusive and exhaustive—that are of interest. A *gamble* on Ω is a bounded, real-valued function on the domain Ω. It can be interpreted as an uncertain reward: if the true state of the world turns out to be ω then the (possibly negative) reward is $X(\omega)$. The reward X is uncertain because it is uncertain which element of Ω is the true state. In order to simplify the discussion, and to separate the issues of uncertainty modeling and utility, I shall concentrate on modeling uncertainty, and assume throughout that rewards are expressed in units of some linear utility, called *utiles*. See Walley (1991, Section 2.2) and Kadane et al. (1999, Part 2) for discussion and motivation.

I use the notation $\mathcal{L}(\Omega)$ for the set of all gambles on Ω. In more traditional approaches to uncertainty modeling, the emphasis is on *events* rather than gambles. Here I focus on gambles, because the language of gambles allows for much more expressive imprecise probability models than that of events.[2] An event A is a subset of Ω, and it can be interpreted as a zero-one valued gamble, yielding one utile if it occurs and zero if it doesn't. In other words, an event A can be identified with its indicator function I_A, which is a gamble.

A subject's uncertainty about a domain Ω can be measured through his attitudes to gambles defined on Ω. There are a number of equivalent ways in which such attitudes can be represented mathematically: (i) by lower and upper previsions; (ii) by sets of linear previsions; and (iii) by sets of almost-desirable gambles.

2.1. **Lower and upper previsions.** One way to measure a subject's uncertainty is by eliciting his *lower prevision*, or supremum acceptable buying price, $\underline{P}(X)$ and his *upper prevision*, or infimum acceptable selling price, $\overline{P}(X)$ for gambles X.

This approach is based on the idea of buying and selling gambles. The transaction in which a gamble X is bought for a price x has reward $X - x$, and this is a new gamble. A subject's *supremum acceptable buying price* $\underline{P}(X)$ for X is the largest real number c such that he is committed to accept the gamble $X - x$ for all $x < c$. Similarly, his *infimum acceptable selling price* $\overline{P}(X)$ for X is the smallest real number d such that he is committed to accept the gamble $x - X$ for all $x > d$. Since buying a gamble X for price x is the same thing as selling $-X$ for price $-x$, it is a rationality requirement that $\overline{P}(-X) = -\underline{P}(X)$, so we can in principle determine upper previsions from lower previsions, and vice versa. I shall consistently use lower rather than upper previsions in developing the present theory.

As stated above, *events* are subsets of Ω and they can be identified with their indicator functions, which are gambles. For an event $A \subseteq \Omega$, buying and selling prices for its indicator function I_A can be regarded as betting rates on and against A. The lower prevision $\underline{P}(I_A)$ of the gamble I_A is also called the *lower probability* of A, and denoted simply as $\underline{P}(A)$. Similarly, upper probabilities are just upper previsions of (indicator functions of) events.[3]

Lower previsions represent a subject's dispositions to buy gambles, and as such they should satisfy a number of consistency, or rationality, criteria, which can be summarized by the following requirement. Assume that the subject specifies a lower prevision $\underline{P}(X)$ for all gambles X in a subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$. In order to identify its possibility space and domain, I also denote this lower prevision by $(\Omega, \mathcal{K}, \underline{P})$. Then $\underline{P}$ is called *coherent* if for

[2] For *precise* probability models, it turns out to be immaterial which of the two languages is chosen.

[3] In standard statistical practice, it is costumary to use the letter P for 'probability' (defined on events), and the letter E for 'expectation' (defined on gambles, or random variables). In this paper, I do not follow this practice: I use the symbol P to refer to 'prevision' (defined on *both* events and gambles), and 'E' further on to refer to 'extension', following de Finetti (1974) and Walley (1991), respectively.

any natural numbers $n \geq 0$ and $m \geq 0$, and for any gambles $X_o, X_1, \ldots, X_n$ in $\mathcal{K}$:

$$\sup_{\omega \in \Omega} \left[\sum_{k=1}^{n} [X_k(\omega) - \underline{P}(X_k)] - m[X_o(\omega) - \underline{P}(X_o)] \right] \geq 0. \tag{2.1}$$

If the condition (2.1) holds for $m = 0$, then $\underline{P}$ is said to *avoid sure loss*: in that case there is no finite number of buying transactions, which the subject is committed to accept as a result of his specifying $\underline{P}$, such that the net result of these combined transactions is always smaller than some strictly negative amount of utility. If the condition holds for all $m > 0$, the subject's lower prevision $\underline{P}$ is consistent in the sense that by the combination of acceptable buying transactions he cannot be induced to pay more for a gamble than he has specified in his lower prevision for it. See Walley (1991, Chapter 2) for a more detailed discussion and justification of these rationality criteria.

We have so far been concerned with lower previsions defined on subsets $\mathcal{K}$ of $\mathcal{L}(\Omega)$. *Natural extension* allows us to 'extend' a lower prevision $\underline{P}$ on $\mathcal{K}$ that avoids sure loss to a coherent lower prevision on all gambles by taking only two things into account: (a) the information contained in $\underline{P}$, and (b) the requirement of coherence. Consider any gamble X on Ω. Assume that p is our subject's supremum acceptable buying price for X. If this new assessment is to be coherent with the lower prevision assessments $\underline{P}$ made previously, it is necessary that $p \geq \underline{E}(X)$, where

$$\underline{E}(X) = \sup_{n, \lambda_k, X_k} \inf_{\omega \in \Omega} \left[X(\omega) - \sum_{k=1}^{n} \lambda_k [X_k(\omega) - \underline{P}(X_k)] \right]. \tag{2.2}$$

Here and later in the paper I denote by $\sup_{n,\lambda_k,X_k}$ the supremum over integer $n \geq 0$, real $\lambda_k \geq 0$ and gambles $X_k \in \mathcal{K}$, for $k = 1, \ldots, n$. The lower prevision $\underline{E}$ defined by (2.2) is called the *natural extension* of the lower prevision $\underline{P}$. It is defined for any gamble X on Ω. Natural extension derives its importance from the following result, proven by Walley (1991, Theorem 3.1.2).

Theorem 2.1. *Let $\underline{P}$ be a lower prevision on a set of gambles $\mathcal{K} \subseteq \mathcal{L}(\Omega)$ that avoids sure loss, and let $(\Omega, \mathcal{L}(\Omega), \underline{E})$ be its natural extension. The following statements hold.*

1. $\inf_{\omega \in \Omega} X(\omega) \leq \underline{E}(X)$ *for all* $X \in \mathcal{L}(\Omega)$.
2. $\underline{E}$ *is a coherent lower prevision on* $\mathcal{L}(\Omega)$.
3. $\underline{E}$ *dominates* $\underline{P}$ *on* $\mathcal{K}$: $\underline{E}(X) \geq \underline{P}(X)$ *for all* $X \in \mathcal{K}$.
4. $\underline{E}$ *coincides with* $\underline{P}$ *on* $\mathcal{K}$ *if and only if* $\underline{P}$ *is coherent.*
5. $\underline{E}$ *is the (point-wise) smallest coherent lower prevision on* $\mathcal{L}(\Omega)$ *that dominates* $\underline{P}$ *on* $\mathcal{K}$.
6. *If* $\underline{P}$ *is coherent then* $\underline{E}$ *is the (point-wise) smallest coherent lower prevision on* $\mathcal{L}(\Omega)$ *that coincides with* $\underline{P}$ *on* $\mathcal{K}$.

This shows that natural extension is *least committal*: any coherent extension of the coherent lower prevision $\underline{P}$ implies a disposition to buy gambles X for a price that is at least as high as $\underline{E}(X)$, and therefore has behavioral implications that are at least as strong. Moreover, if $\underline{P}$ is a lower prevision that avoids sure loss but is not coherent, natural extension corrects and extends it to a coherent lower prevision on all gambles, again in a manner which has minimal behavioral implications.

2.2. **Linear previsions.** The supremum acceptable buying price and the infimum acceptable selling price for a gamble may differ because the subject is indecisive or because he has little information about the gamble. The difference between these buying and selling prices typically decreases as the amount of relevant information increases. In the special case where every gamble X has a 'fair price', meaning that the supremum acceptable buying price agrees with the infimum acceptable selling price, we obtain the theory of linear previsions of de Finetti (1974). A *linear prevision* P on a set of gambles $\mathcal{K}$ is a map taking $\mathcal{K}$ to the set of real numbers $\mathbb{R}$, such that for all $m \geq 0$ and $n \geq 0$, and for any $X_1, \dots,$ X_n and $Y_1, \dots, Y_m$ in $\mathcal{K}$,

$$\sup_{\omega\in\Omega}\left[\sum_{k=1}^{n}[X_k(\omega)-P(X_k)]-\sum_{k=1}^{m}[Y_k(\omega)-P(Y_k)]\right]\geq 0.$$

A linear prevision $(\Omega, \mathcal{K}, P)$ is therefore coherent, both when interpreted as a lower and as an upper prevision on $\mathcal{K}$. Linear previsions are the *precise* probability models, and they provide a link with the more traditional approaches to probability theory: a linear prevision defined on a field of events is simply a finitely additive probability measure. Its natural extension to (measurable) gambles is nothing but the expectation associated with that measure.

