Ambiguity Aversion, Malevolent Nature, and the Variational Representation of Preferences*

Fabio Maccheroni[a] Massimo Marinacci[b] Aldo Rustichini[c]

[a]Istituto di Metodi Quantitativi and IGIER, Università Bocconi

[b]Dipartimento di Statistica e Matematica Applicata and ICER, Università di Torino

[c]Department of Economics, University of Minnesota

June 2004

Abstract

In the classic Anscombe and Aumann decision setting, we give necessary and sufficient conditions that guarantee the existence of a *utility function* u on outcomes and an *ambiguity index* c on the set of all probabilities on the states of the world such that, for all acts f and g,

$$f \succsim g \Leftrightarrow \min_p \left(\int u(f)\, dp + c(p) \right) \geq \min_p \left(\int u(g)\, dp + c(p) \right).$$

The function u represents the decision maker's risk attitudes, while the index c captures his ambiguity attitudes.

The preferences we characterize include as special cases the *multiple priors preferences* of Gilboa and Schmeidler, the *multiplier preferences* of Hansen and Sargent, and the *mean-variance preferences* of Markowitz and Tobin. In this way we are able to provide a rigorous decision-theoretic foundation on the latter two models, which have been widely used in macroeconomics and finance.

*We thank Erio Castagnoli, Rose-Anne Dana, Larry Epstein, Peter Klibanoff, Bart Lipman, Mark Machina, Jianjun Miao, Sujoy Mukerji, and Emre Ozdenoren for helpful discussions and suggestions. We also thank seminar audiences at Bocconi, Boston Univ., Columbia, INSEAD, Messina, Michigan, NYU, Paris 1, Roma "La Sapienza", and UCL. Maccheroni and Marinacci gratefully acknowledge the financial support of the Ministero dell'Istruzione, dell'Università e della Ricerca; Rustichini gratefully acknowledges the financial support of the National Science Foundation (grant # 0136556).

1 Introduction

In the past few years there has been a growing dissatisfaction in macroeconomics toward the so-called communism of models imposed by the rational expectations hypothesis. Under this assumption all agents share the same probabilistic model (i.e., the same probability distribution on some relevant economic phenomenon) and this model has to be correct, that is, it has to be the model governing the given phenomenon.

This is a strong requirement as agents can have different models, each of them being only an approximation of the underlying "true" model. To deal with this misspecification issue, robust control models have recently received a great deal of attention, starting with the work of Hansen and Sargent (see, e.g., [17] and [18]).

In the robust approach agents' objective functions take into account the possibility that their model may not be the correct one, but only an approximation. Specifically, agents rank payoff profiles f according to the following choice criterion

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \theta R(p \| q) \right\}, \tag{1}$$

where Δ is the set of all probabilities and $R(\cdot \| q) : \Delta \to [0, \infty]$ is the relative entropy (see Section 4.2 for the definition). Preferences represented by criterion (1) are called *multiplier preferences*.

Agents behaving according to this choice criterion are considering the possibility that q may not be the appropriate law governing the phenomenon which they are interested in, and for this reason they take into account other possible models p.

The relative likelihood of these alternative models p is measured by the relative entropy, while the positive parameter θ reflects the weight that agents are giving to the possibility that q might not be the correct model – that is, the extent to which they view q as a mere approximation. Since $R(p\|q) = 0$ if and only if $p = q$, when the parameter θ becomes bigger, agents focus more only on q as the correct model, thus giving less importance to possible alternative models p.

As Hansen and Sargent [17] have pointed out, model uncertainty can be viewed as the outcome of ambiguity, resulting from the possibly poor quality of the information on which agents base the choice of the model they use. Ambiguity is a classic issue in Decision Theory since the seminal work of Ellsberg [11]. A popular class of preferences dealing with it are the *multiple priors preferences* axiomatized by Gilboa and Schmeidler [15] (also known as Maxmin Expected Utility preferences). Agents having such preferences rank payoff profiles according to the following criterion

$$V(f) = \min_{p \in C} \int u(f)\, dp, \tag{2}$$

where C is a given convex subset of the set Δ of all probabilities. The set C is interpreted as a set of priors held by agents, and ambiguity is reflected by its possibly non-singleton nature.

As discussed at length by Hansen and Sargent [17], the motivation behind multiplier preferences is closely connected to the one underlying multiple priors preferences. In this paper, we intend to make such connection precise by presenting a general class of preferences that includes both multiplier and multiple prior preferences as special cases. To see in more detail our approach, observe that the multiple priors criterion (2) can be written as follows

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \delta_C(p) \right\},$$

where $\delta_C : \Delta \to [0, \infty]$ is the indicator function of Convex Analysis given by

$$\delta_C(p) = \begin{cases} 0 & \text{if } p \in C, \\ \infty & \text{otherwise.} \end{cases}$$

Like the relative entropy, also the indicator function is a convex function defined on the simplex Δ.

All this suggests the following general representation

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p) \right\}, \tag{3}$$

where $c : \Delta \to [0, \infty]$ is a convex function on the simplex. The study of this representation, which includes both (1) and (2) as special cases, is the subject matter of this paper.

We first axiomatize the representation (3), by showing how it rests on a simple set of axioms that generalizes the multiple priors axiomatization of Gilboa and Schmeidler [15]. We then show how to interpret in a rigorous way the function c as an index of ambiguity aversion: the lower is c, the higher is the ambiguity aversion exhibited by the agent. The relative entropy $\theta R(\cdot \| q)$ and the indicator function $\delta_C(p)$ can thus be viewed as special instances of ambiguity indices. All this clarifies the conceptual and mathematical connections between the choice functionals (1) and (2).

As a dividend of our analysis, we then show that a third classic class of preferences, the *mean-variance preferences* of Markowitz [21] and Tobin [25], can be viewed as a special case of our setting. Recall that mean-variance preferences are represented by the preference functional

$$V(f) = \int f dq - \frac{\theta}{2} \mathrm{Var}(f). \tag{4}$$

We show that, on the domain of monotonicity of V, the following equality holds:

$$\int f dq - \frac{\theta}{2} \mathrm{Var}(f) = \min_{p \in \Delta} \left\{ \int f dp + \frac{1}{2\theta} G(p \| q) \right\},$$

where $G(\cdot\|q) : \Delta \to [0, \infty]$ is the relative Gini concentration index (see Section 4.3 for the definition). As a result, the mean-variance preference functional (4) is a special case of our representation (3). Interestingly, the associated index of ambiguity aversion is the relative version of the classic Gini concentration index. After Shannon's entropy, a second classic concentration index thus comes up in our analysis.

The interpretation of the mean-variance choice functional as reflecting ambiguity is very similar to the one used for the robust functional. Here we can view agents as considering q only as an approximation of the correct model; for this reason they take into account other possible models, whose relative likelihood is now determined by the relative Gini index. Again, the parameter θ measures the extent to which agents consider q only as an approximation.

Summing up, in this paper we generalize a popular class of preferences dealing with ambiguity, the multiple priors preferences, and in this way we are able to provide a rigorous decision-theoretic foundation on two widely used classes of preferences, the multiplier preferences of Hansen and Sargent and the mean-variance preferences of Markowitz and Tobin. As a secondary contribution, we provide a setting in which the two most classic concentration indices, Shannon's entropy and Gini's index, have a natural decision-theoretic interpretation.

1.1 Ambiguity and Nature

There are two notions of ambiguity, proposed by Epstein [12] and Ghirardato and Marinacci [14]. The key difference in the two approaches lies in the different notion of ambiguity neutrality they use: while [14] identifies ambiguity neutrality with subjective expected utility preferences, [12] more generally identifies ambiguity neutrality with probabilistically sophisticated preferences. In a nutshell, [14] claims that, unless the setting is rich enough, probabilistically sophisticated preferences may be compatible with behavior that intuitively can be viewed as generated by ambiguity. For this reason, they only consider subjective expected utility preferences as ambiguity neutral preferences.[1]

The ambiguity interpretation of our preference functional (3) relies on the notion of ambiguity proposed by [14]. This motivates the notions of absolute and comparative ambiguity aversion we will consider in Subsection 3.3, and therefore our interpretation of the preference functional (3) as reflecting ambiguity. In contrast, under the approach of [12] some specifications of (3) (as well as of (2)) may not reflect ambiguity. For

[1] We refer to [12] and [14] for a detailed presentation and motivation of their approaches. Notice that [14]'s notion of ambiguity aversion is what provides a foundation for the standard comparative statics exercises in ambiguity for multiple priors preferences that are based on the size of the set of priors.

example, this is the case of Hansen and Sargent's multiplier preferences, which are easily seen to be probabilistically sophisticated.

There is, however, an alternative interpretation of our preferences, which is in a sense more fundamental than ambiguity. Consider an agent that has to make choices having only limited information, without a full understanding of what is going on. A cautious agent may believe that "out there" there might be somebody/something that in some way might take advantage of this uncertainty and turn it against him. When this attitude is extreme it may result in paranoid behavior. But, normally, it would lead to cautious behavior, without any pathological feature.

This is a fundamental psychological attitude, which can be relevant in many choice situations – including Ellsberg-type choice situations – and our preference functional (3) can be viewed as modelling such psychological treat. In fact, agents ranking payoff profiles according to (3) can be viewed as believing they are playing a zero-sum game against (a malevolent) Nature. In such game, Nature is choosing the true model in order to make agents the most possible worst off; therefore, agents' play is determined by

$$\max_{f} \min_{p} \left\{ \int u(f)\, dp + c(p) \right\}, \tag{5}$$

and strategies consist of pairs (f, p), where Nature chooses model p and agents choose act f.

This interpretation is suggested by Hansen and Sargent [17] as a possible rationale for their multiplier preferences, and it can be used more generally for our variational preferences. Under this interpretation, the ambiguity index c can be viewed as the cost that Nature faces in choosing among models. Higher functions c now mean bigger costs for Nature in making its choices.

Needless to say, this game-theoretic view should not be taken at face value, but as a way to model the fundamental psychological treat we discussed before. The modelling idea here is that agents having such attitude can be modelled *as if* they were thinking of playing a game against Nature. In this regard, observe that the cost function $c(p)$ is a subjective construction of the agents, as it is the structure – given by (5) – of the game against Nature in which they feel engaged. Alternatively, the subjective cost function $c(p)$ can be viewed as reflecting the agents' perception of the relative plausibility of the different Nature's choices, that is, of the different models p that Nature can choose.

The game-theoretic interpretation of our setting would be especially useful in a dynamic extension of it, which is the natural next step in our analysis. In the current static setting it is just a matter of interpretation, in which the reader who adopts this game-theoretic perspective should regard c as a cost function throughout the paper. In terms of notation, this is a reason why we denote by c the index of ambiguity aversion. Another reason for this choice of notation is that it reminds the set C of priors featured by multiple priors preferences.

The paper is organized as follows. After introducing the set up in Section 2, we present the main representation result in Section 3. In the same section we discuss the ambiguity attitudes featured by the preferences we axiomatize. In Section 4 we show that our preferences include as special cases multiple priors preferences, multiplier preferences, and mean-variance preferences. Proofs and related material are collected in the Appendices.

2 Set Up

Consider a set S of *states of the world*, an algebra Σ of subsets of S called *events*, and a set X of *consequences*. We denote by $\mathcal{F}$ the set of all the *(simple) acts*: finite-valued functions $f : S \to X$ which are Σ-measurable. Moreover, we denote by $B_0(\Sigma)$ the set of all real-valued Σ-measurable simple functions, so that $u(f) \in B_0(\Sigma)$ whenever $u : X \to \mathbb{R}$.[2]

Given any $x \in X$, define $x \in \mathcal{F}$ to be the constant act such that $x(s) = x$ for all $s \in S$. With the usual slight abuse of notation, we thus identify X with the subset of the constant acts in $\mathcal{F}$. If $f \in \mathcal{F}$, $x \in X$, and $A \in \Sigma$, we denote by $xAf \in \mathcal{F}$ the act yielding x if $s \in A$ and $f(s)$ if $s \notin A$.

We assume additionally that X is a convex subset of a vector space. For instance, this is the case if X is the set of all the lotteries on a set of prizes, as it happens in the classic setting of Anscombe and Aumann [2]. Using the linear structure of X we can define as usual for every $f, g \in \mathcal{F}$ and $\alpha \in [0, 1]$ the act $\alpha f + (1 - \alpha)g \in \mathcal{F}$, which yields $\alpha f(s) + (1 - \alpha)g(s) \in X$ for every $s \in S$.

We model the decision maker's *preferences* on $\mathcal{F}$ by a binary relation $\succsim$. As usual, $\succ$ and $\sim$ denote respectively the asymmetric and symmetric parts of $\succsim$. If $f \in \mathcal{F}$, an element $x_f \in X$ is a *certainty equivalent for* f if $f \sim x_f$.

3 Representation

3.1 Axioms

In the sequel we make use of the following properties of $\succsim$.

A.1 *Weak Order.* If $f, g, h \in \mathcal{F}$: (a) either $f \succsim g$ or $g \succsim f$, (b) $f \succsim g$ and $g \succsim h$ imply $f \succsim h$.

[2] $u(f) : S \to \mathbb{R}$ is the function defined by $u(f)(s) = u(f(s))$ for all $s \in S$.

A.2 *Weak Certainty-Independence.* If $f, g \in \mathcal{F}$, $x, y \in X$, and $\alpha \in (0,1)$,

$$\alpha f + (1-\alpha)x \succsim \alpha g + (1-\alpha)x \Rightarrow \alpha f + (1-\alpha)y \succsim \alpha g + (1-\alpha)y.$$

A.3 *Continuity.* If $f, g, h \in \mathcal{F}$, the sets $\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succsim h\}$ and $\{\alpha \in [0,1] : h \succsim \alpha f + (1-\alpha)g\}$ are closed.

A.4 *Monotonicity.* If $f, g \in \mathcal{F}$ and $f(s) \succsim g(s)$ for all $s \in S$, then $f \succsim g$.

A.5 *Uncertainty Aversion.* If $f, g \in \mathcal{F}$ and $\alpha \in (0,1)$,

$$f \sim g \Rightarrow \alpha f + (1-\alpha)\, g \succsim f.$$

A.6 *Non-degeneracy.* $f \succ g$ for some $f, g \in \mathcal{F}$.

Axioms A.1, A.3, A.4 and A.6 are standard assumptions. Axioms A.3 and A.6 are technical assumptions, while A.1 and A.4 require preferences to be transitive, complete, and monotone. The latter requirement is basically a state-independence condition, saying that decision makers always (weakly) prefer acts delivering statewise (weakly) better payoffs, regardless of the state where the better payoffs occur. If a preference relation $\succsim$ satisfies A.1, A.3, and A.4, then each act $f \in \mathcal{F}$ admits a certainty equivalent $x_f \in X$.[3]

Axioms A.2 and A.5 are due to Gilboa and Schmeidler [15]. Axiom A.5 is a smoothing axiom that can be interpreted as an ambiguity aversion axiom, as discussed at length in [15], [24], [12], and [14]. Notice that A.5 is a natural axiom in the malevolent Nature interpretation, as agents may well prefer to randomize over indifferent acts in order to make their behavior less predictable for Nature.

As to A.2, it is a weak independence axiom, which requires independence only with respect to mixing with constant acts, provided the mixing weights are kept constant. It is here in a weaker form than the original axiom of [15], and it is this weakening that makes it possible to go beyond the multiple priors model. Because of its importance for our derivation, we devote to A.2 the rest of this subsection.

Consider the following stronger version of A.2.

A.2' *Certainty Independence.* If $f, g \in \mathcal{F}$, $x \in X$, and $\alpha \in (0,1)$, then

$$f \succsim g \Leftrightarrow \alpha f + (1-\alpha)x \succsim \alpha g + (1-\alpha)\, x.$$

Axiom A.2' is the original axiom of [15]. The next lemma shows how it strenghtens A.2.

[3] See the proof of Lemma 31 in Appendix.

Lemma 1 *A binary relation $\succsim$ on $\mathcal{F}$ satisfies A.2' if and only if, for all $f, g \in \mathcal{F}$, $x, y \in X$, and $\alpha, \beta \in (0,1]$, we have:*

$$\alpha f + (1-\alpha)x \succsim \alpha g + (1-\alpha)x \Rightarrow \beta f + (1-\beta)y \succsim \beta g + (1-\beta)y.$$

Axiom A.2 is therefore the special case of A.2' in which the mixing coefficients α and β are required to be equal. At a conceptual level, Lemma 1 shows that [15]'s Certainty Independence Axiom actually involves two types of independence: independence relative to mixing with constants and independence relative to the weights used in such mixing. Our Axiom A.2 retains the first form of independence, but not the second one. In other words, we allow for preference reversals in mixing with constants unless the weights themselves are kept constant.

This is a significant weakening of the Certainty Independence Axiom and its motivation is best seen when the weights α and β are very different, say α is close to 1 and β is close to 0. Intuitively, acts $\alpha f + (1-\alpha)x$ and $\alpha g + (1-\alpha)x$ can then involve far more uncertainty than acts $\beta f + (1-\beta)y$ and $\beta g + (1-\beta)y$, which are almost constant acts. As a result, we expect that, at least in some situations, the ranking between the genuinely uncertain acts $\alpha f + (1-\alpha)x$ and $\alpha g + (1-\alpha)x$ can well differ from that between the almost constant acts $\beta f + (1-\beta)y$ and $\beta g + (1-\beta)y$.

Needless to say, even though we believe that such reversals can well occur (from both a positive and normative standpoint), the only way to test them, and so the plausibility of A.2 and A.2', is by running experiments. This is possible since both A.2 and A.2' have clear behavioral implications, and this will be the subject of future research. In terms of preference functionals, observe that by Theorems 2 and 12 all preference functionals (3) satisfy A.2 and violate A.2', unless they reduce to the multiple priors form (2).

3.2 Main Result

We can now state our main result, which characterizes preferences satisfying axioms A.1-A.6. Here $\Delta = \Delta(\Sigma)$ denotes the set of all finitely additive probabilities on Σ endowed with the weak* topology,[4] and $c : \Delta \to [0, \infty]$ is said to be *grounded* if its infimum value is zero.

Theorem 2 *Let $\succsim$ be a binary relation on $\mathcal{F}$. The following conditions are equivalent:*

(i) $\succsim$ satisfies conditions A.1-A.6;

[4] That is, the $\sigma(\Delta(\Sigma), B_0(\Sigma))$-topology where a net $\{p_d\}_{d \in D}$ converges to p if and only if $p_d(A) \to p(A)$ for all $A \in \Sigma$.

(ii) *there exists a non-constant affine function* $u : X \to \mathbb{R}$ *and a grounded, convex, and lower semicontinuous function* $c : \Delta \to [0, \infty]$ *such that, for all* $f, g \in \mathcal{F}$

$$f \succsim g \Leftrightarrow \min_{p \in \Delta} \left(\int u(f)\, dp + c(p) \right) \geq \min_{p \in \Delta} \left(\int u(g)\, dp + c(p) \right). \tag{6}$$

For each u *there is a (unique) minimal* $c^\star : \Delta \to [0, \infty]$ *satisfying (6), given by*

$$c^\star(p) = \sup_{f \in \mathcal{F}} \left(u(x_f) - \int u(f)\, dp \right). \tag{7}$$

The variational flavor of the representation (6) motivates the following definition.

Definition 3 *A preference* $\succsim$ *on* $\mathcal{F}$ *is called* variational *if it satisfies axioms A.1-A.6.*

By Theorem 2, variational preferences can be represented by a pair $(u, c^\star)$. From now on, when we consider a variational preference we will write u and $c^\star$ to denote the elements of such a pair. Next we give the uniqueness properties of this representation.

Corollary 4 *Two pairs* $(u_0, c_0^\star)$ *and* $(u, c^\star)$ *represent the same variational preference* $\succsim$ *as in Theorem 2 if and only if there exist* $\alpha > 0$ *and* $\beta \in \mathbb{R}$ *such that* $u = \alpha u_0 + \beta$ *and* $c^\star = \alpha c_0^\star$.

In Theorem 2 we saw that $c^\star$ is the minimal non-negative function on Δ for which the representation (6) holds. More is true when $u(X) = \{u(x) : x \in X\}$ is unbounded (either below or above):

Proposition 5 *Let* $\succsim$ *be a variational preference with* $u(X)$ *unbounded. Then, the function* $c^\star$ *defined in (7) is the unique non-negative, grounded, convex, and lower semicontinuous function on* Δ *for which (6) holds.*

As shown in Lemma 32 in the Appendix, the assumption that $u(X)$ is unbounded is equivalent to the following axiom (see [20]).

A.7 *Unboundedness.* There exist $x \succ y$ in X such that for all $\alpha \in (0, 1)$ there exists $z \in X$ satisfying either $y \succ \alpha z + (1 - \alpha) x$ or $\alpha z + (1 - \alpha) y \succ x$.

We call *unbounded* the variational preferences satisfying axiom A.7.

3.3 Ambiguity Attitudes

We now study the ambiguity attitudes featured by variational preferences. We follow the approach proposed in [14], to which we refer for a detailed discussion of the notions we use.

Begin with a comparative notion: given two preferences $\succsim_1$ and $\succsim_2$, say that $\succsim_1$ is *more ambiguity averse than* $\succsim_2$ if, for all $f \in \mathcal{F}$ and $x \in X$,

$$f \succsim_1 x \Rightarrow f \succsim_2 x. \tag{8}$$

To introduce an absolute notion of ambiguity aversion, as in [14] we consider Subjective Expected Utility (SEU) preferences as benchmarks for ambiguity neutrality. We then say that a preference relation $\succsim$ is *ambiguity averse* if it is more ambiguity averse than some SEU preference.

We now apply these notions to our setting. The first thing to observe is that variational preferences are always ambiguity averse.

Proposition 6 *Each variational preference is ambiguity averse.*

As variational preferences satisfy axiom A.5, and the choice rule resulting from (6) is a max min rule, intuitively it is not surprising that variational preferences always display a negative attitude toward ambiguity. Proposition 6 makes precise this intuition.

