Ambiguity and Robust Statistics*

Simone Cerreia-Vioglio[a] Fabio Maccheroni[a] Massimo Marinacci[a] Luigi Montrucchio[b]

[a]Department of Decision Sciences, Università Bocconi

[b]Collegio Carlo Alberto, Università di Torino

March 2011†

Prior distributions can never be quantified or elicited exactly (i.e., without error), especially in a finite amount of time. Berger (1984, p. 64)

1 Introduction

Since the seminal work of Gilboa and Schmeidler [28, p. 142] a relation between decision making under ambiguity and robust Bayesian statistics has been hinted at, and indeed immediate similarities are quite evident. At the same time, a formal treatment of this topic and a complete characterization of the relation between the two approaches is still missing. The object of this paper is to fill the gap, that is, relating ambiguity (also called Knightian uncertainty or model uncertainty) to prior uncertainty.

Ambiguity refers to the case in which a decision maker does not know the probability distribution governing the stochastic nature of the problem he is facing. This uncertainty is captured by using nonadditive probabilities – capacities – or sets of probability measures over the space of states of the world.[1]

Prior uncertainty, in a parametric statistical model $\{P_\theta\}_{\theta\in\Theta}$, refers to uncertainty about the prior μ on Θ. This is a classical problem in robust Bayesian statistics, where uncertainty is again represented by capacities or sets of priors over the space Θ of parameters.[2]

Clearly, prior uncertainty can be reduced (literally) to ambiguity. Loosely, if ν is a (prior) capacity on the space Θ of parameters, then

$$\bar{\nu}(\cdot) = \int_\Theta P_\theta(\cdot)\, d\nu(\theta) \tag{1}$$

defines a (predictive) capacity on the states of the world. Analogously, if $\mathcal{M}$ is a set of priors on the space Θ of parameters, then

$$\left\{\bar{\mu}(\cdot) = \int_\Theta P_\theta(\cdot)\, d\mu(\theta) : \mu \in \mathcal{M}\right\} \tag{2}$$

defines a set of (predictive) probability measures on the states of the world.

In this paper we address the converse problem. That is, we start from a decision theoretic framework of ambiguity and we show under which conditions the decision problem admits a (suitably unique) rephrasing

*We thank Marcelo Moreira, Ben Polak, Chris Shannon, Piotr Zakrzewski for helpful suggestions. The financial support of the European Research Council (advanced grant, BRSCDP-TEA) is gratefully acknowledged.

†First presented with the title *Model Uncertainty and Sufficient Events* at RUD 2010. Most of the results contained in this paper appeared in the presentation slides, posted in June 2010 and still available on the conference website.

[1]Early classical references are Schmeidler [47], Bewley [7], and Gilboa and Schmeidler [28]. See Gilboa and Marinacci [27] for a recent survey.

[2]See Shafer [49], Berger [6], Wasserman and Kadane [54], as well as Huber and Ronchetti [35].

in terms of prior uncertainty. In this way, we are able to provide an axiomatic and behaviorally falsifiable foundation to (old and new) criteria used in robust Bayesian statistics. We achieve this goal by merging the decision theoretic assumptions that are by now established in the literature of choice under ambiguity with some of the fundamental insights contained in Wald [53], Ellsberg [21], Nehring [42], and Gilboa, Maccheroni, Marinacci, and Schmeidler [26].

We model ambiguity by using a generalized Anscombe-Aumann setting as in Gilboa, Maccheroni, Marinacci, and Schmeidler [26]. We denote by Ω the space of states of the world and we consider it endowed with a σ-algebra $\mathcal{F}$ of events and a set $\mathcal{P}$ of probability measures on $\mathcal{F}$. We call these probability measures *objectively rational beliefs*. The elements of $\mathcal{P}$ represent the probabilistic beliefs on $\mathcal{F}$ that the decision maker is able to justify on the basis of the available information. Incompleteness of information is then captured by the nonsingleton nature of $\mathcal{P}$. In particular, the class of objectively rational beliefs we study is a special case of the one considered by Gilboa, Maccheroni, Marinacci, and Schmeidler [26]. In fact, we further require that $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space. This notion was introduced by Dynkin [19] and is discussed in full detail in Section 3.2, which also collects some important examples. Here we just sketch two of them, which were also suggested by Nehring [42, p. 1060] to justify the inclusion of beliefs among the primitives of a decision problem.[3]

First, if an Ellsberg's three colors urn is given (see Example 1) it is natural to consider as $\mathcal{P}$ the set of all probability measures that conditional on the composition of the urn coincide with the classical probability assignment.[4] In this case, $\mathcal{P}$ is naturally interpreted as a datum since it refers to the possible physical compositions of the urn. In other words, $\mathcal{P}$ is the set of all "objective urn models."

Second, in a very different perspective, consider a coin which is tossed over and over again. Based on subjective similarity considerations, the famous de Finetti's argument imposes that only exchangeable probability measures on the space of all sequences of tosses can be rationally justified. In this case, $\mathcal{P}$ consists of those measures that assign the same probability to all finite sequences of the same length with the same number of heads (and the same number of tails). In comparing this example with the previous one, it is important to notice that here we are in the domain of subjective probability – without any obvious physical counterpart. See Example 2 for details and extensions (standard Borel G-spaces and invariant probability measures).

The decision maker's behavior is described by a *subjectively rational preference* $\succsim$ on the set $B_0(X)$ of all acts, that is, simple and measurable mappings f from Ω to the convex set X of outcomes.[5] The preference relation $\succsim$ represents the actual choices of the decision maker among acts. Our central assumption is that choice is coherent with probabilistic information, and it is expressed by imposing consistency *à la* [26] between objectively rational beliefs $\mathcal{P}$ and subjectively rational preference $\succsim$. Formally, the *consistency* axiom requires that

$$\int_\Omega f dP \succsim \int_\Omega g dP \text{ for all } P \in \mathcal{P} \text{ implies } f \succsim g \tag{3}$$

where $\int_\Omega f dP$ and $\int_\Omega g dP$ belong to X and are the means of f and g under P, respectively. In other words, we consider a decision problem in which information, for example based on physical (urns) or symmetry (coins) considerations, restricts the decision maker's conceivable beliefs to belong to $\mathcal{P}$.[6] If the decision maker could

[3] Many more examples – including Gibbs states, Markov processes with a given transition function, and stationary probability distributions for a given stationary transition function – can be found in Dynkin [19].

[4] That attributes to the extraction of a given color the ratio between the number of balls with that color and the total number of balls.

[5] For the sake of precision, $\succsim$ is a nontrivial and continuous preorder which can be represented by a von Neumann-Morgenstern utility u on X.

[6] In the words of Ellsberg [21, p. 661]: *"Out of the set [$\Delta(\Omega)$] of all possible distributions there remains a set [$\mathcal{P}$] of*

confidently select P in $\mathcal{P}$, he would be a standard Anscombe-Aumann expected utility maximizer. In this perspective, the right hand side of condition (3) means that, no matter what the best estimate P is, if the agent trusted it he would prefer f to g.[7] Consistency imposes that the decision maker's subjective rationality takes this fact into account, by declaring f preferred to g.

In a nutshell, our results show that when uncertainty is represented as above, if subjectively rational preferences are consistent with objectively rational beliefs, then ambiguity is equivalent to prior uncertainty. Next we discuss two of these results, which can be immediately recognized as the counterparts of the reduction procedures described by (1) and (2).

As a consequence of Theorem 7, we have that if the decision maker is a Choquet expected utility maximizer and his preferences are consistent, then $\succsim$ is represented by

$$V(f) = \int_{\mathcal{S}(\mathcal{P})} \left(\int_{\Omega} u(f(\omega))\, dP(\omega) \right) d\nu(P) \text{ for all } f \in B_0(X) \tag{4}$$

where $\mathcal{S}(\mathcal{P})$ is the set of strong extreme points of $\mathcal{P}$ and ν is a unique capacity over $\mathcal{S}(\mathcal{P})$. As shown by Dynkin [19], the set $\mathcal{S}(\mathcal{P})$ can be seen as the set of pure models defined by $\mathcal{P}$, while all the other elements of $\mathcal{P}$ can be seen as mixture models. In particular, in the Ellsberg urn case mentioned above $\mathcal{S}(\mathcal{P}) = \mathcal{P}$ describes all possible urn compositions, while in the de Finetti exchangeable case $\mathcal{S}(\mathcal{P})$ is the set of i.i.d. models. In words, our decision maker is acting like a statistician who, starting from $\mathcal{P}$, is able to identify the pure models $\mathcal{S}(\mathcal{P})$, but is not able to exactly quantify a prior probability over them.

By Schmeidler [47], the preferences of this decision maker are also represented by

$$V(f) = \int_{\Omega} u(f(\omega))\, d\rho(\omega) \text{ for all } f \in B_0(X)$$

where ρ is a unique capacity over Ω (rather than over $\mathcal{S}(\mathcal{P})$). Our result implies that

$$\rho(A) = \int_{\mathcal{S}(\mathcal{P})} P(A)\, d\nu(P) \text{ for all } A \in \mathcal{F}$$

that is, each (consistent) capacity ρ is a predictive capacity. In this way, Theorem 7 presents the ambiguity counterpart of the reduction from priors to predictives described by (1).

As a consequence of Theorem 8, we have that if the decision maker is a Bewley expected utility maximizer and his preferences are consistent, then $\succsim$ is represented by

$$V(f) = \left[\int_{\mathcal{S}(\mathcal{P})} \left(\int_{\Omega} u(f(\omega))\, dP(\omega) \right) d\mu(P) \right]_{\mu \in \mathcal{M}} \text{ for all } f \in B_0(X) \tag{5}$$

where $\mathcal{M}$ is a unique compact and convex set of priors over $\mathcal{S}(\mathcal{P})$; and again our decision maker is acting like a statistician who is not able to exactly quantify a single prior probability.

By Bewley [7] and Gilboa, Maccheroni, Marinacci, and Schmeidler [26], the preferences of this decision maker are also represented by

$$V(f) = \left[\int_{\Omega} u(f(\omega))\, dm(\omega) \right]_{m \in \mathcal{C}} \text{ for all } f \in B_0(X)$$

distributions that still seem 'reasonable,' ... that his information – perceived as scanty, unreliable, ambiguous *– does not permit him confidently to rule out ... he might suspect [that his best estimate P among the elements of $\mathcal{P}$] might vary almost hourly with his mood,"* quoted also by [25].

[7] Notice that $\int_{\Omega} f dP$ and $\int_{\Omega} g dP$ would be the certainty equivalents of f and g, respectively.

where $\mathcal{C}$ is a unique compact and convex set of probability measures over Ω (rather than over $\mathcal{S}(\mathcal{P})$). Our results imply that for every $m \in \mathcal{C}$ there is a unique $\mu \in \mathcal{M}$ such that

$$m(A) = \int_{\mathcal{S}(\mathcal{P})} P(A)\, d\mu(P) \text{ for all } A \in \mathcal{F}$$

that is, each (consistent, compact, and convex) set of probability measures $\mathcal{C}$ is a set of predictives. In this way, Theorem 8 presents the ambiguity counterpart of the reduction from priors to predictives described by (2).[8]

The above results are not peculiar to Choquet or Bewley expected utility preferences. The equivalence between ambiguity and prior uncertainty is a general consequence of the consistency axiom in conjunction with the structure of Dynkin space. For example, Theorem 4 delivers a similar result for the Variational preferences of Maccheroni, Marinacci, and Rustichini [40], and it is then generalized to the Uncertainty Averse preferences of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [10] in the subsequent Theorem 5. Moreover, the structure of Dynkin space and the consistency axiom allow us to provide, in Theorem 6, an axiomatic foundation for the Smooth Ambiguity preferences introduced by Klibanoff, Marinacci, Mukerji [37], which is related to the one of Klibanoff, Mukerji, and Seo [38], as discussed below.

Finally, the relation between consistency and the recognition of symmetry/similarity patterns in the state space (like in the de Finetti's coin example) is characterized by the concluding Proposition 9.

2 Related Literature

Our interest in this topic was inspired by Epstein and Seo [22] and De Castro and Al-Najjar [15]. At the same time, the independent work of Klibanoff, Mukerji, and Seo [38] shares some of our insights. For this reason, next we discuss these works.

Epstein and Seo Here the framework is the exchangeable one discussed in Example 2. In this framework, their Theorem 3.2 corresponds to the maxmin special case of our variational representation (15). This result is obtained under strong exchangeability which is slightly weaker than our unambiguous symmetry (see Section 6 of this paper and Section 7 of [38]).

Their Theorem 5.2, and especially its version for belief functions, developed in Epstein and Seo [23] as Theorem 4.1, is almost perfectly complementary to our Theorem 7. In fact, they obtain

$$\nu(A) = \int_{\mathrm{Bel}(S)} \theta^{\infty}(A)\, d\mu(\theta)$$

that is, the belief function ν on $\Omega = S \times S \times ...$ is an additive average (μ is a probability measure) of i.i.d. belief functions θ^{∞}, while (in the exchangeable framework for belief functions) our Theorem 7 delivers

$$\nu(A) = \int_{\Delta^{\sigma}(S)} \theta^{\infty}(A)\, d\beta(\theta)$$

[8] Notice that, while the functional V in (4) is real valued, $V : B_0(X) \to \mathbb{R}$, the functional with the same name in (5) is $\mathbb{R}^{|\mathcal{M}|}$ valued, $V : B_0(X) \to \mathbb{R}^{|\mathcal{M}|}$. This difference corresponds to the contrast between the completeness of Choquet expected utility and the incompleteness of Bewley expected utility. On the one hand, this incompleteness delivers the purest framework possible: the one in which attitudes toward ambiguity/prior uncertainty do not matter. On the other hand, the same incompleteness is a weak point in the description of actual choices: even if incomparability of two alternatives is naturally interpreted as choice deferral, eventually a course of action will be taken. So, while the main role of Theorem 8 is to better exemplify the connection between ambiguity and prior uncertainty, this theorem (the last in order of presentation) is an exception in that for all other results we impose completeness of $\succsim$. A related result appears in De Castro and Al-Najjar [15] and is discussed in the next section.

that is, the belief function ν on $\Omega = S \times S \times ...$ is a nonadditive average (β is a belief function) of i.i.d. probability measures θ^∞. In other words, their agent perceives experiments as being indistinguishable (but not necessarily identical) and is neutral to prior uncertainty, our agent thinks the experiments are identical, but she is averse to prior uncertainty. In light of their findings, the intersection of the two representations is the celebrated de Finetti Theorem, corresponding to the case in which ν is an exchangeable probability.

De Castro and Al-Najjar Here, we discuss the relation of our work with De Castro and Al-Najjar [15] and the more recent Al-Najjar and De Castro [2]. We first introduce their contributions and then describe the relationship with the present paper.

