

Theoretical Economics 2 (2007), 355–394

1555-7561/20070355

Coarse contingencies and ambiguity

LARRY G. EPSTEIN

Department of Economics, Boston University

MASSIMO MARINACCI

Collegio Carlo Alberto, Università di Torino

KYOUNGWON SEO

Department of Economics, University of Rochester

The paper considers an agent who must choose an action today under uncertainty about the consequence of any chosen action but without having in mind a complete list of all the contingencies that could influence outcomes. She conceives of some relevant (subjective) contingencies but she is aware that these contingencies are coarse—they leave out some details that may affect outcomes. Though she may not be able to describe these finer details, she is aware that they exist and this may affect her behavior.

KEYWORDS. Uncertainty, states of the world, ambiguity, coarse contingencies.

JEL CLASSIFICATION. D81.

1. INTRODUCTION

1.1 *Outline*

Consider an agent who must choose an action today under uncertainty about the consequence of any chosen action but without having in mind a complete list of all the contingencies that could influence outcomes. She conceives of some relevant contingencies or states of the world but she is aware that these contingencies are coarse—they leave out some details that may affect outcomes. Though she may not be able to describe these finer details, she is aware that they exist and this may affect her behavior. How does one model such an agent?

Larry G. Epstein: lepstein@bu.edu

Massimo Marinacci: massimo.marinacci@unito.it

Kyoungwon Seo: kseo@troi.cc.rochester.edu

Epstein gratefully acknowledges the financial support of the NSF (award SES-0611456) and Marinacci that of the Ministero dell'Istruzione, dell'Università e della Ricerca. We are grateful also to two referees, seminar audiences at Chicago and Olin, Fabio Maccheroni, Luigi Montrucchio, and especially Bart Lipman and Todd Sarver for helpful comments and discussions. This supersedes a paper with the title “Coarse Contingencies” by Epstein and Marinacci; the main innovation here is to provide foundations for our second model—see Section 4, particularly Theorem 2.

Copyright © 2007 Larry G. Epstein, Massimo Marinacci, and Kyoungwon Seo. Licensed under the [Creative Commons Attribution-NonCommercial License 2.5](http://creativecommons.org/licenses/by-nc/2.5/). Available at <http://econtheory.org>.

The reason for addressing this question seems clear. Outside of artificial laboratory-style settings, it would seem impossible for any decision-maker to identify all relevant and finely detailed contingencies. A sophisticated agent would be aware of her limited foresight and admit the possibility that her conceptualization of the future is incomplete, and then she would take this into account in her decision-making. Think, for example, of a portfolio choice setting. An investor may identify some factors that are likely to influence returns to financial securities. However, few sophisticated investors would be confident that they have identified all relevant factors. Awareness that her conceptualization is missing some details would presumably influence an investor's choice of portfolio.

The standard Savage approach to modeling uncertainty, using a primitive state space, does not capture the agents we have in mind. It is inappropriate for two reasons. First, each Savage-style state is a complete description of the world—it determines a unique outcome for any chosen action. Second, even if we knew how to model a “coarse or incomplete state” and we redefined the Savage state space accordingly, the resulting approach would still be unsatisfactory if, as in Savage, the state space were adopted as a primitive and thus presumed observable by the modeler. Ideally, the agent's conceptualization of the future should be taken to be subjective—it should be derived from preference, that is, from behavior that in principle is observable.

Kreps (1979, 1992) was the first to raise the modeling problem posed above. He, and also subsequent authors, refer to “unforeseen contingencies,” but it seems that they have in mind what we prefer to call “coarse contingencies.” Kreps' seminal idea was that demand for flexibility is indicative of an individual's awareness of the coarseness of her conception. In order to capture demand for flexibility, his model postulates preference over menus of alternatives (or ex post actions), where the latter are to be chosen at an unmodeled ex post stage. The subjective state space that he derives may be identified with the possible future preferences over alternatives.

Nehring (1999) and Dekel et al. (2001) provide alternative extensions. In the models of Kreps and Nehring, menus consist of alternatives from an abstract (typically finite) set. We focus primarily on the model of Dekel et al. (henceforth DLR), where menus consist of lotteries over alternatives. The richer domain permits, given suitable assumptions, the derivation of a (unique) subjective state space as part of the representation of preference. But are these states coarse? DLR describe (p. 893) the agent they view themselves as modeling: “... she sees some relevant considerations, but knows there may be others that she cannot specify. For simplicity, we assume henceforth that the agent conceives of only one situation, ‘something happens,’ but knows that her conceptualization is incomplete.” Later (pp. 919–20), they describe what is needed for a critique of their model: “... just as Ellsberg identified the role of the sure-thing principle in precluding uncertainty-averse behavior, we believe that one must first find a concrete example of behavior that is a sensible response to unforeseen contingencies but that is precluded by our axioms. An important direction for further research is to see if there is such an Ellsbergian example for this setting and, if so, to explore relaxations of our axioms.” This is the direction we pursue here.

Our central argument is that an agent who is aware of, and averse to, the coarseness of her conceptualization of the future will have an incentive to randomize, and thus will violate Independence. As an example, suppose that her subjective conceptualization is trivial—“something happens.” Suppose further that she is indifferent between committing ex ante to lottery β or to lottery β' . She is aware ex ante that there are unforeseen (finer, or back-of-the-mind) contingencies that could affect the desirability of any action. Though she does not understand these finer details and is not able to describe them, she is nevertheless aware that they exist, and she may feel that some may make β' more desirable ex post and some may make β more desirable.

As explained in the next section, randomization may hedge this uncertainty and thus the mixture might be strictly preferable to either lottery. The value of such hedging via randomization has been emphasized in the literature on preferences under ambiguity, starting with [Schmeidler's \(1989\)](#) classic uncertainty aversion axiom. This connection is not a coincidence. In fact, coarseness can be viewed as a source of ambiguity. We propose, and our model formalizes, that the impossibility of fully describing all relevant contingencies is one reason, an important one in our view, why decision makers may not be able to quantify uncertainty about future payoffs with a single probability measure.

To illustrate, consider again an investor deciding on a portfolio. Ultimately, she need be concerned only with the possible returns to securities and their likelihoods. However, in order to assess such likelihoods, she may think in terms of (economic, political, or other) factors that could influence returns, and first try to assess their likelihoods. The crucial point is that even if she is able to assign probabilities to these “physical contingencies,” a unique probability distribution for returns is *not* implied when contingencies are coarse, because then each contingency does *not* pin down a unique return. For concreteness, think of the extreme case where the investor's ‘theory’ of returns is vacuous—she conceives only that, with probability 1, something will happen tomorrow. Since this theory is consistent with *any* probability distribution over returns, ambiguity about returns is implied. Similarly, in more general choice contexts coarse contingencies induce ambiguity about ex post preferences over alternatives, and thus over the subjective state space.

We are not the first to highlight the connection between coarse perceptions and ambiguity. [Mukerji \(1997\)](#) and [Ghirardato \(2001\)](#) argue that an agent who is aware that she has only a coarse perception of the state space can be thought of as using a non-additive probability measure (or capacity). Their approach, inspired by [Dempster \(1967\)](#) and [Shafer \(1976\)](#), is much different from ours in that they take the agent's coarse perception as a primitive, while rendering it subjective is the heart of our model. (See [Section 4.3](#) for further discussion of the connection to the present paper.) In fact, for reasons outlined above, we go further and dispense also with an exogenously specified Savage-style set of states. Thus we extend the literature on preferences under ambiguity by dropping the primitive state space which is universal in that literature. In particular, we axiomatize a counterpart of [Gilboa and Schmeidler's \(1989\)](#) multiple-priors utility without relying on an exogenously specified state space.

1.2 Utility functions and interpretations

DLR assume the Independence axiom only in their most restrictive model (the additive EU representation). We focus also on their weakest Independence-style axiom, called Indifference to Randomization (IR), which they adopt (either explicitly or implicitly) in all of their representation theorems.¹ We argue that the case for IR is not clear-cut, at least given a particular conception of coarse perceptions. Thus we are led to explore two alternative directions: one which continues to assume IR and relaxes DLR's Independence axiom, and a second model which differs more substantially from DLR—not only is IR dropped, but the domain of preference is expanded to include *random menus*.² In both cases, we describe axioms for preference that arguably capture ambiguity due to coarse perceptions and that characterize functional forms for utility representing preference. The two alternative axiomatic models of preference are the main results of the paper.

In DLR's most restrictive model, the utility of any menu x has the form

$$W^{DLR}(x) = \int \max_{\beta \in x} u(\beta) d\mu(u),$$

where μ is a probability measure on ex post vNM utility functions u . (More precise statements are given below.) Given u , the choice out of x maximizes $u(\beta)$, but ex ante, the agent does not know which preference will prevail ex post. The support of μ , corresponding to the set of ex post preferences that she views as possible, constitutes her subjective state space. To evaluate x , she computes its expected payoff assuming an optimal choice of lottery in each subjective state. The representation suggests that subjective states are foreseen by the agent, and we therefore interpret the DLR model as one where states are foreseen.³

For ease of comparison, we describe informally also the two generalizations of W^{DLR} axiomatized below. In the model where IR is satisfied, utility of a menu is given by

$$W^{MP}(x) = \min_{\pi \in \Pi} \int \max_{\beta \in x} u(\beta) d\pi(u),$$

where Π is a set of probability measures on the set of ex post vNM utility functions u . (DLR is the special case where Π is a singleton.) As in the interpretation of W^{DLR} , the agent observes the realization of some u and then chooses out of the menu x by maximizing u .

¹DLR (p. 911) mention the case where ex post utilities are not vNM in the context of establishing a result regarding minimality of the subjective state space. But such violations are not described as germane to the issue of coarse contingencies.

²Nehring (1996, 1999) also adopts the domain of random menus. For more on the connection between our second model and Nehring's work see Appendix B.2.

³Kreps (1992) suggests that one can interpret the functional form alternatively as an "as if" representation for an agent who does not foresee the states, and that it is impossible to distinguish between these two interpretations by observing only choice between menus.

In our second model, where IR is violated, utility has the form

$$W(x) = \int \max_{\beta \in x} \min_{u \in U} u(\beta) d\mu(U),$$

where μ is a probability measure over suitable sets U of vNM utilities.

The main difference between the two models lies in ex post ambiguity. In the first model, W^{MP} , the agent ex ante expects that ex post, before choosing from a menu, all uncertainty will be resolved and a complete state will be realized. Therefore, there is no ex post ambiguity. However, because of the coarseness of her ex ante perception, the agent is not sure then about the likelihoods of her ex post preferences. Hence, W^{MP} has the multiple-priors functional form axiomatized by Gilboa and Schmeidler (1989), where the likelihoods of subjective states are ambiguous.

In contrast, in the second model, W , the agent ex ante expects that even ex post she will only have an incomplete perception of all relevant contingencies. As a result, she expects to receive ex post only an ambiguous signal about her ex post preferences. In the representation, this ambiguous signal is modeled by the set U . Since points within each U are foreseen, the minimization over U suggests complete ignorance within U —the agent foresees and can describe the details within U but has no idea how likely they are.⁴ From an ex ante perspective, utility W can then be interpreted as reflecting the expectation of future ambiguity, due to the realization of an ambiguous signal.⁵ Ambiguity persists ex post, and this makes ex post randomization valuable, thus violating IR.

The paper proceeds as follows. Next we outline the DLR model and argue that their axioms preclude coarse contingencies and ambiguity. Then we describe two alternative models to capture them. Proofs are relegated to appendices.

2. THE DLR MODEL

The DLR model has the following primitives.

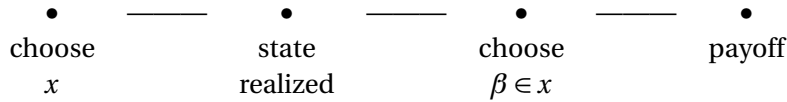
- B : finite set of actions—let $|B| = B$
- $\Delta(B)$: set of probability measures over B , endowed with the weak convergence topology; generic lotteries are $\beta, \beta', \gamma, \dots$
- \mathcal{X} : closed subsets of $\Delta(B)$, endowed with the Hausdorff metric; generic elements are denoted x, x', y, \dots and are called *menus*⁶
- preference \succeq is defined on \mathcal{X} .

⁴Minimizing over U is equivalent to minimizing over all probability measures on U .

⁵If one overlooks the fact that the state space is exogenous only in Gilboa and Schmeidler (1989), then our second model can be viewed as a special case of recursive multiple-priors, a multi-period extension of the Gilboa–Schmeidler model—see Epstein and Schneider (2003).

⁶DLR do not restrict menus to be closed but this difference from their model is unimportant and we overlook it throughout.

The agent ranks menus at time 0 (ex ante) using \succeq with the understanding that at time 1 (ex post), she will choose a lottery from the previously chosen menu (see the following time line).



One can think of a menu as corresponding to an action to be taken ex ante, where the significance of an ex ante action is that it limits options for further action ex post, that is, for the choice of β in $\Delta(B)$. There are no exogenous states of the world, but the agent may envisage some scenarios for time 1. She anticipates learning which scenario is realized before making her choice out of the menu. Thus her subjective conceptualization of the future affects her expected choices out of menus and hence also her ex ante evaluation of menus. In other words, her subjective state space underlies the preference \succeq and (under suitable assumptions) is revealed by it.

For example, the ranking

$$\{\beta, \beta'\} \succ \{\beta\} \succ \{\beta'\}$$

reveals that the agent conceives of a circumstance in which she would strictly prefer β over β' and also another circumstance in which she would strictly prefer β' over β . Under DLR's set of axioms, subjective contingencies concern only the possible ex post preference over lotteries. This is natural—payoffs rather than ex post physical states per se are ultimately all that matter.

DLR assume throughout that preference satisfies the following two axioms.

Order \succeq is complete and transitive.

Continuity For every menu x , the sets $\{y \in \mathcal{X} : y \succeq x\}$ and $\{y \in \mathcal{X} : y \preceq x\}$ are closed.

They occasionally, though not universally, adopt also the next axiom.

Monotonicity For all menus x' and x , $x' \supset x \implies x' \succeq x$.

The axiom states that flexibility has non-negative value. For concreteness, we restrict attention here to models satisfying this property.

The first problematic axiom that we consider is Independence.⁷ It refers to mixtures of two menus as defined by

$$\alpha x + (1 - \alpha)y = \{\alpha\beta + (1 - \alpha)\gamma : \beta \in x, \gamma \in y\}.$$

Formally, the indicated mixture of x and y is another menu and thus when the agent contemplates that menu ex ante, she anticipates choosing out of $\alpha x + (1 - \alpha)y$ ex post. It follows that one should think of the randomization corresponding to the α and $(1 - \alpha)$ weights as taking place at the end—after she has chosen some mixed lottery $\alpha\beta + (1 - \alpha)\gamma$ out of the menu.

⁷DLR use the term Independence to refer to a condition weaker than the one stated below. However, the two axioms are equivalent given their continuity axiom.

Independence For all menus x' , x and y and $0 < \alpha < 1$,

$$x' \succeq x \iff \alpha x' + (1 - \alpha)y \succeq \alpha x + (1 - \alpha)y.$$

In the **introduction** we asserted that the axiom is not intuitive for an agent who is aware of the coarseness of her perception of the relevant contingencies. Here we elaborate. The typical rationale for Independence (DLR, p. 905, for example) relies on two claims. The first claim is the intuitive appeal of the condition

$$x' \succeq x \implies \alpha x' \oplus (1 - \alpha)y \succeq x \oplus (1 - \alpha)y, \quad (1)$$

where $\alpha x \oplus (1 - \alpha)y$ denotes the lottery over menus that delivers x with probability α and y with probability $(1 - \alpha)$, and where the lottery is played out immediately, that is, before any uncertainty is resolved. The second claim is that the agent should be indifferent between the mixture $\alpha x + (1 - \alpha)y$ and the two-stage object $\alpha x \oplus (1 - \alpha)y$. The difference between these two ‘mixtures’ is in the timing of the randomization (or coin toss). In the latter, the coin is tossed immediately—a specific menu is realized before the agent sees a subjective state and chooses from the menu. In contrast, in $\alpha x + (1 - \alpha)y$ the randomization is completed only at the end after choice out of the mixed menu. However, we argue that this difference in timing matters when the agent is aware of the incompleteness of her conceptualization, and that it is intuitive only that

$$\alpha x + (1 - \alpha)y \succeq \alpha x \oplus (1 - \alpha)y, \quad (2)$$

thus refuting the case for the invariance asserted in Independence.