We shall mainly be interested in linear previsions that are defined on the set of all gambles $\mathcal{L}(\Omega)$. These can alternatively be characterized as real-valued maps P on $\mathcal{L}(\Omega)$ that (i) are linear: $P(\lambda X+\mu Y)=\lambda P(X)+\mu P(Y)$ for all gambles X and Y and real numbers λ and μ; (ii) are positive: $X \geq 0 \Rightarrow P(X) \geq 0$ for all gambles X; and (iii) have unit norm: $P(1)=1$. The set of all linear previsions on $\mathcal{L}(\Omega)$ will be denoted by $\mathbb{P}$.

Linear previsions are special cases of coherent lower previsions. Interestingly, they can be used to construct an alternative uncertainty model that is equivalent to a lower prevision: there is a very close relationship between coherent lower previsions and *sets of* linear previsions. To see how this comes about, consider a lower prevision $\underline{P}$ on the set of gambles $\mathcal{K}$. If we define its set of dominating linear previsions by

$$\mathcal{M}(\underline{P})=\{P\in\mathbb{P}\colon(\forall X\in\mathcal{K})(P(X)\geq\underline{P}(X))\},$$

then $\underline{P}$ avoids sure loss if and only if $\mathcal{M}(\underline{P}) \neq \emptyset$ and $\underline{P}$ is coherent if and only if $\underline{P}(X) = \inf\{P(X)\colon P \in \mathcal{M}(\underline{P})\}$ for all $X \in \mathcal{K}$, i.e., $\underline{P}$ is the *lower envelope* of $\mathcal{M}(\underline{P})$. If $\underline{P}$ avoids sure loss, the natural extension $\underline{E}$ of $\underline{P}$ is given by $\underline{E}(X) = \inf\{P(X)\colon P \in \mathcal{M}(\underline{P})\}$ for all gambles X on Ω. On the other hand, the lower envelope of any set of linear previsions is always a coherent lower prevision. This correspondence also establishes a link between the theory of coherent lower previsions and Bayesian sensitivity analysis (Berger, 1984).

2.3. **Almost-desirable gambles.** Yet another completely equivalent formulation of the model of lower and upper previsions can be given in terms of sets of almost-desirable gambles. A gamble X is *almost-desirable* (or *at least marginally* acceptable) to a subject if he is disposed to accept the gamble $X+\epsilon$ for all $\epsilon > 0$, or equivalently, if his lower prevision for X is non-negative: $\underline{P}(X) \geq 0$. Instead of asking the subject for his lower prevision, we may ask him to specify a set $\mathcal{D}$ of gambles that he judges to be almost-desirable. This rather uncommon way of modeling uncertainty is conceptually simple and mathematically elegant, and it allows us to view the subject from a new angle. It is precisely this way of looking at things which led me to the Precision–Imprecision Equivalence results to be proven further on. It will be well to devote some attention to it here.

The rationality requirements for this type of uncertainty model have a simple 'geometrical' interpretation. A set of almost-desirable gambles $\mathcal{D}$ is *coherent* if it is a closed[4] convex cone in the linear space $\mathcal{L}(\Omega)$ that contains all non-negative ($X \geq 0$), and no uniformly negative ($\sup_{\omega\in\Omega} X(\omega) < 0$) gambles. It *avoids sure loss* if it is a subset of some coherent

[4]The topology on $\mathcal{L}(\Omega)$ is the supremum norm topology.

set of almost-desirable gambles, or equivalently, if for all $n \geq 0$ and $X_1, \dots, X_n$ in $\mathcal{D}$, $\sup_{\omega\in\Omega}[\sum_{k=1}^{n} X_k(\omega)] \geq 0$. The *natural extension* of a set of almost-desirable gambles $\mathcal{D}$ is the smallest coherent set of almost-desirable gambles that includes $\mathcal{D}$, or equivalently, the intersection of all coherent sets of almost-desirable gambles that include $\mathcal{D}$.

There is a one-to-one correspondence, with an immediate interpretation, between coherent lower previsions $\underline{P}$ on $\mathcal{L}(\Omega)$ and coherent sets of almost-desirable gambles $\mathcal{D}$, given by

$$\underline{P}(X) = \sup\{\mu\colon X - \mu \in \mathcal{D}\} \text{ and } \mathcal{D} = \{X \in \mathcal{L}(\Omega)\colon \underline{P}(X) \geq 0\}. \tag{2.3}$$

I shall denote by $\mathbb{D}$ the collection of all coherent sets of almost-desirable gambles on Ω. A few of its members are of special interest. The *vacuous set of almost-desirable gambles* $\mathcal{D}_v = \{X \in \mathcal{L}(\Omega)\colon X \geq 0\}$ is the smallest coherent set of almost-desirable gambles: it is included in all others. Through (2.3), it corresponds to the (coherent) *vacuous lower prevision* $\underline{P}_v$, defined by $\underline{P}_v(X) = \inf_{\omega\in\Omega} X(\omega)$ for $X \in \mathcal{L}(\Omega)$. The equivalent models $\underline{P}_v$ and $\mathcal{D}_v$ represent minimal behavioral commitments on the part of the subject: he is only disposed to engage in transactions that are sure to yield a non-negative gain. They are suitable models for the subject's complete ignorance about Ω.

On the other hand, there is no greatest coherent set of almost-desirable gambles, but $\mathbb{D}$ does have maximal elements, in the sense that they are not included in any other coherent set of almost-desirable gambles: adding whatever new gamble to them makes them incoherent. These maximal elements are characterized by the fact that the associated lower previsions, through (2.3), are actually linear previsions. In other words, they take the form:

$$\mathcal{D}_P = \{X \in \mathcal{L}(\Omega)\colon P(X) \geq 0\}, \quad P \in \mathbb{P}.$$

Geometrically speaking, $\mathcal{D}_P$ is the closed half-space of almost-desirable gambles containing the first orthant $\mathcal{D}_v$ (all non-negative gambles) and bounded by the hyperplane $\{X \in \mathcal{L}(\Omega)\colon P(X) = 0\}$ through the origin. Note that the inclusion relation between coherent sets of almost-desirable gambles has the simple interpretation of "is a less precise model than"; $\mathcal{D}_v$ is the most imprecise model, and the $\mathcal{D}_P$ are maximally precise.

3. Lower Desirability Functions

To distinguish between first- and second-order uncertainty, we consider a *subject* who is uncertain about a certain phenomenon of interest, for which he has a possibility space Ω. The second-order uncertainty about the subject's first-order uncertainty model is supposed to be that of a second person, called the *modeler*. To further distinguish between the two, I assume that the subject is male and the modeler female. Below I propose an imprecise probabilistic model for the second-order uncertainty of the modeler about the subject's first-order uncertainty about Ω.

By varying the interpretation of subject and modeler, this formulation can be made to cover most of the second-order uncertainty models in the literature, e.g., *partial elicitation*, where the second-order uncertainty stems from the modeler's failure to completely elicit the subject's beliefs; *partial introspection*, where the modeler is modeling her uncertainty about her own behavior (modeler and subject are the same person); and the *aleatory interpretation*, where the modeler is uncertain about the true probabilities governing a random process (the 'subject'). This last interpretation is encountered in Bayesian statistical problems, where often knowledge about the parameters of the distribution of a random variable is modeled by a probability distribution, which is updated using observations of the variable. (This is not usually considered to be a hierarchical model, though.) A more detailed list of interpretations, and further discussion, can be found in De Cooman and Walley (1999).

Consider the event $D(X)$ that the subject judges the gamble X to be almost-desirable. There is second-order uncertainty when the modeler is uncertain whether or not the event $D(X)$ will occur. We model this uncertainty in terms of a lower desirability function $\underline{\mathfrak{d}}$. The real number $\underline{\mathfrak{d}}(X)$, called the *lower desirability* of X, is the modeler's lower probability for the event $D(X)$, i.e., her supremum acceptable rate for betting on the event that the subject judges the uncertain reward X to be almost-desirable. If the modeler gives assessments $\underline{\mathfrak{d}}(X)$ for all gambles in a subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$, she in fact determines a function $\underline{\mathfrak{d}}$ from the set $\mathcal{K}$ to the unit interval $[0,1]$, called a *lower desirability function*. In order to specify its possibility space Ω and domain $\mathcal{K}$, I also denote this function as $(\Omega,\mathcal{K},\underline{\mathfrak{d}})$.

To give a very simple example, complete ignorance about the subject's behavioral dispositions regarding gambles X in a subset $\mathcal{K}$ of $\mathcal{L}(\Omega)$—apart from the assumption that the subject is rational in that he will at least marginally accept a non-negative gain—can be modeled by the *vacuous lower desirability function* $\underline{\mathfrak{d}}_v$, defined as:

$$\underline{\mathfrak{d}}_v(X) = I_{\mathcal{D}_v}(X) = \begin{cases} 1 & \text{if } X \geq 0 \\ 0 & \text{if } X \not\geq 0. \end{cases}$$

$\underline{\mathfrak{d}}_v$ models minimal behavioral dispositions for the modeler: she is disposed to bet at non-trivial rates only on the event that the subject will at least marginally accept a non-negative gain. Other examples of lower desirability functions are discussed in some detail in Section 7.

