Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Making Decisions under Model Misspecification

Simone Cerreia Vioglio, Lars Peter Hansen, Fabio Maccheroni and Massimo Marinacci

Working Paper n. 668

This Version: August 3, 2020

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Making Decisions under Model Misspecification*

Simone Cerreia–Vioglio[a], Lars Peter Hansen[b], Fabio Maccheroni[a] and Massimo Marinacci[a]

[a]Università Bocconi and Igier, [b]University of Chicago

August 3, 2020

Abstract

We use decision theory to confront uncertainty that is sufficiently broad to incorporate "models as approximations." We presume the existence of a featured collection of what we call "structured models" that have explicit substantive motivations. The decision maker confronts uncertainty through the lens of these models, but also views these models as simplifications, and hence, as misspecified. We extend min-max analysis under model ambiguity to incorporate the uncertainty induced by acknowledging that the models used in decision-making are simplified approximations. Formally, we provide an axiomatic rationale for a decision criterion that incorporates model misspecification concerns.

*We thank the audiences at Advances in Decision Analysis 2019, the Blue Collar Working Group 2.0, SAET 2019 and Caltech for their very useful comments. We thank for the financial support the Alfred P. Sloan Foundation (grant G-2018-11113), the ERC (grants SDDM-TEA and INDIMACRO) and a PRIN grant (2017CY2NCA).

Come l'araba fenice:
che vi sia, ciascun lo dice;
dove sia, nessun lo sa.[1]

1 Introduction

The consequences of a decision may depend on exogenous contingencies and uncertain outcomes that are outside the control of a decision maker. This uncertainty takes on many forms. Economic applications typically feature *risk*, where the decision maker knows probabilities but not necessarily outcomes. Statisticians and econometricians have long wrestled with how to confront *ambiguity* over models or unknown parameters within a model. Each model is itself a simplification or an approximation designed to guide or enhance our understanding of some underlying phenomenon of interest. Thus, the model, by its very nature, is *misspecified*, but in typically uncertain ways. How should a decision maker acknowledge model misspecification in a way that guides the use of purposefully simplified models sensibly? This concern has certainly been on the radar screen of statisticians and control theorists, but it has been largely absent in formal approaches to decision theory.[2] Indeed, the statisticians Box and Cox have both stated the challenge succinctly in complementary ways:

> Since all models are wrong, the scientist must be alert to what is importantly wrong. It is inappropriate to be concerned about mice when there are tigers abroad. Box (1976).
>
> ... it does not seem helpful just to say that all models are wrong. The very word "model" implies simplification and idealization. The idea that complex physical, biological or sociological systems can be exactly described by a few formulae is patently absurd. The construction of idealized representations that capture important stable aspects of such systems is, however, a vital part of general scientific analysis and statistical models, especially substantive ones ... Cox (1995).

While there are formulations of decision and control problems that intend to confront model misspecification, the aim of this paper is: (i) to develop an axiomatic approach that will provide a rigorous guide for applications and (ii) to enrich formal decision theory when applied to environments with uncertainty through the guise of models.

[1] "Like the Arabian phoenix: that it exists, everyone says; where it is, nobody knows." A passage from a libretto of Pietro Metastasio.

[2] In Hansen (2014) and Hansen and Marinacci (2016) three kinds of uncertainty are distinguished based on the knowledge of the decision maker, the most challenging being model misspecification viewed as uncertainty induced by the approximate nature of the models under consideration.

In this paper, we explore formally decision making with multiple models, each of which is allowed to be misspecified. We follow Hansen and Sargent (2020) by referring to these multiple models as "structured models." These structured models are ones that are explicitly motivated or featured, such as ones with substantive motivation or scientific underpinnings, consistent with the use of the term "models" by Box and Cox. They may be based on scientific knowledge relying on empirical evidence and theoretical arguments or on revealing parameterizations of probability models with parameters that are interpretable to the decision maker. In posing decision problems formally, it is often assumed, following Wald (1950), that the correct model belongs to the set of models that decision makers posit. The presumption that a decision maker identifies, among their hypotheses, the correct model is often questionable – recalling the initial quotation, the correct model is often a decision maker phoenix. We embrace, rather than push aside, the "models are approximations" perspective of many applied researchers, as articulated by Box, Cox and others. To explore misspecification formally, we introduce a potentially rich collection of probability distributions that depict possible representations of the data without formal substantive motivation. We refer to these as "unstructured models." We use such alternative models as a way to capture how models could be misspecified.[3]

This distinction between structured and unstructured is central to the analysis in this paper and is used to distinguish aversion to ambiguity over models and aversion to potential model misspecification. At a decision-theoretic level, a proper analysis of misspecification concerns has remained elusive so far. Indeed, some of the few studies dealing with economic agents confronting model misspecification still assume that they are conventional expected utility decision makers who treat model misspecification as if it were model ambiguity, despite being aware of a misspecification issue.[4] We extend the analysis of Hansen and Sargent (2020) by providing an axiomatic underpinning for a corresponding decision theory along with a representation of the implied preferences that can guide applications. In so doing, we show an important connection with the analysis of subjective and objective rationality of Gilboa et al. (2010).

Criterion This paper proposes a first decision-theoretic analysis of decision making under model misspecification. We consider a classic setup in the spirit of Wald (1950), but relative to his seminal work we explicitly remove the assumption that the correct model belongs to the set of posited models and we allow for nonneutrality toward this feature. More formally, we assume that decision makers posit a set Q of *structured* (probabilistic) *models* q on states, motivated by their information, but they are afraid that none of them is correct and so face model misspecification. For this reason, decision makers contemplate what we call

[3] Such a distinction is also present in earlier work by Hansen and Sargent (2007) and Hansen and Miao (2018) but without specific reference to the terms "structured" and "unstructured."

[4] See, e.g., Esponda and Pouzo (2016) and Fudenberg et al. (2017).

unstructured models in ranking acts f, according to a conservative decision criterion[5]

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

To interpret this problem, let

$$c_Q(p) = \min_{q \in Q} c(p, q)$$

where we presume that $c_Q(q) = 0$ when $q \in Q$. In this construction, $c_Q(p)$ is a (Hausdorff) distance between a model p and the posited compact set Q of structured models. This distance is nonzero if and only if p is unstructured, that is, $p \notin Q$. More generally, p's that are closer to the set of structured models Q have a less adverse impact on the preferences, as evident by rewriting (1) as:

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

This representation is a special case of the variational representation axiomatized by Maccheroni et al. (2006). The unstructured models are statistical artifacts that allow the decision maker to assess formally the potential consequences of misspecification as captured by the construction of c_Q. In this paper we provide a formal interpretation of c_Q as an index of misspecification fear: the lower the index, the higher the fear.[6]

A protective belt When c takes the entropic form $\lambda R(p||q)$, with $\lambda > 0$, criterion (1) takes the form

$$\min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{q \in Q} R(p||q) \right\} \tag{2}$$

proposed by Hansen and Sargent (2020). It is the most tractable version of criterion (1), which for a singleton Q further reduces to a standard multiplier criterion a la Hansen and Sargent (2001, 2008). By exchanging orders of minimization, we preserve this tractability and provide a revealing link to this earlier research,

$$\min_{q \in Q} \left\{ \min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda R(p||q) \right\} \right\} \tag{3}$$

The inner minimization problem gives rise to the minimization problem featured by Hansen and Sargent (2001, 2008) to confront the potential misspecification of a given probability model q.[7] Unstructured models lack the substantive motivation of structured models, yet in (1) they act as a protective belt against model misspecification. The importance of

[5] Throughout the paper Δ denotes the set of all probabilities (Section 2.1).

[6] To ease terminology, we often refer to "misspecification" rather than "model misspecification."

[7] The Hansen and Sargent (2001, 2008) formulation of preferences builds on extensive literature in control theory starting with Jacobson (1973)'s deterministic robustness criterion and a stochastic extension given by Petersen et al. (2000), among several others.

their role is proportional (as quantified by λ) to their proximity to the set Q, a measure of their plausibility in view of the decision maker information. The outer minimization over structured models is the counterpart to the Wald (1950) and the more general Gilboa and Schmeidler (1989) max-min criterion.

Our analysis provides a decision-theoretic underpinning for incorporating misspecification concerns in a distinct way from ambiguity aversion. Observe that misspecification fear is absent when the index $\min_{q\in Q} c(p,q)$ equals the indicator function δ_Q of the set of structured models Q, that is,

$$\min_{q\in Q} c(p,q) = \begin{cases} 0 & \text{if } p \in Q \\ +\infty & \text{else} \end{cases}$$

In this case, which corresponds to $\lambda = +\infty$ in (2), criterion (1) takes a max-min form:

$$V(f) = \min_{q\in Q} \int u(f)\, dq \tag{4}$$

This max-min criterion thus characterizes decision makers who confront model misspecification but are not concerned by it, so are misspecification neutral (see Section 4.1). The criterion in (1) may thus be viewed as representing decision makers who use a more prudential variational criterion (1) than if they were to max-minimize over the set of structured models which they posited. In particular, the farther away an unstructured model is from the set Q (so the less plausible it is), the less it is weighted in the minimization.

Axiomatics We use the entropic case (2) to outline our axiomatic approach. Start with a singleton $Q = \{q\}$. Decision makers, being afraid that the reference model q might not be correct, contemplate also unstructured models $p \in \Delta$ and rank acts f according to the multiplier criterion

$$V_{\lambda,q}(f) = \min_{p\in\Delta} \left\{ \int u(f)\, dp + \lambda R(p||q) \right\} \tag{5}$$

Here the positive scalar λ is interpreted as an index of misspecification fear. When decision makers posit a nonsingleton set Q of structured models, but are concerned that none of them is correct, then the multiplier criterion (5) gives only an incomplete *dominance relation*:

$$f \succsim^* g \iff V_{\lambda,q}(f) \geq V_{\lambda,q}(g) \qquad \forall q \in Q \tag{6}$$

With (6), decision makers can safely regard f better than g. This type of ranking has, however, little traction because of the incomplete nature of $\succsim^*$. Nonetheless, the burden of choice will have decision makers to select between alternatives, be they rankable by $\succsim^*$ or not. A cautious way to complete the binary relation $\succsim^*$ is given by the preference $\succsim$ represented by (2), or equivalently by (3), that is,

$$V_{\lambda,Q}(f) = \min_{p\in\Delta} \left\{ \int u(f)\, dp + \lambda \min_{q\in Q} R(p||q) \right\} \tag{7}$$

This criterion thus emerges in our analysis as a cautious completion of a multiplier dominance relation $\succsim^*$. Suitably extended to a general preference pair $(\succsim^*, \succsim)$, this approach permits to axiomatize criterion (1) as the representation of the behavioral preference $\succsim$ and the unanimity criterion

$$f \succsim^* g \Longleftrightarrow \min_{p\in\Delta}\left\{\int u(f)\,dp + c(p,q)\right\} \geq \min_{p\in\Delta}\left\{\int u(g)\,dp + c(p,q)\right\} \qquad \forall q \in Q$$

as the representation of the incomplete dominance relation $\succsim^*$.

2 Preliminaries

2.1 Mathematics

Basic notions We consider a non-trivial *event σ-algebra* Σ of subsets of a *state space* S. We denote by $B_0(\Sigma)$ the space of Σ-measurable simple functions $\varphi : S \to \mathbb{R}$, endowed with the supnorm $\| \ \|_\infty$. The dual of $B_0(\Sigma)$ can be identified with the space $ba(\Sigma)$ of all bounded finitely additive measures on (S, Σ).

We denote by Δ the set of probabilities in $ba(\Sigma)$ and endow Δ and any of its subsets with the weak* topology. In particular, Δ^σ denotes the subset of Δ formed by the countably additive probability measures. Given a subset Q in Δ, we denote by $\Delta(Q)$ the collection of all probabilities p which are absolutely continuous with respect to Q, that is, if $A \in \Sigma$ and $q(A) = 0$ for all $q \in Q$, then $p(A) = 0$. Moreover, $\Delta^\sigma(q)$ denotes the set of elements of Δ^σ which are absolutely continuous with respect to a single $q \in \Delta^\sigma$, i.e., $\Delta^\sigma(q) = \{p \in \Delta^\sigma : p \ll q\}$. Unless otherwise specified, throughout all the subsets of Δ are to be intended non-empty.

The (convex analysis) indicator function $\delta_C : \Delta \to [0, \infty]$ of a convex subset C of Δ is defined by

$$\delta_C(p) = \begin{cases} 0 & \text{if } p \in C \\ +\infty & \text{else} \end{cases}$$

Throughout we adopt the convention $0 \cdot \pm\infty = 0$.

The *effective domain* of $f : C \to (-\infty, \infty]$, denoted by $\operatorname{dom} f$, is the set $\{p \in C : f(p) < \infty\}$ where f takes on a finite value. The function f is:

(i) *grounded* if the infimum of its image is 0, i.e., $\inf_{p\in C} f(p) = 0$;

(ii) *strictly convex* if, given any distinct $p, q \in C$, we have $f(\alpha p + (1-\alpha)q) < \alpha f(p) + (1-\alpha) f(q)$ for all $\alpha \in (0,1)$ such that $\alpha p + (1-\alpha) q \in \operatorname{dom} f$.

Divergences and statistical distances Given a non-empty subset Q of Δ, a function $c : \Delta \times Q \to [0, \infty]$ is a *divergence* (*for the set* Q) if

(i) the sections $c_q : \Delta \to [0, \infty]$ are grounded, lower semicontinuous and convex for each $q \in Q$;

(ii) the function $c_Q : \Delta \to [0, \infty]$ defined by $c_Q(\cdot) = \min_{q \in Q} c(\cdot, q)$ is well defined, grounded, lower semicontinuous and convex;

(iii) $c_Q^{-1}(0) = Q$, that is, $c_Q(p) = 0$ if and only if $p \in Q$.

A divergence c that satisfies the distance property

$$c(p, q) = 0 \Longleftrightarrow p = q \tag{8}$$

is called *statistical distance* (*for the set* Q).[8] In particular, $c_Q(p)$ is now an Hausdorff statistical distance between p and Q.

The next lemma provides a simple condition for a function $c : \Delta \times Q \to [0, \infty]$ to be a statistical distance.

Lemma 1 *Let Q be a compact and convex subset of Δ. A jointly lower semicontinuous and convex function $c : \Delta \times Q \to [0, \infty]$ is a statistical distance if and only if it satisfies the distance property (8).*

Given a continuous strictly convex function $\phi : [0, \infty) \to [0, \infty)$ such that $\phi(1) = 0$ and $\lim_{t \to \infty} \phi(t)/t = \infty$, define a *$\phi$-divergence* $D_\phi : \Delta \times \Delta^\sigma \to [0, \infty]$ by

$$D_\phi(p||q) = \begin{cases} \int \phi\left(\dfrac{dp}{dq}\right) dq & \text{if } p \in \Delta^\sigma(q) \\ \infty & \text{otherwise} \end{cases}$$

Here we adopt the conventions $0/0 = 0$ and $\ln 0 = -\infty$.[9] The most important example of a divergence is the *relative entropy* given by $\phi(t) = t \ln t - t + 1$ and denoted by $R(p||q)$.[10] Another important example is the *Gini relative index* given by the quadratic function $\phi(t) = (t-1)^2/2$ and denoted by $\chi^2(p||q)$.

A ϕ-divergence $D_\phi : \Delta \times \Delta^\sigma \to [0, \infty]$ is jointly lower semicontinuous and convex.[11] Next we show that, when suitably restricted, it is a statistical distance, an important property for our purposes.

Lemma 2 *Let Q be a compact and convex subset of Δ^σ. A restricted ϕ-divergence D_ϕ : $\Delta \times Q \to [0, \infty]$ is a statistical distance.*

[8] By a "statistical distance" we do not restrict ourselves to a metric and in particular, given $p, q \in Q$, $c(p, q)$ is not necessarily equal to $c(q, p)$.

[9] The function dp/dq is any version of the Radon-Nikodym derivative of p with respect to q.

[10] Given the conventions $0/0 = 0 \cdot \pm\infty = 0$, it holds $\phi(0) = 0 \ln 0 - 0 + 1 = 0 \cdot -\infty + 1 = 1$.

[11] See Chapter 1 of Liese and Vajda (1987). We refer to this book for properties of ϕ-divergences.

A ϕ-divergence is an instance of a (universal) statistical distance $c : \Delta \times \Delta^\sigma \to [0, \infty]$ whose restriction to each compact and convex subset Q of Δ^σ is a statistical distance for Q.

Finally, given a coefficient $\lambda \in (0, \infty]$, the function

$$\lambda D_\phi : \Delta \times Q \to [0, \infty]$$

is also a statistical distance. Indeed, when $\lambda = \infty$ we have

$$(\infty) D_\phi (p||q) = \delta_{\{q\}} (p) = \begin{cases} 0 & \text{if } p = q \\ \infty & \text{else} \end{cases}$$

because of the convention $0 \cdot \infty = 0$.

2.2 Decision theory

Setup We consider a generalized Anscombe and Aumann (1963) setup where a decision maker chooses among uncertain alternatives described by (simple) acts $f : S \to X$, which are Σ-measurable simple (i.e., finite valued) functions from a *state space* S to a *consequence space* X. This latter set is assumed to be a non-empty convex subset of a vector space (for instance, X is the set of all simple lotteries defined on a prize space). The triple

$$(S, \Sigma, X) \tag{9}$$

forms an (Anscombe-Aumann) *decision framework*.

Let us denote by $\mathcal{F}$ the set of all acts. Given any consequence $x \in X$, we denote by $x \in \mathcal{F}$ also the constant act that takes value x. Thus, with a standard abuse of notation, we identify X with the subset of constant acts in $\mathcal{F}$. Given a function $u : X \to \mathbb{R}$, we denote by $\operatorname{Im} u$ its image. Observe that $u \circ f \in B_0(\Sigma)$ when $f \in \mathcal{F}$.

A preference $\succsim$ is a binary relation on $\mathcal{F}$ that satisfies the so-called *basic conditions* (cf. Gilboa et al., 2010), i.e., it is:

(i) *reflexive* and *transitive*;

(ii) *monotone*: if $f, g \in \mathcal{F}$ and $f(s) \succsim g(s)$ for all $s \in S$, then $f \succsim g$;

(iii) *continuous*: 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;

(iv) *non-trivial*: there exist $f, g \in \mathcal{F}$ such that $f \succ g$.

Moreover, a preference $\succsim$ is *unbounded* if, for each $x, y \in X$ with $x \succ y$, there exist $z, z' \in X$ such that

$$\frac{1}{2}z + \frac{1}{2}y \succsim x \succ y \succsim \frac{1}{2}x + \frac{1}{2}z'$$

Bets are binary acts that play a key role in decision theory. Formally, given any two prizes $x \succ y$, a bet on an event A is the act xAy defined by

$$xAy\,(s) = \begin{cases} x & \text{if } s \in A \\ y & \text{else} \end{cases}$$

In words, a bet on event A is a binary act that yields a more preferred consequence if A obtains.

Comparative uncertainty aversion As in Ghirardato and Marinacci (2002), given two preferences $\succsim_1$ and $\succsim_2$ on $\mathcal{F}$, we say that $\succsim_1$ is *more uncertainty averse than* $\succsim_2$ if, for each consequence $x \in X$ and act $f \in \mathcal{F}$,

$$f \succsim_1 x \implies f \succsim_2 x$$

In words, a preference is more uncertainty averse than another one if, whenever this preference is "bold enough" to prefer an uncertain alternative over a sure one, so does the other one.

Decision criteria A complete preference $\succsim$ on $\mathcal{F}$ is *variational* if it is represented by a decision criterion $V : \mathcal{F} \to \mathbb{R}$ given by

$$V\,(f) = \min_{p \in \Delta} \left\{ \int u\,(f)\,dp + c\,(p) \right\} \tag{10}$$

where the affine utility function u is non-constant and the index of uncertainty aversion $c : \Delta \to [0, \infty]$ is grounded, lower semicontinuous and convex. In particular, given two unbounded variational preferences $\succsim_1$ and $\succsim_2$ on $\mathcal{F}$ that share the same u, but different indexes c_1 and c_2, we have that $\succsim_1$ is more uncertainty averse than $\succsim_2$ if and only if $c_1 \leq c_2$ (see Maccheroni et al., 2006, Propositions 6 and 8).

When the function c has the entropic form $c\,(p, q) = \lambda R\,(p||q)$ with respect to a reference probability $q \in \Delta^\sigma$, criterion (10) takes the *multiplier* form

$$V_{\lambda,q}\,(f) = \min_{p \in \Delta} \left\{ \int u\,(f)\,dp + \lambda R(p||q) \right\}$$

analyzed by Hansen and Sargent (2001, 2008).[12] If, instead, the function c has the indicator form δ_C, with C compact and convex, criterion (10) takes the *max-min* form

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

axiomatized by Gilboa and Schmeidler (1989).

All these criteria are here considered in their classical interpretation, so Waldean for the max-min criterion, in which the elements of Δ are interpreted as models.

[12] Strzalecki (2011) provides the behavioral assumptions that characterize multiplier preferences among variational preferences.

3 Models and preferences

3.1 Models

The consequences of the acts among which decision makers have to choose depend on exogenous states that are outside their control. They know that states obtain according to a probabilistic model described by a probability measure in Δ, the so-called *true* or *correct model*. If decision makers knew the true model, they would confront only risk, which is the randomness inherent to the probabilistic nature of the model. Our decision makers, unfortunately, may not know the true model. Yet, they are able to posit a set of *structured* probabilistic models Q, based on their information (which might well include existing scientific theories, say economic or physical), that form a set of alternative hypotheses regarding the true model. It is a classical assumption, in the spirit of Wald (1950), in which Q is a set of posited hypotheses about the probabilistic behavior of a, natural or social, phenomenon of interest.

