LEXICOGRAPHIC CHOICE FUNCTIONS

ARTHUR VAN CAMP, GERT DE COOMAN, AND ENRIQUE MIRANDA

ABSTRACT. We investigate a generalisation of the coherent choice functions considered by Seidenfeld et al. (2010), by sticking to the convexity axiom but imposing no Archimedeanity condition. We define our choice functions on vector spaces of options, which allows us to incorporate as special cases both Seidenfeld et al.'s (2010) choice functions on horse lotteries and also pairwise choice—which is equivalent to sets of desirable gambles (Quaeghebeur, 2014)—, and to investigate their connections.

We show that choice functions based on sets of desirable options (gambles) satisfy Seidenfeld's convexity axiom only for very particular types of sets of desirable options, which are exactly those that are representable by lexicographic probability systems that have no non-trivial Savage-null events. We call them lexicographic choice functions. Finally, we prove that these choice functions can be used to determine the most conservative convex choice function associated with a given binary relation.

1. INTRODUCTION

Since the publication of the seminal work of Arrow (1951) and Uzawa (1956), coherent choice functions have been used widely as a model of the rational behaviour of an individual or a group. In particular, Seidenfeld et al. (2010) established an axiomatisation of coherent choice functions, generalising Rubin's (1987) axioms to allow for incomparability. Under this axiomatisation, they proved a representation theorem for coherent choice functions in terms of probability-utility pairs: a choice function C satisfies their coherence axioms if and only if there is some non-empty set S of probability-utility pairs such that $f \in C(A)$ whenever the option f maximises p-expected u-utility over the set of options A for some (p,u) in S.

Allowing for incomparability between options may often be of crucial importance. Faced with a choice between two options, a subject may not have enough information to establish a (strict or weak) preference of one over the other: the two options may be incomparable. This will indeed typically be the case when the available information is too vague or limited. It arises quite intuitively for group decisions, but also for decisions made by a single subject, as was discussed quite thoroughly by Williams (1975), Levi (1980), and Walley (1991), amongst many others. Allowing for incomparability lies at the basis of a generalising approach to probability theory that is often referred to by the term *imprecise probabilities*. It unifies a diversity of well-known uncertainty models, including typically non-linear (or non-additive) functionals, credal sets, and sets of desirable gambles; see the introductory book by Augustin et al. (2014) for a recent overview. Among these, coherent sets of desirable gambles, as discussed by Quaeghebeur (2014), are usually considered to constitute the most general and powerful type of model. Such sets collect the gambles that a given subject considers strictly preferable to the status quo.

Nevertheless, choice functions clearly lead to a still more general model than sets of desirable gambles, because the former's preferences are not necessarily completely determined by the pair-wise comparisons between options that essentially constitute the latter. This was of course already implicit in Seidenfeld et al.'s (2010) work, but was investigated in detail in one of our recent papers (Van Camp et al., 2017), where we

Key words and phrases. Choice functions, coherence, lexicographic probabilities, horse lotteries, maximality, preference relations, convexity, sets of desirable gambles.

zoomed in on the connections between choice functions, sets of desirable gambles, and indifference.

In order to explore the connection between indifference and the strict preference expressed by choice functions, we extended the above-mentioned axiomatisation by Seidenfeld et al. (2010) to choice functions defined on vector spaces of options, rather than convex sets of horse lotteries, and also let go of two of their axioms: (i) the Archimedean one, because it prevents choice functions from modelling the typically non-Archimedean preferences captured by coherent sets of desirable gambles; and (ii) the convexity axiom, because it turns out to be hard to reconcile with Walley–Sen maximality as a decision rule, something that is closely tied in with coherent sets of desirable options (Troffaes, 2007). Although our alternative axiomatisation allows for more leeway, and for an easy comparison with the existing theory of sets of desirable gambles, it also has the drawback of no longer forcing a representation theorem, or in other words, of not leading to a strong belief structure (we refer to De Cooman (2005) for a more detailed discussion of belief models that constitute a strong belief structure). Such a representation is nevertheless interesting because strong belief structures have the advantage that their coherent models are infima (under a partial order implicit in the structure) of their dominating *maximally informative* models. This allows for reasoning with the (typically simpler) maximally informative dominating models, instead of the (possibly more complex) models themselves. In an earlier paper (Van Camp et al., 2017), we discussed a few interesting examples of special 'representable' choice functions, such as the ones from a coherent set of desirable gambles via maximality, or those determined by a set of probability measures via E-admissibility.

The goal of the present paper is twofold: to (i) further explore the connection of our definition of choice functions with Seidenfeld et al.'s (2010); and to (ii) investigate in detail the implications of Seidenfeld et al.'s (2010) convexity axiom in our context. We will prove that, perhaps somewhat surprisingly, for those choice functions that are uniquely determined by binary comparisons, convexity is equivalent to being representable by means of a lexicographic probability measure. This is done by first establishing the implications of convexity in terms of the binary comparisons associated with a choice function, giving rise to what we will call *lexicographic sets of desirable gambles*. These sets include as particular cases the so-called *maximal* and *strictly desirable* sets of desirable gambles. Although in the particular case of binary possibility spaces these are the only two possibilities, for more general spaces lexicographic sets of gambles allow for a greater level of generality, as one would expect considering the above-mentioned equivalence.

A consequence of our equivalence result is that we can consider infima of choice functions associated with lexicographic probability measures, and in this manner subsume the examples of E-admissibility and M-admissibility discussed by Van Camp et al. (2017). It will follow from the discussion that these infima also satisfy the convexity axiom. As one particularly relevant application of these ideas, we prove that the most conservative convex choice function associated with a binary preference relation can be obtained as the infimum of its dominating lexicographic choice functions.

The paper is organised as follows. In Section 2, we recall the basics of coherent choice functions on vector spaces of options as introduced in our earlier work (Van Camp et al., 2015). We motivate our definitions by showing in Section 3 that they include in particular coherent choice functions on horse lotteries, considered by Seidenfeld et al.'s (2010), and we discuss in some detail the connection between the rationality axioms considered by Seidenfeld et al. (2010) and ours.

As a particularly useful example, we discuss in Section 4 those choice functions that are determined by binary comparisons. We have already shown before (Van Camp et al., 2017) that this leads to the model of coherent sets of desirable gambles; here we study the implications of including convexity as a rationality axiom.

In Section 5, we motivate our definition of lexicographic choice functions and study the properties of their associated binary preferences. We prove the connection with lexicographic probability systems and show that the infima of such choice functions can be used when we want to determine the implications of imposing convexity and maximality. We conclude with some additional discussion in Section 6.

2. COHERENT CHOICE FUNCTIONS ON VECTOR SPACES

Consider a real vector space $\mathcal{V}$ provided with the vector addition + and scalar multiplication. We denote the additive identity by 0. For any subsets A_1 and A_2 of $\mathcal{V}$ and any λ in $\mathbb{R}$, we let $\lambda A_1 := \{\lambda u : u \in A_1\}$ and $A_1 + A_2 := \{u + v : u \in A_1 \text{ and } v \in A_2\}$.

Elements of $\mathcal{V}$ are intended as abstract representations of *options* amongst which a subject can express his preferences, by specifying choice functions. Often, options will be real-valued maps on some possibility space, interpreted as uncertain rewards—and therefore also called *gambles*. More generally, they can be *vector-valued gambles*: vector-valued maps on the possibility space. We will see further on that by using such vector-valued gambles, we are able to include as a special case *horse lotteries*, the options considered for instance by Seidenfeld et al. (2010). Also, we have shown (Van Camp et al., 2017) that indifference for choice functions can be studied efficiently by also allowing equivalence classes of indifferent gambles as options; these yet again constitute a vector space, where now the vectors cannot always be identified easily with maps on some possibility space, or gambles. For these reasons, we allow in general any real vector space to serve as our set of (abstract) possible options. We will call such a real vector space an *option space*.

We denote by $\mathcal{Q}(\mathcal{V})$ the set of all non-empty *finite* subsets of $\mathcal{V}$, a strict subset of the power set of $\mathcal{V}$. When it is clear what option space $\mathcal{V}$ we are considering, we will also use the simpler notation $\mathcal{Q}$, and use $\mathcal{Q}_0$ to denote those option sets that include the option that is constant on 0. Elements A of $\mathcal{Q}$ are the option sets amongst which a subject can choose his preferred options.

Definition 1. A *choice function* C on an option space $\mathcal{V}$ is a map

$$C\colon \mathcal{Q} \to \mathcal{Q} \cup \{\emptyset\} \colon A \mapsto C(A) \text{ such that } C(A) \subseteq A.$$

We collect all the choice functions on $\mathcal{V}$ in $\mathcal{C}(\mathcal{V})$, often denoted as $\mathcal{C}$ when it is clear from the context what the option space is.

The idea underlying this simple definition is that a choice function C selects the set $C(A)$ of 'best' options in the *option set* A. Our definition resembles the one commonly used in the literature (Aizerman, 1985; Seidenfeld et al., 2010; Sen, 1977), except perhaps for an also not entirely unusual restriction to *finite* option sets (He, 2012; Schwartz, 1972; Sen, 1971).

Equivalently to a choice function C, we may consider its associated *rejection function* R, defined by $R(A) := A \setminus C(A)$ for all A in $\mathcal{Q}$. It returns the options $R(A)$ that are rejected—not selected—by C.

Another equivalent notion is that of a *choice relation*. Indeed, for any choice function C—and therefore for any rejection function R—the associated choice relation (Seidenfeld et al., 2010, Section 3) is the binary relation $\lhd$ on $\mathcal{Q}$, defined by:

$$A_1 \lhd A_2 \Leftrightarrow A_1 \subseteq R(A_1 \cup A_2) \text{ for all } A_1 \text{ and } A_2 \text{ in } \mathcal{Q}. \tag{1}$$

The intuition behind $\lhd$ is clear: $A_1 \lhd A_2$ whenever every option in A_1 is rejected when presented with the options in $A_1 \cup A_2$.

2.1. **Useful basic definitions and notation.**

We call $\mathbb{N}$ the set of all (positive) integers, $\mathbb{R}_{>0}$ the set of all (strictly) positive real numbers, and $\mathbb{R}_{\geq 0} := \mathbb{R}_{>0} \cup \{0\}$.

Given any subset A of $\mathcal{V}$, we define its *linear hull* $\operatorname{span}(A)$ as the set of all finite linear combinations of elements of A:

$$\operatorname{span}(A) := \left\{ \sum_{k=1}^{n} \lambda_k u_k : n \in \mathbb{N}, \lambda_k \in \mathbb{R}, u_k \in A \right\} \subseteq \mathcal{V},$$

its *positive hull* $\operatorname{posi}(A)$ as the set of all positive finite linear combinations of elements of A:

$$\operatorname{posi}(A) := \left\{ \sum_{k=1}^{n} \lambda_k u_k : n \in \mathbb{N}, \lambda_k \in \mathbb{R}_{>0}, u_k \in A \right\} \subseteq \operatorname{span}(A) \subseteq \mathcal{V},$$

and its *convex hull* $\operatorname{CH}(A)$ as the set of convex combinations of elements of A:

$$\operatorname{CH}(A) := \left\{ \sum_{k=1}^{n} \alpha_k u_k : n \in \mathbb{N}, \alpha_k \in \mathbb{R}_{\geq 0}, \sum_{k=1}^{n} \alpha_k = 1, u_k \in A \right\} \subseteq \operatorname{posi}(A) \subseteq \mathcal{V}.$$

A subset A of $\mathcal{V}$ is called a *convex cone* if it is closed under positive finite linear combinations, i.e. if $\operatorname{posi}(A) = A$. A convex cone $\mathcal{K}$ is called *proper* if $\mathcal{K} \cap -\mathcal{K} = \{0\}$.

With any proper convex cone $\mathcal{K} \subseteq \mathcal{V}$, we can associate an ordering $\preceq_{\mathcal{K}}$ on $\mathcal{V}$, defined for all u and v in $\mathcal{V}$ as follows:

$$u \preceq_{\mathcal{K}} v \Leftrightarrow v - u \in \mathcal{K}.$$

We also write $u \succeq_{\mathcal{K}} v$ for $v \preceq_{\mathcal{K}} u$. The ordering $\preceq_{\mathcal{K}}$ is actually a *vector ordering*: it is a partial order—reflexive, antisymmetric and transitive—that satisfies the following two characteristic properties:

$$u_1 \preceq_{\mathcal{K}} u_2 \Leftrightarrow u_1 + v \preceq_{\mathcal{K}} u_2 + v; \tag{2}$$

$$u_1 \preceq_{\mathcal{K}} u_2 \Leftrightarrow \lambda u_1 \preceq_{\mathcal{K}} \lambda u_2,, \tag{3}$$

for all u_1, u_2 and v in $\mathcal{V}$, and all λ in $\mathbb{R}_{>0}$. Observe, by the way, that as a consequence

$$u \preceq_{\mathcal{K}} v \Leftrightarrow 0 \preceq_{\mathcal{K}} v - u \Leftrightarrow u - v \preceq_{\mathcal{K}} 0$$

for all u and v in $\mathcal{V}$.

Conversely, given any vector ordering $\preceq$, the proper convex cone $\mathcal{K}$ from which it is derived can always be retrieved by $\mathcal{K} = \{u \in \mathcal{V} : u \succeq 0\}$. When the abstract options are gambles, $\preceq$ will typically be the point-wise order $\leq$, but it need not necessarily be.

Finally, with any vector ordering $\preceq$, we associate the strict partial ordering $\prec$ as follows:

$$u \prec v \Leftrightarrow (u \preceq v \text{ and } u \neq v) \Leftrightarrow v - u \in \mathcal{K} \setminus \{0\} \text{ for all } u \text{ and } v \text{ in } \mathcal{V}.$$

We call u *positive* if $u \succ 0$, and collect all positive options in the convex cone $\mathcal{V}_{\succ 0} := \mathcal{K} \setminus \{0\}$.

From now on, we assume that $\mathcal{V}$ is an *ordered vector space*, with a generic but fixed vector ordering $\preceq_{\mathcal{K}}$. We will refrain from explicitly mentioning the actual proper convex cone $\mathcal{K}$ we are using, and simply write $\mathcal{V}$ to mean the ordered vector space, and use $\preceq$ as a generic notation for the associated vector ordering.

2.2. **Rationality axioms.** We focus on a special class of choice functions, which we will call *coherent*.

Definition 2. We call a choice function C on $\mathcal{V}$ *coherent* if for all A, A_1 and A_2 in $\mathcal{Q}$, all u and v in $\mathcal{V}$, and all λ in $\mathbb{R}_{>0}$:

C_1. $C(A) \neq \emptyset$;

C_2. if $u \prec v$ then $\{v\} = C(\{u, v\})$;

C_3. a. if $C(A_2) \subseteq A_2 \setminus A_1$ and $A_1 \subseteq A_2 \subseteq A$ then $C(A) \subseteq A \setminus A_1$;

b. if $C(A_2) \subseteq A_2 \setminus A_1$ and $A \subseteq A_1$ then $C(A_2 \setminus A) \subseteq A_2 \setminus A_1$;

C_4. a. if $A_1 \subseteq C(A_2)$ then $\lambda A_1 \subseteq C(\lambda A_2)$;

b. if $A_1 \subseteq C(A_2)$ then $A_1 + \{u\} \subseteq C(A_2 + \{u\})$;

We collect all the coherent choice functions on $\mathcal{V}$ in $\bar{\mathcal{C}}(\mathcal{V})$, often denoted as $\bar{\mathcal{C}}$ when it is clear from the context what the option space is.

Parts C_3a and C_3b of Axiom C_3 are respectively known as *Sen's condition α* and *Aizerman's condition*. They are more commonly written in terms of the rejection function as, respectively:

$$(A_1 \subseteq R(A_2) \text{ and } A_2 \subseteq A) \Rightarrow A_1 \subseteq R(A), \text{ for all } A, A_1, A_2 \text{ in } \mathcal{Q}, \tag{4}$$

and

$$(A_1 \subseteq R(A_2) \text{ and } A \subseteq A_1) \Rightarrow A_1 \setminus A \subseteq R(A_2 \setminus A), \text{ for all } A, A_1, A_2 \text{ in } \mathcal{Q}. \tag{5}$$

Axiom C_4 has multiple equivalent forms, which we will use throughout:

Proposition 1. *Consider any choice function C. Then the following equivalences hold:*

$$\begin{aligned} C \textit{ satisfies Axiom } C_4\text{a} &\Leftrightarrow (\forall A \in \mathcal{Q}, u \in A, \lambda \in \mathbb{R}_{>0})(u \in C(A) \Rightarrow \lambda u \in C(\lambda A)) && (C_4\text{a.1}) \\ &\Leftrightarrow (\forall A \in \mathcal{Q}, u \in A, \lambda \in \mathbb{R}_{>0})(u \in C(A) \Leftrightarrow \lambda u \in C(\lambda A)) && (C_4\text{a.2}) \\ &\Leftrightarrow (\forall A \in \mathcal{Q}, \lambda \in \mathbb{R}_{>0})\lambda C(A) = C(\lambda A), && (C_4\text{a.3}) \end{aligned}$$

$$\begin{aligned} C \textit{ satisfies Axiom } C_4\text{b} &\Leftrightarrow (\forall A \in \mathcal{Q}, u \in A, v \in \mathcal{V})(u \in C(A) \Rightarrow u + v \in C(A + \{v\})) && (C_4\text{b.1}) \\ &\Leftrightarrow (\forall A \in \mathcal{Q}, u \in A, v \in \mathcal{V})(u \in C(A) \Leftrightarrow u + v \in C(A + \{v\})) && (C_4\text{b.2}) \\ &\Leftrightarrow (\forall A \in \mathcal{Q}, v \in \mathcal{V})C(A) + \{v\} = C(A + \{v\}), && (C_4\text{b.3}) \end{aligned}$$

Proof. For the equivalences involving Axiom C_4a, we will establish the following chain of implications:

$$C \text{ satisfies Axiom } C_4\text{a} \Rightarrow (C_4\text{a.1}) \Rightarrow (C_4\text{a.2}) \Rightarrow (C_4\text{a.3}) \Rightarrow C \text{ satisfies Axiom } C_4\text{a}.$$

To establish the first implication, it suffices to consider $A_1 := C(A_2)$ in the statement of Axiom C_4a, and to consider any u in A_1. That (C_4a.1) implies (C_4a.2) follows by considering $\frac{1}{\lambda} > 0$ instead of λ. That then (C_4a.3) holds is immediate. Finally, to show that (C_4a.3) implies Axiom C_4a, it suffices to note that (C_4a.3) implies in particular that $\lambda C(A) \subseteq C(\lambda A)$, and therefore, by considering any $A_1 \subseteq C(A)$, Axiom C_4a indeed follows.

For the equivalences involving Axiom C_4b, we will establish the following chain of implications:

$$C \text{ satisfies Axiom } C_4\text{b} \Rightarrow (C_4\text{b.1}) \Rightarrow (C_4\text{b.2}) \Rightarrow (C_4\text{b.3}) \Rightarrow C \text{ satisfies Axiom } C_4\text{b}.$$

To establish the first implication, it suffices to consider $A_1 := C(A_2)$ in the statement of Axiom C_4b, and to consider any u in A_1. That (C_4b.1) implies (C_4b.2) follows by considering $-v \in \mathcal{V}$ instead of v. That then (C_4b.3) holds is immediate. Finally, to show that (C_4b.3) implies Axiom C_4b, it suffices to note that (C_4b.3) implies that in particular $C(A)+\{v\} \subseteq C(A+\{v\})$, and therefore, by considering any $A_1 \subseteq C(A)$, Axiom C_4b indeed follows. □

These axioms constitute a subset of the ones introduced by Seidenfeld et al. (2010), duly translated from horse lotteries to our abstract options, which are more general as we will show in Section 3 further on. In this respect, our notion of coherence is less restrictive than theirs. On the other hand, our Axiom C_2 is more restrictive than Seidenfeld et al.'s (2010). This is necessary for the link between coherent choice functions and coherent sets of desirable gambles we will establish in Section 4.