Next we show that comparative ambiguity attitudes for variational preferences are determined by the function $c^\star$. Here $u_1 \approx u_2$ means that there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u_1 = \alpha u_2 + \beta$.

Proposition 7 *Given two variational preferences $\succsim_1$ and $\succsim_2$, the following conditions are equivalent:*

(i) $\succsim_1$ is more ambiguity averse than $\succsim_2$,

(ii) $u_1 \approx u_2$ and $c_1^\star \leq c_2^\star$ (provided $u_1 = u_2$).

Given that $u_1 \approx u_2$, the assumption $u_1 = u_2$ is just a common normalization of the two utility indices. Therefore, Proposition 7 says that more ambiguity averse preference relations are characterized, up to a normalization, by smaller functions $c^\star$. Therefore, the function $c^\star$ can be interpreted as an *index of ambiguity aversion.*

We now give few simple examples illustrating this interpretation of the function $c^\star$.

Example 8 By Proposition 7, maximal ambiguity aversion is characterized by $c^\star(p) = 0$ for each $p \in \Delta$. In this case, (6) becomes:

$$f \succsim g \Leftrightarrow \min_{p \in \Delta} \int u(f)\, dp \geq \min_{p \in \Delta} \int u(g)\, dp,$$

that is,

$$f \succsim g \Leftrightarrow \min_{s\in S} u\left(f\left(s\right)\right) \geq \min_{s\in S} u\left(g\left(s\right)\right),$$

a form that clearly reflects extreme ambiguity aversion (i.e., a paranoid attitude in the malevolent Nature interpretation). ▲

Example 9 Minimal ambiguity aversion corresponds here to ambiguity neutrality as, by Proposition 6, all variational preferences are ambiguity averse. Therefore, the least ambiguity averse functions $c^\star$ are those associated with SEU preferences. As it will be shown in Corollary 13, when a SEU preference is unbounded, $c^\star$ takes the stark form

$$c^\star\left(p\right) = \begin{cases} 0 & \text{if } p = q, \\ \infty & \text{otherwise,} \end{cases}$$

where q is the subjective probability associated with the preference. ▲

Example 10 Denote by $c^\star_q$ the ambiguity index of an unbounded SEU preference with subjective probability q, and by $c^\star_m$ the maximal ambiguity index of Example 8. Given $\alpha \in (0,1)$, suppose the r.h.s. of (6) is:

$$(1-\alpha)\int u\left(f\right) dq + \alpha \min_{s\in S} u\left(f\left(s\right)\right) \geq (1-\alpha)\int u\left(g\right) dq + \alpha \min_{s\in S} u\left(g\left(s\right)\right),$$

which is the well-known ε-contaminated model. In this case,

$$c^\star\left(p\right) = \min_{p_1,p_2\in\Delta}\left\{(1-\alpha)\, c^\star_q\left(p_2\right) + \alpha c^\star_m\left(p_1\right) : (1-\alpha)\, p_2 + \alpha p_1 = p\right\} = \delta_{(1-\alpha)q+\alpha\Delta}\left(p\right),$$

and this is a simple example of an index $c^\star$ not displaying extreme ambiguity attitudes. ▲

We close with a couple of remarks. First, observe that Lemma 34 in the Appendix shows that the set that [14] calls benchmark measures – those probabilities that correspond to SEU preferences less ambiguity averse than $\succsim$ – is given here by $\arg\min c^\star = \{p \in \Delta : c^\star(p) = 0\}$.

Second, notice that by standard convex analysis results (see [23]), the last example can be immediately generalized as follows: the ambiguity index of a convex combination of preference functionals representing unbounded variational preferences is given by the infimal convolution of their ambiguity indices.

3.4 An Extension: Countable Additivity

In Theorem 2 we considered the set Δ of all finitely additive probabilities. In applications, however, it is often important to consider countably additive probabilities, which have very convenient analytical properties. For example, in Section 4 it will be seen that this is the case for the multiplier preferences of Hansen and Sargent [18] and for mean-variance preferences of Markowitz [21] and Tobin [25].

Fortunately, in our setting we can still use the Monotone Continuity axiom introduced by Arrow [3] in order to derive a SEU representation with a countably additive subjective probability (see [6]).

A.8 *Monotone Continuity.* If $f, g \in \mathcal{F}$, $x \in X$, $\{E_n\}_{n\geq 1} \in \Sigma$ with $E_1 \supseteq E_2 \supseteq ...$ and $\bigcap_{n\geq 1} E_n = \emptyset$, then $f \succ g$ implies that there exists $n_0 \geq 1$ such that

$$xE_{n_0} f \succ g.$$

Next we state the countably additive version of Theorem 2. Here $\Delta^\sigma = \Delta^\sigma(\Sigma)$ denotes the set of all countably additive probabilities defined on a σ-algebra Σ, while $\Delta^\sigma(q) = \Delta^\sigma(\Sigma, q)$ denotes the subset of Δ^σ consisting of all probabilities that are absolutely continuous with respect to q; i.e., $\Delta^\sigma(q) = \{p \in \Delta^\sigma : p \ll q\}$.

Theorem 11 *Let $\succsim$ be an unbounded variational preference. The following conditions are equivalent:*

(i) $\succsim$ satisfies A.8;

(ii) $\{p \in \Delta : c^\star(p) \leq t\}$ is a (weakly) compact subset of Δ^σ for each $t \geq 0$.

In this case, there exists $q \in \Delta^\sigma$ such that, for all $f, g \in \mathcal{F}$,

$$f \succsim g \Leftrightarrow \min_{p\in\Delta^\sigma(q)} \left(\int u(f)\, dp + c^\star(p) \right) \geq \min_{p\in\Delta^\sigma(q)} \left(\int u(g)\, dp + c^\star(p) \right). \tag{9}$$

Lemma 33 of the Appendix shows that even when the preference is not unbounded, axiom A.8 still implies the countable additivity of the probabilities involved in the representation.

In view of these results, we call *continuous* the variational preferences satisfying axiom A.8.

4 Special Cases

In this section we show that three important classes of preferences, the multiple priors preferences of Gilboa and Schmeidler [15], the multiplier preferences of Hansen and Sargent [18], and the mean-variance preferences of Markowitz [21] and Tobin [25], are special cases of our variational preferences.

4.1 Multiple Priors Preferences

Begin with the multiple priors choice model axiomatized by Gilboa and Schmeidler [15]. As we have mentioned in Section 3.1, the multiple priors model is characterized by axiom A.2', a stronger version of our independence axiom A.2. Next we show that multiple priors preferences are indeed the special class of variational preferences characterized by A.2'.

In other words, when A.2' replaces A.2, the only probabilities in Δ that "matter" in the representation (6) are the ones to which the decision maker attributes "maximum weight", i.e., the ones in $\arg\min c^{\star}$. The set of priors C used in the multiple priors model is then given by $\{p \in \Delta : c^{\star}(p) = 0\}$.

Proposition 12 *Let $\succsim$ be a variational preference. The following conditions are equivalent:*

(i) $\succsim$ satisfies A.2';

(ii) for all $f \in \mathcal{F}$,

$$\min_{p\in\Delta}\left(\int u(f)\,dp + c^{\star}(p)\right) = \min_{\{p\in\Delta : c^{\star}(p)=0\}} \int u(f)\,dp. \tag{10}$$

Moreover, whenever $\succsim$ is unbounded, (ii) is also equivalent to:

(iii) $c^{\star}$ only takes on values 0 and ∞.

The characterization of the multiple priors model via axioms A.1, A.2', and A.3-A.6 is due to Gilboa and Schmeidler [15]. Proposition 12 shows how the multiple priors model fits in the representation we established in Theorem 2.

As well-known, the standard SEU model is the special case of the multiple priors model characterized by the following stronger version of A.5.

A.5' *Uncertainty Neutrality.* If $f, g \in \mathcal{F}$ and $\alpha \in (0,1)$,

$$f \sim g \Rightarrow \alpha f + (1-\alpha) g \sim f.$$

In terms of our representation, by Theorem 2 and Proposition 12 we have:

Corollary 13 *Let $\succsim$ be a variational preference. The following conditions are equivalent:*

(i) $\succsim$ satisfies A.5';

(ii) $\succsim$ is SEU;

(iii) $\succsim$ satisfies A.2' and $\{p \in \Delta : c^{\star}(p) = 0\}$ is a singleton.

Moreover, whenever $\succsim$ is unbounded, (iii) is also equivalent to:

(iv) there exists $q \in \Delta$ such that $c^{\star}(q) = 0$ and $c^{\star}(p) = \infty$ for every $p \neq q$.

4.2 Multiplier Preferences

In the robust control literature proposed by Hansen and Sargent (see, e.g., [17] and [18]) acts are ranked according to the so-called *multiplier preferences* $\succsim^m$ on $\mathcal{F}$, represented as follows:

$$f \succsim^m g \Leftrightarrow \min_{p \in \Delta^\sigma(q)} \left(\int u(f)\, dp + \theta R(p \| q) \right) \geq \min_{p \in \Delta^\sigma(q)} \left(\int u(g)\, dp + \theta R(p \| q) \right), \quad (11)$$

where $\theta > 0$, $q \in \Delta^\sigma$, $u : X \to \mathbb{R}$ is an affine function, and $R(\cdot \| q) : \Delta \to \mathbb{R}$ is the relative entropy given by:

$$R(p \| q) = \begin{cases} \int \frac{dp}{dq} \left(\log \frac{dp}{dq} \right) dq & \text{if } p \in \Delta^\sigma(q), \\ \infty & \text{otherwise.} \end{cases}$$

Notice that the relative entropy $R(\cdot \| q)$ is non-negative and convex on Δ, strictly convex on its effective domain, with $R(p \| q) = 0$ if and only if $p = q$.

The next result – a version of the Donsker-Varadhan variational formula (see, e.g., [10, p. 36]) – shows that multiplier preferences are continuous variational preferences with $c^\star(p) = \theta R(p \| q)$. In other words, the representation (11) is a special case of (9), which is in turn the countably additive version of (6).

Theorem 14 *Suppose $u(X)$ is unbounded. Then, the multiplier preference $\succsim^m$ is a continuous variational preference. Its index of ambiguity aversion is*

$$c^\star(p) = \theta R(p \| q)$$

for each $p \in \Delta^\sigma(q)$.

In view of Theorem 14, in order to determine the ambiguity attitudes featured by multiplier preferences we can invoke Propositions 6 and 7. By the former result, multiplier preferences are ambiguity averse. As to comparative attitudes, the next simple consequence of Proposition 7 shows that they only depend on the parameter θ, which can therefore be interpreted as a coefficient of ambiguity aversion.

Corollary 15 *Given two multiplier preferences $\succsim_1^m$ and $\succsim_2^m$, the following conditions are equivalent:*

(i) $\succsim_1^m$ is more ambiguity averse than $\succsim_2^m$,

(ii) $u_1 \approx u_2$ and $\theta_1 \leq \theta_2$ (provided $u_1 = u_2$).

In view of Corollary 15, we have maximal ambiguity aversion when the parameter θ goes to 0. The next result shows what is the limit behavior, as θ goes to 0, of the preference functional that represents multiplier preferences in (11). In order to do so, we need a piece of notation. Given $q \in \Delta^\sigma$ and a measurable function $\varphi : S \to \mathbb{R}$, set

$$\operatorname*{ess\,inf}_{s\in S} \varphi(s) = \sup\{t \in \mathbb{R} : q(\{s \in S : \varphi(s) \geq t\}) = 1\}.$$

If φ is bounded below the sup is attained. For example, when q has a finite support $\operatorname{supp}(q)$, we have

$$\operatorname*{ess\,inf}_{s\in S} \varphi(s) = \min_{s\in \operatorname{supp}(q)} \varphi(s). \tag{12}$$

Proposition 16 *For all $f \in \mathcal{F}$,*

$$\lim_{\theta\downarrow 0} \min_{p\in\Delta^\sigma(q)} \left(\int u(f)\,dp + \theta R(p\|q) \right) = \operatorname*{ess\,inf}_{s\in S} u(f(s)). \tag{13}$$

By (13), as θ goes to 0, multiplier preferences tend more and more to rank acts according to the very cautious criterion given by $\operatorname{ess\,inf}_{s\in S} u(f(s))$.

We have shown how multiplier preferences are a special case of our variational preferences. Unlike multiple priors preferences, for multiplier preferences we do not have an additional axiom that on top of axioms A.1-A.6 would deliver them (for multiple priors preferences the needed extra axiom was A.2'). On the other hand, we view multiplier preferences as essentially an analytically convenient specification of variational preferences, much in the same way as, for example, Cobb-Douglas preferences are an analytically convenient specification of homothetic preferences. As a result, we do not expect that in our setting there exist behaviorally significant axioms that would characterize multiplier preferences (as we are not aware of any behaviorally significant axiom characterizing Cobb-Douglas preferences).

We close by discussing the related work of Wang [26]. In a quite different setting [26] recently proposed an axiomatization of a class of preferences that include multiplier preferences as special cases. He considers preferences over triplets (f, C, q), where f is a payoff profile, q is a reference probability, and $C \subseteq \Delta$ is a confidence region. For such preferences he axiomatizes the following representation:

$$V(f, C, q) = \min_{p\in C} \left\{ \int u(f)\,dp + \theta R(p\|q) \right\}. \tag{14}$$

His modelling is very different from ours: in our setup preferences are defined only on acts, and we derive simultaneously both the utility index u and the ambiguity index c; that is, uncertainty is subjective. In [26], both C and q are exogenous, and so uncertainty is objective; moreover, agents' preferences are defined on the significantly larger set of all possible triplets consisting of payoff profile, confidence region, and reference model.

4.3 Mean-Variance Preferences

In this subsection we will show that also mean-variance preferences can be viewed as variational preferences, with index of ambiguity aversion given by the relative Gini concentration index. In this way, after Shannon's entropy, a second classic concentration index pops up in the analysis of variational preferences.

Consider the ordering $\succsim^{mv}$ on $\mathcal{F}$ given by:

$$f \succsim^{mv} g \Leftrightarrow \int u(f)\,dq - \frac{\theta}{2}\mathrm{Var}(u(f)) \geq \int u(g)\,dq - \frac{\theta}{2}\mathrm{Var}(u(g)), \tag{15}$$

where $\theta > 0$, $q \in \Delta^\sigma$, $u : X \to \mathbb{R}$ is an affine function, and Var is the variance with respect to q. When $X \subseteq \mathbb{R}$ and $u(x) = x$ (i.e., when u is risk neutral), $\succsim^{mv}$ reduces to the classic mean-variance ordering

$$f \succsim^{mv} g \Leftrightarrow \int f dq - \frac{\theta}{2}\mathrm{Var}(f) \geq \int g dq - \frac{\theta}{2}\mathrm{Var}(g),$$

of Markowitz [21] and Tobin [25].

The next result shows that the monotone restriction of $\succsim^{mv}$ is a continuous variational preference. Interestingly, its index of ambiguity aversion $c^\star$ turns out to be proportional to the relative Gini concentration index $G(\cdot\|q) : \Delta \to \mathbb{R}$ given by:

$$G(p\|q) = \begin{cases} \displaystyle\int \left(\frac{dp}{dq}\right)^2 dq - 1 & \text{if } p \in \Delta^\sigma(q), \\ \infty & \text{otherwise.} \end{cases}$$

The classic Gini concentration index can be obtained by normalization from the relative one in the same way Shannon's entropy can be obtained from relative entropy. $G(\cdot\|q)$ has properties similar to the relative entropy $R(\cdot\|q)$; in particular, $G(\cdot\|q)$ is non-negative and convex on Δ, strictly convex on its effective domain, with $G(p\|q) = 0$ if and only if $p = q$.

Before stating the announced result, observe that $\succsim^{mv}$ may not be monotone, unless its domain is suitably restricted. To identify such restriction, consider the concave functional $J : B_0(\Sigma) \to \mathbb{R}$ given by $J(\varphi) = \int \varphi dq - (\theta/2)\mathrm{Var}(\varphi)$, and let M be the set in which the Gateaux differential of J is positive (as a linear functional). The convex set M is the domain of monotonicity of the functional J, and so the convex set $\mathcal{G} = \{f \in \mathcal{F} : u(f) \in M\}$ is where the preference $\succsim^{mv}$ does not violate monotonicity.

Theorem 17 *Suppose $u(X)$ is unbounded. Then, the mean-variance preference $\succsim^{mv}$ is a continuous variational preference on $\mathcal{G}$.*[5] *Its index of ambiguity aversion is*

$$c^\star(p) = \frac{1}{2\theta} G(p\|q)$$

[5] Until now we only considered preferences defined on the whole $\mathcal{F}$; then we say that a preference on $\mathcal{G}$ is a continuous variational preference if it admits a continuous variational extension on the entire space $\mathcal{F}$. Again, $c^\star$ is the minimal non-negative function on Δ satisfying (6) for all $f, g \in \mathcal{G}$.

for each $p \in \Delta^{\sigma}(q)$*. In particular, for all* $f, g \in \mathcal{G}$*,*

$$f \succsim^{mv} g \Leftrightarrow \min_{p \in \Delta^{\sigma}(q)} \left(\int u(f)\, dp + \frac{1}{2\theta} G(p \| q) \right) \geq \min_{p \in \Delta^{\sigma}(q)} \left(\int u(g)\, dp + \frac{1}{2\theta} G(p \| q) \right).$$

As we did earlier for multiplier preferences, we can now study the ambiguity attitudes featured by mean-variance preferences. These preferences are ambiguity averse, and their comparative ambiguity attitudes only depend on the parameter θ; again, θ can be interpreted as a coefficient of ambiguity aversion. Specifically, we have:

Corollary 18 *Given two mean-variance preferences* $\succsim_1^{mv}$ *and* $\succsim_2^{mv}$*, the following conditions are equivalent:*

(i) $\succsim_1^{mv}$ *is more ambiguity averse than* $\succsim_2^{mv}$*,*

(ii) $u_1 \approx u_2$ *and* $\theta_1 \geq \theta_2$ *(provided* $u_1 = u_2$*).*

All this provides an ambiguity perspective on mean-variance preferences. As we already discussed in the Introduction, the motivation for this interpretation is similar to the one underlying the ambiguity perspective on multiplier preferences. Under this interpretation we can view the decision maker as considering the base probability q as an approximation of the correct probabilistic model governing the phenomenon in which he is interested. He thus takes into account other possible alternative models p, whose relative likelihood is determined by the relative Gini index. The parameter θ measures the decision maker's confidence in the correctness of q and in the relevance of the alternatives p.

Mean-variance preferences can be easily generalized by replacing the variance either with general indices of variability, which describe how dispersed is a random variable relative to some reference point (the mean in the case of the variance) or with some indices of dispersion, which describe the "intrinsic" dispersion of a random variable (e.g., the Gini mean difference index). In Appendix D we briefly discuss these possible extensions of mean-variance preferences.

A Niveloids

Let $B_0(\Sigma)$ be the set of all real-valued Σ-measurable simple functions and $B(\Sigma)$ its supnorm closure (in the space of all real-valued bounded functions on S). $B_0(\Sigma, K)$ (resp. $B(\Sigma, K)$) is the set of all functions in $B_0(\Sigma)$ (resp. $B(\Sigma)$) taking values in the interval $K \subseteq \mathbb{R}$.

When endowed with the supnorm, $B_0(\Sigma)$ is a normed vector space and $B(\Sigma)$ is a Banach space. The norm dual of $B_0(\Sigma)$ (resp. $B(\Sigma)$) is the space $ba(\Sigma)$ of all bounded and finitely additive set functions $\mu : \Sigma \to \mathbb{R}$ endowed with the total variation norm, the duality being

$$\langle \varphi, \mu \rangle = \int \varphi d\mu$$

for all $\varphi \in B_0(\Sigma)$ (resp. $B(\Sigma)$) and all $\mu \in ba(\Sigma)$ (see, e.g., [9, p. 258]). As it is well known, on $\Delta(\Sigma)$ the $\sigma(ba(\Sigma), B_0(\Sigma))$-topology coincides with the $\sigma(ba(\Sigma), B(\Sigma))$-topology, and they go under the name of *weak* topology*; moreover a subset of $\Delta^{\sigma}(\Sigma)$ is weakly* compact iff it is weakly compact (i.e. compact in the weak topology of the Banach space $ba(\Sigma)$).

For $\varphi, \psi \in B(\Sigma)$ we write $\varphi \geq \psi$ (resp. $\varphi > \psi$) if $\varphi(s) \geq \psi(s)$ (resp. $\varphi(s) > \psi(s)$) for all $s \in S$.

Let Φ be any nonempty collection of elements of $B(\Sigma)$, and Φ_c the constant functions in Φ.[6] We call Φ a *tube* if $\Phi = \Phi + \mathbb{R}$.[7]

Given a functional $I : \Phi \to \mathbb{R}$, we say that I is:

(i) *normalized* if $I(k) = k$ for all $k \in \Phi_c$;

(ii) *monotonic* if $\varphi \geq \psi$ implies $I(\varphi) \geq I(\psi)$ for all $\varphi, \psi \in \Phi$;

(iii) *vertically invariant* if $I(\varphi + c) = I(\varphi) + c$ for all $\varphi \in \Phi$ and $c \in \mathbb{R}$ such that $\varphi + c \in \Phi$;

(iv) a *niveloid* if $I(\varphi) - I(\psi) \leq \sup(\varphi - \psi)$ for all $\varphi, \psi \in \Phi$.[8]

Remark 1 Notice that I is a niveloid iff $I(\psi) - I(\varphi) \geq -\sup(\varphi - \psi) = \inf(\psi - \varphi)$ for all $\varphi, \psi \in \Phi$ iff $\inf(\psi - \varphi) \leq I(\psi) - I(\varphi) \leq \sup(\psi - \varphi)$ for all $\psi, \varphi \in \Phi$. Clearly a niveloid is Lipschitz continuous of rank 1 in the supnorm ($\sup(\varphi - \psi) \leq \|\varphi - \psi\|$).

