Coherent Choice Functions under Uncertainty

Teddy Seidenfeld
Mark J. Schervish
Joseph B. Kadane
Carnegie Mellon University
teddy@stat.cmu.edu
mark@stat.cmu.edu
kadane@stat.cmu.edu

Abstract

We discuss several features of *coherent choice functions* – where the admissible options in a decision problem are exactly those which maximize expected utility for some probability/utility pair in fixed set *S* of probability/utility pairs. In this paper we consider, primarily, normal form decision problems under uncertainty – where only the probability component of *S* is indeterminate. Coherent choice distinguishes between each pair of sets of probabilities. We axiomatize the theory of choice functions and show these axioms are necessary for coherence. The axioms are sufficient for coherence using a set of probability/almost-state-independent utility pairs. We give sufficient conditions when a choice function satisfying our axioms is represented by a set of probability/state-independent utility pairs with a common utility.

Keywords. Choice functions, coherence, *Γ-Maximin*, *Maximality*, uncertainty, state-independent utility.

1 Introduction

In this paper we continue our study of coherent choice functions, which we started in our (2004) "Rubinesque" theory of decision. Coherent choice function theory provides a more general account of Imprecise Probabilities than the theory of coherent strict preference, which we used in our (1995). Coherent choice function theory does not reduce to binary comparisons between options, as Example 1 (below) illustrates. By contrast, coherent strict preference is a binary relation that fails, in principle, to distinguish between some convex sets of probabilities that have the same convex hull.

Specifically, as we show in Section 2, with coherent choice functions, for each two different sets of probabilities it requires only a simple decision problem in order to distinguish by admissibility between them. That is, with coherent choice functions, each set of probabilities has its own footprint of admissible options. In Section 4, we illustrate this added generality with a

non-convex (even a disconnected) set *S* of probabilities that share the common structure that, for each distribution in *S*, two specific events are independent. Coherent choice with respect to the set *S* avoids making information about one event valuable in decisions that depend solely on the other event. This is in sharp contrast with theories that rely on convex sets to depict Imprecise Probabilities

Let *O* be a (closed) set of feasible options. A **choice function *C*(*O*)** identifies the (non-empty) subset of *O* that are the admissible options in the decision problem given by the feasible set *O*. We say that *C*(●) is **coherent** provided that there is a non-empty set *S* of probability/utility pairs *S* = {(*p*,*u*)} such that the admissible options under *C* are precisely those that are Bayes with respect to some probability/utility pair (*p*,*u*) in *S*. That is, for each admissible option, for each $o \in C(O)$, there is a pair $(p,u) \in S$ such that *o* maximizes the *p*-expected *u*-utility over *O*. For short, we will call these the *Bayes-admissible options* in *O* (with respect to *S*).

Aside: If the option set *O* is not closed, then given a set *S* there may be no coherently admissible options in *O*. For example, if utility is linear and increasing in the quantity *X*, then in the decision-under-certainty problem with *O* = $\{0 \leq x < 1\}$, each option is inadmissible with respect to *S*.

In Section 3 we adapt Anscombe-Aumann Horse-lottery theory in order to axiomatize coherent choice functions for cases where only probability (not utility) is indeterminate. This affords a representation of choice functions in the style of our previous work (1995), where we represented coherent strict (binary) preference between options using sets of probabilities and almost-state-independent utilities. One way to understand how the new representation generalizes our previous work is to consider the partial order ⟨ defined on pairs of *sets* of options {*O*₁,*O*₂}: where *O*₁ ⟨ *O*₂ obtains whenever there are no admissible options from set *O*₁ in a choice problem given the combined set of options *O*₁ ∪ *O*₂. When the two sets {*O*₁,*O*₂} are singletons, this relation

reduces to the binary comparison of strict preference between options. Because our (1995) theory leads to a representation in terms of sets of probabilities and almost-state-independent utilities, that feature is inherited by our representation in Section 3.

The use of a coherent choice function coincides with Levi's (1980) principle of *E*-admissibility in cases where the set S is a cross-product of a convex set of probabilities and a convex set of utilities: $S = P \times U$ for convex sets P and U. Also, we find that Savage [1954, pp. 123-124, particularly where he argues that option *b* is "superfluous" for the decision pictured by his Figure 1] endorses a coherent choice rule with S a convex set of probabilities and a common utility. The following example, which we repeat from our ISIPTA-03 paper, illustrates how coherent choice does not reduce to binary comparisons in a setting where only probability is indeterminate.

Example 1: Consider a binary decision problem, $\Omega = \{\omega_1, \omega_2\}$ with three feasible options $O = \{f,g,h\}$, and where utility is determinate: $u(f(\omega_1)) = u(g(\omega_2)) = 0.0$, $u(f(\omega_2)) = u(g(\omega_1)) = 1.0$, and $u(h(\omega_1)) = u(h(\omega_2)) = 0.4$. Let uncertainty over the states be indeterminate, with $P = \{p: 0.25 \leq p(\omega_2) \leq .75\}$. We rehearse three decision rules for this problem.

Γ-Maximin – Maximize minimum expected utility over the feasible options. This rule is well studied in Gilboa and Schmeidler (1989). In brief, *Γ-Maximin* induces a preference ordering over options, but fails the von Neumann-Morgenstern *Independence* postulate. Under *Γ-Maximin* only $\{h\}$ is admissible from the set $\{f,g,h\}$.

Maximality (Sen/Walley) – admissible options are those that are undominated in expectations (over $p \in P$) by any single alternative option. Under *Maximality* all three options are admissible from the set $\{f,g,h\}$ as none dominates the others in pairwise comparisons. *Maximality* does not induce a preference ordering over options; however, admissibility is given by pairwise comparisons. As is evident from Example 1, whether an option (e.g., option *h*) is admissible under *Maximality* depends upon whether the feasible options are closed under mixtures.

Coherent choice. Since the set of probabilities P is convex in this example, coherent choice reduces to Levi's rule of. *E*-admissibility – admissible choices have Bayes' models, i.e., they maximize expected utility for some probability in the (convex) set P. Subset $\{f,g\}$ identifies the Bayes-admissible options from $\{f,g,h\}$ under *Coherent Choice*. This rule does not induce an ordering over options and does not reduce to pairwise comparisons.

Note that *h*, which is never "Bayes" with respect to P, is uniformly dominated by some mixtures of *f* and *g*, e.g., the mixed option given by $.5f \oplus .5g$, with expected utility 0.5 independent of p, uniformly dominates *h*. This is no coincidence, as the following result establishes.

Let $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ be a finite partition of states. Let $O = \{o_1, o_2, ..., o_m\}$ be a finite set of options defined on Ω, such that for $o_i \in O$, $u(o_i(\omega_j)) = u_{ij}$, a cardinal utility of the consequence of o_i when state ω_j obtains. Let P be the class of all probability distributions over Ω. Similarly, let Q be the class of all (simple) mixed acts over O, with a mixed act denoted q.

Theorem 1 (Pearce, 1984, p.1048).
Suppose for each $p \in P$, act $o^* \in O$ fails to maximize expected utility. Then there is a mixed alternative q^* that uniformly, strictly dominates o^*. That is, $u(q^*(\omega_j)) > u(o^*(\omega_j)) + \varepsilon$, for $j = 1, ..., n$, with $\varepsilon > 0$.