More complicated types of lower desirability function could be introduced. For instance, if $\mathcal{D}$ is a set of gambles, we could define $\underline{\mathfrak{d}}(\mathcal{D})$ as the modeler's lower probability for the event that *all* gambles in $\mathcal{D}$ are almost-desirable to the subject. In the present paper I shall not deal with these more general models. But I do want to mention that many of the results proven below can be carried over easily to the more complicated cases.

So far, I have not given a very detailed description of the events $D(X)$. In the next two sections, I describe two possible underlying models that can be used to account in more detail for the events $D(X)$, and that are based on different rationality assumptions about the subject on the part of the modeler.

4. IMPRECISE FIRST-ORDER MODEL

In specifying the numbers $\underline{\mathfrak{d}}(X)$, the modeler will make a number of assumptions about the subject, or rather about his behavioral dispositions. The minimal assumption I try to model in this paper is that the subject is a rational person according to the criteria of coherence described in Section 2.

Assumption 4.1. *The modeler assumes that the subject is rational, in the sense that his behavioral dispositions can be modeled by a coherent set of almost-desirable gambles, or equivalently, by a coherent lower prevision.*

If we recall that $\mathbb{D}$ is the collection of all coherent sets of almost-desirable gambles on Ω, Assumption 4.1 can be reformulated as follows: the modeler assumes that the subject has some coherent set of almost-desirable gambles $\mathcal{D}_T$, which is an element of $\mathbb{D}$. But, the modeler's (second-order) uncertainty about the subject's behavioral dispositions does not allow her to identify $\mathcal{D}_T$ unequivocally. This means that for her uncertainty model, the relevant states of the world are the elements of $\mathbb{D}$. The event $D(X)$ that the subject finds the gamble X almost-desirable is now the event that $X \in \mathcal{D}_T$, and it can be identified with the following subset of the possibility space $\mathbb{D}$:

$$D_i(X) = \{\mathcal{D} \in \mathbb{D} \colon X \in \mathcal{D}\},$$

i.e., the set of all possible states of the world in which X is almost-desirable. $\underline{\mathfrak{d}}(X)$ has been defined as the modeler's lower probability for $D(X)$, so we may interpret her lower desirability function $\underline{\mathfrak{d}}$ as a lower probability defined on the subsets $D_i(X)$ of the possibility space $\mathbb{D}$. In summary, under Assumption 4.1, a lower desirability function $\underline{\mathfrak{d}}$ on the set of gambles $\mathcal{K}$ leads to the specification of a lower probability $\underline{P}_i$ on the set of events $D_i(\mathcal{K}) = \{D_i(X) \colon X \in \mathcal{K}\}$ of the possibility space $\mathbb{D}$, as follows:

$$\underline{P}_i(D_i(X)) = \underline{\mathfrak{d}}(X), \quad X \in \mathcal{K}.$$

The lower probability $(\mathbb{D}, D_i(\mathcal{K}), \underline{P}_i)$ therefore models the modeler's beliefs about the subject's behavioral dispositions regarding the possibility space Ω. It will be called the i*–representation*, or i*–representing lower probability*, of $\underline{\mathfrak{d}}$. The 'i' which appears in this definition refers to the fact that the underlying representation is an *i*mprecise probability model. We may now require that this lower probability should satisfy the consistency requirements of avoiding sure loss and coherence, discussed in Section 2. This leads to the following definitions.

Definition 4.1. Let $\underline{\mathfrak{d}} \colon \mathcal{K} \to [0,1]$ be a lower desirability function, defined on a set of gambles $\mathcal{K} \subseteq \mathcal{L}(\Omega)$. Then $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ is called i*–reasonable* if its i–representing lower probability $(\mathbb{D}, D_i(\mathcal{K}), \underline{P}_i)$ avoids sure loss, or more explicitly, if for all $n \geq 1$ and X_1, ... , X_n in $\mathcal{K}$, there is a $\mathcal{D} \in \mathbb{D}$ such that

$$\sum_{k=1}^{n} \left[I_{\mathcal{D}}(X_k) - \underline{\mathfrak{d}}(X_k)\right] \geq 0. \tag{4.1}$$

Moreover, $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ is called i*–representable* if the lower probability $(\mathbb{D}, D_i(\mathcal{K}), \underline{P}_i)$ is coherent, or more explicitly, if for all $m \geq 0$, $n \geq 0$ and X_o, X_1, ... , X_n in $\mathcal{K}$, there is a $\mathcal{D} \in \mathbb{D}$ such that

$$\sum_{k=1}^{n} \left[I_{\mathcal{D}}(X_k) - \underline{\mathfrak{d}}(X_k)\right] \geq m \left[I_{\mathcal{D}}(X_o) - \underline{\mathfrak{d}}(X_o)\right]. \tag{4.2}$$

An i–representable lower desirability function is always i–reasonable. Conditions (4.1) and (4.2) follow from applying to the lower probability $(\mathbb{D}, D_i(\mathcal{K}), \underline{P}_i)$ the conditions of avoiding sure loss and coherence implicit in (2.1). To see this, notice that for any gamble X and any coherent set of almost-desirable gambles $\mathcal{D}$, $\mathcal{D} \in D_i(X)$ if and only if $X \in \mathcal{D}$. Consequently, $I_{D_i(X)}(\mathcal{D}) = I_{\mathcal{D}}(X)$, and since the indicator functions $I_{\mathcal{D}}$ assume only two values, the supremum in (2.1) is actually achieved.

Let us consider a simple example: the vacuous lower desirability $\underline{\mathfrak{d}}_v$ is i–representable on any domain $\mathcal{K}$. It is (point-wise) dominated by, and therefore more conservative or less committal than, all other i–representable lower desirability functions on $\mathcal{K}$.

If the modeler has specified her lower desirability for a collection of gambles $\mathcal{K}$, we can ask what the consequences of her assessments are for the lower desirability of the other gambles not in $\mathcal{K}$. The answer lies in the concept of natural extension, explained in Section 2.1. If the lower desirability function $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ is i–reasonable, then its i–representation $\underline{P}_i$ avoids sure loss, and we can consider the coherent natural extension $\underline{E}_i$ of $\underline{P}_i$ to the events $D_i(X)$, for all gambles X on Ω and not just for gambles in $\mathcal{K}$. This leads to a new lower desirability function $\underline{\mathfrak{e}}_i$ defined on all gambles X on Ω, as follows:

$$\underline{\mathfrak{e}}_i(X) = \underline{E}_i(D_i(X)) = \sup_{n,\lambda_k,X_k} \inf_{\mathcal{D}\in\mathbb{D}} \left[I_{\mathcal{D}}(X) - \sum_{k=1}^{n} \lambda_k \left[I_{\mathcal{D}}(X_k) - \underline{\mathfrak{d}}(X_k) \right] \right]. \tag{4.3}$$

This formula is obtained after applying (2.2) to calculate the natural extension $\underline{E}_i$ for the lower probability $(\mathbb{D}, D_i(\mathcal{K}), \underline{P}_i)$. I call the lower desirability function $(\Omega, \mathcal{L}(\Omega), \underline{\mathfrak{e}}_i)$ the *i–natural extension* of $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$.

To give a simple example, the i–natural extension of the vacuous lower desirability function $\underline{\mathfrak{d}}_v$ on a set of gambles $\mathcal{K}$ is the vacuous lower desirability function on $\mathcal{L}(\Omega)$.

It is well established that (precise) Bayesian second-order models can be converted or reduced to (precise) Bayesian first-order models by 'integrating out the second-order parameters', as explained in the Introduction. This reduction is actually a consequence of coherence, and therefore it is a form of natural extension (Goldstein, 1983; Walley, 1991). It turns out that something similar can be done for the imprecise second-order models we are considering here. Let us consider the situation where the unknown first-order model is an 'ideal' in that, if the modeler knew what the subject's model $\mathcal{D}_T \in \mathbb{D}$ was, she would adopt it as her own model for making decisions—this happens for instance when modeler and subject coincide. Then, applying the general ideas in Walley (1991, Section 5.10.5), the natural extension of the second-order model $\underline{\mathfrak{d}}$ to a first-order model $\underline{E}_i^1$ is given by

$$\underline{E}_i^1(X) = \underline{E}_i(\widetilde{X}) = \sup_{n,\lambda_k,X_k} \inf_{\mathcal{D}\in\mathbb{D}} \left[\widetilde{X}(\mathcal{D}) - \sum_{k=1}^{n} \lambda_k \left[I_{\mathcal{D}}(X_k) - \underline{\mathfrak{d}}(X_k) \right] \right] \tag{4.4}$$

for all gambles X on Ω, where $\widetilde{X}(\mathcal{D})$ is the supremum acceptable buying price for the gamble X corresponding to the coherent set of almost-desirable gambles $\mathcal{D}$, obtained through (2.3):

$$\widetilde{X}(\mathcal{D}) = \sup\{\mu\colon X - \mu \in \mathcal{D}\}. \tag{4.5}$$

I shall call the lower prevision $\underline{E}_i^1$ the *first-order i–natural extension* of the lower desirability function $\underline{\mathfrak{d}}$. Eq. (4.4) is again found after applying (2.2) to calculate the natural extension $\underline{E}_i$ for the lower probability $(\mathbb{D}, D_i(\mathcal{K}), \underline{P}_i)$. $\underline{E}_i^1(X)$ is the least committal supremum acceptable buying price for X that is still compatible with the modeler's lower desirability assessments $\underline{\mathfrak{d}}$ taking into account the requirements of coherence: it is the supremum price that the modeler can be induced to pay for X by combining finite numbers of acceptable bets on events of the type $D_i(X)$, $X \in \mathcal{K}$. If no other information than $\underline{\mathfrak{d}}$ is available, the modeler should use the *induced first-order lower prevision* $\underline{E}_i^1$ for making decisions, or as an imprecise prior in statistical reasoning. For more details, see Walley (1991, Section 5.10.5), Walley (1997) and De Cooman and Walley (1999).