In De Castro and Al-Najjar [15] the framework is again a special case of the standard Borel G-space setting discussed in Example 2.[9] Like in [22], the transformations in G are interpreted as capturing similarity patterns.

In [15] they show that, denoting by $\{P_\theta\}_{\theta\in\Theta}$ the set of all ergodic measures, there exists a parameterization $\vartheta : \Omega \to \Theta$ of the state space such that for each (subjectively rational) preference $\succsim$ satisfying

$$f \sim \frac{f \circ \pi_1 + ... + f \circ \pi_n}{n} \text{ for all } f \in B_0(X),\ n \in \mathbb{N},\ \pi_1, .., \pi_n \in G \tag{6}$$

then

$$f \succsim g \text{ if and only if } \int_\Omega f dP_\vartheta \succsim \int_\Omega g dP_\vartheta \tag{7}$$

for all acts f and g. That is, recognition of similarity – as captured by (6) – allows to reduce the complexity of comparisons. In fact, while f and g are $\mathcal{F}$-measurable acts, $\int_\Omega f dP_\vartheta$ and $\int_\Omega g dP_\vartheta$ are acts measurable with respect to the invariant σ-algebra. "... In words, the integrals with respect to the parameters [the r.h.s. of Equation (7)] are sufficient summary of how $\succsim$ ranks all acts ..." as they write in [2]. Building on this result, in the exchangeable framework, [15] obtain Bewley expected utility preferences with an underlying set $\mathcal{C}$ consisting of exchangeable probabilities (hence mixtures of i.i.d. models).

In the second part of the subsequent [2],[10] they assume that there exists a parameterization $\vartheta : \Omega \to \Theta$ of the state space such that (7) holds. This assumption allows to relate the original preference $\succsim$ on state-based acts $h : \omega \mapsto h(\omega)$ with a derived relation $\succsim'$, that they call aggregator, on the corresponding parameter-based acts $H : \theta \mapsto \int_\Omega h dP_{\vartheta(\vartheta^{-1}(\theta))} = \int_\Omega h dP_\theta$ by setting

$$F \succsim' G \overset{\text{def}}{\iff} \int_\Omega f dP_\vartheta \succsim \int_\Omega g dP_\vartheta \tag{8}$$

for all acts f and g. Definition (8) together with (7), which is now considered to be an axiom that the preference $\succsim$ must satisfy, allows them to obtain a parametric version of uncertainty averse preferences *à la* Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [10] and a parametric version of second order subjective expected utility *à la* Neilson [43] and Strzalecki [50], by making the corresponding assumptions on the aggregator $\succsim'$.

Two observations clarify the structural relationship and the main differences between the results in De Castro and Al-Najjar [15] and ours. First, denoting by $\mathcal{P}$ the set of all invariant probability measures of a standard Borel G-space, the triple $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space and $\mathcal{S}(\mathcal{P})$ coincides with the set $\{P_\theta\}_{\theta\in\Theta}$ of all ergodic measures (see Example 2). Second, using (7), it can be shown that, in the exchangeable and in the shift frameworks, the *invariance* axiom (6) implies our consistency axiom (3).

[9] Specifically the exchangeable and the shift frameworks discussed at the end of the example.
[10] The first part subsumes [15] and generalizes its results.

Looking at the relationship between representations, their foundation of Bewley expected utility with an underlying set $\mathcal{C}$ consisting of exchangeable probabilities is a special case of our Theorem 8.

Differently from [15], where (7) follows from the invariance axiom (6), in the second part of [2], the existence of a parameterization $\vartheta : \Omega \to \Theta$ is assumed, and especially the *parameterizability* condition (7) becomes the crucial axiom that the preference $\succsim$ must satisfy. This observation is important to understand the two main differences between Al-Najjar and De Castro [2] and the present paper.

First, as a behavioral condition, the consistency axiom (3) can be shown to be weaker than the parameterizability axiom (7), and it may be argued to be more compelling.[11] This implies that falsifying consistency is easier than falsifying parameterizability. Specifically, falsifying our (3) consists in checking whether the relation

$$f \succcurlyeq_{\mathcal{P}} g \overset{\text{def}}{\iff} \int_\Omega f dP \succsim \int_\Omega g dP \text{ for all } P \in \mathcal{P} \tag{9}$$

is a subrelation of the subjectively rational preference $\succsim$; that is, checking whether $f \succcurlyeq_{\mathcal{P}} g$ implies $f \succsim g$. On the other hand, falsifying their (7) consists checking whether the relation

$$f \succcurlyeq_{\vartheta} g \overset{\text{def}}{\iff} \int_\Omega f dP_\vartheta \succsim \int_\Omega g dP_\vartheta \tag{10}$$

coincides with the subjectively rational preference $\succsim$; that is, checking whether $f \succcurlyeq_\vartheta g$ is equivalent to $f \succsim g$. In this respect, notice that, while the integrals $\int_\Omega f dP$ and $\int_\Omega g dP$ in (9) are just elements of X, the integrals $\int_\Omega f dP_\vartheta$ and $\int_\Omega g dP_\vartheta$ in (10) are acts, $\omega \mapsto \int_\Omega f dP_{\vartheta(\omega)}$ and $\omega \mapsto \int_\Omega g dP_{\vartheta(\omega)}$ for all $\omega \in \Omega$.

The other implication of the different strength of consistency and parameterizability is that the derivation of our representation results becomes more delicate and the proofs highly nontrivial.

Second, although some of the functional forms we study are also studied by [2], the key conceptual and formal differences lie in the kind of axiomatic foundations provided. Consider, for example, our general uncertainty aversion case

$$V(f) = \min_{\mu \in \Delta(\mathcal{S}(\mathcal{P}))} G\left(\int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega u(f(\omega))\, dP(\omega) \right) d\mu(P), \mu \right) \tag{11}$$

and the corresponding

$$V(f) = \min_{\mu \in \Delta(\Theta)} G\left(\int_{\Theta} \left(\int_\Omega u(f(\omega))\, dP_\theta(\omega) \right) d\mu(\theta), \mu \right) \tag{12}$$

of [2].

The purpose of our paper is to show how general classes of preferences under ambiguity – like the uncertainty averse ones – admit an equivalent and unique representation in terms of prior uncertainty – like (11). In the example, the goal of our exercise consists in providing an axiomatic foundation to (11) by making the weakest possible assumptions on the primitive preference $\succsim$ describing state-uncertainty.

In [2], the approach is the opposite. For Al-Najjar and De Castro it is "more convenient to introduce assumptions regarding how the decision maker treats parameter-uncertainty directly on acts defined in terms of parameters." These acts are functions from the parameter space Θ to the consequence set X. In this perspective, they start from prior uncertainty (parameter-uncertainty to be precise) and "construct a preference on the underlying state space." For example, in the uncertainty aversion case, they directly assume the axioms of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [10] on the relation $\succsim'$ between parameter-based acts. These axioms deliver a representation $\mathcal{V}$ of $\succsim'$ of the form

$$\mathcal{V}(F) = \min_{\mu \in \Delta(\Theta)} G\left(\int_\Theta u(F(\theta))\, d\mu(\theta), \mu \right)$$

[11] Indeed also the primitives (the data) are different: a set $\mathcal{P}$ of objectively rational beliefs, in our case, a parameterization ϑ, in theirs.

which translates into representation (12) of $\succsim$ by means of the very definition (8) of $\succsim'$. As explained above, our approach is conceptually opposite (and technically very different).

Klibanoff, Mukerji, and Seo (2010) They work in the exchangeable framework of Epstein and Seo [22]. Their objective is to obtain a behavioral definition of relevant (i.i.d.) models and to study their properties.

In so doing, they characterize preferences that admit a representation of the form

$$V(f) = \mathcal{V}\left(\left(\int_{\Omega} u(f)\, d\theta^{\infty}\right)_{\theta\in\Theta}\right)$$

where Θ is the relevant family of measures on $\Delta^{\sigma}(S)$. Their derivation relies on a symmetry condition that they show to be equivalent to our unambiguous symmetry in the exchangeable case. Finally, under additional behavioral assumptions, they characterize the case in which $\mathcal{V}$ describes the α-Maxmin expected utility model and the one in which $\mathcal{V}$ describes the Smooth Ambiguity preference model. The latter characterization is the counterpart in their setting of our Theorem 6.

This paper From an objective rationality point of view, all of the above papers restrict the agent's probabilistic reasoning to symmetry considerations (reflected by the standard Borel G-space structure). In this paper, any kind of information that can be captured by a set $\mathcal{P}$ of probability measures that makes $(\Omega, \mathcal{F}, \mathcal{P})$ a Dynkin space can be consistently included in the decision process. For example, this allows to easily include Ellsberg urns as shown by Example 1. Beyond the relevance in the ambiguity literature of Ellsberg type paradoxes, urn settings are important since they allow to design experiments that make our consistency axiom easily testable (the first experiment of this kind appears in Eliaz and Ortoleva [20]).

3 Preliminaries

3.1 Decision Theory

Consider a set Ω of *states of the world*, a separable (i.e., countably generated) σ-algebra $\mathcal{F}$ of subsets of Ω called *events*, and a convex set X of *consequences*. We denote by $B_0(X) = B_0(\Omega, \mathcal{F}, X)$ the set of all *(simple) acts*: finite-valued functions $f : \Omega \to X$ which are $\mathcal{F}$-measurable.

Given any $x \in X$, define $x \in B_0(X)$ to be the constant act such that $x(\omega) = x$ for all $\omega \in \Omega$. With this slight abuse of notation, it is possible to identify X with the subset of the constant acts in $B_0(X)$. If $x \in X$, $A \in \mathcal{F}$, and $f \in B_0(X)$, we denote by $xAf \in B_0(X)$ the act taking value x if $\omega \in A$ and $f(\omega)$ if $\omega \notin A$.

Using the linear structure of X we can define, as usual, for every $f, g \in B_0(X)$ and $\alpha \in [0,1]$ the act $\alpha f + (1-\alpha)g \in B_0(X)$, which yields $\alpha f(\omega) + (1-\alpha)g(\omega) \in X$ for all $\omega \in \Omega$.

We model the decision maker's *(subjectively rational) preferences* on $B_0(X)$ by a binary relation $\succsim$. As usual, $\succ$ and $\sim$ denote, respectively, the asymmetric and symmetric parts of $\succsim$. If $f \in B_0(X)$, an element $x_f \in X$ is a *certainty equivalent of* f if $f \sim x_f$. Given a probability P on $\mathcal{F}$ and $f \in B_0(X)$, we denote by $\int_{\Omega} f dP$ the mean of f under P, that is

$$\int_{\Omega} f dP = \sum_{x\in X} P\left(f^{-1}(x)\right) x.$$

Notice that such an integral is well defined since f is (measurable and) finite valued and X is convex.

In particular, notice that if the decision maker is an Anscombe-Aumann expected utility maximizer with beliefs represented by P, then $\int_{\Omega} f dP$ is just the certainty equivalent of act f.

3.2 Probability Theory

Let $\Delta(\Omega) = \Delta(\Omega, \mathcal{F})$ be the set of all finitely additive probabilities on $\mathcal{F}$ and $\Delta^{\sigma}(\Omega) = \Delta^{\sigma}(\Omega, \mathcal{F})$ be the set of all probability measures on $\mathcal{F}$. Both sets and any of their subsets will be endowed with the weak* topology.

Definition 1 (Dynkin, 1978) *Let $\mathcal{P}$ be a nonempty subset of $\Delta^{\sigma}(\Omega)$. The triple $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space if and only if there exist a sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, a set $W \in \mathcal{F}$, and a function*

$$\begin{array}{rccc} p: & \mathcal{F} \times \Omega & \to & [0,1] \\ & (A, \omega) & \mapsto & p(A, \omega) \end{array}$$

such that:

(a) for each $P \in \mathcal{P}$ and $A \in \mathcal{F}$, $p(A, \cdot) : \Omega \to [0,1]$ is a version of the conditional probability of A given $\mathcal{G}$;[12]

(b) for each $\omega \in \Omega$, $p(\cdot, \omega) : \mathcal{F} \to [0,1]$ is a probability measure;

(c) $P(W) = 1$ for all $P \in \mathcal{P}$ and $p(\cdot, \omega) \in \mathcal{P}$ for all $\omega \in W$.

Remark 2 *In the rest of the paper, whenever we assume that $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space and we consider a sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, a set $W \in \mathcal{F}$, and a function $p : \mathcal{F} \times \Omega \to [0,1]$, we assume that $\mathcal{G}$, W, and p are as in Definition 1.*

If requirements (a) and (b) of Definition 1 are satisfied, we say that p is a *common regular conditional probability* for $\mathcal{P}$ given $\mathcal{G}$.[13]

Example 0 (Risk and Uncertainty) The smallest possible $\mathcal{P}$ is a singleton $\{P\}$. Define $\mathcal{G} = \{\emptyset, \Omega\}$, take $W = \Omega$ and $p(A, \omega) = P(A)$ for all $A \in \mathcal{F}$ and for all $\omega \in \Omega$. In this case, $(\Omega, \mathcal{F}, \{P\}) = (\Omega, \mathcal{F}, P)$ is the Dynkin space capturing absence of ambiguity.

The largest possible $\mathcal{P}$ is $\Delta^{\sigma}(\Omega)$. Define $\mathcal{G} = \mathcal{F}$, take $W = \Omega$ and $p(A, \omega) = \delta_{\omega}(A)$ for all $A \in \mathcal{F}$ and for all $\omega \in \Omega$. In this case, $(\Omega, \mathcal{F}, \Delta^{\sigma}(\Omega)) = (\Omega, \mathcal{F})$ is the Dynkin space capturing complete ambiguity, the usual framework of decision making under uncertainty. □

Example 1 (Ellsberg Urn) Consider an urn which contains 90 balls. 30 balls are *r*ed while the remaining 60 balls are either *g*reen or *b*lue. A ball is drawn after an agent bets on its color. Taking Savage's definition of state of the world,[14] the state space is $\Omega = \{(n, c) : n = 0, 1, ..., 60,\ c = r, g, b\}$. For $\omega = (n, c) \in \Omega$, the first component of ω is the number of green balls in the urn, and the second is the color of the extracted ball. Given the structure of the problem, there is not a natural "objective probability" on Ω, but rather a set $\mathcal{P}$ of "objective probabilities". In particular, $\mathcal{P}$ is the set of probabilities of the form

$$P_m(n, c) = \begin{cases} 0 & n \neq m \\ \frac{30}{90} & n = m \text{ and } c = r \\ \frac{m}{90} & n = m \text{ and } c = g \\ \frac{60-m}{90} & n = m \text{ and } c = b \end{cases} \tag{13}$$

[12] That is, $p(A, \cdot)$ is $\mathcal{G}$-measurable and $P(A \cap B) = \int_B p(A, \omega)\, dP(\omega)$ for all $B \in \mathcal{G}$.