That the timing of randomization matters can be understood as follows: an agent who is aware of the coarseness of her conception of the future might behave as though she were playing a game against a malevolent nature. She suspects that, after realization of a coarse contingency and after she has chosen an action (or lottery) *ex post*, nature will complete the missing details in a way that is unfavorable for her. (She need not be able to describe the details in order to feel this way.) Then, randomization that is completed immediately, before nature acts, does nothing to impede persecution by nature. However, when randomization is completed only after nature moves, then nature is at a disadvantage. For example, if $x = \{\beta\}$ and $y = \{\gamma\}$, nature may be able to choose underlying details that lead to a low payoff for β and other (*different*) details that lead to an equally low payoff for γ , but then, in general, she cannot be as effective in sabotaging the mixture $\alpha\beta + (1 - \alpha)\gamma$. The agent can expect this to be true in particular if states in which β leads to a high (low) payoff are those in which γ has a low (high) payoff, that is, if β hedges γ . Though the agent cannot describe these fine states, it suffices that the preceding lies beneath her coarse conception and is in the ‘back of her mind.’ Then late randomization will be preferable, consistent with (2), and the latter can be understood as due to the gain from hedging.⁸ The bottom line is that, because of hedging, the agent

⁸It is well-known in the theory of preference over lotteries (see Machina 1984, for example), that the timing of randomization matters for the normative appeal of Independence if the agent must choose an auxiliary action before the risk is resolved. Independence is intuitive if the coin toss corresponding to the randomization is completed before the action choice, but not if it is completed only after the action must be chosen. The argument here about timing is similar except that it is nature, rather than the agent, who takes the auxiliary action.

might exhibit the ranking

$$\{\alpha\beta + (1 - \alpha)\gamma\} \succ \{\beta\} \sim \{\gamma\},$$

violating Independence.

DLR show that in conjunction with Order and Continuity, Independence implies the following axiom.⁹

Indifference to randomization (IR) For every menu x , $x \sim \text{co}(x)$.

To evaluate this axiom, it is important to understand precisely the meaning of the time line sketched above. It describes the agent's ex ante expectations, for example, that ex post she will be able to choose from the menu that is chosen initially. The critical issue is what information she expects to have at that point. In fact, it may very well be that the true complete (Savage-like) state will be realized before she has to choose out of the menu. But since she does not conceive of them ex ante, she cannot be thinking explicitly in terms of the complete states that might be realized ex post. Rather, given her ex ante conceptualization in terms of coarse contingencies, one natural assumption is that she expects only to know which of these is true before choosing out of the menu. In that case, she expects coarseness to persist even ex post. On the other hand, she need not foresee all the complete states in order to believe that one of them will be realized ex post. Thus an alternative assumption is that the agent anticipates that *some* complete state will be realized ex post. The intuitive appeal of IR depends on which of these assumptions is adopted.

If the agent anticipates that some complete state will be realized ex post, then she can be certain that her ex post preference over lotteries will conform to vNM. Thus she anticipates choosing out of the previously chosen menu x in order to maximize a mixture linear utility over lotteries, which means that she will do as well choosing out of x ex post as out of $\text{co}(x)$. Being certain of this ex ante, she will be indifferent between x and $\text{co}(x)$. This is the justification for Indifference to Randomization put forth by DLR.

Suppose, however, that coarseness is expected to persist ex post. Then the agent expects to be concerned ex post not only with how any given lottery β will play out, but also with how (payoff-relevant) back-of-the-mind uncertainty will be resolved eventually. Then, just as described above in the discussion of Independence, randomization may be valuable ex post. Anticipating this ex ante, she might strictly prefer $\text{co}(x)$ to x , violating IR. Both hypotheses concerning the agent's expectations seem to us to be descriptively plausible. Thus we formulate two alternative axiomatic models—one where IR is imposed and one where the axioms reflect coarse perceptions that are expected to persist.

Finally, we describe the most restrictive utility functional form characterized by DLR—the so-called non-negative additive EU representation. To express it, note that each mixture linear $u : \Delta(B) \rightarrow \mathbb{R}^1$ can be identified with a (unique) vector in $N \subset \mathbb{R}^B$, where the role of the subset N is to normalize vNM utilities so that each u corresponds

⁹ $\text{co}(x) = \{\alpha\beta + (1 - \alpha)\beta' : \beta, \beta' \in x, 0 \leq \alpha \leq 1\}$ denotes the convex hull of x . As in the case of Independence, one should think of the randomization as occurring after choice is made out of the menu.

to a unique ordering of lotteries. (DLR's specification of N is not important here; later we adopt a different specification.) The utility of any menu has the form

$$W^{DLR}(x) = \int \max_{\beta \in x} u(\beta) d\mu(u), \quad (3)$$

where μ is a probability measure on N and $u(\beta) = \sum_{b \in B} \beta(b) u_b = u \cdot \beta$.

3. MODEL 1: SHORT-RUN COARSENESS

Here we consider an agent who has coarse contingencies in mind *ex ante*, but who expects to see a complete state *ex post* before choosing out of a menu. As in DLR, preference is defined on the set \mathcal{X} of menus, and satisfies Order, Monotonicity, and Indifference to Randomization. Other axioms are relaxed or modified as we now describe.

First exclude total indifference.

Nondegeneracy $x' \succ x$ for some menus x' and x .

Our principal deviation from DLR is to relax Independence. From the argument surrounding (1)–(2) regarding the gains from hedging, we are led to the following weakening of Independence.

Preference convexity $x' \succeq x \implies \alpha x' + (1 - \alpha)x \succeq x$.

The intuition that ‘hedging’ motives may render randomization valuable recalls the intuition provided in Gilboa and Schmeidler (1989) for their relaxation of Independence designed to accommodate ambiguity aversion. Preference Convexity can in fact be understood as arising from the ambiguity aversion of an agent who is not sure about the likelihoods of her subjective states.

Suppose the agent *ex ante* foresees each possible u , an *ex post* utility function over lotteries. Then she presumably anticipates choosing out of any given menu conditionally on the realization of each u . For example, given x , she anticipates choosing the lottery β_u if u is realized. Thus the menu x is equivalent for her to the (lottery-valued) act given by $u \mapsto \beta_u$. Similarly, x' can be identified with an act $u \mapsto \beta'_u$. Then $x' \succeq x$ translates into the weak preference for the primed act over the unprimed one. If states are ambiguous for her, then, as argued by Gilboa and Schmeidler, she may strictly prefer the α -mixture of these two acts to β_u . But the mixed act is feasible for her by choosing conditionally on each u if she has the menu $\alpha x' + (1 - \alpha)x$, and thus she can do at least as well with the latter menu as with x , which ‘proves’ Preference Convexity.

We adopt two additional axioms that are admittedly “excess baggage”—they express *ex ante* certainty about the payoffs to a specific alternative b_* in B and certainty also that it will be the worst outcome (hence also lottery) *ex post*.

Fix b_* in B . Define a dominance relation on lotteries by

$$\beta \geq_D \beta' \quad \text{if} \quad \beta(b_*) \leq \beta'(b_*) \text{ and } \beta(b) \geq \beta'(b) \text{ for all } b \neq b_*.$$

Thus β may be obtained from β' by shifting probability mass from b_* to other actions. Since the agent is certain that b_* will be worst ex post (and that ex post preference will conform to vNM), she can be certain that she will prefer β to β' and hence that she would not choose β' (alone) if β is available; that is, β' has no flexibility value if β is feasible ex post. We extend this intuition by first extending the dominance relation to menus. Say that x dominates x' , written $x \geq_D x'$, if for every β' in x' there exists β in x such that $\beta \geq_D \beta'$. If $x \geq_D x'$, the agent can be certain of doing as well choosing out of x as out of $x \cup x'$. This explains the next axiom.

Worst For all menus x' and x , if $x \geq_D x'$, then $x \sim x \cup x'$.

Given our other axioms, Worst implies that, for all menus x ,

$$\Delta(B) \sim B \succeq x \succeq \{b_*\} \text{ and } B \succ \{b_*\}.$$

By IR and Monotonicity, $B \sim \Delta(B) \succeq x$. On the other hand, $x \geq_D \{b_*\}$ for any x . Thus Worst and Monotonicity imply that $x \sim x \cup \{b_*\} \succeq \{b_*\}$. We conclude that

$$B \succeq x \succeq \{b_*\}. \quad (4)$$

Finally, $B \sim \{b_*\}$ would imply total indifference, contrary to Nondegeneracy.

Though certain that b_* will be worst and $\Delta(B)$ (or B) best, the agent may nevertheless be uncertain about the cardinal payoffs to each; moreover, cardinal payoffs are important when the agent evaluates menus ex ante and must weigh payoffs across all possible contingencies. We assume that, in fact, expected cardinal payoffs to both b_* and $\Delta(B)$ are certain ex ante, where the payoff to $\Delta(B)$ refers to the payoff to the lottery chosen out of $\Delta(B)$ ex post—the chosen lottery, but not its payoff, may vary with the (back-of-the-mind) state. If the cardinal payoffs to b_* and $\Delta(B)$ are certain, then so are the payoffs to all lotteries of the form $x_p = p\Delta(B) + (1-p)b_*$, for any p in the unit interval.¹⁰ Therefore, mixing with such lotteries provides no hedging gains, which suggests that the invariance required by Independence should be satisfied for such mixtures. This explains:

Certainty independence For all menus x' and x , and $x_p = p\Delta(B) + (1-p)b_*$, and for all $0 < \alpha < 1$,

$$x' \succeq x \iff \alpha x' + (1-\alpha)x_p \succeq \alpha x + (1-\alpha)x_p.$$

Finally, we assume a mild form of continuity à la [Herstein and Milnor \(1953\)](#).

Mild continuity For all menus x , the sets $\{p \in [0, 1] : x_p \succeq x\}$ and $\{p \in [0, 1] : x_p \preceq x\}$ are closed.

¹⁰ $p\Delta(B) + (1-p)b_* = \{pq + (1-p)\delta_{b_*} : q \in \Delta(B)\}$, a menu of lotteries.

To describe the implied functional form, note that any vNM ex post preference that is not degenerate and that ranks b_* as worst can be identified with a unique vector u in N , where

$$N = \{u \in \mathbb{R}_+^B : u(b_*) = 0, \max_B u(b) = 1\}.$$

Note also that¹¹

$$\beta \succeq_D \beta' \iff u \cdot \beta \geq u \cdot \beta' \text{ for all } u \text{ in } N. \tag{5}$$

Our first model is summarized in the following result.

THEOREM 1. *Preference \succeq on \mathcal{X} satisfies Order, Monotonicity, Indifference to Randomization, Nondegeneracy, Preference Convexity, Worst, Certainty Independence, and Mild Continuity, if and only if it admits a representation by $W^{MP} : \mathcal{X} \rightarrow \mathbb{R}$ of the form*

$$W^{MP}(x) = \min_{\pi \in \Pi} \int \max_{\beta \in x} u(\beta) d\pi(u), \tag{6}$$

where Π is a convex and weak*-compact set of Borel probability measures on N .

Moreover, there exists Π that is smallest (in the sense of set inclusion) amongst sets satisfying the above. The set Π is a singleton (and so W^{MP} has the DLR form) if and only if, for all $x, x' \in \mathcal{X}$ and all $\alpha \in [0, 1]$,

$$x \sim x' \implies \alpha x + (1 - \alpha)x' \preceq x. \tag{7}$$

If preference satisfies Independence, then every menu satisfies (7), and the theorem yields (a variant of) the DLR representation result for the non-negative additive EU representation (3).

We pointed out earlier that the axioms, particularly Preference Convexity, are intuitive for an agent who foresees complete states whose likelihoods are ambiguous. Thus the theorem can be viewed as extending Gilboa and Schmeidler’s (1989) multiple-priors utility model by dispensing with an exogenous state space. At a formal level, Theorem 1 is not a trivial variation of the representation result in Gilboa and Schmeidler (1989). Our axioms deliver a (superlinear and translation invariant) preference functional defined only on the convex cone of support functions, a meager subset of the set of all continuous functions on N ; in particular, the cone has an empty interior under the sup-norm topology. For this reason we have to use techniques different from the ones used in Gilboa and Schmeidler (1989), and we exploit the notion of niveloid developed in Maccheroni et al. (2006). The smallness of the domain on which the preference functional is defined results in the non-uniqueness of the set Π of the representation (6), the domain not being big enough to pin down a single set of priors, but only a smallest one.

¹¹If there exists $b' \neq b_*$ such that $\beta(b') < \beta'(b')$, then $u \cdot \beta < u \cdot \beta'$ for u defined by $u_{b'} = 1$ and $u_b = 0$ for $b \neq b'$.

4. MODEL 2: PERSISTENT COARSENESS

Section 2 argued that IR is not intuitive if ex ante coarseness is expected to persist. Here we describe a second model designed to capture this conception of coarseness. We deviate more drastically from DLR, beginning with the adoption of an expanded domain for preference—we assume that the agent ranks not only menus but all *random menus*. In this respect, and also in the nature of our central axiom called Dominance, our approach is closer to that of Nehring (1996, 1999), as developed and generalized by Epstein and Seo (2007).

The sets B , $\Delta(B)$, and \mathcal{X} are as above. Preference \succeq is defined on $\Delta(\mathcal{X})$.¹² Generic elements of $\Delta(\mathcal{X})$ are denoted P, P', Q, \dots and are referred to as random menus. As indicated by the time line below, for any random menu P chosen ex ante, a menu x is realized next, and then, as in DLR and in our first model, the agent expects to see some subjective uncertainty resolved and finally to choose from the realized menu. A difference from our first model is that here, if the agent thinks in terms of coarse contingencies ex ante, then she expects only to see one of those coarse contingencies ex post.



4.1 Utility

Ex post payoff functions are continuous functions $v : \Delta(B) \rightarrow \mathbb{R}$ satisfying suitable normalizations. To express them, fix two actions b_* and b^* , identified with degenerate lotteries, interpreted as the worst and best lotteries ex post. This exacerbates somewhat the problem of “excess baggage” in our first model, since we now assume that not only the worst, but also the best lottery is fixed exogenously. Every vNM ex post preference over lotteries that (is not total indifference and that) ranks b_* as worst and b^* as best may be represented by some u in N^* , where

$$N^* = \{u \in [0, 1]^B : u(b_*) = 0, u(b^*) = 1\}.$$

However, the agent does not conceive of all of these possibilities ex ante—rather she thinks in terms of subsets of N^* . Denote by $\mathcal{K}(N^*)$ the set of closed subsets of N^* , endowed with the Hausdorff metric, which renders it compact metric. Say that $U \subset N^*$ is *comprehensive* if $U = \{u' \in N^* : u' \geq u \text{ for some } u \in U\}$. Denote by $\mathcal{K}^{cc}(N^*)$ the set of closed, convex, and comprehensive subsets of N^* . (Then $\mathcal{K}^{cc}(N^*)$ is also compact metric.)