For the vacuous lower desirability function $\underline{\mathfrak{d}}_v$ on a set of gambles $\mathcal{K}$, the first-order i–natural extension is the vacuous lower prevision $\underline{P}_v$ on $\mathcal{L}(\Omega)$.

5. Precise First-Order Model

It is possible for the modeler to make stronger assumptions about the subject's first-order model than the rationality hypothesis of coherence I discussed in the previous section.

Assumption 5.1. *The modeler assumes that the subject is a Bayesian agent, in the sense that his behavioral dispositions can be modeled by a linear prevision, or equivalently, by a maximal coherent set of almost-desirable gambles.*

This assumption leads to a second-order model that is fairly similar to the one discussed before; only now we shall work with linear previsions $P \in \mathbb{P}$, or equivalently half-spaces $\mathcal{D}_P$ of almost-desirable gambles, rather than the more general coherent sets $\mathcal{D} \in \mathbb{D}$ of almost-desirable gambles. Since the discussion uses essentially the same ideas as in the previous section, I shall dispense with motivation and justification, and limit myself to introducing and stating some definitions.

The modeler assumes in effect that the subject has some linear prevision P_T which is an element of $\mathbb{P}$, and we may interpret her lower desirability function $\underline{\mathfrak{d}}$ as a lower probability on the possibility space $\mathbb{P}$. The event $D(X)$ that the subject finds the gamble X almost-desirable is now the event that $X \in \mathcal{D}_{P_T}$, or equivalently that $P_T(X) \geq 0$, and it can be identified with the following subset of $\mathbb{P}$:

$$D_p(X) = \{P \in \mathbb{P} \colon X \in \mathcal{D}_P\} = \{P \in \mathbb{P} \colon P(X) \geq 0\},$$

the set of all precise first-order models for which the gamble X is almost-desirable. Under Assumption 5.1, a lower desirability function $\underline{\mathfrak{d}}$ on the set of gambles $\mathcal{K}$ leads to the specification of a lower probability $\underline{P}_p$ on the set of events $D_p(\mathcal{K}) = \{D_p(X) \colon X \in \mathcal{K}\}$ of the possibility space $\mathbb{P}$, as follows:

$$\underline{P}_p(D_p(X)) = \underline{\mathfrak{d}}(X), \quad X \in \mathcal{K}.$$

The lower probability $(\mathbb{P}, D_p(\mathcal{K}), \underline{P}_p)$ models the modeler's beliefs about the subject's behavioral dispositions regarding the possibility space Ω, and it will be called the *p–representation*, or *p–representing lower probability* of $\underline{\mathfrak{d}}$. The 'p' which appears in this definition refers to the fact that the underlying representation is a *p*recise probability model. Requiring that this lower probability should satisfy the consistency requirements of avoiding sure loss and coherence leads to the following definitions.

Definition 5.1. Let $\underline{\mathfrak{d}} \colon \mathcal{K} \to [0,1]$ be a lower desirability function, defined on a set of gambles $\mathcal{K} \subseteq \mathcal{L}(\Omega)$. Then $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ is called *p–reasonable* if its p–representing lower probability $(\mathbb{P}, D_p(\mathcal{K}), \underline{P}_p)$ avoids sure loss, or more explicitly, if for all $n \geq 1$ and X_1, ..., X_n in $\mathcal{K}$, there is a $P \in \mathbb{P}$ such that

$$\sum_{k=1}^{n} \left[I_{\mathcal{D}_P}(X_k) - \underline{\mathfrak{d}}(X_k)\right] \geq 0. \tag{5.1}$$

Moreover, $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ is called *p–representable* if the lower probability $(\mathbb{P}, D_p(\mathcal{K}), \underline{P}_p)$ is coherent, or more explicitly, if for all $m \geq 0$, $n \geq 0$ and X_o, X_1, ..., X_n in $\mathcal{K}$, there is a $P \in \mathbb{P}$ such that

$$\sum_{k=1}^{n} \left[I_{\mathcal{D}_P}(X_k) - \underline{\mathfrak{d}}(X_k)\right] \geq m \left[I_{\mathcal{D}_P}(X_o) - \underline{\mathfrak{d}}(X_o)\right]. \tag{5.2}$$

If the lower desirability function $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ is p–reasonable, its p–representation $\underline{P}_p$ avoids sure loss, and we can define its natural extension $\underline{E}_p$ to all gambles on the possibility space $\mathbb{P}$, and in particular to the events $D_p(X)$, for all $X \in \mathcal{L}(\Omega)$. This allows us to define a new lower desirability function $\underline{\mathfrak{e}}_p$ on all gambles X on Ω, as follows:

$$\underline{\mathfrak{e}}_p(X) = \underline{E}_p(D_p(X)) = \sup_{n,\lambda_k,X_k} \inf_{P \in \mathbb{P}} \left[I_{\mathcal{D}_P}(X) - \sum_{k=1}^{n} \lambda_k \left[I_{\mathcal{D}_P}(X_k) - \underline{\mathfrak{d}}(X_k)\right]\right].$$

The lower desirability function $(\Omega, \mathcal{L}(\Omega), \underline{\mathfrak{e}}_p)$ is called the *p–natural extension* of $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$.

The vacuous lower desirability function $\underline{\mathfrak{D}}_v$ defined on a set of gambles $\mathcal{K}$ is always p–representable. Its p–natural extension is the vacuous lower desirability function on $\mathcal{L}(\Omega)$, and coincides with the i–natural extension.

As in the previous section, we can define a first-order natural extension, now based on the assumption that the subject's model is precise. The *first-order p–natural extension* of a lower desirability function $\underline{\mathfrak{D}}$ on $\mathcal{K}$ that is p-reasonable is the coherent lower prevision $\underline{E}_p^1$ defined for all gambles X on Ω by

$$\underline{E}_p^1(X) = \underline{E}_p(X^*) = \sup_{n,\lambda_k,X_k} \inf_{P\in\mathbb{P}} \left[X^*(P) - \sum_{k=1}^{n} \lambda_k \left[I_{\mathcal{D}_P}(X_k) - \underline{\mathfrak{D}}(X_k) \right] \right],$$

where $X^*(P)$ is the prevision, or fair price, for X associated with the precise model P:

$$X^*(P) = P(X). \tag{5.3}$$

For the vacuous lower desirability function $\underline{\mathfrak{D}}_v$ on a set of gambles $\mathcal{K}$, the first-order p–natural extension is the vacuous lower prevision on $\mathcal{L}(\Omega)$, and therefore coincides with the first-order i–natural extension.

6. Precision–Imprecision Equivalence

So far, I have introduced the rationality criteria of reasonability and representability for lower desirability functions, as well as their first- and second-order natural extensions. For each of these notions, there are two definitions, depending on whether the subject's underlying first-order model is assumed to be precise or imprecise. In the present section, I intend to show that these definitions are equivalent, and therefore actually independent of whether the subject's model is precise or not. This is the gist of the Precision–Imprecision Equivalence results in Theorems 6.1 and 6.2.

Theorem 6.1 (Precision–Imprecision Equivalence, Part I)**.** *Let $(\Omega, \mathcal{K}, \underline{\mathfrak{D}})$ be a lower desirability function. Then $(\Omega, \mathcal{K}, \underline{\mathfrak{D}})$ is i–reasonable if and only if it is p–reasonable; and $(\Omega, \mathcal{K}, \underline{\mathfrak{D}})$ is i–representable if and only if it is p–representable.*

Proof. Consider $m \geq 0$, $n \geq 0$, and X_o, ... , X_n in $\mathcal{K}$. To prove the theorem, it is sufficient to show that the existence of a $\mathcal{D} \in \mathbb{D}$ such that (4.2) holds is equivalent to the existence of a $P \in \mathbb{P}$ such that (5.2) holds. For a start, if (5.2) holds for some $P \in \mathbb{P}$, then obviously (4.2) holds for $\mathcal{D}_P \in \mathbb{D}$. Conversely, assume that (4.2) holds for some $\mathcal{D} \in \mathbb{D}$. Denote by $\mathcal{M}(\mathcal{D})$ the set $\{P \in \mathbb{P} \colon (\forall X \in \mathcal{D})(P(X) \geq 0)\}$ of all linear previsions P such that $\mathcal{D} \subseteq \mathcal{D}_P$. This set $\mathcal{M}(\mathcal{D})$ is non-empty because the coherent $\mathcal{D}$ in particular avoids sure loss (Walley, 1991, Theorem 3.8.5). If $m = 0$, then it follows that (5.2) holds for all $P \in \mathcal{M}(\mathcal{D}) \neq \emptyset$, since for such P we have $I_{\mathcal{D}} \leq I_{\mathcal{D}_P}$. This already completes the proof for the first statement. If $m > 0$ there are two possibilities. Either $X_o \in \mathcal{D}$, and then (5.2) holds for all $P \in \mathcal{M}(\mathcal{D}) \neq \emptyset$. Or $X_o \notin \mathcal{D}$, and since $\mathcal{D} = \bigcap_{P\in\mathcal{M}(\mathcal{D})} \mathcal{D}_P$ (Walley, 1991, Theorem 3.8.5), there is a $Q \in \mathcal{M}(\mathcal{D})$ such that $X_o \notin \mathcal{D}_Q$. Then (5.2) holds for Q. This completes the proof for the second statement. □

We may therefore drop the references i and p to the underlying models when speaking about whether a lower desirability function is reasonable or representable. There is an even more striking equivalence result for natural extension.