[13] Clearly, if $\mathcal{P} = \{P\}$ then this amounts to say that p is a *regular conditional probability* for P given $\mathcal{G}$.

[14] That is, a state of the world is *"a description of the world, leaving no relevant aspect undescribed."*

for all $(n,c) \in \Omega$ and $m = 0, 1, ..., 60$. Graphically, each of the above probabilities takes the following form:

	r	g	b
0	**0**	**0**	**0**
1	**0**	**0**	**0**
...	...	...	...
m	$\frac{30}{90}$	$\frac{m}{90}$	$\frac{60-m}{90}$
...	...	...	...
60	**0**	**0**	**0**

Define $\mathcal{G} = \sigma(\{m\} \times \{r, g, b\} : m = 0, 1, ..., 60)$, $W = \Omega$, and $p : 2^{\Omega} \times \Omega \to [0,1]$ by

$$p(A, (m,d)) = \sum_{(n,c)\in A} P_m(n,c) \qquad \forall A \in 2^{\Omega}, \forall (m,d) \in \Omega.$$

Then $(\Omega, 2^{\Omega}, \mathcal{P})$ is the Dynkin space describing the Ellsberg urn situation. □

Example 2 (Standard Borel G-spaces) A triple $((\Omega, \mathcal{F}), (G, \tau), a)$ is a *standard Borel G-space* if and only if

1. $(\Omega, \mathcal{F})$ is a standard Borel space,
2. (G, τ) is a locally compact group satisfying the second axiom of countability,
3. a is a measurable function

$$\begin{array}{rccc} a: & G \times \Omega & \to & \Omega \\ & (\pi, \omega) & \mapsto & \pi\omega \end{array}$$

such that:

- $\pi(\varkappa\omega) = (\pi\varkappa)\omega$ for all $\pi, \varkappa \in G$ and $\omega \in \Omega$,
- $\epsilon\omega = \omega$ for all $\omega \in \Omega$, where ϵ denotes the unit of G.

In this case, $P \in \Delta^{\sigma}(\Omega)$ is an *invariant measure* if and only if $P(\pi^{-1}A) = P(A)$ for all $\pi \in G$ and for all $A \in \mathcal{F}$. An element $B \in \mathcal{F}$ is an *invariant set* if and only if $\pi^{-1}B = B$ for all $\pi \in G$.

Using the results of Varadarajan [51], Dynkin [19], and Becker and Kechris [5], if we denote by $\mathcal{P}$ the set of all invariant probability measures, then any standard Borel G-space generates a Dynkin space $(\Omega, \mathcal{F}, \mathcal{P})$ where $\mathcal{G}$ is the σ-algebra of all invariant events.

The framework of the celebrated de Finetti Exchangeability Theorem is a particular standard Borel G-space. In this case, $\Omega = \{0,1\}^{\mathbb{N}}$, $\mathcal{F}$ is the σ-algebra generated by all cylinders, G is the group of all finite permutations of $\mathbb{N}$, τ is the discrete topology, and for each $\omega = \{\omega_n\}_{n\in\mathbb{N}} \in \Omega$ and $\pi \in G$, the measurable action is such that $\pi\omega = \{\omega_{\pi(n)}\}_{n\in\mathbb{N}}$. Moreover, the invariant probability measures coincide with the exchangeable ones.

Finally, the definitions of invariant measure and invariant set naturally apply to the case in which a single measurable transformation $T : \Omega \to \Omega$,[15] sometimes called *a shift*, is considered. Also in this case, denoting by $\mathcal{P}$ the invariant probability measures and by $\mathcal{G}$ the invariant events, $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space. □

We endow the set $\Delta^{\sigma}(\Omega)$ with the *natural σ-algebra*

$$\mathcal{A} = \sigma(P \mapsto P(A) : A \in \mathcal{F}),$$

[15] Rather than all the transformations $\omega \mapsto \pi\omega$ induced by the elements π of G.

that is, the smallest σ-algebra that makes the functions of the form

$$\begin{array}{rccc} \langle 1_A, \cdot \rangle : & \Delta^\sigma(\Omega, \mathcal{F}) & \to & \mathbb{R} \\ & P & \mapsto & P(A) \end{array}$$

measurable for all $A \in \mathcal{F}$. Any nonempty subset $\mathcal{P}$ of $\Delta^\sigma(\Omega)$ is endowed with the inherited σ-algebra $\mathcal{A}_\mathcal{P} = \mathcal{A} \cap \mathcal{P}$. For all $\mu \in \Delta^\sigma(\mathcal{P}) = \Delta^\sigma(\mathcal{P}, \mathcal{A}_\mathcal{P})$, the *barycenter* $\bar{\mu}$ of μ is the set function defined by

$$\bar{\mu}(A) = \int_\mathcal{P} P(A)\, d\mu(P) \qquad \forall A \in \mathcal{F}.$$

It is easy to see that $\bar{\mu}$ is a well defined probability measure on $\mathcal{F}$, that is, $\bar{\mu} \in \Delta^\sigma(\Omega)$.

Definition 3 *Let $\emptyset \neq \mathcal{P} \subseteq \Delta^\sigma(\Omega)$. An element P of $\mathcal{P}$ is a* strong extreme point *if and only if the only probability measure μ on $\mathcal{P}$ such that P is the barycenter of μ is δ_P, $\mathcal{S}(\mathcal{P})$ denotes the set of all strong extreme points of $\mathcal{P}$.*

Clearly, strong extreme points are extreme points in the sense of convex analysis. The converse is not in general true, even if this is the case in our examples.[16] Indeed, notice that if $\mathcal{P} = \{P\}$ then $\mathcal{S}(\mathcal{P}) = \{P\}$. If $\mathcal{P} = \Delta^\sigma(\Omega, \mathcal{F})$ then $\mathcal{S}(\mathcal{P}) = \{\delta_\omega\}_{\omega \in \Omega}$. For the Ellsberg urn, we have that

$$\mathcal{P} = \mathcal{S}(\mathcal{P}) = \left\{ \frac{30}{90}\delta_{(m,r)} + \frac{m}{90}\delta_{(m,g)} + \frac{60-m}{90}\delta_{(m,b)} : m = 0, 1, ..., 60 \right\}.$$

Thus, each element of $\mathcal{S}(\mathcal{P})$ is a classic "objective urn model". Finally, for a Standard Borel G-space, we have that

$$\mathcal{S}(\mathcal{P}) = \{P \in \mathcal{P} : P(B) = 0 \text{ or } 1 \text{ for each invariant set } B\}.$$

Thus, $\mathcal{S}(\mathcal{P})$ is the set of all *ergodic measures*. This implies that, in the special exchangeable case, $\mathcal{S}(\mathcal{P}) = \{P_\theta\}_{\theta \in [0,1]}$, where, for each $\theta \in [0,1]$, P_θ on each cylinder $\{\omega_i\}_{i=1}^n = \{\omega_1\} \times \{\omega_2\} \times ... \times \{\omega_n\} \times \{0,1\} \times \{0,1\} \times ...$ is defined by

$$P_\theta\left(\{\omega_i\}_{i=1}^n\right) = \theta^j (1-\theta)^{n-j}$$

and $j = \omega_1 + \omega_2 + ... + \omega_n$. That is, $\mathcal{S}(\mathcal{P})$ is the set of all independent and identically distributed probability measures. Again, each element of $\mathcal{S}(\mathcal{P})$ is a classic "objective coin model".

This suggest the interpretation of the elements of $\mathcal{S}(\mathcal{P})$ as the "pure models" corresponding to a Dynkin space $(\Omega, \mathcal{F}, \mathcal{P})$.

4 Subjective Rationality and Consistency

4.1 Subjective Rationality

Here we regroup some behavioral assumptions on the preference $\succsim$ that formalize the concept of subjective rationality discussed in the introduction.

The first two of them are maintained throughout the paper (unless explicitly stated otherwise). The other is a version of the monotone continuity axiom of Villegas [52] and Arrow [4] which, as in [52] and [4], will be invoked to refine finite additivity into countable additivity. We will not expand on them since they are all very common and very well discussed in the literature of decision making under uncertainty.

The usual rationality requirement of most of microeconomics is that preferences are complete, transitive, nontrivial, and continuous. Formally:

[16] See Appendix B for a more precise discussion of strong extreme points and extreme points and of when the two notions coincide.

Basic Conditions: *The relation* $\succsim$ *is complete, transitive, nontrivial, and such that the sets* $\{\lambda \in [0,1] : \lambda f + (1-\lambda)g \succsim h\}$ *and* $\{\lambda \in [0,1] : h \succsim \lambda f + (1-\lambda)g\}$ *are closed in* $[0,1]$ *for all* $f, g, h \in B_0(X)$.

The Anscombe-Aumann framework where X is usually interpreted as a set of simple lotteries (random outcomes) over a set of prizes Z (deterministic outcomes) and the fact that the decision maker we are modelling uses all the available probabilistic information justify the following assumption:

Risk Independence: *If* $x, y, z \in X$ *and* $\alpha \in (0,1)$, *then*

$$x \succsim y \quad \Leftrightarrow \quad \alpha x + (1-\alpha) z \succsim \alpha y + (1-\alpha) z.$$

Risk Independence allows some degree of separation between risk and ambiguity by imposing a standard independence axiom on constant acts, that is, acts that only involve risk and no state uncertainty. Together with the Basic Conditions it amounts to require the existence of an affine utility function $u : X \to \mathbb{R}$ representing the decision maker's preferences over X.[17,18] Virtually all classes of preferences under ambiguity studied in the literature and framed in the Anscombe-Aumann setting share these two assumptions. The main exception is Bewley's model which restricts completeness to constant acts, while extending independence to the whole $B_0(X)$, see Section 5.4 for details.

Finally, monotone continuity requires that vanishing perturbations of acts cannot affect strict preference.

Monotone Continuity: *If* $f \succ g$ *in* $B_0(X)$, $x \in X$, *and* $E_n \downarrow \emptyset$ *in* $\mathcal{F}$, *then* $xE_N f \succ g$ *and* $f \succ xE_N g$ *for some* $N \in \mathbb{N}$.

4.2 Consistency

The core assumption of our work is the Consistency axiom of [26] which connects the family of objectively rational beliefs $\mathcal{P}$ to the decision maker's subjectively rational preference $\succsim$.

Consistency: *If* $f, g \in B_0(X)$, *then*

$$\int_\Omega f dP \succsim \int_\Omega g dP \quad \forall P \in \mathcal{P} \quad \Rightarrow \quad f \succsim g.$$

Recall that, the maintained Basic Conditions and Risk Independence imply the existence of an affine utility function $u : X \to \mathbb{R}$ representing the decision maker's preferences over X. Together with the linearity of means, this implies that Consistency amounts to require that

$$\int_\Omega u(f)\, dP \geq \int_\Omega u(g)\, dP \quad \forall P \in \mathcal{P} \quad \Rightarrow \quad f \succsim g. \tag{14}$$

The objectively rational estimates of our decision maker about the realizations of Ω belongs to $\mathcal{P}$, but he is unsure which one is the best estimate.[19] Nonetheless, the l.h.s. of (14) reveals that, irrespective of the best estimate, f would be preferred to g. Consistency requires that f is preferred to g even if the best estimate is unknown.

From a different perspective, Consistency means that the decision maker chooses act f over act g, if f dominates g in the game in which he chooses an act in $B_0(X)$ and Nature chooses a model in $\mathcal{P}$.

We next discuss Consistency for two important specifications of the space X of consequences.

[17] To remain with Ellsberg [21, p. 661]: *"Let us suppose that an individual must choose among a certain set of actions, to whose possible consequences we can assign 'von Neumann-Morgenstern utilities'."*

[18] Indeed, a much weaker independence axiom on constant acts is necessary and sufficient for the existence of such a function u, see [34], we opted for the strong form above to facilitate the comparison with the Bewley case discussed below.

[19] Moving to Ellsberg [21, p. 662]: *"In this state of mind, searching for additional grounds for choice, he may try new criteria, ask new questions. For any of the probability distributions in the 'reasonably possible' set [$\mathcal{P}$], he can compute an expected value for each of his actions ..."*

Example 3 (Lotteries) In the original Anscombe-Aumann setting, X is the set of all simple lotteries over a set Z of prizes, that is,

$$X = \left\{ x : Z \to [0,1] : x(z) \neq 0 \text{ for finitely many } z\text{'s in } Z \text{ and } \sum_{z \in Z} x(z) = 1 \right\}.$$

Consider an act $f \in B_0(X)$ and a probability measure $P \in \mathcal{P}$. Under P, act f induces a two stage lottery. First an event $\{f = x\}$ occurs, with probability $P(\{\omega \in \Omega : f(\omega) = x\})$, and then a prize z is paid to the decision maker with probability $x(z)$. In this setting, $\int_\Omega f dP$ is the (one stage) lottery obtained by reducing the previous compound lottery. In fact, $y = \int_\Omega f dP$ is the element of X defined by

$$y(z) = \sum_{x \in X} P(\{\omega \in \Omega : f(\omega) = x\}) x(z) \quad \forall z \in Z.$$

Consistency amounts to impose that: whenever the lottery induced by f under P is preferred to the one induced by g under P for all $P \in \mathcal{P}$, then f must be preferred to g. □

Example 4 (Monetary Outcomes) Another important example is the one in which consequences are quantities of one given good, say money. That is, $X = \mathbb{R}$. In this case, $\int_\Omega f dP$ is the usual expectation of act f, and the meaning of Consistency is clear, provided the decision maker preferences are increasing in the amount of good he consumes. □

It is not hard to show that Consistency implies the classical Monotonicity axiom, another fundamental rationality tenet in decision making under uncertainty.[20]

Monotonicity: *If* $f, g \in B_0(X)$, *then* $f(\omega) \succsim g(\omega)$ *for all* $\omega \in \Omega$ *implies* $f \succsim g$.

Actually, Consistency coincides with Monotonicity when $\mathcal{P} = \Delta^\sigma(\Omega, \mathcal{F})$. Finally, observe that preferences that satisfy the Basic Conditions, Risk Independence, and Monotonicity are such that for each act f in $B_0(X)$ there exists a certainty equivalent x_f in X.

5 Robust Bayesian Preferences

This section contains our main representation results. The reader willing to see the formal versions of the Choquet and Bewley cases discussed in the introduction can skip to the last two Subsections 5.3 and 5.4.

In the order of presentation we first discuss more recent models. The interpretation of our results in terms of prior uncertainty is fully discussed only for the first class of preferences we consider, analogous considerations apply to the others. The differences between the various classes pertain to the qualitative reaction to ambiguity/prior uncertainty.