One axiom we omit from our coherence definition, is the Archimedean one. Typically the preference associated with coherent sets of desirable gambles does not have the Archimedean property (Zaffalon and Miranda, 2015, Section 3), so letting go of this axiom is necessary if we want to explore the connection with desirability.

The second axiom that we do not consider as necessary for coherence is what we will call the *convexity axiom*:

C_5. if $A \subseteq A_1 \subseteq \mathrm{CH}(A)$ then $C(A) \subseteq C(A_1)$, for all A and A_1 in $\mathcal{Q}$.

As we will show in Section 4, it is incompatible with Walley–Sen maximality (Walley, 1991; Troffaes, 2007) as a decision rule. Nevertheless, we intend to investigate the connection with desirability for coherent choice functions that do satisfy the convexity axiom.

Two dominance properties are immediate consequences of coherence:

Proposition 2. *Let C be a coherent choice function on $\mathcal{Q}$. Then for all u_1 and u_2 in $\mathcal{V}$ such that $u_1 \leq u_2$, all A in $\mathcal{Q}$ and all v in $A \setminus \{u_1, u_2\}$:*

a. *if $u_2 \in A$ and $v \notin C(A \cup \{u_1\})$ then $v \notin C(A)$;*
b. *if $u_1 \in A$ and $v \notin C(A)$ then $v \notin C(\{u_2\} \cup A \setminus \{u_1\})$.*

Proof. The result is trivial when $u_1 = u_2$, so let us assume that $u_1 < u_2$.

The first statement is again trivial if $u_1 \in A$. When $u_1 \notin A$, it follows from Axiom C_2 that $u_1 \notin C(\{u_1, u_2\})$. By applying Axiom C_3a in the form of Equation (4), we find that $u_1 \notin C(A \cup \{u_1\})$, and then applying Axiom C_3b in the form of Equation (5), together with the assumption that $v \notin C(A \cup \{u_1\})$, we conclude that $v \notin C(A \cup \{u_1\} \setminus \{u_1\}) = C(A)$.

For the second statement, it follows from Axiom C_2 that $u_1 \notin C(\{u_1, u_2\})$. By applying Axiom C_3a in the form of Equation (4), we find that both $u_1 \notin C(A \cup \{u_2\})$ and $v \notin C(A \cup \{u_2\})$, so we can apply Axiom C_3b in the form of Equation (5) to conclude that $v \notin C(\{u_2\} \cup A \setminus \{u_1\})$. □

We are interested in conservative reasoning with choice functions. We therefore introduce a binary relation $\sqsubseteq$ on the set $\mathcal{C}$ of all choice functions, having the interpretation of 'not more informative than', or, in other words, 'at least as uninformative as'.

Definition 3. Given two choice functions C_1 and C_2 in $\mathcal{C}$, we call C_1 *not more informative than* C_2—and we write $C_1 \sqsubseteq C_2$—if $(\forall A \in \mathcal{Q}) C_1(A) \supseteq C_2(A)$.

This intuitive way of ordering choice functions is also used by Bradley (2015, Section 2) and Van Camp et al. (2017, Definition 6). The underlying idea is that a choice function is more informative when it consistently chooses more specifically—or more restrictively—amongst the available options.

Since, by definition, $\sqsubseteq$ is a product ordering of set inclusions, the following result is immediate (Davey and Priestley, 1990).

Proposition 3. *The structure $(\mathcal{C}; \sqsubseteq)$ is a complete lattice:*

(i) *it is a* partially ordered set, *or* poset, *meaning that the binary relation $\sqsubseteq$ on $\mathcal{C}$ is* reflexive, antisymmetric *and* transitive*;*
(ii) *for any subset $\mathcal{C}'$ of $\mathcal{C}$, its infimum $\inf \mathcal{C}'$ and its supremum $\sup \mathcal{C}'$ with respect to the ordering $\sqsubseteq$ exist in $\mathcal{C}$, and are given by $\inf \mathcal{C}'(A) = \bigcup_{C \in \mathcal{C}'} C(A)$ and $\sup \mathcal{C}'(A) = \bigcap_{C \in \mathcal{C}'} C(A)$ for all A in $\mathcal{Q}$.*

The idea underlying these notions of infimum and supremum is that $\inf \mathcal{C}'$ is the most informative model that is not more informative than any of the models in $\mathcal{C}'$, and $\sup \mathcal{C}'$ the least informative model that is not less informative than any of the models in $\mathcal{C}'$.

We have proved elsewhere (Van Camp et al., 2017, Proposition 3) that coherence is preserved under arbitrary non-empty infima. Because of our interest in the additional Axiom C_5, we prove that it also is preserved under arbitrary non-empty infima.

Proposition 4. *Given any non-empty collection $\mathcal{C}'$ of choice functions that satisfy Axiom C_5, its infimum $\inf \mathcal{C}'$ satisfies Axiom C_5 as well.*

Proof. Consider any A and A_1 in $\mathcal{Q}$ such that $A \subseteq A_1 \subseteq \mathrm{CH}(A)$. Then $C(A) \subseteq C(A_1)$ for all C in $\mathcal{C}'$, whence $\inf \mathcal{C}'(A) = \bigcup_{C \in \mathcal{C}'} C(A) \subseteq \bigcup_{C \in \mathcal{C}'} C(A_1) = \inf \mathcal{C}'(A_1)$. □

3. THE CONNECTION WITH OTHER DEFINITIONS OF CHOICE FUNCTIONS

Before we go on with our exploration of choice functions, let us take some time here to explain why we have chosen to define them in the way we did. Seidenfeld et al. (2010) (see also Kadane et al., 2004) define choice functions on *horse lotteries*, instead of options, as this helps them generalise the framework established by Anscombe and Aumann (1963) for binary preferences to non-binary ones.

One reason for our working with the more abstract notion of options—elements of some general vector space—is that they are better suited for dealing with indifference: this involves working with equivalence classes of options, which again constitute a vector space (Van Camp et al., 2017). These equivalence classes can no longer be interpreted easily or directly as gambles, or horse lotteries for that matter. Another reason for using options that are more general than real-valued gambles is that recent work by Zaffalon and Miranda (2015) has shown that a very general theory of binary preference can be constructed using vector-valued gambles, rather than horse lotteries. Such vector-valued gambles again constitute a real vector, or option, space. Here, we show that the conclusions of Zaffalon and Miranda (2015, Section 4) can be extended from binary preferences to choice functions.

We consider an arbitrary possibility space $\mathcal{X}$ of mutually exclusive elementary events, one of which is guaranteed to occur. Consider also a countable set $\mathcal{R}$ of prizes, or rewards.

Definition 4 (Gambles)**.** Any bounded real-valued function on some domain $\mathcal{X}$ is called a *gamble* on $\mathcal{X}$. We collect all gambles on $\mathcal{X}$ in $\mathcal{L}(\mathcal{X})$, often denoted as $\mathcal{L}$ when it is clear from the context what the domain $\mathcal{X}$ is.

When the domain is of the type $\mathcal{X}\times\mathcal{R}$, we call elements f of $\mathcal{L}(\mathcal{X}\times\mathcal{R})$ *vector-valued gambles* on $\mathcal{X}$. Indeed, for each x in $\mathcal{X}$, the partial map $f(x,\cdot)$ is then an element of the vector space $\mathcal{L}(\mathcal{R})$.

The set $\mathcal{L}$, provided with the point-wise addition of gambles, the point-wise multiplication with real scalars, and the point-wise vector ordering $\leq$, constitutes an ordered vector space. We call $\mathcal{L}_{>0} \coloneqq \{f\in\mathcal{L} : f>0\} = \{f\in\mathcal{L} : f\geq 0 \text{ and } f\neq 0\}$ the set of all *positive* (vector-valued) gambles.

Horse lotteries are special vector-valued gambles.

Definition 5 (Horse lotteries)**.** We call *horse lottery* H any map from $\mathcal{X}\times\mathcal{R}$ to $[0,1]$ such that for all x in $\mathcal{X}$, the partial map $H(x,\cdot)$ is a probability mass function over $\mathcal{R}$:[1]

$$(\forall x\in\mathcal{X})\Bigg(\sum_{r\in\mathcal{R}} H(x,r)=1 \text{ and } (\forall r\in\mathcal{R})H(x,r)\geq 0\Bigg).$$

We collect all the horse lotteries on $\mathcal{X}$ with reward set $\mathcal{R}$ in $\mathcal{H}(\mathcal{X},\mathcal{R})$, which is also denoted more simply by $\mathcal{H}$ when it is clear from the context what the possibility space $\mathcal{X}$ and reward set $\mathcal{R}$ are.

Let us, for the remainder of this section, fix $\mathcal{X}$ and $\mathcal{R}$. It is clear that $\mathcal{H}\subseteq\mathcal{L}(\mathcal{X}\times\mathcal{R})$. Seidenfeld et al. (2010) consider choice functions whose domain is $\mathcal{Q}(\mathcal{H})$, the set of all finite subsets of $\mathcal{H}$—choice functions on horse lotteries.[2] We will call them *choice functions on* $\mathcal{H}$. Because of the nature of $\mathcal{H}$, their choice functions are different from ours: they require slightly different rationality axioms. The most significant change is that for Seidenfeld et al. (2010), choice functions need not satisfy Axioms C_4a and C_4b. In fact, choice functions on $\mathcal{H}$ cannot satisfy these axioms, since $\mathcal{H}$ is no linear space: it is not closed under arbitrary linear combinations, only under *convex combinations*. Instead, on their approach a choice function C^* on $\mathcal{H}$ is required to satisfy

C_4^*. $A_1^*\lhd_{C^*} A_2^* \Leftrightarrow \alpha A_1^*+(1-\alpha)\{H\}\lhd_{C^*}\alpha A_2^*+(1-\alpha)\{H\}$ for all α in $(0,1]$, all A_1^* and A_2^* in $\mathcal{Q}(\mathcal{H})$ and all H in $\mathcal{H}$.

[1] Note that $H(x,\cdot)$ defines a (sigma-)additive probability measure, and that this countable additivity property is necessary for Lemma 6 below to hold.

[2] Actually, Seidenfeld et al. (2010) define choice functions on a larger domain: all possibly infinite but *closed* sets of horse lotteries (non-closed sets may not have admissible options). This is an extension we see no need for in the present context.

The binary relation $\lhd_{C^*}$ is the choice relation associated with C^*, defined by Equation (1). Furthermore, for a choice function C^* to be coherent, it needs to additionally satisfy (Seidenfeld et al., 2010):

C_1^*. $C^*(A^*) \neq \emptyset$ for all A^* in $\mathcal{Q}(\mathcal{H})$;

C_2^*. for all A^* in $\mathcal{Q}(\mathcal{H})$, all H_1 and H_2 in $\mathcal{H}$ such that $H_1(\cdot,\top) \leq H_2(\cdot,\top)$ and $H_1(\cdot,r) = H_2(\cdot,r) = 0$ for all r in $\mathcal{R} \setminus \{\bot,\top\}$, and all H in $\mathcal{H} \setminus \{H_1, H_2\}$:
  a. if $H_2 \in A^*$ and $H \in R^*(\{H_1\} \cup A^*)$ then $H \in R^*(A^*)$;
  b. if $H_1 \in A^*$ and $H \in R^*(A^*)$ then $H \in R^*(\{H_2\} \cup A^* \setminus \{H_1\})$;

C_3^*. for all A^*, A_1^* and A_2^* in $\mathcal{Q}(\mathcal{H})$:
  a. if $A_1^* \subseteq R^*(A_2^*)$ and $A_2^* \subseteq A^*$ then $A_1^* \subseteq R^*(A)$;
  b. if $A_1^* \subseteq R^*(A_2^*)$ and $A^* \subseteq A_1^*$ then $A_1^* \setminus A^* \subseteq R^*(A_2^* \setminus A)$;

C_5^*. if $A^* \subseteq A_1^* \subseteq \mathrm{CH}(A)$ then $C^*(A) \subseteq C^*(A_1^*)$, for all A^* and A_1^* in $\mathcal{Q}(\mathcal{H})$;

C_6^*. for all A^*, $A^{*\prime}$, $A^{*\prime\prime}$ $A_i^{*\prime}$ and $A_i^{*\prime\prime}$ (for i in $\mathbb{N}$) in $\mathcal{Q}(\mathcal{H})$ such that the sequence $A_i^{*\prime}$ converges point-wise to $A^{*\prime}$ and the sequence $A_i^{*\prime\prime}$ converges point-wise to $A^{*\prime\prime}$:
  a. If $(\forall i \in \mathbb{N}) A_i^{*\prime\prime} \lhd_{C^*} A_i^{*\prime}$ and $A^{*\prime} \lhd_{C^*} A^*$ then $A^{*\prime\prime} \lhd_{C^*} A^*$;
  b. If $(\forall i \in \mathbb{N}) A_i^{*\prime\prime} \lhd_{C^*} A_i^{*\prime}$ and $A^* \lhd_{C^*} A^{*\prime\prime}$ then $A^* \lhd_{C^*} A^{*\prime}$,

where Seidenfeld et al. (2010) assume that there is a unique worst reward $\bot$ and a unique best reward $\top$ in $\mathcal{R}$. This is a somewhat stronger assumption than we will make: further on in this section, we will only assume that there is a unique worst reward. Axiom C_2^* is the counterpart of Proposition 2 for choice functions on horse lotteries, which is a result of our Axioms C_1–C_4. Seidenfeld et al. (2010) need to impose this property as an axiom, essentially because of the absence in their system of a counterpart for our Axiom C_2. Axioms C_6^*a and C_6^*b are Archimedean axioms, hard to reconcile with desirability (see for instance Zaffalon and Miranda, 2015, Section 4), which is why will not enforce them here.

We now intend to show that under very weak conditions on the rewards set $\mathcal{R}$, choice functions on horse lotteries that satisfy C_4^* are in a one-to-one correspondence with choice functions on a suitably defined option space that satisfy Axioms C_4a and C_4b.

Let us first study the impact of Axiom C_4^*. We begin by showing that an assessment of $H \in C(A)$ for some A in $\mathcal{Q}(\mathcal{H})$ implies other assessments of this type.

Proposition 5. *Consider any choice function C^* on $\mathcal{Q}(\mathcal{H})$ that satisfies Axiom C_4^*, any option sets A^* and $A^{*\prime}$ in $\mathcal{Q}(\mathcal{H})$, and any H in A^* and H' in $A^{*\prime}$. If there are λ and λ' in $\mathbb{R}_{>0}$ such that $\lambda(A^* - \{H\}) = \lambda'(A^{*\prime} - \{H'\})$, then*

$$H \in C^*(A) \Leftrightarrow H' \in C^*(A^{*\prime}).$$

Proof. Fix A^* and $A^{*\prime}$ in $\mathcal{Q}(\mathcal{H})$, H in A^* and H' in $A^{*\prime}$, λ and λ' in $\mathbb{R}_{>0}$, and assume that $\lambda(A^* - \{H\}) = \lambda'(A^{*\prime} - \{H'\})$. We will show that $H \in R^*(A^*) \Leftrightarrow H' \in R^*(A^{*\prime})$. We infer from the assumption that

$$\frac{\lambda}{\lambda+\lambda'} A^* + \frac{\lambda'}{\lambda+\lambda'}\{H'\} = \frac{\lambda'}{\lambda+\lambda'} A^{*\prime} + \frac{\lambda}{\lambda+\lambda'}\{H\}.$$

If we call $\alpha \coloneqq \frac{\lambda}{\lambda+\lambda'}$ to ease the notation along, then $1-\alpha = \frac{\lambda'}{\lambda+\lambda'}$ and $\alpha \in (0,1)$. We now infer from the identity above that $\alpha A^* + (1-\alpha)\{H'\} = (1-\alpha)A^{*\prime} + \alpha\{H\}$. Consider the following chain of equivalences:

$$\begin{aligned}
H \in R^*(A^*) &\Leftrightarrow \{H\} \lhd_{C^*} A^* && \text{by Equation (1)}\\
&\Leftrightarrow \alpha\{H\} + (1-\alpha)\{H'\} \lhd_{C^*} \alpha A^* + (1-\alpha)\{H'\} && \text{using Axiom } C_4^*\\
&\Leftrightarrow \alpha\{H\} + (1-\alpha)\{H'\} \lhd_{C^*} (1-\alpha)A^{*\prime} + \alpha\{H\} &&\\
&\Leftrightarrow \{H'\} \lhd_{C^*} A^{*\prime} && \text{using Axiom } C_4^*\\
&\Leftrightarrow H' \in R^*(A^{*\prime}) && \text{by Equation (1).} \quad \square
\end{aligned}$$

For any r in $\mathcal{R}$, we now introduce $\mathcal{R}_r \coloneqq \mathcal{R} \setminus \{r\}$, the set of all rewards without r. For the connection between choice functions on $\mathcal{H}$ and choice functions on some option space, we need to somehow be able to extend $\mathcal{H}$ to a linear space. The so-called gamblifier φ_r will play a crucial role in this:

Definition 6 (Gamblifier)**.** Consider any r in $\mathcal{R}$. The *gamblifier* φ_r is the linear map

$$\varphi_r \colon \mathcal{L}(\mathcal{X} \times \mathcal{R}) \to \mathcal{L}(\mathcal{X} \times \mathcal{R}_r) \colon f \mapsto \varphi_r f,$$

where $\varphi_r f(x,s) \coloneqq f(x,s)$ for all x in $\mathcal{X}$ and s in $\mathcal{R}_r$.

In particular, the gamblifier φ_r maps any horse lottery H in $\mathcal{H}(\mathcal{X},\mathcal{R})$ to an element $\varphi_r H$ of $\mathcal{L}(\mathcal{X} \times \mathcal{R}_r)$ that satisfies the following two conditions:

$$\varphi_r H(\cdot,\cdot) \geq 0 \text{ and } \sum_{s \in \mathcal{R}_r} \varphi_r H(\cdot,s) \leq 1. \tag{6}$$

Application of φ_r to sets of the form $\lambda(A^* - \{H\})$ essentially leaves the 'information' they contain unchanged:

Lemma 6. *Consider any r in $\mathcal{R}$. Then the following two properties hold:*

(i) *The gamblifier φ_r is one-to-one on $\mathcal{H}$.*
(ii) *For any A^* and $A^{*\prime}$ in $\mathcal{Q}(\mathcal{H})$, any H in A^* and H' in $A^{*\prime}$ and any λ and λ' in $\mathbb{R}_{>0}$:*

$$\lambda(A^* - \{H\}) = \lambda'(A^{*\prime} - \{H'\}) \Leftrightarrow \varphi_r(\lambda(A^* - \{H\})) = \varphi_r(\lambda'(A^{*\prime} - \{H'\})).$$

Proof. We begin with the first statement. Consider any H and H' in $\mathcal{H}$, and assume that $\varphi_r(H) = \varphi_r(H')$. We infer from Definition 6 that

$$H(x,s) = H'(x,s) \text{ for all } x \text{ in } \mathcal{X} \text{ and } s \text{ in } \mathcal{R}_r,$$

and therefore also, since H and H' are horse lotteries,

$$H(x,r) = 1 - \sum_{s \in \mathcal{R}_r} H(x,s) = 1 - \sum_{s \in \mathcal{R}_r} H'(x,s) = H'(x,r) \text{ for all } x \text{ in } \mathcal{X}.$$

Hence indeed $H = H'$.