A *classical decision framework* is described by a quartet:

$$(S, \Sigma, X, Q) \tag{11}$$

in which a set Q of models is added to a standard decision framework (9). The true model might not be in Q, that is, the decision makers information may be unable to pin it down. Throughout the paper we assume that decision makers know this limitation of their information and so confront model misspecification.[13] This is in contrast with Wald (1950) and most of the subsequent decision-theoretic literature, which assumes that decision makers either know the true model and so face risk or, at least, know that the true model belongs to Q and so face model ambiguity.[14]

In what follows we assume that Q is a compact and convex subset of Δ^σ. As usual, convexity significantly simplifies the analysis. Yet, conceptually it is not an innocuous property: a hybrid model that mixes two structured models can only have a less motivation than either of them. Decision criterion (1), however, accounts for the lower appeal of hybrid models when $c(p,q)$ is also convex in q (as, for instance, when c is a ϕ-divergence). To see why, observe that $\min_{p\in\Delta}\left\{\int u(f)\,dp + c(p,q)\right\}$ is, for each act f, convex in q. In turn, this implies that hybrid models negatively affect criterion criterion (1). This negative impact of mixing thus features an "aversion to model hybridization" attitude, behaviorally captured by axiom A.7. Remarkably, (2) the relative entropy criterion turns out to be neutral to model hybridization. In this important special case, the assumption of convexity of Q is actually without any loss of generality (as Appendix A.1.3 clarifies).

[13] Aydogan et al. (2018) propose an experimental setting that reveals the relevance of model misspecification for decision making.

[14] The model ambiguity (or uncertainty) literature is reviewed in Marinacci (2015).

Convexity of Q can be also justified in a robust Bayesian interpretation of our analysis that regards Q as the set of the so-called predictive distributions, which are combinations of "primitive" models (typically extreme points of Q) weighted according to alternative priors over them. For instance, if the primitive models describe states through i.i.d. processes, the elements of Q describe them via exchangeable processes that combine primitive models and priors (as in the Hewitt and Savage, 1955, version of the de Finetti Representation Theorem). Under this interpretation, the p's are introduced to provide a protective shield for each of the predictive distributions constructed from the alternative priors that are considered.

3.2 Preferences

We consider a two-preference setup, as in Gilboa et al. (2010), with a mental preference $\succsim^*$ and a behavioral preference $\succsim$.

Definition 1 *A preference* $\succsim$ *is (*subjectively*)* rational *if it is:*

a. *complete;*

b. *risk independent: if* $x, y, z \in X$ *and* $\alpha \in (0,1)$*, then* $x \sim y$ *implies* $\alpha x + (1-\alpha)z \sim \alpha y + (1-\alpha)z$.

The behavioral preference $\succsim$ governs the decision maker choice behavior and so it is natural to require it to be complete because, eventually, the decision maker has to choose between alternatives (burden of choice). It is subjectively rational because, in an "argumentative" perspective, the decision maker cannot be convinced that it leads to incorrect choices. Risk independence ensures that $\succsim$ is represented on the space of consequences X by an affine utility function $u : X \to \mathbb{R}$, for instance an expected utility functional when X is the set of simple lotteries. So, risk is addressed in a standard way and we abstract from non-expected utility issues.

The mental preference $\succsim^*$ on $\mathcal{F}$ represents the decision maker "genuine" preference over acts, so it has the nature of a dominance relation for the decision maker. As such, it might well not be complete because of the decision maker inability to compare some pairs of acts.

Definition 2 *A preference* $\succsim^*$ *is a* dominance relation *(or is* objectively rational*) if it is:*

a. *c-complete: if* $x, y \in X$*, then* $x \succsim^* y$ *or* $y \succsim^* x$*;*

b. *weak c-independent: if* $f, g \in \mathcal{F}$*,* $x, y \in X$ *and* $\alpha \in (0,1)$*,*

$$\alpha f + (1-\alpha)x \succsim^* \alpha g + (1-\alpha)x \implies \alpha f + (1-\alpha)y \succsim^* \alpha g + (1-\alpha)y$$

c. convex: if $f, g, h \in \mathcal{F}$ and $\alpha \in (0,1)$,

$$f \succsim^* h \text{ and } g \succsim^* h \implies \alpha f + (1-\alpha) g \succsim^* h$$

If $f \succsim^* g$ we say that f *dominates* g (*strictly* if $f \succ^* g$). The dominance relation is, axiomatically, a variational preference which is not required to be complete.[15] It is objectively rational because the decision maker can convince others of its reasonableness, for instance through arguments based on scientific theories (a case especially relevant for our purposes). Momentarily, axiom A.3 will further clarify its nature.

Along with the classical decision framework (11), the preferences $\succsim^*$ and $\succsim$ form a *two-preference classical decision environment*

$$(S, \Sigma, X, Q, \succsim^*, \succsim) \tag{12}$$

The next two assumptions, which we take from Gilboa et al. (2010), connect the two preferences $\succsim^*$ and $\succsim$.

A.1 *Consistency.* For all $f, g \in \mathcal{F}$,

$$f \succsim^* g \Longrightarrow f \succsim g$$

Consistency asserts that, whenever possible, the mental ranking informs the behavioral one. The next condition says that the decision maker opts, by default, for a sure alternative x over an uncertain one f, unless the dominance relation says otherwise.

A.2 *Caution.* For all $x \in X$ and all $f \in \mathcal{F}$,

$$f \not\succsim^* x \Longrightarrow x \succsim f$$

Unlike the previous assumptions, the next two are peculiar to our analysis. They both link Q to the two preferences $\succsim^*$ and $\succsim$ of the decision maker. We begin with the dominance relation $\succsim^*$. Here we write $f \overset{Q}{=} g$ when $q(f = g) = 1$ for all $q \in Q$, i.e., f and g are equal almost everywhere according to each structured model.

A.3 *Objective Q-coherence.* For all $f, g \in \mathcal{F}$,

$$f \overset{Q}{=} g \implies f \sim^* g$$

and $\succsim^*$ is complete when Q is a singleton.

[15] Convexity is stronger than uncertainty aversion a la Schmeidler (1989), which merely requires that $f \sim^* g$ implies $\alpha f + (1-\alpha) g \succsim^* g$. Yet, under completeness of $\succsim^*$ convexity and uncertainty aversion coincide (see, e.g., Lemma 56 of Cerreia-Vioglio et al., 2011).

This axiom first provides a preferential translation of the special status of structured models over unstructured ones: if they all regard two acts to be almost surely identical, the decision maker "genuine" preference $\succsim^*$ follows suit and ranks them indifferent.

The axiom also disciplines the incompleteness of $\succsim^*$ by requiring that model ambiguity, i.e., a nonsingleton Q, is what underlies it. When Q is a singleton, the two preferences $\succsim^*$ and $\succsim$ agree and yet, because of model misspecification, satisfy only a weak form of independence. In other words, in our approach model misspecification may cause violations of the independence axiom for the dominance relation. Later in the paper, Section 4.2 will further discuss this important feature of our analysis.

To introduce the second assumption, recall that a rational preference $\succsim$ admits an affine utility function $u : X \to \mathbb{R}$ because it satisfies risk independence. This permits to define, given a model $p \in \Delta$, a consequence $x_f^p \in X$ for each act f via the equality

$$u(x_f^p) = \int u(f)\, dp$$

We can interpret x_f^p as the certainty equivalent of act f if p were the correct model. This notion of certainty equivalent permits to relate the posited set of models Q with the behavioral preference $\succsim$, here assumed to be rational.

A.4 *Subjective Q-coherence.* For all $f \in \mathcal{F}$ and all $x \in X$, we have

$$x \succ^* x_f^p \Longrightarrow x \succ f$$

if and only if $p \in Q$.

In words, $p \in \Delta$ is a structured model, so belongs to Q, if and only if decision makers take it seriously, that is, they never choose an act f that would be strictly dominated if p were the correct model. Such a salience of p for the decision makers' preference is the preferential footprint of a structured model, which decision makers take seriously under consideration because of its informational, possibly scientific, status (as opposed to an unstructured model, which decision makers regard as a statistical artifact).

4 Representation with given structured information

We now show how the assumptions on the mental and behavioral preferences permit to characterize criterion (1) for a given set Q, that is, for a DM's given structured information. To this end, we say that a divergence $c : \Delta \times Q \to [0, \infty]$ is *uniquely null* if, for all $(p, q) \in \Delta \times Q$, the sets $c_p^{-1}(0)$ and $c_q^{-1}(0)$ are at most singletons. For instance, statistical distances are easily seen to be uniquely null because of the distance property (8).

We are now ready to state our first representation result.

Theorem 1 *Let* $(S, \Sigma, X, Q, \succsim^*, \succsim)$ *be a two-preference classical decision environment, where* (S, Σ) *is a standard Borel space. The following statements are equivalent:*

(i) $\succsim^*$ *is an unbounded dominance relation and* $\succsim$ *is a rational preference that are both* Q*-coherent and jointly satisfy consistency and caution;*

(ii) there exist an onto affine function $u : X \to \mathbb{R}$ *and a divergence* $c : \Delta \times Q \to [0, \infty]$*, with* $\operatorname{dom} c_Q \subseteq \Delta(Q)$*, such that, for all acts* $f, g \in \mathcal{F}$*,*

$$f \succsim^* g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p, q) \right\} \qquad \forall q \in Q \tag{13}$$

and

$$f \succsim g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + \min_{q \in Q} c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + \min_{q \in Q} c(p, q) \right\} \tag{14}$$

If, in addition, c *is uniquely null, then* $c : \Delta \times Q \to [0, \infty]$ *can be chosen to be a statistical distance.*

This result identifies, in particular, the main preferential assumptions underlying a representation of the type

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \min_{q \in Q} c(p, q) \right\} \tag{15}$$

for the preference $\succsim$. While this representation is of interest for a general divergence with respect to a set Q, it is of particular interest when $c : \Delta \times Q \to [0, \infty]$ is a statistical distance. In this case, the partial ordering $\succsim^*$ is more easily interpreted. Though a technical condition of "unique nullity" is imposed to pin down statistical distances, our representation arguably has more general applicability and captures the preferential underpinning of criterion (15).

The Hausdorff statistical distance $\min_{q \in Q} c(p, q)$ between p and Q is strictly positive if and only if p is an unstructured model, i.e., $p \notin Q$. In particular, the more distant from Q is an unstructured model, the more it is penalized as reflected in the minimization problem that criterion (15) features.

A misspecification index A behavioral preference $\succsim$ represented by (15) is variational with index $\min_{q \in Q} c(p, q)$. So, if two unbounded preferences $\succsim_1$ and $\succsim_2$ represented by (15) share the same u but feature different statistical distances $\min_{q \in Q} c_1(p, q)$ and $\min_{q \in Q} c_2(p, q)$, then $\succsim_1$ is more uncertainty averse than $\succsim_2$ if and only if

$$\min_{q \in Q} c_1(p, q) \leq \min_{q \in Q} c_2(p, q)$$

In the present "classical" setting we interpret this comparative result as saying that the lower is $\min_{q \in Q} c(p,q)$, the higher is the fear of misspecification. We thus regard the function

$$p \mapsto \min_{q \in Q} c(p,q) \tag{16}$$

as an index of aversion to model misspecification and we call it, for short, a *misspecification index*. The lower is this index, the higher is the fear of misspecification.

Specifications and computability Two specifications of our representation are noteworthy. First, when c is the entropic statistical distance $\lambda R(p||q)$, with $\lambda \in (0, \infty]$, we have the following important special case of our representation

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{q \in Q} R(p||q) \right\} \tag{17}$$

which gives tractability to our decision criterion under model misspecification. Specifically, for $\lambda \in (0, \infty)$,[16]

$$\min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{q \in Q} R(p||q) \right\} = \min_{q \in Q} -\lambda \log \int e^{-\frac{u(f).}{\lambda}}\, dq \tag{18}$$

This result is well known when Q is a singleton, that is, when (17) is a standard multiplier criterion.[17]

A second noteworthy special case of our representation is the Gini criterion

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{q \in Q} \chi^2(p||q) \right\}.$$

Remarkably, we have

$$\min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{q \in Q} \chi^2(p||q) \right\} = \min_{q \in Q} \left\{ \int u(f)\, dq - \frac{1}{2\lambda} \mathrm{Var}_q(u(f)) \right\}$$

for all acts f for which the *max-min mean-variance* (in utils) criterion on the r.h.s. is monotone. So, the Gini criterion is a monotone version of the max-min mean-variance criterion.

As to computability, in the important case when criterion (1) features a ϕ-divergence, like the specifications just discussed, we need only to know the set Q to compute it, no integral with respect to unstructured models is needed. This is proved in the next result, a consequence of a duality formula of Ben-Tal and Teboulle (2007).[18]

[16] When $\lambda = \infty$, we have $\min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{q \in Q} R(p||q) \right\} = \min_{q \in Q} \int u(f)\, dq$.

[17] See Appendix A.1.3 for the simple proof of (18).

[18] Here ϕ^* denotes the Fenchel conjugate of ϕ.

Proposition 1 *Given $Q \subseteq \Delta^\sigma$ and $\lambda > 0$, for each act $f \in \mathcal{F}$ it holds*

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{q \in Q} D_\phi(p||q) \right\} = \lambda \inf_{q \in Q} \sup_{\eta \in \mathbb{R}} \left\{ \eta - \int \phi^* \left(\eta - \frac{u(f)}{\lambda} \right) dq \right\}$$

for all $u : X \to \mathbb{R}$.

The r.h.s. formula computes criterion (1) for ϕ-divergences by using only integrals with respect to structured models. This formula substantially simplifies computations and thus confirms the analytical tractability of the previous specifications.

4.1 Interpretation of the decision criterion

In the Introduction we outlined a "protective belt" interpretation of decision criterion (15), i.e.,

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

To elaborate, we begin by observing that the misspecification index (16) has the following bounds

$$0 \leq \min_{q \in Q} c(p, q) \leq \delta_Q(p) \qquad \forall p \in \Delta \tag{19}$$

where δ_Q is the indicator function of the set Q of structured models. So, fear of misspecification is absent when the misspecification index is δ_Q – e.g., when $\lambda = +\infty$ in (17) – in which case criterion (15) takes a Wald (1950) max-min form

$$V(f) = \min_{q \in Q} \int u(f)\, dq \tag{20}$$

This max-min criterion characterizes a decision maker who confronts model misspecification but is not concerned by it. In other words, this Waldean decision maker is a natural candidate to be (model) misspecification neutral. The next limit result further corroborates this insight by showing that, when the fear of misspecification vanishes, the decision maker becomes Waldean.[19]

Proposition 2 *For each act $f \in \mathcal{F}$, we have*

$$\lim_{\lambda \uparrow \infty} \min_{p \in \Delta} \left\{ \int u(f)\, dp + \lambda \min_{q \in Q} R(p||q) \right\} = \min_{q \in Q} \int u(f)\, dq$$

These observations, via bounds and limits, call for a proper decision-theoretic analysis of misspecification neutrality. To this end, note that structured models may be incorrect, yet useful as Box (1976) famously remarked. This motivates the next notion. Recall that act xAy, with $x \succ y$, represents a bet on event A.

[19] To ease matters, we state the result in terms of criterion (17). A general version can be easily established via an increasing sequence of misspecification indexes, with $c_Q^n \leq c_Q^{n+1}$ for each n and $\lim c_Q^n(p) = \infty$ for each $p \notin Q$. For example, $c_Q^n(p) = \lambda_n \min_{q \in Q} D_\phi(p||q)$ where $\lambda_n \uparrow \infty$.

Definition 3 *A preference* $\succsim$ *is* bet-consistent *if, given any* $x \succ y$,

$$q(A) \geq q(B) \quad \forall q \in Q \Longrightarrow xAy \succsim xBy$$

for all events $A, B \in \Sigma$.

Under bet-consistency, a decision maker may fear model misspecification yet regards structured models as good enough to choose to bet on events that they unanimously rank as more likely. Preferences that are bet-consistent can be classified as exhibiting a relatively mild form of fear of model misspecification. The following result shows that an important class of preferences, which includes the ones represented by criterion (17), are bet-consistent.

Proposition 3 *If* $\lambda \in (0, \infty)$ *and* $c = \lambda D_\phi$, *then a preference* $\succsim$ *represented by (15) is bet-consistent.*

Next we substantially strengthen bet-consistency by considering all acts, not just bets.

Definition 4 *A rational preference* $\succsim$ *on* $\mathcal{F}$ *is (*model*)* misspecification neutral *if*

$$\int u(f)\,dq \geq \int u(g)\,dq \quad \forall q \in Q \Longrightarrow f \succsim g$$

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

In this case, a decision maker trusts models enough so to follow them when, if correct, they would unanimously rank pairs of acts. Fear of misspecification thus plays no role in the decision maker preference, so it is decision-theoretically irrelevant. For this reason, the decision maker attitude toward model misspecification can be classified as neutral. The next result shows that this may happen if and only if the decision maker adopts the max-min criterion (20).

Proposition 4 *A preference* $\succsim$ *represented by criterion (15) is misspecification neutral if and only if it is represented by the max-min criterion (20).*

This result provides the sought-after decision-theoretic argument for the interpretation of the max-min criterion as the special case of decision criterion (15) that corresponds to aversion to model ambiguity, with no fear of misspecification. As remarked in the Introduction, it suggests that a decision maker using such a criterion may be viewed as a decision maker who, under model ambiguity, would max-minimize over the set of structured models which she posited but that, for fear of misspecification, ends up using the more prudential variational criterion (15). Unstructured models lack the informational status of structured models, yet in the criterion (15) they act as a "protective belt" against model misspecification.

Note that under this interpretation of criterion (15), the special multiplier case of a singleton $Q = \{q\}$ corresponds to a decision maker who, with no fear of misspecification, would adopt the expected utility criterion $\int u(f)\, dq$. In other words, a singleton Q corresponds to an expected utility decision maker who fears misspecification.

Summing up, in our analysis decision makers adopt the max-min criterion (20) if they either confront only model ambiguity (an information trait) or are averse to model ambiguity (a taste trait) with no fear of model misspecification.

4.2 Interpretation of the dominance relation

As just argued, the singleton $Q = \{q\}$ special case

$$\min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \tag{21}$$

of decision criterion (15) is an expected utility criterion under fear of misspecification (of the unique posited q). Via the relation

$$f \succsim^* g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p, q) \right\} \qquad \forall q \in Q \tag{22}$$

the representation theorem thus clarifies the interpretation of $\succsim^*$ as a dominance relation under model misspecification by showing that it amounts to uniform dominance across all structured models with respect to criterion (21).

It is easy to see that strict dominance amounts to (22), with strict inequality for some $q \in Q$. This observation raises a question: is there a notion of dominance that corresponds to strict inequality for all $q \in Q$? To address this question, we introduce a *strong dominance* relation by writing $f \succ\!\!\succ^* g$ if, for all acts $h, l \in \mathcal{F}$,

$$(1 - \delta) f + \delta h \succ^* (1 - \delta) g + \delta l$$

for all small enough $\delta \in [0, 1]$.[20] By taking $h = f$ and $l = g$, we have the basic implication

$$f \succ\!\!\succ^* g \Longrightarrow f \succ^* g$$

Strong dominance is a strengthening of strict dominance in which the decision maker can convince others "beyond reasonable doubt." The next characterization corroborates this interpretation and, at the same time, answers the previous question in the positive.[21]

Proposition 5 *Let $c : \Delta \times Q \to [0, \infty]$ be a divergence, $u : X \to \mathbb{R}$ an onto and affine function and $\succsim^*$ an unbounded dominance relation represented by (22). For all acts $f, g \in \mathcal{F}$, we have $f \succ\!\!\succ^* g$ if and only if there exists $\varepsilon > 0$ such that*

$$\min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p, q) \right\} + \varepsilon \qquad \forall q \in Q$$

[20] Strong dominance has been introduced by Cerreia-Vioglio et al. (2020).

[21] Up to an ε that ensures a needed uniformity of the strict inequality across structured models.

This characterization shows that $\succ^*$ and $\ggcurly^*$ agree on consequences and, more importantly, that

$$f \ggcurly^* g \Longrightarrow \min_{p\in\Delta}\left\{\int u(f)\,dp + c(p,q)\right\} > \min_{p\in\Delta}\left\{\int u(g)\,dp + c(p,q)\right\} \qquad \forall q \in Q$$

In turn, this easily implies

$$f \ggcurly^* g \Longrightarrow f \succ g \tag{23}$$

We can diagram the relationships among the different dominance notions as follows:

$$\begin{array}{ccccc} \ggcurly^* & \Longrightarrow & \succ^* & \not\Longrightarrow & \succ \\ \Downarrow & & \Downarrow & & \\ \succ & \Longrightarrow & \succsim & & \end{array}$$

An instance when

$$f \succ^* g \implies f \succ g \tag{24}$$

may fail is the max-min criterion (20).

We close by discussing misspecification neutrality, which in view of Proposition 4 is characterized by the misspecification index $\min_{q\in Q} c(p,q) = \delta_Q(p)$.