Remark 2 For all $I : \Phi \to \mathbb{R}$, define $\bar{I} : -\Phi \to \mathbb{R}$ by $\bar{I}(\varphi) = -I(-\varphi)$. It is easy to check that: $\overline{(\bar{I})} = I$; I is normalized iff $\bar{I}$ is normalized; I is monotonic iff $\bar{I}$ is monotonic; I is vertically invariant iff $\bar{I}$ is vertically invariant; I is a niveloid iff $\bar{I}$ is a niveloid.

[6] As usual, we write k both for the real number k and for the constant function $k1_S \in B_0(\Sigma)$.

[7] Clearly, if Φ is not a tube, then $\Phi + \mathbb{R}$ is the smallest tube containing Φ.

[8] Dolecki and Greco [8] call *niveloid* a monotonic and vertically invariant functional $T : [-\infty, \infty]^S \to [-\infty, \infty]$. Their Corollary 1.3 and our Lemma 22 explain why we chose to abuse this term.

A.1 Vertically invariant functionals

Next Lemma provides a useful condition for vertical invariance.

Lemma 19 *Let Φ be a convex subset of $B_0(\Sigma)$ (or $B(\Sigma)$) with $0 \in \Phi$ and $I : \Phi \to \mathbb{R}$ be a functional that satisfies*

$$I(\alpha\varphi + (1-\alpha)k) = I(\alpha\varphi) + (1-\alpha)k \tag{16}$$

for all $\varphi \in \Phi$, $k \in \Phi_c$, and $\alpha \in (0,1)$. Then I is vertically invariant provided one of the following conditions holds:

- *Φ is open,*
- *I is continuous and $0 \in \operatorname{int}(\Phi)$,*
- *$\Phi = B_0(\Sigma, K)$ for some interval $K \subseteq \mathbb{R}$ such that $0 \in \operatorname{int}(K)$.*

Proof. If $c = 0$ then $I(\varphi + c) = I(\varphi) + c$ for all $\varphi \in \Phi$. It is sufficient to prove that $I(\varphi + c) = I(\varphi) + c$ for all $\varphi \in \Phi$ and $c > 0$ such that $\varphi + c \in \Phi$.[9]

Let $\varphi, \varphi + c \in \Phi$ and $c > 0$.

Step 1. If $\varphi, \varphi + c \in \operatorname{int}(\Phi)$, then $I(\varphi + c) = I(\varphi) + c$.

There exists $\alpha \in (0,1)$ such that $\varphi/\alpha, (\varphi + c)/\alpha \in \operatorname{int}(\Phi)$. Hence $(\varphi + t)/\alpha \in \operatorname{int}(\Phi)$ for each $t \in [0, c]$. In fact, there exists $\gamma \in [0,1]$ such that $t = \gamma c$ and

$$\frac{\varphi + t}{\alpha} = \frac{\varphi + \gamma c}{\alpha} = \gamma\frac{\varphi + c}{\alpha} + (1-\gamma)\frac{\varphi}{\alpha} \in \operatorname{int}(\Phi).$$

Choose $n \geq 2$ such that $\frac{c/n}{1-\alpha} \in \Phi_c$.[10] Then

$$\begin{aligned}
I(\varphi + c) &= I\left(\varphi + \frac{c}{n} + \dots + \frac{c}{n}\right) \\
&= I\left(\varphi + \frac{c(n-1)}{n} + \frac{c}{n}\right) \\
&= I\left(\alpha\left(\frac{\left(\varphi + \frac{c(n-1)}{n}\right)}{\alpha}\right) + (1-\alpha)\frac{c/n}{1-\alpha}\right) \\
&= I\left(\alpha\left(\frac{\left(\varphi + \frac{c(n-1)}{n}\right)}{\alpha}\right)\right) + (1-\alpha)\frac{c/n}{1-\alpha} \\
&= I\left(\varphi + \frac{c(n-1)}{n}\right) + \frac{c}{n}
\end{aligned}$$

[9] If $c < 0$, set $\psi = \varphi + c$, and $d = -c$. This yields $\psi, \psi + d \in \Phi$ and $d > 0$, then $I(\psi + d) = I(\psi) + d$, that is $I(\varphi) = I(\varphi + c) - c$.

[10] This is possible since in any case $0 \in \operatorname{int}(\Phi)$.

$$\begin{aligned}
&= \cdots \\
&= I\left(\varphi + \frac{c}{n}\right) + \frac{c\,(n-1)}{n} \\
&= I\,(\varphi) + \frac{c}{n} + \frac{c\,(n-1)}{n} \\
&= I\,(\varphi) + c,
\end{aligned}$$

as wanted. □

Step 1 proves the lemma if Φ is open. If I is continuous and $0 \in \operatorname{int}(\Phi)$, since $\varphi, \varphi + c \in \Phi$, then $\left(1 - \frac{1}{n}\right)\varphi$ and $\left(1 - \frac{1}{n}\right)(\varphi + c) \in \operatorname{int}(\Phi)$ for all $n \geq 1$. But, by Step 1, I is vertically invariant on $\operatorname{int}(\Phi)$ and hence

$$\begin{aligned}
I\,(\varphi + c) &= \lim_{n\to\infty} I\left(\left(1 - \frac{1}{n}\right)(\varphi + c)\right) = \lim_{n\to\infty} I\left(\left(1 - \frac{1}{n}\right)\varphi + \left(1 - \frac{1}{n}\right)c\right) \\
&= \lim_{n\to\infty} I\left(\left(1 - \frac{1}{n}\right)\varphi\right) + \left(1 - \frac{1}{n}\right)c = I\,(\varphi) + c.
\end{aligned}$$

It remains to prove the last case when $K = [a, b)$ or $(a, b]$ or $[a, b]$ with $-\infty \leq a < 0 < b \leq \infty$.[11]

Step 2. Assume K contains b (and hence $b < \infty$) and $a < \varphi < \varphi + c \leq b$, then $I(\varphi + c) = I(\varphi) + c$.

Choose $n \geq 2$ such that $b - \frac{c}{n} > 0$ and $\varphi / \left(\frac{b - \frac{c}{n}}{b}\right) > a$. Set $\alpha = \frac{b - \frac{c}{n}}{b} \in (0,1)$. Notice that for all $t \in \left[\frac{c}{n}, \frac{c(n-1)}{n}\right]$

$$\varphi < \varphi + \frac{c}{n} \leq \varphi + t \leq \varphi + \frac{c\,(n-1)}{n} = \varphi + c - \frac{c}{n} \leq b - \frac{c}{n} < b.$$

Divide all the terms by α to obtain

$$a < \varphi/\alpha < \left(\varphi + \frac{c}{n}\right)/\alpha \leq (\varphi + t)/\alpha \leq \left(\varphi + \frac{c\,(n-1)}{n}\right)/\alpha \leq b$$

and hence $(\varphi + t)/\alpha \in B_0(\Sigma, K)$ for each $t \in \left[\frac{c}{n}, \frac{c(n-1)}{n}\right]$. Moreover $\frac{c/n}{1-\alpha} = b \in K$, and

$$\begin{aligned}
I\,(\varphi + c) &= I\left(\varphi + \frac{c\,(n-1)}{n} + \frac{c}{n}\right) \\
&= I\left(\alpha\left(\frac{\left(\varphi + \frac{c(n-1)}{n}\right)}{\alpha}\right) + (1-\alpha)\frac{c/n}{1-\alpha}\right) \\
&= I\left(\alpha\left(\frac{\left(\varphi + \frac{c(n-1)}{n}\right)}{\alpha}\right)\right) + (1-\alpha)\frac{c/n}{1-\alpha} \\
&= I\left(\varphi + \frac{c\,(n-1)}{n}\right) + \frac{c}{n}.
\end{aligned}$$

[11]If $K = (a, b)$, then $B_0(\Sigma, K)$ is open in $B_0(\Sigma)$.

But $a < \varphi < \varphi + \frac{c(n-1)}{n} < b$ implies $\varphi, \varphi + \frac{c(n-1)}{n} \in \operatorname{int}(\Phi)$, and Step 1 guarantees $I\left(\varphi + \frac{c(n-1)}{n}\right) = I(\varphi) + \frac{c(n-1)}{n}$ whence $I(\varphi + c) = I(\varphi) + c$. □

Step 2 concludes the proof if $K = (a, b]$.

Step 3. Assume K contains a (and hence $a > -\infty$) and $a \leq \varphi < \varphi + c < b$, then $I(\varphi + c) = I(\varphi) + c$.

Consider $-K$ and notice that $-b < -\varphi - c < -\varphi \leq -a$, then $\psi = -\varphi - c \in B_0(\Sigma, -K)$, $c > 0$ and $\psi + c = -\varphi \in B_0(\Sigma, -K)$. Moreover, it is immediate to show that $\bar{I}$ satisfies (16), by Step 2

$$\begin{aligned} I(\varphi + c) &= -\bar{I}(-\varphi - c) = -\bar{I}(\psi) \\ &= -\left(\bar{I}(\psi + c) - c\right) = -\bar{I}(-\varphi) + c \\ &= I(\varphi) + c, \end{aligned}$$

as wanted. □

Step 3 concludes the proof if $K = [a, b)$.

If $K = [a, b]$, then $-\infty < a < b < \infty$. If $\varphi, \varphi + c \in B_0(\Sigma, K)$, then $a \leq \varphi < \varphi + \frac{c}{2} < b$ and $a < \varphi + \frac{c}{2} < \varphi + c \leq b$, thus applying Step 2 and Step 3 we obtain

$$I(\varphi + c) = I\left(\varphi + \frac{c}{2}\right) + \frac{c}{2} = I(\varphi) + c.$$

■

Lemma 20 *Let Φ be a convex subset of $B_0(\Sigma)$ (or $B(\Sigma)$) and $I : \Phi \to \mathbb{R}$ a vertically invariant functional that satisfies*

$$I(\alpha\psi + (1 - \alpha)\varphi) \geq I(\varphi) \tag{17}$$

for all $\varphi, \psi \in \Phi$ such that $I(\psi) = I(\varphi)$ and $\alpha \in (0, 1)$. Then Φ is concave provided one of the following conditions holds:

- *I is continuous and* $\operatorname{int}(\Phi)$ *is not empty,*
- *Φ is a tube.*

Proof. Assume I is continuous and $\operatorname{int}(\Phi)$ is not empty. Let $\varphi_0 \in \operatorname{int}(\Phi)$, there exist $\varepsilon > 0$ such that

$$\begin{aligned} N(\varphi_0, \varepsilon) &= \{\psi \in B_0(\Sigma) : \|\varphi_0 - \psi\| \leq \varepsilon\} \\ &= \{\psi \in B_0(\Sigma) : \varphi_0 - \varepsilon \leq \psi \leq \varphi_0 + \varepsilon\} \end{aligned}$$

is contained in $\text{int}(\Phi)$. Moreover - by continuity - there exists $\rho \in \left(0, \frac{1}{3}\right)$ such that $\|\varphi - \varphi_0\| \leq \rho\varepsilon$ implies $|I(\varphi) - I(\varphi_0)| \leq \frac{\varepsilon}{3}$. Then if $\varphi, \psi \in N(\varphi_0, \rho\varepsilon)$, we have

$$|I(\varphi) - I(\psi)| \leq |I(\varphi) - I(\varphi_0)| + |I(\varphi_0) - I(\psi)| \leq \frac{2}{3}\varepsilon$$

and $-\frac{2}{3}\varepsilon \leq I(\varphi) - I(\psi) \leq \frac{2}{3}\varepsilon$. Setting $t = I(\varphi) - I(\psi)$, we get $-\frac{2}{3}\varepsilon \leq t \leq \frac{2}{3}\varepsilon$. Notice that $-\frac{1}{3}\varepsilon \leq -\rho\varepsilon \leq \psi - \varphi_0 \leq \rho\varepsilon \leq \frac{1}{3}\varepsilon$ and $\varphi_0 - \frac{1}{3}\varepsilon \leq \psi \leq \varphi_0 + \frac{1}{3}\varepsilon$. Summing up,

$$\varphi_0 - \varepsilon \leq \psi + t \leq \varphi_0 + \varepsilon$$

and $\psi + t \in \text{int}(\Phi)$. Since $\psi \in \text{int}(\Phi)$ too, then $I(\psi + t) = I(\psi) + t = I(\varphi)$, so that

$$I(\alpha(\psi + t) + (1 - \alpha)\varphi) \geq I(\varphi). \tag{18}$$

Hence,

$$\begin{aligned} I(\varphi) &\leq I(\alpha(\psi + t) + (1 - \alpha)\varphi) = I(\alpha\psi + (1 - \alpha)\varphi + \alpha t) \\ &= I(\alpha\psi + (1 - \alpha)\varphi) + \alpha t \\ &= I(\alpha\psi + (1 - \alpha)\varphi) + \alpha(I(\varphi) - I(\psi)) \end{aligned}$$

and

$$I(\alpha\psi + (1 - \alpha)\varphi) \geq \alpha I(\psi) + (1 - \alpha) I(\varphi). \tag{19}$$

We conclude that I is concave in $N(\varphi_0, \rho\varepsilon)$. As the choice of $N(\varphi_0, \varepsilon)$ was arbitrary, we conclude that I is locally concave on $\text{int}(\Phi)$. A standard result from convex analysis yields concavity on $\text{int}(\Phi)$. Finally, the continuity of I implies its concavity on the whole Φ. This proves the first case. To prove the second, for all $\varphi, \psi \in \Phi$ and $\alpha \in (0, 1)$, set $t = I(\varphi) - I(\psi)$. Since Φ is a tube, $\psi + t \in \Phi$, and $I(\psi + t) = I(\psi) + t = I(\varphi)$. Repeat the argument leading from (18) to (19). ■

Lemma 21 *Let Φ be a nonempty subset of $B(\Sigma)$ and $I : \Phi \to \mathbb{R}$ be a vertically invariant functional. Then there exists a unique vertically invariant functional $\tilde{I} : \Phi + \mathbb{R} \to \mathbb{R}$ extending I to the tube $\Phi + \mathbb{R}$ generated by Φ. Moreover, if Φ is convex and I is concave, then ($\Phi + \mathbb{R}$ is convex and) $\tilde{I}$ is concave.*

Proof. If there exists a vertically invariant functional $\tilde{I} : \Phi + \mathbb{R} \to \mathbb{R}$ extending I on $\Phi + \mathbb{R}$, then for all $\varphi + d \in \Phi + \mathbb{R}$ with $\varphi \in \Phi$ and $d \in \mathbb{R}$, it satisfies

$$\tilde{I}(\varphi + d) = \tilde{I}(\varphi) + d = I(\varphi) + d. \tag{20}$$

In particular it is unique. Next we show that Eq. (20) defines a vertically invariant functional (that obviously extends I). If $\varphi, \psi \in \Phi$, $d, c \in \mathbb{R}$, and $\varphi + d = \psi + c$, then $\varphi = \psi + c - d$. In particular, $\psi \in \Phi$ and $c - d \in \mathbb{R}$ are such that $\psi + (c - d) \in \Phi$, and

$$\begin{aligned} I(\varphi) + d &= I(\psi + c - d) + d \\ &= I(\psi) + c - d + d = I(\psi) + c. \end{aligned}$$

This proves that $\tilde{I}$ is well defined. If $\varphi+d \in \Phi+\mathbb{R}$ (with $\varphi \in \Phi$ and $d \in \mathbb{R}$) and $c \in \mathbb{R}$, then

$$\tilde{I}\left((\varphi+d)+c\right)=\tilde{I}\left(\varphi+d+c\right)=I\left(\varphi\right)+d+c=\tilde{I}\left(\varphi+d\right)+c,$$

that is, $\tilde{I}$ is vertically invariant.

Assume I is concave. Let $\varphi+d, \psi+t \in \Phi+\mathbb{R}$ with $\varphi, \psi \in \Phi$ and $d, t \in \mathbb{R}$ and $\alpha \in (0,1)$. If $\tilde{I}\left(\varphi+d\right)=\tilde{I}\left(\psi+t\right)=c$, then $I\left(\varphi\right)=c-d$ and $I\left(\psi\right)=c-t$. Therefore $I\left(\alpha\varphi+(1-\alpha)\psi\right) \geq \alpha I\left(\varphi\right)+(1-\alpha) I\left(\psi\right)=\alpha\left(c-d\right)+(1-\alpha)\left(c-t\right)$, that is

$$\begin{aligned}\tilde{I}\left(\alpha\left(\varphi+d\right)+(1-\alpha)\left(\psi+t\right)\right) &= \tilde{I}\left(\left(\alpha\varphi+(1-\alpha)\psi\right)+\alpha d+(1-\alpha)t\right)\\ &= I\left(\alpha\varphi+(1-\alpha)\psi\right)+\alpha d+(1-\alpha)t \geq c.\end{aligned}$$

By Lemma 20, since $\Phi+\mathbb{R}$ is a tube, this means that $\tilde{I}$ is concave. ■

A.2 Extensions of niveloids

In this section we obtain some novel results on the extension of niveloids (the first results on this subject appear in [8]).

Lemma 22 *Let Φ be a nonempty subset of $B\left(\Sigma\right)$ and $I:\Phi \to \mathbb{R}$. The following statements are equivalent:*

(i) I is vertically invariant and its unique vertically invariant extension $\tilde{I}$ to $\Phi+\mathbb{R}$ is monotonic.

(ii) I is a niveloid.

In particular, if Φ is a tube, then $I:\Phi \to \mathbb{R}$ is a niveloid iff it is vertically invariant and monotonic (see [8, p. 4]).

Proof. Let I be vertically invariant and $\tilde{I}$ be monotonic. For all $\varphi, \psi \in \Phi$, $\varphi \leq \psi+\sup\left(\varphi-\psi\right)$, but $\varphi, \psi+\sup\left(\varphi-\psi\right) \in \Phi+\mathbb{R}$, then $\tilde{I}\left(\varphi\right) \leq \tilde{I}\left(\psi+\sup\left(\varphi-\psi\right)\right)$ that is

$$I\left(\varphi\right) \leq I\left(\psi\right)+\sup\left(\varphi-\psi\right).$$

Conversely, if I is a niveloid, for all $\varphi \in \Phi$ and $c \in \mathbb{R}$ such that $\varphi+c \in \Phi$

$$c=\inf\left(\left(\varphi+c\right)-\varphi\right) \leq I\left(\varphi+c\right)-I\left(\varphi\right) \leq \sup\left(\left(\varphi+c\right)-\varphi\right)=c$$

that is $I\left(\varphi+c\right)=I\left(\varphi\right)+c$, and I is vertically invariant. Moreover, if $\varphi, \psi \in \Phi$ and $d, t \in \mathbb{R}$ are such that $\psi+t \geq \varphi+d$, then $I\left(\psi\right)-I\left(\varphi\right) \geq \inf\left(\psi-\varphi\right)$ implies

$$\begin{aligned}\tilde{I}\left(\psi+t\right)-\tilde{I}\left(\varphi+d\right) &= I\left(\psi\right)-I\left(\varphi\right)+t-d\\ &\geq \inf\left(\psi-\varphi\right)+t-d\\ &= \inf\left(\left(\psi+t\right)-\left(\varphi+d\right)\right)\\ &\geq 0,\end{aligned}$$

that is $\tilde{I}$ is monotonic. ■

Lemma 23 *A vertically invariant and monotonic functional* $I : B_0(\Sigma, K) \to \mathbb{R}$ *is a niveloid.*

Proof. In view of Lemma 22, we just have to show that $\tilde{I}$ is monotonic. Let $\varphi, \psi \in B_0(\Sigma, K)$ and $d, t, c \in \mathbb{R}$ be such that $\psi + t \geq \varphi + d$. We want to show that $I(\psi) + t \geq I(\varphi) + d$, i.e., that $\psi + c \geq \varphi$ implies $I(\psi) + c \geq I(\varphi)$.

Assume $\sup K = b < \infty$ is not attained. If $c < b - \sup \psi$, then $\varphi \leq \psi + c \leq \sup \psi + c < b$, then $\psi + c \in B_0(\Sigma, K)$ and $I(\varphi) \leq I(\psi + c) = I(\psi) + c$. Else $c \geq b - \sup \psi \geq 0$ and there exists $\varepsilon > 0$ such that $\varphi < b - \varepsilon < b$. A fortiori $c > (b - \varepsilon) - \sup \psi$. There are two subcases:

- $c > (b - \varepsilon) - \inf \psi$, then $I(\psi) + c \geq I(\inf \psi) + c = \tilde{I}(0) + \inf \psi + c > \tilde{I}(0) + \inf \psi + (b - \varepsilon) - \inf \psi \geq \tilde{I}(0) + b - \varepsilon = I(b - \varepsilon) \geq I(\varphi)$.

- $c \leq (b - \varepsilon) - \inf \psi$ (that is $\inf \psi \leq (b - \varepsilon) - c < \sup \psi$), then $\psi + c \geq \varphi$ implies $(\psi + c) \wedge (b - \varepsilon) \geq \varphi$, but $(\psi + c) \wedge (b - \varepsilon) \in B_0(\Sigma, K)$ and $(\psi + c) \wedge (b - \varepsilon) = \min\{\psi + c, b - \varepsilon\} - c + c = \min\{\psi, b - \varepsilon - c\} + c = (\psi \wedge (b - \varepsilon - c)) + c$. Notice that also $\psi \wedge (b - \varepsilon - c) \in B_0(\Sigma, K)$ since $(b - \varepsilon - c) \in [\inf \psi, \sup \psi) \subseteq K$. Therefore

$$\begin{aligned} I(\psi) + c &\geq I(\psi \wedge (b - \varepsilon - c)) + c \\ &= I((\psi \wedge (b - \varepsilon - c)) + c) \\ &= I((\psi + c) \wedge (b - \varepsilon)) \\ &\geq I(\varphi), \end{aligned}$$

  as desired.

Assume that $\sup K = b < \infty$ is attained. If $c \leq b - \sup \psi$, then $\varphi \leq \psi + c \leq \sup \psi + c \leq b$, then $\psi + c \in B_0(\Sigma, K)$ and $I(\varphi) \leq I(\psi + c) = I(\psi) + c$. Else $c > b - \sup \psi \geq 0$ while $\varphi \leq b$. There are two subcases:

- $c \geq b - \inf \psi$, then $I(\psi) + c \geq I(\inf \psi) + c = \tilde{I}(0) + \inf \psi + c \geq \tilde{I}(0) + \inf \psi + b - \inf \psi = I(b) \geq I(\varphi)$.