The utility of any random menu P is given by

$$\mathcal{W}(P) = \int_{\mathcal{X}} \left(\int_{\mathcal{K}^{cc}(N^*)} \max_{\beta \in x} \min_{u \in U} u(\beta) d\mu(U) \right) dP(x), \tag{8}$$

¹²The set \mathcal{X} is compact metric (and so Polish) under the Hausdorff metric. It is endowed with the Borel σ -algebra. Thus $\Delta(\mathcal{X})$ is the set of Borel probability measures. It is also compact Polish under the weak-convergence topology. The set $\Delta(X)$ has similar meaning for any metric space X .

for some Borel probability measure $\mu \in \Delta(\mathcal{X}^{cc}(N^*))$. Say that μ represents the corresponding preference.

To interpret, note first that preference conforms to vNM theory when evaluating lotteries over menus—it has the expected utility form

$$\mathcal{W}(P) = \int_{\mathcal{X}} W(x) dP(x),$$

for the vNM index $W : \mathcal{X} \rightarrow \mathbb{R}$ given by

$$W(x) = \int_{\mathcal{X}^{cc}(N^*)} \max_{\beta \in x} \min_{u \in U} u(\beta) d\mu(U), \quad x \in \mathcal{X}. \quad (9)$$

Since W represents the ranking of (nonrandom) menus, it can be compared with DLR's utility function (3), to which W reduces if μ has support on singletons.¹³ More generally, the functional form (9) might describe an agent who conceives ex ante of the complete states in N^* , but does not expect to see the true state ex post. Rather, she expects only a “signal” U to be realized ex post. There is no prior ambiguity about the likelihoods of signals; however, each signal is “ambiguous”—it will inform the agent that the true subjective state u lies in U , but leave her completely ignorant otherwise.¹⁴

Note that the restriction to sets U that are convex and comprehensive is without loss of generality—for any closed set U , $\min_{u \in U} u(\beta)$ is unchanged if we replace U by its convex hull, or if we add to U points u' in N^* such that $u' \geq u$ for some $u \in U$. The normalization to convex and comprehensive sets permits us to show below that the representing measure μ is unique.

Finally, for perspective, consider the alternative functional form obtained by reversing the order of the max and min appearing inside the integral in (9). That is, consider

$$W^{rev}(x) = \int \min_{u \in U} \max_{\beta \in x} u(\beta) d\mu(U), \quad x \in \mathcal{X}.$$

In general, W^{rev} is ordinally distinct from W . This is suggested by the fact that the mini-max theorem justifying such reversals of order requires that both sets x and U be convex, but menus need not be convex. Preference represented by W^{rev} satisfies all the assumptions of [Theorem 1](#), including IR, and so it is a special case of that model. However, preference represented by W violates IR in general.

4.2 Axioms and representation result

Let \succeq be a preference order on $\Delta(\mathcal{X})$. We assume that it is complete and transitive (Order), continuous in the usual sense (Continuity), and that it is not total indifference (Nondegeneracy).

¹³More precisely, since μ is defined only on comprehensive sets, DLR is the special case where μ has support on sets of the form $\{u' \in N^* : u' \geq u\}$ for some u in N^* .

¹⁴It can be shown that this is a special case of the recursive multiple-priors utility studied by [Epstein and Schneider \(2003\)](#) (though the information structure is exogenous there and subjective here).

We assume also that preference satisfies a form of independence. To distinguish it from the Independence axiom in DLR, we give it another name.

First-stage independence For all random menus P , P' and Q and for all $0 < \alpha < 1$, $P' \succeq P \iff \alpha P' + (1 - \alpha)Q \succeq \alpha P + (1 - \alpha)Q$.

Since a mixture such as $\alpha P + (1 - \alpha)Q$ is a random menu, it follows from the time line described above that a specific menu is realized before the agent sees a subjective state and chooses from the menu. In particular, therefore, all randomization in both component measures P and Q as well as in the mixing is completed before then. As explained in our discussion of the DLR model, this difference in timing differentiates First-Stage Independence from DLR's Independence, and there are no hedging gains from immediate randomization.¹⁵

Our key axiom is adapted from [Nehring \(1996, 1999\)](#) and [Epstein and Seo \(2007\)](#). Its statement makes use of the fact that, by [Aliprantis and Border \(1994, Theorem 3.63\)](#), $\{x \in \mathcal{X} : x \subset G\}$ is open in \mathcal{X} for every open subset $G \subset \Delta(B)$. Therefore, for any menu y ,

$$\{x \in \mathcal{X} : x \cap y \neq \emptyset\} = \mathcal{X} \setminus \{x \in \mathcal{X} : x \subset \Delta(B) \setminus y\}$$

is closed, hence Borel measurable.

Let Y be a set of menus, each of which is interpreted as an upper contour set for some ex post preference order that the agent views as possible. Deferring for a moment the precise specification of Y , consider the following axiom.

Dominance If $P'(\{x \in \mathcal{X} : x \cap y \neq \emptyset\}) \geq P(\{x \in \mathcal{X} : x \cap y \neq \emptyset\})$ for all menus y in Y , then $P' \succeq P$.

Say that P' *dominates* P if the condition in the axiom is satisfied.

Consider two implications of the axiom.¹⁶ First, when $P' = \delta_{x'}$ and $P = \delta_x$ are degenerate, then $\delta_{x'}$ dominates δ_x if $x' \supset x$. Therefore, Dominance implies Monotonicity:

$$x' \supset x \implies x' \succeq x.$$

It also implies (given First-Stage Independence) [Kreps' \(1979\)](#) second key axiom (condition (1.5)): given any menus x and x' , let

$$P' = \frac{1}{2}\delta_x + \frac{1}{2}\delta_{x'} \text{ and } P = \frac{1}{2}\delta_{x \cap x'} + \frac{1}{2}\delta_{x \cup x'}.$$

Then P' dominates P , and thus Dominance implies that

$$\frac{1}{2}\delta_x + \frac{1}{2}\delta_{x'} \succeq \frac{1}{2}\delta_{x \cap x'} + \frac{1}{2}\delta_{x \cup x'}.$$

¹⁵First-Stage Independence corresponds to, and generalizes, condition (1). A difference is that since only here is preference defined on random menus, then only here is the axiom formally meaningful.

¹⁶[Nehring \(1996, 1999\)](#) makes a similar observation for his setting. Note that these implications are valid for any specification of Y , and not just for the one given below.

Deduce from First-Stage Independence that

$$\delta_{x'} \sim \delta_{x \cap x'} \implies \delta_x \succeq \delta_{x \cup x'}.$$

Since Monotonicity is also implied, we have finally that (in the obvious notation)

$$x \cap x' \sim x' \implies x \sim x \cup x',$$

which is equivalent to Kreps' axiom.¹⁷

To understand further the meaning of the axiom, think of each y in Y as an upper contour set for some conceivable ex post preference. Thus lotteries in y are “desirable” according to that ex post preference and $x \cap y \neq \emptyset$ indicates that x contains at least one desirable action, in which case we might refer to x as being desirable. Accordingly, P' dominates P if the probability of the realization of a desirable menu is larger under P' , and if this is true for every set y and hence for every conceivable definition of “desirable.” The set Y specifies which upper contour sets are relevant, or, in other words, Dominance for the given Y implies certainty that ex post upper contour sets lie in Y .

The content of the axiom depends on the specification of Y , which we now describe. Fix actions b_* and b^* , and say that β' is an *elementary improvement* of β if

$$\begin{aligned} \text{either } \beta' - \beta &= \kappa \delta_{b^*} - \kappa \delta_b \text{ for some } b \neq b^* \\ \text{or } \beta' - \beta &= -\kappa \delta_{b_*} + \kappa \delta_b \text{ for some } b \neq b_*. \end{aligned}$$

In the first case, probability mass is shifted from some outcome b to b^* and in the second mass is shifted from b_* to some other outcome. Both shifts are unambiguous improvements if b^* and b_* are the best and worst actions respectively. In particular, if that is the case, then β' dominates β in the sense of the first-order-stochastic dominance relation induced by the ranking of actions.

Assume that Y consists of all nonempty sets y satisfying the following properties.¹⁸

- Y1 y is closed and convex.
- Y2 If β' is an elementary improvement of β , and $\beta \in y$, then $\beta' \in y$.
- Y3 $[\beta^0 \notin y, \beta_q \notin y \text{ for all } q < p] \implies \alpha \beta^0 + (1 - \alpha) \beta_p \notin y \text{ for all } \alpha > 0$.

To interpret these conditions in the context of Dominance, think again of Y as consisting of ex post upper contour sets that the agent thinks possible ex ante. For Y1, y being closed reflects certainty that ex post preferences will be (upper semi) continuous. Convexity expresses the certainty that randomization will be (weakly) valuable. The expectation of indifference to randomization ex post would be captured by adding the requirement that the complement y^c (a strict lower contour set) also be convex.¹⁹ The absence of this requirement permits a role for randomization ex post.

¹⁷The latter is usually stated in the form $x \sim x' \cup x \implies x \cup x'' \sim x' \cup (x \cup x'')$.

¹⁸ β_p denotes the lottery $(b^*, p; b_*, 1 - p)$.

¹⁹Suppose that y^c is convex for every y in Y . Then $y \cap x = \emptyset \implies x \subset y^c \implies \text{co}(x) \subset y^c \implies y \cap \text{co}(x) = \emptyset$, and δ_x dominates $\delta_{\text{co}(x)}$. By Dominance, therefore, $x \succeq \text{co}(x)$. Thus they must be indifferent by Monotonicity.

Property Y2 states that any elementary improvement of a lottery in an upper contour set y is also in y . Thus it expresses certainty that b^* and b_* are the best and worst actions respectively. An immediate implication is that

$$[\beta_q \in y \text{ and } p > q] \implies \beta_p \in y. \quad (10)$$

Iterating Y2 yields

$$b^* \in y \text{ for every } y$$

and

$$b_* \in y \text{ only if } y = \Delta(B).$$

Turn finally to Y3. Let $\beta^0 \notin y$, and suppose also that $\beta_p \notin y$, which, by (10), is stronger than the stated hypothesis. We argue above that randomization can be valuable to hedge coarseness. But suppose that, after the realization of some coarse contingency, the agent is certain not only that b^* will be best ex post, but also that its payoff is constant across all back-of-the-mind details; and similarly for b_* . Thus there is conditional certainty about the *cardinal* payoffs to b^* and b_* , from which it follows that there should be conditional certainty also about the cardinal payoffs to all mixtures $\beta_p = (b^*, p; b_*, 1 - p)$. But then mixing with β_p (or with any β_q) should provide no hedging gains. In that case, the agent should not expect $\alpha\beta^0 + (1 - \alpha)\beta_p$ to be better than both component lotteries, each of which is undesirable (in the sense of not lying in y). In other words, $\alpha\beta^0 + (1 - \alpha)\beta_p \notin y$.

Finally, suppose that β_p does indeed lie in y , but only barely in the sense that $\beta_q \notin y$ for all $q < p$. Since y is closed and $\beta^0 \notin y$, it falls short by a discrete amount from being desirable according to y . Since β_p is just barely in y , mixing with β_p is not enough to compensate for the deficiency in β^0 . Therefore, again we conclude that $\alpha\beta^0 + (1 - \alpha)\beta_p \notin y$.

If the cardinal payoffs to b^* and b_* are expected to be certain conditional on each coarse contingency, then not only should randomization with any β_p be of no value ex post, as just argued in connection with Y3, but also ex ante there should be no hedging gains. But hedging gains were the only reason for arguing that the timing of randomization should matter and hence that indifference in (2) is not intuitive. This explains the final axiom.²⁰

Certainty reversal of order For every menu $x' \in \mathcal{X}$, $0 < \alpha < 1$, and $0 \leq p \leq 1$,

$$\alpha x' \oplus (1 - \alpha)\{\beta_p\} \sim \alpha x' + (1 - \alpha)\{\beta_p\}.$$

We can now state our second main result.

THEOREM 2. \succeq satisfies Order, Continuity, Nondegeneracy, First-Stage Independence, Dominance, and Certainty Reversal of Order if and only if it has a representation of the form (8). Moreover, the representing measure μ in (8) is unique.

²⁰In more accurate but less friendly notation, the axiom asserts that $\alpha\delta_{x'} + (1 - \alpha)\delta_{\{\beta_p\}} \sim \delta_{\alpha x' + (1 - \alpha)\{\beta_p\}}$.

If Certainty Reversal of Order is strengthened so as to apply to all pairs of menus, that is, to require indifference in (2) for all x' and x , then First-Stage Independence implies Independence, and all subjective states are singletons (up to being comprehensive) as in DLR—coarseness is excluded. Alternatively, the latter model is obtained if Dominance is strengthened by shrinking Y so that, in addition to Y1–Y3, every y in Y is required to have a convex complement (which leads to indifference to randomization (IR)).

Finally, with respect to uniqueness of the measure, the theorem establishes uniqueness of μ amongst all representations that use N^* as the state space. However, there is nothing special about our choice of normalizing the lowest and highest utility levels at 0 and 1 respectively. We could equally well have used any other state space for which payoffs to b_* and b^* are constant, that is, where $u_1(b_*) = u_2(b_*) < u_1(b^*) = u_2(b^*)$ for all u_1 and u_2 in the state space. Let N' and N'' be two such sets and let μ' and μ'' be the corresponding representations as delivered by **Theorem 2** (suitably extended). Then there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $N'' = \alpha N' + \beta$, and such that $U'' \in \mathcal{K}^{cc}(N'')$ if and only if $\alpha^{-1}(U'' - \beta) \in \mathcal{K}^{cc}(N')$. Therefore, the uniqueness part of the theorem implies, assuming finite supports for simplicity, that $\mu''(U'') = \mu'(\alpha^{-1}(U'' - \beta))$ for all $U'' \in \mathcal{K}^{cc}(N)$. The representing measure is unique up to such linear transformations, reflecting the choice of normalization for b_* and b^* .

4.3 Dempster and Shafer-style models

Finally, we relate the model (9) to the Dempster–Shafer-style models of **Mukerji (1997)** and **Ghirardato (2001)** mentioned in the **introduction**. They suppose that while there exists a Savage-style state space S , the agent does not conceive of all the complete states in S and has coarse perceptions. These are modeled through an auxiliary epistemic state space Ω and a correspondence Γ from Ω into S . (See the figure below.) There is a probability measure p representing beliefs on Ω .

$$\begin{array}{ccc}
 (\Omega, p) & \xrightarrow{\Gamma} & (S, \nu) \\
 & \searrow \hat{f} & \downarrow f \\
 & & X
 \end{array} \tag{11}$$

Unlike a Savage agent who would view each physical action as an act from S to the outcome set X , and who would evaluate it via its expected utility (using a probability measure on S), the present agent views each action as a (possibly multi-valued) act on Ω .

Ghirardato assumes that each \hat{f} is multi-valued, where the nonsingleton nature of $\hat{f}(\omega)$ reflects her awareness that ω is a coarse contingency. Its utility is given by

$$V^G(\hat{f}) = \int_{\Omega} \left(\min_{x \in \hat{f}(\omega)} u(x) \right) dp.$$

In this formulation, both Ω and the acts \hat{f} are taken to be objective and hence observable to the analyst. One can view our model (9) as one possible way to render them subjective: take $X = \mathbb{R}^1$, $\Omega = \text{Supp}(\mu) \subset \mathcal{K}^{cc}(N^*)$, and $p = \mu$, where μ is the measure appearing

in our representation; and identify each lottery β in $\Delta(B)$ with the multi-valued act $\widehat{\beta}$,

$$\widehat{\beta} : U \longrightarrow \{u(\beta) : u \in U\}.$$

Then

$$V^G(\widehat{\beta}) = \mathcal{W}(\{\beta\}).$$

Turn to the rest of the triangle (11). It is commutative if $\widehat{f}(\omega) = f(\Gamma(\omega))$. This is satisfied in our model if we take $S = N^*$ and $\Gamma(U) = U \subset N^*$.