Lemma 3 *Let c be a statistical distance $c : \Delta\times Q \to [0,\infty]$. We have $\min_{q\in Q} c(p,q) = \delta_Q(p)$ if and only if, for each $q \in Q$, $c(p,q) = \infty$ for all $p \notin Q$.*

In words, misspecification neutrality is characterized by a statistical distance that maximally penalizes unstructured models, which end up playing no role. From a statistical distance angle, this confirms that misspecification neutrality is the attitude of a decision maker who confronts model misspecification, but does not care about it (and so has no use for unstructured models).

This angle becomes relevant here because it shows that, under misspecification neutrality, the representation (22) of the dominance relation becomes

$$f \succsim^* g \Longleftrightarrow \min_{q'\in Q}\left\{\int u(f)\,dq' + c\left(q',q\right)\right\} \geq \min_{q'\in Q}\left\{\int u(g)\,dq' + c\left(q',q\right)\right\} \qquad \forall q \in Q$$

Unstructured models play no role here. Next we show that also statistical distances play no role, so representation (22) further reduces to

$$f \succsim^* g \Longleftrightarrow \int u(f)\,dq \geq \int u(g)\,dq \qquad \forall q \in Q \tag{25}$$

when the dominance relation satisfies the independence axiom. This means, *inter alia*, that fear of model misspecification may cause violations of the independence axiom for such a relation, thus providing a new rationale for violations of this classic axiom.

All this is shown by the next result, which is the version for our setting of the main result of Gilboa et al. (2010).

Proposition 6 *Let* $(S, \Sigma, X, Q, \succsim^*, \succsim)$ *be a two-preference classical decision environment. The following statements are equivalent:*

(i) $\succsim^*$ *is an unbounded dominance relation that satisfies independence and* $\succsim$ *is a rational preference that are both* Q*-coherent and jointly satisfy consistency and caution;*

(ii) there exist an onto affine function $u : X \to \mathbb{R}$ *and a statistical distance* $c : \Delta \times Q \to [0, \infty]$*, with* $c(p, q) = \delta_{\{q\}}(p)$ *for all* $(p, q) \in \Delta \times Q$*, such that (13) and (14) hold, i.e.,*

$$f \succsim^* g \Longleftrightarrow \int u(f)\, dq \geq \int u(g)\, dq \qquad \forall q \in Q$$

and

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

Under independence, the dominance relation $\succsim^*$ thus takes a misspecification neutral form, while the preference $\succsim$ is represented by the max-min criterion.

5 Representation with varying structured information

So far, we carried out our analysis for a given set Q of structured models. Indeed, a two-preference classical decision environment (12) should be more properly written as

$$\left(S, \Sigma, X, Q, \succsim^*_Q, \succsim_Q\right)$$

with the dependence of preferences on Q highlighted. Decision environments, however, may share common state and consequence spaces, but differ on the posited sets of structured models because of different information that decision makers may have. It then becomes important to ensure that decision makers use decision criteria that, across such environments, are consistent.

To address this issue, in this section we consider a family

$$\left\{\left(S, \Sigma, X, Q, \succsim^*_Q, \succsim_Q\right)\right\}_{Q \in \mathcal{Q}}$$

of classical decision environments that differ in the set Q of posited models and we introduce axioms on the family $\left\{\succsim^*_Q\right\}_{Q \in \mathcal{Q}}$ that connect these environments. In keeping with what assumed so far, $\mathcal{Q}$ is the collection of compact and convex subsets of Δ^σ.

A.5 *Monotonicity* (*in model ambiguity*). If $Q' \subseteq Q$ then, for all $f, g \in \mathcal{F}$,

$$f \succsim^*_Q g \Longrightarrow f \succsim^*_{Q'} g$$

According to this axiom, if the "structured" information underlying a set Q is good enough for the decision maker to establish that an act dominates another one, a better information which decreases model ambiguity can only confirm such judgement. Its reversal would be, indeed, at odds with the objective rationality spirit of the dominance relation.

Next we consider a separability assumption.

A.6 *Q-separability*. For all $f, g \in \mathcal{F}$,

$$f \succsim_q^* g \quad \forall q \in Q \implies f \succsim_Q^* g$$

In words, an act dominates another one when it does, separately, through the lenses of each structured model. In this axiom the incompleteness of $\succsim_Q^*$ arises as that of a Paretian order over the, complete but possibly misspecification averse, preferences $\succsim_q^*$ determined by the elements of Q.

These two assumptions, paired with the ones of Theorem 1, guarantee that all dominance relations $\succsim_Q^*$ agree on X. We can thus just write $\succsim^*$, dropping the subscript Q. To state the next axiom, we need a last piece of notation: we denote by $x_{f,q}$ the certainty equivalent of act f for preference $\succsim_q^*$.

A.7 *Model hybridization aversion.* Given any $q, q' \in \Delta^\sigma$,

$$\lambda x_{f,q} + (1-\lambda)\, x_{f,q'} \succsim^* x_{f,\lambda q+(1-\lambda)q'}$$

for all $\lambda \in (0,1)$ and all $f \in \mathcal{F}$.

According to this axiom, the decision maker dislike, *ceteris paribus*, facing a hybrid structured model $\lambda q + (1-\lambda)\, q'$ that, by mixing two structured models q and q', could only have a less substantive motivation (cf. Section 3.1).

We close with a continuity axiom.

A.8 *Lower semicontinuity.* For all $x \in X$ and all $f \in \mathcal{F}$, the set $\{q \in \Delta^\sigma : x \succsim^* x_{f,q}\}$ is closed.

We can now state the extension of Theorem 1 to families of decision environments.

Theorem 2 *Let*

$$\left\{\left(S, \Sigma, X, Q, \succsim_Q^*, \succsim_Q\right)\right\}_{Q \in \mathcal{Q}}$$

be a family of two-preference classical decision environments. The following statements are equivalent:

(i) $\left\{\succsim_Q^*\right\}_{Q\in\mathcal{Q}}$ *is monotone, Q-separable, lower semicontinuous, averse to model hybridization and, for each $Q \in \mathcal{Q}$, the preferences $\succsim_Q^*$ and $\succsim_Q$ satisfy the hypotheses of Theorem 1;*

(ii) there exist an onto affine function $u : X \to \mathbb{R}$ and a jointly lower semicontinuous and convex statistical distance $c : \Delta \times \Delta^{\sigma} \to [0, \infty]$, with $\operatorname{dom} c_Q \subseteq \Delta(Q)$ for all $Q \in \mathcal{Q}$, such that, for all acts $f, g \in \mathcal{F}$,

$$f \succsim_Q^* g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p, q) \right\} \qquad \forall q \in Q \tag{26}$$

and

$$f \succsim_Q g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + \min_{q \in Q} c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + \min_{q \in Q} c(p, q) \right\} \tag{27}$$

Moreover, u is unique up to a positive affine transformation and, given u, c is unique.

This theorem ensures that the decision maker uses consistently criterion (1) across decision environments. In particular, the same statistical distance function is used (e.g., the relative entropy). Moreover, axioms A.5-A.7 further clarify the nature of structured models and their connection with the dominance relation.

Besides its broader scope, this theorem improves Theorem 1 on two counts. First, it features a statistical distance without the need of a unique nullity condition. Second, it contains a sharp uniqueness part. The cost of these improvements is a less parsimonious setting in which the set Q is permitted to vary across the collection $\mathcal{Q}$ of compact and convex subsets of Δ^{σ}.

6 Admissibility

A *two-preference classical decision problem* is a septet

$$\left(F, S, \Sigma, X, Q, \succsim_Q^*, \succsim_Q\right) \tag{28}$$

where $F \subseteq \mathcal{F}$ is a non-empty choice set formed by the acts among which a decision maker has actually to choose, $\succsim_Q^*$ and $\succsim_Q$ are preferences represented by (26) and (27).

Given a compact and convex set Q in Δ^{σ}, the decision maker chooses the best act in F according to $\succsim_Q$. In particular, the *value function* $v : \mathcal{Q} \to (-\infty, \infty]$ is given by

$$v(Q) = \sup_{f \in F} \min_{p \in \Delta} \left\{ \int u(f)\, dp + \min_{q \in Q} c(p, q) \right\} \tag{29}$$

Yet, it is the dominance relation $\succsim_Q^*$ that permits to introduce admissibility.

Definition 5 *An act $f \in F$ is (*weakly*)* admissible *if there is no act $g \in F$ that (strongly) strictly dominates f.*

To relate this notion to the usual notion of admissibility,[22] observe that $g \succ_Q^* f$ amounts to

$$\min_{p\in\Delta}\left\{\int u(g)\,dp + c(p,q)\right\} \geq \min_{p\in\Delta}\left\{\int u(f)\,dp + c(p,q)\right\} \qquad \forall q \in Q$$

with strict inequality for some $q \in Q$. We are thus purposefully defining admissibility in terms of the structured models Q, not the larger class of models Δ, with a model-by-model adjustment for misspecification that makes our notion different from the usual one.

The next result relates optimality and admissibility.

Proposition 7 *Consider a decision problem (28).*

(i) Optimal acts are weakly admissible. They are admissible provided (24) holds.

(ii) Unique optimal acts are admissible.

Optimal acts (if exist) might not be admissible because the max-min nature of decision criterion (15) may lead to violations of (24). Yet, the last result ensures that they belong to the collection of weakly admissible acts

$$F_Q^* = \left\{f \in F : \nexists g \in F, g \succ\!\!\succ_Q^* f\right\}$$

Next we build on this property to establish a comparative statics exercise across decision problems (28) that differ on the posited set Q of structured models.

Proposition 8 *We have*

$$Q \subseteq Q' \Longrightarrow v(Q) \geq v\left(Q'\right)$$

and

$$v(Q) = \max_{f\in F_Q^*} \min_{p\in\Delta}\left\{\int u(f)\,dp + \min_{q\in Q} c(p,q)\right\}$$

provided the sup *in (29) is achieved.*

Smaller sets of structured models are, thus, more valuable. Indeed, in decision problems that feature a larger set of structured models – so, a more discordant information – the decision maker exhibits, *ceteris paribus*, a higher fear of misspecification:

$$Q \subseteq Q' \Longrightarrow \min_{q\in Q} c(p,q) \geq \min_{q\in Q'} c(p,q)$$

In turn, this easily implies $v(Q) \geq v(Q')$.

The decision maker thus dislikes information discordance. In a finite state space,[23] information discordance is maximal, so information is inconclusive, when $Q = \Delta$. Indeed, by the

[22] See, e.g., Ferguson (1967) p. 54.

[23] Infinite state spaces require some technicalities.

distance property (8) we have $\min_{q\in Q} c(p,q) = 0$ for all $p \in \Delta$ if and only if $Q = \Delta$. So, the simplex case represents maximal misspecification fear, given any c (so, any attitude toward model misspecification). Criterion (15) then takes an extreme statewise max-minimization form

$$V(f) = \min_{s\in S} u(f(s))$$

which embodies a form of the precautionary principle that, here, thus emerges out of inconclusive information (e.g., based on inconclusive scientific knowledge). In contrast, information discordance is absent when Q is a singleton.

7 A divergence twist

In our analysis a notion of set divergence naturally arises. Specifically, denoting by $\mathcal{Q}$ the collection of all compact and convex subsets of Δ^σ, say that a function $C : \Delta \times \mathcal{Q} \to [0,\infty]$ is a *set divergence* if

(i) $C(\cdot, Q) : \Delta \to [0,\infty]$ is grounded, lower semicontinuous and convex for each $Q \in \mathcal{Q}$;

(ii) $C(p, Q) = 0$ if and only if $p \in Q$.

If, for each $Q \in \mathcal{Q}$, we consider a lower semicontinuous and convex statistical distance $c : \Delta \times Q \to [0,\infty]$, by Lemma 1 we can define a set divergence by setting $C(p,Q) = \min_{q\in Q} c(p,q)$. In particular, $C(p,\{q\}) = c(p,q)$. This is the Hausdorff-type set divergence that characterizes our decision criterion (1). Yet, for a generic set divergence C, not necessarily pinned down by an underlying statistical distance c, our criterion generalizes to

$$V_Q(f) = \min_{p\in\Delta} \left\{ \int u(f)\, dp + C(p,Q) \right\}$$

Since $C(p,Q) \leq \delta_Q(p)$ for all $p \in \Delta$, this variational criterion still represents a preference that is more uncertainty averse than the max-min one (4). Though the analysis of this general criterion is beyond the scope of this paper, this brief discussion should help to put our exercise in a better perspective.

8 Conclusion

Quantitative researchers use models to enhance their understanding of economic phenomena and to make policy assessments. In essence, each model tells its own quantitative story. We refer to such models as "structured models." Typically, there are more than just one such type of model, with each giving rise to a different quantitative story. Statistical and economic decision theories have addressed how best to confront the ambiguity among structured models. Such structured models are, by their very nature, misspecified. Nevertheless, the decision maker seeks to use such models in sensible ways. This problem is well recognized by applied researchers, but it is typically not part of formal decision theory. In this paper, we extend decision theory to confront model misspecification concerns. In so doing, we recover a variational representation of preferences that includes penalization based on discrepancy measures between "unstructured alternatives" and the set of structured probability models.

A Proofs and related analysis

In the appendix, we provide the proofs of our main results plus some ancillary results. Appendix A.1 contains all the results pertaining statistical ϕ-divergences and distances. Appendix A.3 contains the proofs of our representation results (Theorems 1 and 2). Appendix A.4 contains the proofs of the remaining results.

A.1 Preamble

A.1.1 Proof of Lemma 1

We substantially need to prove that the function $c_Q : \Delta \to [0, \infty]$, defined by $c_Q(p) = \min_{q \in Q} c(p, q)$, is well defined, grounded, lower semicontinuous and convex. This fact follows from the following version of a well known result (see, e.g., Fiacco and Kyparisis, 1986).

Lemma 4 *Let Q be a compact and convex subset of Δ. If $c : \Delta \times Q \to [0, \infty]$ is a jointly lower semicontinuous and convex function such that there exist $\bar{p} \in \Delta$ and $\bar{q} \in Q$ such that $c(\bar{p}, \bar{q}) = 0$, then $c_Q : \Delta \to [0, \infty]$ defined by*

$$c_Q(p) = \min_{q \in Q} c(p, q) \qquad \forall p \in \Delta$$

is well defined, grounded, lower semicontinuous and convex.

Proof Since c is lower semicontinuous and Q is non-empty and compact, c_Q is well defined. Moreover, we have that $0 \geq c(\bar{p}, \bar{q}) \geq c_Q(\bar{p}) \geq 0$, proving that c_Q is grounded. We next show that c_Q is lower semicontinuous. Consider $\tilde{U} = \{p \in \Delta : c_Q(p) > \alpha\}$ where $\alpha \in \mathbb{R}$. If $\tilde{U}$ is empty, then it is open. Otherwise, consider $\bar{p} \in \tilde{U}$. It follows that

$$(\bar{p}, q) \in \left\{ \left(p', q'\right) \in \Delta \times Q : c\left(p', q'\right) > \alpha \right\} = \bar{U} \qquad \forall q \in Q$$

Since c is jointly lower semicontinuous, then $\bar{U}$ is open in the product topology. Thus, for each $q \in Q$ there exist two neighborhoods U_q and V_q such that

$$(\bar{p}, q) \in U_q \times V_q \subseteq \bar{U}$$

Since $q \in V_q$ for all $q \in Q$, we have that $\{V_q\}_{q\in Q}$ is an open cover of Q. Since Q is compact, it admits a finite subcover $\{V_{q_i}\}_{i=1}^n$. Define the open set $U = \cap_{i=1}^n U_{q_i}$. Since $\bar{p} \in U_q$ for all $q \in Q$, note that $\bar{p} \in U$. Consider $p \in U$ and $q' \in Q$. It follows that $q' \in V_{q_i}$ for some $i \in \{1, ..., n\}$. This implies that $(p, q') \in U_{q_i} \times V_{q_i} \subseteq \bar{U}$. We can conclude that $c(p, q') > \alpha$. Since p and q' were arbitrarily chosen in U and Q, we have that $c_Q(p) = \min_{q'\in Q} c(p, q') > \alpha$ for all $p \in U$, proving that $\bar{p} \in U \subseteq \tilde{U}$ and so lower semicontinuity of c_Q.

If $p_1, p_2 \in \Delta$, then define $q_1, q_2 \in Q$ to be such that

$$c(p_1, q_1) = \min_{q\in Q} c(p_1, q) = c_Q(p_1) \text{ and } c(p_2, q_2) = \min_{q\in Q} c(p_2, q) = c_Q(p_2)$$

Consider $\lambda \in (0, 1)$. Define $p_\lambda = \lambda p_1 + (1-\lambda) p_2$ and $q_\lambda = \lambda q_1 + (1-\lambda) q_2 \in Q$. Since c is jointly convex, it follows that

$$\begin{aligned} c_Q(p_\lambda) &= \min_{q\in Q} c(p_\lambda, q) \leq c(p_\lambda, q_\lambda) \leq \lambda c(p_1, q_1) + (1-\lambda) c(p_2, q_2) \\ &= \lambda c_Q(p_1) + (1-\lambda) c_Q(p_2) \end{aligned}$$

proving convexity. ■

Proof of Lemma 1 We first prove the "If" part. We need to prove that c is a divergence that satisfies (8). In particular, we need to show that c_Q and c_q are well defined, grounded, lower semicontinuous and convex for all $q \in Q$. As for c_q, since c is jointly lower semicontinuous and convex, so is c_q and we only need to prove that c_q is grounded. Since $c \geq 0$ satisfies (8), we have that $c_q(q) = c(q, q) = 0$, proving that $c_q \geq 0$ is grounded. By Lemma 4 and since Q is compact and convex and c is jointly lower semicontinuous and convex and such that $c(q, q) = 0$ for all $q \in Q$, then $c_Q : \Delta \to [0, \infty]$ is well defined, grounded, lower semicontinuous and convex. Finally, since c satisfies (8), note that $c_Q(p) = 0$ if and only if $c(p, q) = 0$ for some $q \in Q$ if and only if $p = q$ for some $q \in Q$ if and only if $p \in Q$.

As for the "Only if" part, it is trivial since a statistical distance function, by definition, satisfies (8). ■

A.1.2 Proof of Lemma 2

We actually prove a more complete result.[24] A piece of notation: we write $p \sim Q$ if there exists a control measure $q \in Q$ such that $p \sim q$.[25]

[24] Though a routine result, for the sake of completeness, we provide a proof since we did not find one allowing S to be infinite (see Topsoe, 2001, p. 178 for the finite case).

[25] A probability $q \in Q$ is a *control measure* of Q if $q' \ll q$ for all $q' \in Q$. When Q is a compact and convex subset of Δ^σ, Q has a control measure (see, e.g., Maccheroni and Marinacci, 2001). Such a measure might not be unique, yet any two control measures of Q are equivalent. So, the notion $p \sim Q$ is well defined and independent of the chosen control measure.

Lemma 5 *Let Q be a compact and convex subset of Δ^σ. A restricted ϕ-divergence D_ϕ : $\Delta \times Q \to [0,\infty]$ is a statistical distance. Moreover,*

(i) if $q \in Q$, then $D_\phi(\cdot||q) : \Delta \to [0,\infty]$ is strictly convex;

(ii) if $p \in \Delta^\sigma$ and $p \sim Q$, then $D_\phi(p||\cdot) : Q \to [0,\infty]$ is strictly convex.

Proof It is well known that on $\Delta \times \Delta^\sigma$ the function D_ϕ is jointly lower semicontinuous and convex and satisfies the property

$$D_\phi(p||q) = 0 \iff p = q$$

By Lemma 1, it follows that $D_\phi : \Delta \times Q \to [0,\infty]$ is a statistical distance. We next prove points (i) and (ii).

(i). Consider $q \in Q$. Let $p', p'' \in \Delta$ and $\alpha \in (0,1)$ be such that $p' \neq p''$ and $D_\phi(\alpha p' + (1-\alpha)p''||q) < \infty$. If either $D_\phi(p'||q)$ or $D_\phi(p''||q)$ are not finite, we trivially conclude that $D_\phi(\alpha p' + (1-\alpha)p''||q) < \infty = \alpha D_\phi(p'||q) + (1-\alpha) D_\phi(p''||q)$. Let us then assume that both $D_\phi(p'||q)$ and $D_\phi(p''||q)$ are finite. By definition of D_ϕ and since $\Delta^\sigma(q)$ is convex, this implies that $p', p'' \in \Delta^\sigma(q)$ as well as $\alpha p' + (1-\alpha)p'' \in \Delta^\sigma(q)$. Since p' and p'' are distinct, we have that dp'/dq and dp''/dq take different values on a set of strictly positive q-measure: call it $\tilde{S}$. Since ϕ is strictly convex, it follows that

$$\phi\left(\alpha\frac{dp'}{dq}(s) + (1-\alpha)\frac{dp''}{dq}(s)\right) < \alpha\phi\left(\frac{dp'}{dq}(s)\right) + (1-\alpha)\phi\left(\frac{dp''}{dq}(s)\right) \quad \forall s \in \tilde{S}$$

By definition of D_ϕ, this implies that

$$\begin{aligned}
D_\phi\left(\alpha p' + (1-\alpha)p''||q\right) &= \int_S \phi\left(\frac{d\left[\alpha p' + (1-\alpha)p''\right]}{dq}(s)\right) dq \\
&= \int_S \phi\left(\alpha\frac{dp'}{dq}(s) + (1-\alpha)\frac{dp''}{dq}(s)\right) dq \\
&= \int_{\tilde{S}} \phi\left(\alpha\frac{dp'}{dq}(s) + (1-\alpha)\frac{dp''}{dq}(s)\right) dq \\
&+ \int_{S\setminus\tilde{S}} \phi\left(\alpha\frac{dp'}{dq}(s) + (1-\alpha)\frac{dp''}{dq}(s)\right) dq \\
&< \alpha\int_S \phi\left(\frac{dp'}{dq}(s)\right) dq + (1-\alpha)\int_S \phi\left(\frac{dp''}{dq}(s)\right) dq \\
&= \alpha D_\phi\left(p'||q\right) + (1-\alpha) D_\phi\left(p''||q\right)
\end{aligned}$$

We conclude that $D_\phi(\cdot||q) : \Delta \to [0,\infty]$ is strictly convex.