- $c < b - \inf \psi$ (that is $\inf \psi < b - c < \sup \psi$), then $\psi + c \geq \varphi$ implies $(\psi + c) \wedge b \geq \varphi$, but $(\psi + c) \wedge b \in B_0(\Sigma, K)$ and $(\psi + c) \wedge b = \min\{\psi + c, b\} - c + c = \min\{\psi, b - c\} + c = (\psi \wedge (b - c)) + c$. Notice that also $\psi \wedge (b - c) \in B_0(\Sigma, K)$

since $(b-c) \in (\inf \psi, \sup \psi) \subseteq K$. Therefore

$$\begin{aligned} I(\psi) + c &\geq I(\psi \wedge (b-c)) + c \\ &= I((\psi \wedge (b-c)) + c) \\ &= I((\psi + c) \wedge b) \\ &\geq I(\varphi), \end{aligned}$$

as desired.

Finally, if $\sup K = \infty$, and $\varphi \leq \psi + c$, then $\psi + c \in B_0(\Sigma, K)$ and $I(\varphi) \leq I(\psi + c) = I(\psi) + c$. ■

Lemma 24 *Let $I : \Phi \to \mathbb{R}$ be a niveloid on a nonempty subset Φ of $B(\Sigma)$, and set*

$$\mathcal{L} = \left\{ \varphi \in \Phi + \mathbb{R} : \tilde{I}(\varphi) \geq 0 \right\} + B\left(\Sigma, \mathbb{R}^+\right).$$

The functional defined on $B(\Sigma)$ by

$$\hat{I}(\varphi) = \sup \{c \in \mathbb{R} : \varphi - c \in \mathcal{L}\} \quad \forall \varphi \in B(\Sigma)$$

is the minimum niveloid on $B(\Sigma)$ that extends I. Moreover, if Φ is convex and I is concave, then $\hat{I}$ is concave.

Before entering the proof's details, notice that if I is a niveloid on a tube Φ, then for all $\varphi \in \Phi$, $I(\varphi) = \sup \{c \in \mathbb{R} : c \leq I(\varphi)\} = \sup \{c \in \mathbb{R} : \varphi - c \in \{I \geq 0\}\}$, where $\{I \geq 0\} = \{\varphi \in \Phi : I(\varphi) \geq 0\}$ (see also [13, p. 160]).

Proof. If $\varphi \in \Phi + \mathbb{R}$, and $\varphi \in \mathcal{L}$, then $\varphi \geq \psi$ for some $\psi \in \left\{ \tilde{I} \geq 0 \right\}$, whence $\tilde{I}(\varphi) \geq \tilde{I}(\psi) \geq 0$, that is $\varphi \in \left\{ \tilde{I} \geq 0 \right\}$. This proves that $\varphi \in \Phi + \mathbb{R}$ belongs to $\mathcal{L}$ iff it belongs to $\left\{ \tilde{I} \geq 0 \right\}$. As a consequence, for all $\varphi \in \Phi + \mathbb{R}$, we have

$$\begin{aligned} \tilde{I}(\varphi) &= \sup \left\{ c \in \mathbb{R} : \varphi - c \in \left\{ \tilde{I} \geq 0 \right\} \right\} \\ &= \sup \{c \in \mathbb{R} : \varphi - c \in \mathcal{L}\} \\ &= \hat{I}(\varphi). \end{aligned}$$

Then $\hat{I} : B(\Sigma) \to [-\infty, \infty]$ extends $\tilde{I}$, *a fortiori* I.

Notice that:

- If $\varphi \in \mathcal{L}$ and $\psi \geq \varphi$, then $\psi \in \mathcal{L} + B(\Sigma, \mathbb{R}^+) = \mathcal{L}$.
- If $\varphi \notin \mathcal{L}$ and $\psi \leq \varphi$, then $\psi \notin \mathcal{L}$.

- If $\psi_0 \in \Phi$, then $\psi_0 + d \in \mathcal{L}$ iff $\tilde{I}(\psi_0 + d) \geq 0$ iff $d \geq -I(\psi_0)$. In particular, $\psi_0 - I(\psi_0) \in \mathcal{L}$ and $\psi_0 - I(\psi_0) - 1 \notin \mathcal{L}$.

Let $\psi_0 \in \Phi$. For all $\varphi \in B(\Sigma)$, $\varphi - (\inf \varphi - \sup \psi_0 + I(\psi_0)) \geq \psi_0 - I(\psi_0) \in \mathcal{L}$, hence $\{c \in \mathbb{R} : \varphi - c \in \mathcal{L}\} \neq \emptyset$; therefore $\hat{I}(\varphi) > -\infty$. For all $\varphi \in B(\Sigma)$ and all $c \geq \sup \varphi - \inf \psi_0 + I(\psi_0) + 1$, $\varphi - c \leq \varphi - (\sup \varphi - \inf \psi_0 + I(\psi_0) + 1) \leq \psi_0 - I(\psi_0) - 1 \notin \mathcal{L}$ implies $\varphi - c \notin \mathcal{L}$, and $\{c \in \mathbb{R} : \varphi - c \in \mathcal{L}\}$ is bounded above; therefore $\hat{I}(\varphi) < \infty$. We conclude that $\hat{I} : B(\Sigma) \to \mathbb{R}$.

If $\psi \geq \varphi$ and $\varphi - c \in \mathcal{L}$, then $\psi - c \geq \varphi - c$ implies $\psi - c \in \mathcal{L}$. It follows that $\{c \in \mathbb{R} : \varphi - c \in \mathcal{L}\} \subseteq \{c \in \mathbb{R} : \psi - c \in \mathcal{L}\}$ and $\hat{I}(\psi) \geq \hat{I}(\varphi)$, i.e., $\hat{I}$ is monotonic.

Let $d \in \mathbb{R}$ and $\varphi \in B(\Sigma)$, $\varphi - c \in \mathcal{L}$ iff $(\varphi + d) - (c + d) \in \mathcal{L}$, that is

$$\{c \in \mathbb{R} : \varphi - c \in \mathcal{L}\} + d = \{t \in \mathbb{R} : (\varphi + d) - t \in \mathcal{L}\}$$

and

$$\begin{aligned}
\hat{I}(\varphi + d) &= \sup\{t \in \mathbb{R} : (\varphi + d) - t \in \mathcal{L}\} \\
&= \sup(\{c \in \mathbb{R} : \varphi - c \in \mathcal{L}\} + d) \\
&= \hat{I}(\varphi) + d.
\end{aligned}$$

That is $\hat{I}$ is a niveloid.

Notice that $\left\{\hat{I} \geq 0\right\} = \overline{\mathcal{L}}$. In fact, if $\hat{I}(\varphi) \geq 0$, then for all $\varepsilon > 0$ we have $\hat{I}(\varphi + \varepsilon) > 0$, i.e.

$$\sup\{c \in \mathbb{R} : \varphi + \varepsilon - c \in \mathcal{L}\} > 0,$$

therefore there exists $c > 0$ such that $\varphi + \varepsilon - c \in \mathcal{L}$. This implies $\varphi + \varepsilon \in \mathcal{L}$ (since $\varphi + \varepsilon \geq \varphi + \varepsilon - c \in \mathcal{L}$). Since this is true for all ε, it follows $\varphi = \lim_n \left(\varphi + \frac{1}{n}\right) \in \overline{\mathcal{L}}$, and we conclude $\left\{\hat{I} \geq 0\right\} \subseteq \overline{\mathcal{L}}$. Conversely, for all $\varphi \in \mathcal{L}$, $\varphi - 0 \in \mathcal{L}$ guarantees $\hat{I}(\varphi) \geq 0$; the continuity of $\hat{I}$ implies $\overline{\mathcal{L}} \subseteq \left\{\hat{I} \geq 0\right\}$.

Let $\hat{J}$ be a niveloid on $B(\Sigma)$ that extends I, then $\hat{J}$ coincides with $\tilde{I}$ on $\Phi + \mathbb{R}$. For all $\psi \in \mathcal{L}$ there exists $\varphi \in \left\{\tilde{I} \geq 0\right\}$ such that $\psi \geq \varphi$, therefore $\hat{J}(\psi) \geq \hat{J}(\varphi) = \tilde{I}(\varphi) \geq 0$. Then $\left\{\hat{I} \geq 0\right\} = \overline{\mathcal{L}} \subseteq \left\{\hat{J} \geq 0\right\}$, and this implies that for all $\varphi \in B(\Sigma)$

$$\begin{aligned}
\hat{I}(\varphi) &= \sup\left\{c \in \mathbb{R} : \varphi - c \in \left\{\hat{I} \geq 0\right\}\right\} \\
&\leq \sup\left\{c \in \mathbb{R} : \varphi - c \in \left\{\hat{J} \geq 0\right\}\right\} = \hat{J}(\varphi).
\end{aligned}$$

This shows that $\hat{I}$ is the minimum niveloid on $B(\Sigma)$ that extends I.

Assume Φ is convex and I is concave, then $\tilde{I}$ is concave and $\left\{\tilde{I} \geq 0\right\}$ is convex. So $\mathcal{L} = \left\{\tilde{I} \geq 0\right\} + B(\Sigma, \mathbb{R}^+)$ and $\left\{\hat{I} \geq 0\right\} = \overline{\mathcal{L}}$ are convex. This implies that $\hat{I}$ is concave. In fact, for all $\varphi, \psi \in B(\Sigma)$ such that $\hat{I}(\varphi) = \hat{I}(\psi) = c$, and $\alpha \in (0,1)$, since $\varphi - c, \psi - c \in \left\{\hat{I} \geq 0\right\}$, then

$$\begin{aligned} \hat{I}(\alpha\varphi + (1-\alpha)\psi) - \hat{I}(\varphi) &= \hat{I}(\alpha\varphi + (1-\alpha)\psi - c) \\ &= \hat{I}(\alpha(\varphi - c) + (1-\alpha)(\psi - c)) \geq 0, \end{aligned}$$

and Lemma 20 guarantees concavity. ∎

Inspection of the proof shows that for a non-empty subset Φ of $B_0(\Sigma)$ setting

$$\mathcal{L}_0 = \left\{\varphi \in \Phi : \tilde{I}(\varphi) \geq 0\right\} + B_0\left(\Sigma, \mathbb{R}^+\right)$$

we could obtain the minimum niveloid extending I to $B_0(\Sigma)$.

The following result shows that $\hat{I}$ coincides with the least niveloidal extension of I as in [8, p. 10-11], and with the standard monotonic extension of $\tilde{I}$ as a monotonic functional.[12]

Proposition 25 *For all $\varphi \in B(\Sigma)$,*

$$\hat{I}(\varphi) = \sup_{\psi \in \Phi}\left[I(\psi) + \inf_{s \in S}(\varphi(s) - \psi(s))\right] = \sup\left\{\tilde{I}(\psi) : \psi \in \Phi + \mathbb{R} \text{ and } \psi \leq \varphi\right\}.$$

Proof. If $A \subseteq \mathbb{R}$, then $\sup\{c \in \mathbb{R} : c \leq a \text{ for some } a \in A\} = \sup A$, therefore

$$\begin{aligned} \hat{I}(\varphi) &= \sup\{c \in \mathbb{R} : \varphi - c \geq \psi_0 + d \text{ for some } \psi_0 \in \Phi \text{ and } d \in \mathbb{R} \text{ s.t. } d \geq -I(\psi_0)\} \\ &= \sup\{c \in \mathbb{R} : \varphi - c \geq \psi_0 - I(\psi_0) \text{ for some } \psi_0 \in \Phi\} \\ &= \sup\left\{c \in \mathbb{R} : c \leq I(\psi_0) + \inf_{s \in S}(\varphi(s) - \psi_0(s)) \text{ for some } \psi_0 \in \Phi\right\} \\ &= \sup_{\psi_0 \in \Phi}\left[I(\psi_0) + \inf_{s \in S}(\varphi(s) - \psi_0(s))\right] \\ &= \sup_{\psi_0 \in \Phi}\left[\sup\left\{I(\psi_0) + c : c \in \mathbb{R} \text{ and } c \leq \inf_{s \in S}(\varphi(s) - \psi_0(s))\right\}\right] \\ &= \sup_{\psi_0 \in \Phi}\left[\sup\{I(\psi_0) + c : c \in \mathbb{R} \text{ and } c \leq \varphi - \psi_0\}\right] \\ &= \sup_{\psi_0 \in \Phi}\left[\sup\left\{\tilde{I}(\psi_0 + c) : c \in \mathbb{R} \text{ and } \psi_0 + c \leq \varphi\right\}\right] \\ &= \sup\left\{\tilde{I}(\psi_0 + c) : \psi_0 \in \Phi, c \in \mathbb{R} \text{ and } \psi_0 + c \leq \varphi\right\} \\ &= \sup\left\{\tilde{I}(\psi) : \psi \in \Phi + \mathbb{R} \text{ and } \psi \leq \varphi\right\}. \end{aligned}$$

∎

[12] We thank Larry Epstein for having pointed out to us this characterization.

A.3 Fenchel conjugates of concave niveloids

Remark 3 If $I : B(\Sigma) \to \mathbb{R}$ is a concave niveloid, direct application of the Fenchel Moreau Theorem (see, e.g., [22, p. 42]) guarantees

$$I(\varphi) = \min_{\mu \in ba(\Sigma)} \left(\langle \varphi, \mu \rangle - I^*(\mu) \right)$$

where $I^*(\mu) = \inf_{\psi \in B(\Sigma)} \left(\langle \psi, \mu \rangle - I(\psi) \right)$ is the *Fenchel conjugate* of I. If μ is not positive, there exists $\varphi \geq 0$ such that $\langle \varphi, \mu \rangle < 0$, then $\langle \alpha\varphi, \mu \rangle - I(\alpha\varphi) \leq \alpha \langle \varphi, \mu \rangle - I(0)$ for all $\alpha \geq 0$, whence $I^*(\mu) = -\infty$. If $\mu(S) \neq 1$, choose $\psi \in B(\Sigma)$, then $\langle \psi + c, \mu \rangle - I(\psi + c) = \langle \psi, \mu \rangle - I(\psi) + c(\mu(S) - 1)$ for all $c \in \mathbb{R}$, and so $I^*(\mu) = -\infty$. That is,

$$I(\varphi) = \min_{\mu \in \Delta(\Sigma)} \left(\langle \varphi, \mu \rangle - I^*(\mu) \right).$$

In this section Φ is a (non-empty) convex subset of $B(\Sigma)$ and $I : \Phi \to \mathbb{R}$ is a concave niveloid. We set

$$\partial_\pi I(\varphi) = \{ p \in \Delta(\Sigma) : I(\psi) - I(\varphi) \leq \langle \psi - \varphi, p \rangle \text{ for each } \psi \in \Phi \}.$$

Notice that

$$\partial_\pi I(\varphi) = \left\{ p \in \Delta(\Sigma) : \tilde{I}(\psi) - \tilde{I}(\varphi) \leq \langle \psi - \varphi, p \rangle \text{ for each } \psi \in \Phi + \mathbb{R} \right\}.$$

Lemma 26 *Let $I : \Phi \to \mathbb{R}$ be a concave niveloid. Then, $\partial_\pi I(\varphi) \neq \varnothing$ for all $\varphi \in \Phi$.*

Proof. By Lemma 24, there exists a concave niveloid $\hat{I}$ on $B(\Sigma)$ such that $\hat{I}_{|\Phi} = I$. Let $\varphi \in \Phi$. Being a niveloid, $\hat{I}$ is Lipschitz continuous, then, its standard superdifferential at φ

$$\partial \hat{I}(\varphi) = \left\{ \mu \in ba(\Sigma) : \hat{I}(\psi) - \hat{I}(\varphi) \leq \langle \psi - \varphi, \mu \rangle \text{ for each } \psi \in B(\Sigma) \right\}$$

is nonempty (see, e.g., [22, p. 6-7]).

For all $c \in \mathbb{R}$ and $\mu \in \partial \hat{I}(\varphi)$ we have

$$\hat{I}(\varphi) + c = \hat{I}(\varphi + c) \leq \hat{I}(\varphi) + \langle \varphi + c - \varphi, \mu \rangle = \hat{I}(\varphi) + c\mu(S),$$

and so $c \leq c\mu(S)$. This implies $\mu(S) = 1$.

For all $\psi \geq 0$ and $\mu \in \partial \hat{I}(\varphi)$ we have

$$\langle \psi, \mu \rangle = \langle \varphi + \psi, \mu \rangle - \langle \varphi, \mu \rangle \geq \hat{I}(\varphi + \psi) - \hat{I}(\varphi) \geq 0,$$

this implies $\mu \in ba^+(\Sigma)$.

Therefore, $\partial \hat{I}(\varphi) \subseteq \partial_\pi I(\varphi)$ and we conclude that $\partial_\pi I(\varphi) \neq \varnothing$. ■

Lemma 27 *Let Φ be a convex subset of $B(\Sigma)$ such that $\Phi_c \neq \varnothing$, and $I : \Phi \to \mathbb{R}$ be a concave and normalized niveloid. Then:*

(i) For each $\varphi \in \Phi$,

$$I(\varphi) = \min_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - I^\star(p)) = \min_{p \in \bigcup_{\psi \in \Phi} \partial_\pi I(\psi)} (\langle \varphi, p \rangle - I^\star(p)) \tag{21}$$

where $I^\star : \Delta(\Sigma) \to [-\infty, 0]$ is given by

$$I^\star(p) = \inf_{\psi \in \Phi} (\langle \psi, p \rangle - I(\psi)) \quad \forall p \in \Delta(\Sigma).$$

(ii) $I^\star$ is the maximal functional $R : \Delta(\Sigma) \to [-\infty, 0]$ such that

$$I(\varphi) = \inf_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - R(p)) \quad \forall \varphi \in \Phi. \tag{22}$$

(iii) $I^\star$ coincides with the Fenchel conjugate $\hat{I}^*$ of $\hat{I}$ on $\Delta(\Sigma)$ and

$$\hat{I}(\varphi) = \min_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - I^\star(p)) \quad \forall \varphi \in B(\Sigma). \tag{23}$$

(iv) If (22) holds, and $\Psi \subseteq \Phi$ is such that $\sup_{s \in S} \psi(s) - \inf_{s \in S} \psi(s) < c$ for all $\psi \in \Psi$, then

$$I(\psi) = \inf_{\{p \in \Delta(\Sigma) : R(p) \geq -c\}} (\langle \varphi, p \rangle - R(p)) \quad \forall \psi \in \Psi. \tag{24}$$

Proof. Notice that $I^\star(p) \leq 0$ for all $p \in \Delta(\Sigma)$. For, if we take a constant $k \in \Phi_c$ we have $\langle k, p \rangle = I(k) = k$.

By definition of $I^\star$, for all $\varphi \in \Phi$ and $p \in \Delta(\Sigma)$

$$I(\varphi) \leq \langle \varphi, p \rangle - I^\star(p); \tag{25}$$

moreover,

$$\begin{aligned} p \in \partial_\pi I(\varphi) \quad &\Leftrightarrow \quad I(\varphi) \geq I(\psi) - \langle \psi, p \rangle + \langle \varphi, p \rangle \quad \forall \psi \in \Phi \\ &\Leftrightarrow \quad I(\varphi) \geq \sup_{\psi \in \Phi} (I(\psi) - \langle \psi, p \rangle) + \langle \varphi, p \rangle \\ &\Leftrightarrow \quad I(\varphi) \geq \langle \varphi, p \rangle - \inf_{\psi \in \Phi} (\langle \psi, p \rangle - I(\psi)) \\ &\Leftrightarrow \quad I(\varphi) \geq \langle \varphi, p \rangle - I^\star(p) \\ &\Leftrightarrow \quad I(\varphi) = \langle \varphi, p \rangle - I^\star(p). \end{aligned}$$

Therefore, for all $\varphi \in \Phi$

$$\begin{aligned} I(\varphi) &= \min_{p \in \partial_\pi I(\varphi)} (\langle \varphi, p \rangle - I^\star(p)) \geq \inf_{p \in \bigcup_{\psi \in \Phi} \partial_\pi I(\psi)} (\langle \varphi, p \rangle - I^\star(p)) \\ &\geq \inf_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - I^\star(p)) \geq I(\varphi), \end{aligned}$$

which implies (21). This proves (i). For later use, notice that if P is a subset of $\Delta(\Sigma)$ such that $\partial_\pi I(\varphi) \cap P \neq \emptyset$ for all $\varphi \in \Phi$, then the above argument yields

$$I(\varphi) = \min_{p \in P} \left(\langle \varphi, p \rangle - I^\star(p) \right). \tag{26}$$

Let $R : \Delta(\Sigma) \to [-\infty, 0]$ be such that $I(\varphi) = \inf_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - R(p))$ for all $\varphi \in \Phi$. Then,

$$R(p) \leq \langle \varphi, p \rangle - I(\varphi) \quad \forall p \in \Delta(\Sigma), \varphi \in \Phi,$$

and hence

$$R(p) \leq \inf_{\varphi \in \Phi} (\langle \varphi, p \rangle - I(\varphi)) = I^\star(p) \quad \forall p \in \Delta(\Sigma).$$

This proves (ii).

For all $\varphi \in B(\Sigma)$, set $\hat{J}(\varphi) = \inf_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - I^\star(p))$, $\hat{J}$ is a normalized and concave niveloid on $B(\Sigma)$ that extends I. By (ii) applied to $\hat{J}$, we obtain

$$\begin{aligned} \hat{J}^*(p) &= \hat{J}^\star(p) \geq I^\star(p) = \inf_{\varphi \in \Phi} (\langle \varphi, p \rangle - I(\varphi)) \\ &\geq \inf_{\varphi \in B(\Sigma)} \left(\langle \varphi, p \rangle - \hat{I}(\varphi) \right) \quad \text{(this is } \hat{I}^*(p)\text{)} \\ \text{(since } \hat{I} \leq \hat{J}\text{)} \quad &\geq \inf_{\varphi \in B(\Sigma)} \left(\langle \varphi, p \rangle - \hat{J}(\varphi) \right) = \hat{J}^*(p) \end{aligned}$$

that is $\hat{J}^*(p) = I^\star(p) = \hat{I}^*(p)$ for all $p \in \Delta(\Sigma)$. Apply (i) (or Remark 3) to $\hat{I}$ to obtain

$$\hat{I}(\varphi) = \min_{p \in \Delta(\Sigma)} \left(\langle \varphi, p \rangle - \hat{I}^*(p) \right) = \min_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - I^\star(p)) \quad \forall \varphi \in B(\Sigma).$$

This completes the proof of (iii).