The direct implication in the second statement is trivial; let us prove the converse. Assume that $\varphi_r(\lambda(A^* - \{H\})) = \varphi_r(\lambda'(A^{*\prime} - \{H'\}))$. We may write, without loss of generality, that $A^* = \{H, H_1, \dots, H_n\}$ and $A^{*\prime} = \{H', H'_1, \dots, H'_m\}$ for some n and m in $\mathbb{N}$. Now, consider any element H_i in A^*, then $\varphi_r(\lambda(H_i - H)) \in \varphi_r(\lambda(A^* - \{H\}))$. Consider any j in $\{1,\dots,m\}$ such that $\varphi_r(\lambda(H_i - H)) = \varphi_r(\lambda'(H'_j - H'))$. It follows from the assumption that there is at least one such j. The proof is complete if we can show that $\lambda(H_i - H) = \lambda'(H'_j - H')$. By Definition 6, we already know that

$$\lambda(H_i(\cdot,s) - H(\cdot,s)) = \lambda'(H'_j(\cdot,s) - H'(\cdot,s)) \text{ for all } s \text{ in } \mathcal{R}_r,$$

and therefore, since H, H', H_i and H'_j are horse lotteries, also

$$\begin{aligned}\lambda(H_i(\cdot,r) - H(\cdot,r)) &= \lambda\Big(\sum_{s \in \mathcal{R}_r} H(\cdot,s) - \sum_{s \in \mathcal{R}_r} H_i(\cdot,s)\Big) = \sum_{s \in \mathcal{R}_r} \lambda(H(\cdot,s) - H_i(\cdot,s)) \\ &= \sum_{s \in \mathcal{R}_r} \lambda'(H'(\cdot,s) - H'_j(\cdot,s)) = \lambda'\Big(\sum_{s \in \mathcal{R}_r} H'(\cdot,s) - \sum_{s \in \mathcal{R}_r} H'_j(\cdot,s)\Big) \\ &= \lambda'(H'_j(\cdot,r) - H'(\cdot,r)),\end{aligned}$$

whence indeed $\lambda(H_i - H) = \lambda'(H'_j - H')$. □

We now lift the gamblifier φ_r to a map $\tilde{\varphi}_r$ that turns choice functions on gambles into choice functions on horse lotteries:

$$\tilde{\varphi}_r \colon \mathcal{C}(\mathcal{L}(\mathcal{X} \times \mathcal{R}_r)) \to \mathcal{C}(\mathcal{H}(\mathcal{X},\mathcal{R})) \colon C \mapsto \tilde{\varphi}_r C, \tag{7}$$

where $\tilde{\varphi}_r C(A^*) := \varphi_r^{-1} C(\varphi_r A^*)$ for every A^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$. This definition makes sense because we have proved in Lemma 6 that φ_r is one-to-one on $\mathcal{H}$, and therefore invertible on $\varphi_r\mathcal{H}$. The result of applying $\tilde{\varphi}_r$ to a choice function C on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ is a choice function $\tilde{\varphi}_r C$ on $\mathcal{H}(\mathcal{X},\mathcal{R})$. Observe that we can equally well make $\tilde{\varphi}_r$ apply to rejection functions R, and that for every A^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$:

$$\tilde{\varphi}_r R(A^*) := \varphi_r^{-1} R(\varphi_r A^*) = \varphi_r^{-1}(\varphi_r A^* \setminus C(\varphi_r A^*)) = A^* \setminus \varphi_r^{-1} C(\varphi_r A^*) = A^* \setminus \tilde{\varphi}_r C(A^*),$$

so $\tilde{\varphi}_r R$ is the rejection function associated with the choice function $\tilde{\varphi}_r C$, when R is the rejection function for C.

One property of the transformation $\tilde{\varphi}_r$ that will be useful in our subsequent proofs is the following:

Lemma 7. *Consider any r in $\mathcal{R}$ and any A in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$, and define g by $g(x,s) := \sum_{f\in A}|f(x,s)|$ for all x in $\mathcal{X}$ and s in $\mathcal{R}_r$. Consider any λ in $\mathbb{R}$ such that*

$$\lambda > \max\left\{\max_{x\in\mathcal{X}} \sum_{s\in\mathcal{R}_r} h(x,s) : h \in A + \{g\}\right\} \geq 0.$$

Then $\frac{1}{\lambda}(A+\{g\}) = \varphi_r A^*$ for some A^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$.

Proof. Consider any h in $A+\{g\}$, and let us show that $\frac{1}{\lambda}h$ satisfies the conditions in Equation (6). The first one is satisfied because $\lambda > 0$ and $h = f+g$ for some f in A, so $h = f+g = f+\sum_{f'\in A}|f'| \geq f+|f| \geq 0$ and therefore indeed $\frac{1}{\lambda}h \geq 0$. For the second condition, recall that $\lambda \geq \sum_{s\in\mathcal{R}_r} h(\cdot,s)$ by construction and therefore indeed $\sum_{s\in\mathcal{R}_r}\frac{1}{\lambda}h(\cdot,s) \leq 1$. □

Proposition 8. *Consider any r in $\mathcal{R}$. The operator $\tilde{\varphi}_r$ is a bijection between the choice functions on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ that satisfy Axioms* C_4a *and* C_4b*, and the choice functions on $\mathcal{H}(\mathcal{X},\mathcal{R})$ that satisfy Axiom* C_4^*.

Proof. We first show that $\tilde{\varphi}_r$ is injective. Assume *ex absurdo* that it is not, so there are choice functions C and C' on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ that satisfy Axioms C_4a and C_4b, such that $\tilde{\varphi}_r C = \tilde{\varphi}_r C'$ but nevertheless $C \neq C'$. The latter means that there are A in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$ and f in A such that $f \in C(A)$ and $f \notin C'(A)$. Use Lemma 7 to find some λ in $\mathbb{R}_{>0}$ and g in $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ such that $\frac{1}{\lambda}(A+\{g\}) = \varphi_r A^*$ for some A^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$. If we now apply Axioms C_4a and C_4b we find that $\frac{f+g}{\lambda} \in C(\frac{1}{\lambda}(A+\{g\}))$, or equivalently, $\varphi_r^{-1}(\frac{f+g}{\lambda}) \in \tilde{\varphi}_r C(A^*)$. Similarly, we find that $\frac{f+g}{\lambda} \notin C'(\frac{1}{\lambda}(A+\{g\}))$, or equivalently, $\varphi_r^{-1}(\frac{f+g}{\lambda}) \notin \tilde{\varphi}_r C'(A^*)$. But this contradicts our assumption that $\tilde{\varphi}_r C = \tilde{\varphi}_r C'$.

We now show that application of $\tilde{\varphi}_r$ to any choice function C on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ that satisfies Axioms C_4a and C_4b, results in a choice function $\tilde{\varphi}_\perp C$ that satisfies Axiom C_4^*. Consider any A_1^* and A_2^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$, any H in $\mathcal{H}(\mathcal{X},\mathcal{R})$, and any α in $(0,1]$. Infer the following chain of equivalences:

$$
\begin{aligned}
&A_1^* \lhd_{\tilde{\varphi}_r C} A_2^* \\
&\quad\Leftrightarrow A_1^* \subseteq \tilde{\varphi}_r R(A_1^* \cup A_2^*) && \text{by Equation (1)}\\
&\quad\Leftrightarrow \varphi_r A_1^* \subseteq R(\varphi_r(A_1^* \cup A_2^*)) && \text{by Equation (7)}\\
&\quad\Leftrightarrow \varphi_r \alpha A_1^* \subseteq R(\varphi_r \alpha(A_1^* \cup A_2^*)) && \text{by Axiom } C_4\text{a}\\
&\quad\Leftrightarrow \varphi_r(\alpha A_1^* + (1-\alpha)\{H\}) \subseteq R(\varphi_r(\alpha(A_1^* \cup A_2^*) + (1-\alpha)\{H\})) && \text{by Axiom } C_4\text{b}\\
&\quad\Leftrightarrow \alpha A_1^* + (1-\alpha)\{H\} \subseteq \tilde{\varphi}_r R(\alpha(A_1^* \cup A_2^*) + (1-\alpha)\{H\}) && \text{by Equation (7)}\\
&\quad\Leftrightarrow (\alpha A_1^* + (1-\alpha)\{H\}) \lhd_{\tilde{\varphi}_r C} (\alpha A_2^* + (1-\alpha)\{H\}) && \text{by Equation (1)},
\end{aligned}
$$

which tells us that indeed $\tilde{\varphi}_r C$ satisfies Axiom C_4^*.

The proof is complete if we also show that $\tilde{\varphi}_r$ is surjective—that for every choice function C^* on $\mathcal{H}(\mathcal{X},\mathcal{R})$ that satisfies Axiom C_4^*, there is a choice function C on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ that satisfies Axioms C_4a and C_4b such that $\tilde{\varphi}_r C = C^*$. So consider any choice function C^* on $\mathcal{H}(\mathcal{X},\mathcal{R})$ that satisfies Axiom C_4^*. We will show that the special choice function C on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ based on C^*, defined as

$$f \in C(A) \Leftrightarrow (\exists\lambda\in\mathbb{R}_{>0}, A^*\in\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R})), H\in A^*) \\ \big(\varphi_r(A^*-\{H\}) = \lambda(A-\{f\}) \text{ and } H\in C^*(A^*)\big) \quad (8)$$

for all A in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$ and f in A, satisfies Axioms C_4a and C_4b and $\tilde{\varphi}_r C = C^*$. We first show that C satisfies Axioms C_4a and C_4b. For Axiom C_4a, we use its equivalent form (C_4a.1). Consider any A in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$, any f in A, and any μ in $\mathbb{R}_{>0}$, and assume that $f\in C(A)$. To show that then $\mu f\in C(\mu A)$, it suffices to consider $\lambda' := \frac{\lambda}{\mu}$ in Equation (8), and note that $\lambda(A-\{f\}) = \lambda'(\mu A-\{\mu f\})$; then the desired statement follows at once from Equation (8). For Axiom C_4b, we use its equivalent form (C_4b.1). Consider any A in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$, any f in A, and any g in $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$, and assume that $f\in C(A)$. We show that then $f+g\in C(A+\{g\})$. To this end, it suffices to note that $A-\{f\} = (A+\{g\})-\{f+g\}$; then the desired statement follows at once from Equation (8). So C as defined in Equation (8) does indeed satisfy Axioms C_4a and C_4b, and is therefore a suitable candidate for showing that $\tilde{\varphi}_r C = C^*$.

We now finish the proof by showing that $\tilde{\varphi}_r C = C^*$. To do so, we consider any A^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$, and show that $\tilde{\varphi}_r C(A^*)\subseteq C^*(A^*)$ and $C^*(A^*)\subseteq\tilde{\varphi}_r C(A^*)$. To show that $\tilde{\varphi}_r C(A^*)\subseteq C^*(A^*)$, consider any H in $\tilde{\varphi}_r C(A^*)$. By the definition of $\tilde{\varphi}_r$ (Equation (7)) then $H\in\varphi_r^{-1}(C(\varphi_r A^*))$, and therefore $\varphi_r H\in C(\varphi_r A^*)$. Using Equation (8) [with $A = \varphi_r A^*$ and $f = \varphi_r H$], we find that then

$$\varphi_r(A^{*\prime}-\{H'\}) = \lambda'(\varphi_r A^*-\{\varphi_r H\}) \text{ and } H'\in C^*(A^{*\prime})$$

for some λ' in $\mathbb{R}_{>0}$, $A^{*\prime}$ in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$ and $H'\in A^{*\prime}$. By Lemma 6 and since $\lambda'(\varphi_r A^* - \{\varphi_r H\}) = \varphi_r(\lambda'(A^*-\{H\}))$, infer that then

$$A^{*\prime}-\{H'\} = \lambda'(A^*-\{H\}) \text{ and } H'\in C^*(A^{*\prime}),$$

and because C^* satisfies Axiom C_4^*, and using Proposition 5 this means that indeed $H\in C^*(A^*)$. So we have shown that $H\in C^*(A^*)$, and since the choice of H was arbitrary in $\tilde{\varphi}_r C(A^*)$, therefore indeed $\tilde{\varphi}_r C(A^*)\subseteq C^*(A^*)$.

To show that $C^*(A^*)\subseteq\tilde{\varphi}_r C(A^*)$, consider any H in $C^*(A^*)$. Let $A := \varphi_r A^*$, $f := \varphi_r H$ and $\lambda := 1$, then

$$\varphi_r(A^*-\{H\}) = \lambda(A-\{f\}),$$

whence by Equation (8) $\varphi_r H = f\in C(A) = C(\varphi_r A^*)$. Since φ_r is one-to-one on $\mathcal{H}(\mathcal{X},\mathcal{R})$ (see Lemma 6), therefore indeed $H\in\tilde{\varphi}_r C(A^*)$. So we have shown that $\tilde{\varphi}_r C = C^*$, which completes the proof. □

Specifying a choice function C^* on $\mathcal{H}$ induces a strict preference relation on the reward set, as follows. With any reward r in $\mathcal{R}$ we can associate the constant and degenerate lottery H_r by letting

$$H_r(x,s) := \begin{cases} 1 & \text{if } s = r \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } x \text{ in } \mathcal{X} \text{ and } s \text{ in } \mathcal{R}.$$

This is the lottery that associates the certain reward r with all states. Then a reward r is strictly preferred to a reward s when $H_s\in R^*(\{H_r,H_s\})$.

Definition 7 (C^* has worst reward r)**.** Consider any reward r in $\mathcal{R}$, and any choice function C^* on $\mathcal{H}(\mathcal{X},\mathcal{R})$. We say that C^* *has worst reward* r if $H_r\in R^*(\{H,H_r\})$ for all H in $\mathcal{H}(\mathcal{X},\mathcal{R})\setminus\{H_r\}$.

The worst reward is unique when C^* satisfies Axiom C_1^*: indeed, if there are two different worst rewards r and s, by Definition 7 then $\{H_r,H_s\} = R^*(\{H_r,H_s\})$, contradicting Axiom C_1^*.

The notion of *having worst reward* is closely related with what would be the natural translation of Axiom C_2 to choice functions C^* on $\mathcal{H}(\mathcal{X},\mathcal{R})$: if C^* satisfies

$$(\forall H_1,H_2\in\mathcal{H})\Big(\big(H_1\neq H_2 \text{ and } (\forall s\in\mathcal{R}_r)(H_1(\cdot,s)\leq H_2(\cdot,s))\big)\Rightarrow H_1\in R^*(\{H_1,H_2\})\Big) \tag{9}$$

for some r in $\mathcal{R}$, then we say that C^* *satisfies the dominance relation for worst reward* r.

Proposition 9. *Consider any r in $\mathcal{R}$ and any choice function C on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ that satisfies Axiom* C_4b. *Then $\tilde{\varphi}_rC$ satisfies the dominance relation for worst reward r (Equation* (9)*) if and only if $\tilde{\varphi}_rC$ has worst reward r.*

Proof. For the direct implication, consider any H in $\mathcal{H}(\mathcal{X},\mathcal{R})\setminus\{H_r\}$. Then $H_r(\cdot,s)=0\leq H(\cdot,s)$ for all s in $\mathcal{R}_r$, and also $H\neq H_r$, whence indeed $H_r\in\tilde{\varphi}_rR(\{H,H_r\})$, because by assumption $\tilde{\varphi}_rC$ satisfies Equation (9) for r.

For the converse implication, consider any H_1 and H_2 in $\mathcal{H}(\mathcal{X},\mathcal{R})$ such that $H_1\neq H_2$ and $H_1(\cdot,s)\leq H_2(\cdot,s)$ for all s in $\mathcal{R}_r$. Then $\varphi_rH_1<\varphi_rH_2$, whence $0<\varphi_r(H_2-H_1)$. Observe that for the horse lottery H' in $\mathcal{H}(\mathcal{X},\mathcal{R})$ defined by

$$H'(\cdot,s):=\begin{cases}H_2(\cdot,s)-H_1(\cdot,s) & \text{if } s\in\mathcal{R}_r\\ 1-\sum_{s\in\mathcal{R}_r}\big(H_2(\cdot,s)-H_1(\cdot,s)\big) & \text{if } s=r,\end{cases}$$

we have that $\varphi_rH'=\varphi_r(H_2-H_1)$. Because $\tilde{\varphi}_rC$ is assumed to have worst reward r, we know that in particular $H_r\in\tilde{\varphi}_rR(\{H',H_r\})$, so we infer from Equation (7) that $0=\varphi_rH_r\in R(\{\varphi_rH_r,\varphi_rH'\})=R(\{0,\varphi_rH_2-\varphi_rH_1\})$. Now use Axiom C_4b to infer that $\varphi_rH_1\in R(\{\varphi_rH_1,\varphi_rH_2\})$, whence indeed $H_1\in\tilde{\varphi}_rR(\{H_1,H_2\})$, by Equation (7). □

Applying the lifting $\tilde{\varphi}_r$ furthermore preserves coherence:

Theorem 10. *Consider any reward r in $\mathcal{R}$, and any choice function C on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ that satisfies Axioms* C_4a *and* C_4b. *Then the following statements hold:*

(i) *C satisfies Axiom* C_1 *if and only if $\tilde{\varphi}_rC$ satisfies Axiom* C_1^**;*
(ii) *C satisfies Axiom* C_2 *if and only if $\tilde{\varphi}_rC$ has worst reward r;*
(iii) *C satisfies Axiom* C_3a *if and only if $\tilde{\varphi}_rC$ satisfies Axiom* C_3^*a*;*
(iv) *C satisfies Axiom* C_3b *if and only if $\tilde{\varphi}_rC$ satisfies Axiom* C_3^*b*;*
(v) *$\tilde{\varphi}_rC$ satisfies Axiom* C_4^**;*
(vi) *C satisfies Axiom* C_5 *if and only if $\tilde{\varphi}_rC$ satisfies Axiom* C_5^*.

Proof. For the direct implication of (i), assume that C satisfies Axiom C_1. Consider any A^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$. Then $\tilde{\varphi}_rC(A^*)=\varphi_r^{-1}C(\varphi_rA)\neq\emptyset$.

For the converse implication, assume that $\tilde{\varphi}_rC$ satisfies Axiom C_1^*. Consider any A in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$. By Lemma 7, there are λ in $\mathbb{R}_{>0}$ and g in $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ such that $\frac{1}{\lambda}(A+\{g\})=\varphi_rA^*$ for some A^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$. Applying Axioms C_4a and C_4b and the definition of $\tilde{\varphi}_r$ [Equation (7)], we infer that indeed

$$C(A)=\lambda C\Big(\frac{1}{\lambda}(A+\{g\})\Big)-\{g\}=\lambda C(\varphi_rA^*)-\{g\}=\lambda\varphi_r\tilde{\varphi}_rC(A^*)-\{g\}\neq\emptyset.$$

For the direct implication of (ii), assume that C satisfies Axiom C_2. Consider any H_1 and H_2 in $\mathcal{H}(\mathcal{X},\mathcal{R})$ such that $H_1\neq H_2$ and $H_1(\cdot,s)\leq H_2(\cdot,s)$ for all s in $\mathcal{R}_r$. Then $\varphi_rH_1<\varphi_rH_2$, so Axiom C_2 guarantees that $\varphi_rH_1\in R(\{\varphi_rH_1,\varphi_rH_2\})$. Equation (7) now turns this into $H_1\in\tilde{\varphi}_rR(\{H_1,H_2\})$. Proposition 9 now tells us that $\tilde{\varphi}_rC$ has worst reward r.

For the converse implication, assume that $\tilde{\varphi}_rC$ has worst reward r. Consider any f_1 and f_2 in $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ such that $f_1<f_2$. Let

$$\lambda:=\max_{x\in\mathcal{X}}\sum_{s\in\mathcal{R}_r}(f_2(x,s)-f_1(x,s))>0.$$

Then clearly $\frac{1}{\lambda}(f_2-f_1)=\varphi_rH$ for some H in $\mathcal{H}(\mathcal{X},\mathcal{R})$. Also, $H\neq H_r$ because $f_1\neq f_2$. Using the assumption that $\tilde{\varphi}_rC$ has worst reward r, we find that then $H_r\in\tilde{\varphi}_rR(\{H_r,H\})$. As a consequence, by Equation (7), we find that $0=\varphi_r H_r \in R(\{0,\varphi_r H\}) = R(\{0,\frac{1}{\lambda}(f_2-f_1)\})$. Using Axiom C_4a we infer that $0\in R(\{0,f_2-f_1\})$, and using Axiom C_4b that indeed $f_1\in R(\{f_1,f_2\})$.