Aside: With this result we are able to apply the strict standard of de Finetti's "incoherence" (= uniform, strict dominance) to a broad class of decisions under uncertainty, analogous to traditional Complete Class Theorems for Bayes decisions (Wald, 1950). The standard of incoherence used here is notably stronger than the mere *inadmissibility* (= weak dominance) of non-Bayes decisions, as is used in those Complete Class theorems.

Let $H(O)$ denote the result of taking the closed, convex hull of the option set O. That is, $H(O)$ is the set of all (simple) mixed acts based on O. Since O is finite, q^* of Theorem 1 may be taken to be an option that also is Bayes for some $p^* \in P$. That is, in Theorem 1 we may choose $q^* \in H(O)$ such that $q^* \in C(H(O))$ for a coherent choice function using the set P of all probability distributions on Ω.

Aside: Theorem 1 generalizes to infinite states spaces Ω and infinite, closed options sets O by using Theorem 2.1 of Kindler (1983) to replace Pearce's use of von Neumann's Minimax Theorem, which does not generalize to infinite games.

In terms of Theorem 1, in Example 1 with $o^* = h$, then $q_x^* = xf \oplus (1-x)g$ for $.4 < x < .6$ uniformly dominates o^*. But each such q_x^* is Bayes with respect to $H(O)$ for precisely for one probability on Ω: $p(\omega_1) = .5$. We use this fact, next, to establish that each set of probabilities has its own *unique* coherent choice function.

2. Distinguishing sets of probabilities by their coherent choice functions

Consider a finite state space $\Omega = \{\omega_1, ..., \omega_n\}$ with the class of all options given by horse lotteries (Anscombe and Aumann, 1963) defined on two consequences **1** and

0. In general horse-lottery theory there is a denumerable set of prizes, $\{r_1, r_2, ... \}$. A (simple) *horse lottery* is a function from states to (simple) probability distributions over the set of prizes. In this section we use decision problems involving horse lotteries defined on only two consequences, **0** and **1**, with a strict preference for the constant act **1** over the constant act **0**, as explained below. And we consider coherent choice using a state-independent utility, *u* where $u(1) = 1$ and $u(0) = 0$ in each state, ω. Our goal is to show that if *P* and *P*' are different sets of probabilities, the coherent choice function based on $P \times \{u\}$ is different from the coherent choice function based on $P' \times \{u\}$.

Let *P* be a set of probabilities. For a (closed) set *O*, *C*(*O*) is the non-empty set of Bayes-admissible options. Let $R(O) = O \backslash C(O)$ be the associated *Bayes-rejection function* that identifies the *inadmissible* options in *O*. So, we assume that $\{0\} = R\{0, 1\}$.

Let $p = (p_1, …, p_n)$ be a probability distribution on Ω . Let $\underline{p}$ be the smallest nonzero coordinate of *p*.
Define the *constant* horse lottery act $a = \underline{p}\mathbf{1} + (1-\underline{p})\mathbf{0}$.
For each $j = 1, .., n$, *define* the act h_j by

$$
\begin{aligned}
h_j(\omega_i) &= & 1 & \quad \text{if } i = j \text{ and } p_j = 0,\\
&= & a & \quad \text{if } i \neq j \text{ and } p_j = 0,\\
&= & (\underline{p}/p_j)\mathbf{1} + (1- \underline{p}/p_j)\mathbf{0} & \\
& & & \quad \text{if } i = j \text{ and } p_j > 0,\\
&= & 0 & \quad \text{if } i \neq j \text{ and } p_j > 0.
\end{aligned}
$$

Define the option set $O_p = \{a, h_1, …, h_n\}$.
Theorem 2*: $p \in P$ if and only if $R(O_p) = \varnothing$.*

Proof: First, note that for all *j* and every utility *u*, $E_p(u(h_j)) = \underline{p} = E_p(u(a))$. For the "only if" direction, assume that $(p; u) \in S$ for some utility *u*. Then by this equality, every element of O_p is Bayes with respect to $(p; u)$ and $R(O_p) = \phi$. For the "if" direction, assume that $R(O_p) = \phi$. Notice that $E_q(v(a)) = \underline{p}$ for every probability /utility pair (q, v). Let (q, v) be a probability/utility pair with $q \neq p$. First, consider the case with $\underline{p} < 1$. Then there exists *j* with $q_j > p_j$. So,

$$
E_q(v(h_j)) =
\begin{cases}
q_j \underline{p} / p_j > \underline{p} & \text{if } p_j > 0,\\
q_j + (1-q_j)\underline{p} > \underline{p} & \text{if } p_j = 0.
\end{cases}
$$

Hence, for each (q, v) with $q \neq p$, $E_q(v(h_j)) > E_q(v(a))$. It follows that $a \in R(O_p)$ unless $(p, u) \in S$ for some utility *u*. Finally, consider the case with $\underline{p} = 1$. In this case, $O_p = \{\mathbf{1}, h_j\}$ where $p_j = 1$. So, $E_q(v (h_j)) = q_j < 1 = E_q(v (a))$ for every probability/utility pair (q, v) with $q \neq p$. It follows that $h_j \in R(O_p)$ unless $(p, u) \in S$ for some utility *u*.◊

Corollary Let P_1 and P_2 be two distinct (nonempty) sets of probabilities with corresponding Bayes rejection functions R_1 and R_2. There exists a finite option set O_p, as above, such that $R_1(O_p) \neq R_2(O_p)$.

Thus, each set of probabilities *P* has its own distinct pattern of Bayes rejection functions with respect to option sets O_p for $p \in P$.

Aside: This is a generalization of Theorem 1 that appears at the end of our (2004) paper. That Theorem 1 is the restriction of the corollary to pairs of convex sets of probabilities.

3. Axiomatizing coherent choice functions
We turn, next, to a system of axioms for choice functions that are necessary for coherence, and which are jointly sufficient for a representation of choice by a set *S* of probability/almost-state-independent utility pairs, as explained below. We provide sufficient conditions when these pairs have a common state-independent utility. In such a case the coherent choice function corresponds to choice under indeterminate uncertainty with a determinate utility.

We continue within the framework of the previous section: horse lotteries over a finite state space $\Omega = \{\omega_1, …, \omega_n\}$. In that we are using choice functions over sets of options, the theory presented here extends our (1995) work, which deals solely with binary choice problems. Thus, results that follow from binary choice problems are available also within this theory. For example, it follows from Section II.6 of our (1995) theory that each agreeing cardinal utility for the choice function $C(\bullet)$, if one exists, is a bounded utility function.

Aside: The aspects of the theory given here that compel the use of almost-state-independent utilities parallel the same issues that arise in Section IV of our (1995) representation for partially ordered preferences. In the context of this paper, that theory, which addresses binary choice only, can be taken to axiomatize choice under the Maximality rule.

In this paper, we focus on a representation for choice when utility is determinate, i.e., regarding the two distinguished prizes **1** and **0**, the constant act **1** is better than, and the constant act **0** is worse than, all other constant acts. Also, we assume that all cardinal utilities are scaled so that $u(1) = 1$ and $u(0) = 0$.