Finally, we can write

$$V^G(\widehat{f}) = \int_S u(f) d\nu(s),$$

where ν is the non-additive measure or capacity²¹

$$\nu(A) = \mu(\{\omega : \Gamma(\omega) \subset A\}),$$

and the integral on the right is in the sense of Choquet (see [Schmeidler 1989](#)). Since Schmeidler's Choquet expected utility model was devised in order to accommodate ambiguity, this demonstrates once again the close connection between coarse perceptions and ambiguity.

Though there are differences in detail, similar remarks apply to [Mukerji \(1997\)](#); in particular, our model can be viewed as a way to endogenize the state spaces Ω and S , as well as the correspondence Γ , all of which are taken as primitives by Mukerji.

APPENDIX

A. APPENDIX: PROOF OF [THEOREM 1](#)

This appendix proves [Theorem 1](#). Necessity is immediate; for example, IR is satisfied because

$$\max_{\beta \in x} u \cdot \beta = \max_{\beta \in \text{co}(x)} u \cdot \beta \quad \text{for any } x.$$

The proof of sufficiency is quite long. We provide the complete argument here, including preliminary results on niveloids which are formulated for a setting more abstract than ours and which extend some results of [Maccheroni et al. \(2006\)](#). The proof of the minimality of the set Π is based on a result of [Ergin and Sarver \(2005\)](#).

A.1 Preliminaries

We begin with some preliminaries. Let C be a convex subset of a normed vector space. A function $h : C \rightarrow \mathbb{R}$ is quasiconvex if

$$h(t\beta' + (1-t)\beta'') \leq \max\{h(\beta'), h(\beta'')\} \quad \forall \beta', \beta'' \in C.$$

It is quasiconcave if $-f$ is quasiconvex, and it is quasimonotone if it is both quasiconvex and quasiconcave.

The following result is due to [Greenberg and Pierskalla \(1971, p. 1559\)](#).

²¹More precisely, it corresponds to the special case where ν is a belief function.

LEMMA A.1. *Let $h : C \rightarrow \mathbb{R}$ be continuous. Then h is quasimonotone if and only if the sets $\{\beta : h(\beta) = c\}$ are convex for all $c \in \mathbb{R}$.*

Suppose K is a compact set in some topological space, and let $h : C \rightarrow \mathbb{R}$ be given by

$$h(\beta) = \min_{y \in K} T(\beta, y), \tag{A.1}$$

where $T : C \times K \rightarrow \mathbb{R}$ is continuous on K and affine on C , i.e., for each $\beta', \beta'' \in C$,

$$T(t\beta' + (1-t)\beta'', y) = tT(\beta', y) + (1-t)T(\beta'', y) \quad \forall y \in K.$$

The function $h : C \rightarrow \mathbb{R}$ defined in (A.1) is easily seen to be concave.

Define

$$\Theta(\beta) = \operatorname{argmin}_{y \in K} T(\beta, y) \quad \forall \beta \in C.$$

Say that h is affine on some convex subset $Q \subseteq C$ if $h(\lambda\beta_1 + (1-\lambda)\beta_2) = \lambda h(\beta_1) + (1-\lambda)h(\beta_2)$ for all $\lambda \in [0, 1]$ and all $\beta_1, \beta_2 \in Q$.

LEMMA A.2. (i) *For any finite collection $\{\beta_i\}_{i \in I} \subseteq C$,*

$$h\left(\sum_{i=1}^n \lambda_i \beta_i\right) = \sum_{i=1}^n \lambda_i h(\beta_i) \tag{A.2}$$

for some collection $\{\lambda_i\}_{i \in I}$ with $\lambda_i \in (0, 1)$ and $\sum_{i \in I} \lambda_i = 1$ if and only if $\cap_{i \in I} \Theta(\beta_i) \neq \emptyset$. In this case, $\cap_{i \in I} \Theta(\beta_i) = \Theta(\sum_{i=1}^n \lambda_i \beta_i)$.

(ii) *Let $Q \subseteq C$ be a convex set. Then h is affine on Q if and only if $\cap_{\beta \in Q} \Theta(\beta) \neq \emptyset$.*

(iii) *Given $c \in \mathbb{R}$, the set $h^{-1}(c) = \{\beta \in C : h(\beta) = c\}$ is convex only if $\cap_{\beta \in h^{-1}(c)} \Theta(\beta) \neq \emptyset$.*

PROOF. In points (i) and (ii) we prove the “only if,” the converse being trivial.

(i) Let $\hat{y} \in \Theta(\sum_{i=1}^n \lambda_i \beta_i)$ and $\hat{y}_i \in \Theta(\beta_i)$ for $i \in I$. As $T(\cdot, \hat{y})$ is affine, by (A.2) we have

$$\sum_{i=1}^n \lambda_i T(\beta_i, \hat{y}) = T\left(\sum_{i=1}^n \lambda_i \beta_i, \hat{y}\right) = \sum_{i=1}^n \lambda_i T(\beta_i, \hat{y}_i). \tag{A.3}$$

On the other hand, $\hat{y}_i \in \Theta(\beta_i)$ implies

$$T(\beta_i, \hat{y}_i) \leq T(\beta_i, \hat{y}) \quad \forall i \in I.$$

Hence, by (A.3) we have

$$\sum_{i=1}^n \lambda_i T(\beta_i, \hat{y}) = \sum_{i=1}^n \lambda_i T(\beta_i, \hat{y}_i),$$

which in turn implies $T(\beta_i, \hat{y}) = T(\beta_i, \hat{y}_i)$ for each $i \in I$. This shows that $\Theta(\sum_{i=1}^n \lambda_i \beta_i) \subseteq \cap_{i \in I} \Theta(\beta_i)$. The converse inclusion is trivial, and we conclude that $\Theta(\sum_{i=1}^n \lambda_i \beta_i) = \cap_{i \in I} \Theta(\beta_i) \neq \emptyset$.

(ii) Suppose h is affine on Q . As Q is convex, this implies that $h(\sum_{i=1}^n \lambda_i \beta_i) = \sum_{i=1}^n \lambda_i h(\beta_i)$ for any finite collection $\{\beta_i\}_{i \in I} \subseteq Q$ and any $\{\lambda_i\}_{i \in I}$ with $\lambda_i \in [0, 1]$ and $\sum_{i \in I} \lambda_i = 1$. By the previous point, $\cap_{i \in I} \Theta(\beta_i) \neq \emptyset$. As all sets $\Theta(\beta)$ are compact, the Finite Intersection Property implies that $\cap_{\beta \in Q} \Theta(\beta) \neq \emptyset$.

(iii) Let $\beta_1, \beta_2 \in h^{-1}(c)$. As $h^{-1}(c)$ is convex,

$$h(\lambda\beta_1 + (1 - \lambda)\beta_2) = \lambda h(\beta_1) + (1 - \lambda)h(\beta_2) \quad \forall \lambda \in [0, 1].$$

By the previous point, $\cap_{\beta \in h^{-1}(c)} \Theta(\beta) \neq \emptyset$. □

LEMMA A.3. *Let $Q \subseteq C$ be a convex set and suppose h is continuous. Then h is affine on Q if and only if h is quasiconvex on Q and there exists $\gamma \in Q$ such that, for all $\lambda \in [0, 1]$ and all $\beta \in Q$,*

$$h(\beta) \geq h(\gamma) \quad \text{and} \quad h(\lambda\beta + (1 - \lambda)\gamma) = \lambda h(\beta) + (1 - \lambda)h(\gamma).$$

In this case, there exists $y \in K$ such that

$$h(\beta) = T(\beta, y) \quad \forall \beta \in Q. \tag{A.4}$$

PROOF. For “only if,” take $\gamma \in \operatorname{argmin}_{\beta \in \Delta(B)} h(\beta)$. Consider now the “if” part. First observe that h is quasimonotone, and so by **Lemma A.1** its level sets $h^{-1}(c) = \{\beta : h(\beta) = c\}$ are convex for all $c \in \mathbb{R}$. By **Lemma A.2(iii)**,

$$\cap_{\beta \in h^{-1}(c)} \Theta(\beta) \neq \emptyset \quad \forall c \in \mathbb{R}. \tag{A.5}$$

Set $\tilde{Q} = \{\beta \in Q : h(\beta) > h(\gamma)\}$. Let $\beta_1, \beta_2 \in \tilde{Q}$ and suppose $h(\beta_1) \geq h(\beta_2)$. If $h(\beta_1) = h(\beta_2)$, then

$$h(\lambda\beta_1 + (1 - \lambda)\beta_2) = \lambda h(\beta_1) + (1 - \lambda)h(\beta_2) \quad \forall \lambda \in [0, 1]$$

since h is quasimonotone. Suppose $h(\beta_1) > h(\beta_2)$. There exists $\lambda \in (0, 1)$ such that $h(\beta_2) = \lambda h(\beta_1) + (1 - \lambda)h(\gamma) = h(\lambda\beta_1 + (1 - \lambda)\gamma)$. Set $c = h(\beta_2)$. By (A.5), there is $\hat{y} \in \Theta(\beta_2) \cap \Theta(\lambda\beta_1 + (1 - \lambda)\gamma)$. Hence

$$\begin{aligned} h(\lambda\beta_1 + (1 - \lambda)\gamma) &= T(\lambda\beta_1 + (1 - \lambda)\gamma, \hat{y}) \\ &\geq \lambda T(\beta_1, \hat{y}) + (1 - \lambda)T(\gamma, \hat{y}) \\ &\geq \lambda h(\beta_1) + (1 - \lambda)h(\gamma) \\ &= h(\lambda\beta_1 + (1 - \lambda)\gamma), \end{aligned}$$

so that $T(\beta_1, \hat{y}) = h(\beta_1)$. Hence $\hat{y} \in \Theta(\beta_1)$, and we conclude that $\Theta(\beta_1) \cap \Theta(\beta_2) \neq \emptyset$. By **Lemma A.2(i)**, $h(\lambda\beta_1 + (1 - \lambda)\beta_2) = \lambda h(\beta_1) + (1 - \lambda)h(\beta_2)$ for all $\lambda \in [0, 1]$, and so the function h is affine on \tilde{Q} .

Let $\bar{\beta} \in Q$ be such that $h(\bar{\beta}) = h(\gamma)$. If $\beta^* \in Q$ is also such that $h(\beta^*) = h(\gamma)$, then $h(t\beta^* + (1 - t)\bar{\beta}) = t h(\beta^*) + (1 - t)h(\bar{\beta})$ since h is quasimonotone. Suppose that $\beta^* \in \tilde{Q}$. Given $\lambda \in [0, 1]$, set $\beta_\lambda = \lambda\beta^* + (1 - \lambda)\bar{\beta}$. Since h is concave and $h(\beta^*) > h(\bar{\beta})$, we have

$$h(\beta_\lambda) \geq \lambda h(\beta^*) + (1 - \lambda)h(\bar{\beta}) > h(\bar{\beta}) \quad \forall \lambda \in (0, 1]$$

and so $\beta_\lambda \in \tilde{Q}$ for each $\lambda \in (0, 1]$. By the continuity of h , we then have

$$\begin{aligned} t h(\beta^*) + (1-t)h(\bar{\beta}) &= \lim_{\lambda \rightarrow 0} t h(\beta^*) + (1-t)h(\beta_\lambda) \\ &= \lim_{\lambda \rightarrow 0} h(t\beta^* + (1-t)\beta_\lambda) \\ &= h(t\beta^* + (1-t)\bar{\beta}), \end{aligned}$$

and so we can conclude that h is affine on Q . By Lemma A.2(ii), $\cap_{\beta \in Q} \Theta(\beta) \neq \emptyset$. Any $y \in \cap_{\beta \in Q} \Theta(\beta)$ satisfies (A.4). \square

A.2 Niveloids

Let $(E, \geq, \|\cdot\|)$ be a normed Riesz space and let H be a convex cone in E containing an order unit e .²² Say that $\|\cdot\|$ is an e -norm if there exists $k > 0$ such that $|f| \leq k\|f\|e$ for all $f \in E$. Throughout we consider only e -norms.

EXAMPLE A.1. Each normed Riesz space has a natural e -norm, called the e -uniform Riesz norm, given by

$$\|f\|_e = \inf\{k \geq 0 : |f| \leq ke\}.$$

In this case, $|f| \leq \|f\|_e e$ for all $f \in E$ (see de Jonge and van Rooij 1977, p. 103). For example, if E is a function space and e is 1_Ω , then $\|\cdot\|_e$ is the supnorm. \diamond

LEMMA A.4. If $h_1, h_2 \in H$, then $h_1 + h_2 \in H$.

A functional $I : H \rightarrow \mathbb{R}$ with $I(0) = 0$ is an e -niveloid if it is monotone, $I(e) = 1$, and

$$I(h + \alpha e) = I(h) + \alpha \text{ for all } h \in H \text{ and } \alpha \geq 0.$$

If the preceding is true also for all $\alpha < 0$ such that $h + \alpha e \in H$, say that I is e -translation invariant.

LEMMA A.5. Any e -niveloid $I : H \rightarrow \mathbb{R}$ is Lipschitz continuous and e -translation invariant. Moreover, I is concave if and only if it is quasiconcave.

REMARK. Observe that Lemma A.5 applies to any e -niveloid on a convex cone.

PROOF. For translation invariance, if $\alpha < 0$, then

$$I(h) + \alpha = I(h + \alpha e - \alpha e) + \alpha = I(h + \alpha e) - \alpha + \alpha = I(h + \alpha e).$$

For the Lipschitz property, argue as follows. If $f, g \in H$, then $f - g \leq |f - g|$. As $\|\cdot\|$ is an e -norm, we have $|f - g| \leq k\|f - g\|e$, and so $f \leq g + k\|f - g\|e$. By the monotonicity of I , we have $I(f) - I(g) \leq k\|f - g\|$.

²²That is, E is a lattice under the order \geq and the norm $\|\cdot\|$ is such that, for all $f, g \in E$, $\|f\| \leq \|g\|$ whenever $|f| \leq |g|$. Recall that $|f| = f^+ + f^- = f \vee 0 + (-f) \vee 0$, and that $e \in E_+$ is an order unit if for each $f \in E$ there is $\alpha > 0$ such that $|f| \leq \alpha e$.

Finally, suppose I is quasiconcave. Let $f, g \in H$ with $I(f) \geq I(g)$ and set $\alpha = I(f) - I(g)$. Then I is concave, because, for all $t \in [0, 1]$,

$$\begin{aligned} I(tf + (1-t)g) + (1-t)\alpha &= I(tf + (1-t)(g + \alpha e)) \\ &\geq tI(f) + (1-t)I(g + \alpha e) \\ &= tI(f) + (1-t)I(g) + (1-t)\alpha. \quad \square \end{aligned}$$

Let $(E, \geq, \|\cdot\|)$ be a Banach lattice with topological dual E^* . Denote by E_+^* the set of all monotone elements in E^* and let $\Delta = \{L \in E_+^* : L(e) = 1\}$. Recall that a functional $I : H \rightarrow \mathbb{R}$ is superlinear if it is positively homogeneous and superadditive, i.e., $I(f + g) \leq I(f) + I(g)$ for all $f, g \in H$. Superlinearity and concavity are equivalent properties for positively homogeneous functionals.