For the direct implication of (iii), assume that C satisfies Axiom C_3a. Consider any A^*, A_1^* and A_2^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$ and assume that $A_1^*\subseteq\tilde{\varphi}_r R(A_2^*)$ and $A_2^*\subseteq A^*$. Then $\varphi_r A_1^*\subseteq R(\varphi_r A_2^*)$ by Equation (7), and $\varphi_r A_1^*\subseteq\varphi_r A^*$. Use version (4) of Axiom C_3a to infer that then $\varphi_r A_1^*\subseteq R(\varphi_r A^*)$, whence indeed $A_1^*\subseteq\tilde{\varphi}_r R(A^*)$ by Equation (7).

For the converse implication, assume that $\tilde{\varphi}_r C$ satisfies Axiom C_3^*a. Consider any A, A_1 and A_2 in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$ and assume that $A_1\subseteq R(A_2)$ and $A_2\subseteq A$. Use Lemma 7 to find λ in $\mathbb{R}_{>0}$ and g in $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ such that $\frac{1}{\lambda}(A+\{g\})=\varphi_r A^*$ for some A^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$. Analogously, we find that $\frac{1}{\lambda}(A_2+\{g\})=\varphi_r(A_2^*)$ for some $A_2^*\subseteq A^*$. $A_1\subseteq R(A_2)$ implies $A_1\subseteq A_2$, so also $\frac{1}{\lambda}(A_1+\{g\})=\varphi_r(A_1^*)$ for some $A_1^*\subseteq A_2^*$. Using Axioms C_4a and C_4b, we infer from the assumptions that $\frac{1}{\lambda}(A_1+\{g\})\subseteq R\big(\frac{1}{\lambda}(A_2+\{g\})\big)$, or in other words, $\varphi_r A_1^*\subseteq R(\varphi_r A_2^*)$. Equation (7) then yields that $A_1^*\subseteq\tilde{\varphi}_r R(A_2^*)$. As a result, using Axiom C_3^*a, $A_1^*\subseteq\tilde{\varphi}_r R(A^*)$, which, again applying Equation (7), results in $\frac{1}{\lambda}(A_1+\{g\})=\varphi_r A_1^*\subseteq R(\varphi_r A^*)=R\big(\frac{1}{\lambda}(A+\{g\})\big)$, and as a consequence, by Axioms C_4a and C_4b, we find eventually that indeed $A_1\subseteq R(A)$.

For the direct implication of (iv), assume that C satisfies Axiom C_3b. Consider any A^*, A_1^* and A_2^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$ and assume that $A_1^*\subseteq\tilde{\varphi}_r R(A_2^*)$ and $A^*\subseteq A_1^*$. Then $\varphi_r A_1^*\subseteq R(\varphi_r A_2^*)$ by Equation (7), and $\varphi_r A^*\subseteq\varphi_r A_1^*$. Use version (5) of Axiom C_3b to infer that then $\varphi_r(A_1^*\setminus A^*)=(\varphi_r A_1^*)\setminus(\varphi_r A)\subseteq R((\varphi_r A_2^*)\setminus(\varphi_r A^*))=R(\varphi_r(A_2^*\setminus A^*))$, whence indeed $A_1^*\setminus A^*\subseteq\tilde{\varphi}_r R(A_2^*\setminus A^*)$.

For the converse implication, assume that $\tilde{\varphi}_r C$ satisfies Axiom C_3^*b. Consider any A, A_1 and A_2 in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$ and assume that $A_1\subseteq R(A_2)$ and $A\subseteq A_1$. Use Lemma 7 to find λ in $\mathbb{R}_{>0}$ and g in $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ such that $\frac{1}{\lambda}(A_2+\{g\})=\varphi_r A_2^*$ for some A_2^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$. $A_1\subseteq R(A_2)$ implies $A_1\subseteq A_2$, whence $\frac{1}{\lambda}(A_1+\{g\})=\varphi_r(A_1^*)$ for some $A_1^*\subseteq A_2^*$, and analogously, $\frac{1}{\lambda}(A+\{g\})=\varphi_r(A^*)$ for some $A^*\subseteq A_1^*$. Using Axioms C_4a and C_4b we find that $\frac{1}{\lambda}(A_1+\{g\})\subseteq R\big(\frac{1}{\lambda}(A_2+\{g\})\big)$, or in other words, $\varphi_r A_1^*\subseteq R(\varphi_r A_2^*)$. Equation (7) then tells us that $A_1^*\subseteq\tilde{\varphi}_r R(A_2^*)$, which, using Axiom C_3^*b, results in $A_1^*\setminus A^*\subseteq\tilde{\varphi}_r R(A_2^*\setminus A^*)$. Again applying Equation (7) results in

$$
\begin{aligned}
\frac{1}{\lambda}((A_1\setminus A)+\{g\}) = \frac{1}{\lambda}(A_1+\{g\})\setminus\frac{1}{\lambda}(A+\{g\}) &= (\varphi_r A_1^*)\setminus(\varphi_r A^*) = \varphi_r(A_1^*\setminus A^*)\\
&\subseteq R(\varphi_r(A_2^*\setminus A^*)) = R((\varphi_r A_2^*)\setminus(\varphi_r A^*))\\
&= R(\frac{1}{\lambda}(A_2+\{g\})\setminus\frac{1}{\lambda}(A+\{g\})) = R(\frac{1}{\lambda}((A_2\setminus A)+\{g\})),
\end{aligned}
$$

and as a consequence, by Axioms C_4a and C_4b, we find eventually that indeed $A_1\setminus A\subseteq R(A_2\setminus A)$.

For (v), since by Proposition 8, $\tilde{\varphi}_r$ is a bijection between the choice functions on $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ that satisfy Axioms C_4a and C_4b, and the choice functions on $\mathcal{H}(\mathcal{X},\mathcal{R})$ that satisfy Axiom C_4^*, therefore indeed $\tilde{\varphi}_r C$ satisfies Axiom C_4^*.

For the direct implication of (vi), assume that C satisfies Axiom C_5. Consider any A^* and A_1^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$ and assume that $A^*\subseteq A_1^*\subseteq\mathrm{CH}(A^*)$. Then $\varphi_r A^*\subseteq\varphi_r A_1^*\subseteq\mathrm{CH}(\varphi_r A^*)$, whence $C(\varphi_r A^*)\subseteq C(\varphi_r A_1^*)$ by Axiom C_5. Use Equation (7) to infer that then indeed $\tilde{\varphi}_r C(A^*)\subseteq\tilde{\varphi}_r C(A_1^*)$.

For the converse implication, assume that $\tilde{\varphi}_r C$ satisfies Axiom C_5^*. Consider any A and A_1 in $\mathcal{Q}(\mathcal{L}(\mathcal{X}\times\mathcal{R}_r))$ and assume that $A\subseteq A_1\subseteq\mathrm{CH}(A)$. Use Lemma 7 to find λ in $\mathbb{R}_{>0}$ and g in $\mathcal{L}(\mathcal{X}\times\mathcal{R}_r)$ such that $\frac{1}{\lambda}(A_1+\{g\})=\varphi_r A_1^*$ for some A_1^* in $\mathcal{Q}(\mathcal{H}(\mathcal{X},\mathcal{R}))$, and analogously, $\frac{1}{\lambda}(A+\{g\})=\varphi_r(A^*)$ for some $A^*\subseteq A_1^*$. From $A_1\subseteq\mathrm{CH}(A)$ infer that $\frac{1}{\lambda}(A_1+\{g\})\subseteq\mathrm{CH}(\frac{1}{\lambda}(A+\{g\}))$, or in other words, $\varphi_r A_1^*\subseteq\mathrm{CH}(\varphi_r A^*)$. Then we claim that $A_1^*\subseteq\mathrm{CH}(A^*)$. To prove this, consider any H in A_1^*. Then there are n in $\mathbb{N}$, H_i in A, and $\alpha_i \geq 0$ such that $\sum_{i=1}^n \alpha_i = 1$ and $H(\cdot,s) = \sum_{i=1}^n \alpha_i H_i(\cdot,s)$ for all s in $\mathcal{R}_r$. Moreover,

$$\begin{aligned} H(\cdot,r) &= 1 - \sum_{s\in\mathcal{R}_r} H(\cdot,s) = 1 - \sum_{s\in\mathcal{R}_r}\sum_{i=1}^n \alpha_i H_i(\cdot,s) \\ &= \sum_{i=1}^n \alpha_i - \sum_{i=1}^n \alpha_i \sum_{s\in\mathcal{R}_r} H_i(\cdot,s) = \sum_{i=1}^n \alpha_i\Big(1 - \sum_{s\in\mathcal{R}_r} H_i(\cdot,s)\Big) = \sum_{i=1}^n \alpha_i H_i(\cdot,r), \end{aligned}$$

so indeed $H \in \mathrm{CH}(A^*)$. Use Axiom C_5^* to infer that then $\tilde{\varphi}_r C(A^*) \subseteq \tilde{\varphi}_r C(A_1^*)$. Equation (7) turns this into $C\big(\frac{1}{\lambda}(A+\{g\})\big) = C(\varphi_r A^*) \subseteq C(\varphi_r A_1^*) = C\big(\frac{1}{\lambda}(A_1+\{g\})\big)$, which by Axioms C_4a and C_4b, results in $C(A) \subseteq C(A_1)$. □

We conclude that our discussion of choice functions on linear spaces subsumes the treatment of choice functions on horse lotteries satisfying Axiom C_4^*. Using the connections established above, all the results that we will prove later on are also applicable to choice functions on horse lotteries that satisfy the corresponding rationality axioms.

4. THE LINK WITH DESIRABILITY

Van Camp et al. (2017) have studied in some detail how the coherent choice functions in the sense of Definition 2 can be related to coherent sets of desirable options (gambles). As an example, given a coherent set of desirable options D, the choice function that identifies the undominated—under the preference relation induced by D—options, is coherent. This choice rule is called *maximality* (see Equation (12) further on). There are other rules that induce coherent choice functions, such as E-admissibility—those choice functions identify the options whose (precise) expectation is maximal for at least one probability mass function in the credal set induced by D. Since we have shown in earlier work (Van Camp et al., 2017, Proposition 13) that maximality leads to the most conservative coherent choice function that reflects the binary choices represented by D (see also Bradley, 2015, Theorem 3), we focus on maximality as the connecting tool between desirability and choice functions. Here, we investigate what remains of this connection when we require in addition that our choice functions should satisfy Axiom C_5.

We recall that a set of desirable options is simply a subset of the vector space $\mathcal{V}$. The underlying idea is that a subject strictly prefers each option in this set to the status quo 0. As for choice functions, we pay special attention to *coherent* sets of desirable options.

Definition 8. A set of desirable options D is called *coherent* if for all u and v in $\mathcal{V}$, and all λ in $\mathbb{R}_{>0}$:
D_1. $0 \notin D$;
D_2. $\mathcal{V}_{>0} \subseteq D$;
D_3. if $u \in D$ then $\lambda u \in D$;
D_4. if $u, v \in D$ then $u + v \in D$.
We collect all coherent sets of desirable options in the set $\bar{\mathcal{D}}$.

More details can be found in a number of papers and books (Walley, 1991, 2000; Moral, 2005; Miranda and Zaffalon, 2010; Couso and Moral, 2011; De Cooman and Quaeghebeur, 2012; De Cooman and Miranda, 2012; Quaeghebeur, 2014; Quaeghebeur et al., 2015; De Bock and De Cooman, 2015).

Axioms D_3 and D_4 guarantee that a coherent D is a convex cone. This convex cone induces a strict partial order $\prec_D$ on $\mathcal{V}$, by letting

$$u \prec_D v \Leftrightarrow 0 \prec_D v - u \Leftrightarrow v - u \in D, \tag{10}$$

so $D = \{u \in \mathcal{V} : 0 \prec_D u\}$ (De Cooman and Quaeghebeur, 2012; Quaeghebeur, 2014). D and $\prec_D$ are mathematically equivalent: given one of D or $\prec_D$, we can determine the other unequivocally using the formulas above. When it is clear from the context which set of desirable options D we are working with, we often refrain from mentioning the explicit reference to D in $\prec_D$ and then we simply write $\prec$. Coherence for sets of desirable options transfers to binary relations $\prec$ as follows: $\prec$ must be a strict partial order—meaning that it is irreflexive and transitive—such that $< \subseteq \prec$, and must satisfy the two characteristic properties of Equations (2) and (3).

What is the relationship between choice functions and sets of desirable options? Since we have just seen that sets of desirable options represent binary preferences, we see that we can associate a set of desirable options D_C with every given choice function C by focusing on its binary choices:

$$u \prec_{D_C} v \Leftrightarrow v - u \in D_C \Leftrightarrow u \in R(\{u, v\}) \text{ for all } u, v \text{ in } \mathcal{V}. \tag{11}$$

D_C is a coherent set of desirable options if C is a coherent choice function (Van Camp et al., 2017, Proposition 12). Conversely (Van Camp et al., 2017, Proposition 13), if we start out with a coherent set of desirable options D then the set $\{C \in \bar{\mathcal{C}} : D_C = D\}$ of all coherent choice functions whose binary choices are represented by D, is non-empty, and its smallest, or least informative, element $C_D \coloneqq \inf\{C \in \bar{\mathcal{C}} : D_C = D\}$ is given by:

$$C_D(A) \coloneqq \{u \in A : (\forall v \in A)\, v - u \notin D\} = \{u \in A : (\forall v \in A)\, u \nprec_D v\} \text{ for all } A \text{ in } \mathcal{Q}. \tag{12}$$

It selects all options from A that are undominated, or maximal, under the ordering $\prec_D$, or in other words, it is the corresponding choice function based on Walley–Sen maximality. This C_D is easy to characterise:

Proposition 11. *Given any coherent set of desirable options D, then*

$$0 \in C_D(\{0\} \cup A) \Leftrightarrow D \cap A = \emptyset \textit{ for all } A \textit{ in } \mathcal{Q}.$$

Proof. By Equation (12), $0 \in C_D(\{0\} \cup A) \Leftrightarrow (\forall v \in \{0\} \cup A)\, v \notin D \Leftrightarrow (\{0\} \cup A) \cap D = \emptyset$, which is equivalent to $A \cap D = \emptyset$, because $0 \notin D$ for any coherent D. □

Although C_D is coherent when D is, it does not necessarily satisfy the additional Axiom C_5, as the following counterexample shows.

Example 1. Consider the two-dimensional vector space $\mathcal{V} = \mathbb{R}^2$. We provide it with the component-wise vector ordering $\leq$, and consider the vacuous set of desirable options $D = \{u \in \mathcal{V} : u > 0\} = \mathcal{V}_{>0}$, which is coherent. By Proposition 11, $0 \in C_D(\{0\} \cup A) \Leftrightarrow A \cap \mathcal{V}_{>0} = \emptyset$ for all A in $\mathcal{Q}$. To show that C_D does not satisfy Axiom C_5, consider $A = \{0, u, v\}$, where $u \coloneqq (-1, 2)$ and $v \coloneqq (2, -1)$. We find that $0 \in C_D(A)$ because $\{u, v\} \cap \mathcal{V}_{>0} = \emptyset$, since $u \ngtr 0$ and $v \ngtr 0$.

However, for the option set $A_1 = A \cup \{\frac{u+v}{2}\} \subseteq \mathrm{CH}(A)$, we find that $\frac{u+v}{2} = (1/2, 1/2) > 0$ and therefore $0 \notin C_D(A_1)$, meaning that Axiom C_5 is not satisfied. ◊

For the specific coherent set of desirable options D considered in Example 1, the corresponding choice function C_D fails to satisfy C_5. However, there are other sets of desirable options D for which C_D does satisfy the convexity axiom. They are identified in the next proposition.

Proposition 12. *Consider any coherent set of desirable options D. Then the corresponding coherent choice function C_D satisfies Axiom* C_5 *if and only if D^c is a convex cone, or in other words, if and only if* $\mathrm{posi}(D^c) = D^c$*, or equivalently,* $\mathrm{posi}(D^c) \cap D = \emptyset$.

Proof. Van Camp et al. (2017, Proposition 13) have already shown that C_D is a coherent choice function.

For necessity, assume that $\mathrm{posi}(D^c) \neq D^c$, or equivalently, that $\mathrm{posi}(D^c) \cap D \neq \emptyset$. Then there is some option u in D such that $u \in \mathrm{posi}(D^c)$, meaning that there are n in $\mathbb{N}$, λ_k in $\mathbb{R}_{>0}$ and u_k in D^c such that $u = \sum_{k=1}^{n} \lambda_k u_k$. Let $A \coloneqq \{0, u_1, \ldots, u_n\}$ and $A_1 \coloneqq A \cup \{u\}$. Due to the coherence of D [more precisely Axiom D_3], we can rescale $u \in D$ while keeping the u_k fixed, in such a way that we achieve that $\sum_{k=1}^{n} \lambda_k = 1$, whence $A \subseteq A_1 \subseteq \mathrm{CH}(A)$.

We find that $0 \in C_D(A)$ by Proposition 11, because $A \cap D = \emptyset$, but $0 \notin C_D(A_1)$ because $u \in D$, so $A_1 \cap D \neq \emptyset$. This tells us that C_D does not satisfy Axiom C$_5$, because clearly $C_D(A) \nsubseteq C_D(A_1)$.

For sufficiency, assume that C_D does not satisfy Axiom C$_5$. Therefore, there are A and A_1 in $\mathcal{Q}$ such that $A \subseteq A_1 \subseteq \mathrm{CH}(A)$ and $C_D(A) \nsubseteq C_D(A_1)$, or, in other words, such that $u \in C_D(A)$ and $u \notin C_D(A_1)$ for some u in A. Consider such A and A_1 in $\mathcal{Q}$, and u in A. Due to Axiom C$_4$b, we find that $0 \in C_D(A - \{u\})$ and $0 \notin C_D(A_1 - \{u\})$, or equivalently, by Proposition 11, that $A - \{u\} \subseteq D^c$ and $A_1 - \{u\} \cap D \neq \emptyset$. But $A_1 - \{u\} \subseteq \mathrm{CH}(A) - \{u\} = \mathrm{CH}(A - \{u\}) \subseteq \mathrm{posi}(A - \{u\}) \subseteq \mathrm{posi}(D^c)$, so $\mathrm{posi}(D^c) \cap D \neq \emptyset$. □

This proposition seems to indicate that there is something special about coherent sets of desirable options whose complement is a convex cone too. We give them a special name that will be motivated and explained in the next section.

Definition 9. A coherent set of desirable options D is called *lexicographic* if

$$\mathrm{posi}(D^c) = D^c, \text{ or, equivalently, if } \mathrm{posi}(D^c) \cap D = \emptyset.$$

We collect all the lexicographic coherent sets of desirable options in $\bar{\mathcal{D}}_{\mathrm{L}}$.