Finally, as to (iv), the monotonicity of $\hat{I}$ implies that $\inf_{s \in S} \psi(s) = \hat{I}(\inf_{s \in S} \psi(s)) \leq I(\psi), \langle \psi, p \rangle \leq \sup_{s \in S} \psi(s) = \hat{I}(\sup_{s \in S} \psi(s))$ for all $p \in \Delta(\Sigma)$ and all $\psi \in \Psi$. For each $\psi \in \Psi$, there exists $\varepsilon > 0$ such that $\sup_{s \in S} \psi(s) - \inf_{s \in S} \psi(s) + \varepsilon < c$. For all $p \in \Delta(\Sigma)$ such that $R(p) < -c$, we have

$$\begin{aligned} R(p) &< -\sup_{s \in S} \psi(s) + \inf_{s \in S} \psi(s) - \varepsilon, \text{ and} \\ \sup_{s \in S} \psi(s) + \varepsilon &< \inf_{s \in S} \psi(s) - R(p), \text{ i.e.} \\ I(\psi) + \varepsilon &\leq \sup_{s \in S} \psi(s) + \varepsilon < \inf_{s \in S} \psi(s) - R(p) \leq \langle \varphi, p \rangle - R(p). \end{aligned}$$

On the other hand,

$$\begin{aligned} I(\psi) &= \inf_{p \in \Delta(\Sigma)} (\langle \psi, p \rangle - R(p)) \\ &= \min \left(\inf_{p \in \{R < -c\}} (\langle \psi, p \rangle - R(p)), \inf_{p \in \{R \geq -c\}} (\langle \psi, p \rangle - R(p)) \right) \end{aligned}$$

which concludes the proof, since $\inf_{p \in \{R < -c\}} (\langle \varphi, p \rangle - R(p)) \geq I(\psi) + \varepsilon$. ■

Remark 4 Inspection of the proof shows that: (i) $\partial_\pi I(\varphi) = \arg\min_{p\in\Delta(\Sigma)}(\langle\varphi,p\rangle - I^\star(p))$. (ii) If $k \in \Phi_c$, then $\partial_\pi I(k) = \{I^\star = 0\} = \arg\max_{p\in\Delta(\Sigma)} I^\star(p)$. (iii) $I^\star$ is concave and weakly* upper semicontinuous.

Corollary 28 *Let Φ be a convex subset of $B_0(\Sigma)$ (resp. $B(\Sigma)$) such that $\Phi_c \neq \emptyset$ and $\Phi + \mathbb{R} = B_0(\Sigma)$ (resp. $\Phi + \mathbb{R} = B(\Sigma)$),[13] and $I : \Phi \to \mathbb{R}$ be a concave and normalized niveloid. Then, $I^\star$ is the Fenchel conjugate of the unique niveloid $\tilde{I}$ extending I to $B_0(\Sigma)$ (resp. $B(\Sigma)$). In this case $I^\star$ is the unique concave and weakly* upper semicontinuous function $R : \Delta(\Sigma) \to [-\infty, 0]$ such that*

$$I(\varphi) = \inf_{p\in\Delta(\Sigma)} (\langle\varphi,p\rangle - R(p)) \quad \forall\varphi \in \Phi.$$

Proof. The equality

$$\begin{aligned} I^\star(p) &= \inf_{\psi\in\Phi}(\langle\psi,p\rangle - I(\psi)) = \inf_{\substack{\psi\in\Phi\\ c\in\mathbb{R}}}(\langle\psi,p\rangle + c - I(\psi) - c) \\ &= \inf_{\substack{\psi\in\Phi\\ c\in\mathbb{R}}}\left(\langle\psi + c,p\rangle - \tilde{I}(\psi + c)\right) = \inf_{\psi\in B_0(\Sigma)}\left(\langle\psi,p\rangle - \tilde{I}(\psi)\right) \end{aligned}$$

yields the first part of the statement. Let $R : \Delta(\Sigma) \to [-\infty, 0]$ be a concave and weakly* upper semicontinuous functional such that $I(\varphi) = \inf_{p\in\Delta(\Sigma)}(\langle\varphi,p\rangle - R(p))$ for all $\varphi \in \Phi$. Then, $\tilde{I}(\varphi) = \inf_{p\in\Delta(\Sigma)}(\langle\varphi,p\rangle - R(p))$ for all $\varphi \in B_0(\Sigma)$. By the Fenchel-Moreau Theorem

$$\begin{aligned} R(p) &= R^{**}(p) = \inf_{\varphi\in B_0(\Sigma)}(\langle\varphi,p\rangle - R^*(\varphi)) \\ &= \inf_{\varphi\in B_0(\Sigma)}\left(\langle\varphi,p\rangle - \left(\inf_{p\in\Delta(\Sigma)}(\langle\varphi,p\rangle - R(p))\right)\right) \\ &= \inf_{\varphi\in B_0(\Sigma)}\left(\langle\varphi,p\rangle - \tilde{I}(\varphi)\right) = \tilde{I}^*(p) = I^\star(p), \end{aligned}$$

as desired. ■

A.4 Monotone continuous niveloids

Proposition 29 *Let $I : B_0(\Sigma, K) \to \mathbb{R}$ be a normalized concave niveloid, with K unbounded and Σ a σ-algebra. Then, the following conditions are equivalent:*

(i) If $\varphi, \psi \in B_0(\Sigma, K)$, $k \in K$, $\{E_n\}_{n\geq1} \in \Sigma$ with $E_1 \supseteq E_2 \supseteq ...$ and $\bigcap_{n\geq1} E_n = \emptyset$, then $I(\varphi) > I(\psi)$ implies that there exists $n_0 \geq 1$ such that

$$I\left(k1_{E_{n_0}} + \varphi 1_{E^c_{n_0}}\right) > I(\psi).$$

[13] E.g. $\Phi = B_0(\Sigma, K)$ with K an unbounded interval.

(ii) If $R : \Delta(\Sigma) \to [-\infty, 0]$ *is such that*

$$I(\varphi) = \inf_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - R(p)) \quad \forall \varphi \in B_0(\Sigma, K),$$

then $\{p \in \Delta(\Sigma) : R(p) > -\infty\} \subseteq \Delta^\sigma(\Sigma)$.

(iii) $\{p \in \Delta(\Sigma) : I^\star(p) \geq c\}$ *is a weakly compact subset of* $\Delta^\sigma(\Sigma)$ *for each* $c \leq 0$.

(iv) There exists $q \in \Delta^\sigma(\Sigma)$ *such that* $\{p \in \Delta(\Sigma) : I^\star(p) \geq c\}$ *is a weakly compact subset of* $\Delta^\sigma(\Sigma, q)$ *for each* $c \leq 0$*, and for each* $\varphi \in B_0(\Sigma, K)$*,*

$$I(\varphi) = \min_{p \in \Delta^\sigma(\Sigma, q)} (\langle \varphi, p \rangle - I^\star(p)). \tag{27}$$

(v) There exists $r \in \Delta^\sigma(\Sigma)$ *and* $T : \Delta^\sigma(\Sigma, r) \to [-\infty, 0]$ *such that* $\{p \in \Delta(\Sigma) : T(p) \geq c\}$ *is a weakly compact subset of* $\Delta^\sigma(\Sigma, r)$ *for each* $c \leq 0$*, and for each* $\varphi \in B_0(\Sigma, K)$*,*

$$I(\varphi) = \inf_{p \in \Delta^\sigma(\Sigma, r)} (\langle \varphi, p \rangle - T(p)).$$

Proof. (i) $\Rightarrow$ (ii). For all $\varphi \in \Phi + \mathbb{R}$, $\tilde{I}(\varphi) = \inf_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - R(p))$. Moreover, if $\varphi, \psi \subseteq B_0(\Sigma)$, $c \in \mathbb{R}$, $\Sigma \ni E_n \downarrow \emptyset$, and $\tilde{I}(\varphi) > \tilde{I}(\psi)$, there exists $a \in \mathbb{R}$ such that $\varphi - a, \psi - a \subseteq B_0(\Sigma, K)$ and $c - a \in K$. Then $\tilde{I}(\varphi - a + a) > \tilde{I}(\psi - a + a)$ implies $I(\varphi - a) > I(\psi - a)$ and (i) implies that there exists $n_0 \geq 1$ such that

$$I\left((c - a) 1_{E_{n_0}} + (\varphi - a) 1_{E_{n_0}^c}\right) > I(\psi - a).$$

The definition of $\tilde{I}$ implies

$$\tilde{I}\left(c 1_{E_{n_0}} + \varphi 1_{E_{n_0}^c}\right) > \tilde{I}(\psi). \tag{28}$$

Therefore we can assume $K = \mathbb{R}$.

We begin by proving that

$$\{R \geq c\} = \{p \in \Delta(\Sigma) : R(p) \geq c\} \subseteq \Delta^\sigma(\Sigma)$$

for each $c \in \mathbb{R}$.

Let $\Sigma \ni E_n \downarrow \emptyset$. We show that $I\left(a 1_{E_n^c}\right) \uparrow a$ for all $a > 0$. Monotonicity guarantees that $\lim_n I\left(a 1_{E_n^c}\right) \leq I(a) = a$. If $b = \lim_n I\left(a 1_{E_n^c}\right) < a$, then $I(a) > I(b)$, but $I\left(a 1_{E_n^c}\right) \leq I(b) < I(a)$ for all $n \geq 1$ contradicting (i).

Let $p \in \{R \geq c\}$. Then, $\langle a 1_{E_n^c}, p \rangle - I\left(a 1_{E_n^c}\right) \geq R(p) \geq c$ for each $n \geq 1$ and each $a > 0$, and so

$$1 \geq \lim_{n \to \infty} p(E_n^c) \geq \lim_{n \to \infty} \left(\frac{c}{a} + \frac{I\left(a 1_{E_n^c}\right)}{a}\right) = \frac{c}{a} + 1.$$

Letting $a \to \infty$, we conclude that $\lim_n p(E_n) = 0$, so that $p \in \Delta^\sigma(\Sigma)$. In particular, $\{R > -\infty\} = \bigcup_{n \geq 1} \{R \geq -n\} \subseteq \Delta^\sigma(\Sigma)$, as desired (we borrowed this last argument from [13, p. 169-170]).

(ii) $\Rightarrow$ (iii) Clearly (ii) implies $\{I^\star > -\infty\} \subseteq \Delta^\sigma(\Sigma)$. Since $I^\star$ is weakly* upper semicontinuous, $\{I^\star \geq c\}$ is a weakly* closed subset of $\Delta(\Sigma)$ consisting of probability measures for each $c \leq 0$, it follows that $\{I^\star \geq c\}$ is weakly compact.

(iii) $\Rightarrow$ (iv) Since $\{I^\star \geq -n\}$ is weakly compact, there exists $q_n \in \Delta^\sigma(\Sigma)$ such that $p \ll q_n$ for each $p \in \{I^\star \geq -n\}$. Set $q = \sum_{n \geq 1} 2^{-n} q_n$, $q \in \Delta^\sigma(\Sigma)$ and $p \ll q$ for each $p \in \{I^\star > -\infty\}$. Therefore $\{p \in \Delta(\Sigma) : I^\star(p) \geq c\}$ is a (weakly compact) subset of $\Delta^\sigma(\Sigma, q)$ for each $c \leq 0$. Consider the set $\bigcup_\psi \partial_\pi I(\psi)$. It is easily seen to be included in $\{I^\star > -\infty\}$. Hence, $\bigcup_\psi \partial_\pi I(\psi) \subseteq \Delta^\sigma(\Sigma, q)$. By Lemma 27, we conclude that (27) holds.

(iv) $\Rightarrow$ (v) Take $r = q$ and $T = I^\star_{|\Delta^\sigma(\Sigma,q)}$.

(v) $\Rightarrow$ (i) Let T' be the extension of T to $\Delta(\Sigma)$ obtained by setting $T'(p) = -\infty$ if $p \notin \Delta^\sigma(\Sigma, r)$. Clearly, T' is weakly* upper semicontinuous and

$$I(\varphi) = \inf_{p \in \Delta^\sigma(\Sigma,r)} (\langle \varphi, p \rangle - T(p)) = \min_{p \in \Delta(\Sigma)} (\langle \varphi, p \rangle - T'(p)).$$ [14]

Let $\varphi, \psi \subseteq B_0(\Sigma, K)$, $k \in K$, $\Sigma \ni E_n \downarrow \emptyset$, and $I(\varphi) > I(\psi)$. Set $E_0 = \emptyset$, $a = \min\{\inf_{s \in S} \varphi(s), k\}$, $b = \max\{\sup_{s \in S} \varphi(s), k\}$, and $\varphi_n = a1_{E_n} + \varphi 1_{E_n^c}$ for all $n \geq 0$. Notice that $\varphi_0 = \varphi$ and

$$a \leq \varphi_n \leq k1_{E_n} + \varphi 1_{E_n^c} \leq b \quad \forall n \geq 0.$$

By Lemma 27(iv), setting $c = b - a + 1$, for all $n \geq 0$ we have:

$$I(\varphi_n) = \inf_{p \in \{T' \geq -c\}} (\langle \varphi_n, p \rangle - T'(p)) = \inf_{p \in \{T \geq -c\}} (\langle \varphi_n, p \rangle - T(p)).$$

On the other hand, $\{T \geq -c\}$ is a weakly compact subset of $\Delta^\sigma(\Sigma, r)$ and the functions

$$\phi_n : \begin{array}{ccc} \{T \geq -c\} & \to & \mathbb{R} \\ p & \mapsto & \langle \varphi_n, p \rangle - T(p) \end{array}$$

are weakly lower semicontinuous for all $n \geq 0$ (since T is weakly upper semicontinuous). Moreover, by the Monotone Convergence Theorem, $\phi_n \uparrow \phi_0$. The Dini-Cartan Lemma (see, e.g., [7]) guarantees that $\inf_{p \in \{T \geq -c\}} \phi_n(p) \to \inf_{p \in \{T \geq -c\}} \phi_0(p)$, that is

$$I(\varphi_n) \to I(\varphi). \tag{29}$$

Monotonicity and (29) guarantee that eventually

$$I\left(k1_{E_n} + \varphi 1_{E_n^c}\right) \geq I(\varphi_n) > I(\psi)$$

and (i) holds. ■

[14] By Corollary 28, $T' = I^\star$.

B Mean-Variance and relative Gini Index

Let Σ be a σ-algebra and $q \in \Delta^{\sigma}(\Sigma)$. Recall that $\mathrm{Var}_q(\varphi) = \int \left(\varphi - \int \varphi dq\right)^2 dq$ for all $\varphi \in L^2(\Sigma, q)$ (endowed with the usual $\|\cdot\|_2$). Notice that the functional $J : L^2(\Sigma, q) \to \mathbb{R}$ defined by

$$J(\varphi) = \int \varphi dq - \frac{\theta}{2}\mathrm{Var}_q(\varphi),$$

as well as its restrictions to $B(\Sigma)$ and $B_0(\Sigma)$, are concave and Gateaux differentiable (each in its domain). Concavity is trivial.

To prove Gateaux differentiability, let $L \in \{L^2(\Sigma, q), B(\Sigma), B_0(\Sigma)\}$. For all $\varphi, \psi \in L$ and $t \in \mathbb{R}$

$$\begin{aligned} J(\varphi + t\psi) &= \int (\varphi + t\psi)\, dq - \frac{\theta}{2}\mathrm{Var}_q(\varphi + t\psi) \\ &= -\frac{\theta}{2}\mathrm{Var}_q(\psi)\, t^2 + \left[\int \psi \left(1 - \theta\left(\varphi - \int \varphi dq\right)\right) dq\right] t + J(\varphi), \end{aligned}$$

then

$$DJ(\varphi; \psi) = \lim_{t \to 0} \frac{J(\varphi + t\psi) - J(\varphi)}{t} = \int \psi \left(1 - \theta\left(\varphi - \int \varphi dq\right)\right) dq. \tag{30}$$

If $L = L^2(\Sigma, q)$, the usual identification of $L^2(\Sigma, q)$ with its dual and (30) yield that the Gateaux derivative of J at φ is

$$\nabla J(\varphi) = 1 - \theta\left(\varphi - \int \varphi dq\right).$$

Else, (30) implies that the Gateaux derivative of the restriction of J to $B(\Sigma)$ (resp. $B_0(\Sigma)$) belongs to $\Delta^{\sigma}(\Sigma, q)$ and its Radon-Nikodym derivative w.r.t. q is $1-\theta\left(\varphi - \int \varphi dq\right)$.

In any case, the *domain of monotonicity* of J on L is the convex tube

$$M(\theta, L) = \left\{\varphi \in L : q\left(\left\{s \in S : \varphi(s) - \int \varphi dq \leq \frac{1}{\theta}\right\}\right) = 1\right\}.$$

Therefore, if $L \in \{B(\Sigma), B_0(\Sigma)\}$, the restriction I of J to $M(\theta, L)$ is a niveloid, and $\partial_{\pi} I(\varphi) \cap \Delta^{\sigma}(\Sigma, q)$ is not empty for all $\varphi \in M(\theta, L)$.[15] Setting $P = \Delta^{\sigma}(\Sigma, q)$ in (26) of the proof of Lemma 27 yields

$$I(\varphi) = \min_{p \in \Delta^{\sigma}(\Sigma, q)} \left(\langle \varphi, p\rangle - I^{\star}(p)\right) \quad \text{for all } \varphi \in M(\theta, L). \tag{31}$$

[15] It contains the probability measure defined by

$$r(A) = \int_A 1 - \theta\left(\varphi - \int \varphi dq\right) dq \quad \forall A \in \Sigma.$$

The functional defined for all $p \in \Delta(\Sigma)$ by

$$G(p\|q) = \begin{cases} \int (dp/dq)^2 \, dq - 1 & \text{if } p \in \Delta^\sigma(\Sigma, q), \\ \infty & \text{else.} \end{cases}$$

is called *relative Gini concentration index* of p with respect to q. Next theorem summarizes the properties of $G(p\|q)$. In particular, point (v) shows that

$$-I^\star(p) = \frac{1}{2\theta} G(p\|q). \tag{32}$$

Theorem 30 *Let $\theta > 0$.*

(i) The proper domain of $G(\cdot\|q)$ is $D = \{p \in \Delta^\sigma(\Sigma, q) : dp/dq \in L^2(\Sigma, q)\}$.[16]

(ii) For all $p \in D$,

$$\frac{1}{2\theta} G(p\|q) = - \min_{\varphi \in L^2(\Sigma, q)} \left\{ \int \varphi dp - \int \varphi dq + \frac{\theta}{2} \mathrm{Var}_q(\varphi) \right\}, \tag{33}$$

and the min *is attained in $M(\theta, L^2(\Sigma, q))$.*

(iii) For all $p \in \Delta^\sigma(\Sigma)$,

$$\begin{aligned} \frac{1}{2\theta} G(p\|q) &= - \inf_{\varphi \in B(\Sigma)} \left\{ \int \varphi dp - \int \varphi dq + \frac{\theta}{2} \mathrm{Var}_q(\varphi) \right\} \\ &= - \inf_{\varphi \in B_0(\Sigma)} \left\{ \int \varphi dp - \int \varphi dq + \frac{\theta}{2} \mathrm{Var}_q(\varphi) \right\}. \end{aligned} \tag{34}$$

(iv) $G(\cdot\|q)$ is convex, non-negative, and strictly convex on D. In particular, $G(p\|q) = 0$ iff $p = q$.

(v) For all $p \in \Delta^\sigma(\Sigma, q)$,

$$\frac{1}{2\theta} G(p\|q) = - \inf_{\varphi \in M(\theta, B_0(\Sigma))} \left\{ \int \varphi dp - \int \varphi dq + \frac{\theta}{2} \mathrm{Var}_q(\varphi) \right\}. \tag{35}$$

(vi) For all $p \in \Delta^\sigma(\Sigma)$,

$$G(p\|q) = \sup_{\mathcal{P}} \sum_{A \in \mathcal{P}} \frac{p(A)^2}{q(A)} - 1, \tag{36}$$

where $\mathcal{P}$ ranges over all finite partitions of S in Σ, and the summands equal 0 if $p(A) = 0$ and equal ∞ if $p(A) > 0$ and $q(A) = 0$.

(vii) For all $t \in \mathbb{R}$, $\{p \in \Delta(\Sigma) : G(p\|q) \leq t\}$ is a weakly compact subset of $\Delta^\sigma(\Sigma)$.

[16] Notice that, if $p \in D$, then $G(p\|q) = \mathrm{Var}_q(dp/dq)$.

Proof. (i) is trivial.