THEOREM A.1. *Suppose E is a separable Banach lattice with $E = \overline{H - H}$. Let $I : H \rightarrow \mathbb{R}$ be a quasiconcave and positively homogeneous e -niveloid. Then there exists a greatest superlinear e -niveloid $\hat{I} : E \rightarrow \mathbb{R}$ that extends I , that is, $J \leq \hat{I}$ pointwise for any extension $J : E \rightarrow \mathbb{R}$ of I . In particular, there exists a smallest convex and weak*-compact set $\Gamma \subseteq \Delta$ such that*

$$I(f) = \min_{L \in \Gamma} L(f) \quad \forall f \in H.$$

The set Γ is a singleton if and only if I is additive.

PROOF. Define $I' : E \rightarrow \mathbb{R}$ by

$$I'(f) = \sup_{g \in H, \alpha \in \mathbb{R}} \{I(g) + \alpha : g + \alpha e \leq f\}.$$

Then I' is a superlinear e -niveloid that extends I : It is readily verified that $I' = I$ on H , and that I' is monotone and e -translation invariant. In addition,

- I' is superadditive: Suppose that $f_1, f_2 \in E$ and $I'(f_1 + f_2) < I'(f_1) + I'(f_2)$. Then there exist $g_1, g_2 \in H$ and $\alpha_1, \alpha_2 \in \mathbb{R}$, with $g_1 + \alpha_1 e \leq f_1$ and $g_2 + \alpha_2 e \leq f_2$, such that

$$I'(f_1 + f_2) < I(g_1) + \alpha_1 + I(g_2) + \alpha_2 \leq I(g_1 + g_2) + \alpha_1 + \alpha_2,$$

a contradiction because $g_1 + g_2 \in H$ and $g_1 + g_2 + (\alpha_1 + \alpha_2)e \leq f_1 + f_2$.

- I' is positively homogeneous: Given any $\beta \geq 0$ and $f \in E$, we have

$$\begin{aligned} I'(\beta f) &= \sup_{g \in H, \alpha \in \mathbb{R}} \{I(g) + \alpha : g + \alpha e \leq \beta f\} \\ &= \sup_{g \in H, \alpha \in \mathbb{R}} \left\{ I\left(\frac{g}{\beta}\right)\beta + \frac{\alpha}{\beta}\beta : \frac{g}{\beta} + \frac{\alpha}{\beta}e \leq f \right\} \\ &= \beta \sup_{g \in H, \alpha \in \mathbb{R}} \left\{ I\left(\frac{g}{\beta}\right) + \frac{\alpha}{\beta} : \frac{g}{\beta} + \frac{\alpha}{\beta}e \leq f \right\} \\ &= \beta \sup_{h \in H, \gamma \in \mathbb{R}} \{I(h) + \gamma : h + \gamma e \leq f\} = \beta I'(f). \end{aligned}$$

Let $\tilde{I} : E \rightarrow \mathbb{R}$ be any superlinear e -niveloid that extends I . The superdifferential $\partial \tilde{I}(f)$ at $f \in E$ is given by

$$\partial \tilde{I}(f) = \{L \in E^* : \tilde{I}(g) \leq \tilde{I}(f) + L(g - f) \forall g \in E\}. \tag{A.6}$$

Since \tilde{I} is concave and Lipschitz continuous (by Lemma A.5), $\partial \tilde{I}(f)$ is nonempty, convex and weak*-compact for each $f \in E$ (see Phelps 1993, Proposition 1.11).

We now show that $\tilde{I}(f) = \min_{L \in \partial \tilde{I}(0)} L(f)$: Let $L \in \partial \tilde{I}(f)$. If we take $g = 0$ in (A.6), we get $\tilde{I}(f) \geq L(f)$, while if we take $g = 2f$, then we get $\tilde{I}(f) \leq L(f)$. We conclude that $\tilde{I}(f) = L(f)$. This implies that $\partial \tilde{I}(f) = \{L \in \partial \tilde{I}(0) : L(f) = \tilde{I}(f)\}$ and $\tilde{I}(f) = \min_{L \in \partial \tilde{I}(0)} L(f)$.

Next we show that $\partial \tilde{I}(0) \subseteq \Delta$. Fix $f \in E$ and $L \in \partial \tilde{I}(0)$. Then $\tilde{I}(f + \alpha e) \leq L(f + \alpha e)$ for all $\alpha \in \mathbb{R}$, and so $\tilde{I}(f) \leq L(f) + \alpha(L(e) - \tilde{I}(e))$ for all $\alpha \in \mathbb{R}$. This contradicts $\tilde{I}(f) > -\infty$ unless $L(e) = \tilde{I}(e) = 1$. It remains to prove that $\partial \tilde{I}(0) \subseteq E_+^*$: $L \in \partial \tilde{I}(0) \implies L(f) \geq \tilde{I}(f) \geq 0$ for every $f \in E_+$ by the monotonicity of \tilde{I} . Therefore $\partial \tilde{I}(0) \subseteq \Delta$.

Define $I^* : E^* \rightarrow \mathbb{R}$ by $I^*(L) = \inf_{h \in H} \{L(h) - I(h)\}$. Then $I^*(L) = 0$ for $L \in \partial \tilde{I}(0)$. Therefore, for each $f \in E$,

$$\tilde{I}(f) = \min_{L \in \partial \tilde{I}(0)} L(f) = \min_{L \in \partial \tilde{I}(0)} \{L(f) - I^*(L)\}.$$

By Ergin and Sarver (2005, Propn. 2), there is a convex weak*-compact set Γ such that²³

$$I(h) = \min_{L \in \Gamma} \{L(h) - I^*(L)\} \quad \forall h \in H,$$

and such that Γ is smallest in this respect, that is, $\Gamma \subseteq \Phi$ for any convex weak*-compact set Φ satisfying

$$I(h) = \min_{L \in \Phi} \{L(h) - I^*(L)\} \quad \forall h \in H.$$

Since

$$I(h) = \tilde{I}(h) = \min_{L \in \partial \tilde{I}(0)} L(h) = \min_{L \in \partial \tilde{I}(0)} \{L(h) - I^*(L)\} \quad \forall h \in H,$$

it follows that $\Gamma \subseteq \partial \tilde{I}(0)$. Hence the functional $\hat{I} : E \rightarrow \mathbb{R}$ given by

$$\hat{I}(f) = \min_{L \in \Gamma} L(f) \quad \forall f \in E$$

is the greatest superlinear e -niveloid that extends I on E , that is, $\hat{I} \geq \tilde{I}$ for any superlinear e -niveloid \tilde{I} extending I .

As a result, Γ is the desired smallest set. In fact, let $\Gamma' \subseteq \Delta$ be any other convex and weak*-compact set such that $I(f) = \min_{L \in \Gamma'} L(f)$ for all $f \in H$. Define $\tilde{I} : E \rightarrow \mathbb{R}$ by $\tilde{I}(f) = \min_{L \in \Gamma'} L(f)$. The functional \tilde{I} is a superlinear e -niveloid that extends I , and so $\hat{I} \geq \tilde{I}$, which in turn implies $\Gamma \subseteq \Gamma'$.

Finally, suppose I is additive. By Lemma A.5, I is Lipschitz and so by the Hahn-Banach Theorem it admits a unique linear extension $J : E \rightarrow \mathbb{R}$. Hence $J \leq \hat{I}$ and

$$J(f) = -J(-f) \geq -\hat{I}(-f) \geq \hat{I}(f) \geq J(f) \quad \forall f \in E,$$

which implies $J = \hat{I}$ and $\Gamma = \{J\}$. □

²³Ergin and Sarver's result shows the existence of a smallest weak*-compact set. Its closed convex hull is then the smallest convex and weak*-compact set.

Given a convex subset G containing both 0 and e , let $\langle G \rangle$ be the vector subspace it generates.

COROLLARY A.1. *Let G be a convex subset containing both 0 and e of a separable Banach lattice E , with $E = \overline{\langle G \rangle}$. Let $W : G \rightarrow \mathbb{R}$ be quasiconcave and monotone, such that $W(0) = 0$ and*

$$W(tf + (1-t)\gamma e) = tW(f) + (1-t)\gamma \quad \forall f \in G, \forall t, \gamma \in [0, 1].$$

Then there exists a smallest convex and weak-compact set $\Gamma \subseteq \Delta$ such that*

$$W(f) = \min_{L \in \Gamma} L(f) \quad \forall f \in G. \tag{A.7}$$

The set Γ is a singleton if and only if W is affine.

PROOF. Let $H = \bigcup_{\alpha > 0} \alpha G$ be the cone generated by G . For each $h \in H$, there exists $\alpha > 0$ such that $h/\alpha \in G$. Define $I : H \rightarrow \mathbb{R}$ by $I(h) = \alpha W(h/\alpha)$.

The functional I is well defined: suppose that, for a given $h \in H$, there exist $\alpha, \beta > 0$ such that $h/\alpha, h/\beta \in G$. Without loss of generality suppose $\beta \leq \alpha$. Then

$$W\left(\frac{h}{\alpha}\right) = W\left(\frac{\beta}{\alpha} \frac{h}{\beta}\right) = W\left(\frac{\beta}{\alpha} \frac{h}{\beta} + \left(1 - \frac{\beta}{\alpha}\right) 0\right) = \frac{\beta}{\alpha} W\left(\frac{h}{\beta}\right),$$

as desired. Observe that $h/\beta \in G$ and $\beta \leq \alpha$ imply $h/\alpha \in G$. Hence, given any $h_1, h_2 \in H$, there exists $\alpha \geq 1$ such that $h_i/\alpha \in G$ and $I(h_i) = \alpha W(h_i/\alpha)$ for each $i = 1, 2$. This property will be tacitly used in the sequel.

The functional I is clearly positively homogeneous and is an e -niveloid. We now show that it is quasiconcave on H . Given any $h_1, h_2 \in H$, there exists $\alpha > 0$ such that $h_1/\alpha, h_2/\alpha \in G$. For any $t \in [0, 1]$,

$$\begin{aligned} I(th_1 + (1-t)h_2) &= I\left(t\alpha \frac{h_1}{\alpha} + (1-t)\alpha \frac{h_2}{\alpha}\right) \\ &= \alpha I\left(t \frac{h_1}{\alpha} + (1-t) \frac{h_2}{\alpha}\right) \\ &= \alpha W\left(t \frac{h_1}{\alpha} + (1-t) \frac{h_2}{\alpha}\right) \\ &\geq \alpha \min\left\{W\left(\frac{h_1}{\alpha}\right), W\left(\frac{h_2}{\alpha}\right)\right\} \\ &= \min\{I(h_1), I(h_2)\}. \end{aligned}$$

Thus I is quasiconcave. By **Theorem A.1**, there is a smallest convex and weak*-compact set $\Gamma \subseteq \Delta$ such that

$$I(f) = \min_{L \in \Gamma} L(f) \quad \forall f \in H.$$

In particular, Γ is a singleton if and only if I is additive. Hence, Γ is such that (A.7) holds, and it is the smallest such set: Suppose $\Gamma' \subseteq \Delta$ is another set for which (A.7) holds. Define

$\tilde{I} : E \rightarrow \mathbb{R}$ by $\tilde{I}(f) = \min_{L \in \Gamma} L(f)$. Then $\tilde{I}(f) = I(f)$ for all $f \in H$, and by **Theorem A.1** we then have

$$\min_{L \in \Gamma} L(f) \geq \min_{L \in \Gamma'} L(f) \quad \forall f \in E,$$

which implies $\Gamma \subseteq \Gamma'$. Since I is additive if and only if W is affine, this completes the proof. \square

A.3 Application

Let $\Delta(N)$ be the set of all Borel probability measures on N and $\mathcal{C}(N)$ be the set of all continuous functions on N . Endow $\mathcal{C}(N)$ with the supnorm $\|\cdot\|_s$ and define an order \geq on $\mathcal{C}(N)$ by $f \geq g$ if $f(u) \geq g(u)$ for all $u \in N$.

The triple $(\mathcal{C}(N), \geq, \|\cdot\|_s)$ forms a separable Banach lattice; its dual is given by the set of all bounded Borel measures and the set Δ of **Theorem A.1** is given by $\Delta(N)$. Moreover, 1_N is an order unit e that makes $\|\cdot\|_s$ an e -norm, while the 0 is the function on N that is identically zero.

Denote by Σ the set of all support functions $\sigma_x : N \rightarrow \mathbb{R}$, given by $\sigma_x(u) = \max_{\beta \in x} \beta \cdot u$ for each $u \in N$ and $x \in \mathcal{X}$. Then Σ is a convex subset of $\mathcal{C}(N)$ containing both e and 0 . For the latter, note that $\sigma_{\Delta(B)} = 1_N$ and $\sigma_{\{b_*\}} = 0$.

For this setting we have the following version of a classic result of **Hörmander (1955)**.

LEMMA A.6. $\mathcal{C}(N)$ is the supnorm closure of $\langle \Sigma \rangle$.

PROOF. Since $0 \in \Sigma$, we have $\langle \Sigma \rangle = \cup_{\alpha > 0} \alpha \Sigma - \cup_{\alpha > 0} \alpha \Sigma$. By proceeding as in **Hörmander (1955, p. 185)**, we can show that $\langle \Sigma \rangle$ is a Riesz subspace of $\mathcal{C}(N)$. Moreover, given $u, u' \in N$ with $u \neq u'$ there is $\beta \in \Delta(N)$ such that $\beta \cdot u \neq \beta \cdot u'$. Hence, $\sigma_{\{\beta\}}(u) \neq \sigma_{\{\beta\}}(u')$ and so $\langle \Sigma \rangle$ separates the points of N . Since $\langle \Sigma \rangle$ also contains 1_N , application of the lattice version of the Stone–Weierstrass Theorem completes the proof. \square

The next result is an immediate consequence of **Corollary A.1** and **Lemmas A.3, A.5, and A.6**. Observe that a weak*-compact subset of $\Delta(N)$ is weakly compact.

COROLLARY A.2. Let $W : \Sigma \rightarrow \mathbb{R}$ be monotone and quasiconcave and satisfy the conditions $W(0) = 0$, $W(e) = 1$, and

$$W(\alpha \sigma_x + (1 - \alpha) \gamma e) = \alpha W(\sigma_x) + (1 - \alpha) \gamma \quad \forall \sigma_x \in \Sigma \forall \alpha, \gamma \in [0, 1].$$

There exists a smallest convex and weakly compact subset $\Pi \subseteq \Delta(N)$ such that

$$W(\sigma_x) = \min_{\pi \in \Pi} \int_N \sigma_x(u) d\pi \quad \forall \sigma_x \in \Sigma.$$

The set Π is a singleton if and only if W is quasimonotone.

We turn finally to the proof of **Theorem 1**. Adopt the hypotheses stated there.

LEMMA A.7. *There exists $\mathcal{W} : \mathcal{X} \rightarrow \mathbb{R}$ that represents \succeq and such that, for each $x, x' \in \mathcal{X}$, $\alpha \in [0, 1]$, and $x_p \in C$,*

$$\begin{aligned} \mathcal{W}(\alpha x + (1 - \alpha)x_p) &= \alpha \mathcal{W}(x) + (1 - \alpha)p \\ \mathcal{W}(\alpha x + (1 - \alpha)x') &\geq \min\{\mathcal{W}(x), \mathcal{W}(x')\}. \end{aligned}$$

The functional \mathcal{W} is unique up to positive affine transformations.