Theorem 6.2 (Precision–Imprecision Equivalence, Part II)**.** *Let $(\Omega, \mathcal{K}, \underline{\mathfrak{D}})$ be a lower desirability function. If $\underline{\mathfrak{D}}$ is reasonable, then its i–natural extension and its p–natural extension agree everywhere: $\underline{\mathfrak{e}}_i(X) = \underline{\mathfrak{e}}_p(X)$ for all $X \in \mathcal{L}(\Omega)$. Moreover, if $\underline{\mathfrak{D}}$ is representable* *then, its first-order i–natural extension and its first-order p–natural extension agree everywhere: $\underline{E}^1_i(X)=\underline{E}^1_p(X)$ for all $X\in\mathcal{L}(\Omega)$.*

Proof. As in the proof of the previous theorem, let $\mathcal{M}(\mathcal{D})$ be the (non-empty) set of linear previsions $\{P\in\mathbb{P}\colon(\forall X\in\mathcal{D})(P(X)\geq 0)\}$, for all $\mathcal{D}\in\mathbb{D}$. Let Y be a gamble on the set $\mathbb{P}$, and use it to define a gamble $Y^\uparrow$ on the set $\mathbb{D}$ as follows: for all $\mathcal{D}\in\mathbb{D}$, $Y^\uparrow(\mathcal{D})=\inf_{P\in\mathcal{M}(\mathcal{D})}Y(P)$. We first prove that $\underline{E}_i(Y^\uparrow)=\underline{E}_p(Y)$, or equivalently,

$$\sup_{n,\lambda_k,X_k}\inf_{\mathcal{D}\in\mathbb{D}}\left[Y^\uparrow(\mathcal{D})-\sum_{k=1}^n\lambda_k\left[I_{\mathcal{D}}(X_k)-\underline{\mathfrak{d}}(X_k)\right]\right]$$
$$=\sup_{n,\lambda_k,X_k}\inf_{P\in\mathbb{P}}\left[Y(P)-\sum_{k=1}^n\lambda_k\left[I_{\mathcal{D}_P}(X_k)-\underline{\mathfrak{d}}(X_k)\right]\right].$$

Consider $n\geq 0$, non-negative real $\lambda_1,\dots,\lambda_n$, and $X_1,\dots,X_n$ in $\mathcal{K}$. It suffices to show that the corresponding infima in the above expression are equal. Since $\{\mathcal{D}_P\colon P\in\mathbb{P}\}\subseteq\mathbb{D}$, and since $\mathcal{M}(\mathcal{D}_P)=\{P\}$ whence $Y^\uparrow(\mathcal{D}_P)=Y(P)$, it follows at once that

$$\inf_{\mathcal{D}\in\mathbb{D}}\left[Y^\uparrow(\mathcal{D})-\sum_{k=1}^n\lambda_k\left[I_{\mathcal{D}}(X_k)-\underline{\mathfrak{d}}(X_k)\right]\right]$$
$$\leq\inf_{P\in\mathbb{P}}\left[Y(P)-\sum_{k=1}^n\lambda_k\left[I_{\mathcal{D}_P}(X_k)-\underline{\mathfrak{d}}(X_k)\right]\right].$$

Conversely, we see that

$$\inf_{\mathcal{D}\in\mathbb{D}}\left[Y^\uparrow(\mathcal{D})-\sum_{k=1}^n\lambda_k\left[I_{\mathcal{D}}(X_k)-\underline{\mathfrak{d}}(X_k)\right]\right]$$
$$=\inf_{\mathcal{D}\in\mathbb{D}}\inf_{P\in\mathcal{M}(\mathcal{D})}\left[Y(P)-\sum_{k=1}^n\lambda_k\left[I_{\mathcal{D}}(X_k)-\underline{\mathfrak{d}}(X_k)\right]\right]$$
$$\geq\inf_{P\in\mathbb{P}}\left[Y(P)-\sum_{k=1}^n\lambda_k\left[I_{\mathcal{D}_P}(X_k)-\underline{\mathfrak{d}}(X_k)\right]\right],$$

since $I_{\mathcal{D}}\leq I_{\mathcal{D}_P}$ for any $P\in\mathcal{M}(\mathcal{D})$, and moreover $\mathbb{P}=\bigcup_{\mathcal{D}\in\mathbb{D}}\mathcal{M}(\mathcal{D})$. This proves that $\underline{E}_i(Y^\uparrow)=\underline{E}_p(Y)$.

It remains to be proven that both statements follow from this equality. Consider any gamble X on Ω. Observe that both $I_{D_p(X)}$ and X^* are gambles on $\mathbb{P}$. Since for any $P\in\mathbb{P}$ and $\mathcal{D}\in\mathbb{D}$, $I_{D_p(X)}(P)=I_{\mathcal{D}_P}(X)$ and $I_{D_i(X)}(\mathcal{D})=I_{\mathcal{D}}(X)$, and moreover $\mathcal{D}=\bigcap\{\mathcal{D}_P\colon P\in\mathcal{M}(\mathcal{D})\}$ (Walley, 1991, Theorem 3.8.5) we see that

$$I^\uparrow_{D_p(X)}(\mathcal{D})=\inf_{P\in\mathcal{M}(\mathcal{D})}I_{D_p(X)}(P)=\inf_{P\in\mathcal{M}(\mathcal{D})}I_{\mathcal{D}_P}(X)=I_{\mathcal{D}}(X)=I_{D_i(X)}(\mathcal{D}),$$

whence $I^\uparrow_{D_p(X)}=I_{D_i(X)}$ and $\underline{\mathfrak{e}}_i(X)=\underline{E}_i(I_{D_i(X)})=\underline{E}_i(I^\uparrow_{D_p(X)})=\underline{E}_p(I_{D_p(X)})=\underline{\mathfrak{e}}_p(X)$. Similarly, combine (4.5) and $\mathcal{D}=\bigcap\{\mathcal{D}_P\colon P\in\mathcal{M}(\mathcal{D})\}$ to find $X^{*\uparrow}=\widetilde{X}$, whence $\underline{E}^1_i(X)=\underline{E}_i(\widetilde{X})=\underline{E}_i(X^{*\uparrow})=\underline{E}_p(X^*)=\underline{E}^1_p(X)$. □

In discussing natural extension, we may therefore drop the references i and p to the underlying models as well.

Remark 6.1. Precision–imprecision equivalence is not a trivial, or vacuous, result. This becomes clear when we consider the notion of upper desirability, for which, maybe surprisingly, there is no such equivalence. The modeler's *upper desirability* $\overline{\mathfrak{d}}(X)$ for a gamble X can be defined as her *upper* probability that the subject will find X almost-desirable. By replacing 'lower probability' by 'upper probability' in Definitions 4.1 and 5.1, we may define i/p–reasonability and i/p–representability for upper desirability functions. But these definitions are no longer equivalent. To see this, consider the upper desirability function $\overline{\mathfrak{d}}$ defined for all gambles X on Ω by

$$\overline{\mathfrak{d}}(X) = I_{\mathcal{D}_v}(X) = \begin{cases} 1 & \text{if } X \geq 0 \\ 0 & \text{elsewhere.} \end{cases}$$

It expresses that the modeler is absolutely sure that the subject's model is the vacuous $\mathcal{D}_v$, and it is therefore clearly i–representable. But it is not even p-reasonable: it is completely incompatible with Assumption 5.1 that the subject has a precise model.

The following result outlines a number of general properties of the natural extension of a lower desirability function. Notice the close analogy with Theorem 2.1.