(ii). Before starting, we make three observations.

a. Since Q is a non-empty, compact and convex subset of Δ^σ, note that there exists $\bar{q} \in Q$ such that $q \ll \bar{q}$ for all $q \in Q$. Since $p \sim Q$, we have that $p \sim \bar{q}$. This implies also that $q \ll p$ for all $q \in Q$.

b. If $q \sim p$, then $(dp/dq)^{-1}$ is well defined almost everywhere (with respect to either p or q) and can be chosen (after defining arbitrarily the function over a set of zero measure) to be the Radon-Nikodym derivative dq/dp.

c. Since ϕ is strictly convex, if we define $\phi^{\star} : (0,\infty) \to [0,\infty)$ by $\phi^{\star}(x) = x\phi(1/x)$ for all $x > 0$, then also $\phi^{\star}$ is strictly convex. By point b, if $p \in \Delta^{\sigma}$ and $q \in Q$ are such that $p \sim q$ and we define $\dot{p} = dp/dq$, then $q(\{\dot{p} = 0\}) = 0 = p(\{\dot{p} = 0\})$ and

$$\begin{aligned} D_{\phi}(p||q) &= \int_S \phi\left(\frac{dp}{dq}\right) dq = \int_{\{\dot{p}=0\}} \phi\left(\frac{dp}{dq}\right) dq + \int_{\{\dot{p}>0\}} \phi\left(\frac{dp}{dq}\right) dq \\ &= \int_{\{\dot{p}>0\}} \phi\left(\frac{1}{\left(\frac{dp}{dq}\right)^{-1}}\right) dq = \int_{\{\dot{p}>0\}} \phi^{\star}\left(\frac{dq}{dp}\right)\frac{dp}{dq} dq \\ &= \int_{\{\dot{p}>0\}} \phi^{\star}\left(\frac{dq}{dp}\right) dp \end{aligned}$$

We can now prove the statement. Let $q', q'' \in Q$ and $\alpha \in (0,1)$ be such that $q' \neq q''$ and $D_{\phi}(p||\alpha q' + (1-\alpha)q'') < \infty$. If either $D_{\phi}(p||q')$ or $D_{\phi}(p||q'')$ are not finite, we trivially conclude that $D_{\phi}(p||\alpha q' + (1-\alpha)q'') < \infty = \alpha D_{\phi}(p||q') + (1-\alpha)D_{\phi}(p||q'')$. Let us then assume that both $D_{\phi}(p||q')$ and $D_{\phi}(p||q'')$ are finite. By definition of D_{ϕ}, we can conclude that $p \ll q'$ and $p \ll q''$. By point a, this yields that $q' \sim p \sim q''$ and $p \sim \alpha q' + (1-\alpha)q''$. Since q' and q'' are distinct, we have that dq'/dp and dq''/dp take different values on a set of strictly positive p-measure: call it $\tilde{S}$. By point c, we have that

$$p\left(\left\{\frac{dp}{d[\alpha q' + (1-\alpha)q'']} = 0\right\}\right) = p\left(\left\{\frac{dp}{dq'} = 0\right\}\right) = p\left(\left\{\frac{dp}{dq''} = 0\right\}\right) = 0$$

Thus, by point c and since dq'/dp and dq''/dp take different values on a set of strictly positive p-measure, there exists a p-measure 1 set $\tilde{S}$ such that

$$\begin{aligned} D_{\phi}\left(p||\alpha q' + (1-\alpha)q''\right) &= \int_{\tilde{S}} \phi^{\star}\left(\frac{d[\alpha q' + (1-\alpha)q'']}{dp}\right) dp \\ &< \alpha \int_{\tilde{S}} \phi^{\star}\left(\frac{dq'}{dp}\right) dp + (1-\alpha)\int_{\tilde{S}} \phi^{\star}\left(\frac{dq''}{dp}\right) dp \\ &= \alpha D_{\phi}\left(p||q'\right) + (1-\alpha)D_{\phi}\left(p||q''\right) \end{aligned}$$

proving point (ii).

A.1.3 Non-convex set of structured models

Let us consider two decision makers who adopt criterion (17), the first one posits a, possibly non-convex, set of structured models Q and the second one posits its closed convex hull $\overline{\mathrm{co}}\, Q$. So, the second decision maker considers also all the mixtures of structured models posited by the first decision maker. Next we show that their preferences over acts actually agree. It is thus without loss of generality to assume that the set of posited structured models is convex, as it was assumed in the main text. Before doing so we prove formula (18). Observe that given a compact subset $Q \subseteq \Delta^{\sigma}$, be that convex or not, we have

$$\begin{aligned}\min_{p\in\Delta}\left\{\int u(f)\,dp+\lambda\min_{q\in Q}R(p||q)\right\} &= \min_{p\in\Delta}\min_{q\in Q}\left\{\int u(f)\,dp+\lambda R(p||q)\right\} \\ &= \min_{q\in Q}\min_{p\in\Delta}\left\{\int u(f)\,dp+\lambda R(p||q)\right\} \\ &= \min_{q\in Q}\phi_{\lambda}^{-1}\left(\int \phi_{\lambda}(u(f))\,dq\right)\end{aligned}$$

where $\phi_{\lambda}(t) = -e^{-\frac{1}{\lambda}t}$ for all $t \in \mathbb{R}$ where $\lambda > 0$.

Proposition 9 *If $Q \subseteq \Delta^{\sigma}$ is compact, then for each $f \in \mathcal{F}$*

$$\min_{p\in\Delta}\left\{\int u(f)\,dp+\lambda\min_{q\in Q}R(p||q)\right\} = \min_{p\in\Delta}\left\{\int u(f)\,dp+\lambda\min_{q\in\overline{\mathrm{co}}\,Q}R(p||q)\right\}$$

Proof First observe that $\overline{\mathrm{co}}\,Q \subseteq \Delta^{\sigma}$. Indeed, since Q is a compact subset of Δ^{σ}, the set function $\nu : \Sigma \to [0,1]$, defined by $\nu(E) = \min_{q\in Q} q(E)$ for all $E \in \Sigma$ is an exact capacity which is continuous at S. This implies that $Q \subseteq \mathrm{core}\,\nu \subseteq \Delta^{\sigma}$, yielding that $\overline{\mathrm{co}}\,Q \subseteq \mathrm{core}\,\nu \subseteq \Delta^{\sigma}$. Given what we have shown before we can conclude that

$$\begin{aligned}\min_{p\in\Delta}\left\{\int u(f)\,dp+\lambda\min_{q\in Q}R(p||q)\right\} &= \min_{q\in Q}\phi_{\lambda}^{-1}\left(\int \phi_{\lambda}(u(f))\,dq\right) \\ &= \phi_{\lambda}^{-1}\left(\min_{q\in Q}\left(\int \phi_{\lambda}(u(f))\,dq\right)\right) \\ &= \phi_{\lambda}^{-1}\left(\min_{q\in\overline{\mathrm{co}}\,Q}\left(\int \phi_{\lambda}(u(f))\,dq\right)\right) \\ &= \min_{q\in\overline{\mathrm{co}}\,Q}\phi_{\lambda}^{-1}\left(\int \phi_{\lambda}(u(f))\,dq\right) \\ &= \min_{p\in\Delta}\left\{\int u(f)\,dp+\lambda\min_{q\in\overline{\mathrm{co}}\,Q}R(p||q)\right\}\end{aligned}$$

proving the statement. ■

A.2 Proof of Proposition 1

The result follows from the following lemma. Here, as usual, ϕ is extended to $\mathbb{R}$ by setting $\phi(t) = +\infty$ if $t \notin [0,+\infty)$. In particular, ϕ^* is non-decreasing.

Lemma 6 *For each $Q \subseteq \Delta^{\sigma}$ and each $\lambda > 0$,*

$$\inf_{p\in\Delta}\left\{\int u(f)\,dp+\lambda\inf_{q\in Q}D_{\phi}(p||q)\right\} = \lambda\inf_{q\in Q}\sup_{\eta\in\mathbb{R}}\left\{\eta-\int\phi^*\left(\eta-\frac{u(f)}{\lambda}\right)dq\right\}$$

for all $u : X \to \mathbb{R}$ and all $f : S \to X$ such that $u \circ f$ is bounded and measurable.

Proof By Theorem 4.2 of Ben-Tal and Teboulle (2007), for each $q \in \Delta^\sigma$ it holds

$$\inf_{p\in\Delta}\left\{\int \xi dp + D_\phi(p||q)\right\} = \sup_{\eta\in\mathbb{R}}\left\{\eta - \int \phi^*\left(\eta - \xi\right) dq\right\}$$

for all $\xi \in L^\infty(q)$. Then, if $u \circ f$ is bounded and measurable, from $u \circ f \in L^\infty(q)$ for all $q \in \Delta^\sigma$, it follows that, for all $\lambda > 0$,

$$\begin{aligned}\inf_{p\in\Delta}\left\{\int u(f)\, dp + \lambda D_\phi(p||q)\right\} &= \lambda \inf_{p\in\Delta}\left\{\int \frac{u(f)}{\lambda} dp + D_\phi(p||q)\right\} \\ &= \lambda \sup_{\eta\in\mathbb{R}}\left\{\eta - \int \phi^*\left(\eta - \frac{u(f)}{\lambda}\right) dq\right\}\end{aligned}$$

for all $\lambda > 0$, as desired. ■

Proof of Proposition 1 In view of the last lemma, it is enough to observe that, if $f : S \to X$ is simple and measurable, then $u \circ f$ is simple and measurable for all $u : X \to \mathbb{R}$. ■

A.3 Representation results

The proof of Theorem 1 is based on four key steps. We first provide a representation for an unbounded and objectively Q-coherent dominance relation $\succsim^*$ (Appendix A.3.1). Second, we provide a representation for a pair of binary relations $(\succsim^*, \succsim)$ which satisfy all of the assumptions of Theorem 1 with the exception of subjective Q-coherence (Appendix A.3.2). Third, we provide two results regarding variational preferences which will help isolate the set of structured models Q in the main representation (Appendix A.3.3). Finally, we merge these three steps to prove our first representation result (Appendix A.3.4). The proof of Theorem 2 instead is presented as one result and it relies on some of the aforementioned results. In what follows, given a function $c : \Delta \times Q \to [0, \infty]$, where Q is a compact and convex subset of Δ^σ, we say that c is a *weak divergence* (*for the set* Q) if it satisfies the first two properties defining a divergence.

A.3.1 A Bewley-type representation

The next result is a multi-utility (variational) representation for unbounded dominance relations.

Lemma 7 *Let $\succsim^*$ be a binary relation on $\mathcal{F}$, where (S, Σ) is a standard Borel space. The following statements are equivalent:*

(i) $\succsim^*$ is an unbounded dominance relation which satisfies objective Q-coherence;

(ii) *there exist an onto affine function* $u : X \to \mathbb{R}$ *and a weak divergence* $c : \Delta \times Q \to [0, \infty]$ *such that* $\operatorname{dom} c(\cdot, q) \subseteq \Delta(Q)$ *for all* $q \in Q$ *and*

$$f \succsim^* g \iff \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p, q) \right\} \quad \forall q \in Q \tag{30}$$

To prove this result, we need to introduce one mathematical object. Let $\succeq^*$ be a binary relation on $B_0(\Sigma)$. We say that $\succeq^*$ is *convex niveloidal* if and only if $\succeq^*$ is a preorder that satisfies the following five properties:

1. For each $\varphi, \psi \in B_0(\Sigma)$ and for each $k \in \mathbb{R}$
$$\varphi \succeq^* \psi \implies \varphi + k \succeq^* \psi + k$$
2. If $\varphi, \psi \in B_0(\Sigma)$ and $\{k_n\}_{n \in \mathbb{N}} \subseteq \mathbb{R}$ are such that $k_n \uparrow k$ and $\varphi - k_n \succeq^* \psi$ for all $n \in \mathbb{N}$, then $\varphi - k \succeq^* \psi$;
3. For each $\varphi, \psi \in B_0(\Sigma)$
$$\varphi \geq \psi \implies \varphi \succeq^* \psi$$
4. For each $k, h \in \mathbb{R}$ and for each $\varphi \in B_0(\Sigma)$
$$k > h \implies \varphi + k \succ^* \varphi + h$$
5. For each $\varphi, \psi, \xi \in B_0(\Sigma)$ and for each $\lambda \in (0, 1)$
$$\varphi \succeq^* \xi \text{ and } \psi \succeq^* \xi \implies \lambda\varphi + (1 - \lambda)\psi \succeq^* \xi$$

Lemma 8 *If* $\succsim^*$ *is an unbounded dominance relation, then there exists an onto affine function* $u : X \to \mathbb{R}$ *such that*

$$x \succsim^* y \iff u(x) \geq u(y) \tag{31}$$

Proof Since $\succsim^*$ is a non-trivial preorder on $\mathcal{F}$ that satisfies c-completeness, continuity and weak c-independence, it is immediate to conclude that $\succsim^*$ restricted to X satisfies weak order, continuity and risk independence.[26] By Herstein and Milnor (1953), it follows that there exists an affine function $u : X \to \mathbb{R}$ that satisfies (31). Since $\succsim^*$ is a non-trivial preorder

[26] To prove that $\succsim^*$ satisfies risk independence, it suffices to deploy the same technique of Lemma 28 of Maccheroni et al. (2006) and observe that $\succsim^*$ is a complete preorder on X. This yields that

$$x \sim^* y \implies \frac{1}{2}x + \frac{1}{2}z \sim^* \frac{1}{2}y + \frac{1}{2}z \quad \forall z \in X$$

By Theorem 2 of Herstein and Milnor (1953) and since $\succsim^*$ satisfies continuity, we can conclude that $\succsim^*$ satisfies risk independence.

on $\mathcal{F}$ that satisfies monotonicity, we have that $\succsim^*$ is non-trivial on X. By Lemma 59 of Cerreia-Vioglio et al. (2011) and since $\succsim^*$ is non-trivial on X and satisfies unboundedness, we can conclude that u is onto. ■

Since u is affine and onto, note that $\{u(f) : f \in \mathcal{F}\} = B_0(\Sigma)$. In light of this observation, we can define a binary relation $\succeq^*$ on $B_0(\Sigma)$ by

$$\varphi \succeq^* \psi \iff f \succsim^* g \text{ where } u(f) = \varphi \text{ and } u(g) = \psi \tag{32}$$

Lemma 9 *If $\succsim^*$ is an unbounded dominance relation, then $\succeq^*$, defined as in (32), is a well defined convex niveloidal binary relation. Moreover, if $\succsim^*$ is objectively Q-coherent, then $\varphi \overset{Q}{=} \psi$ implies $\varphi \sim^* \psi$.*

Proof We begin by showing that $\succeq^*$ is well defined. Assume that $f_1, f_2, g_1, g_2 \in \mathcal{F}$ are such that $u(f_i) = \varphi$ and $u(g_i) = \psi$ for all $i \in \{1, 2\}$. It follows that $u(f_1(s)) = u(f_2(s))$ and $u(g_1(s)) = u(g_2(s))$ for all $s \in S$. By Lemma 8, this implies that $f_1(s) \sim^* f_2(s)$ and $g_1(s) \sim^* g_2(s)$ for all $s \in S$. Since $\succsim^*$ is a preorder that satisfies monotonicity, this implies that $f_1 \sim^* f_2$ and $g_1 \sim^* g_2$. Since $\succsim^*$ is a preorder, if $f_1 \succsim^* g_1$, then

$$f_2 \succsim^* f_1 \succsim^* g_1 \succsim^* g_2 \implies f_2 \succsim^* g_2$$

that is, $f_1 \succsim^* g_1$ implies $f_2 \succsim^* g_2$. Similarly, we can prove that $f_2 \succsim^* g_2$ implies $f_1 \succsim^* g_1$. In other words, $f_1 \succsim^* g_1$ if and only if $f_2 \succsim^* g_2$, proving that $\succeq^*$ is well defined. It is immediate to prove that $\succeq^*$ is a preorder. We next prove properties 1–5.

1. Consider $\varphi, \psi \in B_0(\Sigma)$ and $k \in \mathbb{R}$. Assume that $\varphi \succeq^* \psi$. Let $f, g \in \mathcal{F}$ and $x, y \in X$ be such that $u(f) = 2\varphi$, $u(g) = 2\psi$, $u(x) = 0$ and $u(y) = 2k$. Since u is affine, it follows that
$$\begin{aligned} u\left(\frac{1}{2}f + \frac{1}{2}x\right) &= \frac{1}{2}u(f) + \frac{1}{2}u(x) = \varphi \succeq^* \psi \\ &= \frac{1}{2}u(g) + \frac{1}{2}u(x) = u\left(\frac{1}{2}g + \frac{1}{2}x\right) \end{aligned}$$
proving that $\frac{1}{2}f + \frac{1}{2}x \succsim^* \frac{1}{2}g + \frac{1}{2}x$. Since $\succsim^*$ satisfies weak c-independence and u is affine, we have that $\frac{1}{2}f + \frac{1}{2}y \succsim^* \frac{1}{2}g + \frac{1}{2}y$, yielding that
$$\begin{aligned} \varphi + k &= \frac{1}{2}u(f) + \frac{1}{2}u(y) = u\left(\frac{1}{2}f + \frac{1}{2}y\right) \succeq^* u\left(\frac{1}{2}g + \frac{1}{2}y\right) \\ &= \frac{1}{2}u(g) + \frac{1}{2}u(y) = \psi + k \end{aligned}$$

2. Consider $\varphi, \psi \in B_0(\Sigma)$ and $\{k_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ such that $k_n \uparrow k$ and $\varphi - k_n \succeq^* \psi$ for all $n \in \mathbb{N}$. We have two cases:

(a) $k > 0$. Consider $f, g, h \in \mathcal{F}$ such that

$$u(f) = \varphi,\ u(g) = \varphi - k \text{ and } u(h) = \psi$$

Since $k > 0$ and $k_n \uparrow k$, there exists $\bar{n} \in \mathbb{N}$ such that $k_n > 0$ for all $n \geq \bar{n}$. Define $\lambda_n = 1 - k_n/k$ for all $n \in \mathbb{N}$. It follows that $\lambda_n \in [0, 1]$ for all $n \geq \bar{n}$. Since u is affine, for each $n \geq \bar{n}$

$$u(\lambda_n f + (1 - \lambda_n) g) = \lambda_n u(f) + (1 - \lambda_n) u(g) = \varphi - k_n \succeq^* \psi = u(h)$$

yielding that $\lambda_n f + (1 - \lambda_n) g \succsim^* h$ for all $n \geq \bar{n}$. Since $\succsim^*$ satisfies continuity and $\lambda_n \to 0$, we have that $g \succsim^* h$, that is,

$$\varphi - k = u(g) \succeq^* u(h) = \psi$$

(b) $k \leq 0$. Since $\{k_n\}_{n\in\mathbb{N}}$ is convergent, $\{k_n\}_{n\in\mathbb{N}}$ is bounded. Thus, there exists $h > 0$ such that $k_n + h > 0$ for all $n \in \mathbb{N}$. Moreover, $k_n + h \uparrow k + h > 0$. By point 1, we also have that $\varphi - (k_n + h) = (\varphi - k_n) - h \succeq^* \psi - h$ for all $n \in \mathbb{N}$. By subpoint a, we can conclude that $(\varphi - k) - h = \varphi - (k + h) \succeq^* \psi - h$. By point 1, we obtain that $\varphi - k \succeq^* \psi$.

3. Consider $\varphi, \psi \in B_0(\Sigma)$ such that $\varphi \geq \psi$. Let $f, g \in \mathcal{F}$ be such that $u(f) = \varphi$ and $u(g) = \psi$. It follows that $u(f(s)) \geq u(g(s))$ for all $s \in S$. By Lemma 8, this implies that $f(s) \succsim^* g(s)$ for all $s \in S$. Since $\succsim^*$ satisfies monotonicity, this implies that $f \succsim^* g$, yielding that $\varphi = u(f) \succeq^* u(g) = \psi$.