(ii) If $p \in D$, $dp/dq \in L^2(\Sigma, q)$ and the functional $F : L^2(\Sigma, q) \to \mathbb{R}$ given by

$$F(\varphi) = \int \varphi dp - \int \varphi dq + \frac{\theta}{2}\mathrm{Var}_q(\varphi)$$

is well defined, convex, and Gateaux differentiable. Its Gateaux derivative is

$$\nabla F(\varphi) = \frac{dp}{dq} - 1 + \theta\varphi - \theta \int \varphi dq. \tag{37}$$

Notice that $\hat{\varphi}$ solves $\nabla F(\hat{\varphi}) = 0$ iff $\hat{\varphi} + a$ does (for all $a \in \mathbb{R}$). W.l.o.g. we can assume $\int \hat{\varphi} dq = -\theta^{-1}$ and obtain $\hat{\varphi} = -\theta^{-1} dp/dq$. This implies

$$\begin{aligned}
\min_{\varphi \in L^2(\Sigma,q)} F(\varphi) &= F(\hat{\varphi}) \\
&= \int -\frac{1}{\theta}\frac{dp}{dq} dp - \int -\frac{1}{\theta}\frac{dp}{dq} dq + \frac{\theta}{2}\left(\int\left(-\frac{1}{\theta}\frac{dp}{dq}\right)^2 dq - \left(\int -\frac{1}{\theta}\frac{dp}{dq} dq\right)^2\right) \\
&= -\frac{1}{\theta}\left(\int \frac{dp}{dq} dp - 1\right) + \frac{1}{2\theta}\left(\int\left(\frac{dp}{dq}\right)^2 dq - 1\right) \\
&= -\frac{1}{2\theta}\left(\int\left(\frac{dp}{dq}\right)^2 dq - 1\right) = -\frac{1}{2\theta} G(p\|q).
\end{aligned}$$

Notice that $\hat{\varphi} - \int \hat{\varphi} dq = -\frac{1}{\theta}\frac{dp}{dq} + \frac{1}{\theta} \leq \frac{1}{\theta}$, that is, the min is attained in $M(\theta, L^2(\Sigma, q))$.

(iii) If $p \in D$, since $B(\Sigma)$ is dense in $(L^2(\Sigma, q), \|\cdot\|_2)$, then

$$\inf_{\varphi \in B(\Sigma)}\left\{\int \varphi dp - \int \varphi dq + \frac{\theta}{2}\mathrm{Var}_q(\varphi)\right\} = \min_{\varphi \in L^2(\Sigma,q)} F(\varphi) = -\frac{1}{2\theta} G(p\|q).$$

If $p \in \Delta^\sigma(\Sigma, q)$, but $dp/dq \notin L^2(\Sigma, q)$ (i.e., $\int (dp/dq)\, dp = \int (dp/dq)^2 dq = \infty$, we only have to show that

$$\inf_{\varphi \in B(\Sigma)}\left\{\int \varphi dp - \int \varphi dq + \frac{\theta}{2}\mathrm{Var}_q(\varphi)\right\} = -\infty.$$

For all $n \geq 1$,

$$\begin{aligned}
&\inf_{\varphi \in B(\Sigma)}\left\{\int \varphi dp - \int \varphi dq + \frac{\theta}{2}\mathrm{Var}_q(\varphi)\right\} \\
\leq\ & \int -\frac{1}{\theta}\left(\frac{dp}{dq} \wedge n\right) dp - \int -\frac{1}{\theta}\left(\frac{dp}{dq} \wedge n\right) dq + \frac{\theta}{2}\mathrm{Var}_q\left(-\frac{1}{\theta}\left(\frac{dp}{dq} \wedge n\right)\right) \\
=\ & -\frac{1}{\theta}\left(\int\left(\frac{dp}{dq} \wedge n\right)\frac{dp}{dq} dq - \int\left(\frac{dp}{dq} \wedge n\right) dq\right) + \frac{1}{2\theta}\left(\int\left(\frac{dp}{dq} \wedge n\right)^2 dq - \left(\int\left(\frac{dp}{dq} \wedge n\right) dq\right)^2\right) \\
=\ & a_n.
\end{aligned}$$

By the Beppo Levi Monotone Convergence Theorem,

$$b_n = \frac{1}{\theta}\int\left(\frac{dp}{dq}\wedge n\right)dq - \frac{1}{2\theta}\left(\int\left(\frac{dp}{dq}\wedge n\right)dq\right)^2 \to \frac{1}{2\theta}$$

while

$$\begin{aligned} c_n &= -\frac{1}{\theta}\int\left(\frac{dp}{dq}\wedge n\right)\frac{dp}{dq}dq + \frac{1}{2\theta}\int\left(\frac{dp}{dq}\wedge n\right)^2 dq \\ &= -\frac{1}{2\theta}\int\left(\frac{dp}{dq}\wedge n\right)\frac{dp}{dq}dq - \frac{1}{2\theta}\int\left(\frac{dp}{dq}\wedge n\right)\frac{dp}{dq}dq + \frac{1}{2\theta}\int\left(\frac{dp}{dq}\wedge n\right)^2 dq \\ &= -\frac{1}{2\theta}\int\left(\frac{dp}{dq}\wedge n\right)\frac{dp}{dq}dq - \frac{1}{2\theta}\int\left(\frac{dp}{dq}\wedge n\right)\left(\frac{dp}{dq} - \left(\frac{dp}{dq}\wedge n\right)\right)dq\ . \end{aligned}$$

The first term goes to $-\infty$ and the second is negative, so that $c_n \to -\infty$. We conclude that $a_n = b_n + c_n \to -\infty$ and

$$\inf_{\varphi\in B(\Sigma)}\left\{\int\varphi dp - \int\varphi dq + \frac{\theta}{2}\mathrm{Var}_q(\varphi)\right\} = -\infty.$$

Finally, if p is not absolutely continuous w.r.t. q, then there exists $A\in\Sigma$ such that $q(A)=0$ and $p(A)>0$, then for all $t\in\mathbb{R}$,

$$\inf_{\varphi\in B(\Sigma)}\left\{\int\varphi dp - \int\varphi dq + \frac{\theta}{2}\mathrm{Var}_q(\varphi)\right\} \le \int t1_A dp - \int t1_A dq + \frac{\theta}{2}\mathrm{Var}_q(t1_A) = tp(A)\,.$$

This obviously implies $\inf_{\varphi\in B(\Sigma)}\left\{\int\varphi dp - \int\varphi dq + \frac{\theta}{2}\mathrm{Var}_q(\varphi)\right\} = -\infty$. This concludes the proof of (34). The second equality holds since the functional

$$\varphi \mapsto \int\varphi dp - \int\varphi dq + \frac{\theta}{2}\mathrm{Var}_q(\varphi)$$

is continuous in the supnorm of $B(\Sigma)$, for all $p\in\Delta^\sigma(\Sigma)$ and $B_0(\Sigma)$ is dense in $B(\Sigma)$.

(iv) Clearly (34) implies convexity and non-negativity. Assume $p, r\in D$ are distinct and $\alpha\in(0,1)$. Then $dp/dq, dr/dq \in L^2(\Sigma,q)$ and since $\int(dp/dq)\,dq = \int(dr/dq)\,dq = 1$, they cannot be linearly dependent. By the Cauchy-Schwartz inequality

$$\begin{aligned} G(\alpha p + (1-\alpha)r||q) + 1 &= \int\left(\alpha\frac{dp}{dq} + (1-\alpha)\frac{dr}{dq}\right)^2 dq \\ &= \alpha^2\left\|\frac{dp}{dq}\right\|_2^2 + (1-\alpha)^2\left\|\frac{dr}{dq}\right\|_2^2 + 2\alpha(1-\alpha)\int\left(\frac{dp}{dq}\frac{dr}{dq}\right)dq \\ &< \alpha^2\left\|\frac{dp}{dq}\right\|_2^2 + (1-\alpha)^2\left\|\frac{dr}{dq}\right\|_2^2 + 2\alpha(1-\alpha)\left\|\frac{dp}{dq}\right\|_2\left\|\frac{dr}{dq}\right\|_2 \\ &= \left(\alpha\left\|\frac{dp}{dq}\right\|_2 + (1-\alpha)\left\|\frac{dr}{dq}\right\|_2\right)^2 \\ &\le \alpha\left\|\frac{dp}{dq}\right\|_2^2 + (1-\alpha)\left\|\frac{dr}{dq}\right\|_2^2 \\ &= \alpha(G(p||q)+1) + (1-\alpha)(G(r||q)+1) \\ &= \alpha G(p||q) + (1-\alpha)G(r||q) + 1. \end{aligned}$$

(v) Let $p \in \Delta^\sigma(\Sigma, q)$. For each finite partition $\mathcal{P}$ of S in Σ, denote by $\Sigma_\mathcal{P}$ the algebra generated by $\mathcal{P}$; moreover, $p_\mathcal{P}$ and $q_\mathcal{P}$ denote the restrictions on $\Sigma_\mathcal{P}$ of p and q, respectively. Notice that $p_\mathcal{P} \ll q_\mathcal{P}$ and $dp_\mathcal{P}/dq_\mathcal{P} \in L^2(\Sigma_\mathcal{P}, q_\mathcal{P}) = B_0(\Sigma_\mathcal{P}) = \mathbb{R}^\mathcal{P}$. Then,

$$\begin{aligned}
&\inf_{\varphi \in M(\theta, B_0(\Sigma))} \left\{ \int \varphi dp - \int \varphi dq + \frac{\theta}{2} \mathrm{Var}_q(\varphi) \right\} \qquad (38)\\
&= \inf_{\mathcal{P}} \inf_{\varphi \in M(\theta, B_0(\Sigma_\mathcal{P}))} \left\{ \int \varphi dp_\mathcal{P} - \int \varphi dq_\mathcal{P} + \frac{\theta}{2} \mathrm{Var}_{q_\mathcal{P}}(\varphi) \right\} \\
&= \inf_{\mathcal{P}} \inf_{\varphi \in M(\theta, L^2(\Sigma_\mathcal{P}, q_\mathcal{P}))} \left\{ \int \varphi dp_\mathcal{P} - \int \varphi dq_\mathcal{P} + \frac{\theta}{2} \mathrm{Var}_{q_\mathcal{P}}(\varphi) \right\} \\
&= \inf_{\mathcal{P}} \inf_{\varphi \in L^2(\Sigma_\mathcal{P}, q_\mathcal{P})} \left\{ \int \varphi dp_\mathcal{P} - \int \varphi dq_\mathcal{P} + \frac{\theta}{2} \mathrm{Var}_{q_\mathcal{P}}(\varphi) \right\} \\
&= \inf_{\mathcal{P}} \inf_{\varphi \in B_0(\Sigma_\mathcal{P})} \left\{ \int \varphi dp_\mathcal{P} - \int \varphi dq_\mathcal{P} + \frac{\theta}{2} \mathrm{Var}_{q_\mathcal{P}}(\varphi) \right\} \\
&= \inf_{\varphi \in B_0(\Sigma)} \left\{ \int \varphi dp - \int \varphi dq + \frac{\theta}{2} \mathrm{Var}_q(\varphi) \right\} = -\frac{1}{2\theta} G(p \| q),
\end{aligned}$$

that implies (35).

(vi) Setting $\theta = 1/2$ in (38) yields

$$G(p \| q) = -\inf_{\mathcal{P}} \inf_{\varphi \in B_0(\Sigma_\mathcal{P})} \left\{ \int \varphi dp_\mathcal{P} - \int \varphi dq_\mathcal{P} + \frac{1}{4} \mathrm{Var}_{q_\mathcal{P}}(\varphi) \right\}. \qquad (39)$$

In turn, (iii) yields

$$\begin{aligned}
\inf_{\varphi \in B_0(\Sigma_\mathcal{P})} \left\{ \int \varphi dp_\mathcal{P} - \int \varphi dq_\mathcal{P} + \frac{1}{4} \mathrm{Var}_{q_\mathcal{P}}(\varphi) \right\} &= -G(p_\mathcal{P} \| q_\mathcal{P}) = 1 - \int \left(\frac{dp_\mathcal{P}}{dq_\mathcal{P}} \right)^2 dq_\mathcal{P} \\
&= 1 - \int \left(\sum_{A \in \mathcal{P}} \frac{p(A)}{q(A)} 1_A \right)^2 dq_\mathcal{P} \\
&= 1 - \sum_{A \in \mathcal{P}} \frac{p(A)^2}{q(A)},
\end{aligned}$$

then

$$G(p \| q) = -\inf_{\mathcal{P}} \left(1 - \sum_{A \in \mathcal{P}} \frac{p(A)^2}{q(A)} \right) = \sup_{\mathcal{P}} \sum_{A \in \mathcal{P}} \frac{p(A)^2}{q(A)} - 1,$$

that is (36) if $p \in \Delta^\sigma(\Sigma, q)$.

If p is not absolutely continuous w.r.t. q, then there exists $B \in \Sigma$ such that $q(B) = 0$ and $p(B) > 0$, then

$$\sup_{\mathcal{P}} \sum_{A \in \mathcal{P}} \frac{p(A)^2}{q(A)} - 1 \geq \frac{p(B)^2}{q(B)} + \frac{p(B^c)^2}{q(B^c)} - 1 = \infty = G(p \| q).$$

(vii) Clearly we can restrict our attention to $t \geq 0$. Set $C = \{p \in \Delta(\Sigma) : G(p\|q) \leq t\} \subseteq \Delta^{\sigma}(\Sigma)$. We show that (a) $\lim_{q(B)\to 0} p(B) = 0$ uniformly w.r.t. $p \in C$, and (b) $\{p_n\}_{n\geq 1} \subseteq C$ and $p_n(B) \to p(B)$ for all $B \in \Sigma$, then $p \in C$. Then, a classical result of Bartle, Dunford, and Schwartz guarantees that C is weakly compact. (See, e.g., [9, p. 306].)

(a) For all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\frac{2\delta}{\varepsilon}(t+1) < \frac{1}{2}\varepsilon$$

For all B such that $q(B) < \delta$,

$$\begin{aligned} p(B) &= \int_B \frac{dp}{dq}dq = \int_{B\cap\left\{\frac{dp}{dq}\leq\frac{\varepsilon}{2\delta}\right\}} \frac{dp}{dq}dq + \int_{B\cap\left\{\frac{dp}{dq}>\frac{\varepsilon}{2\delta}\right\}} \frac{dp}{dq}dq \\ &\leq \frac{1}{2}\frac{\varepsilon}{\delta}q(B) + \int_{B\cap\left\{\frac{dp}{dq}>\frac{\varepsilon}{2\delta}\right\}} \left(\frac{dp}{dq}\right)^2\left(\frac{dp}{dq}\right)^{-1} dq \leq \frac{1}{2}\varepsilon + \frac{2\delta}{\varepsilon}\int_{B\cap\left\{\frac{dp}{dq}>\frac{\varepsilon}{2\delta}\right\}} \left(\frac{dp}{dq}\right)^2 dq \\ &\leq \frac{1}{2}\varepsilon + \frac{2\delta}{\varepsilon}(t+1) < \varepsilon \end{aligned}$$

for all $p \in C$.

(b) For any finite partition $\mathcal{P}$ of S in Σ, define on $\Delta^{\sigma}(\Sigma)$ the function

$$G_{\mathcal{P}}(p\|q) = \sum_{A\in\mathcal{P}} \frac{p(A)^2}{q(A)} - 1$$

where the summands equal 0 if $p(A) = 0$ and equal ∞ if $p(A) > 0$ and $q(A) = 0$, and set $C_{\mathcal{P}} = \{p \in \Delta^{\sigma}(\Sigma) : G_{\mathcal{P}}(p\|q) \leq t\}$. We show that $\{p_n\}_{n\geq 1} \subseteq C_{\mathcal{P}}$ and $p_n(B) \to p(B)$ for all $B \in \Sigma$, then $p \in C_{\mathcal{P}}$. First notice that $p \in \Delta^{\sigma}(\Sigma)$ (for the Vitali-Hahn-Saks Theorem). For all $A \in \mathcal{P}$ such that $q(A) = 0$, then $p_n(A) = 0$ for all $n \geq 1$ (else $G_{\mathcal{P}}(p_n\|q) = \infty$), and hence

$$\frac{p_n(A)^2}{q(A)} \to 0 = \frac{p(A)^2}{q(A)}.$$

For all $A \in \mathcal{P}$ such that $q(A) > 0$, then

$$\frac{p_n(A)^2}{q(A)} \to \frac{p(A)^2}{q(A)}$$

we conclude that

$$\sum_{A\in\mathcal{P}} \frac{p(A)^2}{q(A)} - 1 = \lim_{n\to\infty}\sum_{A\in\mathcal{P}} \frac{p_n(A)^2}{q(A)} - 1 \leq t,$$

as wanted. Now (b) descends from the observation that

$$\mathcal{C} = \{p \in \Delta^{\sigma}(\Sigma) : G(p\|q) \leq t\} = \left\{p \in \Delta^{\sigma}(\Sigma) : \sup_{\mathcal{P}} G_{\mathcal{P}}(p\|q) \leq t\right\} = \bigcap_{\mathcal{P}} \mathcal{C}_{\mathcal{P}}.$$

■

C Proofs of the Results in the main text

The standard proof of Lemma 1 is omitted.

Lemma 31 *A binary relation $\succsim$ on $\mathcal{F}$ satisfies A.1-A.4 and A.6 if and only if there exist a non-constant affine function $u : X \to \mathbb{R}$ and a normalized niveloid $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ such that*

$$f \succsim g \Leftrightarrow I(u(f)) \geq I(u(g)).$$

Proof. Assume $\succsim$ on $\mathcal{F}$ satisfies A.1-A.4 and A.6. Let $x, y \in X$ be such that $x \sim y$. If there exists $z \in X$ such that $\frac{1}{2}x + \frac{1}{2}z \nsim \frac{1}{2}y + \frac{1}{2}z$, w.l.o.g. $\frac{1}{2}x + \frac{1}{2}z \succ \frac{1}{2}y + \frac{1}{2}z$; by A.2 (we can replace z with x to obtain) $\frac{1}{2}x + \frac{1}{2}x \succ \frac{1}{2}y + \frac{1}{2}x$ and (we can replace z with y to obtain) $\frac{1}{2}x + \frac{1}{2}y \succ \frac{1}{2}y + \frac{1}{2}y$, and conclude $x \succ y$, which is absurd. Then the premises of the Mixture Space Theorem (Hernstein and Milnor [19]) are satisfied, and there exists an affine function $u : X \to \mathbb{R}$ such that $x \succsim y$ iff $u(x) \geq u(y)$. By A.6 there exist $f, g \in \mathcal{F}$ such that $f \succ g$. Let $x, y \in X$ be such that $x \succsim f(s)$ and $g(s) \succsim y$ for all $s \in S$, then $x \succsim f \succ g \succsim y$ implies $x \succ y$, and u cannot be constant. Moreover u is unique up to positive affine transformations and we can assume $0 \in \operatorname{int}(u(X))$.

For all $f \in \mathcal{F}$, let $x, y \in X$ be such that $x \succsim f(s) \succsim y$ for all $s \in S$, then $x \succsim f \succsim y$. By A.3 the sets $\{\alpha \in [0,1] : \alpha x + (1-\alpha)y \succsim f\}$ and $\{\alpha \in [0,1] : f \succsim \alpha x + (1-\alpha)y\}$ are closed; they are nonempty since 1 belongs to the first and 0 to the second; their union is the whole $[0,1]$. Since $[0,1]$ is connected, their intersection is not empty, hence there exists $\beta \in [0,1]$ such that $\beta x + (1-\beta)y \sim f$. In particular, any act f admits a certainty equivalent $x_f \in X$.

If $f \sim x_f$, set $U(f) = u(x_f)$. U is well defined since $f \sim x_f$ and $f \sim y_f$ with $x_f, y_f \in X$ implies $x_f \sim y_f$ and $u(x_f) = u(y_f)$. Clearly, $f \succsim g$ iff $x_f \succsim x_g$ iff $u(x_f) \succsim u(x_g)$ iff $U(f) \geq U(g)$. Therefore U represents $\succsim$.

If $f \in \mathcal{F}$ then $u(f) \in B_0(\Sigma, u(X))$. Conversely, if $\varphi \in B_0(\Sigma, u(X))$,

$$\varphi(s) = u(x_i) \text{ if } s \in A_i$$

for suitable $x_1, ..., x_N \in X$ and a partition $\{A_1, A_2, ..., A_N\}$ of S in Σ. Therefore, setting

$$f(s) = x_i \text{ if } s \in A_i$$

we have $\varphi = u(f)$. We can conclude that $B_0(\Sigma, u(X)) = \{u(f) : f \in \mathcal{F}\}$. Moreover, $u(f) = u(g)$ iff $u(f(s)) = u(g(s))$ for all $s \in S$ iff $f(s) \sim g(s)$ for all $s \in S$, and by A.4, $f \sim g$ or equivalently $U(f) = U(g)$.

Define $I(\varphi) = U(f)$ if $\varphi = u(f)$. By what we have just observed, $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ is well defined. If $\varphi = u(f)$ and $\psi = u(g) \in B_0(\Sigma, u(X))$ and $\varphi \geq \psi$, then $u(f(s)) \geq u(g(s))$ for all $s \in S$, and $f(s) \succsim g(s)$ for all $s \in S$, so $f \succsim g$, $U(f) \geq U(g)$, and $I(\varphi) = I(u(f)) = U(f) \geq U(g) = I(u(g)) = I(\psi)$. Therefore, I is monotonic. Take $k \in u(X)$, say $k = u(x)$, $I(k1_S) = I(u(x) 1_S) = U(x) = u(x) = k$. Therefore, I is normalized.

Take $\alpha \in (0,1)$, $\varphi = u(f) \in B_0(\Sigma, u(X))$, $k = u(x_k) \in u(X)$; denote by x_0 an element in X such that $u(x_0) = 0$. Choose $x, y \in X$ such that $x \succsim f(s) \succsim y$ for all $s \in S$, then $\alpha x + (1-\alpha) x_0 \succsim \alpha f(s) + (1-\alpha) x_0 \succsim \alpha y + (1-\alpha) x_0$ for all $s \in S$. The technique used in the second paragraph of this proof yields the existence of $\beta \in [0,1]$ such that

$$\beta(\alpha x + (1-\alpha) x_0) + (1-\beta)(\alpha y + (1-\alpha) x_0) \sim \alpha f + (1-\alpha) x_0,$$

that is $\alpha z + (1-\alpha) x_0 \sim \alpha f + (1-\alpha) x_0$, where $z = \beta x + (1-\beta) y \in X$. Then, by A.2, $\alpha z + (1-\alpha) x_k \sim \alpha f + (1-\alpha) x_k$, and

$$\begin{aligned} I(\alpha\varphi + (1-\alpha) k) &= I(u(\alpha f + (1-\alpha) x_k)) \\ &= u(\alpha z + (1-\alpha) x_k) \\ &= \alpha u(z) + (1-\alpha) k \\ &= \alpha u(z) + (1-\alpha) 0 + (1-\alpha) k \\ &= u(\alpha z + (1-\alpha) x_0) + (1-\alpha) k \\ &= I(u(\alpha f + (1-\alpha) x_0)) + (1-\alpha) k \\ &= I(\alpha\varphi) + (1-\alpha) k. \end{aligned}$$

By Lemma 19, I is vertically invariant, and we already proved that it is monotonic. By Lemma 23 it is a niveloid, and we already proved that it is normalized.