5.1 Variational Preferences

The class of Variational preferences was introduced by Maccheroni, Marinacci, and Rustichini [40] and it encompasses the Maxmin expected utility preferences of Gilboa and Schmeidler [28] as well as the Multiplier preferences of Hansen and Sargent [33], later axiomatized by Strzalecki [50]. Variational preferences are characterized by the following extra key assumptions:

Uncertainty Aversion: *If* $f, g \in B_0(X)$ *and* $\alpha \in (0,1)$, *then* $f \sim g$ *implies* $\alpha f + (1-\alpha) g \succsim f$.

[20] Under the maintained hypotheses that $\succsim$ satisfies the Basic Conditions and Risk Independence. See also Lemma 28.

Weak Certainty Independence: *If* $f, g \in B_0(X)$, $x, y \in X$, *and* $\alpha \in (0,1)$, *then*

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

The Uncertainty Aversion axiom, introduced by Schmeidler [47] and sometimes called Ambiguity Aversion, is a central axiom of the literature on decision making under ambiguity, and it can be seen as a preference for hedging against ambiguity. See [10] for a detailed discussion.

On the other hand, Weak Certainty Independence is in between Risk Independence (which requires stability of the preference w.r.t. mixing between constant acts only) and full blown Independence (which requires stability of the preference w.r.t. mixing between all acts). In this respect, notice that, for each $h \in B_0(X)$, $x, y \in X$, and $\alpha \in (0,1)$, the graphs of $\alpha f + (1-\alpha) x$ and $\alpha f + (1-\alpha) y$ are congruent up to a translation. That is, convex combinations with different constants (and fixed weights) do not affect the variability of consequences in the different states. See [40] for a detailed discussion.

We further assume that there are either arbitrarily good or arbitrarily bad outcomes. For example, this is automatically satisfied when X contains the set of simple lotteries over $\mathbb{R}$ (see Example 3) and the decision maker is risk averse (or risk neutral, or risk loving).

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

Both in the original representation result of [40] and in the following one, this extra assumption implies that the utility function $u : X \to \mathbb{R}$ is unbounded, which in turn delivers uniqueness of the representation. We are ready to state our first representation result.

Theorem 4 *Let* $(\Omega, \mathcal{F}, \mathcal{P})$ *be a Dynkin space and let* $\succsim$ *be a binary relation on* $B_0(X)$. *The following conditions are equivalent:*

(i) $\succsim$ *satisfies the Basic Conditions, Consistency, Uncertainty Aversion, Weak Certainty Independence, and Unboundedness;*

(ii) there exist an unbounded affine function $u : X \to \mathbb{R}$ *and a grounded lower semicontinuous convex function* $\gamma : \Delta(\mathcal{S}(\mathcal{P})) \to [0, \infty]$ *such that the functional defined by*

$$V(f) = \min_{\mu \in \Delta(\mathcal{S}(\mathcal{P}))} \left\{ \int_{\mathcal{S}(\mathcal{P})} \left(\int_{\Omega} u(f(\omega))\, dP(\omega) \right) d\mu(P) + \gamma(\mu) \right\} \quad \forall f \in B_0(X) \tag{15}$$

represents $\succsim$.

Moreover,

1. u *is cardinally unique and, given* u, γ *is unique;*

2. $\succsim$ *satisfies Monotone Continuity if and only if*

$$\gamma(\mu) = \infty \quad \forall \mu \notin \Delta^{\sigma}(\mathcal{S}(\mathcal{P})).$$

The interpretation of Theorem 4 in terms of prior uncertainty is quite straightforward: starting from the set $\mathcal{P}$ of objectively rational beliefs, our decision maker identifies the pure models $\mathcal{S}(\mathcal{P})$ and considers all the possible priors over the models, the elements of $\Delta(\mathcal{S}(\mathcal{P}))$.

If, as it is the case in Bayesian statistics, he is able to identify a single prior $\mu_0 \in \Delta^\sigma(\mathcal{S}(\mathcal{P}))$, he simply evaluates each act according to its expected utility

$$V(f) = \int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega u(f(\omega))\, dP(\omega) \right) d\mu_0(P) \tag{16}$$

which is the counterpart of the celebrated de Finetti Exchangeability Theorem, and corresponds to select

$$\gamma(\mu) = \begin{cases} 0 & \mu = \mu_0 \\ \infty & \mu \neq \mu_0 \end{cases}$$

in representation (15). In other words, all priors μ, except μ_0, are excluded and given the highest possible penalty $\gamma(\mu) = \infty$.[21]

When, as it happens in robust Bayesian statistics, see Huber and Strassen [36] and Huber and Ronchetti [35], there is some degree of uncertainty about the prior μ_0, the decision maker can consider a (compact and convex) neighborhood $\mathcal{N}$ of μ_0, for example an ε-contamination neighborhood or a neighborhood relative to some convex statistical distance. Then, uncertainty aversion induces him to consider the minimum expected utility over $\mathcal{N}$. This leads to the evaluation

$$V(f) = \min_{\mu \in \mathcal{N}} \int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega u(f(\omega))\, dP(\omega) \right) d\mu(P)$$

and corresponds to select

$$\gamma(\mu) = \begin{cases} 0 & \mu \in \mathcal{N} \\ \infty & \mu \notin \mathcal{N} \end{cases}$$

in representation (15). In other words, all priors μ, except those in $\mathcal{N}$, are excluded and given the highest possible penalty $\gamma(\mu) = \infty$.[22]

In the same robustness perspective, the decision maker can choose a convex statistical distance, say the relative entropy $R(\mu||\mu_0)$, and penalize alternative priors proportionally to their distance from μ_0.[23] This corresponds to select

$$\gamma(\mu) \propto R(\mu||\mu_0)$$

in representation (15).[24]

Summing up, the robust approach here is incarnated by minimization of expected utility over prior probabilities where the soft constraint is given by the penalty function γ, which thus captures the degree of confidence of the decision maker on the various priors.

5.1.1 Uncertainty Averse Preferences

The robust Bayesian representation discussed for Variational preferences admits an extension that only relies on Uncertainty Aversion, but requires a strengthening of the Unboundedness assumption. The new requirement is that there are arbitrarily good *and* arbitrarily bad outcomes. Formally:

Full Unboundedness: *There exist $x \succ y$ in X such that for all $\alpha \in (0,1)$ there exist $z, z' \in X$ that satisfy $y \succ \alpha z + (1-\alpha) x$ and $\alpha z' + (1-\alpha) y \succ x$.*

[21] In terms of behavioral assumptions this amounts to add Independence (see Subsection 5.4) and Monotone Continuity.

[22] In terms of behavioral assumptions this amounts to add Certainty Independence (see Gilboa and Schmeidler [28]).

[23] Recall that $R(\mu\|\mu_0) = \int_{\mathcal{S}(\mathcal{P})} \frac{d\mu}{d\mu_0} \log\left(\frac{d\mu}{d\mu_0}\right) d\mu_0$ if μ is countably additive and absolutely continuous with respect to μ_0 and $R(\mu\|\mu_0) = \infty$ otherwise.

[24] In terms of behavioral assumptions this amounts to add the assumptions of Theorem 6 (see [10]).

In the representation this extra assumption implies that the utility function $u : X \to \mathbb{R}$ is onto (i.e., $u(X) = \mathbb{R}$) which again delivers the uniqueness of the representation.

In order to state the main result, we have to introduce some notions related to quasiconcave duality.

We define $\mathcal{L}(\mathbb{R} \times \Delta(\mathcal{S}(\mathcal{P})))$ to be the class of functions $G : \mathbb{R} \times \Delta(\mathcal{S}(\mathcal{P})) \to (-\infty, \infty]$ such that

- G is quasiconvex and lower semicontinuous;
- $G(\cdot, \mu)$ is increasing for all $\mu \in \Delta(\mathcal{S}(\mathcal{P}))$;
- $t = \min_{\mu \in \Delta(\mathcal{S}(\mathcal{P}))} G(t, \mu)$ for all $t \in \mathbb{R}$.

Given a function $G \in \mathcal{L}(\mathbb{R} \times \Delta(\mathcal{S}(\mathcal{P})))$, we say that G is linearly continuous if and only if the functional $I : B_0(\mathbb{R}) \to \mathbb{R}$ defined by

$$I(\varphi) = \min_{\mu \in \Delta(\mathcal{S}(\mathcal{P}))} G\left(\int_{\mathcal{S}(\mathcal{P})} \left(\int_{\Omega} \varphi(\omega)\, dP(\omega)\right) d\mu(P), \mu\right) \quad \forall \varphi \in B_0(\mathbb{R})$$

is continuous.

Theorem 5 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and let $\succsim$ be a binary relation on $B_0(X)$. The following conditions are equivalent:*

(i) $\succsim$ satisfies the Basic Conditions, Consistency, Uncertainty Aversion, Risk Independence, and Full Unboundedness;

(ii) there exist an onto affine function $u : X \to \mathbb{R}$ and a linearly continuous function G in $\mathcal{L}(\mathbb{R} \times \Delta(\mathcal{S}(\mathcal{P})))$ such that the functional defined by

$$V(f) = \min_{\mu \in \Delta(\mathcal{S}(\mathcal{P}))} G\left(\int_{\mathcal{S}(\mathcal{P})} \left(\int_{\Omega} u(f(\omega))\, dP(\omega)\right) d\mu(P), \mu\right) \quad \forall f \in B_0(X) \tag{17}$$

represents $\succsim$.

Moreover,

1. *u is cardinally unique and, given u, G is unique;*
2. *$\succsim$ satisfies Monotone Continuity if and only if*

$$G(t, \mu) = \infty \quad \forall (t, \mu) \notin \mathbb{R} \times \Delta^{\sigma}(\mathcal{S}(\mathcal{P})).$$

Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [10] show that the functions G in $\mathcal{L}(\mathbb{R} \times \Delta(\mathcal{S}(\mathcal{P})))$ can be interpreted as games against Nature, thus Theorem 5 can be seen as providing the basis for a very general maxmin approach to robust Bayesian decision theory.[25] Moreover, in line with the results contained in [10], it can be shown that G captures the comparative ambiguity/prior uncertainty attitudes of the decision maker.

[25] As Huber and Ronchetti [35, p. 17] write *"... as we defined robustness to mean insensitivity with regard to small deviations from the assumptions, any quantitative measure of robustness must somehow be concerned with the maximum degradation of performance possible for an ε-deviation from the assumptions. An optimally robust procedure then minimizes this maximum degradation and hence will be a minimax procedure of some kind."* See also Grünwald and Dawid [31].

5.2 Smooth Ambiguity Preferences

The structure of Dynkin space and the axiom of Consistency allow us to provide an axiomatic foundation of the Smooth Ambiguity preference model proposed by Klibanoff, Marinacci, Mukerji [37].[26]

Theorem 6 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and let $\succsim$ be a binary relation on $B_0(X)$. The following conditions are equivalent:*

(i) $\succsim$ satisfies the Basic Conditions, Consistency, Risk Independence, Monotone Continuity, and its restriction to $\mathcal{G}$-measurable acts satisfies P2-P6 of Savage [45];

(ii) there exist a nonconstant affine function $u : X \to \mathbb{R}$, a strictly increasing and continuous function $\phi : u(X) \to \mathbb{R}$, and a nonatomic $\mu \in \Delta^{\sigma}(\mathcal{S}(\mathcal{P}))$ such that the functional defined by

$$V(f) = \int_{\mathcal{S}(\mathcal{P})} \phi\left(\int_{\Omega} u(f(\omega))\, dP(\omega)\right) d\mu(P) \quad \forall f \in B_0(X) \tag{18}$$

represents $\succsim$.

Moreover,

1. *u is cardinally unique, μ is unique, and, given u, ϕ is cardinally unique;*
2. *ϕ is concave if and only if $\succsim$ satisfies Uncertainty Aversion.*

Notice that, in the perspective of this paper, standard expected utility (16) corresponds to the case in which ϕ is the identity. Thus, if we interpret $\int_{\Omega} u(f(\omega))\, dP(\omega)$ as the payoff associated by the decision maker to act f if the pure model P is true, this means that he his neutral to model risk (that is the risk involved by the fact that only the prior μ over pure models is known to him). Theorem 6, corresponds to a decision maker who, being able to form a single prior, is not neutral to model risk, but has an attitude towards it that is captured by the curvature of ϕ (as exemplified by point 2. of the statement).

Finally, observe that in the original version of [37], the outer integral (w.r.t. μ) in (18) is over the entire set $\Delta^{\sigma}(\Omega)$. The restriction to $\mathcal{S}(\mathcal{P})$ is natural given our interpretation of $\mathcal{S}(\mathcal{P})$ in terms of pure models and, mathematically, corresponds to constraining the support of μ in [37].

5.3 Choquet Expected Utility Preferences

In the theory of choice under ambiguity it is impossible to overestimate the importance in terms of innovative import of Choquet expected utility. Developed by Schmeidler [47], it is arguably the first behaviorally and mathematically sound reply to Ellsberg's critique.

Recall that two acts f and g are said to be comonotonic if and only if $f(\omega) \succ f(v)$ and $g(v) \succ g(\omega)$ for no $\omega, v \in \Omega$.

Comonotonic Independence: *If $f, g, h \in B_0(X)$ are pairwise comonotonic and $\alpha \in (0,1)$, then*

$$f \succ g \quad \Rightarrow \quad \alpha f + (1-\alpha)h \succ \alpha g + (1-\alpha)h. \tag{19}$$

Theorem 7 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and let $\succsim$ be a binary relation on $B_0(X)$. The following conditions are equivalent:*

[26] For an axiomatic foundation of the Smooth Ambiguity preference model, in a different setting, see also Seo [48].

(i) $\succsim$ *satisfies the Basic Conditions, Consistency, and Comonotonic Independence on* $\mathcal{G}$*-measurable acts;*

(ii) there exist a nonconstant affine function $u : X \to \mathbb{R}$ *and a capacity* ν *on* $\mathcal{A}_{\mathcal{S}(\mathcal{P})}$ *such that the functional defined by*

$$V(f) = \int_{\mathcal{S}(\mathcal{P})} \left(\int_{\Omega} u(f(\omega))\, dP(\omega) \right) d\nu(P) \quad \forall f \in B_0(X) \tag{20}$$

represents $\succsim$.

Moreover,

1. u *is cardinally unique and* ν *is unique;*
2. ν *is convex if and only if* $\succsim$ *satisfies Uncertainty Aversion;*
3. ν *is continuous if and only if* $\succsim$ *satisfies Monotone Continuity.*

Here prior uncertainty is captured by the lack of additivity of the "prior" capacity, exactly like in Huber and Strassen [36]. A weakening of the axioms allows to obtain a more general representation where model-contingent payoffs $\int_{\Omega} u(f(\omega))\, dP(\omega)$ are weighted by a function ϕ as in Theorem 6.[27]

Choquet integrals over sets of probability measures have been studied by Giraud [29] and Amarante [3]. Their setups and assumptions are quite far from our Dynkin space setting, with its objective rationality interpretation. As a consequence, their representations and analyses are very different.