4. Consider $k, h \in \mathbb{R}$ and $\varphi \in B_0(\Sigma)$. We first assume that $k > h$ and $k = 0$. By point 3, we have that $\varphi = \varphi + k \succeq^* \varphi + h$. By contradiction, assume that $\varphi \not\succ^* \varphi + h$. It follows that $\varphi \sim^* \varphi + h$, yielding that $I = \{w \in \mathbb{R} : \varphi \sim^* \varphi + w\}$ is a non-empty set which contains 0 and h. We next prove that I is an unbounded interval, that is, $I = \mathbb{R}$. First, consider $w_1, w_2 \in I$. Without loss of generality, assume that $w_1 \geq w_2$. By point 3 and since $w_1, w_2 \in I$, we have that for each $\lambda \in (0, 1)$

$$\varphi \succeq^* \varphi + w_1 \succeq^* \varphi + (\lambda w_1 + (1 - \lambda) w_2) \succeq^* \varphi + w_2 \succeq^* \varphi$$

proving that $\varphi \sim^* \varphi + (\lambda w_1 + (1 - \lambda) w_2)$, that is, $\lambda w_1 + (1 - \lambda) w_2 \in I$. Next, we observe that $I \cap (-\infty, 0) \neq \emptyset \neq I \cap (0, \infty)$. Since $h \in I$ and $h < 0$, we have that $I \cap (-\infty, 0) \neq \emptyset$. Since I is an interval and $0, h \in I$, we have that $h/2 \in I$. By point 1 and since $\varphi \sim^* \varphi + h/2$, we have that $\varphi - h/2 \sim^* (\varphi + h/2) - h/2 = \varphi$, proving that $0 < -h/2 \in I \cap (0, \infty)$. By definition of I, note that if $w \in I \backslash \{0\}$, then $\varphi + w \sim^* \varphi$. By point 1 and since $w/2 \in I$ and $\succeq^*$ is a preorder, we have that $(\varphi + w) + w/2 \sim^* \varphi + w/2 \sim^* \varphi$, that is, $\frac{3}{2}w, \frac{1}{2}w \in I$. Since I is an interval, we have that either $\left(\frac{3}{2}w, \frac{1}{2}w\right) \subseteq I$ if $w < 0$ or $\left(\frac{1}{2}w, \frac{3}{2}w\right) \subseteq I$ if $w > 0$. This will help us in proving that I is unbounded from below and above. By contradiction, assume that I is bounded from below and define $m = \inf I$. Since $I \cap (-\infty, 0) \neq \emptyset$, we have that $m < 0$. Consider $\{w_n\}_{n\in\mathbb{N}} \subseteq I \cap (-\infty, 0)$ such that $w_n \downarrow m$. Since $\left(\frac{3}{2}w_n, \frac{1}{2}w_n\right) \subseteq I$ for all $n \in \mathbb{N}$, it follows that $m \leq \frac{3}{2}w_n$ for all $n \in \mathbb{N}$. By passing to the limit, we obtain that $m \leq \frac{3}{2}m < 0$, a contradiction. By contradiction, assume that I is bounded from above and define $M = \sup I$. Since $I \cap (0, \infty) \neq \emptyset$, we have that $M > 0$. Consider $\{w_n\}_{n\in\mathbb{N}} \subseteq I \cap (0, \infty)$ such that $w_n \uparrow M$. Since $\left(\frac{1}{2}w_n, \frac{3}{2}w_n\right) \subseteq I$ for all $n \in \mathbb{N}$, it follows that $M \geq \frac{3}{2}w_n$ for all $n \in \mathbb{N}$. By passing to the limit, we obtain that $M \geq \frac{3}{2}M > 0$, a contradiction. To sum up, I is a non-empty unbounded interval, that is, $I = \mathbb{R}$. This implies that $\varphi \sim^* \varphi + w$ for all $w \in \mathbb{R}$. In particular, select $w_1 = \|\varphi\|_\infty + 1$ and $w_2 = -\|\varphi\|_\infty - 1$. Since $\succeq^*$ is a preorder, we have that $\varphi + w_1 \sim^* \varphi + w_2$. Moreover, $\varphi + w_1 \geq 1 > -1 \geq \varphi + w_2$. By point 3, this implies that $\varphi + w_1 \succeq^* 1 \succeq^* -1 \succeq^* \varphi + w_2$. Since $\succeq^*$ is a preorder and $\varphi + w_1 \sim^* \varphi + w_2$, we can conclude that $1 \sim^* -1$. Note also that there exist $x, y \in X$ such that $u(x) = 1$ and $u(y) = -1$. By Lemma 8, this implies that $x \succ^* y$. By definition of $\succeq^*$ and since $u(x) = 1 \sim^* -1 = u(y)$, we also have that $y \succsim^* x$, a contradiction. Thus, we proved that if $k > h$ and $k = 0$, then $\varphi + k \succ^* \varphi + h$. Assume simply that $k > h$. This implies that $0 > h - k$ and $\varphi \succ^* \varphi + (h - k)$. By point 1, we can conclude that $\varphi + k \succ^* \varphi + (h - k) + k = \varphi + h$.

5. Consider $\varphi, \psi, \xi \in B_0(\Sigma)$ and $\lambda \in (0, 1)$. Assume that $\varphi \succeq^* \xi$ and $\psi \succeq^* \xi$. Let $f, g, h \in \mathcal{F}$ be such that $u(f) = \varphi$, $u(g) = \psi$ and $u(h) = \xi$. By assumption and definition of $\succeq^*$, we have that $f \succsim^* h$ and $g \succsim^* h$. Since $\succsim^*$ satisfies convexity and u is affine, this implies that $\lambda f + (1 - \lambda) g \succsim^* h$, yielding that $\lambda\varphi + (1 - \lambda)\psi = \lambda u(f) + (1 - \lambda) u(g) = u(\lambda f + (1 - \lambda) g) \succeq^* u(h) = \xi$.

Points 1–5 prove the first part of the statement. Finally, consider $\varphi, \psi \in B_0(\Sigma)$. Note that there exist a partition $\{A_i\}_{i=1}^n$ of S and $\{\alpha_i\}_{i=1}^n$ and $\{\beta_i\}_{i=1}^n$ in $\mathbb{R}$ such that

$$\varphi = \sum_{i=1}^n \alpha_i 1_{A_i} \text{ and } \psi = \sum_{i=1}^n \beta_i 1_{A_i}$$

Note that $\{s \in S : \varphi(s) \neq \psi(s)\} = \cup_{i\in\{1,...,n\}:\alpha_i\neq\beta_i} A_i$. Since $\varphi \overset{Q}{=} \psi$, we have that $q(A_i) = 0$ for all $q \in Q$ and for all $i \in \{1, ..., n\}$ such that $\alpha_i \neq \beta_i$. Since u is unbounded, define $\{x_i\}_{i=1}^n \subseteq X$ to be such that $u(x_i) = \alpha_i$ for all $i \in \{1, ..., n\}$. Since u is unbounded, define $\{y_i\}_{i=1}^n \subseteq X$ to be such that $y_i = x_i$ for all $i \in \{1, ..., n\}$ such that $\alpha_i = \beta_i$ and $u(y_i) = \beta_i$ otherwise. Define $f, g : S \to X$ by $f(s) = x_i$ and $g(s) = y_i$ for all $s \in A_i$ and for all $i \in \{1, ..., n\}$. It is immediate to see that $f \overset{Q}{=} g$ as well as $u(f) = \varphi$ and $u(g) = \psi$. Since $\succsim^*$ is objectively Q-coherent, we have that $f \sim^* g$, yielding that $\varphi \sim^* \psi$ and proving the second part of the statement. ■

The next three results (Lemmas 10 and 11 as well as Proposition 10) will help us representing $\succsim^*$. This paired with Lemma 8 and Proposition 11 will yield the proof of Lemma 7.

Lemma 10 *Let $\succsim^*$ be a convex niveloidal binary relation. If $\psi \in B_0(\Sigma)$, then $U(\psi) = \{\varphi \in B_0(\Sigma) : \varphi \succsim^* \psi\}$ is a non-empty convex set such that:*

1. *$\psi \in U(\psi)$;*
2. *if $\varphi \in B_0(\Sigma)$ and $\{k_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ are such that $k_n \uparrow k$ and $\varphi - k_n \in U(\psi)$ for all $n \in \mathbb{N}$, then $\varphi - k \in U(\psi)$;*
3. *if $k > 0$, then $\psi - k \notin U(\psi)$;*
4. *if $\varphi_1 \geq \varphi_2$ and $\varphi_2 \in U(\psi)$, then $\varphi_1 \in U(\psi)$;*
5. *if $k \geq 0$ and $\varphi_2 \in U(\psi)$, then $\varphi_2 + k \in U(\psi)$.*

Proof Since $\succsim^*$ is reflexive, we have that $\psi \in U(\psi)$, proving that $U(\psi)$ is non-empty and point 1. Consider $\varphi_1, \varphi_2 \in U(\psi)$ and $\lambda \in (0,1)$. By definition, we have that $\varphi_1 \succsim^* \psi$ and $\varphi_2 \succsim^* \psi$. Since $\succsim^*$ satisfies convexity, we have that $\lambda\varphi_1 + (1-\lambda)\varphi_2 \succsim^* \psi$, proving convexity of $U(\psi)$. Consider $\varphi \in B_0(\Sigma)$ and $\{k_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ such that $k_n \uparrow k$ and $\varphi - k_n \in U(\psi)$ for all $n \in \mathbb{N}$. It follows that $\varphi - k_n \succsim^* \psi$ for all $n \in \mathbb{N}$, then $\varphi - k \succsim^* \psi$, that is, $\varphi - k \in U(\psi)$, proving point 2. If $k > 0$, then $0 > -k$ and $\psi = \psi + 0 \succ^* \psi - k$, that is, $\psi - k \notin U(\psi)$, proving point 3. Consider $\varphi_1 \geq \varphi_2$ such that $\varphi_2 \in U(\psi)$, then $\varphi_1 \succsim^* \varphi_2$ and $\varphi_2 \succsim^* \psi$, yielding that $\varphi_1 \succsim^* \psi$ and, in particular, $\varphi_1 \in U(\psi)$, proving point 4. Finally, to prove point 5, it is enough to set $\varphi_1 = \varphi_2 + k$ in point 4. ■

Before stating the next result, we define few properties that will turn out to be useful later on. A functional $I : B_0(\Sigma) \to \mathbb{R}$ is:

1. a niveloid if $I(\varphi) - I(\psi) \leq \sup_{s\in S}(\varphi(s) - \psi(s))$ for all $\varphi, \psi \in B_0(\Sigma)$;
2. normalized if $I(k) = k$ for all $k \in \mathbb{R}$;[27]
3. monotone if for each $\varphi, \psi \in B_0(\Sigma)$
$$\varphi \geq \psi \implies I(\varphi) \geq I(\psi)$$
4. $\succsim^*$ consistent if for each $\varphi, \psi \in B_0(\Sigma)$
$$\varphi \succsim^* \psi \implies I(\varphi) \geq I(\psi)$$

[27] With the usual abuse of notation, we denote by k both the real number and the constant function taking value k.

5. concave if for each $\varphi, \psi \in B_0(\Sigma)$ and $\lambda \in (0,1)$

$$I(\lambda\varphi + (1-\lambda)\psi) \geq \lambda I(\varphi) + (1-\lambda) I(\psi)$$

6. translation invariant if for each $\varphi \in B_0(\Sigma)$ and $k \in \mathbb{R}$

$$I(\varphi + k) = I(\varphi) + k$$

Lemma 11 *Let $\succsim^*$ be a convex niveloidal binary relation. If $\psi \in B_0(\Sigma)$, then the functional $I_\psi : B_0(\Sigma) \to \mathbb{R}$, defined by*

$$I_\psi(\varphi) = \max\{k \in \mathbb{R} : \varphi - k \in U(\psi)\} \qquad \forall \varphi \in B_0(\Sigma)$$

is a concave niveloid which is $\succsim^*$ consistent and such that $I_\psi(\psi) = 0$. Moreover, we have that:

1. *The functional $\bar{I}_\psi = I_\psi - I_\psi(0)$ is a normalized concave niveloid which is $\succsim^*$ consistent.*

2. *If $\succsim^*$ satisfies*
$$\psi \overset{Q}{=} \psi' \implies \psi \sim^* \psi'$$
then
$$\psi \overset{Q}{=} \psi' \implies I_\psi = I_{\psi'} \text{ and } \bar{I}_\psi = \bar{I}_{\psi'}$$

Proof Consider $\varphi \in B_0(\Sigma)$. Define $C_\varphi = \{k \in \mathbb{R} : \varphi - k \in U(\psi)\}$. Note that C_φ is non-empty. Indeed, if we set $k = -\|\varphi\|_\infty - \|\psi\|_\infty$, then we obtain that $\varphi - k = \varphi + \|\varphi\|_\infty + \|\psi\|_\infty \geq 0 + \|\psi\|_\infty \geq \psi \in U(\psi)$. By property 4 of Lemma 10, we can conclude that $\varphi - k \in U(\psi)$, that is, $k \in C_\varphi$. Since $U(\psi)$ is convex, it follows that C_φ is an interval. Since $\varphi \in B_0(\Sigma)$, note that there exists $\hat{k} \in \mathbb{R}$ such that $\psi \geq \varphi - \hat{k}$. It follows that $\psi \succsim^* \varphi - \hat{k}$. In particular, we can conclude that $\psi \succ^* \varphi - \left(\hat{k} + \varepsilon\right)$ for all $\varepsilon > 0$. This yields that C_φ is bounded from above. Finally, assume that $\{k_n\}_{n\in\mathbb{N}} \subseteq C_\varphi$ and $k_n \uparrow k$. By property 2 of Lemma 10, we can conclude that $k \in C_\varphi$. To sum up, C_φ is a non-empty bounded from above interval of $\mathbb{R}$ that satisfies the property

$$\{k_n\}_{n\in\mathbb{N}} \subseteq C_\varphi \text{ and } k_n \uparrow k \implies k \in C_\varphi \tag{33}$$

The first part yields that $\sup\{k \in \mathbb{R} : \varphi - k \in U(\psi)\} = \sup C_\varphi \in \mathbb{R}$ is well defined. By (33), we also have that $\sup C_\varphi \in C_\varphi$, that is, $\sup C_\varphi = \max C_\varphi$, proving that I_ψ is well defined. Next, we prove that I_ψ is a concave niveloid. We first show that I_ψ is monotone and translation invariant. By Proposition 2 of Cerreia-Vioglio et al. (2014), this implies that I_ψ is a niveloid. Rather than proving monotonicity, we prove that I_ψ is $\succsim^*$ consistent.[28]

[28] Since if $\varphi_1 \geq \varphi_2$, then $\varphi_1 \succsim^* \varphi_2$, it follows that $\succsim^*$ consistency implies monotonicity.

Consider $\varphi_1, \varphi_2 \in B_0(\Sigma)$ such that $\varphi_1 \succeq^* \varphi_2$. By the properties of $\succeq^*$ and definition of I_ψ, we have that

$$\varphi_1 - I_\psi(\varphi_2) \succeq^* \varphi_2 - I_\psi(\varphi_2) \text{ and } \varphi_2 - I_\psi(\varphi_2) \in U(\psi)$$

and, in particular, $\varphi_2 - I_\psi(\varphi_2) \succeq^* \psi$. Since $\succeq^*$ is a preorder, this implies that $\varphi_1 - I_\psi(\varphi_2) \succeq^* \psi$, that is, $\varphi_1 - I_\psi(\varphi_2) \in U(\psi)$ and $I_\psi(\varphi_2) \in C_{\varphi_1}$, proving that $I_\psi(\varphi_1) \geq I_\psi(\varphi_2)$. We next prove translation invariance. Consider $\varphi \in B_0(\Sigma)$ and $k \in \mathbb{R}$. By definition of I_ψ, we can conclude that

$$(\varphi + k) - (I_\psi(\varphi) + k) = \varphi - I_\psi(\varphi) \in U(\psi)$$

This implies that $I_\psi(\varphi) + k \in C_{\varphi+k}$ and, in particular, $I_\psi(\varphi + k) \geq I_\psi(\varphi) + k$. Since k and φ were arbitrarily chosen, we have that

$$I_\psi(\varphi + k) \geq I_\psi(\varphi) + k \qquad \forall \varphi \in B_0(\Sigma), \forall k \in \mathbb{R}$$

This yields that $I_\psi(\varphi + k) = I_\psi(\varphi) + k$ for all $\varphi \in B_0(\Sigma)$ and for all $k \in \mathbb{R}$.[29]

We move to prove that I_ψ is concave. Consider $\varphi_1, \varphi_2 \in B_0(\Sigma)$ and $\lambda \in (0,1)$. By definition of I_ψ, we have that

$$\varphi_1 - I_\psi(\varphi_1) \in U(\psi) \text{ and } \varphi_2 - I_\psi(\varphi_2) \in U(\psi)$$

Since $U(\psi)$ is convex, we have that

$$\begin{aligned}&(\lambda\varphi_1 + (1-\lambda)\varphi_2) - (\lambda I_\psi(\varphi_1) + (1-\lambda) I_\psi(\varphi_2)) \\ &= \lambda(\varphi_1 - I_\psi(\varphi_1)) + (1-\lambda)(\varphi_2 - I_\psi(\varphi_2)) \in U(\psi)\end{aligned}$$

yielding that $\lambda I_\psi(\varphi_1) + (1-\lambda) I_\psi(\varphi_2) \in C_{\lambda\varphi_1+(1-\lambda)\varphi_2}$ and, in particular, $I_\psi(\lambda\varphi_1 + (1-\lambda)\varphi_2) \geq \lambda I_\psi(\varphi_1) + (1-\lambda) I_\psi(\varphi_2)$.

Finally, since $\psi \in U(\psi)$, note that $0 \in C_\psi$ and $I_\psi(\psi) \geq 0$. By definition of I_ψ, if $I_\psi(\psi) > 0$, then $\psi - I_\psi(\psi) \in U(\psi)$, a contradiction with property 3 of Lemma 10.

1. It is routine to check that $\bar{I}_\psi$ is a normalized concave niveloid which is $\succeq^*$ consistent.

2. Clearly, we have that if $\psi \sim^* \psi'$, then $U(\psi) = U(\psi')$, yielding that $I_\psi = I_{\psi'}$ and, in particular, $I_\psi(0) = I_{\psi'}(0)$ as well as $\bar{I}_\psi = \bar{I}_{\psi'}$. The point trivially follows. ■

Proposition 10 *Let $\succeq^*$ be a binary relation on $B_0(\Sigma)$. The following statements are equivalent:*

(i) $\succeq^*$ is convex niveloidal;

[29] Observe that if $\varphi \in B_0(\Sigma)$ and $k \in \mathbb{R}$, then $-k \in \mathbb{R}$ and

$$I_\psi(\varphi) = I_\psi((\varphi + k) - k) \geq I_\psi(\varphi + k) - k$$

yielding that $I_\psi(\varphi + k) \leq I_\psi(\varphi) + k$.

(ii) there exists a family of concave niveloids $\{I_\alpha\}_{\alpha\in A}$ *on* $B_0(\Sigma)$ *such that*

$$\varphi \succeq^* \psi \iff I_\alpha(\varphi) \geq I_\alpha(\psi) \qquad \forall \alpha \in A \tag{34}$$

(iii) there exists a family of normalized concave niveloids $\{\bar{I}_\alpha\}_{\alpha\in A}$ *on* $B_0(\Sigma)$ *such that*

$$\varphi \succeq^* \psi \iff \bar{I}_\alpha(\varphi) \geq \bar{I}_\alpha(\psi) \qquad \forall \alpha \in A \tag{35}$$

Proof (iii) implies (i). It is trivial.

(i) implies (ii). Let $A = B_0(\Sigma)$. We next show that

$$\varphi_1 \succeq^* \varphi_2 \iff I_\psi(\varphi_1) \geq I_\psi(\varphi_2) \qquad \forall \psi \in B_0(\Sigma)$$

where I_ψ is defined as in Lemma 11 for all $\psi \in B_0(\Sigma)$. By Lemma 11, we have that I_ψ is $\succeq^*$ consistent for all $\psi \in B_0(\Sigma)$. This implies that

$$\varphi_1 \succeq^* \varphi_2 \implies I_\psi(\varphi_1) \geq I_\psi(\varphi_2) \qquad \forall \psi \in B_0(\Sigma)$$

Vice versa, consider $\varphi_1, \varphi_2 \in B_0(\Sigma)$. Assume that $I_\psi(\varphi_1) \geq I_\psi(\varphi_2)$ for all $\psi \in B_0(\Sigma)$. Let $\psi = \varphi_2$. By Lemma 11, we have that

$$I_{\varphi_2}(\varphi_1) \geq I_{\varphi_2}(\varphi_2) = 0$$

yielding that $\varphi_1 \geq \varphi_1 - I_{\varphi_2}(\varphi_1) \in U(\varphi_2)$. By point 4 of Lemma 10, this implies that $\varphi_1 \in U(\varphi_2)$, that is, $\varphi_1 \succeq^* \varphi_2$.

(ii) implies (iii). Given a family of concave niveloids $\{I_\alpha\}_{\alpha\in A}$, define $\bar{I}_\alpha = I_\alpha - I_\alpha(0)$ for all $\alpha \in A$. It is immediate to verify that $\bar{I}_\alpha$ is a normalized concave niveloid for all $\alpha \in A$. It is also immediate to observe that

$$I_\alpha(\varphi_1) \geq I_\alpha(\varphi_2) \quad \forall \alpha \in A \iff \bar{I}_\alpha(\varphi_1) \geq \bar{I}_\alpha(\varphi_2) \quad \forall \alpha \in A$$

proving the implication. ■

Remark 1 Given a convex niveloidal binary relation $\succeq^*$ on $B_0(\Sigma)$, we call *canonical* (resp., *canonical normalized*) the representation $\{I_\psi\}_{\psi\in B_0(\Sigma)}$ (resp., $\{\bar{I}_\psi\}_{\psi\in B_0(\Sigma)}$) obtained from Lemma 11 and the proof of Proposition 10. By the previous proof, clearly, $\{I_\psi\}_{\psi\in B_0(\Sigma)}$ and $\{\bar{I}_\psi\}_{\psi\in B_0(\Sigma)}$ satisfy (34) and (35) respectively.