Conversely, assume there exist a non-constant affine function $u: X \to \mathbb{R}$ and a normalized niveloid $I: B_0(\Sigma, u(X)) \to \mathbb{R}$ such that

$$f \succsim g \Leftrightarrow I(u(f)) \geq I(u(g)).$$

Choose $c \in \mathbb{R}$ such that $0 \in \operatorname{int}(u(X) + c)$ and set $v = u + c$. Define $J: B_0(\Sigma, v(X)) \to \mathbb{R}$ by $J(\varphi) = I(\varphi - c) + c$. Notice that J is a normalized niveloid,[17] and

$$\begin{aligned} f \succsim g &\Leftrightarrow I(u(f)) \geq I(u(g)) \\ &\Leftrightarrow I(u(f) + c - c) + c \geq I(u(g) + c - c) + c \\ &\Leftrightarrow I(v(f) - c) + c \geq I(v(g) - c) + c \\ &\Leftrightarrow J(v(f)) \geq J(v(g)). \end{aligned}$$

[17] In fact, for all $\varphi, \psi \in B_0(\Sigma, v(X))$,

$$J(\varphi) - J(\psi) = I(\varphi - c) + c - I(\psi - c) - c \leq \sup((\varphi - c) - (\psi - c)) = \sup(\varphi - \psi).$$

Moreover, for all $t \in v(X)$,

$$J(t) = I(t - c) + c = t - c + c = t.$$

Clearly, $\succsim$ satisfies A.1.

If $f, g \in \mathcal{F}$, $x, y \in X$, $\alpha \in (0,1)$, then $\alpha v(h), (1-\alpha)v(z), \alpha v(h) + (1-\alpha)v(z) \in B_0(\Sigma, v(X))$ for $h = f, g$ and $z = x, y$; moreover

$$\begin{aligned}
\alpha f + (1-\alpha)x &\succsim \alpha g + (1-\alpha)x \Rightarrow \\
J(\alpha v(f) + (1-\alpha)v(x)) &\geq J(\alpha v(g) + (1-\alpha)v(x)) \Rightarrow \\
J(\alpha v(f)) + (1-\alpha)v(x) &\geq J(\alpha v(g)) + (1-\alpha)v(x) \Rightarrow \\
J(\alpha v(f)) + (1-\alpha)v(y) &\geq J(\alpha v(g)) + (1-\alpha)v(y) \Rightarrow \\
J(\alpha v(f) + (1-\alpha)v(y)) &\geq J(\alpha v(g) + (1-\alpha)v(y)) \Rightarrow \\
\alpha f + (1-\alpha)y &\succsim \alpha g + (1-\alpha)y
\end{aligned}$$

and A.2 holds.

If $f, g, h \in \mathcal{F}$, $\alpha \in [0,1]$, and there exists $\alpha_n \in [0,1]$ such that $\alpha_n \to \alpha$ and $\alpha_n f + (1-\alpha_n)g \succsim h$ for all $n \geq 1$; then $v(\alpha_n f + (1-\alpha_n)g) = \alpha_n v(f) + (1-\alpha_n)v(g)$converges uniformly to $\alpha v(f) + (1-\alpha)v(g) = v(\alpha f + (1-\alpha)g)$. $J(v(\alpha_n f + (1-\alpha_n)g)) \geq J(v(h))$ for all $n \geq 1$, and the continuity of J guarantee $J(v(\alpha f + (1-\alpha)g)) \geq J(v(h))$. Therefore $\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succsim h\}$ is closed. A similar argument shows that $\{\alpha \in [0,1] : h \succsim \alpha f + (1-\alpha)g\}$ is closed too, and A.3 holds.

Given $f, g \in \mathcal{F}$, $f(s) \succsim g(s)$ for all $s \in S$ iff $J(v(f(s))) \geq J(v(g(s)))$ for all s iff $v(f(s)) \geq v(g(s))$ for all s, then monotonicity of J yields $J(v(f)) \geq J(v(g))$. This shows A.4.

Finally, since v is not constant and it represents $\succsim$ on X, there exist $x \succ y$, and A.6 holds too. ■

Proof of Theorem 2 (and Proposition 5). Assume $\succsim$ satisfies A.1-A.6. Lemma 31, guarantees that there exist a non-constant affine function $u : X \to \mathbb{R}$ and a normalized niveloid $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ such that

$$f \succsim g \Leftrightarrow I(u(f)) \geq I(u(g)).$$

Next we show that A.5 implies that $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ is concave. Let $\varphi, \psi \in B_0(\Sigma, u(X))$ be such that $I(\varphi) = I(\psi)$ and $\alpha \in (0,1)$. If $f, g \in \mathcal{F}$ are such that $\varphi = u(f)$ and $\psi = u(g)$, then $f \sim g$ and, by A.5, $\alpha f + (1-\alpha)g \succsim f$, that is

$$\begin{aligned}
I(\alpha\varphi + (1-\alpha)\psi) &= I(\alpha u(f) + (1-\alpha)u(g)) = I(u(\alpha f + (1-\alpha)g)) \\
&\geq I(u(f)) = I(\varphi).
\end{aligned}$$

Lemma 20 guarantees concavity of I.

The functional $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ is, therefore, a concave and normalized niveloid. For all $p \in \Delta(\Sigma)$, set

$$\begin{aligned}
c^\star(p) &= -I^\star(p) \\
&= -\inf_{\psi \in B_0(\Sigma, u(X))} (\langle \psi, p \rangle - I(\psi)) = \sup_{f \in \mathcal{F}} \left(u(x_f) - \int u(f)\, dp \right)
\end{aligned}$$

(where x_f is a certainty equivalent for f). Lemma 27 and Remark 4 guarantee that

$$I(\varphi) = \min_{p\in\Delta(\Sigma)} \left(\int \varphi dp + c^\star(p) \right) \quad \forall \varphi \in B_0(\Sigma, u(X)),$$

and that $c^\star$ is non-negative, grounded, convex, and weakly* lower semicontinuous.

Let $c : \Delta(\Sigma) \to [0,\infty]$ be a grounded, convex, and weakly* lower semicontinuous function such that

$$f \succsim g \Leftrightarrow \min_{p\in\Delta(\Sigma)} \left(\int u(f)\, dp + c(p) \right) \geq \min_{p\in\Delta(\Sigma)} \left(\int u(g)\, dp + c(p) \right).$$

For all $\varphi = u(f) \in B_0(\Sigma, u(X))$,

$$\begin{aligned} I(\varphi) &= I(u(f)) = u(x_f) = \min_{p\in\Delta(\Sigma)} \left(\int u(x_f)\, dp + c(p) \right) \\ &= \min_{p\in\Delta(\Sigma)} \left(\int u(f)\, dp + c(p) \right) = \min_{p\in\Delta(\Sigma)} \left(\int \varphi dp + c(p) \right). \end{aligned}$$

Lemma 27 yields $c^\star \leq c$ (this concludes the proof that (i) implies (ii)). Moreover, if $u(X)$ is unbounded, Corollary 28 guarantees $c = c^\star$ (this proves Proposition 5).

For the converse, notice that u and $I(\varphi) = \min_{p\in\Delta(\Sigma)} \left(\int \varphi dp + c(p) \right)$ are a non-constant affine function and a normalized niveloid representing $\succsim$. By Lemma 31, $\succsim$ satisfies A.1-A.4 and A.6. Concavity of I guarantees A.5.[18] ■

Proof of Corollary 4. Let $(u_0, c_0^\star)$ represent $\succsim$ as in Theorem 2. If $(u, c^\star)$ is another representation of $\succsim$ (as in Theorem 2), by (6) u and u_0 are affine representations of the restriction of $\succsim$ to X. Hence, by standard uniqueness results there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u = \alpha u_0 + \beta$. By (7),

$$c^\star(p) = \sup_{f\in\mathcal{F}} \left(u(x_f) - \int u(f)\, dp \right) = \sup_{f\in\mathcal{F}} \left(\alpha u_0(x_f) + \beta - \int (\alpha u_0(f) + \beta)\, dp \right) = \alpha c_0^\star(p),$$

as desired. The converse is trivial. ■

Lemma 32 *Let $\succsim$ be a binary relation on X represented by an affine function $u : X \to \mathbb{R}$. $u(X)$ is unbounded (either below or above) iff $\succsim$ satisfies A.7.*

Proof. If $u(X)$ is unbounded below, w.l.o.g. $u(X) \supseteq (-\infty, 0]$. Let $x \in u^{-1}(0)$, $y \in u^{-1}(-1)$, and for all $\alpha \in (0,1)$ choose $z = z(\alpha) \in u^{-1}\left(-\frac{2}{\alpha}\right)$, to obtain

$$u(y) = -1 > -2 = \alpha u(z) + (1-\alpha) u(x) = u(\alpha z + (1-\alpha) x).$$

[18]If $f \sim g$ and $\alpha \in (0,1)$, then

$$I(u(\alpha f + (1-\alpha) g)) = I(\alpha u(f) + (1-\alpha) u(g)) \geq \alpha I(u(f)) + (1-\alpha) I(u(g)) = I(u(f)).$$

If $u(X)$ is unbounded above, w.l.o.g. $u(X) \supseteq [0, \infty)$. Let $x \in u^{-1}(1)$, $y \in u^{-1}(0)$, and for all $\alpha \in (0,1)$ choose $z = z(\alpha) \in u^{-1}\left(\frac{2}{\alpha}\right)$, to obtain

$$u(x) = 1 < 2 = \alpha u(z) + (1-\alpha) u(y) = u(\alpha z + (1-\alpha) y).$$

This proves sufficiency.

If $\succsim$ satisfies A.7, then there exist $x, y \in X$ such that $x \succ y$ – w.l.o.g. $u(x) = 1$ and $u(y) = 0$ – and for all $n \in \mathbb{N}$ there exists $z_n \in X$ such that

$$\begin{aligned} &\text{either } y \succ \frac{1}{n} z_n + \left(1 - \frac{1}{n}\right) x \text{ or } \frac{1}{n} z_n + \left(1 - \frac{1}{n}\right) y \succ x \text{ i.e.} \\ &\text{either } 0 > \frac{1}{n} u(z_n) + \left(1 - \frac{1}{n}\right) \text{ or } \frac{1}{n} u(z_n) > 1 \text{ i.e.} \\ &\text{either } u(z_n) < 1 - n \text{ or } u(z_n) > n. \end{aligned}$$

This obviously implies that $\{u(z_n) : n \in \mathbb{N}\}$ is an unbounded subset of $u(X)$. ■

Proposition 6 is part of Lemma 34.

Proof of Proposition 7. Let $(u_i, c_i^\star)$ represent $\succsim_i$ as in Theorem 2, $i = 1, 2$, and set $I_i(\varphi) = \min_{p \in \Delta(\Sigma)} \left(\int \varphi dp + c_i^\star(p)\right)$ for all $\varphi \in B_0(\Sigma, u_i(X))$.

(i) implies (ii). By (8), we can choose $u_1 = u_2 = u$. For all $f \in \mathcal{F}$, if $f \sim_1 x$, then $f \succsim_2 x$; therefore

$$I_1(u(f)) = u(x) \leq I_2(u(f)).$$

This implies $I_1 \leq I_2$, and

$$c_1^\star(p) = \sup_{\varphi \in B_0(\Sigma, u(X))} \left(I_1(\varphi) - \int \varphi dp\right) \leq \sup_{\varphi \in B_0(\Sigma, u(X))} \left(I_2(\varphi) - \int \varphi dp\right) = c_2^\star(p)$$

for all $p \in \Delta(\Sigma)$.

(ii) implies (i). Let $u_1 = u_2 = u$. For all $f \in \mathcal{F}$ and $x \in X$, if $f \succsim_1 x$, then

$$\min_{p \in \Delta(\Sigma)} \left(\int u(f) \, dp + c_1^\star(p)\right) \geq u(x),$$

but $c_1^\star \leq c_2^\star$ implies $\min_{p \in \Delta(\Sigma)} \left(\int u(f) \, dp + c_2^\star(p)\right) \geq \min_{p \in \Delta(\Sigma)} \left(\int u(f) \, dp + c_1^\star(p)\right)$, finally $f \succsim_2 x$. ■

Proof of Theorem 11. Let $(u, c^\star)$ represent $\succsim$ as in Theorem 2, and set, for all $\varphi \in B_0(\Sigma, u(X))$, $I(\varphi) = \min_{p \in \Delta(\Sigma)} \left(\int \varphi dp + c^\star(p)\right)$. It is easy to check that $\succsim$ satisfies A.8 on $\mathcal{F}$ iff I satisfies the condition (i) of Proposition 29 on $B_0(\Sigma, u(X))$. Unboundedness of $u(X)$ and the relation $c^\star = -I^\star$ allow to apply Proposition 29 and obtain the desired equivalence. ■

Lemma 33 *Let $\succsim$ be a variational preference that satisfies axiom A.8. Then, for all $f, g \in \mathcal{F}$,*

$$f \succsim g \Leftrightarrow \inf_{p\in\Delta^{\sigma}(\Sigma)} \left(\int u(f)\,dp + c^{\star}(p)\right) \geq \inf_{p\in\Delta^{\sigma}(\Sigma)} \left(\int u(f)\,dp + c^{\star}(p)\right).$$

Proof. Let $(u, c^{\star})$ represent $\succsim$ as in Theorem 2, and set, for all $\varphi \in B_0(\Sigma, u(X))$, $I(\varphi) = \min_{p\in\Delta(\Sigma)} \left(\int \varphi dp + c^{\star}(p)\right)$. Let $\varphi \in \operatorname{int}(B_0(\Sigma, u(X)))$, $E_n \downarrow \emptyset$, and $\varepsilon > 0$ such that $\varphi - \varepsilon \in B_0(\Sigma, u(X))$, then

$$I\left(\varphi 1_{E_n^c} + (\min\varphi - \varepsilon) 1_{E_n}\right) - I(\varphi) \leq I(\varphi - \varepsilon 1_{E_n}) - I(\varphi) \leq -\varepsilon p(E_n) \leq 0$$

for all $p \in \partial_{\pi} I(\varphi)$. Consider a sequence $\{k_j\}_{j\geq 1}$ in $u(X)$ such $k_j < I(\varphi)$ and $k_j \uparrow I(\varphi)$. By A.8, I satisfies (i) of Proposition 29 on $B_0(\Sigma, u(X))$, then for all $j \geq 1$ there exists $n_0 \geq 1$ such that $k_j < I\left(\varphi 1_{E_{n_0}^c} + (\min\varphi - \varepsilon) 1_{E_{n_0}}\right)$. Since the sequence $I\left(\varphi 1_{E_n^c} + (\min\varphi - \varepsilon) 1_{E_n}\right)$ is increasing, this implies

$$\lim_{n\to\infty} I\left(\varphi 1_{E_n^c} + (\min\varphi - \varepsilon) 1_{E_n}\right) > k_j.$$

Passing to the limit for $j \to \infty$, we obtain

$$\lim_{n\to\infty} I\left(\varphi 1_{E_n^c} + (\min\varphi - \varepsilon) 1_{E_n}\right) = I(\varphi), \tag{40}$$

and $p(E_n) \to 0$ (uniformly w.r.t. $p \in \partial_{\pi} I(\varphi)$), that is $\partial_{\pi} I(\varphi)$ is a (weakly compact) subset of $\Delta^{\sigma}(\Sigma)$. By Lemma 27 and Remark 4

$$I(\varphi) = \min_{p\in\Delta(\Sigma)} \left(\int \varphi dp + c^{\star}(p)\right) \text{ and } J(\varphi) = \inf_{p\in\Delta^{\sigma}(\Sigma)} \left(\int \varphi dp + c^{\star}(p)\right)$$

coincide on $\operatorname{int}(B_0(\Sigma, u(X)))$, being continuous, they coincide on $B_0(\Sigma, u(X))$. ■

Proof of Proposition 12. Let $(u, c^{\star})$ represent $\succsim$ as in Theorem 2, w.l.o.g., assume $[-1, 1] \subseteq u(X)$.

(i) ⇒ (ii) By [15, Thm 1], there is a weakly* compact and convex set $C \subseteq \Delta(\Sigma)$ such that $u(x_f) = \min_{p\in C} \int u(f)\,dp$ for all $f \in \mathcal{F}$ and each $x_f \sim f$. By Theorem 2,

$$c^{\star}(p) = \sup_{f\in\mathcal{F}} \left(\min_{q\in C} \int u(f)\,dq - \int u(f)\,dp\right) \qquad \forall p \in \Delta(\Sigma). \tag{41}$$

Suppose $p \in C$, then $c^{\star}(p) \leq 0$, since $c^{\star}$ is non-negative, we have $c^{\star}(p) = 0$. Next, suppose $p_0 \notin C$, by the Separating Hyperplane Theorem, there is a simple measurable function $\varphi : S \to u(X)$ such that $\int \varphi dp > \int \varphi dp_0$ for each $p \in C$. Hence, taking $f \in \mathcal{F}$ such that $\varphi = u(f)$, $\min_{p\in C} \int u(f)\,dp - \int u(f)\,dp_0 > 0$, which in turn implies $c^{\star}(p_0) > 0$. We conclude that $c^{\star}(p) = 0$ if and only if $p \in C$. Therefore, for all $f \in \mathcal{F}$

$$\begin{aligned} \min_{p\in\Delta(\Sigma)} \left(\int u(f)\,dp + c^{\star}(p)\right) &= u(x_f) = \min_{p\in C} \int u(f)\,dp \\ &= \min_{p\in\{c^{\star}=0\}} \int u(f)\,dp. \end{aligned}$$

(ii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (ii) are trivial.

Assume $\succsim$ is unbounded above (resp. below).

(ii) $\Rightarrow$ (iii) For all $p \in \Delta(\Sigma)$

$$\begin{aligned} c^{\star}(p) &= \sup_{f \in \mathcal{F}} \left(\min_{q \in \{c^{\star}=0\}} \int u(f)\, dq - \int u(f)\, dp \right) \\ &= \sup_{\varphi \in B_0(\Sigma, u(X))} \left(\min_{q \in \{c^{\star}=0\}} \int \varphi dq - \int \varphi dp \right) \end{aligned}$$

Suppose, $c^{\star}(p_0) > 0$. There exist a non-negative (resp. non-positive), simple, measurable function $\varphi : S \to u(X)$ and $\varepsilon > 0$ such that $\min_{p \in \{c^{\star}=0\}} \int \varphi dp - \int \varphi dp_0 > \varepsilon$, but $n\varphi \in B_0(\Sigma, u(X))$ for all $n \in \mathbb{N}$ and

$$c^{\star}(p_0) \geq \min_{p \in \{c^{\star}=0\}} \int n\varphi dp - \int n\varphi dp_0 \geq n\varepsilon$$

for all $n \in \mathbb{N}$. We conclude $c^{\star}(p_0) = \infty$. ■

The proof of Corollary 13 is omitted. Just notice that A.5' can be used to obtain affinity of the functional I appearing in Lemma 31 in the same way in which A.5 is used to obtain its concavity at the beginning of the proof of Theorem 2.

Lemma 34 *Let $\succsim$ be a variational preference represented by $(u, c^{\star})$ as in Theorem 2 and $q \in \Delta(\Sigma)$. The following conditions are equivalent:*

(i) q corresponds to a SEU preference less ambiguity averse than $\succsim$;

(ii) $c^{\star}(q) = 0$;

(iii) $q \in \partial I(k)$ for some (all) $k \in u(X)$, where $I(\varphi) = \min_{p \in \Delta(\Sigma)} \left(\int \varphi dp + c^{\star}(p) \right)$ for all $\varphi \in B_0(\Sigma, u(X))$.

In particular, any variational preference is ambiguity averse.

Proof. (i) implies (ii). Suppose $\succsim_0$ is a SEU preference, with associated subjective probability q and utility index u_0 such that $\succsim$ is more ambiguity averse than $\succsim_0$. By (8), we can assume $u_0 = u$. By Proposition 7, $c^{\star} \leq c_0^{\star}$, by Corollary 13 $c_0^{\star}(q) = 0$, and hence $0 \leq c^{\star}(q) \leq c_0^{\star}(q) = 0$.

(ii) implies (iii). $c^{\star}(q) = 0$ iff $\sup_{\varphi \in B_0(\Sigma, u(X))} \left(I(\varphi) - \int \varphi dq \right) = 0$ iff $I(\varphi) \leq \int \varphi dq$ for all $\varphi \in B_0(\Sigma, u(X))$ iff $I(\varphi) - I(k) \leq \int \varphi dq - \int k dq$ for all $\varphi \in B_0(\Sigma, u(X))$ and all $k \in u(X)$ iff $q \in \partial I(k)$ for all $k \in u(X)$.

(iii) implies (i). If $q \in \partial I(k)$ for some $k \in u(X)$, then $I(\varphi) - I(k) \leq \int \varphi dq - \int k dq$ for all $\varphi \in B_0(\Sigma, u(X))$, then $I(\varphi) \leq \int \varphi dq$ for all $\varphi \in B_0(\Sigma, u(X))$. Denote by $\succsim_0$ the SEU preference, with associated subjective probability q and utility index u.

Notice that for all $f \in \mathcal{F}$ and $x \in X$: $f \succsim x$ implies $I(u(f)) \geq u(x)$, *a fortiori* $\int u(f)\, dq \geq u(x)$ and $f \succsim_0 x$.