Another important subclass $\hat{\mathcal{D}}$ of coherent sets of desirable options collects all the maximally informative, or *maximal*, ones:

$$\hat{\mathcal{D}} := \{D \in \bar{\mathcal{D}} : (\forall D' \in \bar{\mathcal{D}}) D \subseteq D' \Rightarrow D = D'\}.$$

The sets of desirable options in $\hat{\mathcal{D}}$ are the undominated elements of the complete infimum-semilattice $(\bar{\mathcal{D}}, \subseteq)$. Couso and Moral (2011) have proved the following elegant and useful characterisation of these maximal elements:

Proposition 13. *Given any coherent set of desirable options D and any non-zero option $u \notin D$, then $\mathrm{posi}(D \cup \{-u\})$ is a coherent set of desirable options. As a consequence, a coherent set of desirable options D is maximal if and only if*

$$(\forall u \in \mathcal{V} \setminus \{0\})(u \in D \text{ or } -u \in D).$$

De Cooman and Quaeghebeur (2012) have proved that the set of all coherent sets of desirable options is *dually atomic*, meaning that any coherent set of desirable options D is the infimum of its non-empty set of dominating maximal coherent sets of desirable options:

Proposition 14. *For any coherent set of desirable options D, its set of dominating maximal coherent sets of desirable options $\hat{\mathcal{D}}_D := \{\hat{D} \in \hat{\mathcal{D}} : D \subseteq \hat{D}\}$ is non-empty, and $D = \bigcap \hat{\mathcal{D}}_D$.*

Any maximal coherent set of desirable options is also a lexicographic one: $\hat{\mathcal{D}} \subseteq \bar{\mathcal{D}}_{\mathrm{L}}$ [3] Consider a maximal D and arbitrary n in $\mathbb{N}$, u_k in D^c and $\lambda_k \in \mathbb{R}_{>0}$ for $k \in \{1, \dots, n\}$. Then since all $-u_k \in D \cup \{0\}$ by Proposition 13, we infer that $-\sum_{k=1}^n \lambda_k u_k \in D \cup \{0\}$, because the coherent D is in particular a convex cone. If $\sum_{k=1}^n \lambda_k u_k = 0$, then $\sum_{k=1}^n \lambda_k u_k \in D^c$ by Axiom D$_1$. If $\sum_{k=1}^n \lambda_k u_k \neq 0$, then $-\sum_{k=1}^n \lambda_k u_k \in D$, and since coherence [more specifically, a combination of Axioms D$_1$ and D$_4$] implies that a coherent set of gambles cannot include both a gamble and its opposite, we conclude that, here too, $\sum_{k=1}^n \lambda_k u_k \in D^c$. Therefore, D^c is indeed a convex cone.

[3] As pointed out by a reviewer, this can actually be obtained as a corollary to a result by Hammer (1955, Theorem 2), by taking into account that maximal sets of desirable gambles are *semispaces* and that lexicographic sets of desirable gambles correspond to *hemispaces*. To make this paper more self-contained, we give a proof using the coherence axioms of sets of desirable gambles we are employing in this paper.

5. LEXICOGRAPHIC CHOICE FUNCTIONS

In this section, we embark on a more detailed study of lexicographic sets of desirable options, and amongst other things, explain where their name comes from. We will restrict ourselves here to the special case where $\mathcal{V}$ is the linear space $\mathcal{L}(\mathcal{X})$ of all gambles on a *finite* possibility space $\mathcal{X}$, provided with the component-wise order $\leq$ as its vector ordering.

We first show that the lower expectation functional associated with a lexicographic D is actually a linear prevision (Walley, 1991; Troffaes and De Cooman, 2014).

Proposition 15. *For any D in $\bar{\mathcal{D}}_{\mathrm{L}}$, the coherent lower prevision $\underline{P}_D$ on $\mathcal{L}(\mathcal{X})$ defined by*

$$\underline{P}_D(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in D\} \text{ for all } f \text{ in } \mathcal{L}(\mathcal{X})$$

is a linear prevision*: a real linear functional that is positive—so $(\forall f \geq 0)\underline{P}_D(f) \geq 0$—and normalised—meaning that $\underline{P}_D(1) = 1$.*

Proof. Consider any f in $\mathcal{L}$ and ϵ in $\mathbb{R}_{>0}$, then we first prove that $f \in D$ or $\epsilon - f \in D$. Assume *ex absurdo* that $f \notin D$ and $\epsilon - f \notin D$. Then, because by assumption $\operatorname{posi}(D^c) = D^c$ is a convex cone, we also have that $f + \epsilon - f = \epsilon \notin D$, which contradicts Axiom D_2. Now, Walley (1991, Theorem 3.8.3) guarantees that for any such D, the corresponding functional $\underline{P}_D$ is indeed a linear prevision. □

To get some feeling for what these lexicographic models represent, we first look at the special case of binary possibility spaces $\{a, b\}$, leading to a two-dimensional option space $\mathcal{V} = \mathcal{L}(\{a, b\})$ provided with the point-wise order. It turns out that lexicographic sets of desirable options (gambles) are easy to characterise there, so we have a simple expression for $\bar{\mathcal{D}}_{\mathrm{L}}$.

Proposition 16. *All lexicographic coherent sets of desirable gambles on the binary possibility space $\{a, b\}$ are given by (see also Figure 1):*

$$\bar{\mathcal{D}}_{\mathrm{L}} := \{D_\rho, D^a_\rho, D^b_\rho : \rho \in (0,1)\} \cup \{D_0, D_1\} = \{D_\rho : \rho \in (0,1)\} \cup \hat{\mathcal{D}},$$

where

$$\begin{aligned}
D_\rho &:= \{\lambda(\rho - \mathbb{I}_{\{a\}}) : \lambda \in \mathbb{R}\} + \mathcal{V}_{>0} = \operatorname{span}(\{\rho - \mathbb{I}_{\{a\}}\}) + \mathcal{V}_{>0} && \text{for all } \rho \text{ in } (0,1)\\
D^a_\rho &:= D_\rho \cup \{\lambda(\rho - \mathbb{I}_{\{a\}}) : \lambda \in \mathbb{R}_{<0}\} = D_\rho \cup \operatorname{posi}(\{\mathbb{I}_{\{a\}} - \rho\}) && \text{for all } \rho \text{ in } (0,1)\\
D^b_\rho &:= D_\rho \cup \{\lambda(\rho - \mathbb{I}_{\{a\}}) : \lambda \in \mathbb{R}_{>0}\} = D_\rho \cup \operatorname{posi}(\{\rho - \mathbb{I}_{\{a\}}\}) && \text{for all } \rho \text{ in } (0,1).\\
D_0 &:= \{f \in \mathcal{V} : f(b) > 0\} \cup \mathcal{V}_{>0}\\
D_1 &:= \{f \in \mathcal{V} : f(a) > 0\} \cup \mathcal{V}_{>0}.
\end{aligned}$$

Proof. We first observe that every set of desirable options in $\bar{\mathcal{D}}_{\mathrm{L}}$ is coherent. Indeed, for any ρ in $(0,1)$, D_ρ is the smallest coherent set of desirable gambles corresponding to the linear prevision E_p, with $p := (\rho, 1-\rho)$, while D^a_ρ, D^b_ρ are maximal coherent sets of desirable gambles corresponding to the same linear prevision E_p. Finally, D_0 is the maximal (and only) coherent set of desirable gambles corresponding to E_p with $p := (0,1)$, while D_1 is the maximal (and only) coherent set of desirable gambles corresponding to E_p with $p := (1,0)$.

We now prove that we recover all lexicographic coherent sets of desirable gambles in this way. Consider any lexicographic coherent set of desirable gambles D^*. Then $\underline{P}_{D^*}$ is a linear prevision, by Proposition 15, so $\underline{P}_{D^*}$ is characterised (i) by the mass function $(1,0)$, (ii) by the mass function $(0,1)$, or (iii) by the mass function $(\rho^*, 1-\rho^*)$ for some ρ^* in $(0,1)$. If (i), the only coherent set of desirable gambles that induces the linear prevision with mass function $(1,0)$ is $D_1 \in \bar{\mathcal{D}}_{\mathrm{L}}$. If (ii), the only coherent set of desirable gambles that induces the linear prevision with mass function $(0,1)$ is $D_0 \in \bar{\mathcal{D}}_{\mathrm{L}}$. If (iii), there are only three coherent sets of desirable gambles that induce the linear prevision with mass function $(\rho^*, 1-\rho^*)$: D_{ρ^*}, $D^a_{\rho^*}$ and $D^b_{\rho^*}$, and all are elements of $\bar{\mathcal{D}}_{\mathrm{L}}$. □

In the language of sets of desirable gambles (see for instance Quaeghebeur, 2014), this means that in the binary case lexicographic sets of desirable gambles are either *maximal* or *strictly desirable* with respect to a linear prevision.

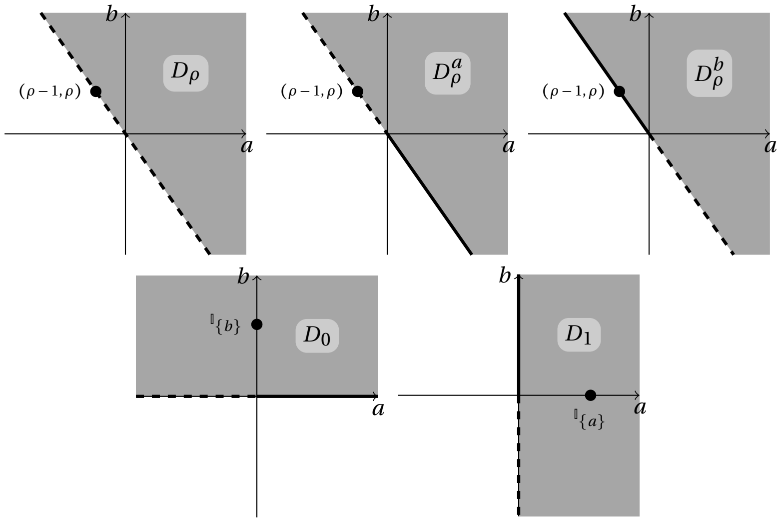

FIGURE 1. The lexicographic coherent sets of desirable gambles on the binary possibility space $\{a,b\}$, with $\rho \in (0,1)$.

We now turn to the more general finite-dimensional case. Recall that a *lexicographic order* $<_{\mathrm{L}}$ with $\ell \in \mathbb{N}$ *layers* on a vector space $\mathcal{V}$ of finite dimension n is defined by

$$u <_{\mathrm{L}} v \Leftrightarrow (\exists k \in \{1,\dots,\ell\})(u_k < v_k \text{ and } (\forall j \in \{1,\dots,k-1\})u_j = v_j),$$

and denote, as usual, its reflexive version $\leq_{\mathrm{L}}$ as $u \leq_{\mathrm{L}} v \Leftrightarrow (u <_{\mathrm{L}} v \text{ or } u = v)$ for any two vectors $u = (u_1,\dots,u_n)$ and $v = (v_1,\dots,v_n)$ in $\mathcal{V}$. A *lexicographic probability system* is an ℓ-tuple $p \coloneqq (p_1,\dots,p_\ell)$ of probability mass functions on a possibility space $\mathcal{X}$ of cardinality n. We associate with this tuple p an expectation operator $\mathrm{E}_p \coloneqq (\mathrm{E}_{p_1},\dots,\mathrm{E}_{p_\ell})$, and a (strict) preference relation $\prec_p$ on $\mathcal{L}(\mathcal{X})$, defined by:

$$f \prec_p g \Leftrightarrow \mathrm{E}_p(f) <_{\mathrm{L}} \mathrm{E}_p(g), \text{ for all } f,g \text{ in } \mathcal{L}(\mathcal{X}). \tag{13}$$

We refer to work by Blume et al. (1991), Fishburn (1982) and Seidenfeld et al. (1990) for more details on generic lexicographic probability systems. The connection between lexicographic probability systems and sets of desirable gambles has also been studied by Cozman (2015) and Benavoli et al. (2017), and the connection with full conditional measures by Halpern (2010) and Hammond (1994). Below, we first recall a number of relevant basic properties of lexicographic orders in Propositions 18 and 19. We then provide a characterisation of lexicographic sets of desirable gambles in terms of lexicographic orders in Theorem 23.

Remark that the reflexive version $\preceq_p$ of $\prec_p$—defined by $f \preceq_p g \Leftrightarrow \mathrm{E}_p(f) \leq_{\mathrm{L}} \mathrm{E}_p(g)$ for all f and g in $\mathcal{L}(\mathcal{X})$—is a total order on $\mathcal{L}(\mathcal{X})$ (Blume et al., 1991).

In what follows, we will restrict our attention to lexicographic probability systems p that satisfy the following condition:

$$(\forall x \in \mathcal{X})(\exists k \in \{1,\dots,\ell\})p_k(x) > 0. \tag{14}$$

This condition requires that there should be no possible outcome in $\mathcal{X}$ that has zero probability in every layer. It is closely related to the notion of a *Savage-null event* (Savage, 1972, Section 2.7):

Definition 10. An event $B \subseteq \mathcal{X}$ is called *Savage-null* if $(\forall f, g \in \mathcal{L}(\mathcal{X}))\mathbb{I}_B f \leq_p \mathbb{I}_B g$. The event $\emptyset$ is always Savage-null, and is called the *trivial Savage-null event*.

An important feature of preference relations $\prec_p$ based on lexicographic probability systems is the *incomparability relation* $\|_p$, defined by: $f \|_p g$ if and only if $f \not\prec_p g$ and $g \not\prec_p f$ for all f and g in $\mathcal{L}(\mathcal{X})$. Since $\leq_p$ is a total order, it follows that

$$f \|_p g \Leftrightarrow \mathrm{E}_p(f) = \mathrm{E}_p(g). \tag{15}$$

Finally, it also follows that

$$f \not\prec_p g \Leftrightarrow g \prec_p f \text{ or } g \|_p f \Leftrightarrow \mathrm{E}_p(g) \leq_{\mathrm{L}} \mathrm{E}_p(f). \tag{16}$$

Proposition 17. *Consider any lexicographic probability system $p = (p_1, \dots, p_\ell)$. Then Condition* (14) *holds if and only if there are no non-trivial Savage-null events.*

Proof. For the direct implication, consider any lexicographic probability system p that satisfies Condition (14), and consider any non-empty event $B \subseteq \mathcal{X}$. Consider any x in B, then $\mathbb{I}_B \geq \mathbb{I}_{\{x\}}$ so $\mathrm{E}_{p_k}(\mathbb{I}_B) \geq \mathrm{E}_{p_k}(\mathbb{I}_{\{x\}})$ for every $k \in \{1, \dots, \ell\}$. Also, $\mathrm{E}_p(0) <_{\mathrm{L}} \mathrm{E}_p(\mathbb{I}_{\{x\}})$ by Condition (14), so $0\mathbb{I}_B \prec_p 1\mathbb{I}_B$ whence $1\mathbb{I}_B \not\leq_p 0\mathbb{I}_B$ and hence, by Definition 10, B is indeed no Savage-null event.

For the converse implication, consider any lexicographic probability system p and assume that Condition (14) does not hold. Then there is some x^* in $\mathcal{X}$ such that $p_k(x^*) = 0$ for all k in $\{1, \dots, \ell\}$, and therefore $\mathrm{E}_{p_k}(f\mathbb{I}_{\{x^*\}}) = 0 = \mathrm{E}_{p_k}(g\mathbb{I}_{\{x^*\}})$ for all f and g in $\mathcal{L}(\mathcal{X})$ and k in $\{1, \dots, \ell\}$, so $\mathrm{E}_p(f\mathbb{I}_{\{x^*\}}) = \mathrm{E}_p(g\mathbb{I}_{\{x^*\}})$ for all f and g in $\mathcal{L}(\mathcal{X})$. This implies that $f\mathbb{I}_{\{x^*\}} \leq_p g\mathbb{I}_{\{x^*\}}$ for all f and g in $\mathcal{L}(\mathcal{X})$, so indeed there is a non-trivial Savage-null event $\{x^*\}$. □

Proposition 18. *Consider any lexicographic probability system p with ℓ layers. Then $\prec_p$ is a* strict weak order, *meaning that $\prec_p$ is irreflexive, and both $\prec_p$ and $\|_p$ are transitive. As a consequence, the relation $\not\prec_p$ is transitive as well.*

Proof. This is a consequence of Equations (13), (15) and (16), taking into account that $<_{\mathrm{L}}$ and $\leq_{\mathrm{L}}$ are transitive, and that $<_{\mathrm{L}}$ is irreflexive. □

We now link the lexicographic orderings $\prec_p$ with the preference relation $\lhd_D$ based on desirability, given by Equation (10). We begin with an auxiliary result:[4]

Proposition 19. *Consider any lexicographic probability system p with ℓ layers, and consider any coherent set of desirable gambles D. Then $\prec_p$ and $\lhd_D$ are (strict) vector orders compatible with $<$: they are irreflexive, transitive and*

(i) $f \prec_p g \Leftrightarrow f + h \prec_p g + h \Leftrightarrow \lambda f \prec_p \lambda g$;
(ii) *if there are no non-trivial Savage-null events, then* $f < g \Rightarrow f \prec_p g$;
(iii) $f \lhd_D g \Leftrightarrow f + h \lhd_D g + h \Leftrightarrow \lambda f \lhd_D \lambda g$;
(iv) $f < g \Rightarrow f \lhd_D g$,

for all f, g and h in $\mathcal{L}(\mathcal{X})$ and λ in $\mathbb{R}_{>0}$.

Proof. It is clear from Proposition 18 that $\prec_p$ is irreflexive and transitive. To show that $\lhd_D$ is irreflexive, infer from $f - f = 0 \notin D$ [Axiom D_1] that indeed $f \not\lhd_D f$ for all f in $\mathcal{L}(\mathcal{X})$. To show that $\lhd_D$ is transitive, assume that $f \lhd_D g$ and $g \lhd_D h$. Then $g - f \in D$ and $h - g \in D$, by Equation (10), and hence $h - f = g - f + (h - g) \in D$, by Axiom D_4. Using Equation (10) again, we find that then indeed $f \lhd_D h$. Let us now prove the remaining statements.

[4]Except for the second statement, most of the items in this propositions are well-known (Quaeghebeur, 2014, Section 1.4.1); we include a simple proof for completeness.

(i) This follows from the definition of $\prec$ and the linearity of the expectation operator.
(ii) Assume that there are no non-trivial Savage-null events. Use Proposition 17 to infer that Condition (14) holds. Consider any f in $\mathcal{L}(\mathcal{X})$ such that $0 < f$. Then $0 \leq f$—so $0 \leq \mathrm{E}_{p_k}(f)$ for every k in $\{1,\ldots,\ell\}$—and $0 < f(x^*)$ for some x^* in $\mathcal{X}$. Then $p_k(x^*) > 0$ for some k in $\{1,\ldots,\ell\}$ by Condition (14), so $0 <_{\mathrm{L}} \mathrm{E}_p(\mathbb{I}_{\{x^*\}})$. Use $f(x^*)\mathbb{I}_{\{x^*\}} \leq f$ to infer that then also $0 <_{\mathrm{L}} \mathrm{E}_p(f)$, whence indeed $0 \prec_p f$.
(iii) The first equivalence follows immediately from Equation (10), while the second is a consequence of the scaling axiom of coherent sets of desirable options.
(iv) Assume that $f < g$. Then $0 < g - f$, whence $g - f \in D$ by Axiom D_2. Using Equation (10), we find that then indeed $f \prec_D g$. □

Next we establish a link between lexicographic probability systems and preference relations associated with lexicographic sets of desirable gambles. We refer to papers by Cozman (2015, Section 2.1) and Seidenfeld et al. (1990) for other relevant discussion on the connection between lexicographic probabilities and partial preference relations. Our proof is somewhat reminiscent of the representation of conditional probabilities by Krauss (1968), and will make repeated use of the following separation theorem (Holmes, 1975), in the form stated by Walley (1991, Appendix E1):

Theorem 20 (Separating hyperplane theorem)**.** *Let $\mathcal{W}_1$ and $\mathcal{W}_2$ be two convex subsets of a finite-dimensional linear topological space $\mathcal{B}$. If $0 \in \mathcal{W}_1 \cap \mathcal{W}_2$ and $\operatorname{int}(\mathcal{W}_1) \cap \mathcal{W}_2 = \emptyset$, then there is a non-zero continuous linear functional Λ on $\mathcal{B}$ such that*

$$\Lambda(w) \geq 0 \text{ for all } w \text{ in } \mathcal{W}_1 \text{ and } \Lambda(w') \leq 0 \text{ for all } w' \text{ in } \mathcal{W}_2.$$

If $\mathcal{W}_1$ and $\mathcal{W}_2$ are finite, $\mathcal{W}_1$ non-empty, and $\sum_{i=1}^m \lambda_i w_i - \sum_{k=1}^n \mu_k w'_k \neq 0$ for all m and n in $\mathbb{N}$, all $\lambda_1, \ldots, \lambda_m$ in $\mathbb{R}_{\geq 0}$ with $\lambda_i > 0$ for at least one i in $\{1,\ldots,m\}$, all $\mu_1, \ldots, \mu_n$ in $\mathbb{R}_{\geq 0}$, all $w_1, \ldots, w_m$ in $\mathcal{W}_1$, and all $w'_1, \ldots, w'_n$ in $\mathcal{W}_2$, then there is a non-zero continuous linear functional Λ on $\mathcal{B}$ such that

$$\Lambda(w) > 0 \text{ for all } w \text{ in } \mathcal{W}_1 \text{ and } \Lambda(w') \leq 0 \text{ for all } w' \text{ in } \mathcal{W}_2.$$

Two clarifications here are (i) that we will apply the theorem to linear subsets of $\mathcal{L}(\mathcal{X})$, which is a linear topological space (Walley, 1991, Appendix D) that is finite-dimensional because $\mathcal{X}$ is finite, and (ii) that when the linear topological space is finite-dimensional, the assumption $\operatorname{int}(\mathcal{W}_1) \neq \emptyset$ that Walley (1991, Appendix E1) mentions is not necessary for the separating hyperplane theorem to hold, as shown by Holmes (1975, Theorem 4B).