Given a strict preference between these two prizes, the Anscombe-Aumann (1963) theory of horse-lotteries is given by four substantive axioms, which we summarize as follows.
A-A Axiom 1: Choice over horse lotteries reduces to a pairwise comparison of options since binary preference satisfies *ordering*.
A-A Axiom 2: Preference satisfies the von Neumann-Morgenstern postulate of *Independence*.
A-A Axiom 3: An Archimedean postulate is introduced in order to assure that preference has a real-valued

representation, thus insuring also a real-valued representation for subjective probability over Ω and a real-valued cardinal utility over prizes.

A-A Axiom 4: To insure existence of a state-independent utility representation for preference over horse lotteries, a final axiom requires that the decision maker's preference for constant horse lotteries reproduces under each non-null state in the form of called-off horse lotteries.

We adapt our presentation here to match these four axioms.

Axiom 1a (Sen's property *alpha*):
If $O_2 \subseteq R(O_1)$ and $O_1 \subseteq O_3$, then $O_2 \subseteq R(O_3)$.
You cannot promote an unacceptable option into an acceptable option by adding to the choice set of options.

Axiom 1b (a variant of Aizerman's 1985 condition):
If $O_2 \subseteq R(O_1)$ and $O_3 \subseteq O_2$, then $O_2 \backslash O_3 \subseteq R(closure[O_1 \backslash O_3])$.
You cannot promote an unacceptable option into an acceptable option by deleting unacceptable options from the option set. (We require *closure*$[O_1 \backslash O_3]$ since $O_1 \backslash O_3$ may not be closed, despite closure of O_1 and of O_3.)

With Axioms 1a and 1b, define a strict partial order $\langle$ on sets of options as follows. Let O_1 and O_2 be two option sets.

Defn: $O_1 \langle O_2$ *if and only if* $O_1 \subseteq R[O_1 \cup O_2]$.

Lemma 1 of our (2004) establishes that given Axioms 1a and 1b, the binary relation $\langle$ is a strict partial order over pairs of sets of options: $\langle$ is *transitive* and *anti-symmetric*.

The role of mixtures is captured in the following pair of axioms for $\langle$. With O_1 an option set and o an option, the notation $\alpha O_1 \oplus (1-\alpha)o$ denotes the set of pointwise mixtures, $\alpha o_1 \oplus (1-\alpha)o$ for $o_1 \in O_1$.

Axiom 2a Independence is formulated for the relation $\langle$ over sets of options. Let o be an option and $0 < \alpha \le 1$.
$O_1 \langle O_2$ *if and only if* $\alpha O_1 \oplus (1-\alpha)o \langle \alpha O_2 \oplus (1-\alpha)o$.
Axiom 2b Mixtures If $o \in O$ and $o \in R[H(O)]$, then $o \in R[O]$.
Axiom 2b asserts that unacceptable options from a mixed set remain so even before mixing.

Aside: *Independence* (Axiom 2a) fails in Γ-*Maximin* theory. *Mixing* (Axiom 2b) fails for the choice function determined by *Maximality*.

The Archimedean condition requires a technical adjustment, as the canonical form used by, e.g. von Neumann-Morgenstern theory or Anscombe-Aumann theory is too restrictive in this setting. (See section II.4 of our 1995.) The reformulated version of the Archimedean condition is as a continuity principle compatible with strict preference as a strict partial order. It reads as follows.

Let A_n and B_n ($n = 1, \dots$) be sets of options converging pointwise, respectively, to the option sets A and B. Let N be an option set.
Axiom 3a: If, for each n, $B_n \langle A_n$ and $A \langle N$, then $B \langle N$.
Axiom 3b: If, for each n, $B_n \langle A_n$ and $N \langle B$, then $N \langle A$.

State-neutrality / dominance is captured by the following relations. Consider horse lotteries h_1 and h_2, with $h_i(\omega_j) = \beta_{ij}\mathbf{1} \oplus (1-\beta_{ij})\mathbf{0}$; $i = 1,2$ $j = 1, \dots, n$.

Definition: h_2 **weakly dominates** h_1 if $\beta_{2j} \ge \beta_{1j}$ for $j = 1, \dots, n$.

Axiom 4: Assume that o_2 weakly dominates o_1, and that a is an option different from each of these two.
4a: If $o_2 \in O$ and $a \in R(\{o_1\} \cup O)$ then $a \in R(O)$.
4b: If $o_1 \in O$ and $a \in R(O)$ then $a \in R(\{o_2\} \cup O\backslash o_1)$.

In words, Axiom *4a* says that when a weakly dominated option is removed from the set of options, other inadmissible options remain inadmissible. So, by Axiom 1, when an option is replaced in the option set by one that it weakly dominates, other admissible options remain admissible.

Axiom *4b* says that when an option is replaced by one that weakly dominates it, (other) inadmissible options remain inadmissible. Trivially by Axiom 1, merely adding a weakly dominating option cannot promote an inadmissible option into one that is admissible.

Axiom 4 captures key aspects of what Savage's postulate **P3** asserts about state-independent utility of the prizes **1** and **0** without assuming states are not-null. That is, the intended representation for the choice function $C(\bullet)$ is by the expected utility rule applied with a set of probability distributions P. However, it may be that for each state ω_j there is a probability distribution $p_j \in P$ such that $P_j(\omega_j) = 0$. In the language of our (1995) paper, each state in Ω is potentially null under P. Then Savage's **P3** (or the corresponding axiom of Anscombe-Aumann theory) is vacuous. Nonetheless, Axiom 4 reports two facts about weakly dominated lotteries that obtain even when each state is potentially null.

Theorem 3: Axioms 1–4 are necessary for a coherent choice function.
That is, let S be a non-empty set of pairs of probability/state-independent utilities, and let $C_S(\bullet)$ be the coherent choice function defined by setting the admissible options in feasible set O to be exactly those that are Bayes-admissible with respect to S. Then $C_S(\bullet)$ satisfies Axioms 1–4.

Proof: The argument for the necessity of Axioms 1–3 is given in our (2004). That Axiom 4 is necessary as well follows immediately by noting that whenever o_2 weakly dominates o_1 then for each $(p,u) \in S$, $E_p(u(o_2)) \geq E_p(u(o_1))$. ◊

The following result is helpful in linking our theory with Theorem 1.
Definition: h_2 *strongly dominates* h_1 if $\beta_{2j} > \beta_{1j}$ for $j = 1, ..., n$.
Lemma–Inadmissibility of strongly dominated options: If h_2 strongly dominates h_1 then $\{h_1\} = R(\{h_1, h_2\})$.

Proof: The strategy of the proof is as follows: Use the Independence axiom to convert the problem with option set $O = \{h_1, h_2\}$ into an equivalent problem $O` = \{h`_1, h`_2\}$, where $h`_1$ is a constant horse lottery, and where $h`_2$ strongly dominates $h`_1$. Then we show that $h`_2$ weakly dominates another constant horse lottery, $h``_2$ which also strongly dominates $h`_1$. Then, by Independence $\{h`_1\} = R(\{h`_1, h``_2\})$ and by Axiom *4b*, $\{h`_1\} = R(\{h`_1, h`_2\})$ Last, by Independence, $\{h_1\} = R(\{h_1, h_2\})$.