The next result clarifies what the relation is between any representation of $\succeq^*$ and the canonical ones. This will be useful in establishing an extra property of $\{\bar{I}_\psi\}_{\psi\in B_0(\Sigma)}$ in Corollary 1.

Lemma 12 *Let $\succeq^*$ be a convex niveloidal binary relation. If B is an index set and $\{J_\beta\}_{\beta\in B}$ is a family of normalized concave niveloids such that*

$$\varphi \succeq^* \psi \iff J_\beta(\varphi) \geq J_\beta(\psi) \qquad \forall \beta \in B$$

then for each $\psi \in B_0(\Sigma)$

$$I_\psi(\varphi) = \inf_{\beta\in B} \left(J_\beta(\varphi) - J_\beta(\psi)\right) \qquad \forall \varphi \in B_0(\Sigma) \tag{36}$$

and

$$\bar{I}_\psi(\varphi) = \inf_{\beta\in B} \left(J_\beta(\varphi) - J_\beta(\psi)\right) + \sup_{\beta\in B} J_\beta(\psi) \qquad \forall \varphi \in B_0(\Sigma) \tag{37}$$

Proof Fix $\varphi \in B_0(\Sigma)$ and $\psi \in B_0(\Sigma)$. By definition, we have that

$$I_\psi(\varphi) = \max\{k \in \mathbb{R} : \varphi - k \in U(\psi)\}$$

Since $\{J_\beta\}_{\beta\in B}$ represents $\succeq^*$ and each J_β is translation invariant, note that for each $k \in \mathbb{R}$

$$\begin{aligned}
\varphi - k \in U(\psi) &\iff \varphi - k \succeq^* \psi \iff J_\beta(\varphi - k) \geq J_\beta(\psi) \quad \forall \beta \in B \\
&\iff J_\beta(\varphi) - k \geq J_\beta(\psi) \quad \forall \beta \in B \iff J_\beta(\varphi) - J_\beta(\psi) \geq k \quad \forall \beta \in B \\
&\iff \inf_{\beta\in B} \left(J_\beta(\varphi) - J_\beta(\psi)\right) \geq k
\end{aligned}$$

Since $\varphi - I_\psi(\varphi) \in U(\psi)$, this implies that $I_\psi(\varphi) = \inf_{\beta\in B}(J_\beta(\varphi) - J_\beta(\psi))$. Since φ and ψ were arbitrarily chosen, (36) follows. Since $\bar{I}_\psi = I_\psi - I_\psi(0)$, we only need to compute $-I_\psi(0)$. Since each J_β is normalized, we have that $-I_\psi(0) = -\inf_{\beta\in B}(J_\beta(0) - J_\beta(\psi)) = -\inf_{\beta\in B}(-J_\beta(\psi)) = \sup_{\beta\in B} J_\beta(\psi)$, proving (37). ■

Corollary 1 *If $\succeq^*$ is a convex niveloidal binary relation, then $\bar{I}_0 \leq \bar{I}_\psi$ for all $\psi \in B_0(\Sigma)$.*

Proof By Lemma 12 and Remark 1 and since each $\bar{I}_{\psi'}$ is a normalized concave niveloid, we have that

$$\bar{I}_0(\varphi) = \inf_{\psi'\in B_0(\Sigma)} \left(\bar{I}_{\psi'}(\varphi) - \bar{I}_{\psi'}(0)\right) + \sup_{\psi'\in B_0(\Sigma)} \bar{I}_{\psi'}(0) = \inf_{\psi'\in B_0(\Sigma)} \bar{I}_{\psi'}(\varphi) \leq \bar{I}_\psi(\varphi) \quad \forall \varphi \in B_0(\Sigma)$$

for all $\psi \in B_0(\Sigma)$, proving the statement. ■

The next result will be instrumental in providing a niveloidal multi-representation of $\succsim^*$ when $|Q| \geq 2$. In order to discuss it, we need a piece of terminology. We denote by V the quotient space $B_0(\Sigma)/M$ where M is the vector subspace $\left\{\varphi \in B_0(\Sigma) : \varphi \overset{Q}{=} 0\right\}$. Recall that the elements of V are equivalence classes $[\psi]$ with $\psi \in B_0(\Sigma)$ where $\psi', \psi'' \in [\psi]$ if and only if $\psi \overset{Q}{=} \psi' \overset{Q}{=} \psi''$.

Proposition 11 *If (S,Σ) is a standard Borel space and $|Q|\geq 2$, then there exists a bijection $f:V\rightarrow Q$.*

Proof We begin by observing that:

$$|ca(\Sigma)|\leq|ca_{+}(\Sigma)\times ca_{+}(\Sigma)|=|ca_{+}(\Sigma)|=|(0,\infty)\times\Delta^{\sigma}|=|\Delta^{\sigma}|$$

The first inequality holds because the map $g:ca(\Sigma)\rightarrow ca_{+}(\Sigma)\times ca_{+}(\Sigma)$, defined by $\mu\mapsto(\mu^{+},\mu^{-})$, is injective. Since Σ is non-trivial, $ca_{+}(\Sigma)$ is infinite and a bijection justifying the first equality exists by Theorem 1.4.5 of Srivastava (1998). As to the second equality, the map $g:ca_{+}(\Sigma)\setminus\{0\}\rightarrow(0,\infty)\times\Delta^{\sigma}$, defined by $\mu\mapsto(\mu(S),\mu/\mu(S))$, is a bijection and so $|ca_{+}(\Sigma)\setminus\{0\}|=|(0,\infty)\times\Delta^{\sigma}|$; by Theorem 1.3.1 of Srivastava (1998), $|ca_{+}(\Sigma)|=|ca_{+}(\Sigma)\setminus\{0\}|=|(0,\infty)\times\Delta^{\sigma}|$. As to the last equality, by Theorem 1.4.5 and Exercise 1.5.1 of Srivastava (1998), being $|(0,\infty)|=|(0,1)|\leq|\Delta^{\sigma}|$, we have $|\Delta^{\sigma}|\leq|(0,\infty)\times\Delta^{\sigma}|=|(0,1)\times\Delta^{\sigma}|\leq|\Delta^{\sigma}\times\Delta^{\sigma}|=|\Delta^{\sigma}|$, yielding that $|(0,\infty)\times\Delta^{\sigma}|=|\Delta^{\sigma}|$.

We conclude that $|ca(\Sigma)|\leq|\Delta^{\sigma}|$, that is, there exists an injective map $g:ca(\Sigma)\rightarrow\Delta^{\sigma}$. Since Q is a compact and convex subset of Δ^{σ}, there exists $\bar{q}\in Q$ such that $q\ll\bar{q}$ for all $q\in Q$. We define $h:V\rightarrow ca(\Sigma)$ by

$$h([\psi])(A)=\int_{A}\psi d\bar{q}\qquad\forall A\in\Sigma$$

Note that h is well defined. For, if $\psi'\in[\psi]$, that is, $\psi\overset{Q}{=}\psi'$, then $\psi\overset{\bar{q}}{=}\psi'$, yielding that $\int_{A}\psi d\bar{q}=\int_{A}\psi' d\bar{q}$ for all $A\in\Sigma$. Similarly, $h([\psi])=h([\psi'])$ implies that $\psi\overset{\bar{q}}{=}\psi'$. Since $q\ll\bar{q}$ for all $q\in Q$, this implies that $\psi\overset{Q}{=}\psi'$ and $[\psi]=[\psi']$, proving h is injective. This implies that $\tilde{f}=g\circ h$ is a well defined injective function from V to Δ^{σ}. Clearly, we have that $|\Delta^{\sigma}|\geq\left|\tilde{f}(V)\right|\geq|[0,1]|$. Since (S,Σ) is a standard Borel space and Q is convex and $|Q|\geq 2$, we also have that $|[0,1]|\geq|\Delta^{\sigma}|\geq|Q|\geq[0,1]$. This implies that $|V|=\left|\tilde{f}(V)\right|=|Q|$, proving the statement. ■

Proof of Lemma 7 (ii) implies (i). It is trivial.

(i) implies (ii). Since $\succsim^{*}$ is objectively Q-coherent, if $|Q|=1$, that is $Q=\{\bar{q}\}$, then $\succsim^{*}$ is complete. By Maccheroni et al. (2006) and since $\succsim^{*}$ is unbounded, it follows that there exists an onto and affine $u:X\rightarrow\mathbb{R}$ and a grounded, lower semicontinuous and convex $c_{\bar{q}}:\Delta\rightarrow[0,\infty]$ such that $V:\mathcal{F}\rightarrow\mathbb{R}$ defined by

$$V(f)=\min_{p\in\Delta}\left\{\int u(f)\,dp+c_{\bar{q}}(p)\right\}\qquad\forall f\in\mathcal{F}$$

represents $\succsim^{*}$. If we define $c:\Delta\times Q\rightarrow[0,\infty]$ by $c(p,q)=c_{\bar{q}}(p)$ for all $(p,q)\in\Delta\times Q$, then we have that c is a weak divergence. By Lemma 15 and since $\succsim^{*}$ is objectively Q-coherent, it follows that $c(p,q)=\infty$ for all $p\in\Delta\backslash\Delta(Q)$ and for all $q\in Q$, proving the implication.

Assume $|Q| > 1$. By Lemma 8, there exists an onto affine function $u : X \to \mathbb{R}$ which represents $\succsim^*$ on X. By Lemma 9, this implies that we can consider the convex niveloidal binary relation $\succeq^*$ defined as in (32). By definition of $\succeq^*$ and Proposition 10 (and its proof), we have that

$$f \succsim^* g \iff u(f) \succeq^* u(g) \iff \bar{I}_\psi(u(f)) \geq \bar{I}_\psi(u(g)) \quad \forall \psi \in B_0(\Sigma)$$

where each $\bar{I}_\psi$ is a normalized concave niveloid. As before, consider $V = B_0(\Sigma)/M$ where M is the vector subspace $\left\{\varphi \in B_0(\Sigma) : \varphi \overset{Q}{=} 0\right\}$. For each equivalence class $[\psi]$, select exactly one $\psi' \in B_0(\Sigma)$ such that $\psi' \in [\psi]$. We denote this subset of $B_0(\Sigma)$ by $\tilde{V}$. Clearly, we have that

$$\bar{I}_\psi(u(f)) \geq \bar{I}_\psi(u(g)) \quad \forall \psi \in B_0(\Sigma) \implies \bar{I}_\psi(u(f)) \geq \bar{I}_\psi(u(g)) \quad \forall \psi \in \tilde{V}$$

Vice versa, assume that $\bar{I}_\psi(u(f)) \geq \bar{I}_\psi(u(g))$ for all $\psi \in \tilde{V}$. Consider $\hat{\psi} \in B_0(\Sigma)$. It follows that there exists $[\psi]$ in V such that $\hat{\psi} \in [\psi]$. Similarly, consider $\psi' \in \tilde{V}$ such that $\psi' \in [\psi]$. It follows that $\hat{\psi} \overset{Q}{=} \psi'$. By Lemmas 9 and 11 and since $\succsim^*$ is objectively Q-coherent, then $\bar{I}_{\hat{\psi}} = \bar{I}_{\psi'}$, yielding that $\bar{I}_{\hat{\psi}}(u(f)) \geq \bar{I}_{\hat{\psi}}(u(g))$. Since $\hat{\psi}$ was arbitrarily chosen $\bar{I}_\psi(u(f)) \geq \bar{I}_\psi(u(g))$ for all $\psi \in B_0(\Sigma)$. By construction, observe that there exists a bijection $\tilde{f} : \tilde{V} \to V$. By Proposition 11, we have that there exists a bijection $f : V \to Q$. Define $\bar{f} = f \circ \tilde{f}$. By Corollary 1, if we define $\hat{I}_q = \bar{I}_{\bar{f}^{-1}(q)}$ for all $q \in Q$, then we have that

$$\hat{I}_{\bar{f}(0)} \leq \hat{I}_q \qquad \forall q \in Q$$

and

$$\begin{aligned} f \succsim^* g &\iff \bar{I}_\psi(u(f)) \geq \bar{I}_\psi(u(g)) \quad \forall \psi \in B_0(\Sigma) \iff \bar{I}_\psi(u(f)) \geq \bar{I}_\psi(u(g)) \quad \forall \psi \in \tilde{V} \\ &\iff \hat{I}_q(u(f)) \geq \hat{I}_q(u(g)) \quad \forall q \in Q \end{aligned}$$

Since each $\hat{I}_q$ is a normalized concave niveloid, we have that for each $q \in Q$ there exists a function $c_q : \Delta \to [0, \infty]$ which is grounded, lower semicontinuous, convex and such that

$$\hat{I}_q(\varphi) = \min_{p \in \Delta} \left\{ \int \varphi dp + c_q(p) \right\} \qquad \forall \varphi \in B_0(\Sigma)$$

If we define $c : \Delta \times Q \to [0, \infty]$ by $c(p, q) = c_q(p)$ for all $(p, q) \in \Delta \times Q$, then c satisfies the first property defining a divergence and (30) holds. By Lemma 15 and (30) and since $\succsim^*$ is objectively Q-coherent, it follows that $c(p, q) = \infty$ for all $p \notin \Delta \backslash \Delta(Q)$ and for all $q \in Q$. Finally, recall that

$$c(p, q) = \sup_{\varphi \in B_0(\Sigma)} \left\{ \hat{I}_q(\varphi) - \int \varphi dp \right\} \qquad \forall q \in Q, \forall p \in \Delta$$

Consider $q \in Q$ and $p \in \Delta \backslash \Delta(Q)$. It follows that there exists $A \in \Sigma$ such that $q(A) = 0$ for all $q \in Q$ and $p(A) > 0$. By Lemma 9 and since $\succsim^*$ is objectively Q-coherent, we have that $\lambda 1_A \sim^* 0$ for all $\lambda \in \mathbb{R}$. Since $\hat{I}_{\bar{f}(0)} \leq \hat{I}_q$ all $q \in Q$, we have that for each $q \in Q$

$$c\left(p, \bar{f}(0)\right) = \sup_{\varphi \in B_0(\Sigma)} \left\{ \hat{I}_{\bar{f}(0)}(\varphi) - \int \varphi dp \right\} \leq \sup_{\varphi \in B_0(\Sigma)} \left\{ \hat{I}_q(\varphi) - \int \varphi dp \right\} = c(p, q) \quad \forall p \in \Delta$$

Since $c\left(\cdot, \bar{f}(0)\right)$ is grounded, lower semicontinuous and convex and $\bar{f}(0) \in Q$, this implies that $c_Q(\cdot) = \min_{q \in Q} c(\cdot, q) = c\left(\cdot, \bar{f}(0)\right)$ is well defined and shares the same properties, proving that c is a weak divergence. ■

A.3.2 A parametric representation

Lemma 13 *Let $(\succsim^*, \succsim)$ be two binary relations on $\mathcal{F}$, where (S, Σ) is a standard Borel space. The following statements are equivalent:*

(i) $\succsim^*$ is an unbounded dominance relation satisfying objective Q-coherence and $\succsim$ is a rational preference that jointly satisfy consistency and caution;

(ii) there exist an onto affine function $u : X \to \mathbb{R}$ and a weak divergence $c : \Delta \times Q \to [0, \infty]$ such that $\operatorname{dom} c(\cdot, q) \subseteq \Delta(Q)$ for all $q \in Q$ and

$$f \succsim^* g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p, q) \right\} \qquad \forall q \in Q$$

as well as

$$f \succsim g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + c_Q(p) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c_Q(p) \right\}$$

Proof (i) implies (ii). We proceed by steps. Before starting, we make one observation. By Lemma 7 and since $\succsim^*$ is an unbounded dominance relation which is objectively Q-coherent there exist an onto affine function $u : X \to \mathbb{R}$ and a weak divergence $c : \Delta \times Q \to [0, \infty]$ such that $\succsim^*$ is objectively Q-coherent, it follows that $\operatorname{dom} c(\cdot, q) \subseteq \Delta(Q)$ for all $q \in Q$ and

$$f \succsim^* g \iff \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p, q) \right\} \quad \forall q \in Q$$

We are left to show that $c_Q : \Delta \to [0, \infty]$ is such that

$$f \succsim g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + c_Q(p) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c_Q(p) \right\} \tag{38}$$

To prove this we assume that c is as in the proof of (i) implies (ii) in Lemma 7. This covers both cases $|Q| = 1$ and $|Q| > 1$. In particular, for each $q \in Q$ define $\hat{I}_q : B_0(\Sigma) \to \mathbb{R}$ by

$$\hat{I}_q(\varphi) = \min_{p \in \Delta} \left\{ \int \varphi dp + c(p, q) \right\} \qquad \forall \varphi \in B_0(\Sigma)$$

and recall that there exists $\hat{q}(=\bar{f}(0)$ when $|Q|>1)$ such that $c(\cdot,\hat{q})\le c(\cdot,q)$, thus $\hat{I}_{\hat{q}}\le \hat{I}_q$, for all $q\in Q$.

Step 1. $\succsim$ agrees with $\succsim^*$ on X. In particular, $u:X\to\mathbb{R}$ represents $\succsim^*$ and $\succsim$.

Proof of the Step Note that $\succsim^*$ and $\succsim$ restricted to X are continuous weak orders that satisfy risk independence. Moreover, by the observation above, $\succsim^*$ is represented by u. By Herstein and Milnor (1953) and since $\succsim$ is non-trivial, it follows that there exists a non-constant and affine function $v:X\to\mathbb{R}$ that represents $\succsim$ on X. Since $(\succsim^*,\succsim)$ jointly satisfy consistency, it follows that for each $x,y\in X$

$$u(x)\ge u(y)\Longrightarrow v(x)\ge v(y)$$

By Ghirardato et al. (2004), u and v are equal up to an affine and positive transformation, hence the statement. We can set $v=u$. □

Step 2. There exists a normalized, monotone and continuous functional $I:B_0(\Sigma)\to\mathbb{R}$ such that

$$f\succsim g\Longleftrightarrow I(u(f))\ge I(u(g))$$

Proof of the Step By Cerreia-Vioglio et al. (2011) and since $\succsim$ is a rational preference relation, the statement follows. □

Step 3. $I(\varphi)\le\inf_{q\in Q}\hat{I}_q(\varphi)$ for all $\varphi\in B_0(\Sigma)$.

Proof of the Step Consider $\varphi\in B_0(\Sigma)$. Since each $\hat{I}_q$ is normalized and monotone and u is onto, we have that $\hat{I}_q(\varphi)\in[\inf_{s\in S}\varphi(s),\sup_{s\in S}\varphi(s)]\subseteq \operatorname{Im}u$ for all $q\in Q$. Since $\varphi\in B_0(\Sigma)$, it follows that there exists $f\in\mathcal{F}$ such that $\varphi=u(f)$ and $x\in X$ such that $u(x)=\inf_{q\in Q}\hat{I}_q(\varphi)$. For each $\varepsilon>0$ there exists $x_\varepsilon\in X$ such that $u(x_\varepsilon)=u(x)+\varepsilon$. Since $\inf_{q\in Q}\hat{I}_q(\varphi)=u(x)$, it follows that for each $\varepsilon>0$ there exists $q\in Q$ such that $\hat{I}_q(u(f))=\hat{I}_q(\varphi)<u(x_\varepsilon)=\hat{I}_q(u(x_\varepsilon))$, yielding that $f\not\succsim^* x_\varepsilon$. Since $(\succsim^*,\succsim)$ jointly satisfy caution, we have that $x_\varepsilon\succsim f$ for all $\varepsilon>0$. By Step 2, this implies that

$$u(x)+\varepsilon=u(x_\varepsilon)=I(u(x_\varepsilon))\ge I(u(f))=I(\varphi)\quad\forall\varepsilon>0$$

that is, $\inf_{q\in Q}\hat{I}_q(\varphi)=u(x)\ge I(\varphi)$, proving the step. □

Step 4. $I(\varphi)\ge\inf_{q\in Q}\hat{I}_q(\varphi)$ for all $\varphi\in B_0(\Sigma)$.

Proof of the Step Consider $\varphi\in B_0(\Sigma)$. We use the same objects and notation of Step 3. Note that for each $q'\in Q$

$$\hat{I}_{q'}(u(f))=\hat{I}_{q'}(\varphi)\ge\inf_{q\in Q}\hat{I}_q(\varphi)=u(x)=\hat{I}_{q'}(u(x))$$

that is, $f \succsim^* x$. Since $(\succsim^*, \succsim)$ jointly satisfy consistency, we have that $f \succsim x$. By Step 2, this implies that

$$I(\varphi) = I(u(f)) \geq I(u(x)) = u(x) = \inf_{q \in Q} \hat{I}_q(\varphi)$$

proving the step. $\square$

Step 5. $I(\varphi) = \min_{p\in\Delta} \left\{ \int \varphi dp + c_Q(p) \right\}$ *for all* $\varphi \in B_0(\Sigma)$.

Proof of the Step By Steps 3 and 4 and since $\hat{I}_{\hat{q}} \leq \hat{I}_q$ for all $q \in Q$, we have that

$$I(\varphi) = \min_{q\in Q} \hat{I}_q(\varphi) = \hat{I}_{\hat{q}}(\varphi) \qquad \forall \varphi \in B_0(\Sigma)$$

Since $c(\cdot, \hat{q}) = c_Q(\cdot)$, it follows that for each $\varphi \in B_0(\Sigma)$

$$I(\varphi) = \hat{I}_{\hat{q}}(\varphi) = \min_{p\in\Delta} \left\{ \int \varphi dp + c(p, \hat{q}) \right\} = \min_{p\in\Delta} \left\{ \int \varphi dp + c_Q(p) \right\}$$

proving the step. $\square$

Thus, (38) follows from Steps 2 and 5, this completes the proof.