Our proof will also make use of the following two lemmas.

Lemma 21. *Consider any coherent set D of desirable gambles on a finite possibility space $\mathcal{X}$, and consider any linear subspace $\Lambda \subseteq \mathcal{L}(\mathcal{X})$. Then $\operatorname{int}(\operatorname{cl}(D \cap \Lambda)) \cap D^c = \emptyset$, where* int *is the topological interior and* cl *the topological closure.*

Proof. We first prove $\operatorname{int}(\operatorname{cl}(D)) \cap D^c = \emptyset$. To show that, we will use the fact that D, and therefore also $\operatorname{cl}(D)$, is a convex set. Since the interior of a convex set is always included in the relative interior ri of that convex set (see Brøndsted, 1983, Section 1.3), we find that $\operatorname{int}(\operatorname{cl}(D)) \subseteq \operatorname{ri}(\operatorname{cl}(D))$. A well-known result (Brøndsted, 1983, Theorem 3.4(d)) states that $\operatorname{ri}(\operatorname{cl}(C)) = \operatorname{ri}(C)$ for any convex set C in a finite-dimensional vector space, whence $\operatorname{int}(\operatorname{cl}(D)) \subseteq \operatorname{ri}(D)$. But $\operatorname{ri}(D)$ is a subset of D, so $\operatorname{int}(\operatorname{cl}(D)) \subseteq D$, and hence indeed $\operatorname{int}(\operatorname{cl}(D)) \cap D^c = \emptyset$.

Now consider $D \cap \Lambda$, a subset of D. Since both cl and int respect set inclusion, we find that $\operatorname{int}(\operatorname{cl}(D \cap \Lambda)) \subseteq \operatorname{int}(\operatorname{cl}(D)) \subseteq D$, whence indeed $\operatorname{int}(\operatorname{cl}(D \cap \Lambda)) \cap D^c = \emptyset$. □

Lemma 22. *Consider a non-zero real linear functional Λ_1 on the n-dimensional real vector space $\mathcal{L}(\mathcal{X})$, and a sequence of non-zero real linear functionals Λ_k defined on the $n-k+1$-dimensional real vector space $\ker \Lambda_{k-1}$ for all k in $\{2,\ldots,\ell\}$, where $\ell \in \{2,\ldots,n\}$. Assume that all Λ_k are positive in the sense that $(\forall f \in \mathcal{L}_{\geq 0} \cap \operatorname{dom} \Lambda_k)(\Lambda_k(f) \geq 0)$, for all*

$k \in \{1, \dots, \ell\}$. *Then for each k in $\{2, \dots, \ell\}$ the real linear functional Λ_k on $\ker \Lambda_{k-1}$ can be extended to a real linear functional Γ_k on $\mathcal{L}(\mathcal{X})$ with the following properties:*

(i) *For all f in $\mathcal{L}_{\geq 0}$: $\Gamma_k(f) \geq 0$;*
(ii) $\Gamma_k(1) > 0$;
(iii) $\ker \Gamma_k \cap \ker \Lambda_{k-1} = \ker \Lambda_k$;
(iv) *For all f in $\ker \Lambda_{k-1}$: $\Gamma_k(f) > 0 \Leftrightarrow \Lambda_k(f) > 0$.*

Proof. Fix any k in $\{2, \dots, \ell\}$. Since the real functional Λ_k on the $n-k+1$-dimensional real vector space $\ker \Lambda_{k-1}$ is non-zero, there is some h_k in $\ker \Lambda_{k-1}$ such that $\Lambda_k(h_k) > 0$. We will consider the quotient space $\mathcal{L}(\mathcal{X})/\ker \Lambda_k$, a k-dimensional vector space whose elements $f/\ker \Lambda_k = f + \ker \Lambda_k$ are the affine subspaces through f, parallel to the subspace $\ker \Lambda_k$, for $f \in \mathcal{L}(\mathcal{X})$. We first show that it follows from Theorem 20 that there is a non-zero linear functional $\tilde{\Gamma}_k$ on $\mathcal{L}(\mathcal{X})/\ker \Lambda_k$ such that

$$\tilde{\Gamma}_k(u) \leq 0 \text{ for all } u \text{ in } \mathcal{W}_k^2 := \left\{-\mathbb{I}_{\{x\}}/\ker \Lambda_k : x \in \mathcal{X}_k\right\}, \text{ and}$$
$$\tilde{\Gamma}_k(u) > 0 \text{ for all } u \text{ in } \mathcal{W}_k^1 := \{h_k/\ker \Lambda_k\} \cup \left\{\mathbb{I}_{\{x\}}/\ker \Lambda_k : x \in \mathcal{X}_k\right\}, \quad (17)$$

where we let $\mathcal{X}_k := \{x \in \mathcal{X} : \mathbb{I}_{\{x\}} \notin \ker \Lambda_k\} \subseteq \mathcal{X}$. The set $\mathcal{X}_k$ is non-empty: since $\ker \Lambda_k$ is $n-k$-dimensional, at most $n-k$ of the linearly independent indicators $\mathbb{I}_{\{x\}}$, $x \in \mathcal{X}$ may lie in $\ker \Lambda_k$, so $|\mathcal{X}_k| \geq k$. To show that we can apply Theorem 20, we prove that the condition for it is satisfied: $\sum_{i=1}^n \lambda_i w_i^1 - \sum_{k=1}^m \mu_k w_k^2 \neq 0$ for all m and n in $\mathbb{N}$, all λ_1, …, λ_m in $\mathbb{R}_{\geq 0}$ with $\lambda_i > 0$ for at least one i in $\{1, \dots, m\}$, all μ_1, …, μ_n in $\mathbb{R}_{\geq 0}$, all w_1^1, …, w_n^1 in $\mathcal{W}_k^1$, and all w_1^2, …, w_m^2 in $\mathcal{W}_k^2$. Since $\mathcal{W}_k^1$ and $\mathcal{W}_k^2$ are finite, it is not difficult to see that it suffices to consider $\sum_{i=1}^n \lambda_i w_i^1 = \lambda h_k/\ker \Lambda_k + \sum_{x \in \mathcal{X}_k} \lambda_x \mathbb{I}_{\{x\}}/\ker \Lambda_k$ and $\sum_{j=1}^m \mu_j w_j^2 = -\sum_{x \in \mathcal{X}_k} \mu_x \mathbb{I}_{\{x\}}/\ker \Lambda_k$. So assume *ex absurdo* that $\lambda h_k/\ker \Lambda_k + \sum_{x \in \mathcal{X}_k}(\lambda_x + \mu_x)\mathbb{I}_{\{x\}}/\ker \Lambda_k = 0$, or equivalently, that $\lambda h_k + \sum_{x \in \mathcal{X}_k}(\lambda_x + \mu_x)\mathbb{I}_{\{x\}} \in \ker \Lambda_k$ for some $\mu_x \geq 0$, $\lambda_x \geq 0$ and $\lambda \geq 0$ for all x in $\mathcal{X}_k$, where λ or at least one of $\{\lambda_x : x \in \mathcal{X}_k\}$ are positive. Let $\mathcal{X}'_k := \{x \in \mathcal{X}_k : \lambda_x + \mu_x > 0\}$ and $g := \sum_{x \in \mathcal{X}'_k}(\lambda_x + \mu_x)\mathbb{I}_{\{x\}}$, then we know that $\lambda h_k + g \in \ker \Lambda_k$.

There are now a number of possibilities. The first is that $\lambda = 0$, whence $\mathcal{X}'_k \neq \emptyset$ and therefore $g \in \ker \Lambda_k \subseteq \dots \subseteq \ker \Lambda_1$. This implies that $0 = \Lambda_1(g) = \sum_{x \in \mathcal{X}'_k}(\lambda_x + \mu_x)\Lambda_1(\mathbb{I}_{\{x\}})$. Since all $\mathbb{I}_{\{x\}} \succ 0$ and Λ_1 is positive, we find that $\mathbb{I}_{\{x\}} \in \ker \Lambda_1 = \operatorname{dom} \Lambda_2$ for all x in $\mathcal{X}'_k$. This in turn allows us to conclude that $0 = \Lambda_2(g) = \sum_{x \in \mathcal{X}'_k}(\lambda_x + \mu_x)\Lambda_2(\mathbb{I}_{\{x\}})$. Since all $\mathbb{I}_{\{x\}} \succ 0$ and Λ_2 is positive, we find that $\mathbb{I}_{\{x\}} \in \ker \Lambda_2 = \operatorname{dom} \Lambda_3$ for all x in $\mathcal{X}'_k$. We can go on in this way until we eventually conclude that $0 = \Lambda_k(g) = \sum_{x \in \mathcal{X}'_k}(\lambda_x + \mu_x)\Lambda_k(\mathbb{I}_{\{x\}})$. Since all $\mathbb{I}_{\{x\}} \succ 0$ and Λ_k is positive, we find that $\mathbb{I}_{\{x\}} \in \ker \Lambda_k$ for all x in $\mathcal{X}'_k$, a contradiction.

The second possibility is that $\lambda > 0$. If now $\mathcal{X}'_k = \emptyset$, we find that $\lambda h_k \in \ker \Lambda_k$, whence $\lambda \Lambda_k(h_k) = 0$, a contradiction. If $\mathcal{X}'_k \neq \emptyset$, we find that $\lambda h_k + g \in \ker \Lambda_k \subseteq \dots \subseteq \ker \Lambda_1$. Since $h_k \in \ker \Lambda_{k-1} \subseteq \dots \subseteq \ker \Lambda_1$, this implies that $g \in \ker \Lambda_{k-1} \subseteq \dots \subseteq \ker \Lambda_1$ too. This implies that $0 = \Lambda_1(g) = \sum_{x \in \mathcal{X}'_k}(\lambda_x + \mu_x)\Lambda_1(\mathbb{I}_{\{x\}})$. Since all $\mathbb{I}_{\{x\}} \succ 0$ and Λ_1 is positive, we find that $\mathbb{I}_{\{x\}} \in \ker \Lambda_1 = \operatorname{dom} \Lambda_2$ for all x in $\mathcal{X}'_k$. This in turn allows us to conclude that $0 = \Lambda_2(g) = \sum_{x \in \mathcal{X}'_k}(\lambda_x + \mu_x)\Lambda_2(\mathbb{I}_{\{x\}})$. Since all $\mathbb{I}_{\{x\}} \succ 0$ and Λ_2 is positive, we find that $\mathbb{I}_{\{x\}} \in \ker \Lambda_2 = \operatorname{dom} \Lambda_3$ for all x in $\mathcal{X}'_k$. We can go on in this way until we eventually conclude that $0 = \Lambda_{k-1}(g) = \sum_{x \in \mathcal{X}'_k}(\lambda_x + \mu_x)\Lambda_{k-1}(\mathbb{I}_{\{x\}})$. Since all $\mathbb{I}_{\{x\}} \succ 0$ and Λ_{k-1} is positive, we find that $\mathbb{I}_{\{x\}} \in \ker \Lambda_{k-1} = \operatorname{dom} \Lambda_k$ for all x in $\mathcal{X}'_k$. This now allows us to rewrite $\lambda h_k + g \in \ker \Lambda_k$ as $0 = \Lambda_k(\lambda h_k + g) = \lambda \Lambda_k(h_k) + \sum_{x \in \mathcal{X}'_k}(\lambda_x + \mu_x)\Lambda_k(\mathbb{I}_{\{x\}})$. Since all $\mathbb{I}_{\{x\}} \succ 0$ and Λ_k is positive, this implies that $\lambda \Lambda_k(h_k) \leq 0$, a contradiction. We conclude that, indeed, there is a non-zero linear functional $\tilde{\Gamma}_k$ on $\mathcal{L}(\mathcal{X})/\ker \Lambda_k$ that satisfies Equation (17).

We now define the new real linear functional Γ_k on $\mathcal{L}(\mathcal{X})$ by letting

$$\Gamma_k(f) := \tilde{\Gamma}_k(f/\ker \Lambda_k) \text{ for all } f \text{ in } \mathcal{L}(\mathcal{X}).$$

Observe that, since $f = \sum_{x\in\mathcal{X}} f(x)\mathbb{I}_{\{x\}}$, this leads to

$$\Gamma_k(f) = \sum_{x\in\mathcal{X}} f(x)\tilde{\Gamma}_k\big(\mathbb{I}_{\{x\}}/\ker\Lambda_k\big) = \sum_{x\in\mathcal{X}_k} f(x)\tilde{\Gamma}_k\big(\mathbb{I}_{\{x\}}/\ker\Lambda_k\big),$$

where the second equality follows from $\mathbb{I}_{\{x\}} \in \ker\Lambda_k$, and therefore $\mathbb{I}_{\{x\}}/\ker\Lambda_k = 0$, for all $x \in \mathcal{X}\setminus\mathcal{X}_k$. If we also take into account Equation (17), this proves in particular that (i) and (ii) hold.

For the rest of the proof, consider any f in $\ker\Lambda_{k-1}$ and $\lambda := \Lambda_k(f)/\Lambda_k(h_k)$, a well-defined real number because $\Lambda_k(h_k) > 0$. Then $0 = \Lambda_k(f)-\lambda\Lambda_k(h_k) = \Lambda_k(f-\lambda h_k)$, so $f-\lambda h_k \in \ker\Lambda_k$. As a result, $f/\ker\Lambda_k = \lambda h_k/\ker\Lambda_k$ and therefore $\Gamma_k(f) = \tilde{\Gamma}_k(f/\ker\Lambda_k) = \tilde{\Gamma}_k(\lambda h_k/\ker\Lambda_k) = \lambda\tilde{\Gamma}_k(h_k/\ker\Lambda_k)$. Substituting back for λ, we get the equality:

$$\Gamma_k(f)\Lambda_k(h_k) = \tilde{\Gamma}_k(h_k/\ker\Lambda_k)\Lambda_k(f).$$

Since both $\Lambda_k(h_k) > 0$ and $\tilde{\Gamma}_k(h_k/\ker\Lambda_k) > 0$ [by Equation (17)], we see that $\Gamma_k(f)$ and $\Lambda_k(f)$ are either both zero, both (strictly) positive, or both (strictly) negative. This proves (iii) and (iv). □

Theorem 23. *Given a lexicographic probability system $p = (p_1,\dots,p_\ell)$ that has no non-trivial Savage-null events, the set of desirable gambles $D_p := \{f \in \mathcal{L}(\mathcal{X}) : 0 \prec_p f\}$ corresponding with the preference relation $\prec_p$, is an element of $\bar{\mathcal{D}}_{\mathrm{L}}$—a coherent and lexicographic set of desirable gambles. Conversely, given a lexicographic set of desirable gambles D in $\bar{\mathcal{D}}_{\mathrm{L}}$, its corresponding preference relation $\lessdot_D$ is a preference relation based on some lexicographic probability system $p = (p_1,\dots,p_\ell)$ that has no non-trivial Savage-null events.*

Proof. We begin with the first statement. We first show that D_p is coherent. For Axiom D_1, infer from $0\not\prec_p 0$ by the irreflexivity of $\prec_p$ [see Proposition 18] that indeed $0 \notin D_p$. For Axiom D_2, consider any f in $\mathcal{L}_{>0}$. Use Proposition 19 to infer that $0 \prec_p f$, whence indeed $f \in D_p$. For Axiom D_3, consider any f in D_p and λ in $\mathbb{R}_{>0}$. Then $0 \prec_p f$, and hence $0 \prec_p \lambda f$ using Proposition 19. Then indeed $\lambda f \in D_p$. For Axiom D_4, consider any f and g in D_p, whence $0 \prec_p f$ and $0 \prec_p g$. From $0 \prec_p g$ infer that $f \prec_p f+g$ by Proposition 19, and using $0 \prec_p f$, that $0 \prec_p f+g$ by the transitivity of $\prec_p$ [see Proposition 18]. Then indeed $f+g \in D_p$.

So it only remains to show that $\operatorname{posi}(D_p^c) = D_p^c$. Consider any f and g in D_p^c and any λ_1 and λ_2 in $\mathbb{R}_{>0}$, then we must prove that $\lambda_1 f + \lambda_2 g \in D_p^c$. Since by assumption $0\not\prec_p f$ and $0\not\prec_p g$, Equation (16) guarantees that

$$\mathrm{E}_p(f) \leq_{\mathrm{L}} \mathrm{E}_p(0) = 0 \text{ and } \mathrm{E}_p(g) \leq_{\mathrm{L}} \mathrm{E}_p(0).$$

By the linearity of the expectation operator,

$$\mathrm{E}_p(\lambda_1 f + \lambda_2 g) \leq_{\mathrm{L}} \mathrm{E}_p(0) = 0,$$

whence $0\not\prec_p \lambda_1 f + \lambda_2 g$. Then indeed $\lambda_1 f + \lambda_2 g \in D_p^c$.