Here are the details. Let $0 \leq \beta_* = min\{\beta_{1j}\}$ and $1 > \beta^* = max\{\beta_{1j}\}$. Let $h_3(\omega_j) = \beta_{3j}\mathbf{1} \oplus (1-\beta_{3j})\mathbf{0}$, where $\beta_{3j} = \beta^* + \beta_* - \beta_{1j}$. Then the horse lottery $h`_1 = .5h_1 \oplus .5h_3$ is the constant (von Neumann-Morgenstern) lottery with $\beta`_{1j} = (\beta^* + \beta_*)/2$. Define $h`_2 = .5h_2 \oplus .5h_3$. The Independence axiom asserts that $\{h_1\} = R(\{h_1, h_2\})$ if and only if $\{h_1`\} = R(\{h_1`, h_2`\})$. But $h_2`$ strongly dominates $h_1`$, because h_2 strongly dominates h_1. In fact, $\beta`_{2j} - \beta`_{1j} = (\beta_{2j} - \beta_{1j})/2 > 0$. So, let $0 < \delta = min\{\beta_{2j} - \beta_{1j}\}$, and then $\delta/2 = min\{\beta`_{2j} - \beta`_{1j}\}$. Let $h``_2$ be the constant (von Neumann-Morgenstern) lottery defined with $\beta``_{2j} = \beta`_{1j} + \delta/2 = (\beta^* + \beta_* + \delta)/2 > \beta`_{1j}$. Observe, also, that $h`_2$ weakly dominates $h``_2$.

Then, as announced before, by Independence $\{h`_1\} = R(\{h`_1, h``_2\})$ and by Axiom *4b*, $\{h`_1\} = R(\{h`_1, h`_2\})$, and by another application of Independence, $\{h_1\} = R(\{h_1, h_2\})$. ◊

Next we introduce two concepts central to our argument for representing coherent choice functions.

Definitions: The pair (p,u) is *a local Bayes model for option* o provided that o maximizes (p,u)-expected utility with respect to the options in set O.

The pair (p,u) is *a global Bayes model for the choice function* $C(\bullet)$ provided that, for each option set O, if $o \in O$ maximizes (p,u)-expected utility with respect to the options in set O then $o \in C(O)$.

We adapt the concept of a set of *almost state-independent utilities*, presented in our (1995, Definition 31), as follows. Let $\{r_1, ..., r_m\}$ be a set of rewards and assume that for each constant horse lottery $r \in \{r_1, ..., r_m\}$, $\{\mathbf{0}\} \langle \{r\} \langle \{\mathbf{1}\}$, so that the constant acts $\mathbf{0}$ and $\mathbf{1}$ strictly bound the value of the other constant acts.

The set of probability/utility pairs $S^{\#} = \{(p_j, u_j): j = 1, ...\}$ form a set of *almost state independent utilities* for $\{r_1, ..., r_m\}$ provided that for each $\varepsilon > 0$, there is a pair $(p_\varepsilon, u_\varepsilon) \in S^{\#}$ and a set of states $\Omega_{(1-\varepsilon)} \subseteq \Omega$ with $p_\varepsilon(\Omega_{(1-\varepsilon)}) \geq 1-\varepsilon$ such that for each $r \in \{r_1, ..., r_m\}$

$$max_{\omega_i, \omega_j \in \Omega_{1-\varepsilon}} | u_{\varepsilon,\omega_i}(r) - u_{\varepsilon,\omega_j}(r) | \leq \varepsilon.$$

The remaining theorem we seek is this one.

Theorem 4: A choice function $C(\bullet)$ defined on the class $\mathcal{H}$ of simple Anscombe-Aumann Horse-lotteries using (at least) three prizes $\{\mathbf{0}, r, \mathbf{1}\}$, with $\{\mathbf{0}\} \langle \{r\} \langle \{\mathbf{1}\}$, satisfies our 4 axioms only if it is represented by a set S of probability/almost-state-independent utility pairs.

A sufficient condition is given at the end of Appendix 2 for the global Bayes models of S to be comprised solely of probability/state-independent utility pairs.

This theorem follows from three lemmas, described next.

Lemma 1: For each choice set O and admissible option $o \in C(O)$, o has at least one local Bayes model.
Proof: By Theorem 1, an option lacking a local Bayes model is strongly dominated by a finite mixture of other options already available in the same choice problem. Then, Axiom 3 and the Lemma on inadmissibility of strongly dominated acts demonstrates Lemma 1.

Aside: Let $o \in C(O)$. If (p,u) and (p',u) both are local Bayes models for o, then so too is each pair (q,u) of the form $q = xp + (1-x)p'$ ($0 \leq x \leq 1$). Likewise, if each of (p_j,u) ($j = 1, ...$) is a local Bayes model for o and the sequences of distributions $\{p_j\}$ converges to distribution q, then also (q,u) is a local Bayes model for o. Hence, we have the following corollary

Corollary: The set of local Bayes models for $o \in C(O)$ with a common utility u is given by a non-empty, closed, convex set of probabilities.

Next, following the ideas presented in Section 2, given a distribution p, we identify a special choice problem O^* so that precisely when all of its options are admissible, then p is a global Bayes model for the choice function. Thus, the notation for the special choice problem should be 'O^*_p' with the subscript identifying the distribution p. To make the proofs readable, that subscript is suppressed here.

Lemma 2: Suppose that $C(O^*) = O^*$. Then p is a global Bayes model for the choice function $C(\bullet)$.
Proof: See Appendix 1.

Lemma 3: For each admissible option $o \in C(O)$ at least one of its local Bayes models is a global Bayes model or else there is a set of probability/almost-state-independent utility pairs that represent C.
Proof: See Appendix 2.

4. An example of coherent choice using a non-convex set P reflecting "expert" opinion
In this section we illustrate how coherent choices may represent "expert" opinions while preserving independence between two events.

Example 2: Consider a decision problem among three options – three treatment plans $\{T_1, T_2, T_3\}$ defined over 4 states $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ with determinate utility outcomes given in the following table. That is, the numbers in the table are the utility outcomes for the options (rows) in the respective states (columns)

	ω_1	ω_2	ω_3	ω_4
T_1	0.00	0.00	1.00	1.00
T_2	1.00	1.00	0.00	0.00
T_3	0.99	-0.01	-0.01	0.99

Let a convex set P of probabilities be generated by two extreme points, distributions p_1 and p_2, defined by the following table. Distribution p_3 is the .50-.50 (convex) mixture of p_1 and p_2.

	ω_1	ω_2	ω_3	ω_4
p_1	0.08	0.32	0.12	0.48
p_2	0.48	0.12	0.32	0.08
p_3	0.28	0.22	0.22	0.28

Note that (for $i = 1, 2, 3$) under probability p_i, only option T_i is Bayes-admissible from the option set of $\{T_1, T_2, T_3\}$.

Without convexity – that is, using only the set comprised by the two (extreme) distributions $\{p_1, p_2\}$ – option T_3 is the sole Bayes-<u>inadmissible</u> option from among the three options $\{T_1, T_2, T_3\}$.

Now, interpret these states as the cross product of two binary partitions: a medical event A (patient allergic) and its complementary event NA (patient not-allergic), with a binary meteorological partition. S (sunny) and NS (cloudy). Specifically: $\omega_1 = A\&S$ $\omega_2 = A\&NS$ $\omega_3 = NA\&S$ $\omega_4 = NA\&NS$.

Under probability distribution p_1, the two partitions are independent events with $p_1(A) = .4$ and $p_1(S) = .2$. Likewise, under probability distribution p_2, the two partitions are independent events with $p_2(A) = .6$ and $p_2(S) = .8$. And under distribution p_3 the events A and S are positively correlated: $.56 = p_3(A \mid S) > p_3(A) = .5$, as happens with each distribution q that is a non-trivial mixture of p_1 and p_2.