PROOF. The set $C = \{x_p \equiv p\Delta(B) + (1 - p)b_* : p \in [0, 1]\}$ is a convex subset of the vector space $\{\lambda\Delta(B) + \nu b_* : \lambda, \nu \in \mathbb{R}\}$. In fact, for all $x_p, x_q \in C$ and all $\alpha \in [0, 1]$,

$$\begin{aligned} \sigma_{\alpha x_p + (1 - \alpha)x_q} &= \alpha \sigma_{x_p} + (1 - \alpha) \sigma_{x_q} = \alpha \sigma_{p\Delta(B) + (1 - p)b_*} + (1 - \alpha) \sigma_{q\Delta(B) + (1 - q)b_*} \\ &= \alpha(p\sigma_{\Delta(B)} + (1 - p)\sigma_{b_*}) + (1 - \alpha)(q\sigma_{\Delta(B)} + (1 - q)\sigma_{b_*}) \\ &= (\alpha p + (1 - \alpha)q)\sigma_{\Delta(B)} + (1 - \alpha p - (1 - \alpha)q)\sigma_{b_*} \\ &= \sigma_{(\alpha p + (1 - \alpha)q)\Delta(B) + (1 - \alpha p - (1 - \alpha)q)b_*} = \sigma_{x_{\alpha p + (1 - \alpha)q}} \end{aligned}$$

and so $\alpha x_p + (1 - \alpha)x_q = x_{\alpha p + (1 - \alpha)q}$.

Because \succeq satisfies the vNM axioms on C (for example, Mild Continuity implies the Archimedean axiom on C), there exists an affine function $u : C \rightarrow \mathbb{R}$, unique up to positive affine transformations, such that $x_p \succeq x_{p'}$ if and only if $u(x_p) \geq u(x_{p'})$. Normalize u so that $u(b_*) = 0$ and $u(\Delta(B)) = 1$. Hence,

$$u(x_p) = u(p\Delta(B) + (1 - p)b_*) = pu(\Delta(B)) + (1 - p)u(b_*) = p. \tag{A.8}$$

Any $x \in \mathcal{X}$ satisfies (4), that is, $\Delta(B) \succeq x \succeq b_*$. By Mild Continuity there exists a $p \in [0, 1]$ such that $x_p \sim x$. Such p is unique since, by (A.8), $x_p \sim x_q$ if and only if $p = q$.

Set $\mathcal{W}(x) = u(x_p) = p$. Clearly, $x \succeq x'$ if and only if $\mathcal{W}(x) \geq \mathcal{W}(x')$, and \mathcal{W} is the unique functional on \mathcal{X} representing \succeq that reduces to u on C .

Consider $x \in \mathcal{X}$ and $x_p \in C$. There exists $x_q \in C$ such that $x \sim x_q$. By Certainty Independence,

$$x \sim x_q \iff \alpha x + (1 - \alpha)x_p \sim \alpha x_q + (1 - \alpha)x_p \quad \forall \alpha \in [0, 1],$$

and so

$$\begin{aligned} \mathcal{W}(\alpha x + (1 - \alpha)x_p) &= \mathcal{W}(\alpha x_q + (1 - \alpha)x_p) \\ &= \alpha \mathcal{W}(x_q) + (1 - \alpha)\mathcal{W}(x_p) \\ &= \alpha \mathcal{W}(x) + (1 - \alpha)\mathcal{W}(x_p). \end{aligned}$$

Finally, quasiconcavity of \mathcal{W} is a direct consequence of Preference Convexity. □

For any menu x , define its \geq_D -hull by

$$\text{hull}(x) = \{\beta' \in \Delta(B) : \beta \geq_D \beta' \text{ for some } \beta \in x\}.$$

If x is convex, then so is $\text{hull}(x)$.

LEMMA A.8. (i) For any convex x and $\beta^0 \in \Delta(B) \setminus \text{hull}(x)$, there exists u in N such that

$$\sigma_x(u) < u(\beta^0). \quad (\text{A.9})$$

(ii) For any convex x and y ,

$$\sigma_y = \sigma_x \implies y \sim x$$

$$\sigma_y \geq \sigma_x \implies y \succeq x$$

PROOF. (i) Define \succeq_D^* on \mathbb{R}^B by $\beta \succeq_D^* \beta' \iff u \cdot \beta \geq u \cdot \beta'$ for all u in N . By (5), \succeq_D^* agrees with \succeq_D on $\Delta(B)$; hence we denote both orders by \succeq_D . Let

$$H = \{\beta \in \mathbb{R}^B : \sum_b \beta_b = 1\}$$

and

$$\text{hull}_H(x) = \{\beta' \in H : \beta \succeq_D \beta' \text{ for some } \beta \in x\}.$$

Then $\beta^0 \notin \text{hull}_H(x)$, a closed convex set. Therefore, there exists a v in \mathbb{R}^B such that

$$\sup_{\beta \in \text{hull}_H(x)} v \cdot \beta < v \cdot \beta^0. \quad (\text{A.10})$$

Since $b_* \in \text{hull}_H(x)$, it follows that $v \cdot b_* < v \cdot \beta^0$. This implies that v is not a constant.

Without loss of generality let $v_{b_*} = 0$, so that $v \cdot \beta^0 > 0$. It remains only to show that

$$v_b \geq 0 \text{ for all } b \neq b_*. \quad (\text{A.11})$$

In that case, we can renormalize v to obtain $u \in N$ satisfying (A.9).

To prove (A.11), given any $b \neq b_*$ and $\kappa < 0$, define

$$\gamma_\kappa \equiv \kappa \delta_b + (1 - \kappa) \delta_{b_*}.$$

If $\kappa < 0$, then $\beta \succeq_D \gamma_\kappa$ for all $\beta \in \Delta(B)$, and hence $\gamma_\kappa \in \text{hull}_H(x)$. By (A.10),

$$v_b = \frac{1}{\kappa}(v_b \kappa) = \frac{1}{\kappa}(v \cdot \beta_\kappa) > \frac{1}{\kappa}(v \cdot \beta^0).$$

Since this holds for every $\kappa < 0$, $v_b \geq 0$.

(ii) We have

$$x \succeq_D \text{hull}(x) \text{ and } y \succeq_D \text{hull}(y).$$

By Worst,

$$x \sim \text{hull}(x) \text{ and } y \sim \text{hull}(y). \quad (\text{A.12})$$

Thus it suffices to show that

$$\sigma_x = \sigma_y \implies \text{hull}(x) = \text{hull}(y).$$

Suppose that $\beta^0 \in \text{hull}(x) \setminus \text{hull}(y)$. Then by (i) there exists u in N such that

$$\sigma_y(u) < \beta^0 \cdot u \leq \sigma_x(u), \quad (\text{A.13})$$

a contradiction. We conclude that $\text{hull}(x) = \text{hull}(y)$.

Finally, let $\sigma_y \geq \sigma_x$ on N . By (A.12), it is enough to show that $\text{hull}(x) \subseteq \text{hull}(y)$. Otherwise, there exists $\beta^0 \in \text{hull}(x) \setminus \text{hull}(y)$, which implies (A.13), contradicting our hypothesis. \square

Define $W : \Sigma \rightarrow \mathbb{R}$ by

$$W(\sigma_x) = \mathcal{W}(x) \text{ for each } x \in \mathcal{X}.$$

Then W is well-defined because, by Lemmas A.7–A.8 and IR,

$$\begin{aligned} \sigma_x = \sigma_{x'} &\implies \sigma_{\text{co}(x)} = \sigma_{\text{co}(x')} \implies \text{co}(x) \sim \text{co}(x') \implies \\ \mathcal{W}(\text{co}(x)) &= \mathcal{W}(\text{co}(x')) \implies \mathcal{W}(x) = \mathcal{W}(x'). \end{aligned}$$

Also by Lemma A.8, W is monotone, with $W(0) = 0$ and $W(e) = 1$. Moreover,

$$\begin{aligned} W(\alpha\sigma_x + (1-\alpha)\gamma\sigma_B) &= W(\sigma_{\alpha x + (1-\alpha)x_\gamma}) = \mathcal{W}(\alpha x + (1-\alpha)x_\gamma) \\ &= \alpha\mathcal{W}(x) + (1-\alpha)\mathcal{W}(x_\gamma) \\ &= \alpha W(\sigma_x) + (1-\alpha)\gamma W(\sigma_B) \end{aligned}$$

and

$$\begin{aligned} W(\alpha\sigma_x + (1-\alpha)\sigma_{x'}) &= W(\sigma_{\alpha x + (1-\alpha)x'}) = \mathcal{W}(\alpha x + (1-\alpha)x') \\ &\geq \min\{\mathcal{W}(x), \mathcal{W}(x')\} = \min\{W(x), W(x')\}. \end{aligned}$$

Hence W satisfies the hypotheses of Corollary A.2, and so there exists a smallest convex and weakly compact subset $\Pi \subseteq \Delta(N)$ such that, for each $\sigma \in \Sigma$,

$$\mathcal{W}(x) = W(\sigma_x) = \min_{\pi \in \Pi} \int_N \sigma_x(u) d\pi = \min_{\pi \in \Pi} \int_N \max_{\beta \in x} u(\beta) d\pi.$$

Finally, if (7) holds, then Preference Convexity implies

$$x \sim x' \implies \alpha x + (1-\alpha)x' \sim x \quad \forall x, x' \in \mathcal{X} \forall \alpha \in [0, 1],$$

and so W is quasimonotone by Lemma A.1. By Corollary A.2, Π is a singleton.

B. APPENDIX: PROOF OF THEOREM 2

B.1 Upper contour sets

Refer to the menu y as an *upper contour set* if

$$y = \{\beta \in \Delta(B) : \min_{u \in U} u \cdot \beta \geq s\} \tag{B.1}$$

for some $0 \leq s \leq 1$ and some closed, convex and comprehensive $U \subset N^*$.

PROPOSITION B.1. *y is an upper contour set if and only if it satisfies Y1–Y3. Moreover, the corresponding pair (U, s) is unique.*

Thus

$$Y = \{y \in \mathcal{X} : y = \{\beta \in \Delta(B) : \min_{u \in U} u \cdot \beta \geq s\}, 0 \leq s \leq 1, U \in \mathcal{K}^{cc}(N^*)\}. \tag{B.2}$$

Since $\mathcal{K}^{cc}(N^*)$ is compact, it is straightforward to show that Y is compact (and hence also measurable).²⁴

We now turn to the proof of the proposition.

²⁴In Lemma B.3, we prove the continuity of $U \times s \mapsto \{\beta \in \Delta(B) : \min_{u \in U} u \cdot \beta \geq s\}$.

LEMMA B.1. *Let y satisfy Y1–Y3. If $\beta^0 \in \Delta(B) \setminus y$, then there exists $u \in N^*$ such that*

$$u \cdot \beta^0 < s \leq u \cdot \beta \text{ for every } \beta \in y, \quad (\text{B.3})$$

where

$$s = \min\{p \in [0, 1] : \beta_p \in y\}. \quad (\text{B.4})$$

PROOF. Define the partial order \geq_D on $\Delta(B)$ as the smallest transitive relation satisfying the condition that $\beta' \geq_D \beta$ if β' is an elementary improvement of β . Note that, for all lotteries β ,

$$b^* \geq_D \beta \geq b_* \text{ for all } \beta \in \Delta(B), \text{ and} \quad (\text{B.5})$$

$$[\beta' \geq_D \beta, \beta \in y] \implies \beta' \in y. \quad (\text{B.6})$$

As a result, $\{p \in [0, 1] : \beta_p \in y\}$ is nonempty because it contains $\beta_1 = b^*$, and is closed by Y1. Therefore, s is well-defined and $\beta_s \in y$. Moreover, by Y3,

$$\text{co}(\{\beta^0, \beta_s\}) \cap y = \{\beta_s\}, \quad (\text{B.7})$$

where $\text{co}(\{\beta^0, \beta_s\}) = \{\alpha\beta^0 + (1 - \alpha)\beta_s : \alpha \in [0, 1]\}$.

The definition of \geq_D admits the obvious extension to all of \mathbb{R}^B . Let

$$H = \{\beta \in \mathbb{R}^B : \sum_b \beta_b = 1\}$$

and

$$\text{Phull}_H(y) = \{\beta \in H : \beta \neq \beta_s, \beta \geq_D \beta' \text{ for some } \beta' \in y\}.$$

It follows from Y1 that $\text{Phull}_H(y)$ is convex. Define also

$$D^0 = \{\beta \in H : \beta' \geq_D \beta \text{ for some } \beta' \in \text{co}(\{\beta^0, \beta_s\})\}.$$

This set is also closed and convex, and it does not intersect $\text{Phull}_H(y)$: if $\beta \in D^0 \cap \text{Phull}_H(y)$, then there exists $\alpha > 0$ and β' in y such that

$$\alpha\beta^0 + (1 - \alpha)\beta_s \geq_D \beta \geq_D \beta'.$$

Since \geq_D is transitive, $\alpha\beta^0 + (1 - \alpha)\beta_s \geq_D \beta'$ and $\alpha\beta^0 + (1 - \alpha)\beta_s \in y$ by (B.6). But this contradicts (B.7).

Therefore, there exists a separating hyperplane v in \mathbb{R}^B such that

$$v \cdot \beta^0 < v \cdot \beta_s \leq v \cdot \beta \text{ for all } \beta \in \text{Phull}_H(y), \text{ and} \quad (\text{B.8})$$

$$v \cdot \beta \leq v \cdot \beta_s \text{ for all } \beta \in D^0. \quad (\text{B.9})$$

Since $b^* \in \text{Phull}_H(y)$, by (B.5), it follows that

$$v \cdot \beta^0 < v \cdot b^* = v_{b^*}.$$

Claim 1. $v_b \leq v_{b^*}$ for all b : Let $v_b > v_{b^*}$ for some b , and consider $\beta' \in H$ of the form

$$\beta'_b = -\kappa, \beta'_{b^*} = 1 + \kappa, \text{ and } \beta'_{b'} = 0 \text{ for } b' \neq b, b^*, \text{ where } \kappa > 0.$$

Then for any lottery β , $\beta' \geq_D b^* \geq_D \beta$. Hence $\beta' \geq_D \beta$ and $\beta' \in \text{Phull}_H(y)$. But $v \cdot \beta' = v_{b^*} + (v_{b^*} - v_b)\kappa < v \cdot \beta^0$ for sufficiently large κ , contradicting (B.8).

Claim 2. $v_{b_*} \leq v_b$ for all b : Let $v_b < v_{b_*}$ for some b and consider $\beta \in H$ of the form

$$\beta_b = -\kappa, \beta_{b_*} = 1 + \kappa, \text{ and } \beta_{b'} = 0 \text{ for } b' \neq b, b_*, \text{ where } \kappa > 0.$$

Then, using (B.5), $\beta^0 \geq_D b_* \geq_D \beta$, which implies $\beta^0 \geq_D \beta$. Hence $\beta \in D^0$. However, $v \cdot \beta = (1 + \kappa)v_{b_*} - \kappa v_b > v \cdot \beta_s$ for sufficiently large κ , contradicting (B.9).

Finally, normalize v to obtain u in N^* ($u_{b_*} = 0, u_{b^*} = 1$). Then $u \cdot \beta_s = s$ and (B.3) is satisfied. □

PROOF OF PROPOSITION B.1. We prove only sufficiency. Define s by (B.4), so that, by (10),

$$\beta_p \in y \text{ if and only if } p \geq s. \tag{B.10}$$

Let

$$U = \{u \in N^* : u \cdot \beta \geq s \quad \forall \beta \in y\}.$$

Then U is closed, convex, and comprehensive. We claim that s and U satisfy (B.1).

That $y \subset \{\beta \in \Delta(B) : \min_{u \in U} u \cdot \beta \geq s\}$ is immediate from the definition of U . Thus it remains to prove that

$$\min_{u \in U} u \cdot \beta \geq s \implies \beta \in y.$$

Suppose to the contrary that

$$\min_{u \in U} u \cdot \beta^0 \geq s \text{ and } \beta^0 \notin y. \tag{B.11}$$

Then by the lemma, there exists $u \in N^*$ such that

$$u \cdot \beta^0 < s \leq u \cdot \beta \text{ for every } \beta \in y. \tag{B.12}$$

Because $s \leq u \cdot \beta$ for every $\beta \in y$, it follows that $u \in U$. By (B.11), $u \cdot \beta^0 \geq s$. But this contradicts (B.12).