Theorem 6.3. *Let* $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ *be a lower desirability function that is reasonable, and let* $(\Omega, \mathcal{L}(\Omega), \underline{\mathfrak{e}})$ *be its natural extension. The following statements hold.*

1. $I_{\mathcal{D}_v}(X) \leq \underline{\mathfrak{e}}(X)$ *for all* $X \in \mathcal{L}(\Omega)$.
2. $\underline{\mathfrak{e}}$ *is a representable lower desirability function on* $\mathcal{L}(\Omega)$.
3. $\underline{\mathfrak{e}}$ *dominates* $\underline{\mathfrak{d}}$ *on* $\mathcal{K}$*:* $\underline{\mathfrak{e}}(X) \geq \underline{\mathfrak{d}}(X)$ *for all* $X \in \mathcal{K}$.
4. $\underline{\mathfrak{e}}$ *coincides with* $\underline{\mathfrak{d}}$ *on* $\mathcal{K}$ *if and only if* $\underline{\mathfrak{d}}$ *is representable.*
5. $\underline{\mathfrak{e}}$ *is the (point-wise) smallest representable lower desirability function on* $\mathcal{L}(\Omega)$ *that dominates* $\underline{\mathfrak{d}}$ *on* $\mathcal{K}$.
6. *If* $\underline{\mathfrak{d}}$ *is representable then* $\underline{\mathfrak{e}}$ *is the (point-wise) smallest representable lower desirability function on* $\mathcal{L}(\Omega)$ *that coincides with* $\underline{\mathfrak{d}}$ *on* $\mathcal{K}$.

Proof. I shall use imprecise first-order models to prove the theorem. The proof involving precise first-order models is completely analogous. The first statement follows from Eq. (4.3): the infimum for $n = 0$ is precisely $I_{\mathcal{D}_v}(X)$, since $\mathcal{D}_v$ is included in any other element of $\mathbb{D}$. The rest of the proof relies rather heavily on Theorem 2.1. The second statement is obvious, since $\underline{\mathfrak{e}}_i$ has i–representation $\underline{E}_i$ by construction, and $\underline{E}_i$ is coherent since $\underline{P}_i$ avoids sure loss, by assumption. Since $\underline{P}_i$ is dominated by its natural extension $\underline{E}_i$ on $D_i(\mathcal{K})$, we have for all $X \in \mathcal{K}$ that $\underline{\mathfrak{e}}_i(X) = \underline{E}_i(D_i(X)) \geq \underline{P}_i(D_i(X)) = \underline{\mathfrak{d}}_i(X)$, which proves the third statement. We now prove the fourth statement. $\underline{\mathfrak{d}}$ is i–representable if and only if its i–representation $(\mathbb{D}, D_i(\mathcal{K}), \underline{P}_i)$ is coherent, or in other words, if and only if $\underline{P}_i$ agrees on its domain $D_i(\mathcal{K})$ with its natural extension $\underline{E}_i$, and this is equivalent to the equality of $\underline{\mathfrak{e}}_i$ and $\underline{\mathfrak{d}}$ on $\mathcal{K}$. To prove the last two statements, consider an i–representable lower desirability function $\underline{\mathfrak{d}}'$ on $\mathcal{L}(\Omega)$, and denote its (coherent) i–representation by $(\mathbb{D}, D_i(\mathcal{L}(\Omega)), \underline{P}_i')$. First, assume that $\underline{\mathfrak{d}}'$ dominates $\underline{\mathfrak{d}}$ on $\mathcal{K}$. Then $\underline{P}_i'$ dominates $\underline{P}_i$ on $D_i(\mathcal{K})$, and therefore everywhere dominates the natural extension $\underline{E}_i$ of $\underline{P}_i$. Consequently $\underline{\mathfrak{e}}_i$ is dominated by $\underline{\mathfrak{d}}'$ on $\mathcal{L}(\Omega)$, which proves the fifth statement. Finally, assume that $\underline{\mathfrak{d}}$ is i-representable, and that $\underline{\mathfrak{d}}$ and $\underline{\mathfrak{d}}'$ agree on $\mathcal{K}$, so the coherent lower previsions $\underline{P}_i$ and $\underline{P}_i'$ agree on $D_i(\mathcal{K})$. It follows from the properties of natural extension that $\underline{P}_i'$ dominates $\underline{E}_i$ everywhere, which implies that $\underline{\mathfrak{d}}'$ dominates $\underline{\mathfrak{e}}_i$ on $\mathcal{L}(\Omega)$. Moreover, since $\underline{P}_i$ is coherent, it agrees with its natural extension $\underline{E}_i$ on its domain $D_i(\mathcal{K})$, and consequently $\underline{\mathfrak{e}}_i$ and $\underline{\mathfrak{d}}$ agree on $\mathcal{K}$. This proves the last statement. □

7. Connections with existing uncertainty models

Lower desirability functions generalize a number of existing uncertainty representations, such as imprecise probabilities and precise Bayesian models. There is also an interesting connection with Bayesian hierarchical representations, and lower desirability functions can be used to provide a behavioral foundation for so-called fuzzy probability models. In this section, I give a brief discussion of these connections.

7.1. **Imprecise probability models.** Consider a set $\mathcal{K}$ of gambles on Ω, and assume that *the modeler is absolutely sure that the subject will judge all the gambles X in $\mathcal{K}$ to be almost-desirable*. She can model this by a lower desirability function $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$, defined as follows:

$$\underline{\mathfrak{d}}(X) = 1, \quad X \in \mathcal{K}. \tag{7.1}$$

There is a very close connection between this type of model and the imprecise probability models studied in Section 2.3, or equivalently, in Section 2.1. I will limit myself to just stating the correspondences, as their proofs are fairly straightforward, given the material in the previous sections and in Walley (1991, Sections 3.7 and 3.8).

(i) $\underline{\mathfrak{d}}$ is reasonable if and only if the set $\mathcal{K}$ of almost-desirable gambles avoids sure loss.

Let us therefore assume that $\mathcal{K}$ indeed avoids sure loss. As explained in Section 2.3, the natural extension $\mathcal{E}$ of $\mathcal{K}$ is the smallest coherent set of almost-desirable gambles that includes $\mathcal{K}$, or equivalently, $\mathcal{E} = \bigcap\{\mathcal{D} \in \mathbb{D} \colon \mathcal{K} \subseteq \mathcal{D}\} = \bigcap\{\mathcal{D}_P \colon P \in \mathbb{P} \text{ and } \mathcal{K} \subseteq \mathcal{D}_P\}$. Any rational subject who judges all gambles in $\mathcal{K}$ to be almost-desirable, will also judge at least all the gambles in $\mathcal{E}$ to be almost-desirable (whether he has a precise model or not). The modeler, who assumes the subject to be rational, can therefore be sure that he will judge all gambles in $\mathcal{E}$ and not just in $\mathcal{K}$ to be almost-desirable. The following facts are therefore not surprising.

(ii) $\underline{\mathfrak{d}}$ is representable on $\mathcal{K}$.
(iii) The natural extension $\underline{\mathfrak{e}}$ of $\underline{\mathfrak{d}}$ is given by $\underline{\mathfrak{e}}(X) = I_{\mathcal{E}}(X)$, for any gamble X on Ω.
(iv) The first-order natural extension $\underline{E}^1$ of $\underline{\mathfrak{d}}$ is the lower prevision associated with the set of almost-desirable gambles $\mathcal{E}$, through Eq. (2.3): $\underline{E}^1(X) = \sup\{\mu \colon X - \mu \in \mathcal{E}\}$, for any gamble X on Ω.

The lower desirability function $\underline{\mathfrak{d}}$ therefore leads to an induced first-order model that is equivalent to the natural extension $\mathcal{E}$ of $\mathcal{K}$.

7.2. **Bayesian first-order models.** To see the connection between lower desirability functions and Bayesian first-order models, or linear previsions, assume that *the modeler knows for sure that the subject has a precise prevision P_S on a set of gambles $\mathcal{K}_S$*, meaning that his lower and upper previsions coincide on $\mathcal{K}_S$. For every $X \in \mathcal{K}_S$, $P_S(X)$ is a *fair price* in the sense of de Finetti (1974): it is the unique number such that the subject accepts to buy the gamble X for any price $x < P_S(X)$ and accepts to sell X for any price $y > P_S(X)$.

The modeler can represent this information in terms of a lower desirability function $\underline{\mathfrak{d}}$ on the domain $\mathcal{K} = \bigcup_{X \in \mathcal{K}_S}[\{X - x \colon x \leq P_S(X)\} \cup \{y - X \colon y \geq P_S(X)\}]$: for $X \in \mathcal{K}_S$ she is sure that the modeler will find $X - x$ almost-desirable for all $x \leq P_S(X)$ and $y - X$ almost-desirable for all $y \geq P_S(X)$, whence, for such x and y:

$$\underline{\mathfrak{d}}(X - x) = 1 \text{ and } \underline{\mathfrak{d}}(y - X) = 1.$$

Note that this is a special case of the assessments in Section 7.1. It is now a straightforward exercise to show that the following statements are equivalent:

(i) the set of almost-desirable gambles $\mathcal{K}$ avoids sure loss;

(ii) $\underline{\mathfrak{D}}$ is reasonable;
(iii) $\underline{\mathfrak{D}}$ is representable; and
(iv) $\mathcal{M}(\mathcal{K}) = \{P \in \mathbb{P}\colon (\forall Y \in \mathcal{K})(P(Y) \geq 0)\} \neq \emptyset$.

Since if we look closer, $\mathcal{M}(\mathcal{K}) = \{P \in \mathbb{P}\colon (\forall X \in \mathcal{K}_S)(P(X) = P_S(X))\}$, these statements are also equivalent to:

(v) P_S is a linear prevision on $\mathcal{K}_S$,

that is, the system of fair prices $P_S(X)$, $X \in \mathcal{K}_S$ is coherent in the sense of de Finetti (1974).