(ii) implies (i). It is routine. $\blacksquare$

A.3.3 Two variational lemmas

The next two lemmas will be key in characterizing subjective and objective Q-coherence.

Lemma 14 *Let* $\succsim$ *be a variational preference represented by* $V : \mathcal{F} \to \mathbb{R}$ *defined by*

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

and let $\bar{p} \in \Delta$. *If* $\succsim$ *is unbounded, then the following conditions are equivalent:*

(i) $c(\bar{p}) = 0$;

(ii) $x_f^{\bar{p}} \succsim f$ *for all* $f \in \mathcal{F}$;

(iii) for each $f \in \mathcal{F}$ *and for each* $x \in X$

$$x \succ x_f^{\bar{p}} \implies x \succ f.$$

Proof We actually prove that (i)$\Longrightarrow$(ii)$\Longleftrightarrow$(iii), with equivalence when $\succsim$ is unbounded.

(i) implies (ii). Let $f \in \mathcal{F}$. It is enough to observe that $c(\bar{p}) = 0$ implies

$$V\left(x_f^{\bar{p}}\right) = u\left(x_f^{\bar{p}}\right) = \int u(f)\, d\bar{p} + c(\bar{p}) \geq \min_{p\in\Delta} \left\{ \int u(f)\, dp + c(p) \right\} = V(f)$$

yielding that $x_f^{\bar{p}} \succsim f$.

(ii) implies (iii). Assume that $x_f^{\bar{p}} \succsim f$ for all $f \in \mathcal{F}$. Since $\succsim$ is complete and transitive, it follows that if $x \succ x_f^{\bar{p}}$, then $x \succ f$.

(iii) implies (ii). By contradiction, suppose that there exists $f \in \mathcal{F}$ such that $f \succ x_f^{\bar{p}}$. Let $x_f \in X$ be such that $x_f \sim f$. This implies that $x_f \succ x_f^{\bar{p}}$ and so $x_f \succ f$, a contradiction.

(ii) implies (i). Let $\succsim$ be unbounded. Assume that $x_f^{\bar{p}} \succsim f$ for all $f \in \mathcal{F}$, i.e., $V(f) \leq \int u(f)\, d\bar{p}$ for all $f \in \mathcal{F}$. So, $\bar{p}$ corresponds to a SEU preference that is more ambiguity averse than $\succsim$. By Lemma 32 of Maccheroni et al. (2006), we can conclude that $c(\bar{p}) = 0$. ■

Lemma 15 *Let $\succsim$ be a variational preference represented by $V : \mathcal{F} \to \mathbb{R}$ defined by*

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

If Q is a compact and convex subset of Δ^{σ} and $\succsim$ is unbounded and such that

$$f \overset{Q}{=} g \implies f \sim g$$

then $\operatorname{dom} c \subseteq \Delta(Q)$.

Proof Let $p \in \Delta \backslash \Delta(Q)$. It follows that there exists $A \in \Sigma$ such that $q(A) = 0$ for all $q \in Q$ as well as $p(A) > 0$. Define $I : B_0(\Sigma) \to \mathbb{R}$ by $I(\varphi) = \min_{p \in \Delta} \left\{ \int \varphi dp + c(p) \right\}$ for all $\varphi \in B_0(\Sigma)$. Since u is unbounded, for each $\lambda \in \mathbb{R}$ there exists $x_\lambda \in X$ such that $u(x_\lambda) = \lambda$. Similarly, there exists $y \in X$ such that $u(y) = 0$. For each $\lambda \in \mathbb{R}$ define $f_\lambda = x_\lambda A y$. By construction, we have that $f_\lambda \overset{Q}{=} y$ for all $\lambda \in \mathbb{R}$. This implies that $I(\lambda 1_A) = V(f_\lambda) = V(y) = I(0) = 0$. By Maccheroni et al. (2006) and since u is unbounded, we have that

$$c(p) = \sup_{\varphi \in B_0(\Sigma)} \left\{ I(\varphi) - \int \varphi dp \right\} \geq \sup_{\lambda \in \mathbb{R}} \left\{ I(\lambda 1_A) - \lambda p(A) \right\} = \infty$$

Since p was arbitrarily chosen, it follows that $\operatorname{dom} c \subseteq \Delta(Q)$. ■

A.3.4 Proof of Theorem 1

We only prove (i) implies (ii), the converse being routine. By Lemma 13, there exist an onto and affine function $u : X \to \mathbb{R}$ and a weak divergence $c : \Delta \times Q \to [0, \infty]$ such that $\operatorname{dom} c(\cdot, q) \subseteq \Delta(Q)$ for all $q \in Q$ and $\succsim^*$ is represented by

$$f \succsim^* g \Leftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p, q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p, q) \right\} \quad \forall q \in Q \tag{39}$$

and $\succsim$ is represented by $V : \mathcal{F} \to \mathbb{R}$ defined by

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + c_Q(p) \right\} \tag{40}$$

By Lemma 14 and since $\succsim$ is subjectively Q-coherent and $\succsim^*$ and $\succsim$ coincide on X, we conclude that $c_Q^{-1}(0) = Q$, proving the implication.

Next, assume that c is uniquely null. Define the correspondence $\Gamma : Q \rightrightarrows Q$ by

$$\Gamma(q) = \{p \in \Delta : c(p,q) = 0\} = \arg\min c_q$$

Since $c_Q \leq c_q$ for all $q \in Q$ and $c_Q^{-1}(0) = Q$, we have that Γ is well defined. Since c_q is grounded, it follows that $\Gamma(q) \neq \emptyset$ for all $q \in Q$. Since c is uniquely null and c_q is grounded, we have that $c_q^{-1}(0)$ is a singleton, that is,

$$c(p,q) = c\left(p',q\right) = 0 \Longrightarrow p = p'$$

This implies that $\Gamma(q)$ is a singleton, therefore Γ is a function. Since $c_Q^{-1}(0) = Q$, observe that

$$\cup_{q\in Q}\Gamma(q) = \cup_{q\in Q} \arg\min c_q = \arg\min c_Q = Q$$

that is, Γ is surjective. Since c is uniquely null, we have that $c_p^{-1}(0)$ is at most a singleton, that is,

$$c(p,q) = c\left(p,q'\right) = 0 \implies q = q'$$

yielding that Γ is injective. To sum up, Γ is a bijection. Define $\tilde{c} : \Delta \times Q \to [0,\infty]$ by $\tilde{c}(p,q) = c\left(p,\Gamma^{-1}(q)\right)$ for all $(p,q) \in \Delta \times Q$. Note that $\tilde{c}(\cdot,q)$ is grounded, lower semicontinuous, convex and $\operatorname{dom}\tilde{c}(\cdot,q) \subseteq \Delta(Q)$ for all $q \in Q$. Next, we show that $\tilde{c}_Q = c_Q$. Since c_Q is well defined, for each $p \in \Delta$ there exists $q_p \in Q$ such that

$$\tilde{c}(p,\Gamma(q_p)) = c(p,q_p) = \min_{q\in Q} c(p,q) \leq c\left(p,q'\right) = \tilde{c}\left(p,\Gamma\left(q'\right)\right) \quad \forall q' \in Q$$

Since Γ is a bijection, we have that $\tilde{c}(p,\Gamma(q_p)) \leq \tilde{c}(p,q)$ for all $q \in Q$. Since p was arbitrarily chosen, it follows that

$$c_Q(p) = \min_{q\in Q} c(p,q) = \tilde{c}(p,\Gamma(q_p)) = \min_{q\in Q}\tilde{c}(p,q) = \tilde{c}_Q(p) \quad \forall p \in \Delta$$

To sum up, $\tilde{c}_Q = c_Q$ and $\tilde{c}_Q^{-1}(0) = c_Q^{-1}(0) = Q$. In turn, since c_Q is grounded, lower semicontinuous and convex, this implies that $\tilde{c}_Q$ is grounded, lower semicontinuous and convex. Since Γ is a bijection, we can conclude that (39) holds with $\tilde{c}$ in place of c and (40) holds with $\tilde{c}_Q$ in place of c_Q.

We are left to show that $\tilde{c}(p,q) = 0$ if and only if $p = q$. Since $c_q^{-1}(0)$ is a singleton for all $q \in Q$ and Γ is a bijection, if $\tilde{c}(p,q) = 0$, then $c\left(p,\Gamma^{-1}(q)\right) = 0$, yielding that $p = \Gamma\left(\Gamma^{-1}(q)\right) = q$. On the other hand, $\tilde{c}(q,q) = c\left(q,\Gamma^{-1}(q)\right) = 0$. We can conclude that $\tilde{c}(p,q) = 0$ if and only if $p = q$, proving that $\tilde{c}$ is a statistical distance. ■

A.3.5 Proof of Theorem 2

We only prove (i) implies (ii), the converse being routine. We proceed by steps.

Step 1. $\succsim_Q^*$ *agrees with* $\succsim_{Q'}^*$ *on* X *for all* $Q, Q' \in \mathcal{Q}$*. In particular, there exists an affine and onto function* $u : X \to \mathbb{R}$ *representing* $\succsim_Q^*$ *for all* $Q \in \mathcal{Q}$*.*

Proof of the Step Let $Q, Q' \in \mathcal{Q}$ be such that $Q \supseteq Q'$. Note that $\succsim_Q^*$ and $\succsim_{Q'}^*$, restricted to X, satisfy weak order, continuity and risk independence. By Herstein and Milnor (1953) and since $\succsim_Q^*$ and $\succsim_{Q'}^*$ are non-trivial, there exist two non-constant affine functions $u_Q, u_{Q'} : X \to \mathbb{R}$ which represent $\succsim_Q^*$ and $\succsim_{Q'}^*$, respectively. Since $\left\{\succsim_Q^*\right\}_{Q \in \mathcal{Q}}$ is monotone, we have that

$$u_Q(x) \geq u_Q(y) \Longrightarrow u_{Q'}(x) \geq u_{Q'}(y)$$

By Ghirardato et al. (2004), u_Q and $u_{Q'}$ are equal up to an affine and positive transformation. Next, fix $\bar{q} \in Q$. Set $u = u_{\bar{q}}$. Given any other $q \in \Delta^\sigma$, consider $\bar{Q} = \text{co}\{\bar{q}, q\}$. By the previous part, it follows that $u_{\bar{Q}}$, u_q and $u_{\bar{q}}$ are equal up to an affine transformation. Given that q was arbitrarily chosen, we can set $u = u_q$ for all $q \in Q$. Similarly, given a generic $Q \in \mathcal{Q}$, select $q \in Q$. Since $Q \supseteq \{q\}$, it follows that we can set $u = u_Q$. Since each $\succsim_Q^*$ is unbounded for all $Q \in \mathcal{Q}$, we have that u is onto. $\square$

Step 2. For each $q \in \Delta^\sigma$ *there exists a normalized, monotone, translation invariant and concave functional* $I_q : B_0(\Sigma) \to \mathbb{R}$ *such that*

$$f \succsim_q^* g \Longleftrightarrow I_q(u(f)) \geq I_q(u(g))$$

Moreover, there exists a unique grounded, lower semicontinuous and convex function $c_q : \Delta \to [0, \infty]$ *such that*

$$I_q(\varphi) = \min_{p \in \Delta} \left\{ \int \varphi dp + c_q(p) \right\} \qquad \forall \varphi \in B_0(\Sigma) \tag{41}$$

Proof of the Step Fix $q \in \Delta^\sigma$. Since $\succsim_q^*$ is an unbounded dominance relation which is complete, we have that $\succsim_q^*$ is a variational preference. By the proof of Theorem 3 and Proposition 6 of Maccheroni et al. (2006), there exists a normalized, monotone, translation invariant and concave functional $I_q : B_0(\Sigma) \to \mathbb{R}$ such that

$$f \succsim_q^* g \iff I_q(u(f)) \geq I_q(u(g))$$

Moreover, we have that there exists a unique grounded, lower semicontinuous and convex function $c_q : \Delta \to [0, \infty]$ satisfying (41). $\square$

Define $c : \Delta \times \Delta^\sigma \to [0, \infty]$ by $c(p,q) = c_q(p)$ for all $(p,q) \in \Delta \times \Delta^\sigma$. Define the map $J : B_0(\Sigma) \times \Delta^\sigma \to \mathbb{R}$ by $J(\varphi, q) = I_q(\varphi)$. Observe that, for all $(p,q) \in \Delta \times \Delta^\sigma$,

$$c(p,q) = c_q(p) = \sup_{\varphi \in B_0(\Sigma)} \left\{ I_q(\varphi) - \int \varphi dp \right\} = \sup_{\varphi \in B_0(\Sigma)} \left\{ J(\varphi, q) - \int \varphi dp \right\} \tag{42}$$

Step 3. J is convex and lower semicontinuous in the second argument.

Proof of the Step Note that for each $\varphi \in B_0(\Sigma)$ and for each $q \in \Delta^\sigma$

$$J(\varphi, q) = I_q(\varphi) = u(x_{f,q}) \qquad \text{where } f \in \mathcal{F} \text{ is s.t. } \varphi = u(f)$$

Fix $\varphi \in B_0(\Sigma)$ and $t \in \mathbb{R}$. By Step 1 and since $\left\{\succsim_Q^*\right\}_{Q \in \mathcal{Q}}$ is lower semicontinuous on Δ^σ, the set

$$\{q \in \Delta^\sigma : J(\varphi, q) \leq t\} = \{q \in \Delta^\sigma : u(x) \geq u(x_{f,q})\} = \{q \in \Delta^\sigma : x \succsim^* x_{f,q}\}$$

is closed where $x \in X$ and $f \in \mathcal{F}$ are such that $u(x) = t$ as well as $u(f) = \varphi$. Since φ and t were arbitrarily chosen, this yields that J is lower semicontinuous in the second argument. Fix $\varphi \in B_0(\Sigma)$, $q, q' \in \Delta^\sigma$ and $\lambda \in (0,1)$. Since $\left\{\succsim_Q^*\right\}_{Q \in \mathcal{Q}}$ is averse to model hybridization and u is affine,

$$\begin{aligned} J\left(\varphi, \lambda q + (1-\lambda) q'\right) &= u\left(x_{f, \lambda q + (1-\lambda) q'}\right) \leq u\left(\lambda x_{f,q} + (1-\lambda) x_{f,q'}\right) \\ &= \lambda u(x_{f,q}) + (1-\lambda) u\left(x_{f,q'}\right) = \lambda J(\varphi, q) + (1-\lambda) J\left(\varphi, q'\right) \end{aligned}$$

where $f \in \mathcal{F}$ is such that $u(f) = \varphi$. Since φ, q, q' and λ were arbitrarily chosen, this yields that J is convex in the second argument. $\square$

Step 4. c is jointly lower semicontinuous and convex. Moreover, its q-sections are grounded, lower semicontinuous and convex.

Proof of the Step By Step 3, the map $(p,q) \mapsto J(\varphi, q) - \int \varphi dp$, defined over $\Delta \times \Delta^\sigma$, is jointly lower semicontinuous and convex. By (42) and the definition of c, we conclude that c is jointly lower semicontinuous and convex. By Step 1, the rest of the statement follows.$\square$

Step 5. For each $Q \in \mathcal{Q}$ we have that $f \succsim_Q^* g$ if and only if $f \succsim_q^* g$ for all $q \in Q$. In particular, we have that

$$f \succsim_Q^* g \iff \min_{p \in \Delta} \left\{ \int u(f) \, dp + c(p,q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g) \, dp + c(p,q) \right\} \quad \forall q \in Q \tag{43}$$

Proof of the Step Fix $Q \in \mathcal{Q}$. Since $\left\{\succsim_Q^*\right\}_{Q \in \mathcal{Q}}$ is monotone, we have that

$$f \succsim_Q^* g \implies f \succsim_q^* g \qquad \forall q \in Q$$

Since $\left\{\succsim_Q^*\right\}_{Q\in\mathcal{Q}}$ is Q-separable, we can conclude that $f \succsim_Q^* g$ if and only if $f \succsim_q^* g$ for $q \in Q$. By Step 2 and the definition of c, this implies (43). $\square$

Step 6. $\succsim_Q^*$ agrees with $\succsim_Q$ on X for all $Q \in \mathcal{Q}$. Moreover, $\succsim_Q$ is represented by the function u of Step 1.

Proof of the Step Fix $Q \in \mathcal{Q}$. Note that $\succsim_Q^*$ and $\succsim_Q$, restricted to X, satisfy weak order, continuity and risk independence. By Herstein and Milnor (1953) and since $\succsim_Q$ is non-trivial, there exist a non-constant affine function v_Q which represents $\succsim_Q$. By Step 1, $\succsim_Q^*$ is represented by u. Since $\left(\succsim_Q^*, \succsim_Q\right)$ jointly satisfy consistency, it follows that for each $x, y \in X$

$$u(x) \geq u(y) \Longrightarrow v_Q(x) \geq v_Q(y)$$

By Ghirardato et al. (2004), v_Q and u are equal up to an affine and positive transformation. So we can set $v_Q = u$, proving the statement. $\square$

Step 7. For each $Q \in \mathcal{Q}$ we have that

$$f \succsim_Q g \iff \min_{p\in\Delta}\left\{\int u(f)\,dp + \min_{q\in Q} c(p,q)\right\} \geq \min_{p\in\Delta}\left\{\int u(g)\,dp + \min_{q\in Q} c(p,q)\right\} \tag{44}$$

Moreover, the function $c_Q : \Delta \to [0,\infty]$, defined by $c_Q(p) = \min_{q\in Q} c(p,q)$ for all $p \in \Delta$, is well defined, grounded, lower semicontinuous and convex.

Proof of the Step Fix $Q \in \mathcal{Q}$. By Cerreia-Vioglio et al. (2011) and since $\succsim_Q$ is a rational preference relation, there exists a normalized, monotone and continuous functional $I_Q : B_0(\Sigma) \to \mathbb{R}$ such that

$$f \succsim_Q g \iff I_Q(u(f)) \geq I_Q(u(g)) \tag{45}$$

We next show that $I_Q(\varphi) \leq \inf_{q\in Q} I_q(\varphi)$ for all $\varphi \in B_0(\Sigma)$. Consider $\varphi \in B_0(\Sigma)$. Since each I_q is normalized and monotone and u is onto, we have that $I_q(\varphi) \in [\inf_{s\in S}\varphi(s), \sup_{s\in S}\varphi(s)] \subseteq \operatorname{Im} u$ for all $q \in Q$. Since $\varphi \in B_0(\Sigma)$, it follows that there exists $f \in \mathcal{F}$ such that $\varphi = u(f)$ and $x \in X$ such that $u(x) = \inf_{q\in Q} I_q(\varphi)$. For each $\varepsilon > 0$ there exists $x_\varepsilon \in X$ such that $u(x_\varepsilon) = u(x) + \varepsilon$. Since $\inf_{q\in Q} I_q(\varphi) = u(x)$, it follows that for each $\varepsilon > 0$ there exists $q \in Q$ such that $I_q(u(f)) = I_q(\varphi) < u(x_\varepsilon) = I_q(u(x_\varepsilon))$, yielding that $f \not\succsim_Q^* x_\varepsilon$. Since $\left(\succsim_Q^*, \succsim_Q\right)$ jointly satisfy caution, we have that $x_\varepsilon \succsim_Q f$ for all $\varepsilon > 0$. By (45), this implies that

$$u(x) + \varepsilon = u(x_\varepsilon) = I_Q(u(x_\varepsilon)) \geq I_Q(u(f)) = I_Q(\varphi) \quad \forall \varepsilon > 0$$

that is, $\inf_{q\in Q} I_q(\varphi) = u(x) \geq I_Q(\varphi)$. We next prove that $I_Q(\varphi) \geq \inf_{q\in Q} I_q(\varphi)$ for all $\varphi \in B_0(\Sigma)$. Consider $\varphi \in B_0(\Sigma)$. We use the same objects of before. Note that for each $q' \in Q$

$$I_{q'}(u(f)) = I_{q'}(\varphi) \geq \inf_{q\in Q} I_q(\varphi) = u(x) = I_{q'}(u(x))$$

that is, $f \succsim_Q^* x$. Since $\left(\succsim_Q^*, \succsim_Q\right)$ jointly satisfy consistency, we have that $f \succsim_Q x$. By (45), this implies that

$$I_Q(\varphi) = I_Q(u(f)) \geq I_Q(u(x)) = u(x) = \inf_{q \in Q} I_q(\varphi)$$

proving that $I_Q = \inf_{q \in Q} I_q$. By (43) and since c is jointly lower semicontinuous and convex, we can conclude that

$$\begin{aligned} I_Q(\varphi) &= \inf_{q \in Q} \min_{p \in \Delta} \left\{ \int \varphi dp + c(p,q) \right\} = \min_{q \in Q} \min_{p \in \Delta} \left\{ \int \varphi dp + c(p,q) \right\} \\ &= \min_{p \in \Delta} \min_{q \in Q} \left\{ \int \varphi dp + c(p,q) \right\} = \min_{p \in \Delta} \left\{ \int \varphi dp + \min_{q \in Q} c(p,q) \right\} \quad \forall \varphi \in B_0(\Sigma) \end{aligned}$$

By (45), this implies that (44) holds. Finally, by Lemma 4, we have that the function $p \mapsto \min_{q \in Q} c(p,q)$ is well defined, grounded, lower semicontinuous and convex. □

Step 8. $c_Q^{-1}(0) = Q$ *for all* $Q \in \mathcal{Q}$. *Moreover,* $c(p,q) = 0$ *if and only if* $p = q$.