For the second statement, we consider any D in $\bar{\mathcal{D}}_{\mathrm{L}}$, and we construct a lexicographic probability system p with no non-trivial Savage-null events and such that $\prec_p$ equals $\lessdot_D$. Define the real functional Λ_1 on $\mathcal{L}(\mathcal{X})$ by letting $\Lambda_1(f) := \sup\{\alpha \in \mathbb{R} : f-\alpha \in D\}$ for all f in $\mathcal{L}(\mathcal{X})$. Proposition 15 guarantees that Λ_1 is a linear functional. Its kernel $\ker\Lambda_1$ is an $n-1$-dimensional linear space,[5] where n is the finite dimension of the real vector space $\mathcal{L}(\mathcal{X})$—the cardinality of $\mathcal{X}$. Since both D^c and $\ker\Lambda_1$ are convex cones, so is their intersection $D^c \cap \ker\Lambda_1$, and it contains 0 because $0 \in D^c$ and $0 \in \ker\Lambda_1$. Using similar arguments, we see that $D \cap \ker\Lambda_1$ is either a convex cone or empty. When $D\cap\ker\Lambda_1 = \emptyset$,

[5]To see that $\ker\Lambda_1$ is a linear space, note that by Proposition 15 $\Lambda_1 = \underline{P}_D$ is a linear prevision—so Λ_1 is a linear map from the n-dimensional linear space $\mathcal{L}(\mathcal{X})$ to $\mathbb{R}$. Since Λ_1 is a linear map, its kernel is closed under addition and scalar multiplication, so it is a linear space, and by the rank-nullity theorem, its dimension is $(\dim\mathcal{L}(\mathcal{X})) - \dim\mathbb{R} = n-1$.

let $\ell := 1$, and stop. When $D \cap \ker \Lambda_1 \neq \emptyset$, it follows from Theorem 20 that there is some non-zero (continuous) linear functional Λ_2 on $\ker \Lambda_1$ such that

$$\Lambda_2(f) \leq 0 \text{ for all } f \text{ in } D^c \cap \ker \Lambda_1 \text{ and } \Lambda_2(f) \geq 0 \text{ for all } f \text{ in } D \cap \ker \Lambda_1.$$

[Apply Theorem 20 with $\mathcal{B} = \ker \Lambda_1$, $\mathcal{W}_2 = D^c \cap \ker \Lambda_1$ and $\mathcal{W}_1 = \mathrm{cl}(D \cap \ker \Lambda_1)$ (the topological closure of $D \cap \ker \Lambda_1$ in $\ker \Lambda_1$); then $\mathrm{int}(\mathcal{W}_1) \cap \mathcal{W}_2 = \emptyset$ by Lemma 21, and $0 \in \mathcal{W}_1 \cap \mathcal{W}_2$] $\ker \Lambda_2$ is a $n-2$-dimensional linear space. Also, $D \cap \ker \Lambda_2$ is either empty or a non-empty convex cone. If it is empty, let $\ell := 2$; otherwise, we repeat the same procedure again: it follows from Theorem 20 that there is some non-zero (continuous) linear functional Λ_3 on $\ker \Lambda_2$ such that

$$\Lambda_3(f) \leq 0 \text{ for all } f \text{ in } D^c \cap \ker \Lambda_2 \text{ and } \Lambda_3(f) \geq 0 \text{ for all } f \text{ in } D \cap \ker \Lambda_2.$$

[Apply Theorem 20 with $\mathcal{B} = \ker \Lambda_2$, $\mathcal{W}_2 = D^c \cap \ker \Lambda_2$ and $\mathcal{W}_1 = \mathrm{cl}(D \cap \ker \Lambda_2)$ (the topological closure of $D \cap \ker \Lambda_2$ in $\ker \Lambda_2$); then $\mathrm{int}(\mathcal{W}_1) \cap \mathcal{W}_2 = \emptyset$ by Lemma 21, and $0 \in \mathcal{W}_1 \cap \mathcal{W}_2$] $\ker \Lambda_3$ is a $n-3$-dimensional linear space. Also, $D \cap \ker \Lambda_3$ is either empty or a non-empty convex cone. If it is empty, let $\ell := 3$; if not, continue in the same vein. This leads to successive linear functionals Λ_k defined on the $n-k+1$-dimenional linear spaces $\ker \Lambda_{k-1}$ such that

$$\Lambda_k(f) \leq 0 \text{ for all } f \text{ in } D^c \cap \ker \Lambda_{k-1} \text{ and } \Lambda_k(f) \geq 0 \text{ for all } f \text{ in } D \cap \ker \Lambda_{k-1}. \tag{18}$$

This sequence stops as soon as $D \cap \ker \Lambda_k = \emptyset$, and we then let $\ell := k$. Because the finite dimensions of the successive $\ker \Lambda_k$ decrease with 1 at each step, we are guaranteed to stop after at most n repetitions: should $D \cap \ker \Lambda_k \neq \emptyset$ for all $k \in \{1, \ldots, n-1\}$ then $\ker \Lambda_n$ will be the 0-dimensional linear space $\{0\}$, and then necessarily $D \cap \ker \Lambda_n = \emptyset$. For the last functional Λ_ℓ, we have moreover that

$$\Lambda_\ell(f) > 0 \text{ for all } f \text{ in } D \cap \ker \Lambda_{\ell-1}. \tag{19}$$

To see this, recall that by construction $\Lambda_\ell(f) \geq 0$ for all f in $D \cap \ker \Lambda_{\ell-1}$, and that $D \cap \ker \Lambda_\ell = \emptyset$.

In this fashion we obtain ℓ linear functionals $\Lambda_1, \ldots, \Lambda_\ell$, each defined on the kernel of the previous functional—except for the domain $\mathcal{L}(\mathcal{X})$ of Λ_1. We now show that we can turn the $\Lambda_2, \ldots, \Lambda_\ell$ into expectation operators: positive and normalised linear functionals on the linear space $\mathcal{L}(\mathcal{X})$. Indeed, consider their respective extensions $\Gamma_2, \ldots, \Gamma_\ell$ to $\mathcal{L}(\mathcal{X})$ from Lemma 22, and let $\Gamma_1 := \Lambda_1$. They satisfy $\Gamma_k(1) > 0$ for all $k \in \{1, \ldots, \ell\}$; see Proposition 15 and Lemma 22(ii). Now consider the real linear functionals on $\mathcal{L}(\mathcal{X})$ defined by $\mathrm{E}_1 := \Gamma_1$, and $\mathrm{E}_k(f) := \Gamma_k(f)/\Gamma_k(1)$ for all k in $\{2, \ldots, \ell\}$ and f in $\mathcal{L}(\mathcal{X})$. It is obvious from Proposition 15 and Lemma 22(i) that these linear functionals are normalised and positive, and therefore expectation operators on $\mathcal{L}(\mathcal{X})$. Indeed each E_k is the expectation operator associated with the mass function p_k defined by $p_k(x) := \mathrm{E}_k(\mathbb{I}_{\{x\}})$ for all x in $\mathcal{X}$. In this way, $p := (p_1, \ldots, p_\ell)$ defines a lexicographic probability system.

We now prove that p has no non-trivial Savage-null events, using Proposition 17. Assume *ex absurdo* that there is some x^* in $\mathcal{X}$ such that $p_k(x^*) = \mathrm{E}_k(\mathbb{I}_{\{x^*\}}) = 0$ for all k in $\{1, \ldots, \ell\}$. Then $\mathbb{I}_{\{x^*\}} \in \ker \Gamma_1 = \ker \Lambda_1$ and $\mathbb{I}_{\{x^*\}} \in \Gamma_k$ for all k in $\{2, \ldots, \ell\}$. Invoke Lemma 22(iii) to find that $\mathbb{I}_{\{x^*\}} \in \ker \Lambda_1 \cap \ker \Gamma_2 = \ker \Lambda_2$. Repeated application of this same lemma eventually leads us to conclude that $\mathbb{I}_{\{x^*\}} \in \ker \Lambda_{\ell-1}$ and $\mathbb{I}_{\{x^*\}} \in \ker \Lambda_\ell$. Since also $\mathbb{I}_{\{x^*\}} \in D$ and hence $\mathbb{I}_{\{x^*\}} \in D \cap \ker \Lambda_{\ell-1}$ [Axiom D_2], Equation (19) implies that $\Lambda_\ell(\mathbb{I}_{\{x^*\}}) > 0$, a contradiction.

It now only remains to prove that $\prec_D$ is the lexicographic ordering with respect to *this* lexicographic probability system, or in other words that

$$f \in D \Leftrightarrow 0 <_{\mathrm{L}} (\mathrm{E}_1(f), \ldots, \mathrm{E}_\ell(f)) \text{ for all } f \text{ in } \mathcal{L}(\mathcal{X}).$$

For necessity, assume that $f \in D$. Then $\mathrm{E}_1(f) \geq 0$ by the definition of Λ_1. If $\mathrm{E}_1(f) > 0$, then we are done. So assume that $\mathrm{E}_1(f) = 0$. Then $f \in \ker \Lambda_1$ and $\Lambda_2(f) \geq 0$ by Equation (18). Again, if $\Lambda_2(f) > 0$, we can invoke Lemma 22(iv) to find that $\Gamma_2(f) > 0$ and hence $\mathrm{E}_2(f) > 0$, and we are done. So assume that $\Lambda_2(f) = 0$. Then $f \in \ker\Lambda_2$ and $\Lambda_3(f) \geq 0$ by Equation (18). We can go on in this way, and we call k the largest number for which $\mathrm{E}_j(f) = 0$ for all j in $\{1,\dots,k-1\}$, or in other words, the smallest number for which $\mathrm{E}_k(f) > 0$. Then $k \leq \ell$ by construction—see Equation (19))—, whence indeed $0 <_{\mathrm{L}} (\mathrm{E}_1(f),\dots,\mathrm{E}_\ell(f))$.

For sufficiency, assume that $0 <_{\mathrm{L}} (\mathrm{E}_1(f),\dots,\mathrm{E}_\ell(f))$, meaning that there is some k in $\{1,\dots,\ell\}$ for which $\mathrm{E}_j(f) = 0 = \Gamma_j(f)$ for all j in $\{1,\dots,k-1\}$ and $\mathrm{E}_k(f) > 0$, whence also $\Gamma_k(f) > 0$. So $f \in \ker\Gamma_j$ for all $j \in \{1,\dots,k-1\}$ and therefore repeated application of Lemma 22(iii) tells us that $f \in \ker\Lambda_j$ for all $j \in \{1,\dots,k-1\}$. Since $\Gamma_k(f) > 0$, we infer from Lemma 22(iv) that also $\Lambda_k(f) > 0$, whence indeed $f \in D$ by Equation (18). □

We conclude that the sets of desirable options in $\bar{\mathcal{D}}_{\mathrm{L}}$ are exactly the ones that are representable by a lexicographic probability system that has no non-trivial Savage-null events. This is, of course, the reason why we have called the coherent sets of desirable options in $\bar{\mathcal{D}}_{\mathrm{L}} := \{D \in \bar{\mathcal{D}} : \mathrm{posi}(D^c) = D^c\}$ *lexicographic*.

This does not mean, however, that the correspondence between the two families is bijective:[6] Indeed, consider the binary possibility space $\mathcal{X} = \{a,b\}$ and the lexicographic systems $p = (p_1, p_2)$ and $p' = (p_1, p_2')$, associated with the respective mass functions $p_1 = (1/2, 1/2)$, $p_2 = (0,1)$ and $p_2' = (1/4, 3/4)$ on $\{a,b\}$. Then p and p' have no non-trivial Savage-null events. However, $D_p = D_{p'} = \{f : f(a)+f(b) > 0 \text{ or } f(b) = -f(a) > 0\}$, so we see that there are two different lexicographic probability systems that map to the same lexicographic set of desirable gambles. Note, however, that $\prec_p$ and $\prec_{p'}$ are both equal to $\lessdot_D$—and therefore $\prec_p = \prec_{p'}$—, as is also guaranteed by the following corollary.

Corollary 24. *Consider any coherent lexicographic set of desirable gambles D in $\bar{\mathcal{D}}_{\mathrm{L}}$ and any lexicographic probability system p that has no non-trivial Savage-null events. Then $\lessdot_D = \prec_p \Leftrightarrow D = D_p$. As a consequence, given any coherent lexicographic set of desirable gambles D in $\bar{\mathcal{D}}_{\mathrm{L}}$, then $\lessdot_D = \prec_p$ for all lexicographic probability systems p that have no non-trivial Savage-null events such that $D_p = D$.*

Proof. For the first statement, infer the following chain of equivalences:

$$\begin{aligned}
\lessdot_D = \prec_p &\Leftrightarrow (\forall f \in \mathcal{L})(0 \lessdot_D f \Leftrightarrow 0 \prec_p f) && \text{by Proposition 19(i) and (iii)}\\
&\Leftrightarrow (\forall f \in \mathcal{L})(f \in D \Leftrightarrow f \in D_p) && \text{by the definitions of } \lessdot_D \text{ and } D_p\\
&\Leftrightarrow D = D_p
\end{aligned}$$

The second statement now follows immediately. □

This corollary is important, since it guarantees that $\lessdot_p$ and $\lessdot_{p'}$ differ if and only if D_p and $D_{p'}$ differ: it rules out that two different preference relations $\lessdot_p$ and $\lessdot_{p'}$ (based on two different lexicographic probability systems that have no non-trivial Savage-null events) map to the same coherent lexicographic set of desirable gambles $D_p = D_{p'}$. For more information about this relation—and also taking updating into account—can be found in work by Benavoli et al. (2017), which builds on the important lexicographic separation theorem by Martínez-Legaz (1983).

Remark 1. As pointed out by a reviewer, the second part of Theorem 23 can also be obtained as a consequence of an earlier result by Martínez-Legaz and Vicente-Pérez (2012, Corollary 3.5), considering that lexicographic sets of desirable gambles are hemispaces and the representation of lexicographic probability systems as stochastic matrices. This same result was also used by Benavoli et al. (2017) in their study of the connection between sets of desirable gambles and sets of lexicographic probability systems. It makes

[6] That two different lexicographic systems (represented as stochastic matrices of full rank) may be associated with the same coherent set of desirable gambles can also be inferred from an example by Benavoli et al. (2017, Example 1).

use of a *lexicographic* separation result (Martínez-Legaz, 1983) that is more directly suited for our purpose than the separation results (see Theorem 20) we borrowed from Walley (1991) . However, we feel that there is value in our giving a more direct proof, since it is more directly tailored to, and 'translated' in, the language of sets of desirable gambles. ◊

Lexicographic probability systems can now also be related to specific types of choice functions, through Proposition 12: given a coherent set of desirable options D, the most conservative coherent choice function C_D whose binary choices are represented by $D_C = D$ satisfies the convexity axiom C_5 if and only if D is a lexicographic set of desirable options. We will call $\bar{\mathcal{C}}_{\mathrm{L}} := \{C_D : D \in \bar{\mathcal{D}}_{\mathrm{L}}\}$ the set of *lexicographic choice functions*.

Looking first at the most conservative coherent choice function that corresponds to D and then checking whether it is 'convex', leads rather restrictively to lexicographic choice functions, and is only possible for lexicographic D: convexity and choice based on Walley–Sen maximality are only compatible for lexicographic binary choice. But suppose we turn things around somewhat, first restrict our attention to all 'convex' coherent choice functions from the outset, and then look at the most conservative such choice function that makes the same binary choices as present in some given D:

$$\inf\{C \in \bar{\mathcal{C}} : C \text{ satisfies Axiom } \mathrm{C}_5 \text{ and } D_C = D\}.$$

We infer from Proposition 4 that this infimum is still 'convex' and coherent. It will, of course, no longer be lexicographic, unless D is. The following proposition tells us it still is an infimum of lexicographic choice functions.

Proposition 25. *Consider an arbitrary coherent set of desirable options D. The most conservative coherent choice function C that satisfies Axiom C_5 and $D_C = D$ is the infimum of all lexicographic choice functions $C_{D'}$ with D' in $\bar{\mathcal{D}}_{\mathrm{L}}$ such that $D \subseteq D'$:*

$$\inf\{C \in \bar{\mathcal{C}} : C \textit{ satisfies Axiom } \mathrm{C}_5 \textit{ and } D_C = D\} = \inf\{C_{D'} : D' \in \bar{\mathcal{D}}_{\mathrm{L}} \textit{ and } D \subseteq D'\}.$$

Proof. Denote the choice function on the left-hand side by C_{left}, and the one on the right-hand side by C_{right}. Both are coherent, and so by Axiom C_4b completely characterised by the option sets from which 0 is chosen. Consider any A in $\mathcal{Q}_0$, then we have to show that $0 \in C_{\mathrm{left}}(\{0\} \cup A) \Leftrightarrow 0 \in C_{\mathrm{right}}(\{0\} \cup A)$.

For the direct implication, we assume that $0 \in C_{\mathrm{left}}(\{0\} \cup A)$, meaning that there is some C^* in $\bar{\mathcal{C}}$ that satisfies Axiom C_5, $D_{C^*} = D$ and $0 \in C^*(\{0\} \cup A)$. We have to prove that there is some D^* in $\bar{\mathcal{D}}_{\mathrm{L}}$ such that $D \subseteq D^*$ and $D^* \cap A = \emptyset$ [by Proposition 11], and we will do so by constructing a suitable lexicographic probability system, by a repeated application of an appropriate version of the separating hyperplane theorem [Theorem 20], as in the proof of Theorem 23.

To prepare for this, we prove that $\mathrm{posi}(\{0\} \cup A) \cap D = \emptyset$. Indeed, assume *ex absurdo* that $\mathrm{posi}(\{0\} \cup A) \cap D \neq \emptyset$, so there is some $f \in D$ such that $f \in \mathrm{posi}(\{0\} \cup A)$. Then there is some λ in $\mathbb{R}_{>0}$ such that $g := \lambda f \in \mathrm{CH}(\{0\} \cup A)$. Let $A' := A \cup \{g\}$, so $\{0\} \cup A' \subseteq \mathrm{CH}(\{0\} \cup A)$, whence $0 \in C^*(\{0\} \cup A')$ by Axiom C_5, if we recall that $0 \in C^*(\{0\} \cup A)$. But $f \in D$ implies that $g \in D$, and since $D_{C^*} = D$, also that $g \in D_{C^*}$, or equivalently, $0 \in R^*(\{0, g\})$, by Proposition 11. Version (4) of Axiom C_3a then guarantees that $0 \in R^*(\{0\} \cup A')$, a contradiction.

It follows from this observation that we can apply Theorem 20 to show that there is some non-zero linear functional Λ_1 on $\mathcal{L}$ such that

$$\Lambda_1(f) \leq 0 \text{ for all } f \text{ in } \mathrm{posi}(\{0\} \cup A) \text{ and } \Lambda_1(f) \geq 0 \text{ for all } f \text{ in } D. \tag{20}$$

[Apply Theorem 20 with $\mathcal{B} = \mathcal{L}(\mathcal{X})$, $\mathcal{W}_2 = \mathrm{posi}(\{0\} \cup A)$ and $\mathcal{W}_1 = D \cup \{0\}$, then $\mathrm{int}(\mathcal{W}_1) \cap \mathcal{W}_2 = \emptyset$ since $\mathrm{int}(\mathcal{W}_1) \subseteq D$, and $0 \in \mathcal{W}_1 \cap \mathcal{W}_2$.] Its kernel $\ker \Lambda_1$ is an $n-1$-dimensional linear space, where n is the dimension of $\mathcal{L}(\mathcal{X})$—the cardinality of $\mathcal{X}$. Since both D and $\ker \Lambda_1$ are convex cones, their intersection $\ker \Lambda_1 \cap D$ is either empty or a convex cone. When $\ker \Lambda_1 \cap D = \emptyset$, we let $\ell := 1$, and stop.

When $\ker\Lambda_1 \cap D \neq \emptyset$, it follows from the same version of the separating hyperplane theorem that there is some non-zero linear functional Λ_2 on $\ker\Lambda_1$ such that

$$\Lambda_2(f) \leq 0 \text{ for all } f \text{ in } \ker\Lambda_1 \cap \operatorname{posi}(\{0\} \cup A) \text{ and } \Lambda_2(f) \geq 0 \text{ for all } f \text{ in } \ker\Lambda_1 \cap D.$$

[Apply Theorem 20 with $\mathcal{B} = \ker\Lambda_1$, $\mathcal{W}_2 = \operatorname{posi}(\{0\} \cup A) \cap \ker\Lambda_1$ and $\mathcal{W}_1 = (\ker\Lambda_1 \cap D) \cup \{0\}$, then $\operatorname{int}(\mathcal{W}_1) \cap \mathcal{W}_2 = \emptyset$ since $\mathcal{W}_2 \subseteq \operatorname{posi}(\{0\} \cup A)$ and $\operatorname{int}(\mathcal{W}_1) \subseteq D$, and $0 \in \mathcal{W}_1 \cap \mathcal{W}_2$.] $\ker\Lambda_2$ is a $n-2$-dimensional linear space. As before, $D \cap \ker\Lambda_2$ is either empty or a non-empty convex cone. If it is empty, let $\ell := 2$; otherwise, repeat the same procedure over and over again, leading to successive non-zero linear functionals Λ_k on $\ker\Lambda_{k-1}$ such that

$$\Lambda_k(f) \leq 0 \text{ for all } f \text{ in } \ker\Lambda_{k-1} \cap \operatorname{posi}(\{0\} \cup A) \text{ and } \Lambda_k(f) \geq 0 \text{ for all } f \text{ in } \ker\Lambda_{k-1} \cap D, \tag{21}$$

until eventually we get to the first k such that $D \cap \ker\Lambda_k = \emptyset$, and then let $\ell := k$ and stop. We are guaranteed to stop after at most n repetitions, since $\ker\Lambda_n$ is the 0-dimensional linear space $\{0\}$, for which $D \cap \ker\Lambda_n = \emptyset$. For the last functional Λ_ℓ, we have that

$$\Lambda_\ell(f) > 0 \text{ for all } f \text{ in } D \cap \ker\Lambda_{\ell-1}. \tag{22}$$

To see this, recall that by construction $\Lambda_\ell(f) \geq 0$ for all f in $D \cap \ker\Lambda_{\ell-1}$, and that $D \cap \ker\Lambda_\ell = \emptyset$.