We now turn to uniqueness. Suppose that

$$y = \{\beta \in \Delta(B) : \min_{u \in U} u \cdot \beta \geq s\} = \{\beta \in \Delta(B) : \min_{u' \in U'} u' \cdot \beta \geq s'\}, \tag{B.13}$$

where U and s are as above and where U' is closed, convex, and comprehensive. That $s' = s$ follows from (B.10). Evidently, $U' \subset U$. To prove equality, let $u \in U \setminus U'$. Separate u and U' by some $\gamma \in \mathbb{R}^B$:

$$u \cdot \gamma < u' \cdot \gamma \text{ for all } u' \in U'.$$

Suppose that there exists $\gamma \in \Delta(B)$ satisfying the preceding (existence of such a γ follows from a separation argument in \mathbb{R}^B and from comprehensiveness of U'). Then, by mixing γ suitably with b_* and b^* , it is without loss of generality to assume that $u \cdot \gamma < s < u' \cdot \gamma$ for all $u' \in U'$. But this is impossible: by (B.13), $\gamma \in y$ (using the U' -representation) and $\gamma \notin y$ (using the U -representation). □

B.2 An intermediate result

In the sufficiency part of the proof, we use some results from [Epstein and Seo \(2007\)](#). They prove a representation result for preferences over random menus satisfying axioms that are collectively weaker than Order, Continuity, Nondegeneracy, and First-Stage Independence, plus a form of Dominance whose formal statement is identical to the one in this paper, but where the set Y of upper contour sets is required to satisfy a set of regularity conditions much weaker than Y1–Y3. (One qualification is that Epstein and Seo permit $\mathcal{B} \in Y$, but this is a matter of a harmless difference in normalizations. Otherwise, their conditions are trivial implications of (B.2).)

Therefore, deduce from [Epstein and Seo \(2007\)](#) that \succeq satisfying our axioms has a representation:

$$\mathcal{W}(P) = \int m(\{y \in \mathcal{X} : x \cap y \neq \emptyset\}) dP(x) \text{ for all } P \in \Delta(\mathcal{X}), \tag{B.14}$$

for some countably additive probability measure $m \in \Delta(\mathcal{X})$, viewed as a measure over upper contour sets, such that $m(Y) = 1$.²⁵

The rationale for (B.14) can be understood roughly as follows: Order, Continuity, and First-Stage Independence imply that there exists an expected utility representation for \succeq , that is,

$$\mathcal{W}(P) = \int W(x) dP(x)$$

for some continuous vNM index $W : \mathcal{X} \rightarrow [0, 1]$. The normalizations

$$W(\{b_*\}) = 0 \text{ and } W(\{b^*\}) = 1$$

can be shown to be without loss of generality given the remaining axioms. It is well-known that, under suitable assumptions, a set function, such as W , can be expressed in the form

$$W(x) = m(\{y \in \mathcal{X} : x \cap y \neq \emptyset\}) \quad \forall x \in \mathcal{X} \tag{B.15}$$

for some unique countably additive probability measure on the Borel σ -algebra generated by the Hausdorff topology on \mathcal{X} (see [Choquet 1954](#), Theorems 50.1 and 51.1, as well as the more recent works surveyed in [Marinacci and Montrucchio 2004](#)). The key property of W that permits such a representation is that it is *infinitely alternating*: for any finite set of menus $\{x_i\}_{i=1}^n$,

$$W(\cap_{i=1}^n x_i) \leq \sum_{\{I: \emptyset \neq I \subseteq \{1, \dots, n\}\}} (-1)^{|I|+1} W(\cup_{i \in I} x_i). \tag{B.16}$$

²⁵They prove also uniqueness of m , but we rely on [Epstein and Seo \(2007\)](#) only for the existence of m . In order to be as self-contained as possible, the uniqueness property that we need here is proven below—see [Lemma B.6](#).

One way to understand the representation result (B.14) in Epstein and Seo (2007) is that it shows that the axioms for preference \succeq adopted there imply that W is infinitely alternating.²⁶

Finally, Dominance leads to the restriction $m(Y) = 1$ (thanks to a version of (B.15) derived in Epstein and Seo 2007). This is because, as described in the text, Dominance imposes ex ante certainty that only upper contour sets in Y will be relevant ex post.

We use (B.14), including the restriction $m(Y) = 1$, heavily below where we derive the further implications of Certainty Reversal of Order. We also use the following lemma from Aliprantis and Border (1994, Theorem 14.69).

LEMMA B.2. *The Borel σ -algebra on \mathcal{X} coincides with the σ -algebra generated by the sets of the form*

$$\{y \in \mathcal{X} : x \cap y \neq \emptyset\} \text{ for } x \in \mathcal{X}.$$

B.3 *The remainder of the proof*

Necessity: Only Dominance requires proof. The utility function \mathcal{W} can be rewritten in the form (B.14). (Such a representation is implied not only by the axioms, as noted above, but also by the functional form (8); the measure m is described in (B.23) below.) Therefore,

$$\begin{aligned} \mathcal{W}(P) &= \int_{\mathcal{X}} \int_Y \max_{\beta \in x} \mathbf{1}_y(\beta) d m(y) d P(x) = \int \int \max_{\beta \in x} \mathbf{1}_y(\beta) d P(x) d m(y) \\ &= \int_Y P(\{x : x \cap y \neq \emptyset\}) d m(y) \end{aligned}$$

and

$$\mathcal{W}(P') - \mathcal{W}(P) = \int_Y (P'(\{x : x \cap y \neq \emptyset\}) - P(\{x : x \cap y \neq \emptyset\})) d m(y),$$

which is non-negative if P' dominates P .

The remainder of the proof establishes *sufficiency* of the axioms, and then the asserted uniqueness.

We sometimes adopt the abbreviation

$$\mathcal{B} = \Delta(B).$$

Set

$$h_U(\beta) = \min_{u \in U} u \beta, \tag{B.17}$$

and define $\Psi : \mathcal{K}^{cc}(N^*) \times [0, 1] \rightarrow Y$ by

$$\Psi(U, s) = \{\beta : h_U(\beta) \geq s\}. \tag{B.18}$$

²⁶Nehring (1996) was the first to show a connection between an ordinal property analogous to Dominance (that he calls Indirect Stochastic Dominance) and the cardinal property “infinitely alternating.” He derives a representation similar to (B.14) for his setting where menus consist of alternatives from the finite set B rather than lotteries over B .

Given $D \subset \mathcal{K}^{cc}(N^*)$, $\Psi(D, [0, s])$ is the image of $D \times [0, s]$, that is,

$$\Psi(D, [0, s]) = \{\Psi(U, s') : U \in D, s' \in [0, s]\}.$$

By **Proposition B.1**, Ψ is one-to-one and onto.

LEMMA B.3. *For any Borel set $D \subset \mathcal{K}^{cc}(N^*)$, $\Psi(D, [0, s])$ is a Borel-measurable subset of \mathcal{X} .*

PROOF. First observe that

$$(\beta, U) \mapsto h_U(\beta) \text{ is continuous.} \tag{B.19}$$

By the Maximum Theorem, it suffices to show that (i) $(\beta, u) \mapsto u\beta$ is jointly continuous, and (ii) the correspondence $\Gamma : \mathcal{B} \times \mathcal{K}^{cc}(N^*) \rightsquigarrow N^*$ defined by

$$\Gamma(\beta, U) = U$$

is continuous. Condition (i) is clear. For (ii), when viewed as a function from $\mathcal{B} \times \mathcal{K}^{cc}(N^*)$ into $\mathcal{K}^{cc}(N^*)$, Γ is continuous under the Hausdorff metric. Therefore, (ii) follows from **Aliprantis and Border (1994, Theorem 14.16)**.

We claim that Ψ is continuous. By **Aliprantis and Border (1994, Theorem 14.16)**, it suffices to show that Ψ is a continuous correspondence.

To show that Ψ is upper hemicontinuous, we let $(U_n, s_n) \rightarrow (U, s)$ and $\beta_n \in \Psi(U_n, s_n)$ and prove that β_n has a limit point in $\Psi(U, s)$. Since \mathcal{B} is compact, there is a convergent subsequence β_{n_k} . Let $\beta_{n_k} \rightarrow \beta$. Then $h_{U_{n_k}}(\beta_{n_k}) \geq s_{n_k}$ for every k , which, by **(B.19)**, implies that $h_U(\beta) \geq s$.

We now show that Ψ is lower hemicontinuous. Let $(U_n, s_n) \rightarrow (U, s)$ and $\beta \in \Psi(U, s)$. We need to show that there is a sequence $\beta_n \in \Psi(U_n, s_n)$ such that $\beta_n \rightarrow \beta$. Define

$$t_n = \begin{cases} \min \left\{ \frac{1 - s_n}{1 - h_{U_n}(\beta)}, 1 \right\} & \text{if } h_{U_n}(\beta) < 1 \\ 1 & \text{if } h_{U_n}(\beta) = 1 \end{cases}$$

$$\beta_n = t_n\beta + (1 - t_n)b^*.$$

If $h_{U_n}(\beta) < 1$, then

$$\begin{aligned} h_{U_n}(\beta_n) &= t_n h_{U_n}(\beta) + (1 - t_n) \\ &= -t_n(1 - h_{U_n}(\beta)) + 1 \\ &\geq -\frac{1 - s_n}{1 - h_{U_n}(\beta)}(1 - h_{U_n}(\beta)) + 1 \geq s_n, \end{aligned}$$

and if $h_{U_n}(\beta) = 1$, then

$$h_{U_n}(\beta_n) = h_{U_n}(\beta) = 1 \geq s_n.$$

Hence, $\beta_n \in \Psi(U_n, s_n)$. And by **(B.19)**, $\lim_n(1 - h_{U_n}(\beta)) \leq \lim_n(1 - s_n)$, which implies $t_n \rightarrow 1$ and $\beta_n \rightarrow \beta$.

Finally, fix $s \in [0, s]$ and let \mathcal{D} be the set of measurable sets D such that $\Psi(D, [0, s])$ is Borel measurable. Since Ψ is one-to-one, \mathcal{D} is easily seen to be a σ -algebra. Moreover, when D is closed, it is compact and hence, by continuity of Ψ , $\Psi(D, [0, s])$ is also compact and hence measurable. Thus \mathcal{D} contains every closed set and thus also every Borel-measurable set. \square

LEMMA B.4. *Let $m \in \Delta(\mathcal{X})$ be the probability measure satisfying (B.14). Then Certainty Reversal of Order implies that, for each measurable $D \subset \mathcal{X}^{cc}(N^*)$,*

$$m(\Psi(D, [0, s])) = s \cdot m(\Psi(D, [0, 1])).$$

PROOF. Recall that $m(Y) = 1$. Therefore, by Proposition B.1, it suffices to consider upper contour sets of the form $y = \{\beta : h_U(\beta) \geq s\}$. For any such set y and $0 \leq t \leq 1$, define $t * y = \{\beta : h_U(\beta) \geq ts\}$. For $A \subset Y$ and $0 \leq t \leq 1$, let

$$t * A = \{t * y : y \in A\}.$$

STEP 1. $t * \{y \in Y : y \cap x \neq \emptyset\} = \{y \in Y : y \cap (tx + (1-t)\{b_*\}) \neq \emptyset\}$

PROOF. When $t = 0$, both sets are empty. Let $0 < t \leq 1$. It suffices to show that

$$\begin{aligned} y' \in Y \text{ and } y' = t * y \text{ for some } y \in Y \text{ such that } y \cap x \neq \emptyset \\ \iff [y' \in Y \text{ and } y' \cap (tx + (1-t)\{b_*\}) \neq \emptyset]. \end{aligned}$$

\implies : Assume that $y' = t * y$ and $y \cap x \neq \emptyset$. Let $\hat{\beta} \in y \cap x$ and $\beta' = t\hat{\beta} + (1-t)b_*$. Then,

$$\min_{u \in U} u\beta' = \min_{u \in U} u(t\hat{\beta} + (1-t)b_*) = t \min_{u \in U} u\hat{\beta} \geq ts.$$

Hence $\beta' \in y'$. Also, $t\hat{\beta} + (1-t)b_* \in tx + (1-t)\{b_*\}$. Therefore, $y' \cap (tx + (1-t)\{b_*\}) \neq \emptyset$.

\impliedby : Let $y' \in Y$ and $y' \cap (tx + (1-t)\{b_*\}) \neq \emptyset$. By the latter condition, there exists $\hat{\beta} \in x$ such that

$$\beta' = t\hat{\beta} + (1-t)b_* \in y'.$$

Since $y' \in Y$, there exist $U \in \mathcal{X}^{cc}(N^*)$ and $s \in [0, 1]$ such that $y' = \{\beta : h_U(\beta) \geq s\}$. Moreover $s \leq t$, because

$$s \leq \min_{u \in U} u\beta' = \min_{u \in U} u \cdot (t\hat{\beta} + (1-t)b_*) = t \min_{u \in U} u\hat{\beta} \leq t.$$

Let $y = \{\beta : h_U(\beta) \geq s/t\}$. Then

$$t * y = \{\beta : \min_{u \in U} u\beta \geq t \cdot (s/t)\} = y'.$$

Moreover, $\hat{\beta} \in y$ because

$$t \min_{u \in U} u\hat{\beta} = \min_{u \in U} u \cdot (t\hat{\beta} + (1-t)b_*) = \min_{u \in U} u\beta' \geq s.$$

Hence $\hat{\beta} \in x \cap y \neq \emptyset$. This completes the step. \triangleleft

STEP 2. $m(t * \{y \in Y : y \cap x \neq \emptyset\}) = t \cdot m(\{y \in Y : y \cap x \neq \emptyset\})$

PROOF. By Step 1, (B.14), and Certainty Reversal of Order,

$$\begin{aligned} m(t * \{y \in Y : y \cap x \neq \emptyset\}) &= m(\{y \in Y : y \cap (tx + (1-t)\{b_*\}) \neq \emptyset\}) \\ &= m(\{y \in \mathcal{X} : y \cap (tx + (1-t)\{b_*\}) \neq \emptyset\}) \\ &= W(tx + (1-t)\{b_*\}) \\ &= t \cdot W(x) = t \cdot m(\{y \in Y : y \cap x \neq \emptyset\}). \quad \triangleleft \end{aligned}$$

STEP 3. For any Borel set $A \subset Y$, we have $m(t * A) = t \cdot m(A)$

PROOF. Let Σ be the collection of all Borel sets $A \subset Y$ satisfying the noted condition. It is easy to verify that Σ is a Dynkin system: (i) $Y \in \Sigma$; (ii) if $A, A' \in \Sigma$ and $A \subset A'$, then $A' \setminus A \in \Sigma$; (iii) if a sequence $\{A_n\}$ in Σ is such that $A_n \nearrow A$, then $A \in \Sigma$. The proofs are as follows.

(i) $m(t * Y) = m(t * \{y \in Y : y \cap \mathcal{B} \neq \emptyset\}) = t \cdot m(\{y \in Y : y \cap \mathcal{B} \neq \emptyset\}) = t \cdot m(Y)$, where the second equality is by Step 2.

(ii) Since $t * (A' \setminus A) = (t * A') \setminus (t * A)$ and $t * A \subset t * A'$,

$$\begin{aligned} m(t * (A' \setminus A)) &= m(t * A') - m(t * A) \\ &= t \cdot m(A') - t \cdot m(A) = t \cdot m(A' \setminus A). \end{aligned}$$

(iii) $m(t * A) = \lim m(t * A_n) = \lim t \cdot m(A_n) = t \cdot m(A)$.

Let \mathcal{F} be the π -system generated by sets of the form $\{y \in Y : y \cap x \neq \emptyset\}$. That is, \mathcal{F} is the smallest family of subsets such that \mathcal{F} is closed under finite intersections and $\{y \in Y : y \cap x \neq \emptyset\} \in \mathcal{F}$.