7.3. **Justifying fuzzy probability models.** In the examples discussed above, the model $\underline{\mathfrak{D}}$ assumes only the value one and its natural extension $\underline{\mathfrak{e}}$ only the values zero and one. But what makes lower desirability functions especially interesting is that they may assume values between zero and one, and allow us to express more nuance in assessments: in fact, they allow for a 'much more continuous' transition between the two extremes of absolute certainty (lower probability one) and complete ignorance (lower probability zero). In the present section, I give an indication of how this may be applied.

Consider a gamble X on Ω and the corresponding class of gambles $\mathcal{K}_X = \{X - x\colon x \in R\}$, where R is some subset of the set $\mathbb{R}$ of real numbers. We also consider the following lower desirability function:

$$\underline{\mathfrak{D}}(X - x) = g_X(x), \quad x \in R,$$

where g_X is some map from the set R to the real unit interval $[0, 1]$. Note that $g_X(x)$ is the modeler's lower probability for the event that the subject will buy the gamble X for any price $x - \epsilon$, $\epsilon > 0$, or in other words, that the subject's supremum acceptable buying price, or lower prevision, for X is at least x. In order not to unduly complicate matters, I shall assume that $R = \mathbb{R}$ and that

$(g0)$ g_X is left-continuous.

Fairly similar, but slightly more complicated, results can be derived in the more general case as well.

Since the modeler can expect a rational subject's willingness to buy a gamble to become smaller as its price increases, we assume that:

$(g1)$ g_X is non-increasing.

Moreover, if $x \leq \inf_{\omega\in\Omega} X(\omega)$, then the subject cannot lose from buying X for the price x: the buying transaction always results in a non-negative gain. Since the modeler can be certain that a rational subject will (at least marginally) accept a non-negative gain, we make the following assumption:

$(g2)$ if $x \leq \inf_{\omega\in\Omega} X(\omega)$ then $g_X(x) = 1$.

Finally, if $x > \sup_{\omega\in\Omega} X(\omega)$ then buying X for price x results in a sure loss. Any rational subject will avoid this, so our modeler can be sure that he will not engage in such a transaction. This implies the following assumption, which states that it is completely plausible to the modeler that the subject will not buy X for x:

$(g3)$ if $x > \sup_{\omega\in\Omega} X(\omega)$ then $g_X(x) = 0$.

The p–representation of $\underline{\mathfrak{D}}$ is the lower probability $(\mathbb{P}, D_p(\mathcal{K}_X), \underline{P}_p)$ that is defined by $\underline{P}_p(D_p(X - x)) = g_X(x)$. Note that $D_p(\mathcal{K}_X)$ is a chain of sets. We can use the results in De Cooman and Aeyels (1999) to arrive at the following conclusions.

If $(g0)$–$(g3)$ hold, $\underline{P}_p$ is coherent and so $\underline{\mathfrak{D}}$ is representable. The natural extension $\underline{E}_p$ of $\underline{P}_p$ to all events is a necessity measure, that is, the conjugate lower probability of a possibility measure $\overline{E}_p$ (De Cooman, 1997; De Cooman and Aeyels, 1999) with possibility distribution $\pi\colon \mathbb{P} \to [0,1]$ given by $\pi(P) = 1 - g_X^+(P(X))$. Here $g_X^+\colon \mathbb{R} \to [0,1]$ is the right-continuous non-decreasing mapping defined by $g_X^+(x) = g_X(x+) = \sup\{g_X(x+\epsilon)\colon \epsilon > 0\}$. Note that $\pi(P)$ is the modeler's upper probability that P is the subject's true model P_T. For the natural extension $\underline{\mathfrak{e}}$ of $\underline{\mathfrak{d}}$ we find that, for any gamble Y on Ω:

$$\begin{aligned}\underline{\mathfrak{e}}(Y) &= \underline{E}_p(D_p(Y)) = 1 - \overline{E}_p(\mathrm{co}D_p(Y)) \\ &= 1 - \sup\{\pi(P)\colon P \notin D_p(Y)\} = \inf\{g_X^+(P(X))\colon P(Y) < 0\}.\end{aligned}$$

For the first-order natural extension $\underline{E}^1$ of $\underline{\mathfrak{d}}$ we find, since a necessity measure is 2–monotone and the natural extension of a 2–monotone lower probability to gambles can be found by Choquet integration (Walley, 1981; Walley, 1991):

$$\begin{aligned}\underline{E}^1(Y) = \underline{E}_p(Y^*) &= \inf[Y] + \int_{\inf[Y]}^{\sup[Y]} \underline{\mathfrak{e}}(Y-y)\mathrm{d}y \\ &= \inf[Y] + \int_{\inf[Y]}^{\sup[Y]} \inf\{g_X^+(P(X))\colon P(Y) < y\}\mathrm{d}y\end{aligned}$$

where $\inf[Y]$ denotes the infimum value $\inf_{\omega\in\Omega} Y(\omega)$ of Y and $\sup[Y]$ the supremum value $\sup_{\omega\in\Omega} Y(\omega)$ of Y. There is another way of writing this first-order natural extension. Let $\mathcal{M}_\alpha = \{P \in \mathbb{P}\colon \pi(P) \geq \alpha\}$ be the (non-empty) set of precise models P such that the modeler has an upper probability at least α that P is the subject's true model P_T, for $0 \leq \alpha < 1$. The set of linear previsions $\mathcal{M}_\alpha$ corresponds to a coherent lower prevision $\underline{P}_\alpha$, defined by $\underline{P}_\alpha(Y) = \inf\{P(Y)\colon P \in \mathcal{M}_\alpha\}$. The first-order natural extension is a uniform average of the $\underline{P}_\alpha$: $\underline{E}^1(Y) = \int_0^1 \underline{P}_\alpha(Y)\mathrm{d}\alpha$; see Walley (1997) for a proof.

The present discussion focuses on one gamble X. We may follow the same approach for a number of gambles X in a collection $\mathcal{K}_S$. This leads to a lower desirability function $\underline{\mathfrak{d}}$ defined on the set $\mathcal{K} = \bigcup_{X\in\mathcal{K}_S} \mathcal{K}_X$ by $\underline{\mathfrak{d}}(X - x) = g_X(x)$ for all $X \in \mathcal{K}_S$. The formulae above then yield useful approximations for the natural extension, provided we replace $g_X^+(P(X))$ everywhere by $\sup_{X\in\mathcal{K}_S} g_X^+(P(X))$. This is the approach followed by Peter Walley and myself in De Cooman (2000) and De Cooman and Walley (1999). The functions $1-g_X^+$ are very closely related to the so-called *buying functions* introduced there.

The fact that a possibility distribution π on $\mathbb{P}$ arises naturally (as a consequence of coherence) in this context, also allows us to link lower desirability functions with Zadeh's (1984) notion of a fuzzy probability. Indeed, the uncertainty about the first-order model P_T, modeled by the possibility measure $\overline{E}_p$ with distribution π, leads to uncertainty about its values $P_T(X)$ on gambles $X \in \mathcal{L}(\Omega)$, which can be modeled by the induced possibility distributions π_X on $\mathbb{R}$:

$$\pi_X(x) = \overline{E}_p(\{P \in \mathbb{P}\colon P(X) = x\}) = \sup_{P(X)=x} \pi(P), \quad x \in \mathbb{R}.$$

$\pi_X(x)$ is the modeler's upper probability (or plausibility, or possibility) that the subject's fair price for X equals x. The $\mathbb{R} - [0,1]$ mapping π_X is formally a so-called *fuzzy number*, and it could be called the fuzzy prevision of X, or the fuzzy probability of A if X is the indicator function I_A of an event A. This shows that it is possible to give a behavioral definition of, and justification for, fuzzy previsions and probabilities. More details on the connection between fuzzy previsions and (second-order) possibility distributions can be found in De Cooman (1998) and De Cooman and Walley (1999).

7.4. **Bayesian hierarchical models.** Call a lower desirability function $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ *linearly representable* if its p–representation $(\mathbb{P}, D_p(\mathcal{K}), \underline{P}_p)$ is a linear prevision, that is, if and only if there is some linear prevision Π on $\mathbb{P}$ with domain $\mathcal{L}(\mathbb{P})$, or in other words a Bayesian hierarchical (second-order) model, that *coincides* with $\underline{P}_p$ on the events $D_p(X)$, $X \in \mathcal{K}$. Linearly representable lower desirability functions are in particular also representable (and reasonable). Denote by $\mathbb{L}$ the set of the linearly representable lower desirability functions with domain $\mathcal{L}(\Omega)$.