Proof of the Step Fix $Q \in \mathcal{Q}$. Since $\succsim_Q$ is subjectively Q-coherent, it follows that $c_Q^{-1}(0) = Q$. In particular, when $Q = \{q\}$ for some $q \in \Delta^\sigma$, we have that $c(p,q) = 0$ if and only if $c_Q(p) = 0$ if and only if $p \in Q$ if and only if $p = q$. □

Step 9. $\operatorname{dom} c(\cdot, q) \subseteq \Delta(Q)$ *for all* $q \in Q$ *and for all* $Q \in \mathcal{Q}$.

Proof of the Step By the previous part of the proof, we have that $\succsim_q^*$ coincides with $\succsim_q$ on $\mathcal{F}$ for all $q \in \Delta^\sigma$. By Lemma 15 and since $\succsim_q^*$ is objectively $\{q\}$-coherent, we can conclude that $\operatorname{dom} c(\cdot, q) \subseteq \Delta(q) \subseteq \Delta(Q)$ for all $q \in Q$ and for all $Q \in \mathcal{Q}$. □

Steps 4, 7, 8 and 9 prove that c is a statistical distance which is jointly lower semicontinuous and convex such that $\operatorname{dom} c(\cdot, q) \subseteq \Delta(Q)$ for all $q \in Q$ and for all $Q \in \mathcal{Q}$, yielding that $\operatorname{dom} c_Q \subseteq \Delta(Q)$ for all $Q \in \mathcal{Q}$. Steps 5 and 7 prove, respectively, (26) and (27). As for uniqueness, assume that the function $\tilde{c} : \Delta \times \Delta^\sigma \to [0, \infty]$ is a statistical distance which is jointly lower semicontinuous and convex and such that $\operatorname{dom} c_Q \subseteq \Delta(Q)$ for all $Q \in \mathcal{Q}$ and that satisfies (26) and (27). By Proposition 6 of Maccheroni et al. (2006), it follows that $\tilde{c}(\cdot, q) = c(\cdot, q)$ for all $q \in \Delta^\sigma$, yielding that $c = \tilde{c}$. ■

A.4 Other proofs

Proof of Proposition 2 First, note that $\min_{q \in Q} R(p||q) = 0$ if and only if $p \in Q$. Indeed, we have that

$$\min_{q \in Q} R(p||q) = 0 \iff \exists \bar{q} \in Q \text{ s.t. } R(p||\bar{q}) = 0 \iff \exists \bar{q} \in Q \text{ s.t. } p = \bar{q}$$

Define $\lambda_n = n$ for all $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, we have $\lambda_n \min_{q\in Q} R(p||q) = 0$ if and only if $p \in Q$. So, for each $p \in \Delta$,

$$\lim_n \lambda_n \min_{q\in Q} R(p||q) = \begin{cases} 0 & \text{if } p \in Q \\ \infty & \text{if } p \notin Q \end{cases}$$

Since $\lambda_n \min_{q\in Q} R(p||q) = 0$ for each $n \in \mathbb{N}$ if and only if $p \in Q$, by Proposition 12 of Maccheroni et al. (2006) we have

$$\lim_n \min_{p\in\Delta} \left\{ \int u(f)\,dp + \lambda_n \min_{q\in Q} R(p||q) \right\} = \min_{q\in Q} \int u(f)\,dq \quad \forall f \in \mathcal{F}$$

Finally, by (19), we have that for each $f \in \mathcal{F}$

$$\begin{aligned} \min_{q\in Q} \int u(f)\,dq &\leq \lim_n \min_{p\in\Delta} \left\{ \int u(f)\,dp + \lambda_n \min_{q\in Q} R(p||q) \right\} \\ &\leq \lim_{\lambda\uparrow\infty} \min_{p\in\Delta} \left\{ \int u(f)\,dp + \lambda \min_{q\in Q} R(p||q) \right\} \leq \min_{q\in Q} \int u(f)\,dq \end{aligned}$$

yielding the statement. ■

Proof of Proposition 3 Note that $c(\cdot, q) = \lambda D_\phi(\cdot||q)$ is Shur convex (with respect to q) for all $q \in Q$. Consider $A, B \in \Sigma$. Assume that $q(A) \geq q(B)$ for all $q \in Q$. Let $q \in Q$. Consider $x, y \in X$ such that $x \succ y$. It follows that

$$\int v(u(xAy))\,dq \geq \int v(u(xBy))\,dq$$

for each $v : \mathbb{R} \to \mathbb{R}$ strictly increasing and concave. By Theorem 2 of Cerreia-Vioglio et al. (2012) and since q was arbitrarily chosen, it follows that

$$\min_{p\in\Delta} \left\{ \int u(xAy)\,dp + \lambda D_\phi(p||q) \right\} \geq \min_{p\in\Delta} \left\{ \int u(xBy)\,dp + \lambda D_\phi(p||q) \right\} \quad \forall q \in Q$$

yielding that $xAy \succsim^* xBy$ and, in particular, $xAy \succsim xBy$. ■

Proof of Proposition 4 We prove the "only if", the converse being obvious. Define $\gtrsim^*$ by $f \gtrsim^* g$ if and only if $\int u(f)\,dq \geq \int u(g)\,dq$ for all $q \in Q$. By hypothesis, the pair $(\gtrsim^*, \succsim)$ satisfies consistency. Let $f \not\gtrsim^* x$. Then, there exists $q \in Q$ such that $u(x_f^q) = \int u(f)\,dq < u(x)$. Hence, $x \succ x_f^q$. Since $c_Q^{-1}(0) = Q$, by Lemma 14 we have $x \succ f$. So, the pair $(\gtrsim^*, \succsim)$ satisfies default to certainty. By Theorem 4 of Gilboa et al. (2010), this pair admits the representation

$$f \gtrsim^* g \Longleftrightarrow \int u(f)\,dq \geq \int u(g)\,dq \quad \forall q \in Q$$

and

$$f \succsim g \Longleftrightarrow \min_{q\in Q} \int u(f)\,dq \geq \min_{q\in Q} \int u(g)\,dq$$

Note that, in the notation of Gilboa et al. (2010), we have $C = Q$ because C is unique up to closure and convexity and Q is closed and convex. ■

Proof of Proposition 5 For each $q \in Q$ define $I_q : B_0(\Sigma) \to \mathbb{R}$ by

$$I_q(\varphi) = \min_{p \in \Delta} \left\{ \int u(\varphi)\, dp + c(p, q) \right\}$$

Recall that $f \succ\!\!\succ^* g$ if and only if for each $h, l \in \mathcal{F}$ there exists $\varepsilon > 0$ such that

$$(1-\delta) f + \delta h \succ^* (1-\delta) g + \delta l \qquad \forall \delta \in [0, \varepsilon] \tag{46}$$

Moreover, given $f, g \in \mathcal{F}$, define $k_* = \inf_{s \in S} u(f(s))$ and $k^* = \sup_{s \in S} u(g(s))$.

"Only if." Assume that $f \succ\!\!\succ^* g$. Let $\hat{\varepsilon} > 0$. Consider $u(x) = k_* - \hat{\varepsilon}$ and $u(y) = k^* + \hat{\varepsilon}$. By definition, there exists $\varepsilon > 0$ such that

$$(1-\delta) f + \delta x \succ^* (1-\delta) g + \delta y \qquad \forall \delta \in [0, \varepsilon]$$

Note that for each $q \in Q$ and for each $\delta \in [0, 1]$

$$\begin{aligned} I_q(u((1-\delta) f + \delta x)) &= I_q((1-\delta) u(f) + \delta u(x)) = I_q(u(f) - \delta u(f) + \delta u(x)) \\ &\leq I_q(u(f) - \delta k_* + \delta(k_* - \hat{\varepsilon})) = I_q(u(f)) - \delta\hat{\varepsilon} \end{aligned}$$

and

$$\begin{aligned} I_q(u((1-\delta) g + \delta y)) &= I_q((1-\delta) u(g) + \delta u(y)) = I_q(u(g) - \delta u(g) + \delta u(y)) \\ &\geq I_q(u(g) - \delta k^* + \delta(k^* + \hat{\varepsilon})) = I_q(u(g)) + \delta\hat{\varepsilon} \end{aligned}$$

It follows that for each $q \in Q$ and for each $\delta \in [0, \varepsilon]$

$$I_q(u(f)) - I_q(u(g)) - 2\delta\hat{\varepsilon} \geq I_q(u((1-\delta) f + \delta x)) - I_q(u((1-\delta) g + \delta y)) \geq 0$$

If we set $\delta = \varepsilon > 0$, then $I_q(u(f)) \geq I_q(u(g)) + 2\varepsilon\hat{\varepsilon}$, proving the statement.

"If." Let $f, g \in \mathcal{F}$. Assume there exists $\varepsilon > 0$ such that $I_q(u(f)) \geq I_q(u(g)) + \varepsilon$ for all $q \in Q$. Without loss of generality, we can assume that $\geq$ holds with strict inequality.[30] Consider $h, l \in \mathcal{F}$. Define $k_\star = \inf_{s \in S} u(h(s))$ and $k^\star = \sup_{s \in S} u(l(s))$. Define also $k^\sim = \sup_{s \in S} u(f(s))$ and $k_\sim = \inf_{s \in S} u(g(s))$. Note that for each $q \in Q$ and for each $\delta \in [0, 1]$

$$\begin{aligned} I_q(u((1-\delta) f + \delta h)) &= I_q((1-\delta) u(f) + \delta u(h)) = I_q(u(f) - \delta u(f) + \delta u(h)) \\ &= I_q(u(f) + \delta(u(h) - u(f))) \\ &\geq I_q(u(f) + \delta(k_\star - k^\sim)) = I_q(u(f)) + \delta(k_\star - k^\sim) \end{aligned}$$

[30] It is enough to replace ε with $\varepsilon/2$.

and

$$
\begin{aligned}
I_q\left(u\left((1-\delta)g+\delta l\right)\right) &= I_q\left((1-\delta)u(g)+\delta u(l)\right) = I_q\left(u(g)-\delta u(g)+\delta u(l)\right) \\
&= I_q\left(u(g)+\delta\left(u(l)-u(g)\right)\right) \\
&\leq I_q\left(u(g)+\delta\left(k^\star-k_\sim\right)\right) = I_q\left(u(g)\right)+\delta\left(k^\star-k_\sim\right)
\end{aligned}
$$

It follows that for each $q \in Q$ and for each $\delta \in [0,1]$

$$
\begin{aligned}
I_q\left(u\left((1-\delta)f+\delta h\right)\right) - I_q\left(u\left((1-\delta)g+\delta l\right)\right) &\geq I_q\left(u(f)\right)+\delta\left(k_\star-k^\sim\right)-I_q\left(u(g)\right)-\delta\left(k^\star-k_\sim\right) \\
&\geq \varepsilon+\delta\hat{\varepsilon}
\end{aligned}
$$

where $\hat{\varepsilon} = k_\star - k^\sim - k^\star + k_\sim$. We have two cases:

1. $\hat{\varepsilon} \geq 0$. In this case, $I_q\left(u\left((1-\delta)f+\delta h\right)\right) - I_q\left(u\left((1-\delta)g+\delta l\right)\right) > 0$ for all $\delta \in [0,1]$ and all $q \in Q$, proving (46).

2. $\hat{\varepsilon} < 0$. In this case, $I_q\left(u\left((1-\delta)f+\delta h\right)\right) - I_q\left(u\left((1-\delta)g+\delta l\right)\right) > 0$ for all $\delta \in [0, -\varepsilon/2\hat{\varepsilon}]$ and all $q \in Q$, proving (46).

This completes the proof of the result. ■

Proof of Lemma 3 Given $q \in Q$, if $c(p,q) = \infty$ for all $p \notin Q$, then $c_Q(p) = \infty$ for all $p \notin Q$. Since $c_Q(q) = 0$ for all $q \in Q$, we conclude that $c_Q(p) = \delta_Q(p)$ for all $p \in \Delta$. Conversely, for each $q \in Q$ we have $c(p,q) \geq c_Q(p) = \delta_Q(p) = \infty$ for all $p \notin Q$. ■

Proof of Proposition 6 (i) implies (ii). By Proposition 2 of Cerreia-Vioglio (2016) and since $\succsim^*$ is unbounded, there exists a compact and convex set $C \subseteq \Delta$ and an affine and onto map $u : X \to \mathbb{R}$ such that

$$
f \succsim^* g \Longleftrightarrow \int u(f)\,dq \geq \int u(g)\,dq \qquad \forall q \in C \tag{47}
$$

and

$$
f \succsim g \iff \min_{q\in C} \int u(f)\,dq \geq \min_{q\in C} \int u(g)\,dq \tag{48}
$$

By Lemma 14 and since $\succsim$ is subjectively Q-coherent and $\succsim^*$ and $\succsim$ coincide on X, we can conclude that $C = Q$. If we set $c : \Delta \times Q \to [0,\infty]$ to be $c(p,q) = \delta_{\{q\}}(p)$ for all $(p,q) \in \Delta \times Q$, then it is immediate to see that c is a statistical distance. By (47) and (48) and since $C = Q$, (13) and (14) follow.

(ii) implies (i). It is trivial. ■

Proof of Proposition 7 (i) Let $\hat{f} \in F$ be optimal. By (23), if there is $g \in F$ such that $g \succ\!\!\succ^*_Q \hat{f}$, then $g \succ_Q \hat{f}$, a contradiction with $\hat{f}$ being optimal. We conclude that $\hat{f}$ is weakly admissible. A similar argument proves that there is no $g \in F$ such that $g \succ_Q^* \hat{f}$ when (24) holds.

(ii) Suppose $\hat{f} \in F$ is the unique optimal act, that is, $\hat{f} \succ_Q f$ for all $\hat{f} \neq f \in F$. If $g \in F$ is such that $g \succ_Q^* \hat{f}$, then $g \succsim_Q \hat{f}$. In turn, this implies $g \succsim_Q \hat{f} \succ_Q g$, a contradiction. We conclude that $\hat{f}$ is admissible. ■

Proof of Proposition 8 By Lemma 1, c restricted to $\Delta \times Q$ (resp., $\Delta \times Q'$) is a statistical distance function. Since $Q \subseteq Q'$, it follows that $\min_{q \in Q} c(p,q) \geq \min_{q \in Q'} c(p,q)$ for all $p \in \Delta$. We thus have

$$\min_{p \in \Delta} \left\{ \int u(f)\, dp + \min_{q \in Q} c(p,q) \right\} \geq \min_{p \in \Delta} \left\{ \int u(f)\, dp + \min_{q \in Q'} c(p,q) \right\} \qquad \forall f \in F$$

yielding that $v(Q) \geq v(Q')$. Next, fix Q and assume that the sup in (29) is achieved. Let $\bar{f} \in F$ be such that

$$\min_{p \in \Delta} \left\{ \int u\left(\bar{f}\right) dp + \min_{q \in Q} c(p,q) \right\} = v(Q)$$

By contradiction, assume that $\bar{f} \in F/F_Q^*$. By Proposition 5 and since $\bar{f} \notin F_Q^*$ and $\bar{f} \in F$, there exists $g \in F$ such that $g \ggcurly_Q^* \bar{f}$, that is, there exists $\varepsilon > 0$ such that

$$\min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p,q) \right\} \geq \min_{p \in \Delta} \left\{ \int u\left(\bar{f}\right) dp + c(p,q) \right\} + \varepsilon \qquad \forall q \in Q$$

This implies that

$$\begin{aligned}
v(Q) &\geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + \min_{q \in Q} c(p,q) \right\} = \min_{p \in \Delta} \min_{q \in Q} \left\{ \int u(g)\, dp + c(p,q) \right\} \\
&\geq \inf_{q \in Q} \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p,q) \right\} \geq \inf_{q \in Q} \min_{p \in \Delta} \left\{ \int u\left(\bar{f}\right) dp + c(p,q) \right\} + \varepsilon \\
&\geq \min_{p \in \Delta} \min_{q \in Q} \left\{ \int u\left(\bar{f}\right) dp + c(p,q) \right\} + \varepsilon = \min_{p \in \Delta} \left\{ \int u\left(\bar{f}\right) dp + \min_{q \in Q} c(p,q) \right\} + \varepsilon \\
&= v(Q) + \varepsilon
\end{aligned}$$

a contradiction. ■

References

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

[2] I. Aydogan, L. Berger, V. Bosetti and N. Liu, Three layers of uncertainty and the role of model misspecification: an experiment, IGIER WP 623, 2018.

[3] A. Ben-Tal and M. Teboulle, An old-new concept of convex risk measures: The optimized certainty equivalent, *Mathematical Finance*, 17, 449–476, 2007.

[4] G. E. P. Box, Science and statistics, *Journal of the American Statistical Association*, 71, 791–799, 1976.

[5] S. Cerreia-Vioglio, Objective rationality and uncertainty averse preferences, *Theoretical Economics*, 11, 523–545, 2016.

[6] S. Cerreia-Vioglio, P. Ghirardato, F. Maccheroni, M. Marinacci and M. Siniscalchi, Rational preferences under ambiguity, *Economic Theory*, 48, 341–375, 2011.

[7] S. Cerreia-Vioglio, A. Giarlotta, S. Greco, F. Maccheroni and M. Marinacci, Rational preference and rationalizable choice, *Economic Theory*, 69, 61–105, 2020.

[8] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci and L. Montrucchio, Uncertainty averse preferences, *Journal of Economic Theory*, 146, 1275–1330, 2011.

[9] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci and L. Montrucchio, Probabilistic sophistication, second order stochastic dominance and uncertainty aversion, *Journal of Mathematical Economics*, 48, 271–283, 2012.

[10] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci and A. Rustichini, Niveloids and their extensions: Risk measures on small domains, *Journal of Mathematical Analysis and Applications*, 413, 343–360, 2014.

[11] D. R. Cox, Comment on "Model uncertainty, data mining and statistical inference", *Journal of the Royal Statistical Society*, Series A, 158, 455–456, 1995.

[12] I. Esponda and D. Pouzo, Berk-Nash equilibrium: A framework for modeling agents with misspecified models, *Econometrica*, 84, 1093–1130, 2016.

[13] T. S. Ferguson, *Mathematical statistics: A decision theoretic approach*, Academic Press, New York, 1967.

[14] A. V. Fiacco and J. Kyparisis, Convexity and concavity properties of the optimal value function in parametric nonlinear programming, *Journal of Optimization Theory and Applications*, 48, 95–126, 1986.

[15] D. Fudenberg, G. Romanyuk and P. Strack, Active learning with a misspecified prior, *Theoretical Economics*, 12, 1155–1189, 2017.

[16] P. Ghirardato, F. Maccheroni and M. Marinacci, Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory*, 118, 133–173, 2004.

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

[18] I. Gilboa, F. Maccheroni, M. Marinacci and D. Schmeidler, Objective and subjective rationality in a multiple prior model, *Econometrica*, 78, 755–770, 2010.

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

[20] L. P. Hansen, Nobel lecture: Uncertainty outside and inside economic models, *Journal of Political Economy*, 122, 945–987, 2014.

[21] L. P. Hansen and M. Marinacci, Ambiguity aversion and model misspecification: An economic perspective, *Statistical Science*, 31, 511–515, 2016.

[22] L. P. Hansen and J. Miao, Aversion to ambiguity and model misspecification in dynamic stochastic environments, *Proceedings of the National Academy of Sciences*, 115, 9163–9168, 2018.

[23] L. P. Hansen and T. J. Sargent, Robust control and model uncertainty, *American Economic Review*, 91, 60–66, 2001.

[24] L. P. Hansen and T. J. Sargent, Recursive robust estimation and control without commitment, *Journal of Economic Theory*, 136, 1–27, 2007.

[25] L. P. Hansen and T. J. Sargent, *Robustness*, Princeton University Press, Princeton, 2008.

[26] L. P. Hansen and T. J. Sargent, Structured uncertainty and model misspecification, *mimeo*, 2020.

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

[28] E. Hewitt and L. J. Savage, Symmetric measures on Cartesian products, *Transactions of the American Mathematical Society*, 80, 470–501, 1955.

[29] F. Liese and I. Vajda, *Convex statistical distances*, Teubner, Leipzig, 1987.

[30] D. H. Jacobson, Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games, *IEEE Transactions on Automatic Control*, AC-18, 124–131, 1973.

[31] F. Maccheroni and M. Marinacci, A Heine-Borel Theorem for $ba(\Sigma)$, *RISEC*, 48, 353–362, 2001.

[32] F. Maccheroni, M. Marinacci and A. Rustichini, Ambiguity aversion, robustness, and the variational representation of preferences, *Econometrica*, 74, 1447–1498, 2006.

[33] M. Marinacci, Model uncertainty, *Journal of the European Economic Association*, 13, 1022–1100, 2015.

[34] I. R. Petersen, M. R. James and P. Dupuis, Minimax optimal control of stochastic uncertain systems with relative entropy constraints, *IEEE Transactions on Automatic Control*, 45, 398–412, 2000.

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

[36] S. M. Srivastava, *A course on Borel sets*, Springer, New York, 1998.

[37] T. Strzalecki, Axiomatic foundations of multiplier preferences, *Econometrica*, 79, 47–73, 2011.

[38] F. Topsoe, Basic concepts, identities and inequalities – the toolkit of information theory, *Entropy*, 3, 162–190, 2001.

[39] A. Wald, *Statistical decision functions*, John Wiley & Sons, New York, 1950.