We show that $\mathcal{F} \subset \Sigma$. Let $\mathcal{F}_n = \{\cap_{i=1}^n \{y \in Y : y \cap x_i \neq \emptyset\} : x_i \in \mathcal{X}\}$. It suffices to show that $\mathcal{F}_n \subset \Sigma$ for all $n \geq 1$. We argue by induction. By Step 2, $\mathcal{F}_1 \subset \Sigma$. We suppose $\mathcal{F}_n \subset \Sigma$ and show that $\mathcal{F}_{n+1} \subset \Sigma$. Let $A_n = \cap_{i=1}^n \{y \in Y : y \cap x_i \neq \emptyset\}$ and $A_{n+1} = A_n \cap \{y \in Y : y \cap x_{n+1} \neq \emptyset\} \in \mathcal{F}_{n+1}$. Then

$$\begin{aligned} m(t * A_{n+1}) &= m(t * A_n) + m(t \{y \in Y : y \cap x_{n+1} \neq \emptyset\}) \\ &\quad - m[t * (A_n \cup \{y \in Y : y \cap x_{n+1} \neq \emptyset\})]. \end{aligned}$$

The last term can be rewritten as

$$\begin{aligned} m[t * ((\cap_{i=1}^n \{y \in Y : y \cap x_i \neq \emptyset\}) \cup \{y \in Y : y \cap x_{n+1} \neq \emptyset\})] \\ &= m[t * (\cap_{i=1}^n \{y \in Y : y \cap x_i \neq \emptyset\} \cup \{y \in Y : y \cap x_{n+1} \neq \emptyset\})] \\ &= m[t * (\cap_{i=1}^n \{y \in Y : y \cap (x_i \cup x_{n+1}) \neq \emptyset\})] \\ &= t \cdot m[\cap_{i=1}^n \{y \in Y : y \cap (x_i \cup x_{n+1}) \neq \emptyset\}] \text{ (by } \mathcal{F}_n \subset \Sigma) \\ &= t \cdot m[(\cap_{i=1}^n \{y \in Y : y \cap x_i \neq \emptyset\}) \cup \{y \in Y : y \cap x_{n+1} \neq \emptyset\}] \\ &= t \cdot m[A_n \cup \{y \in Y : y \cap x_{n+1} \neq \emptyset\}]. \end{aligned}$$

Thus

$$m(t * A_{n+1}) = t \cdot m(A_n) + t \cdot m(\{y \in Y : y \cap x_{n+1} \neq \emptyset\}) - t \cdot m[A_n \cup \{y \in Y : y \cap x_{n+1} \neq \emptyset\}] = t \cdot m(A_{n+1}).$$

Hence $\mathcal{F} \subset \Sigma$.

Therefore, by Dynkin's Lemma (Aliprantis and Border 1994, Theorem 8.10), Σ includes the σ -algebra generated by \mathcal{F} , which is the Borel σ -algebra (Lemma B.2). \triangleleft

Finally, note that $\Psi(D, [0, s]) = s * \Psi(D, [0, 1])$ and $\Psi(D, [0, 1]) \subset Y$. Therefore, by Lemma B.3 and Step 3, $m(\Psi(D, [0, s])) = s \cdot m(\Psi(D, [0, 1]))$. \square

The desired representation will be established using the measure $\mu \in \Delta(\mathcal{K}^{cc}(N^*))$ defined by

$$\mu(D) = m(\Psi(D, [0, 1])) \text{ for each Borel set } D \subset \mathcal{K}^{cc}(N^*).$$

LEMMA B.5. *The function \mathcal{W} in (B.14) satisfies*

$$\mathcal{W}(P) = \int \int \max_{\beta \in x} \min_{u \in U} u \cdot \beta \, d\mu(U) \, dP(x) \quad \text{for each } P \in \Delta(\mathcal{K}(\mathcal{B})).$$

PROOF. First we claim that, for any $x \in \mathcal{K}(\mathcal{B})$,²⁷

$$\{\Psi(U, s) : \max_{\beta \in x} h_U(\beta) \geq s\} = \{y \in Y : y \cap x \neq \emptyset\}. \tag{B.20}$$

\subset : $\max_{\beta \in x} h_U(\beta) \geq s \implies \exists \beta \in x, h_U(\beta) \geq s \implies \Psi(U, s) \in \{y \in Y : y \cap x \neq \emptyset\}$.

\supset : If $y \in Y$ and $y \cap x \neq \emptyset$, then $y = \Psi(U, s)$ for some $(U, s) \in \mathcal{K}^{cc}(N^*) \times [0, 1]$, and $\exists \beta \in \Psi(U, s) \cap x$. Then $h_U(\beta) \geq s$. Thus $\max_{\beta \in x} h_U(\beta) \geq s$ and $y = \Psi(U, s) \in \{\Psi(U', s') : \max_{\beta \in x} h_{U'}(\beta) \geq s'\}$.

Since $U \mapsto \max_{\beta \in x} \min_{u \in U} u \cdot \beta$ is bounded and continuous, it is μ -integrable. Thus there is a decreasing sequence of step functions $\phi_n(U) = \sum_{i=1}^n \alpha_{n,i} I_{D_{n,i}}(U)$ such that ϕ_n converges to the function $U \mapsto \max_{\beta \in x} h_U(\beta)$ and such that

$$\int \max_{\beta \in x} h_U(\beta) \, d\mu(U) = \lim \int \phi_n \, d\mu.$$

²⁷Recall that h_U and Ψ are defined in (B.17) and (B.18).

Then

$$\begin{aligned}
& \int \max_{\beta \in x} h_U(\beta) d\mu(U) \\
&= \lim \sum_{i=1}^n \alpha_{n,i} \mu(D_{n,i}) \\
&= \lim \sum_{i=1}^n \alpha_{n,i} m(\Psi(D_{n,i}, [0, 1])) \text{ (by definition of } \mu) \\
&= \lim \sum_{i=1}^n m(\Psi(D_{n,i}, [0, \alpha_{n,i}])) \text{ (by Lemma B.4)} \\
&= \lim m\left(\bigcup_i^n \Psi(D_{n,i}, [0, \alpha_{n,i}])\right) \text{ (the sets } \Psi(D_{n,i}, [0, \alpha_{n,i}]) \text{ are disjoint)} \\
&= \lim m(\{\Psi(U, s) : s \leq \phi_n(U)\}) \text{ (by definition of } \Psi \text{ and } \phi_n) \\
&= m\left(\bigcap_n \{\Psi(U, s) : s \leq \phi_n(U)\}\right) \text{ (since } m \text{ is c.a.)}
\end{aligned}$$

Since

$$\bigcap_n \{\Psi(U, s) : s \leq \phi_n(U)\} = \{\Psi(U, s) : \max_{\beta \in x} h_U(\beta) \geq s\},$$

we have

$$\begin{aligned}
\int \max_{\beta \in x} h_U(\beta) d\mu(U) &= m(\{\Psi(U, s) : \max_{\beta \in x} h_U(\beta) \geq s\}) \\
&= m(\{y \in Y : y \cap x \neq \emptyset\}) \text{ (by (B.20)).}
\end{aligned}$$

Therefore, by (B.14),

$$\int \int \max_{\beta \in x} h_U(\beta) d\mu(U) dP(x) = \int m(\{y \in Y : y \cap x \neq \emptyset\}) dP(x) = \mathcal{W}(P). \quad \square$$

It remains to prove the uniqueness assertion.

LEMMA B.6. *Given any $\mu \in \Delta(\mathcal{X}^{cc}(N^*))$ such that preference \succeq is represented by the utility function in (8), define $W : \mathcal{X} \rightarrow \mathbb{R}$ by*

$$W(x) = \int_{\mathcal{X}^{cc}(N^*)} \max_{\beta \in x} h_U(\beta) d\mu(U). \quad (\text{B.21})$$

Then there exists a unique Borel probability measure m_μ on \mathcal{X} such that

$$W(x) = m_\mu(\{y : x \cap y \neq \emptyset\}) \quad \forall x \in \mathcal{X}. \quad (\text{B.22})$$

Further, for all Borel measurable sets B ,

$$m_\mu(B) = \int_0^1 \mu(\{U : \Psi(U, s) \in B\}) d\lambda(s), \tag{B.23}$$

where λ denotes the Lebesgue measure.

PROOF. The set function W is infinitely alternating, that is, it satisfies (B.16). Moreover, since μ is countably additive, the Monotone Convergence Theorem implies that $W(x_n) \downarrow W(x)$ whenever $x_n \downarrow x$. Thus the existence of a unique measure m_μ satisfying (B.22) follows from Choquet (1954, Theorems 50.1, 51.1). (In fact, existence of m_μ was asserted in Section B.2 and was used in Lemma B.4—the new claim here is uniqueness.)

For each $s \in [0, 1]$, define $\Psi_s : \mathcal{X}^{cc}(N^*) \rightarrow Y$ by $\Psi_s(U) = \Psi(U, s)$. As a correspondence, Ψ_s is compact-valued and upper hemicontinuous. By Aliprantis and Border (1994, Theorem 14.11), it is a closed correspondence. Hence, by Aliprantis and Border (1994, Theorem 14.68) it is measurable, and so by Aliprantis and Border (1994, Corollary 14.70), when viewed as a function, Ψ_s is Borel measurable, that is, $\Psi_s^{-1}(B) = \{U : \Psi(U, s) \in B\}$ is Borel measurable for every measurable $B \subset \mathcal{X}$.

Therefore, $s \mapsto \mu \circ \Psi_s^{-1}(B)$ is well-defined for each measurable B ; it is also Lebesgue integrable (see Epstein and Seo 2007, Lemma B.2). Define, for every measurable B ,

$$m'_\mu(B) = \int_0^1 \mu \circ \Psi_s^{-1}(B) d\lambda(s).$$

By the Monotone Convergence Theorem, m'_μ is easily seen to be a Borel measure. Moreover, for all $x \in \mathcal{X}$,

$$\begin{aligned} W(x) &= \int_V \max_{\beta \in x} h_U(\beta) d\mu(U) = \int_0^1 \mu(\{U : \max_{\beta \in x} h_U(\beta) \geq s\}) d\lambda(s) \\ &= \int_0^1 \mu(\{U : x \cap \{\beta \in \mathcal{B} : h_U(\beta) \geq s\} \neq \emptyset\}) d\lambda(s) \\ &= \int_0^1 \mu \circ \Psi_s^{-1}(\tilde{x}) d\lambda = m'_\mu(\tilde{x}), \end{aligned}$$

where $\tilde{x} = \{y \in Y : x \cap y \neq \emptyset\}$. Therefore, by (B.22), $m'_\mu(\tilde{x}) = m_\mu(\tilde{x})$ for all $x \in \mathcal{X}$. By Lemma B.2 and the uniqueness property established in first part of the proof, we conclude that $m'_\mu(B) = m_\mu(B)$ for all measurable B . □

LEMMA B.7. Suppose that μ and μ' represent the same preference as in (8). Then $\mu = \mu'$.

PROOF. Define W and W' as in (B.21). Since both W and W' are vNM indices for expected utility functions representing \succeq , they must be identical (given also that they agree on $\{b_*\}$ and $\{b^*\}$). Let m_μ and $m_{\mu'}$ be the Borel probability measures on \mathcal{X} corresponding to μ and μ' as in Lemma B.6. Then by Lemma B.6, $m_\mu = m_{\mu'}$. Therefore, for each

Borel set $D \subset \mathcal{H}^{cc}(N^*)$,

$$\begin{aligned}\mu(D) &= \int_0^1 \mu(D) ds = \int_0^1 \mu(\{U : \Psi(U, s) \in \Psi(D, [0, 1])\}) ds \\ &= m_\mu(\Psi(D, [0, 1])) = m_{\mu'}(\Psi(D, [0, 1])) = \mu'(D).\end{aligned}\quad \square$$

REFERENCES

- Aliprantis, Charalambos D. and Kim C. Border (1994), *Infinite Dimensional Analysis*. Springer, Berlin. [368, 386, 387, 390, 392]
- Choquet, Gustave (1954), “Theory of capacities.” *Annales de l’Institut Fourier*, 5, 131–295. [385, 392]
- de Jonge, E. and A. C. M. van Rooij (1977), *Introduction to Riesz spaces*. Mathematisch Centrum, Amsterdam. [375]
- Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini (2001), “Representing preferences with a unique subjective state space.” *Econometrica*, 69, 891–934. [356]
- Dempster, Arthur P. (1967), “Upper and lower probabilities induced by a multivalued mapping.” *Annals of Mathematical Statistics*, 38, 325–339. [357]
- Epstein, Larry G. and Martin Schneider (2003), “Recursive multiple-priors.” *Journal of Economic Theory*, 113, 1–31. [359, 367]
- Epstein, Larry G. and Kyoungwon Seo (2007), “Subjective states: a more robust model.” Unpublished paper, Boston University. [366, 368, 385, 386, 392]
- Ergin, Haluk and Todd Sarver (2005), “A unique costly contemplation representation.” Unpublished paper, Massachusetts Institute of Technology. [372, 377]
- Ghirardato, Paolo (2001), “Coping with ignorance: Unforeseen contingencies and non-additive uncertainty.” *Economic Theory*, 17, 247–276. [357, 371]
- Gilboa, Itzhak and David Schmeidler (1989), “Maxmin expected utility with non-unique prior.” *Journal of Mathematical Economics*, 18, 141–153. [357, 359, 363, 365]
- Greenberg, Harvey J. and William P. Pierskalla (1971), “A review of quasi-convex functions.” *Operations Research*, 19, 1553–1570. [372]
- Herstein, I. N. and John Milnor (1953), “An axiomatic approach to measurable utility.” *Econometrica*, 21, 291–297. [364]
- Hörmander, Lars (1955), “Sur la fonction d’appui des ensembles convexes dans un espace localement convexe.” *Arkiv för Matematik*, 3, 181–186. [379]
- Kreps, David M. (1979), “A representation theorem for ‘preference for flexibility’.” *Econometrica*, 47, 565–577. [356, 368]

- Kreps, David M. (1992), "Static choice in the presence of unforeseen contingencies." In *Economic analysis of markets and games* (Partha Dasgupta, Douglas Gale, Oliver Hart, and Eric Maskin, eds.), 258–281, MIT Press, Cambridge. [356, 358]
- Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini (2006), "Ambiguity aversion, robustness, and the variational representation of preferences." *Econometrica*, 74, 1447–1498. [365, 372]
- Machina, Mark J. (1984), "Temporal risk and the nature of induced preferences." *Journal of Economic Theory*, 33, 199–231. [361]
- Marinacci, Massimo and Luigi Montrucchio (2004), "Introduction to the mathematics of ambiguity." In *Uncertainty in Economic Theory* (Itzhak Gilboa, ed.), 46–107, Routledge, London. [385]
- Mukerji, Sujoy (1997), "Understanding the nonadditive probability decision model." *Economic Theory*, 9, 23–46. [357, 371, 372]
- Nehring, Klaus (1996), "Preference for flexibility and freedom of choice in a Savage framework." Working Paper, University of California, Davis. [358, 366, 368, 386]
- Nehring, Klaus (1999), "Preference for flexibility in a Savage framework." *Econometrica*, 67, 101–119. [356, 358, 366, 368]
- Phelps, Robert R. (1993), *Convex Functions, Monotone Operators and Differentiability*, second edition, volume 1364 of *Lecture Notes in Mathematics*. Springer, Berlin. [377]
- Schmeidler, David (1989), "Subjective probability and expected utility without additivity." *Econometrica*, 57, 571–587. [357, 372]
- Shafer, Glenn (1976), *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, New Jersey. [357]