5.4 Bewley Expected Utility Preferences

In this section, we consider another classical model of ambiguity, which was proposed by Bewley in the same years in which Schmeidler was refining Choquet expected utility.

The main difference with the rest of the paper is that in Bewley's view ambiguity can be seen as a source of incompleteness of the preference $\succsim$. In terms of behavior, this can be revealed by the decision of postponing the choice between some pairs of alternatives (which are thus incomparable, due to lack or poor quality of information).[28] For sake of brevity, in this subsection, we will say that $\succsim$ satisfies the *Weak Basic Conditions* if and only if it satisfies the Basic Conditions where the assumption of completeness on $B_0(X)$ is replaced by the weaker assumption of completeness on X. Moreover, we will say that it satisfies *Binary Monotone Continuity* if and only if it satisfies Monotone Continuity when f and g are constant acts. At the same time, Risk Independence is replaced by the full blown independence axiom.

Independence: *If* $f, g, h \in B_0(X)$ *and* $\alpha \in (0,1)$, *then*

$$f \succsim g \quad \Leftrightarrow \quad \alpha f + (1-\alpha) h \succsim \alpha g + (1-\alpha) h.$$

Theorem 8 *Let* $(\Omega, \mathcal{F}, \mathcal{P})$ *be a Dynkin space and let* $\succsim$ *be a binary relation on* $B_0(X)$. *The following conditions are equivalent:*

(i) $\succsim$ *satisfies the Weak Basic Conditions, Consistency, Independence, and Binary Monotone Continuity;*

[27] This is actually one of the objects of the authors' current research.

[28] See, for example, Danan and Ziegelmeyer [14] and Kopylov [39].

(ii) there exist a nonconstant affine function $u : X \to \mathbb{R}$ and a nonempty, compact, and convex set $\mathcal{M}$ of $\Delta^\sigma(\mathcal{S}(\mathcal{P}))$ such that for each f and g in $B_0(X)$

$$f \succsim g \Leftrightarrow \int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega u(f(\omega))\, dP(\omega) \right) d\mu(P) \geq \int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega u(g(\omega))\, dP(\omega) \right) d\mu(P) \tag{21}$$

for all $\mu \in \mathcal{M}$.

Moreover, u is cardinally unique and $\mathcal{M}$ is unique.

We are in the purest multiple priors setting, the one in which the incompleteness of information generates incompleteness of the preference, which in turn is reflected by the impossibility of pinning down a single prior probability on $\mathcal{S}(\mathcal{P})$.

6 Consistency and Unambiguous Symmetry

The role of Consistency is to connect (probabilistic) information and choice behavior. In some special cases, information naturally comes in the form of symmetry/similarity considerations about the state space, and Consistency follows from the recognition of the implied patterns.[29] In this case, it is natural to consider a standard Borel G-space as in Example 2 and the *revealed unambiguous preference* $\succsim^*$ introduced by Ghirardato, Maccheroni, Marinacci [24], that is,

$$f \succsim^* g \quad \Leftrightarrow \quad \alpha f + (1-\alpha) h \succsim \alpha g + (1-\alpha) h \quad \forall \alpha \in [0,1], \forall h \in B_0(X).$$

It is immediate to see that $\succsim^*$ is derived from the primitive $\succsim$. Specifically, it is the part of $\succsim$ which is not affected by hedging considerations.

We next introduce *unambiguous symmetry*, the main behavioral assumption of this section.

Unambiguous Symmetry: *If $f \in B_0(X)$ and $\pi \in G$, then $f \circ \pi \sim^* f$.*

We are ready to state the last result of the paper.

Proposition 9 *Let $((\Omega, \mathcal{F}), (G, \tau), a)$ be a standard Borel G-space, $\mathcal{P}$ be the set of invariant probability measures, and $\succsim$ be a binary relation on $B_0(X)$. If $\succsim$ satisfies the Basic Conditions, Risk Independence, and $\succsim^*$ satisfies Binary Monotone Continuity, then the following statements are equivalent:*

(i) $\succsim$ satisfies Consistency;

(ii) $\succsim$ satisfies Monotonicity and Unambiguous Symmetry.

In other words, in this particular setup, Consistency incorporates subjective similarity assessments into the preferences.

[29] See the discussion of Epstein and Seo [22], De Castro and Al-Najjar [15], and Klibanoff, Mukerji, and Seo [38] in Section 2.

A Common Regular Conditional Probabilities

In this section, we consider $(\Omega, \mathcal{F})$, a measurable space, and $\mathcal{G}$ a sub-σ-field of $\mathcal{F}$. As usual, $\Delta^\sigma(\Omega, \mathcal{F})$ is the set of all probability measures on $\mathcal{F}$. We denote by $B(\Omega)$ the set of all bounded functions from Ω to $\mathbb{R}$ and by $B(\Omega, \mathcal{F})$ the set of all bounded and $\mathcal{F}$-measurable functions. Both spaces are endowed with the supnorm. Given an element $P \in \Delta^\sigma(\Omega, \mathcal{F})$, we denote by $\mathcal{L}^1(\Omega, \mathcal{F}, P)$ the set of all $\mathcal{F}$-measurable and P-integrable functions from Ω to $\mathbb{R}$. Finally, if $P \in \Delta^\sigma(\Omega, \mathcal{F})$ then we denote by $P_\mathcal{G}$ the restriction of P to $\mathcal{G}$. Notice that $P_\mathcal{G} \in \Delta^\sigma(\Omega, \mathcal{G})$.

Let P be an element of $\Delta^\sigma(\Omega, \mathcal{F})$, $\mathcal{G}$ a sub-σ-field of $\mathcal{F}$, and f an element of $\mathcal{L}^1(\Omega, \mathcal{F}, P)$. By [8, Ch. 6], there exists $g \in \mathcal{L}^1(\Omega, \mathcal{G}, P_\mathcal{G})$ such that

$$\int_G f dP = \int_G g dP_\mathcal{G} \qquad \forall G \in \mathcal{G}.$$

Any such g is called *a version of the conditional expected value of f given $\mathcal{G}$*. The set of all these functions g is denoted by $E_P[f|\mathcal{G}]$. $E_P[f|\mathcal{G}]$ forms an equivalence class of $\mathcal{L}^1(\Omega, \mathcal{G}, P_\mathcal{G})$ with respect to the $P_\mathcal{G}$-a.s. equality. If $f = 1_A$ for some $A \in \mathcal{F}$ then $E_P[f|\mathcal{G}]$ is denoted by $P[A|\mathcal{G}]$ and any of its elements is called *a version of the conditional probability of A given $\mathcal{G}$*.

Definition 10 *Let P be an element of $\Delta^\sigma(\Omega, \mathcal{F})$ and $\mathcal{G}$ a sub-σ-field of $\mathcal{F}$. A* regular conditional probability (r.c.p.) for P given $\mathcal{G}$ *is a function $p : \mathcal{F} \times \Omega \to [0, 1]$ such that*

- *for each $A \in \mathcal{F}$, $p(A, \cdot) : \Omega \to [0, 1]$ is a version of the conditional probability of A given $\mathcal{G}$ (that is, $p(A, \cdot) \in P[A|\mathcal{G}]$);*
- *for each $\omega \in \Omega$, $p(\cdot, \omega) : \mathcal{F} \to [0, 1]$ is a probability measure (that is, $p^\omega = p(\cdot, \omega) \in \Delta^\sigma(\Omega, \mathcal{F})$).*

The measurable space $(\Omega, \mathcal{F})$ is a *standard Borel space* if and only if it is isomorphic to a measurable space $(I, \mathcal{B}orel(I))$ for some Borel set I of a Polish space. That is, there exists a bimeasurable bijection $\varphi : (\Omega, \mathcal{F}) \to (I, \mathcal{B}orel(I))$.

Theorem 11 ([30, Ch. 6]) *Let P be an element of $\Delta^\sigma(\Omega, \mathcal{F})$ and $\mathcal{G}$ a sub-σ-field of $\mathcal{F}$. If $(\Omega, \mathcal{F})$ is a standard Borel space then there exists a regular conditional probability for P given $\mathcal{G}$.*

Theorem 12 *Let $(\Omega, \mathcal{F})$ be a standard Borel space and $\mathcal{G}$ a sub-σ-field of $\mathcal{F}$. The following statements are equivalent for $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$:*

(i) $\bigcap_{P \in \mathcal{P}} E_P[f|\mathcal{G}] \neq \emptyset$ *for all* $f \in B(\Omega, \mathcal{F})$.

(ii) $\bigcap_{P \in \mathcal{P}} P[A|\mathcal{G}] \neq \emptyset$ *for all* $A \in \mathcal{F}$.

(iii) The elements of $\mathcal{P}$ admit a common regular conditional probability given $\mathcal{G}$.

In this case, if p and p' are two common regular conditional probabilities for $\mathcal{P}$ given $\mathcal{G}$ then

$$p(\cdot, \omega) = p'(\cdot, \omega) \qquad \mathcal{P}\text{-a.s.},$$

that is, $\{\omega \in \Omega : p(\cdot, \omega) = p'(\cdot, \omega)\} \in \mathcal{G}$ and $P(\{\omega \in \Omega : p(\cdot, \omega) = p'(\cdot, \omega)\}) = 1$ for all $P \in \mathcal{P}$.

Proof. Clearly, (i) implies (ii).

(ii) implies (iii). We proceed by steps.

Step 1: For each $A \in \mathcal{F}$ and for each $g_A \in \bigcap_{P \in \mathcal{P}} P[A|\mathcal{G}]$ we have that $(g_A \vee 0) \wedge 1 \in \bigcap_{P \in \mathcal{P}} P[A|\mathcal{G}]$.

Proof of the Step.

For each $A \in \mathcal{F}$ let g_A be an element of $\bigcap_{P\in\mathcal{P}} P[A|\mathcal{G}]$. It follows that $g_A^+ \in \bigcap_{P\in\mathcal{P}} P[A|\mathcal{G}]$. Indeed, for each $P \in \mathcal{P}$ we have that

$$\int_G g_A dP_{\mathcal{G}} = P(A \cap G) \geq 0 \qquad \forall G \in \mathcal{G}.$$

Thus, we have that $P_{\mathcal{G}}(\{\omega \in \Omega : g_A(\omega) < 0\}) = 0$ and $g_A = g_A^+$ $P_{\mathcal{G}}$-a.s., that is P-a.s., hence $g_A \vee 0 = g_A^+ \in P[A|\mathcal{G}]$. Without loss of generality, we can assume that $g_A(\omega) \geq 0$ for all $\omega \in \Omega$. Next, we show that $g_A \wedge 1 \in \bigcap_{P\in\mathcal{P}} P[A|\mathcal{G}]$. Indeed, for each $P \in \mathcal{P}$

$$\int_G g_A dP_{\mathcal{G}} = P(A \cap G) \leq P(G) = \int_G 1_\Omega dP_{\mathcal{G}} \qquad \forall G \in \mathcal{G}.$$

Thus, we have that $P_{\mathcal{G}}(\{\omega \in \Omega : g_A(\omega) > 1\}) = 0$ and $g_A = g_A \wedge 1$ $P_{\mathcal{G}}$-a.s., that is P-a.s., hence $g_A \wedge 1 \in P[A|\mathcal{G}]$. □

In light of Step 1, for each $A \in \mathcal{F}$ fix an element $\hat{g}_A \in \bigcap_{P\in\mathcal{P}} P[A|\mathcal{G}]$ such that $\hat{g}_A(\omega) \in [0,1]$ for all $\omega \in \Omega$. Given $\hat{g}_\emptyset \in \bigcap_{P\in\mathcal{P}} P[\emptyset|\mathcal{G}]$ and $\hat{g}_\Omega \in \bigcap_{P\in\mathcal{P}} P[\Omega|\mathcal{G}]$, define $N_0 = \{\omega \in \Omega : \hat{g}_\emptyset(\omega) \neq 0\}$ and $N_1 = \{\omega \in \Omega : \hat{g}_\Omega(\omega) \neq 1\}$.

Step 2: Given $i \in \{0,1\}$, we have that $P(N_i) = 0$ for all $P \in \mathcal{P}$.

Proof of the Step.

By the choice of $\hat{g}_\emptyset$ and $\hat{g}_\Omega$, we have that $1_\Omega \geq \hat{g}_\Omega, \hat{g}_\emptyset \geq 0$. Since $\hat{g}_\emptyset \in P[\emptyset|\mathcal{G}] = E_P[0|\mathcal{G}]$ for all $P \in \mathcal{P}$, we have that

$$\int_G \hat{g}_\emptyset dP_{\mathcal{G}} = 0 \qquad \forall G \in \mathcal{G}.$$

This implies that $\hat{g}_\emptyset = 0$ $P_{\mathcal{G}}$-a.s., that is P-a.s., for all $P \in \mathcal{P}$. It follows that $P(N_0) = 0$ for all $P \in \mathcal{P}$.

Since $\hat{g}_\Omega \in P[\Omega|\mathcal{G}] = E_P[1_\Omega|\mathcal{G}]$ for all $P \in \mathcal{P}$, we have that

$$\int_G \hat{g}_\Omega dP_{\mathcal{G}} = 1 \qquad \forall G \in \mathcal{G}.$$

This implies that $\hat{g}_\Omega = 1$ $P_{\mathcal{G}}$-a.s., that is P-a.s., for all $P \in \mathcal{P}$. It follows that $P(N_1) = 0$ for all $P \in \mathcal{P}$. □

By [30, Corollary 3.4 and Theorem 4.3] and since $(\Omega, \mathcal{F})$ is a standard Borel space, there exists an at most countable field $\mathcal{A}$ generating $\mathcal{F}$ such that every finitely additive probability on $\mathcal{A}$ extends (uniquely) to a probability measure on $\mathcal{F}$. Since $\mathcal{A} \times \mathcal{A}$ is at most countable, let $\{(A_n, B_n)\}_{n=2}^\infty$ be the collection of all disjoint pairs of elements of $\mathcal{A}$ and set

$$N_n = \{\omega \in \Omega : \hat{g}_{A_n \cup B_n}(\omega) \neq \hat{g}_{A_n}(\omega) + \hat{g}_{B_n}(\omega)\} \qquad \forall n \geq 2.$$

Step 3: $P(N_n) = 0$ for all $P \in \mathcal{P}$ and for all $n \geq 2$.

Proof of the Step.