Continuing with the example, we see that the three options have the following interpretations: T_1 and T_2 are ordinary medical options for how to treat the patient, with outcomes that depend solely upon the patient's allergic state. T_3 is an option that makes the allocation of medical treatment a function of the meteorological state, with a "fee" of 0.01 utile assessed for that input. That is, T_3 is the option "T_1 if cloudy and T_2 if sunny, while paying a fee of 0.01."

Suppose that p_1 represents the opinion of medical expert 1, and p_2 represents the opinion of medical expert 2. Without convexity of the credal probabilities, T_3 is inadmissible. This captures the shared agreement between the two medical experts that T_3 is unacceptable from the choice of three $\{T_1, T_2, T_3\}$, and it captures the pre-systematic understanding that under T_3 you pay to use medically irrelevant inputs about the weather in order to determine the medical treatment. However, with convexity of the set generated by p_1 and p_2, then T_3 is admissible as well. Convexity of the set of indeterminate probabilities, we note, is required in each of Gilboa and Schmeidler's (1989) version of Γ-*Maximin*, in Walley's (1990) version of Maximality, and in Levi's (1980) account of *E-admissibility*..

Aside: This illustration relies on the fact that normal and extensive form decisions are generally *not* equivalent in decision theories with indeterminate probabilities. Example 2 is in the normal form, as are all the choice problems considered in this paper. In the extensive form of this decision problem, the decision maker has the opportunity to make a terminal choice between T_1 and T_2 first, or to take as a third option a sequential alternative: pay a fee of 0.01 utiles in order to learn the state of the weather before choosing between T_1 and T_2. Under decision rules for extensive form problems that we endorse, and which we believe also are endorsed by Levi, then it is *E-<u>inadmissible</u>* to postpone the immediate medical decision between T_1 and T_2 in order to pay an amount to acquire the irrelevant meteorological evidence. And this holds whether the indeterminate probability set is convex or not. Related results about independence with indeterminate probability are presented in Cozman and Walley (2005).

5. Concluding Remarks
We have discussed *coherent choice functions* – where the admissible options in a decision problem are exactly

those which maximize expected utility for some probability/utility pair in fixed set **S** of probability/utility pairs. All of the decision problems used here to characterize and axiomatize coherent choice functions are *normal form* decision problems. But, as indicated in section 4, normal and extensive form decisions generally are not equivalent when probability (or utility) is indeterminate. One of our future projects is to study coherent choice for extensive form, i.e., sequential decision problems

Also, as noted in Lemma 3, in parallel with our findings about coherent strict partial orders (1995) the axioms are sufficient for coherence using a set of probability/almost-state-independent utility pairs. Though we give sufficient conditions when a choice function satisfying our axioms is represented by a set of probability/state-independent utility pairs with a common utility, also we intend to study how to modify the axioms to avoid the use of almost-state-independent utilities.

Acknowledgements

Appendix 1 – Lemma 2

Lemma 2: Suppose that $C(O^*) = O^*$. Then **p** is a global Bayes model for the choice function $C(\bullet)$.
Proof. Let $p = (p_1, \ldots, p_n)$ be a probability distribution on Ω with $\underline{p}$ its smallest nonzero coordinate. O^* is comprised by a set of acts that span all the elements of $\mathcal{H}$ with **p**-Expected utility $\underline{p}$.

Partition the states in Ω in two sets: $\Omega_1^p = \{\omega_1, \ldots, \omega_k\}$ those that comprise the support of **p** and, $\Omega_2^p = \{\omega_{k+1}, \ldots, \omega_n\}$ those states null under **p**. Clearly, $\Omega_2^p = \phi$ if and only if **p** has full support. We define O^* by two cases, depending whether $\Omega_2^p = \phi$ or not.

Case 1: $\Omega_2^p = \phi$ and **p** has full support. O^* is comprised by n-many acts, $\{a_j : j = 1, \ldots, n\}$ For each $j = 1, \ldots, n$, define the act a_j by

$$a_j(\omega_i) = \frac{p}{p_j}\mathbf{1} \oplus (1-\frac{p}{p_j})\mathbf{0} \quad \text{if } i = j$$
$$= \mathbf{0} \quad \text{if } i \neq j.$$

Case 2: $\Omega_2^p \neq \phi$. O^* is defined by $k(n+2-k)$-many acts which can be understood to be the product of acts defined on $\Omega_1^p \times \Omega_2^p$. With respect to Ω_1^p, O^* contains k-many acts that span horse lotteries defined on

Ω_1^p that have **p**-Expected utility $\underline{p}$, similarly to Case 1. With respect to Ω_2^p, O^* contains $(n+2-k)$-many acts that span all horse lotteries defined on Ω_2^p, including the two constants **0** and **1**.
For each $j = 1, \ldots, k$, and $m = k+1, \ldots, n+2$ define the act a_j^m by

$$a_j^m(\omega_i) = \frac{p}{p_j}\mathbf{1} \oplus (1-\frac{p}{p_j})\mathbf{0} \quad \text{if } i = j$$
$$= \mathbf{1} \text{ if } i = m \ \textbf{or } (m = n+2 \text{ and } i > k)$$
$$= \mathbf{0} \quad \text{otherwise}$$

Note that $a_j^{n+1}(\omega_i) \neq \mathbf{0}$ if and only if $i = j$. In particular, it equals **0** on Ω_2^p. And note that $a_j^{n+2}(\omega_i) \neq \mathbf{0}$ if and only if, either $i = j$ or $i > k$. It equals **1** on Ω_2^p.

Let O^* be the choice problem formed by taking the convex hull of these options:

In Case 1 $O^* = H\{a_j : j = 1, \ldots, n\}$, the convex hull of n-many options. In Case 2 $O^* = H\{a_j^m : j = 1, \ldots, k; m = k+1, \ldots, n+2\}$, the convex hull of $k(n+2-k)$-many options.

Let a_p denote the constant horse lottery that awards the identical von Neumann-Morgenstern lottery in each state, with $a_p = \underline{p}\mathbf{1} \oplus (1-\underline{p})\mathbf{0}$.

Claim 1: $a_p \in O^*$.
Proof: In Case 1, when **p** has full support, $p_1a_1 \oplus p_2a_2 \oplus \ldots \oplus p_na_n$ is the horse lottery a_p. In Case 2, when **p**-null states exist, for each $j = 1, \ldots, k$, define the horse lottery $b_j = (1-\underline{p})a_j^{n+1} \oplus \underline{p}a_j^{n+2}$ with payoffs:

$$b_j(\omega_i) = a_p \quad \text{if } i > k$$
$$b_j(\omega_i) = \frac{p}{p_j}\mathbf{1} \oplus (1-\frac{p}{p_j})\mathbf{0} \quad \text{if } i = j$$
$$b_j(\omega_i) = \mathbf{0} \quad \text{if } i \neq j \text{ and } i \leq k.$$

Then $p_1b_1 \oplus p_2b_2 \oplus \ldots \oplus p_kb_k$ is the horse lottery a_p.
◊–claim 1.

Note that (p,u) is a local Bayes model for each element of O^* as the **p**-Expected utility for each element of O^* is the same value, namely $\underline{p}$.