In this fashion we obtain ℓ linear functionals $\Lambda_1, \dots, \Lambda_\ell$, each defined on the kernel of the previous functional—except for the domain $\mathcal{L}(\mathcal{X})$ of Λ_1. We now show that we can turn the $\Lambda_1, \dots, \Lambda_\ell$ into expectation operators: positive and normalised linear functionals on the linear space $\mathcal{L}(\mathcal{X})$. Indeed, consider their respective extensions $\Gamma_2, \dots, \Gamma_\ell$ to $\mathcal{L}(\mathcal{X})$ from Lemma 22, and let $\Gamma_1 := \Lambda_1$. They satisfy $\Gamma_k(1) > 0$ for all k in $\{1, \dots, \ell\}$; see Proposition 15 and Lemma 22(ii). Now consider the real linear functionals on $\mathcal{L}(\mathcal{X})$ defined by $\mathrm{E}_k(f) := \Gamma_k(f)/\Gamma_k(1)$ for all k in $\{1, \dots, \ell\}$ and f in $\mathcal{L}(\mathcal{X})$. It is obvious from Lemma 22(i) that these linear functionals are normalised and positive, and therefore expectation operators on $\mathcal{L}(\mathcal{X})$. Indeed each E_k is the expectation operator associated with the mass function p_k defined by $p_k(x) := \mathrm{E}_k(\mathbb{I}_{\{x\}})$ for all x in $\mathcal{X}$. In this way, $p := (p_1, \dots, p_\ell)$ defines a lexicographic probability system.

We now prove that p has no non-trivial Savage-null events, using Proposition 17. Assume *ex absurdo* that there is some x^* in $\mathcal{X}$ such that $p_k(x^*) = \mathrm{E}_k(\mathbb{I}_{\{x^*\}}) = 0$ for all k in $\{1, \dots, \ell\}$. Then $\mathbb{I}_{\{x^*\}} \in \ker\Gamma_1 = \ker\Lambda_1$ and $\mathbb{I}_{\{x^*\}} \in \Gamma_k$ for all k in $\{2, \dots, \ell\}$. Invoke Lemma 22(iii) to find that $\mathbb{I}_{\{x^*\}} \in \ker\Lambda_1 \cap \ker\Gamma_2 = \ker\Lambda_2$. Repeated application of this same lemma eventually leads us to conclude that $\mathbb{I}_{\{x^*\}} \in \ker\Lambda_{\ell-1}$ and $\mathbb{I}_{\{x^*\}} \in \ker\Lambda_\ell$. Since also $\mathbb{I}_{\{x^*\}} \in D$ and hence $\mathbb{I}_{\{x^*\}} \in D \cap \ker\Lambda_{\ell-1}$ [Axiom D$_2$], Equation (22) implies that $\Lambda_\ell(\mathbb{I}_{\{x^*\}}) > 0$, a contradiction.

If we now let $D^* := \{f \in \mathcal{L}(\mathcal{X}) : 0 <_{\mathrm{L}} (\mathrm{E}_1(f), \dots, \mathrm{E}_\ell(f))\}$, then $D^* \in \bar{\mathcal{D}}_{\mathrm{L}}$ by Theorem 23. If we can show that $D \subseteq D^*$ and $A \cap D^* = \emptyset$, we are done. So first, consider any f in D. Then $\Lambda_1(f) \geq 0$ by Equation (20). If $\Lambda_1(f) > 0$ then also $\mathrm{E}_1(f) > 0$ by Lemma 22(ii), and therefore $f \in D^*$. If $\Lambda_1(f) = 0$ then $\Lambda_2(f) \geq 0$ by Equation (21). If $\Lambda_2(f) > 0$ then also $\mathrm{E}_2(f) > 0$ by Lemma 22(ii)&(iv), and therefore $f \in D^*$. We can go on in this way until we get to the first k for which $\Lambda_k(f) > 0$, and therefore also $\mathrm{E}_k(f) > 0$ by Lemma 22(ii)&(iv), whence therefore $f \in D^*$. We are guaranteed to find such a k because we infer from Equation (22) that $\Lambda_\ell(f) > 0$. This shows that indeed $D \subseteq D^*$.

Secondly, consider any f in A. Then $\Lambda_1(f) \leq 0$ by Equation (20). If $\Lambda_1(f) < 0$ then also $\mathrm{E}_1(f) < 0$ by Lemma 22(ii), and therefore $f \notin D^*$. If $\Lambda_1(f) = 0$ then $\Lambda_2(f) \leq 0$ by Equation (21). If $\Lambda_2(f) < 0$ then also $\mathrm{E}_2(f) < 0$ by Lemma 22(ii)&(iv), and therefore $f \notin D^*$. If we go on in this way, only two things can happen: either there is a first k for which $\Lambda_k(f) < 0$, and therefore also $\mathrm{E}_k(f) < 0$ by Lemma 22(ii)&(iv), whence therefore $f \notin D^*$. Or we find that $\Lambda_k(f) \leq 0$, and therefore also $\mathrm{E}_k(f) \leq 0$ by Lemma 22(ii)&(iv), for all $k \in \{1, \dots, \ell\}$, whence again $f \notin D^*$. This shows that indeed $A \cap D^* = \emptyset$.

For the converse implication, assume that $0 \in C_{\text{right}}(\{0\} \cup A)$. We must prove that there is some $\tilde{C}$ in $\bar{\mathcal{C}}$ that satisfies Axiom C_5, $D_{\tilde{C}} = D$ and $0 \in \tilde{C}(\{0\} \cup A)$. We claim that $\tilde{C} \coloneqq C_{\text{right}}$ does the job. Because we know by assumption that $0 \in C_{\text{right}}(\{0\} \cup A)$, and from Propositions 12 and 4 that C_{right} is coherent and satisfies Axiom C_5, it only remains to prove that $D_{C_{\text{right}}} = D$. To this end, consider any f in $\mathcal{L}(\mathcal{X})$ and recall the following equivalences:

$$
\begin{aligned}
f \in D_{C_{\text{right}}} &\Leftrightarrow 0 \in R_{\text{right}}(\{0, f\}) && \text{[Equation (11)]} \\
&\Leftrightarrow (\forall D' \in \bar{\mathcal{D}}_{\text{L}})(D \subseteq D' \Rightarrow 0 \in R_{D'}(\{0, f\})) && \text{[definition of inf]} \\
&\Leftrightarrow (\forall D' \in \bar{\mathcal{D}}_{\text{L}})(D \subseteq D' \Rightarrow f \in D') && \text{[Proposition 11]} \\
&\Leftrightarrow f \in D, && \text{[Proposition 14 and } \hat{\mathcal{D}} \subseteq \bar{\mathcal{D}}_{\text{L}}]
\end{aligned}
$$

which completes the proof. □

As a consequence of this result, we also have that, for any coherent set of desirable options D,

$$\inf\{C \in \bar{\mathcal{C}} : C \text{ satisfies Axiom } C_5 \text{ and } D \subseteq D_C\} = \inf\{C_{D'} : D' \in \bar{\mathcal{D}}_{\text{L}} \text{ and } D \subseteq D'\}.$$

6. Discussion and future research

One of the advantages of lexicographic probability systems is that they are more informative than single probability measures, and that they allow us to deal with some of the issues that arise when conditioning on events of probability zero. This is also the underlying idea behind some imprecise probability models, such as sets of desirable gambles. In this paper, we have investigated the connection between the two models, by means of the more general theory of coherent choice functions. We have shown that lexicographic probability systems correspond to the convexity axiom that was considered by Seidenfeld et al. when considering choice functions on horse lotteries. The study of this axiom has led to the consideration of what we have called lexicographic sets of desirable gambles.

In addition, we have also discussed the connection between our notion of coherent choice functions on abstract vectors, and the earlier notion for horse lotteries, developed mostly by Seidenfeld et al. (2010). We have proved that by defining choice functions on arbitrary vector spaces—something which also proves useful when studying the implications of an indifference assessment (Van Camp et al., 2017)—we can include choice functions on horse lotteries as a particular case. This allows us in particular to formulate our results for that framework. Note, nevertheless, that there are some differences between Seidenfeld et al.'s (2010) approach and ours, due to the rationality axioms considered and also to the fact that they deal with possibly infinite (but closed) sets of options, whereas our model assumes that choices are always made between finite sets of alternatives. It would be interesting to investigate the extent to which our results can be generalised to infinite option sets.

One of the advantages of Seidenfeld et al.'s (2010) approach is that it leads to a representation theorem, in the sense that any coherent choice function can be obtained as the infimum of an arbitrary family of more informative convex coherent choice functions that essentially correspond to probability mass functions. Based on the results and conclusions derived here, it seems natural to wonder if a similar result can be established in our framework. Unfortunately, the answer to this question is negative: it turns out that in addition to convexity we need another axiom, which we have called *weak Archimedeanity*. With this extra axiom, at least for binary possibility spaces, it turns out a similar representation result can be proved: every such choice function is an infimum of its *lexicographic* dominating choice functions, showing the importance of a study of lexicographic choice functions also from another angle of perspective. The observation that we need an Archimedean axiom is in agreement with Seidenfeld et al.'s (2010) need of their *Archimedean axiom*, which is—-unlike our *weak* Archimedeanity—difficult to join with desirability. We intend to report on these results elsewhere.

On the other hand, we would also like to combine our results with the discussion by Van Camp et al. (2017), and investigate indifference and conditioning for the special case of lexicographic choice functions. In particular, this should allow us to link our work with Blume et al.'s (1991) discussion of conditioning lexicographic probabilities. In this context, Benavoli et al. (2017) have recently made a similar connection between sets of desirable gambles and sets of lexicographic probabilities. Given our characterisation of the special subclass of lexicographic sets of desirable options (or gambles for that matter)—those sets of desirable gambles that are representable by a single lexicographic probability—, we believe that exploring this might provide additional insight into this relation. Finally, it would be of obvious interest to extend our results in Section 5 to lexicographic probability systems defined on infinite spaces.

ACKNOWLEDGMENTS

Gert de Cooman's research was partly funded through project number 3G012512 of the Research Foundation Flanders (FWO). Enrique Miranda acknowledges financial support by project TIN2014-59543-P. The authors would like to express their gratitude to Alessandro Facchini and the anonymous reviewers of both this paper and an earlier conference version for their constructive feedback and relevant references, that led to a number of improvements.

REFERENCES

Mark A Aizerman. New problems in the general choice theory. *Social choice and Welfare*, 2(4):235–282, 1985. doi: 10.1007/BF00292690.

Francis J. Anscombe and Robert J. Aumann. A definition of subjective probability. *The Annals of Mathematical Statistics*, 34:199–205, 1963. doi: 10.1214/aoms/1177704255. URL http://www.jstor.org/stable/2991295.

Kenneth J. Arrow. *Social choice and individual values.* Cowles Foundation Monographs Series. Yale University Press, 1951.

Thomas Augustin, Frank P. A. Coolen, Gert de Cooman, and Matthias C. M. Troffaes, editors. *Introduction to Imprecise Probabilities.* John Wiley & Sons, 2014. ISBN 978-0-470-97381-3.

Alessio Benavoli, Alessandro Facchini, Marco Zaffalon, and José Vicente-Pérez. A polarity theory for sets of desirable gambles. In Alessandro Antonucci, Giorgio Corani, Inés Couso, and Sébastien Destercke, editors, *Proceedings of the Tenth International Symposium on Imprecise Probability: Theories and Applications*, volume 62 of *Proceedings of Machine Learning Research*, pages 37–48. PMLR, 10–14 Jul 2017. URL http://proceedings.mlr.press/v62/benavoli17b.html.

Lawrence Blume, Adam Brandenburger, and Eddie Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica: Journal of the Econometric Society*, pages 61–79, 1991.

Seamus Bradley. How to choose among choice functions. In Thomas Augustin, Serena Doria, Enrique Miranda, and Erik Quaeghebeur, editors, *ISIPTA '15: Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications*, pages 57–66, Pescara, 2015. SIPTA.

Arne Brøndsted. Graduate Texts in Mathematics. Springer, 1983. ISBN 9783540907220,354090722X,038790722X,9780387907222.

Inés Couso and Serafín Moral. Sets of desirable gambles: conditioning, representation, and precise probabilities. *International Journal of Approximate Reasoning*, 52(7): 1034–1055, 2011. doi: 10.1016/j.ijar.2011.04.004.

Fabio Cozman. Some remarks on sets of lexicographic probabilities and sets of desirable gambles. In Thomas Augustin, Serena Doria, Enrique Miranda, and Erik Quaeghebeur, editors, *ISIPTA '15: Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications*, pages 87–96, Pescara, 2015. SIPTA.

Brian A. Davey and Hilary A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, Cambridge, 1990. doi: 10.1017/CBO9780511809088.

Jasper De Bock and Gert de Cooman. Credal networks under epistemic irrelevance: the sets of desirable gambles approach. *International Journal of Approximate Reasoning*, 56(B):178–207, 2015. ISSN 0888-613X. doi: 10.1016/j.ijar.2014.07.002.

Gert de Cooman. Belief models: an order-theoretic investigation. *Annals of Mathematics and Artificial Intelligence*, 45(1–2):5–34, 2005. doi: 10.1007/s10472-005-9006-x.

Gert de Cooman and Enrique Miranda. Irrelevance and independence for sets of desirable gambles. *Journal of Artificial Intelligence Research*, 45:601–640, 2012. doi: 10.1613/jair.3770.

Gert de Cooman and Erik Quaeghebeur. Exchangeability and sets of desirable gambles. *International Journal of Approximate Reasoning*, 53(3):363–395, 2012. doi: 10.1016/j.ijar.2010.12.002. Precisely imprecise: A collection of papers dedicated to Henry E. Kyburg, Jr.

Peter C. Fishburn. *The Foundations of Expected Utility*. Theory and Decision Library 31. Springer Netherlands, 1982. ISBN 978-90-481-8373-9,978-94-017-3329-8.

Joseph Y. Halpern. Lexicographic probability, conditional probability, and nonstandard probability. *Games and Economic Behavior*, 68, 2010. doi: 10.1016/j.geb.2009.03.013.

Preston C. Hammer. Maximal convex sets. *Duke Mathematical Journal*, 22:103–106, 1955. doi: 10.1215/S0012-7094-55-02209-2.

Peter J. Hammond. Elementary non-Archimedean representations of probability for decision theory and games. *Patrick Suppes: Scientific Philosopher*, pages 25–61, 1994. doi: 10.1007/978-94-011-0774-7_2.

Junnan He. A generalized unification theorem for choice theoretic foundations: Avoiding the necessity of pairs and triplets. Economics Discussion Paper 2012-23, Kiel Institute for the World Economy, 2012. URL http://www.economics-ejournal.org/economics/discussionpapers/2012-23.

Richard B. Holmes. *Geometric Functional Analysis and Its Applications*. Springer-Verlag, New York, 1975.

Joseph B. Kadane, Mark J. Schervish, and Teddy Seidenfeld. A Rubinesque theory of decision. *Institute of Mathematical Statistics Lecture Notes-Monograph Series*, 45:45–55, 2004. doi: 10.1214/lnms/1196285378. URL http://www.jstor.org/stable/4356297.

Peter H. Krauss. Representation of conditional probability measures on boolean algebras. *Acta Mathematica Academiae Scientiae Hungaricae*, 19(3-4):229–241, 1968. doi: 10.1007/BF01894506.

Isaac Levi. Potential surprise: its role in inference and decision making. In L. J. Cohen and M. Hesse, editors, *Applications of Inductive Logic*, pages 1–27. Clarendon Press, Oxford, 1980.

Juan Enrique Martínez-Legaz. Exact quasiconvex conjugation. *Zeitschrift für Operations-Research*, 27(1):257–266, Dec 1983. ISSN 1432-5217. doi: 10.1007/BF01916921. URL https://doi.org/10.1007/BF01916921.

Juan Enrique Martínez-Legaz and José Vicente-Pérez. Lexicographical representation of convex sets. *Journal of Convex Analysis*, 19(2):485–496, 2012.

Enrique Miranda and Marco Zaffalon. Notes on desirability and coherent lower previsions. *Annals of Mathematics and Artificial Intelligence*, 60(3–4):251–309, 2010.

Serafín Moral. Epistemic irrelevance on sets of desirable gambles. *Annals of Mathematics and Artificial Intelligence*, 45(1–2):197–214, 2005. doi: 10.1007/s10472-005-9011-0.

Erik Quaeghebeur. Desirability. In Thomas Augustin, Frank P. A. Coolen, Gert de Cooman, and Matthias C. M. Troffaes, editors, *Introduction to Imprecise Probabilities*, chapter 1, pages 1–27. John Wiley & Sons, 2014. doi: 10.1002/9781118763117.ch1.

Erik Quaeghebeur, Gert de Cooman, and Filip Hermans. Accept & reject statement-based uncertainty models. *International Journal of Approximate Reasoning*, 57:69–102, 2015. doi: 10.1016/j.ijar.2014.12.003.

Herman Rubin. A weak system of axioms for “rational" behavior and the nonseparability of utility from prior. *Statistics & Risk Modeling*, 5(1-2):47–58, 1987. doi: 10.1524/strm.1987.5.12.47.

Leonard J. Savage. *The Foundations of Statistics*. Dover, New York, 1972. Second revised edition, first published 1954.

Thomas Schwartz. Rationality and the myth of the maximum. *Noûs*, 6(2):97–117, 1972. doi: 10.2307/2216143.

Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane. Decisions without ordering. In W. Sieg, editor, *Acting and reflecting*, volume 211 of *Synthese Library*, pages 143–170. Kluwer, Dordrecht, 1990.

Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane. Coherent choice functions under uncertainty. *Synthese*, 172(1):157–176, 2010. doi: 10.1007/s11229-009-9470-7.

Amartya Sen. Choice functions and revealed preference. *The Review of Economic Studies*, 38(3):307–317, July 1971. doi: 10.2307/2296384.

Amartya Sen. Social choice theory: A re-examination. *Econometrica*, 45:53–89, 1977. doi: 10.2307/1913287.

Matthias C. M. Troffaes. Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning*, 45(1):17–29, 2007. doi: 10.1016/j.ijar.2006.06.001.

Matthias C. M. Troffaes and Gert de Cooman. *Lower Previsions*. Wiley, 2014.

Hirofumi Uzawa. Note on preference and axioms of choice. *Annals of the Institute of Statistical Mathematics*, 8:35–40, 1956. doi: 10.1007/BF02863564.

Arthur Van Camp, Gert de Cooman, Enrique Miranda, and Erik Quaeghebeur. Modelling indifference with choice functions. In Thomas Augustin, Serena Doria, Enrique Miranda, and Erik Quaeghebeur, editors, *ISIPTA ’15: Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications*, pages 305–314, Pescara, 2015. Aracne.

Arthur Van Camp, Gert de Cooman, Enrique Miranda, and Erik Quaeghebeur. Coherent choice functions, desirability and indifference. *Fuzzy Sets and Systems*, 2017. doi: 10.1016/j.fss.2017.05.019. In press.

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

Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2–3):125–148, 2000. doi: 10.1016/S0888-613X(00)00031-1.

Peter M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975. Revised journal version: Williams (2007).

Peter M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44(3):366–383, 2007. Revised journal version of Williams (1975).

Marco Zaffalon and Enrique Miranda. Desirability and the birth of incomplete preferences. *ArXiv e-prints*, June 2015. URL https://arxiv.org/abs/1506.00529.

GHENT UNIVERSITY, IDLAB, TECHNOLOGIEPARK–ZWIJNAARDE 914, 9052 ZWIJNAARDE, BELGIUM
E-mail address: arthur.vancamp@ugent.be

GHENT UNIVERSITY, IDLAB, TECHNOLOGIEPARK–ZWIJNAARDE 914, 9052 ZWIJNAARDE, BELGIUM
E-mail address: gert.decooman@ugent.be

DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH, UNIVERSITY OF OVIEDO, SPAIN
E-mail address: mirandaenrique@uniovi.es