By contradiction, assume that there exists a $P \in \mathcal{P}$ and $n \geq 2$ such that $P(N_n) > 0$. Since $\hat{g}_{A_n \cup B_n} \in P[A_n \cup B_n|\mathcal{G}]$, for each $G \in \mathcal{G}$ we have that

$$\begin{aligned}\int_G \hat{g}_{A_n \cup B_n} dP_{\mathcal{G}} &= P((A_n \cup B_n) \cap G) = P((A_n \cap G) \cup (B_n \cap G)) \\ &= P(A_n \cap G) + P(B_n \cap G) = \int_G \hat{g}_{A_n} dP_{\mathcal{G}} + \int_G \hat{g}_{B_n} dP_{\mathcal{G}} \\ &= \int_G (\hat{g}_{A_n} + \hat{g}_{B_n}) dP_{\mathcal{G}},\end{aligned}$$

which implies that $\hat{g}_{A_n} + \hat{g}_{B_n} \in P[A_n \cup B_n|\mathcal{G}]$. Thus, we have that $\hat{g}_{A_n} + \hat{g}_{B_n} = \hat{g}_{A_n \cup B_n}$ $P_{\mathcal{G}}$-a.s., that is P-a.s., which contradicts $P(N_n) > 0$. □

Let $U = \left(\bigcup_{n=0}^{\infty} N_n\right)^c$. By Step 2 and Step 3, it follows that $U \in \mathcal{G}$ and $P(U) = 1$ for all $P \in \mathcal{P}$. Moreover, we have that for each $\omega \in U$:

$$\begin{aligned} \hat{g}_A(\omega) &\in [0,1] \qquad \forall A \in \mathcal{A}, \\ \hat{g}_\emptyset(\omega) &= 0, \\ \hat{g}_\Omega(\omega) &= 1, \\ \hat{g}_{A\cup B}(\omega) &= \hat{g}_A(\omega) + \hat{g}_B(\omega) \qquad \forall A, B \in \mathcal{A} \text{ such that } A \cap B = \emptyset. \end{aligned}$$

That is, for each $\omega \in U$ the mapping $A \mapsto \hat{g}_A(\omega)$ is a finitely additive probability measure on $\mathcal{A}$. Let Q be a fixed element of $\mathcal{P}$ and set for each $A \in \mathcal{A}$ and $\omega \in \Omega$

$$\tilde{g}(A,\omega) = \begin{cases} \hat{g}_A(\omega) & \text{if } \omega \in U \\ Q(A) & \text{if } \omega \in U^c \end{cases}.$$

For all $P \in \mathcal{P}$, $\tilde{g} : \mathcal{A} \times \Omega \to [0,1]$ is a function such that:

- for each $A \in \mathcal{A}$, $\tilde{g}(A,\cdot)$ is $\mathcal{G}$-measurable and $\tilde{g}(A,\cdot) = g_A(\cdot)$ $P_{\mathcal{G}}$-a.s., that is, $\tilde{g}(A,\cdot) \in P[A|\mathcal{G}]$;
- for each $\omega \in \Omega$, $\tilde{g}(\cdot,\omega) : \mathcal{A} \to [0,1]$ is a finitely additive probability measure on $\mathcal{A}$.

For each $\omega \in \Omega$ denote by $\hat{g}(\cdot,\omega)$ the unique probability measure that extends $\tilde{g}(\cdot,\omega)$ from $\mathcal{A}$ to $\mathcal{F}$.

Step 4: The mapping $(A,\omega) \mapsto \hat{g}(A,\omega)$ is a regular conditional probability for all $P \in \mathcal{P}$ given $\mathcal{G}$.

Proof of the Step.

Since $\mathcal{F}$ is a standard Borel Space and by [30, Corollary 3.4 and Theorem 4.3], $\hat{g}$ is well defined.[30] Let $\mathcal{M} = \{A \in \mathcal{F} : \hat{g}(A,\cdot) \in P[A|\mathcal{G}]\}$. By construction, notice that:

- for each $A \in \mathcal{A}$, $\hat{g}(A,\omega) = \tilde{g}(A,\omega)$ for all $\omega \in \Omega$ and $\hat{g}(A,\cdot) = \tilde{g}(A,\cdot) \in P[A|\mathcal{G}]$;
- for each $\omega \in \Omega$, $\hat{g}(\cdot,\omega) : \mathcal{F} \to [0,1]$ is a probability measure on $\mathcal{F}$.

This implies that $\mathcal{A} \subseteq \mathcal{M}$. Consider $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}$ and $A_n \nearrow A$ (resp., $\searrow$). Since $\mathcal{F}$ is a σ-algebra and $\hat{g}(\cdot,\omega)$ is a measure for all $\omega \in \Omega$, we have that $A \in \mathcal{F}$ and $\hat{g}(A_n,\omega) \nearrow \hat{g}(A,\omega)$ (resp., $\searrow$) for all $\omega \in \Omega$. It follows that $\hat{g}(A_n,\cdot) \nearrow \hat{g}(A,\cdot)$ (resp., $\searrow$). Since $\hat{g}$ is bounded, we thus obtain that $\hat{g}(A,\cdot) \in \mathcal{L}^1(\Omega,\mathcal{G},P_{\mathcal{G}})$. By the Monotone Convergence Theorem, we conclude that for each $G \in \mathcal{G}$,

$$\int_G \hat{g}(A,\cdot)\, dP_{\mathcal{G}} = \lim_{n\to\infty} \int_G \hat{g}(A_n,\cdot)\, dP_{\mathcal{G}} = \lim_{n\to\infty} P(A_n \cap G) = P(A \cap G).$$

That is, we obtain that $\hat{g}(A,\cdot) \in P[A|\mathcal{G}]$. By the Monotone Class Lemma (see, e.g., [8, Theorem 3.4]) and since $\mathcal{M}$ is a monotone class containing $\mathcal{A}$, it follows that $\mathcal{M} \supseteq \sigma(\mathcal{A}) = \mathcal{F}$. □

Step 4 proves the implication.

(iii) implies (i). Let $p : \mathcal{F} \times \Omega \to [0,1]$ be a common regular conditional probability for all the elements P of $\mathcal{P}$ given $\mathcal{G}$. Then, for each $P \in \mathcal{P}$ we have that:

- for each $A \in \mathcal{F}$, $p(A,\cdot) \in P[A|\mathcal{G}]$;
- for each $\omega \in \Omega$, $p^\omega \in \Delta^\sigma(\Omega,\mathcal{F})$.

[30] Since $(\Omega,\mathcal{F})$ is standard, $\tilde{g}(\cdot,\omega)$ is a countably additive probability measure on $\mathcal{A}$ for each $\omega \in \Omega$. By the Caratheodory theorem (see, e.g., [8, Theorem 3.3 and Theorem 11.2]), it extends uniquely to $\mathcal{F} = \sigma(\mathcal{A})$.

Define

$$\begin{array}{rccc} ^\star: & B(\Omega,\mathcal{F}) & \to & B(\Omega,\mathcal{G}) \\ & f & \mapsto & f^\star \end{array}$$

where $f^\star(\omega) = \int_\Omega f dp^\omega$ for all $\omega \in \Omega$ and for all $f \in B(\Omega, \mathcal{F})$. In the rest of the proof, we prove that such a mapping is well defined, linear, (Lipschitz) continuous, and such that $f^\star \in E_P[f|\mathcal{G}]$ for all $f \in B(\Omega, \mathcal{F})$ and for all $\mathcal{P} \in P$.

Claim: The mapping $^\star : B(\Omega, \mathcal{F}) \to B(\Omega)$ *is linear and continuous.*

Proof of the Claim.

We first observe that the mapping $^\star : B(\Omega, \mathcal{F}) \to B(\Omega)$ is well defined and linear. Indeed,

$$|f^\star(\omega)| = \left| \int_\Omega f dp^\omega \right| \leq \|f\| \qquad \forall \omega \in \Omega.$$

Moreover, for each $\alpha, \beta \in \mathbb{R}$ and for each $f, g \in B(\Omega, \mathcal{F})$, we have that $\alpha f + \beta g \in B(\Omega, \mathcal{F})$ and

$$(\alpha f + \beta g)^\star(\omega) = \int_\Omega (\alpha f + \beta g)\, dp^\omega = \alpha \int_\Omega f dp^\omega + \beta \int_\Omega g dp^\omega = \alpha f^\star(\omega) + \beta g^\star(\omega) \qquad \forall \omega \in \Omega,$$

that is, $(\alpha f + \beta g)^\star = \alpha f^\star + \beta g^\star$. Second, if $\{f_n\}_{n\in\mathbb{N}} \subseteq B(\Omega, \mathcal{F})$ and $f \in B(\Omega, \mathcal{F})$ are such that $f_n \to f$ (where the convergence is uniform) then

$$|f^\star(\omega) - f_n^\star(\omega)| = \left| \int_\Omega f dp^\omega - \int_\Omega f_n dp^\omega \right| \leq \|f_n - f\| \qquad \forall \omega \in \Omega.$$

It follows that $f_n^\star \to f^\star$, proving that $^\star : B(\Omega, \mathcal{F}) \to B(\Omega)$ is Lipschitz continuous. □

Fix $P \in \mathcal{P}$. We next show that $f^\star \in E_P[f|\mathcal{G}]$ for all $f \in B(\Omega, \mathcal{F})$. Let $V = \{f \in B(\Omega, \mathcal{F}) : f^\star \in E_P[f|\mathcal{G}]\}$. Since the mapping $^\star : B(\Omega, \mathcal{F}) \to B(\Omega)$ is linear, if $f, g \in V$ and $\alpha, \beta \in \mathbb{R}$ then

$$(\alpha f + \beta g)^\star = \alpha f^\star + \beta g^\star \in \alpha E_P[f|\mathcal{G}] + \beta E_P[g|\mathcal{G}] \subseteq E_P[\alpha f + \beta g|\mathcal{G}].$$

This implies that V is a vector subspace of $B(\Omega, \mathcal{F})$. Similarly, since $^\star : B(\Omega, \mathcal{F}) \to B(\Omega)$ is continuous, if $\{f_n\}_{n\in\mathbb{N}} \subseteq V$ and $f_n \to f$ then $\{f_n^\star\}_{n\in\mathbb{N}} \subseteq B(\Omega, \mathcal{G})$ converges to $f^\star$. It follows that $f^\star \in B(\Omega, \mathcal{G})$ and that

$$\int_G f^\star dP = \lim_n \int_G f_n^\star dP = \lim_n \int_G f_n dP = \int_G f dP \qquad \forall G \in \mathcal{G}.$$

This implies that $f^\star \in E_P[f|\mathcal{G}]$ and that V is closed. Finally, observe that if $A \in \mathcal{F}$ and $f = 1_A$ then for each $\omega \in \Omega$

$$f^\star(\omega) = (1_A)^\star(\omega) = \int_\Omega 1_A dp^\omega = p(A, \omega),$$

that is, $f^\star = p(A, \cdot) \in P[A|\mathcal{G}] = E_P[1_A|\mathcal{G}]$. We can conclude that V contains all the indicator functions of elements of $\mathcal{F}$. Since V is a vector subspace of $B(\Omega, \mathcal{F})$, this implies that $B_0(\Omega, \mathcal{F}) \subseteq V \subseteq B(\Omega, \mathcal{F})$. Since V is closed, it follows that $V = B(\Omega, \mathcal{F})$.

Since P in $\mathcal{P}$ was arbitrarily chosen, it follows that $f^\star \in \bigcap_{P\in\mathcal{P}} E_P[f|\mathcal{G}]$ for all $f \in B(\Omega, \mathcal{F})$. Moreover, notice that $^\star : B(\Omega, \mathcal{F}) \to B(\Omega)$ is a linear and (Lipschitz) continuous operator and we can restrict its target space to $B(\Omega, \mathcal{G})$.

Finally, we prove $\mathcal{P}$-a.s. uniqueness of the common regular conditional probability for $\mathcal{P}$ given $\mathcal{G}$. Since $(\Omega, \mathcal{F})$ is a standard Borel space there exists an at most countable algebra, $\mathcal{A} = \{A_n\}_{n\in\mathbb{N}}$,[31] generating $\mathcal{F}$.

[31] In particular, $\mathcal{A}$ is a π-class.

Let p and p' be two common regular conditional probabilities for $\mathcal{P}$ given $\mathcal{G}$. Fix $P \in \mathcal{P}$. By assumption, for each $n \in \mathbb{N}$ we have that $p(A_n, \cdot)$ and $p'(A_n, \cdot)$ are elements of $P[A_n|\mathcal{G}]$. It follows that

$$U_n = \{\omega \in \Omega : p(A_n, \omega) = p'(A_n, \omega)\} \in \mathcal{G} \text{ and } P(U_n) = 1.$$

This implies that $U = \bigcap_{n\in\mathbb{N}} U_n \in \mathcal{G}$ and $P(U) = 1$. We can conclude that

$$\begin{aligned} U &= \{\omega \in \Omega : p(A_n, \omega) = p'(A_n, \omega), \quad \forall n \in \mathbb{N}\} \\ &= \{\omega \in \Omega : p(A, \omega) = p'(A, \omega), \quad \forall A \in \mathcal{A}\} \\ &= \left\{\omega \in \Omega : p(\cdot, \omega)_{|\mathcal{A}} = p'(\cdot, \omega)_{|\mathcal{A}}\right\} \\ &= \{\omega \in \Omega : p(\cdot, \omega) = p'(\cdot, \omega)\} \end{aligned}$$

where the last equality follows by Dynkin lemma (see, e.g., [8, Theorem 3.3]) given that $\mathcal{F} = \sigma(\mathcal{A})$. The proof is complete since P in $\mathcal{P}$ was arbitrarily chosen. ■

Remark 13 *Notice that:*

- *The implications (i) implies (ii) and (iii) implies (i) hold for* any *measurable space. The standardness assumption is used* only *in proving (ii) implies (iii). On the other hand, the uniqueness part of the statement just requires that $\mathcal{F}$ is countably generated.*
- *Under (iii), the operator $^\star : B(\Omega, \mathcal{F}) \to B(\Omega, \mathcal{G})$, defined for each $f \in B(\Omega, \mathcal{F})$ by*

$$f^\star(\omega) = \int_\Omega f dp^\omega \qquad \forall \omega \in \Omega,$$

  is well defined, linear, continuous, and such that:

  - $(1_A)^\star = p(A, \cdot) \in \bigcap_{P\in\mathcal{P}} P[A|\mathcal{G}]$ *for all* $A \in \mathcal{F}$;
  - $f^\star \in \bigcap_{P\in\mathcal{P}} E_P[f|\mathcal{G}]$ *for all* $f \in B(\Omega, \mathcal{F})$.

- *For a singleton set $\mathcal{P} = \{P\}$, point (i) of Theorem 12 is trivially satisfied. Thus, Theorem 12 yields Theorem 11 as a corollary. Moreover, the proof of (iii) implies (i) shows that if p is a regular conditional probability for P given $\mathcal{G}$ then for each $f \in B(\Omega, \mathcal{F})$ the function $f^\star$, defined by*

$$f^\star(\omega) = \int_\Omega f dp^\omega \qquad \forall \omega \in \Omega,$$

  is a version of the conditional expected value of f given $\mathcal{G}$. Theorem 12 also shows P-a.s. uniqueness of the regular conditional probability for P given $\mathcal{G}$.