Claim 2: If $\underline{p} < 1$ then (p,u) is the only local Bayes model for a_p

Proof: Note that regardless the distribution q on Ω, a_p has q-Expected utility $\underline{p}$. We argue by cases that when $\underline{p}$ < 1, q is not a local model for a_p with respect to O^*.

If p has full support ($\Omega_2^p = \phi$), the q-Expected utility of

$$a_j = q_j \frac{p}{p_j} > \underline{p}. \text{ And if } j = m > k, \text{ so that } p_j = 0 \text{ and}$$

$q(\Omega_2^p) > 0$, then the q-Expected utility of

$$p_1 a_1^{n+2} \oplus p_2 a_2^{n+2} \oplus \ldots \oplus p_k a_k^{n+2}$$
$$= q(\Omega_1)\underline{p} + q(\Omega_2) > \underline{p}.$$

Hence, (q,u) is not a local Bayes model for a_p. $\diamond$-claim 2

Note also that for the case $p_1 = \underline{p} = 1$, $a_p = 1$ and then O^* = $H\{1, a_1^2, \ldots a_1^{n+2}\}$. In which case if $q \neq p$, q is not a local Bayes model for a_1^{n+1}, which has a q-expected value of $q_1 < 1$. Thus, we have
Proposition:
$\qquad p$ is the sole local Bayes model for all of O^*.

Claim 3: O^* contains all the horse lotteries in $\mathcal{H}$ with p-expected utility equal to $\underline{p}$.
Proof: Let o be a horse lottery with p-Expected utility $\underline{p}$. Write $o(\omega_j) = \alpha_j 1 \oplus (1 - \alpha_j)0$, $j = 1, \ldots, n$.
Case 1 (p has full support.): For $\omega_i \in \Omega = \Omega_1^p$ we have that $\sum_i p_i \alpha_i = \underline{p}$ and $0 \leq \alpha_i \leq 1$. The set of α-vectors satisfying these two equations is closed and convex, with extreme points given by the acts $\{a_j\}$. That is, if $\alpha^* = \langle \alpha^*_1, \ldots, \alpha^*_n \rangle$ is an extreme point of this set of α-vectors, then $\alpha^* = a_j$ for some $1 \leq j \leq n$. Since a closed, convex set is identified by its extreme points, this establishes that $o \in O^*$.
Case 2 (There are null states under p.**):** The reasoning is similar to Case 1, noting that O^* spans all horse lotteries defined over Ω_2^p. $\diamond$–Claim 3.

We complete the proof of Lemma 2, as follows. Let O be a choice set and let $\phi \neq O_p \subseteq O$ be those options for which p is a local Bayes model. So, each $a \in O_p$ maximizes the p-Expected utility of options in O at common value $\mathbf{r}$. There are two cases, depending upon whether $\mathbf{r} \geq \underline{p}$ or $\mathbf{r} < \underline{p}$.
In the former case, mix $\mathbf{0}$ into each act in O to form the

choice set O $= \dfrac{p}{r} O \oplus (1 - \dfrac{p}{r})\mathbf{0}$, with the isomorphism

between O and O that associates each $o \in O$ with $o' \in$

O, where o $= \dfrac{p}{r} o \oplus (1 - \dfrac{p}{r})\mathbf{0}$.

In case $\mathbf{r} < \underline{p}$ mix $\mathbf{1}$ into each act in O to form the choice

set O $= \dfrac{1-p}{1-r} O \oplus (\dfrac{p-r}{1-r})\mathbf{1}$, with the isomorphism

between O and O that associates each $o \in O$ with $o' \in$

O, where o $= \dfrac{1-p}{1-r} o \oplus (\dfrac{p-r}{1-r})\mathbf{1}$.

The argument continues in parallel between the two cases. By the Axiom 2, $a \in C(O)$ if and only if $a'\in C(O$`$)$. Also evident is the fact that for each $a$` $\in O$`$_p$ the $p$-Expected utility of $a$` equals $\underline{p}$. Thus, by Claim 3, for each a` $\in O$`$_p$, a`$\in C(O^*)$.

Claim 4: Let o`$\in O$` and o`$\notin O$`$_p$. Then each local Bayes model q for o` with respect to $O^* \cup \{o$`$\}$ is singular with respect to p, i.e., $\Omega_1^q \cap \Omega_1^p = \phi$.
Proof: Because o`$\notin O$`$_p$ then $\mathbf{E}_p(o$`$) < \underline{p}$ and, trivially, $p$ is not a local Bayes model for $o$`. Fix a distribution $q \neq p$ where $\Omega_1^q \cap \Omega_1^p \neq \phi$. We argue indirectly that q is not a local Bayes model for o' with respect to $O^* \cup \{o$`$\}$.
First consider the case where $\Omega_1^q \subseteq \Omega_1^p$, that is where q is absolutely continuous with respect to p. Within the n-1 dimensional simplex of distributions on Ω, let L_{pq} be the line determined by the two points p and q, having endpoints denoted q_* and q^*. Identify these endpoints by placing q in the closed line segment $[q^*, p]$, and thus p lies in the closed line segment $[q, q_*]$, from which we know that $p \neq q^*$, though it is possible that $q = q^*$.
Moreover, since $\Omega_2^q \supseteq \Omega_2^p$ we have that $p \neq q_*$, since each endpoint of L_{pq} has some null-state not shared as a null state with any other point on that line. So, p is internal to the line L_{pq}. Because q_* is an endpoint of L_{pq}, as just argued, $\Omega_2^{q*} \cap \Omega_1^p \neq \phi$. Assume that $\omega_k \in \Omega_2^{q*} \cap \Omega_1^p$. Since p lies on the line $[q_*, q^*]$, $\omega_k \in \Omega_1^{q^*}$.
Consider the act a_k^{n+1} (or the act a_k if p has full support). Since $\mathbf{E}_q(o$`$) \geq \mathbf{E}_q(a_k^{n+1})$ and $\mathbf{E}_p(o$`$) < \mathbf{E}_p(a_k^{n+1}) = \underline{p}$, there exists a unique distribution r_k situated on the line L_{pq} and between p and q (possibly with $r_k = q$), such that $E_{r_k}(o$`$) = E_{r_k}(a_k^{n+1})$. Because expected utility is linear in probability, for each distribution $t$ in the half open interval $(r_k, q_*]$, $\mathbf{E}_t(o$`$) < \mathbf{E}_t(a_k^{n+1})$. But $E_{q*}[a_k^{n+1}] = 0 > E_{q*}[o$`$]$, which is a contradiction as no act has a negative expected value.
This completes the argument when q is absolutely continuous with respect to p.
Next, assume that $\Omega_1^q \cap \Omega_1^p \neq \phi$ and write

$$q(\bullet) = q(\bullet \mid \Omega_1^p)q(\Omega_1^p) + q(\bullet \mid \Omega_2^p)q(\Omega_2^p),$$

where $q(\Omega_1^p) > 0$. So, $\mathbf{q}(\bullet|\Omega_1^p)$ is absolutely continuous with respect to p.