Ambiguity aversion of $\succsim$ now follows from the observation that $\arg\min_{p\in\Delta(\Sigma)} c^\star(p)$ is nonempty and $\min_{p\in\Delta(\Sigma)} c^\star(p) = 0$ (or, equivalently, from Lemma 26 that guarantees $\partial I(k) \cap \Delta(\Sigma) = \partial_\pi I(k) \neq \emptyset$ for all $k \in u(X)$). ■

Proof of Theorem 14. Theorem 2 guarantees that $\succsim^m$ is a variational preference. Moreover, the sets $\{p \in \Delta(\Sigma) : R(p\|q) \leq c\}$ are weakly* compact (and so weakly compact) subsets of $\Delta^\sigma(\Sigma)$ for all $c \geq 0$ (see [16, p. 557]). Therefore $\theta R(\cdot\|q)$ is a non-negative, grounded, convex and weakly* lower semicontinuous function on $\Delta(\Sigma)$ for which (6) holds. Thus Proposition 5 guarantee that $c^\star(\cdot) = \theta R(\cdot\|q)$. Finally, Theorem 11 guarantees that $\succsim$ satisfies A.8. ■

Proof of Proposition 16. First notice that: if J is a set of indexes, $\alpha_j, \beta_j \in \mathbb{R}$ for all $j \in J$, and $\inf_{j\in J}(\alpha_j\theta + \beta_j) > -\infty$ for all $\theta \in [0,1]$, then the function $h(\theta) = \inf_{j\in J}(\alpha_j\theta + \beta_j)$ - being concave and upper semicontinuous - is continuous on $[0,1]$ (see e.g. [10, p. 432]).

In particular, for all $f \in \mathcal{F}$,

$$\begin{aligned}\lim_{\theta\downarrow 0}\min_{p\in\Delta^\sigma(\Sigma,q)}\left(\int u(f)\,dp + \theta R(p\|q)\right) &= \lim_{\theta\downarrow 0}\inf_{p\in\mathrm{Dom}\,R(\cdot\|q)}\left(\int u(f)\,dp + \theta R(p\|q)\right)\\ &= \inf_{p\in\mathrm{Dom}\,R(\cdot\|q)}\int u(f)\,dp,\end{aligned}$$

where $\mathrm{Dom}\,R(\cdot\|q)$ is the effective domain of $R(\cdot\|q)$.

Next we show that, for all $\varphi \in B_0(\Sigma)$,

$$\inf_{p\in\mathrm{Dom}\,R(\cdot\|q)}\int \varphi dp = \inf_{p\in\Delta^\sigma(\Sigma,q)}\int \varphi dp. \tag{42}$$

Say $\varphi = \sum_{j=1}^n \alpha_j 1_{A_j}$ with $\alpha_1, ..., \alpha_n \in \mathbb{R}$ and $A_1, ..., A_n$ a partition of S in Σ. For every $p \in \Delta^\sigma(\Sigma, q)$, define $\psi_p \in B_0(\Sigma)$ by

$$\psi_p(s) = \begin{cases} \dfrac{p(A_j)}{q(A_j)} & \text{if } s \in A_j \text{ and } q(A_j) > 0,\\ 1 & \text{otherwise.}\end{cases}$$

It is easy to see that ψ_p is non-negative and $\int \psi_p dq = 1$. Call p' the element of $\Delta^\sigma(\Sigma, q)$ such that $dp'/dq = \psi_p$. Then $R(p'\|q) < \infty$ (i.e. $p' \in \mathrm{Dom}\,R(\cdot\|q)$) and $\int \varphi dp' = \int \varphi dp$. Therefore

$$\left\{\int \varphi dp : p \in \Delta^\sigma(\Sigma, q)\right\} \subseteq \left\{\int \varphi dp : p \in \mathrm{Dom}\,R(\cdot\|q)\right\},$$

which yields (42) since the converse inclusion is trivial.

For all $f \in \mathcal{F}$, (42) yields

$$\lim_{\theta \downarrow 0} \min_{p \in \Delta^{\sigma}(\Sigma, q)} \left(\int u(f)\, dp + \theta R(p \| q) \right) = \inf_{p \in \Delta^{\sigma}(\Sigma, q)} \int u(f)\, dp.$$

The proof is concluded by showing that

$$\inf_{p \in \Delta^{\sigma}(\Sigma, q)} \int \varphi dp = \operatorname{ess\,inf}_{s \in S} \varphi(s),$$

for all $\varphi \in L^{\infty}(\Sigma, q)$. In fact, $L^{\infty}(\Sigma, q)$ is (isometrically isomorphic to) the norm dual of $ca(\Sigma, q)$.[19] Therefore, for all $\varphi \in L^{\infty}(\Sigma, q)^{-}$ we have

$$\begin{aligned} -\operatorname{ess\,inf}_{s \in S} \varphi(s) &= -\sup \{ t \in \mathbb{R} : q(\{ s \in S : \varphi(s) \geq t \}) = 1 \} \\ &= -\sup \{ t \in \mathbb{R} : q(\{ s \in S : -\varphi(s) \leq -t \}) = 1 \} \\ &= -\sup - \{ t \in \mathbb{R} : q(\{ s \in S : -\varphi(s) \leq t \}) = 1 \} \\ &= \inf \{ t \in \mathbb{R} : q(\{ s \in S : -\varphi(s) \leq t \}) = 1 \} \\ &= \| -\varphi \|_{\infty} \\ &= \sup_{\substack{\mu \in ca(\Sigma, q) \\ \|\mu\|_v = 1}} \int -\varphi d\mu \\ &= \sup_{\substack{\mu \in ca(\Sigma, q) \\ \|\mu\|_v = 1}} \int \psi d\mu, \end{aligned}$$

where $\psi = -\varphi \in L^{\infty}(\Sigma, q)^{+}$. For all $\mu \in ca(\Sigma, q)$ with $\|\mu\|_v = 1$, $\mu = \mu^{+} - \mu^{-}$ for some $\mu^{+}, \mu^{-} \in ca(\Sigma, q)^{+}$ such that $\mu^{+}(S) + \mu^{-}(S) = 1$. Then,

$$\int \psi d\mu = \int \psi d\mu^{+} - \int \psi d\mu^{-} \leq \int \psi d\mu^{+}.$$

If $\mu^{+}(S) = 0$, then $\int \psi d\mu \leq \int \psi dq$; otherwise,

$$\int \psi d\mu \leq \int \psi d\mu^{+} \leq \frac{1}{\mu^{+}(S)} \int \psi d\mu^{+} = \int \psi d\left(\frac{\mu^{+}}{\mu^{+}(S)} \right).$$

In any case, there exists $p \in \Delta^{\sigma}(\Sigma, q)$ such that $\int \psi d\mu \leq \int \psi dp$. We thus conclude that

$$\sup_{\substack{\mu \in ca(\Sigma, q) \\ \|\mu\|_v = 1}} \int \psi d\mu = \sup_{p \in \Delta^{\sigma}(\Sigma, q)} \int \psi dp$$

and

$$\operatorname{ess\,inf}_{s \in S} \varphi(s) = -\sup_{p \in \Delta^{\sigma}(\Sigma, q)} \int -\varphi dp = \inf_{p \in \Delta^{\sigma}(\Sigma, q)} \int \varphi dp.$$

[19] The space of all countably additive set functions on Σ that are absolutely continuous w.r.t. q, endowed with the total variation norm $\|\cdot\|_v$ ([9, p. 176]).

Finally, for each $\varphi \in L^{\infty}(\Sigma, q)$, there exists $c \in \mathbb{R}$ such that $\varphi + c \in L^{\infty}(\Sigma, q)^{-}$, then

$$\begin{aligned}
\operatorname*{ess\,inf}_{s\in S} \varphi(s) &= \operatorname*{ess\,inf}_{s\in S} (\varphi(s) + c) - c \\
&= \left(\inf_{p\in\Delta^{\sigma}(\Sigma,q)} \int (\varphi + c)\, dp \right) - c \\
&= \inf_{p\in\Delta^{\sigma}(\Sigma,q)} \int \varphi dp,
\end{aligned}$$

as wanted. ∎

Proof of Theorem 17. Notice that $M = M(\theta, B_0(\Sigma))$. Equation (31) and Theorem 30.v imply that

$$\int \varphi dq - \frac{\theta}{2} \mathrm{Var}_q(\varphi) = \min_{p\in\Delta^{\sigma}(\Sigma,q)} \left(\int \varphi dp + \frac{1}{2\theta} G(p\|q) \right) \quad \text{for all } \varphi \in M.$$

Therefore the definition of $\succsim^{mv}$, yields

$$f \succsim^{mv} g \Leftrightarrow \min_{p\in\Delta^{\sigma}(\Sigma,q)} \left(\int u(f)\, dp + \frac{1}{2\theta} G(p\|q) \right) \geq \min_{p\in\Delta^{\sigma}(\Sigma,q)} \left(\int u(g)\, dp + \frac{1}{2\theta} G(p\|q) \right)$$

for all $f, g \in \mathcal{G}$.

Consider the preference defined on $\mathcal{F}$ by

$$f \succsim g \Leftrightarrow \min_{p\in\Delta^{\sigma}(\Sigma,q)} \left(\int u(f)\, dp + \frac{1}{2\theta} G(p\|q) \right) \geq \min_{p\in\Delta^{\sigma}(\Sigma,q)} \left(\int u(g)\, dp + \frac{1}{2\theta} G(p\|q) \right).$$

Clearly $\succsim$ and $\succsim^{mv}$ coincide on $\mathcal{G}$. Moreover, the sets $\{p \in \Delta(\Sigma) : G(p\|q) \leq t\}$ are weakly compact subsets of $\Delta^{\sigma}(\Sigma)$ for all $t \geq 0$, by Theorem 30.vii. Therefore $(1/2\theta)\, G(\cdot\|q)$ is a non-negative, grounded, convex and weakly* lower semicontinuous function on $\Delta(\Sigma)$ for which (6) holds. Theorem 2 guarantees that $\succsim$ is a variational preference, while Theorem 11 guarantees that $\succsim$ satisfies A.8.

Notice that $M = \{u(f) + t : f \in \mathcal{G}, t \in \mathbb{R}\}$. In fact, since $u(X)$ is unbounded, for all $\varphi \in M$, there exists $t \in \mathbb{R}$ such that $\varphi - t \in B_0(\Sigma, u(X))$, then there exists $f \in \mathcal{F}$ such that $u(f) = \varphi - t$, since M is a tube, $\varphi - t \in M$, that is $f \in \mathcal{G}$ and $\varphi = u(f) + t$. The converse inclusion is obvious. Set, for all $\varphi \in M$, $I(\varphi) = \min_{p\in\Delta(\Sigma)} \left(\int \varphi dp + (1/2\theta)\, G(p\|q)\right) = \min_{p\in\Delta^{\sigma}(\Sigma,q)} \left(\int \varphi dp + (1/2\theta)\, G(p\|q)\right)$.

Let c be a non-negative, grounded, convex and weakly* lower semicontinuous function on $\Delta(\Sigma)$ such that

$$f \succsim^{mv} g \Leftrightarrow \min_{p\in\Delta(\Sigma)} \left(\int u(f)\, dp + c(p) \right) \geq \min_{p\in\Delta(\Sigma)} \left(\int u(g)\, dp + c(p) \right) \tag{43}$$

for all $f, g \in \mathcal{G}$. Since $X \subseteq \mathcal{G}$, for all $\varphi = u(f)$ with $f \in \mathcal{G}$,

$$\begin{aligned}
I(\varphi) &= I(u(f)) = u(x_f) = \min_{p\in\Delta(\Sigma)} \left(\int u(x_f)\, dp + c(p) \right) \\
&= \min_{p\in\Delta(\Sigma)} \left(\int u(f)\, dp + c(p) \right) = \min_{p\in\Delta(\Sigma)} \left(\int \varphi dp + c(p) \right).
\end{aligned}$$

Then, $I(\varphi) = \min_{p \in \Delta(\Sigma)} \left(\int \varphi dp + c(p) \right)$ for all $\varphi \in M$, and Lemma 27 yields $-c \leq I^\star$ and $-I^\star \leq c$. This implies that $-I^\star$ is the minimal functional satisfying (43). Theorem 30.v implies that $-I^\star(p) = (1/2\theta) G(p \| q)$ for all $p \in \Delta^\sigma(\Sigma, q)$. This concludes the proof. ■

Remark 5 Notice that the ambiguity index $c^\star = -I^\star$ obtained in the above proof also satisfies a "$\mathcal{G}$-restricted version" of (7). In fact,

$$\begin{aligned} c^\star(p) &= \sup_{\varphi \in M} \left(I(\varphi) - \int \varphi dp \right) = \sup_{f \in \mathcal{G}, t \in \mathbb{R}} \left(I(u(f) + t) - \int (u(f) + t) \, dp \right) \\ &= \sup_{f \in \mathcal{G}} \left(I(u(f)) - \int u(f) \, dp \right) = \sup_{f \in \mathcal{G}} \left(u(x_f) - \int u(f) \, dp \right) \end{aligned}$$

for all $p \in \Delta(\Sigma)$.

The simple proof of Corollary 18 is omitted.

D Beyond the Variance: dispersion and variability

The variance is a classic example of a measure of dispersion. The purpose of such measures is to describe how dispersed is a random variable relative to some reference point (the mean in the case of the variance). We now present a general definition of measure of dispersion, that suggests a natural way to generalize mean-variance preferences.

Let $B(\Sigma)$ be the set of all bounded real-valued Σ-measurable functions defined on S. Given a base probability $q \in \Delta^\sigma(\Sigma)$, consider a vertically invariant reference functional $\tau_q : B(\Sigma) \to \mathbb{R}$, like the mean functional $\tau_q(\varphi) = \int \varphi dq$ for each $\varphi \in B(\Sigma)$. A measure of dispersion $H_q : B(\Sigma) \to \mathbb{R}$ about τ_q is a convex functional satisfying the following conditions, discussed in Bickel and Lehmann [4] and [5]:

(i) $H_q(0) = 0$ and $H_q(\varphi + k) = H_q(\varphi)$ for each $\varphi \in B(\Sigma)$ and each constant function $k \in B(\Sigma)$;

(ii) $H_q(\varphi_1) \geq H_q(\varphi_2)$ if $|\varphi_1 - \tau_q(\varphi_1)|$ stochastically dominates $|\varphi_2 - \tau_q(\varphi_2)|$, that is, if $F_{|\varphi_1 - \tau_q(\varphi_1)|}(t) \leq F_{|\varphi_2 - \tau_q(\varphi_2)|}(t)$ for each $t \geq 0$.[20]

For example, important classes of measures of dispersion about the mean $\overline{\varphi} \equiv \int \varphi dq$ are given by:

$$H_q^v(\varphi) = \int_0^1 \left(F^{-1}_{|\varphi - \overline{\varphi}|}(t) \right)^\gamma dt \qquad \text{and} \qquad H_q^\sigma(\varphi) = \left(\int_0^1 \left(F^{-1}_{|\varphi - \overline{\varphi}|}(t) \right)^\gamma dt \right)^{\frac{1}{\gamma}},$$

[20] $F_{|\varphi - \tau_q(\varphi)|} : \mathbb{R}_+ \to [0,1]$ is the cumulative distribution of $|\varphi - \tau_q(\varphi)| \in B(\Sigma)$ w.r.t. the base probability q; i.e., $F_{|\varphi - \tau_q(\varphi)|}(t) = q(\{s \in S : |\varphi(s) - \tau_q(\varphi)| \leq t\})$ for all $t \geq 0$.

where $\gamma \geq 1$ is a positive constant.

Variance $\mathrm{Var}_q(\varphi)$ and standard deviation $SD_q(\varphi)$ are the special cases of, respectively, H_q^v and H_q^σ when $\gamma = 2$. Specifically,

$$\begin{aligned}\mathrm{Var}_q(\varphi) &= \int_0^1 \left(F^{-1}_{|\varphi-\overline{\varphi}|}(t)\right)^2 dt = \int (\varphi(s) - \overline{\varphi})^2 \, dq, \\ SD_q(\varphi) &= \left(\int_0^1 \left(F^{-1}_{|\varphi-\overline{\varphi}|}(t)\right)^2 dt\right)^{\frac{1}{2}} = \sqrt{\int (\varphi(s) - \overline{\varphi})^2 \, dq}.\end{aligned}$$

A preference $\succsim$ on the set of all acts $\mathcal{F}$ on S is a *dispersion preference* if there is a dispersion measure $H_q : B(\Sigma) \to \mathbb{R}$ such that, for all $f, g \in \mathcal{F}$,

$$f \succsim g \Leftrightarrow \int u(f)\, dq - \theta H_q(u(f)) \geq \int u(g)\, dq - \theta H_q(u(g)), \tag{44}$$

where $\theta > 0$ and $u : X \to \mathbb{R}$ is an affine function. For example, the "power" ordering

$$f \succsim g \Leftrightarrow \int u(f)\, dq - \theta \int \left(u(f) - \int u(f)\, dq\right)^\gamma dq \geq \int u(g)\, dq - \theta \int \left(u(g) - \int u(g)\, dq\right)^\gamma dq$$

is a dispersion preference, which for $\gamma = 2$ reduces to the mean-variance preference we studied before.

While dispersion measures describe the dispersion of a random variable about some reference point, variability measures try to describe the "intrinsic" dispersion of a random variable. For example, a classic variability measure is Gini's mean difference:

$$G(\varphi) = \int_{S \times S} (\varphi(s) - \varphi(s'))^2 \, d(q \otimes q).$$

In the same way dispersion measures induce dispersion preferences, variability measures as well can be used to introduce variability preferences. For brevity, here we just give a simple example involving Gini's mean difference.

Example 35 Suppose $S = \{s_1, s_2\}$, and $\Sigma = 2^S$. Given a base probability $q \in \Delta(\Sigma)$, define a preference $\succsim^v$ on $\mathcal{F}$ as follows: for all $f, g \in \mathcal{F}$,

$$f \succsim^v g \Leftrightarrow \int u(f)\, dq - \frac{\theta}{2}(u(f(s_1)) - u(f(s_2))) \geq \int u(g)\, dq - \frac{\theta}{2}(u(g(s_1)) - u(g(s_2))),$$

where $\theta > 0$ and $u : X \to \mathbb{R}$ is an affine function. The preference $\succsim^v$ may not be monotone, unless its domain is suitably restricted. Set

$$M(\theta) = \left\{\varphi \in \mathbb{R}^2 : \varphi_2 - \frac{1 - q_1}{\theta} \leq \varphi_1 \leq \frac{q_1}{\theta} + \varphi_2\right\}.$$

It is easy to check that $\succsim^v$ is a continuous variational preference on $\mathcal{F}_\theta = \{f : u(f) \in M(\theta)\}$, with index of ambiguity aversion $c^\star$ given by:

$$c^\star(p) = -\frac{(p_1 - q_1)^2}{2\theta}.$$

Hence,

$$f \succsim^v g \Leftrightarrow \min_{p \in \Delta^\sigma(\Sigma, q)} \left(\int u(f)\, dp + \frac{1}{2\theta}(p_1 - q_1)^2 \right) \geq \min_{p \in \Delta^\sigma(\Sigma, q)} \left(\int u(g)\, dp + \frac{1}{2\theta}(p_1 - q_1)^2 \right)$$

for all $f, g \in \mathcal{F}$. ▲

References

[1] C.D. Aliprantis and K.C. Border (1999), *Infinite dimensional analysis*, Springer, New York.

[2] F. J. Anscombe and R. J. Aumann, A definition of subjective probability, *Annals of Mathematical Statistics*, 34, 199-205, 1963.

[3] K. Arrow, *Essays in the theory of risk-bearing*, North-Holland, Amsterdam, 1970.

[4] P. J. Bickel and E. L. Lehmann, Descriptive statistics for nonparametric models I: introduction, *Annals of Statistics*, 3, 1038-1044, 1975.

[5] P. J. Bickel and E. L. Lehmann, Descriptive statistics for nonparametric models III: dispersion, *Annals of Statistics*, 4, 1139-1158, 1976.

[6] A. Chateauneuf, F. Maccheroni, M. Marinacci and J-M. Tallon, Monotone continuous multiple priors, *Economic Theory*, forthcoming.

[7] C. Dellacherie and P-A. Meyer, *Probabily and potential C,* North Holland, Amsterdam, 1988.

[8] Dolecki, S. and G.H. Greco (1995). Niveloids, *Topological Methods in Nonlinear Analysis,* 5, 1-22.

[9] N. Dunford and J.T. Schwartz, *Linear Operators, Part I*, Wiley, New York, 1958.

[10] P. Dupuis and R. S. Ellis, *A weak convergence approach to the theory of large deviations*, Wiley, New York, 1997.

[11] D. Ellsberg, Risk, ambiguity, and the Savage axioms, *Quarterly Journal of Economics*, 75, 643-669, 1961.

[12] L.G. Epstein, A definition of uncertainty aversion, *Review of Economic Studies*, 66, 579-608, 1999.

[13] H. Follmer and A. Schied, *Stochastic finance*, de Gruyter, Berlin, 2002.

[14] P. Ghirardato and M. Marinacci, Ambiguity made precise: a comparative foundation, *Journal of Economic Theory*, 102, 251-289, 2002.

[15] I. Gilboa and D. Schmeidler, Maxmin expected utility with a non-unique prior, *Journal of Mathematical Economics*, 18, 141-153, 1989.

[16] P. Groeneboom, J. Oosterhoff, and F. H. Ruymgaart, Large deviation theorems for empirical probability measures, *Annals of Probability*, 7, 553-586, 1979.

[17] L. Hansen and T. Sargent, Wanting robustness in macroeconomics, mimeo, 2000.

[18] L. Hansen and T. Sargent, Robust control and model uncertainty, *American Economic Review*, 91, 60-66, 2001.

[19] I.N. Herstein and J. Milnor, An axiomatic approach to measurable utility, *Econometrica*, 21, 291-297, 1953.

[20] I. Kopylov, Procedural rationality in the multiple prior model, mimeo, 2001.

[21] H. M. Markowitz, Portfolio selection, *Journal of Finance*, 7, 77-91, 1952.

[22] R. R. Phelps, *Convex functions, monotone operators and differentiability*, Springer-Verlag, New York, 1992.

[23] R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, 1970.

[24] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica*, 57, 571-587, 1989.

[25] J. Tobin, Liquidity preference as behavior toward risk, *Review of Economic Studies*, 25, 65-86, 1958.

[26] T. Wang, A class of multi-prior preferences, mimeo, 2003.