B Dynkin Spaces

In this section, we maintain the notation and terminology introduced in the previous one. We denote by $ba(\Omega, \mathcal{F})$ the space of bounded and finitely additive set functions from $\mathcal{F}$ to $\mathbb{R}$. We consider $ba(\Omega, \mathcal{F})$ endowed with the weak* sigma algebra,

$$\mathcal{A}_{ba(\Omega,\mathcal{F})} = \sigma(\langle f, \cdot\rangle : f \in B(\Omega, \mathcal{F})),$$

and each of its subsets with the induced sigma algebra. It is well known that $ba(\Omega, \mathcal{F})$ is isometrically isomorphic to the norm dual, $B(\Omega, \mathcal{F})^*$, of $B(\Omega, \mathcal{F})$. We endow $ba(\Omega, \mathcal{F})$ with the weak* topology and each of its subsets with the relative (weak*) topology.

Fact 1 *Let $(\Omega, \mathcal{F})$ be a measurable space.* $\mathcal{A}_{ba(\Omega,\mathcal{F})} = \sigma(\langle 1_A, \cdot\rangle : A \in \mathcal{F})$ *and for each* $\mathcal{C} \subseteq ba(\Omega, \mathcal{F})$

$$\mathcal{A}_{\mathcal{C}} = \mathcal{A}_{ba(\Omega,\mathcal{F})} \cap \mathcal{C} = \sigma\left(\langle f, \cdot\rangle_{|\mathcal{C}} : f \in B(\Omega, \mathcal{F})\right) = \sigma\left(\langle 1_A, \cdot\rangle_{|\mathcal{C}} : A \in \mathcal{F}\right).$$

If $\mathcal{C}$ is a norm bounded subset of $ba(\Omega, \mathcal{F})$ and μ is a finitely additive probability on $\mathcal{C}$, that is on $\mathcal{A}_{\mathcal{C}}$, then the *barycenter* $\bar{\mu}$ *of* μ is the functional defined for each $f \in B(\Omega, \mathcal{F})$ by

$$\bar{\mu}(f) = \int_{\mathcal{C}} \langle f, m\rangle \, d\mu(m).$$

Notice that $\bar{\mu}$ is well defined. Indeed, since $\mathcal{C}$ is norm bounded, the mapping $m \mapsto \langle f, m\rangle$ is bounded and $\mathcal{A}_{\mathcal{C}}$ measurable. Moreover, let us define $\tilde{\mu} : \mathcal{F} \to \mathbb{R}$ to be such that $\tilde{\mu}(A) = \bar{\mu}(1_A)$ for all $A \in \mathcal{F}$. Then, we have the following

Fact 2 *Let $(\Omega, \mathcal{F})$ be a measurable space, $\mathcal{C}$ a norm bounded subset of $ba(\Omega, \mathcal{F})$, and $\mu : \mathcal{A}_{\mathcal{C}} \to [0, 1]$ a finitely additive probability. Then, $\tilde{\mu} \in ba(\Omega, \mathcal{F})$ and $\bar{\mu} \in B(\Omega, \mathcal{F})^*$. Moreover, we have that*

$$\int_{\mathcal{C}} \langle f, m\rangle \, d\mu(m) = \bar{\mu}(f) = \int_{\Omega} f d\tilde{\mu} \qquad \forall f \in B(\Omega, \mathcal{F}).$$

For this reason, we will just, equivalently, say that $\tilde{\mu}$ is the barycenter of μ and we will denote it by $\bar{\mu}$. By the Hahn-Banach Theorem, $\bar{\mu}$ always belongs to the closed convex hull of $\mathcal{C}$ in $ba(\Omega, \mathcal{F})$.[32] Conversely, the closed convex hull of $\mathcal{C}$ consists exactly of all barycenters of all finitely additive probabilities on $\mathcal{C}$. Indeed, if $m \in \overline{\text{co}}^*(\mathcal{C})$ then there exists a net $\{m_\alpha\}_{\alpha \in A}$ of convex combinations of elements of $\mathcal{C}$ that converges to m. That is, there exists a net $\{\mu_\alpha\}_{\alpha \in A}$ of simple probability measures on $\mathcal{C}$ such that $\{\bar{\mu}_\alpha\}_{\alpha \in A}$ converges to m and $\bar{\mu}_\alpha = m_\alpha$ for all $\alpha \in A$. Since the set of finitely additive probabilities on $\mathcal{C}$ is a compact set, it follows that there exists a subnet $\{\mu_{\alpha_\beta}\}_{\beta \in B}$ of $\{\mu_\alpha\}_{\alpha \in A}$ which converges to μ. It is immediate to check that $\bar{\mu} = m$.

As a consequence of the Monotone Convergence Theorem, we have the following fact:

Fact 3 *If $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$ and $\mu \in \Delta^\sigma(\mathcal{P}, \mathcal{A}_{\mathcal{P}})$ then $\bar{\mu} \in \Delta^\sigma(\Omega, \mathcal{F})$.*

Definition 14 *A subset $\mathcal{P}$ of $\Delta^\sigma(\Omega, \mathcal{F})$ is* measure convex *if and only if $\bar{\mu} \in \mathcal{P}$ for all $\mu \in \Delta^\sigma(\mathcal{P}, \mathcal{A}_{\mathcal{P}})$.*

Measure convex subsets of $\Delta^\sigma(\Omega, \mathcal{F})$ are convex.[33] Conversely, by the Hahn-Banach Theorem, compact convex subsets of $\Delta^\sigma(\Omega, \mathcal{F})$ are measure convex. On the other hand, $\Delta^\sigma(\Omega, \mathcal{F})$ is measure convex without being necessarily compact. Finally, the intersection of measure convex sets is measure convex. These last two facts allow to naturally define the *measure convex hull* $\text{mco}(\mathcal{P})$ of any $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$, that is, $\text{mco}(\mathcal{P})$ is the intersection of all measure convex subsets of $\Delta^\sigma(\Omega, \mathcal{F})$ that further contain $\mathcal{P}$. From the previous discussion, it follows that

$$\text{co}(\mathcal{P}) \subseteq \text{mco}(\mathcal{P}) \subseteq \overline{\text{co}}^*(\mathcal{P}) \cap \Delta^\sigma(\Omega, \mathcal{F}).$$

Definition 15 *Let $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$. An element P of $\mathcal{P}$ is:*

[32] If $\bar{\mu} \notin \overline{\text{co}}^*(\mathcal{C})$ then there would exist $f \in B(\Omega, \mathcal{F})$, $\alpha \in \mathbb{R}$, and $\varepsilon > 0$ such that $\langle f, \bar{\mu}\rangle \leq \alpha < \alpha + \varepsilon \leq \langle f, m'\rangle$ for all $m' \in \overline{\text{co}}^*(\mathcal{C})$. This would imply that $\langle f, \bar{\mu}\rangle = \int_{\mathcal{C}} \langle f, m\rangle \, d\mu(m) \geq \alpha + \varepsilon > \langle f, \bar{\mu}\rangle$, a contradiction.

[33] In fact, for each $P_1, P_2 \in \mathcal{P}$ and each $\alpha \in [0, 1]$,

$$\begin{aligned}\overline{(\alpha\delta_{P_1} + (1-\alpha)\delta_{P_2})}(A) &= \int_{\mathcal{P}} \langle 1_A, \cdot\rangle \, d(\alpha\delta_{P_1} + (1-\alpha)\delta_{P_2}) \\ &= \alpha \int_{\mathcal{P}} \langle 1_A, \cdot\rangle \, d\delta_{P_1} + (1-\alpha)\int_{\mathcal{P}} \langle 1_A, \cdot\rangle \, d\delta_{P_2} \\ &= \alpha P_1(A) + (1-\alpha) P_2(A) \qquad \forall A \in \mathcal{F}.\end{aligned}$$

1. *an* extreme point *of* $\mathcal{P}$ *if and only if for each* $P', P'' \in \mathcal{P}$ *and* $\alpha \in (0,1)$ *such that* $P = \alpha P' + (1-\alpha) P''$ *we have that* $P' = P = P''$*.* $\mathcal{E} = \mathcal{E}(\mathcal{P})$ *denotes the set of all extreme points of* $\mathcal{P}$*;*

2. *a* strong extreme point *of* $\mathcal{P}$ *if and only if the only probability measure* μ *on* $\mathcal{P}$ *such that* P *is the barycenter of* μ *is* δ_P*.* $\mathcal{S} = \mathcal{S}(\mathcal{P})$ *denotes the set of all strong extreme points of* $\mathcal{P}$*.*

Clearly, strong extreme points are extreme points. The two notions coincide when $\mathcal{P}$ is measure convex and $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space. We here recall the definition of Dynkin space.

Definition 16 (Dynkin, 1978) *Let* $\mathcal{P}$ *be a nonempty subset of* $\Delta^\sigma(\Omega, \mathcal{F})$ *where* $(\Omega, \mathcal{F})$ *is a separable measurable space.* $(\Omega, \mathcal{F}, P)$ *is a Dynkin space if and only if there exists a sub-*σ*-algebra* $\mathcal{G} \subseteq \mathcal{F}$*, a set* $W \in \mathcal{F}$*, and a function*

$$\begin{array}{rccc} p: & \mathcal{F} \times \Omega & \to & [0,1] \\ & (A, \omega) & \mapsto & p(A, \omega) \end{array}$$

such that:

(a) for each $P \in \mathcal{P}$ *and* $A \in \mathcal{F}$*,* $p(A, \cdot) : \Omega \to [0,1]$ *is a version of the conditional probability of* A *given* $\mathcal{G}$*;*

(b) for each $\omega \in \Omega$*,* $p(\cdot, \omega) : \mathcal{F} \to [0,1]$ *is a probability measure;*

(c) $P(W) = 1$ *for all* $P \in \mathcal{P}$ *and* $p(\cdot, \omega) \in \mathcal{P}$ *for all* $\omega \in W$*.*

Given $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$ and a sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, we denote

$$\Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P}) = \{Q \in \Delta^\sigma(\Omega, \mathcal{G}) : Q(B) = 0 \text{ if } P(B) = 0 \text{ for all } P \in \mathcal{P}\}.$$

Theorem 17 ([19, Theorem 3.1]) *Let* $(\Omega, \mathcal{F}, \mathcal{P})$ *be a Dynkin space. Then:*

1. $\mathcal{S}(\mathcal{P}) = \{P \in \mathcal{P} : P(\mathcal{G}) = \{0,1\}\}$*.*

2. $\mathcal{S}(\mathcal{P})$ *is measurable, that is,* $\mathcal{S}(\mathcal{P}) \in \mathcal{A}_{\mathcal{P}}$*.*

3. *For each* $P \in \mathcal{P}$ *there exists a unique measure* μ_P *concentrated on* $\mathcal{S}(\mathcal{P})$ *such that* $P = \bar{\mu}_P$*.*

4. *If* $\mathcal{P}$ *is measure convex then the restriction map* $P \mapsto P_{\mathcal{G}}$ *from* $\mathcal{P}$ *to* $\Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$ *is affine, bijective, and* $\mathcal{S}(\mathcal{P}) = \mathcal{E}(\mathcal{P})$*.*

If p *is a common regular conditional probability for* $\mathcal{P}$ *given* $\mathcal{G}$ *then*

$$\mu_P(\Gamma) = P(\{\omega \in \Omega : p^\omega \in \Gamma\}), \tag{22}$$

for all $\Gamma \in \mathcal{A}_{\mathcal{P}}$*. Moreover,*

$$\mathcal{S}(\mathcal{P}) = \{P \in \mathcal{P} : P(\{\omega \in \Omega : p^\omega = P\}) = 1\}, \tag{23}$$

particularly, $\mathcal{S}(\mathcal{P})$ *is a subset of* $\{p^\omega\}_{\omega \in \Omega}$*.*

We next present a characterization of the measure convex hull of a set $\mathcal{P}$. Before proving it, we present some ancillary notions and facts.

Proposition 18 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space. Then:*

$$\begin{aligned}
\mathrm{mco}(\mathcal{P}) &= \left\{ \int_\Omega p(\cdot,\omega)\, dQ(\omega) : Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P}) \right\} \\
&= \{ R \in \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P}) : p \text{ is a r.c.p. given } \mathcal{G} \text{ for } R \} \\
&= \{ \bar{\mu}_{Q^W} : Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P}) \} \\
&= \{ \bar{\mu} : \mu \in \Delta^\sigma(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}) \} \\
&= \{ \bar{\mu} : \mu \in \Delta^\sigma(\mathcal{P}, \mathcal{A}_{\mathcal{P}}) \}.
\end{aligned}$$

Moreover,

1. *The restriction map $R \mapsto R_\mathcal{G}$ from $\mathrm{mco}(\mathcal{P})$ to $\Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$ is an affine bijection. The inverse image of $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$ is $\bar{\mu}_{Q^W}$.*

2. *The map $Q \mapsto \mu_{Q^W}$ from $\Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$ to $\Delta^\sigma(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})})$ is an affine bijection. The inverse image of λ is $\bar{\lambda}_\mathcal{G}$.*

3. *$\mathrm{mco}(\mathcal{P})$ is closed in $\Delta^\sigma(\Omega, \mathcal{F})$ with respect to the relative weak* topology.*

Before proving this result, we make few observations. Consider a Dynkin space $(\Omega, \mathcal{F}, \mathcal{P})$. It is not hard to check that the mapping $\omega \mapsto p^\omega$ is $\mathcal{G} - \mathcal{A}_{\Delta^\sigma(\Omega,\mathcal{F})}$ measurable. Since $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space, let $W \in \mathcal{F}$ be such that $P(W) = 1$ for all $P \in \mathcal{P}$ and $p^\omega \in \mathcal{P}$ for all $\omega \in W$. Then, the map

$$p^\cdot_{|W} : W \to \mathcal{P}$$

is $\mathcal{G} \cap W$-$\mathcal{A}_\mathcal{P}$ measurable. In fact, if $\Gamma \in \mathcal{A}_\mathcal{P}$ then there is $\Pi \in \mathcal{A}_{\Delta^\sigma(\Omega,\mathcal{F})}$ such that $\Gamma = \Pi \cap \mathcal{P}$. This implies that

$$\begin{aligned}
\{\omega \in \Omega : p^\omega \in \Pi\} \in \mathcal{G} &\Rightarrow \{\omega \in W : p^\omega \in \Pi \text{ and } p^\omega \in \mathcal{P}\} = \{\omega \in W : p^\omega \in \Pi\} = \{\omega \in \Omega : p^\omega \in \Pi\} \cap W \in \mathcal{G} \cap W \\
&\Rightarrow \left\{\omega \in W : p^\omega_{|W} \in \Gamma\right\} = \{\omega \in W : p^\omega \in \Pi \cap \mathcal{P}\} \in \mathcal{G} \cap W.
\end{aligned}$$