$\mathbf{E}_q(\bullet) = \mathbf{E}_q(\bullet \mid \Omega_1^p)q(\Omega_1^p) + \mathbf{E}_q(\bullet \mid \Omega_2^p)q(\Omega_2^p)$. Since $a_k^{n+2} \in \boldsymbol{O^*}$ and $\mathbf{E}_q(o\text{`}) \geq \mathbf{E}_q(a_k^{n+2})$, it follows that $\mathbf{E}_q(o\text{`} \mid \Omega_1^p) \geq \mathbf{E}_q(a_k^{n+2} \mid \Omega_1^p) = \mathbf{E}_q(a_k^{n+1} \mid \Omega_1^p)$. However, as $q(\bullet \mid \Omega_1^p)$ is absolutely continuous with respect to p, we have the same situation involving $q(\bullet \mid \Omega_1^p)$ and p as when q is absolutely continuous with respect to p, completing the proof. ◊-Claim 4

Next, we show that if there is a local Bayes model for $o\boldsymbol{'}$ with respect to $\boldsymbol{O^*} \cup \{o\text{`}\}$, then no element of $\boldsymbol{O^*}$ becomes inadmissible by adding option $o\text{`}$.

Claim 5: Assume that $a \in \boldsymbol{C(O^*)}$, $o\text{`} \in \boldsymbol{O`}$ but $o\text{`} \notin \boldsymbol{O`}_p$, and let $o\text{`}$ have a local Bayes model $\boldsymbol{q}$ with respect to $\boldsymbol{O^*} \cup \{o\text{`}\}$. Then $a \in \boldsymbol{C(O^* \cup \{o\text{`}\})}$.
Proof: Assume the premise. In the light of Axiom 4 we are done proving Claim 5 if we identify an act $a^* \in \boldsymbol{O^*}$ such that a^* weakly dominates $o\boldsymbol{'}$. This we do as follows.

By Claim 4, $\boldsymbol{q}$ is singular with respect to p. Consider an act a_k^{n+2} for $\omega_k \in \Omega_1^p$.
Definition: For $W \subseteq \Omega$ and act o, define the act $o|W$ by:
$$o(\omega)|W = o(\omega), \text{ for } \omega \in W,$$
and $\quad o(\omega)|W = \boldsymbol{0}, \text{ otherwise.}$

Write $o\text{`}$ as an sum of three acts $o\text{`} = o\text{`}|\Omega_1^q + o\text{`}|(\Omega_2^p \cap \Omega_2^q) + o\text{`}|\Omega_1^p$, and likewise for $a_k^{n+2} = a_k^{n+2}|\Omega_1^q + a_k^{n+2}|(\Omega_2^p \cap \Omega_2^q) + a_k^{n+2}|\Omega_1^p$. Because $a_k^{n+2}(\omega) = \boldsymbol{1}$ for $\omega \in \Omega_2^p$, then $a_k^{n+2}|\Omega_1^q$ weakly dominates $o\text{`}|\Omega_1^q$, and likewise $a_k^{n+2}|(\Omega_2^p \cap \Omega_2^q)$ weakly dominates $o\text{`}|(\Omega_2^p \cap \Omega_2^q)$. By Claim 4, $o\text{`}|\Omega_1^p$ fails to have a local Bayes model with respect to $\boldsymbol{O^*} \cup \{o\text{`}|\Omega_1^p\}$. So, by Lemma 1, there exists an option $b \in \boldsymbol{H(O^*)}$ that uniformly dominates $o\text{`}|\Omega_1^p$. Let $a^* = a_k^{n+2}|\Omega_1^q + a_k^{n+2}|(\Omega_2^p \cap \Omega_2^q) + b|\Omega_1^q$. Then a^* weakly dominates $o\text{`}$ and, as $\mathbf{E}_p[a^*] = \mathbf{E}_p[b|\Omega_1^q] = \underline{p}$, we have $a^* \in \boldsymbol{O^*}$. ◊-Claim 5

Assume that $a\text{`} \in \boldsymbol{C(O^*)}$. Let $\boldsymbol{N`} = \{o\text{`}: o\text{`} \in \boldsymbol{O`} \text{ and } o\text{`} \notin \boldsymbol{O`}_p$ but $o\text{`}$ has no local Bayes model with respect to $\boldsymbol{O^*} \cup \{o\text{`}\}\}$. Then by Lemma 1, $o\text{`} \in \boldsymbol{R(O^* \cup N`)}$. By

Axiom 1, as $a\text{`} \in \boldsymbol{C(O^*)}$ then $a\text{`} \in \boldsymbol{C}(\boldsymbol{O^* \cup N`})$. If $o\text{`} \in \boldsymbol{O`} \backslash \boldsymbol{N`}$, then using Claim 5, $a\text{`} \in \boldsymbol{C(O^* \cup N` \cup o`)}$. By a simple induction on an arbitrary well-ordering of $\boldsymbol{O`} \backslash \boldsymbol{N`}$, then $a\text{`} \in \boldsymbol{C(O^* \cup N` \cup O` \backslash N`)} = \boldsymbol{C(O^* \cup O`)}$. By Axiom 1, if $a\text{`} \in \boldsymbol{O`}$ then $a\text{`} \in \boldsymbol{C(O`)}$. Finally, by Axiom 2, $a \in \boldsymbol{C(O)}$. ◊-Lemma 2

Appendix 2 – Lemma 3

Lemma 3: For each admissible option $o \in \boldsymbol{C(O)}$ at least one of its local Bayes models is a global Bayes model or else there is a set of probability/almost-state-independent utility pairs that serve as a global Bayes-model.
Proof: The next claim, which we use to establish Lemma 3, extends the idea of Axiom 4 to the strict partial order $\langle$.
Claim 6: Suppose that for option sets $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{D}$, $\boldsymbol{B} \langle \boldsymbol{A}$ and $\boldsymbol{B} \cap \boldsymbol{C(D)} \neq \phi$. Then $\boldsymbol{A} \cap \boldsymbol{C}(closure\{\boldsymbol{D \backslash B} \cup \boldsymbol{A}\}) \neq \phi$.
Proof (indirect): Suppose that $\boldsymbol{A} \subseteq \boldsymbol{R}(closure\{\boldsymbol{D \backslash B} \cup \boldsymbol{A}\})$. By Axiom 1 applied twice, $\boldsymbol{A} \subseteq \boldsymbol{R(D \cup A)}$ and $\boldsymbol{A} \subseteq \boldsymbol{R(D \cup A \cup B)}$. Since $\boldsymbol{B} \langle \boldsymbol{A}$, likewise $\boldsymbol{B} \subseteq \boldsymbol{R(D \cup A \cup B)}$. Thus, $\boldsymbol{A} \cup \boldsymbol{B} \langle \boldsymbol{D}$. By transitivity, $\boldsymbol{B} \langle \boldsymbol{D}$ and so $\boldsymbol{B} \cap \boldsymbol{C(D)} = \phi$. ◊-Claim 6

Given $o \in \boldsymbol{C(O)}$ and following the ideas we used in (1995, Definition 19), we introduce the notion of a **target set** $\boldsymbol{T(o,O)}$ of probability distributions for o with respect to choice problem $\boldsymbol{O}$. The target set for o is a subset of the local Bayes models for o which, we show, contains all of its global Bayes models. We demonstrate that whenever the target set includes a boundary point, that boundary point is a global Bayes model.

Given a probability distribution $\boldsymbol{p}$, recall the decision problem $\boldsymbol{O_p} = \{a^p, h_1^p, ..., h_1^p\}$ defined in Section 2. We state without proof that whenever $\boldsymbol{C(O_p)} = \boldsymbol{O_p}$ then $\boldsymbol{C(O^*)} = \boldsymbol{O^*}$ for $\boldsymbol{O^*}$ defined with respect to $\boldsymbol{p}$ as in Lemma 2, and so $\boldsymbol{p}$ is a global Bayes model.