Using the results relating lower previsions to linear previsions mentioned in Section 2.2, it is straightforward to prove the following statements. Let $(\Omega, \mathcal{K}, \underline{\mathfrak{d}})$ be a lower desirability function, and let

$$\mathcal{M}(\underline{\mathfrak{d}}) = \{\underline{\mathfrak{m}} \in \mathbb{L} \colon (\forall X \in \mathcal{K})(\underline{\mathfrak{d}}(X) \leq \underline{\mathfrak{m}}(X))\}$$

be its set of point-wise dominating linearly representable lower desirability functions. Then

(i) $\underline{\mathfrak{d}}$ is reasonable if and only if it is point-wise dominated on its domain by a linearly representable lower desirability function, i.e., $\mathcal{M}(\underline{\mathfrak{d}}) \neq \emptyset$;
(ii) $\underline{\mathfrak{d}}$ is representable if and only if it is the point-wise infimum of its set $\mathcal{M}(\underline{\mathfrak{d}})$ of point-wise dominating linearly representable lower desirability functions: for all $X \in \mathcal{K}$, $\underline{\mathfrak{d}}(X) = \inf\{\underline{\mathfrak{m}}(X) \colon \underline{\mathfrak{m}} \in \mathcal{M}(\underline{\mathfrak{d}})\}$;
(iii) the natural extension $(\Omega, \mathcal{L}(\Omega), \underline{\mathfrak{e}})$ of $\underline{\mathfrak{d}}$ is given by $\underline{\mathfrak{e}}(X) = \inf\{\underline{\mathfrak{m}}(X) \colon \underline{\mathfrak{m}} \in \mathcal{M}(\underline{\mathfrak{d}})\}$, for all $X \in \mathcal{L}(\Omega)$.

Moreover, if Π is a Bayesian second-order model, the corresponding first-order linear prevision, i.e., its first-order natural extension, E^1, on $\mathcal{L}(\Omega)$ is given by $E^1(X) = \Pi(X^*)$. For the first-order natural extension $\underline{E}^1$ of $\underline{\mathfrak{d}}$, we find correspondingly

$$\underline{E}^1(Y) = \inf\{\Pi(Y^*) \colon (\forall X \in \mathcal{K})(\underline{\mathfrak{d}}(X) \leq \Pi(D_p(X)))\}$$

for all gambles Y on Ω. These observations show that the link between lower desirability functions and Bayesian second-order models is very similar to the one between imprecise probabilities and Bayesian first-order models (linear previsions).

8. Conclusion

The message I want to convey in this paper should now be clear: for an interesting class of imprecise second-order uncertainty models, it does not matter whether the underlying first-order model is assumed to be precise or imprecise, at least as far as their behavioral consequences, modeled by natural extension, are concerned. This result, which I find surprising at the level of ideas, also has interesting practical implications: it tells us that in this context we lose no generality in working with underlying first-order models that are precise. These have the advantage that they are much easier to work with mathematically than their imprecise counterparts.

The discussion has focused on lower desirability: a modeler's *lower* probability for the event that a gamble will be *almost*-desirable to a subject. The italicized words in the previous sentence indicate two limitations of the present model. There is no reason why we could not also study a notion of *upper* desirability, defined in terms a modeler's upper probability for such an event. The observations in Remark 6.1 tell us however that for upper desirability, no precision–imprecision equivalence should be expected. The obvious critique will then be that precision–imprecision equivalence arises only because we look at a limited class of models: the first-order model is allowed to be very general, but the second-order model consists of lower probabilities, which are much less expressive than, say, lower previsions, and are moreover only defined on very special types of events. That is true, of course, although Section 7 assures us that lower desirability functions generalize (and in a sense include) most of the uncertainty models in the literature: Bayesian first- and second-order models, imprecise probabilities, and fuzzy probabilities. It remains to be investigated to what extent and under what special conditions more general imprecise second-order models, such as the one described by Walley (1991, Section 5.10.5), exhibit some kind of precision–imprecision equivalence.

A second imperfection of the present model is that it is formulated in terms of *almost*-desirability. Recall that a gamble X is almost-desirable to a subject if he accepts $X + \epsilon$ for all $\epsilon > 0$. So almost-desirability involves the acceptance of an infinite number of gambles, which indicates problems for the operationalizability of the model: it may not be possible to 'call the second-order bets', i.e., to verify whether or not the event $D(X)$ that the subject finds X almost-desirable has occurred or not. More effort must go into finding ways to make the present model operational. One way to achieve this is by working with real desirability rather than almost-desirability. This being said, I do not expect the change from almost-desirability to real desirability to radically alter the conclusions of this paper: almost-desirability is after all a close approximation to real desirability, introduced and used mainly because there is a one-to-one connection with the notion of a lower prevision; see Walley (1991, Section 3.8, Chapter 4 and Appendix F) for more details.

(Almost-)desirability has a central part in this paper: roughly speaking, I have introduced a second-order model that is a 'lower probability of desirability', whereas the previous hierarchical models in the literature have tended to focus on probability, and were introduced in terms of 'probabilities of probabilities'. The asymmetry in the present approach is bound to be seen as unattractive, especially since the modeler's lower probabilities are much less expressive than the subject's sets of almost-desirable gambles. As already stated, it would be worthwhile to investigate ways of making the modeler's model more expressive while retaining some form of precision–imprecision equivalence. As it is, using almost-desirability at the level of the subject has two advantages. First, it allows us to move closer towards a behavioral, operational definition of second-order probability. Secondly, sets of almost-desirable gambles are mathematically equivalent to better-known uncertainty models such as lower previsions and (sets of) precise probabilities, but they allow us to look at existing problems in a new and sometimes more revealing light: the Precision–Imprecision Equivalence theorems are an interesting illustration of this.

Acknowledgements

I am indebted to Peter Walley, Jean-Marc Bernard and two anonymous referees, who carefully read an earlier draft of the paper, and whose detailed suggestions helped me improve it.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian state, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with the author.

References

Berger, J. O., 1984. The robust Bayesian viewpoint. In: Kadane, J. B. (Ed.), *Robustness of Bayesian Analyses*, Elsevier Science, Amsterdam.

Berger, J. O, 1985. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.

Berger, J. O., 1994. An overview of robust Bayesian analysis. *Test* **3**, 5–124. With discussion.

Bernardo, J. M., Smith, A. F. M., 1994. *Bayesian Theory*. John Wiley & Sons, Chichester.

De Cooman, G., 1997. Possibility theory I–III. *International Journal of General Systems* **25**, 291–371.

De Cooman, G., 1998. Possibilistic previsions. *Proceedings of IPMU '98*, Vol. I. Éditions EDK, Paris, pp. 2–9.

De Cooman, G., 2000. A behavioural theory of fuzzy probability. In preparation.

De Cooman, G., Aeyels, D., 1999. Supremum preserving upper probabilities. *Information Sciences* **118**, 173–212.

De Cooman, G., Cozman, F. G., Moral, S., Walley, P. (Eds.), 1999. *ISIPTA '99 – Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications*. Imprecise Probabilities Project, Ghent.

De Cooman, G., Walley, P., 1999. An imprecise hierarchical model for behaviour under uncertainty. Submitted for publication.

de Finetti, B., 1974. *Theory of Probability*, Vol. 1. John Wiley & Sons, Chichester. English Translation of *Teoria delle Probabilità*.

Gärdenfors, P., Sahlin, N.-E., 1982. Unreliable probabilities, risk taking, and decision making. *Synthese* **53**, 361–386.

Goldstein, M., 1983. The prevision of a prevision. *Journal of the American Statistical Society* **87**, 817–819.

Goldstein, M., 1985. Temporal coherence. In: Bernardo, J. M., DeGroot, M. H., Lindley, D. V., Smith, A. F. M. (Eds.), *Bayesian Statistics*, Vol. 2. North-Holland, Amsterdam, pp. 231–248. With discussion.

Good, I. J., 1980. Some history of the hierarchical Bayesian methodology. In: Bernardo, J. M., DeGroot, M. H., Lindley, D. V., Smith, A. F. M. (Eds.), *Bayesian Statistics*, Vol. 1. Valencia University Press, Valencia, pp. 489–519.

Kadane, J. B., Schervish, M. J., Seidenfeld, T., 1999. *Rethinking the Foundations of Statistics*. Cambridge University Press, Cambridge.

Levi, I., 1980. *The Enterprise of Knowledge*. MIT Press, London.

Nau, R. F., 1992. Indeterminate probabilities on finite sets. *The Annals of Statistics* **20**,1737–1767.

Robert, C. P., 1994. *The Bayesian Choice*. Springer-Verlag, New York.

Smith, C. A. B., 1961. Consistency in statistical inference and decision. *Journal of the Royal Statistical Society, Series A* **23**, 1–37.

von Winterfeldt, D., Edwards, W., 1986. *Decision Analysis and Behavioral Research*. Cambridge University Press, Cambridge.

Walley, P., 1981. *Coherent Lower (and Upper) Probabilities*, Statistics Research Report 22. University of Warwick, Coventry.

Walley, P., 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.

Walley, P., 1997. Statistical inferences based on a second-order possibility distribution. *International Journal of General Systems* **26**, 337–383.

Watson, S. R., Weiss, J. J., Donnell, M. L., 1979. Fuzzy decision analysis. *IEEE Transactions on Systems, Man, and Cybernetics* **9**, 1–9.

Zadeh, L. A., 1976. The concept of a linguistic variable and its application to approximate reasoning III. *Information Sciences* **9**, 43–80.

Zadeh, L. A., 1984. Fuzzy probabilities. *Information Processing and Management* **20**, 363–372.

Zellner, A., 1971. *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York.

UNIVERSITEIT GENT, ONDERZOEKSGROEP SYSTEMS, TECHNOLOGIEPARK – ZWIJNAARDE 9, 9052 ZWIJNAARDE, BELGIUM

E-mail address: gert.decooman@rug.ac.be