Definition: $\boldsymbol{T(o,O)} = \{\boldsymbol{p}: \boldsymbol{p}$ is local Bayes model for o in choice problem $\boldsymbol{O}$ and $\{h_1^p, ..., h_n^p\} \subseteq \boldsymbol{C(O_p)}\}$

Claim 7: $\boldsymbol{T(o,O)}$ is a non-empty, convex set.
Proof: Without loss of generality, and to simplify the presentation, we give the proof for a binary state space $\Omega = \{\omega_1, \omega_2\}$. Convexity is shown as follows. Note that for $\boldsymbol{p}$ defined by $\boldsymbol{p}(\omega_2) = 0$, $h_2^p \in \boldsymbol{C(O_p)}$, and for $\boldsymbol{p}$ defined by $\boldsymbol{p}(\omega_2) = 1$, $h_1^p \in \boldsymbol{C(O_p)}$. And by Claim 6, if $h_2^p \in \boldsymbol{C(O_p)}$, then for all distributions $\boldsymbol{q}$ with $\boldsymbol{q}(\omega_2) \leq \boldsymbol{p}(\omega_2)$ we have $h_2^q \in \boldsymbol{C(O_q)}$; and if $h_1^p \in \boldsymbol{C(O_p)}$, then for all distributions $\boldsymbol{q}$ with $\boldsymbol{q}(\omega_2) \geq \boldsymbol{p}(\omega_2)$ we have $h_1^q \in$

$C(O_q)$. In the general case, with more than 2 states, the same result follows by noting that $T(o,O)$ is an intersection of half-planes. We show that $T(o,O)$ is non-empty by an indirect argument using the Archimedean axiom. So, assume that for each p, $C\{h_1^p, h_2^p\}$ is a unit set, and by the observation above, let q be the **lub** $\{p(\omega_2): h_2^p \in C\{h_1^p, h_2^p\}\}$. There are two cases.

Case 1: $\{h_2^q\} = C\{h_1^q, h_2^q\}$ So $q(\omega_2) < 1$ and then $h_1^q \langle h_2^q$ and for all $p(\omega_2) > q(\omega_2)$, $h_2^p \langle h_1^p$. But as p approaches q, h_i^p converges to h_i^q for $i = 1, 2$. Then by Axiom 3, $h_1^q \langle h_1^q$.

Case 2: $\{h_1^q\} = C\{h_1^q, h_2^q\}$. So $q(\omega_2) > 0$ and then $h_2^q \langle h_1^q$ and for all $p(\omega_2) < q(\omega_2)$, $h_1^p \langle h_2^p$. But as p approaches q, h_i^p converges to h_i^q for $i = 1, 2$. Then by Axiom 3, $h_2^q \langle h_2^q$. ◊-Claim 7

To complete the proof of Lemma 3 there are two cases to consider.

Case 1: $T(o,O)$ contains at least one of its boundary points. Suppose, e.g., that q is the **lub** $\{p(\omega_2): h_2^p \in C\{h_1^p, h_2^p\}\}$ and that $R\{h_1^q, h_2^q\} = \phi$. Then for each $0 \leq x \leq 1$, $R\{h_1^q, h_2^q, x h_1^q \oplus (1-x) h_2^q\} = \phi$, as the following reasoning establishes.

Assume that $q(\omega_2) < 1$, or we are done. Then for all $p(\omega_2) > q(\omega_2)$, $h_2^p \langle h_1^p$ as before. For $0 < x \leq 1$, by Axiom 2, $h_2^p \langle x h_1^p \oplus (1-x) h_2^p$. As p approaches q, by Axiom 3, then $x h_1^q \oplus (1-x) h_2^q \in C\{h_1^q, h_2^q, x h_1^q \oplus (1-x) h_2^q\}$, on pain of contradiction otherwise that $h_2^q \langle h_2^q$. The reasoning is similar if the target set $T(o,O)$ is closed at the other end. Then, at each point p of closure for $T(o,O)$, $R(O_p) = \phi$ and p is global Bayes model.

Case 2: If the target set is entirely open and there is no $p \in T(o,O)$ such that $R(O_p) = \phi$, we arrive at the parallel situation studied in Section IV.2 of our (1995). That situation is one where, first, a coherent choice function C is induced by a finite set P of linearly independent probabilities on Ω. The convex target sets for C include subsets of P as extreme points, i.e., $R(O_p) = \phi$ for each $p \in P$. Hence, C is represented by the set P of global Bayes models. Then, this choice function C is changed into another C^+ which is formed by adding the strict

preferences associated with finitely many conditions of the form $T(o,O) \cap R(O_p) \neq \phi$. The results established in Section IV.2 of our (1995) show that then C^+ satisfies the axioms. Also, those results show that in a neighborhood of the extreme points of the target sets for C there are sets of probability/almost-state-independent utility pairs that are local Bayes models for C, and which then represent the choice function C^+. These almost-state-independent utilities result by adding at least one new prize $\{r\}$ to the two $\{0, 1\}$ used to create the horse lotteries studied here. ◊-Lemma 3

Corollary: If for each choice problem O and $o \in C(O)$ and the target set $T(o,O)$ includes at least one of its boundary points, then C is represented by a set of probability/state-independent utility pairs.

References

[1] Aizerman, M.A. (1985) "New Problems in General Choice Theory," *Soc Choice Welfare* **2**: 235-282.
[2] Anscombe, F.J. and Aumann, R.J. (1963) "A definition of subjective probability," *Ann. Math. Stat.* **34**: 199-205.
[3] Cozman, F.G. and Walley, P. (2005) "Graphoid properties of epistemic irrelevance and independence," *Ann. Math. and A.I.* **45**: 173-195.
[4] Gilboa, I. and Schmeidler, D. (1989) "Maxmin expected utility with non-unique prior," *J. Math.Econ.* **18**: 141-153.
[5] Kadane, J.B., Schervish, M.J., and Seidenfeld, T. (2004) "A Rubinesque theory of decision*," IMS Lecture Notes Monograph* **45**: 1-11.
[6] Kindler, J. (1983) "A General Solution Concept for Two Person, Zero Sum Games," *J. Optimization Theory and Applications* **40**: 105-119.
[7] Levi, I. (1974) "On indeterminate probabilities" *J.Phil.* **71**: 391-418.
[8] Pearce, D. (1984) "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica* **52**: 1029-1050.
[9] Savage, L.J. (1954) *The Foundations of Statistics.* Wiley, New York.
[10] Schervish, M.J., Seidenfeld, T, Kadane, J.B., and Levi, I. (2003) "Extensions of expected utility theory and some limitations of pairwise comparisons" In *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications* (J-M Bernard, T.Seidenfeld, and M.Zaffalon, eds.) 496-510. Carleton Scientific.
[11] Seidenfeld, T., Schervish, M.J., and Kadane, J.B. (1995) "A representation of partially ordered preferences," *Ann Stat.* **23**: 2168-2217.
[12] Sen, A. (1977) "Social choice theory: a re-examination," *Econometrica* 45: 53-89.
[13] Wald, A. (1950) Statistical Decision Functions. John Wiley, New York.
[14] Walley, P. (1990) *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.