We denote

$$\Delta^\sigma(W, \mathcal{G} \cap W, \mathcal{P}) = \{Q' \in \Delta^\sigma(W, \mathcal{G} \cap W) : Q'(B \cap W) = 0 \text{ if } B \in \mathcal{G} \text{ and } P(B) = 0 \text{ for all } P \in \mathcal{P}\}.$$

For all $Q \in \Delta^\sigma(W, \mathcal{G} \cap W, \mathcal{P})$ set $\mu_Q = Q \circ \left(p^\cdot_{|W}\right)^{-1}$, that is,

$$\mu_Q(\Gamma) = Q(\{\omega \in W : p^\omega \in \Gamma\}) \qquad \forall \Gamma \in \mathcal{A}_\mathcal{P}. \tag{24}$$

Notice that if $P \in \mathcal{P}$ then $P_{\mathcal{G} \cap W} \in \Delta^\sigma(W, \mathcal{G} \cap W, \mathcal{P})$. In this case, we might just write μ_P rather than $\mu_{P_{\mathcal{G}\cap W}}$ and, by (24), we have that

$$\mu_P(\Gamma) = P(\{\omega \in W : p^\omega \in \Gamma\}) \qquad \forall \Gamma \in \mathcal{A}_\mathcal{P}.$$

Notice that $\{\omega \in \Omega : p^\omega \in \Gamma\} = \{\omega \in W : p^\omega \in \Gamma\} \cup \{\omega \in W^c : p^\omega \in \Gamma\}$ might not belong to $\mathcal{F}$. Nevertheless, since $\{\omega \in W : p^\omega \in \Gamma\} \in \mathcal{G} \cap W \subseteq \mathcal{F}$ and $P(W^c) = 0$ for all $P \in \mathcal{P}$, $\{\omega \in \Omega : p^\omega \in \Gamma\}$ must belong to the P completion of $\mathcal{F}$ for all $P \in \mathcal{P}$. Thus, we can write

$$\mu_P(\Gamma) = P(\{\omega \in W : p^\omega \in \Gamma\}) = P(\{\omega \in \Omega : p^\omega \in \Gamma\}) \qquad \forall P \in \mathcal{P}, \forall \Gamma \in \mathcal{A}_\mathcal{P}.$$

In other words, (22) holds modulo completion. Clearly, for each $Q \in \Delta^\sigma(W, \mathcal{G} \cap W, \mathcal{P})$ we have that $\mu_Q \in \Delta^\sigma(\mathcal{P}, \mathcal{A}_\mathcal{P})$. On the other hand for each $f \in B(\Omega, \mathcal{F})$, since $\langle f, \cdot \rangle \in B(\mathcal{P}, \mathcal{A}_\mathcal{P})$ and by the change of variable formula (see, e.g., [8, Theorem 16.13]), we have that

$$\bar{\mu}_Q(f) = \int_\mathcal{P} \langle f, \cdot \rangle \, d\mu_Q = \int_W (\langle f, \cdot \rangle \circ p^\cdot) \, dQ = \int_W \langle f, p^\omega \rangle \, dQ(\omega) = \int_W f^\star(\omega) \, dQ(\omega) = \int_\Omega f^\star(\omega) \, dQ_W(\omega), \tag{25}$$

where $Q_W(B) = Q(B \cap W)$ for all $B \in \mathcal{G}$.

Fact 4 *The map* $_W : \Delta^\sigma(W, \mathcal{G} \cap W, \mathcal{P}) \to \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$ *such that* $Q \mapsto Q_W$ *is an affine bijection. We denote its inverse by* $^W : \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P}) \to \Delta^\sigma(W, \mathcal{G} \cap W, \mathcal{P})$, *that is,* $Q \mapsto Q^W$.

Since $f^\star \in \bigcap_{P \in \mathcal{P}} E_P[f|\mathcal{G}]$ and by (25), if $P \in \mathcal{P}$ then

$$\bar{\mu}_P(f) = \int_W f^\star(\omega) \, dP_{\mathcal{G} \cap W}(\omega) = \int_\Omega f^\star(\omega) \, dP_\mathcal{G}(\omega) = \int_\Omega f(\omega) \, dP(\omega) \qquad \forall f \in B(\Omega, \mathcal{F}). \tag{26}$$

It follows that $\bar{\mu}_P = P$. Moreover, if $f \in B(\Omega, \mathcal{G})$ then

$$P(\{\omega \in \Omega : f(\omega) \neq f^\star(\omega)\}) = 0 \qquad \forall P \in \mathcal{P}.$$

This implies that $f = f^\star$ Q-a.s. for all $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$. Thus, we can conclude that if $f \in B(\Omega, \mathcal{G})$ then for each $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$

$$\int_\Omega f dQ = \int_\Omega f^\star dQ = \int_W f^\star dQ^W = \bar{\mu}_{Q^W}(f). \tag{27}$$

Proposition 19 *Let* $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$. *The following statements are true:*

1. $\bar{\mu}_{Q^W} \in \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$. *Moreover,* $\bar{\mu}_{Q^W} \in \mathcal{P}$ *if* $\mathcal{P}$ *is measure convex;*
2. $\bar{\mu}_{Q^W}(A) = \int_\Omega p(A, \omega) \, dQ(\omega)$ *for all* $A \in \mathcal{F}$;
3. $(\bar{\mu}_{Q^W})_\mathcal{G} = Q$ *and* $(\bar{\mu}_{Q^W})_{\mathcal{G} \cap W} = Q^W$;
4. p *is a regular conditional probability for* $\bar{\mu}_{Q^W}$;
5. $\mu_{Q^W}(\mathcal{S}(\mathcal{P})) = 1$;
6. $\mu_{(P_\mathcal{G})^W} = \mu_{P_{\mathcal{G} \cap W}} =: \mu_P$ *for all* $P \in \mathcal{P}$.

Proof. 1. Since $\mu_{Q^W} \in \Delta^\sigma(\mathcal{P}, \mathcal{A}_\mathcal{P})$ and by Fact 3, $\bar{\mu}_{Q^W} \in \Delta^\sigma(\Omega, \mathcal{F})$ and $\bar{\mu}_{Q^W} \in \mathcal{P}$ if $\mathcal{P}$ is measure convex. Finally, observe that for each $A \in \mathcal{F}$ such that $P(A) = 0$ for all $P \in \mathcal{P}$,

$$\bar{\mu}_{Q^W}(A) = \int_\mathcal{P} P(A) \, d\mu_{Q^W}(P) = \int_\mathcal{P} 0 d\mu_{Q^W}(P) = 0,$$

proving that $\bar{\mu}_{Q^W} \in \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$.

2. For each $A \in \mathcal{F}$, by (25) and Remark 13,

$$\bar{\mu}_{Q^W}(A) = \int_\mathcal{P} \langle 1_A, \cdot \rangle \, d\mu_{Q^W} = \int_\Omega (1_A)^\star(\omega) \, dQ^W_W(\omega) = \int_\Omega (1_A)^\star(\omega) \, dQ(\omega) = \int_\Omega p(A, \omega) \, dQ(\omega).$$

3. For each $B \in \mathcal{G}$, by (27),

$$Q(B) = \int_\Omega 1_B dQ = \bar{\mu}_{Q^W}(B),$$

proving the first part of the statement. Moreover, by point 1. and the previous part of the proof,

$$\left((\bar{\mu}_{Q^W})_{\mathcal{G} \cap W}\right)_W (B) = \bar{\mu}_{Q^W}(B \cap W) = \bar{\mu}_{Q^W}(B) = Q(B) \qquad \forall B \in \mathcal{G}.$$

By Fact 4,we have that $\left(\bar{\mu}_{Q^W}\right)_{\mathcal{G}\cap W} = Q^W$ if and only if $\left(\left(\bar{\mu}_{Q^W}\right)_{\mathcal{G}\cap W}\right)_W = Q$. Then, the statement follows.

4. Since $p(A,\cdot)\,1_B(\cdot) \in B(\Omega,\mathcal{G})$ for all $A \in \mathcal{F}$ and $B \in \mathcal{G}$, we have that

$$\begin{aligned}\bar{\mu}_{Q^W}(A\cap B) &= \int_{\mathcal{P}} P(A\cap B)\, d\mu_{Q^W}(P) = \int_{\mathcal{P}} \left(\int_B p(A,\omega)\, dP(\omega)\right) d\mu_{Q^W}(P) \\ &= \int_{\mathcal{P}} \langle p(A,\cdot)\,1_B(\cdot), P\rangle\, d\mu_{Q^W}(P) = \bar{\mu}_{Q^W}(p(A,\cdot)\,1_B(\cdot)).\end{aligned}$$

Thus, by (27) and point 3., it follows that

$$\begin{aligned}\bar{\mu}_{Q^W}(A\cap B) = \bar{\mu}_{Q^W}(p(A,\cdot)\,1_B(\cdot)) &= \int_{\Omega} p(A,\omega)\,1_B(\omega)\, dQ(\omega) \\ &= \int_B p(A,\cdot)\, d\left(\bar{\mu}_{Q^W}\right)_{\mathcal{G}},\end{aligned}$$

proving that $p(A,\cdot) \in \bar{\mu}_{Q^W}[A|\mathcal{G}]$.

5. Set $\hat{Q} = \bar{\mu}_{Q^W}$. Since $(\Omega,\mathcal{F},\mathcal{P})$ is a Dynkin space, we have that there exists a countable family $\{H_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ generating $\mathcal{F}$. Without loss of generality, we can assume that $\{H_n\}_{n\in\mathbb{N}}$ is a π-class. Fix an arbitrary $n \in \mathbb{N}$. For each $\bar{\omega} \in W$, we have that $p^{\bar{\omega}} \in \mathcal{P}$ and

$$\begin{aligned}\int_{\Omega} \left(p(H_n,\omega) - p^{\bar{\omega}}(H_n)\right)^2 dp^{\bar{\omega}}(\omega) &= \int_{\Omega} \left(p(H_n,\omega) - \int_{\Omega} p(H_n,\omega)\, dp^{\bar{\omega}}(\omega)\right)^2 dp^{\bar{\omega}}(\omega) \\ &= \int_{\Omega} p(H_n,\omega)^2\, dp^{\bar{\omega}}(\omega) - \left(\int_{\Omega} p(H_n,\omega)\, dp^{\bar{\omega}}(\omega)\right)^2 \\ &= \int_{\Omega} p(H_n,\omega)^2\, dp^{\bar{\omega}}(\omega) - p^{\bar{\omega}}(H_n)^2 \\ &= \left(p(H_n,\cdot)^2\right)^{\star}(\bar{\omega}) - p^{\bar{\omega}}(H_n)^2 = \left(p(H_n,\cdot)^2\right)^{\star}(\bar{\omega}) - p(H_n,\bar{\omega})^2.\end{aligned}$$

By point 4. and since $p(H_n,\cdot)$ is $\mathcal{G}$-measurable, $\left(p(H_n,\cdot)^2\right)^{\star}, p(H_n,\cdot)^2 \in E_{\hat{Q}}\left[p(H_n,\cdot)^2|\mathcal{G}\right]$. It follows that

$$\int_{\Omega} \left(p(H_n,\omega) - p^{\bar{\omega}}(H_n)\right)^2 dp^{\bar{\omega}}(\omega) = 0 \qquad \forall \bar{\omega} \in W.$$

By point 1., $\hat{Q}(W) = 1$. By Dynkin Lemma and since n was arbitrarily chosen and $\{H_n\}_{n\in\mathbb{N}}$ is a countable π-class, this implies that

$$\begin{aligned}1 &= \hat{Q}\left(\left\{\bar{\omega} \in W : p^{\bar{\omega}}\left(\left\{\omega \in \Omega : p(H_n,\omega) = p^{\bar{\omega}}(H_n), \quad \forall n \in \mathbb{N}\right\}\right) = 1\right\}\right) \\ &= \hat{Q}\left(\left\{\bar{\omega} \in W : p^{\bar{\omega}}\left(\left\{\omega \in \Omega : p(\cdot,\omega) = p^{\bar{\omega}}\right\}\right) = 1\right\}\right)\end{aligned}$$

By Theorem 17, we can conclude that

$$1 = \hat{Q}\left(\left\{\bar{\omega} \in W : p^{\bar{\omega}}\left(\left\{\omega \in \Omega : p^{\omega} = p^{\bar{\omega}}\right\}\right) = 1\right\}\right) = \hat{Q}\left(\left\{\bar{\omega} \in W : p^{\bar{\omega}} \in \mathcal{S}(\mathcal{P})\right\}\right).$$

Next, since $\mathcal{S}(\mathcal{P}) \in \mathcal{A}_{\mathcal{P}}$ and $p^{\cdot}_{|W}$ is $\mathcal{G}\cap W$-$\mathcal{A}_{\mathcal{P}}$ measurable, observe that $\{\bar{\omega} \in W : p^{\bar{\omega}} \in \mathcal{S}(\mathcal{P})\} \in \mathcal{G}\cap W$. By point 4., $\hat{Q}_{\mathcal{G}\cap W} = \left(\bar{\mu}_{Q^W}\right)_{\mathcal{G}\cap W} = Q^W$, this implies that

$$\mu_{Q^W}(\mathcal{S}(\mathcal{P})) = Q^W\left(\left\{\bar{\omega} \in W : p^{\bar{\omega}} \in \mathcal{S}(\mathcal{P})\right\}\right) = \hat{Q}\left(\left\{\bar{\omega} \in W : p^{\bar{\omega}} \in \mathcal{S}(\mathcal{P})\right\}\right) = 1,$$

proving the statement.

6. Notice that if $P \in \mathcal{P}$ then for each $B \in \mathcal{G}$ we have that

$$P(B) = P(B\cap W) = \left(P_{\mathcal{G}\cap W}\right)_W(B),$$

that is, $\left(P_{\mathcal{G}\cap W}\right)_W = P_{\mathcal{G}}$, $\left(P_{\mathcal{G}}\right)^W = P_{\mathcal{G}\cap W}$. By definition of μ_P, this implies that

$$\mu_{(P_{\mathcal{G}})^W} = \mu_{P_{\mathcal{G}\cap W}} =: \mu_P,$$

proving the statement. ■

Proof of Proposition 18. Point 2. of Proposition 19 shows that

$$\left\{\int_{\Omega} p(\cdot,\omega)\, dQ(\omega) : Q \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})\right\} = \left\{\bar{\mu}_{Q^W} : Q \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})\right\}. \tag{28}$$

Set $\mathcal{C} = \{R \in \Delta^{\sigma}(\Omega,\mathcal{F},\mathcal{P}) : p \text{ is a r.c.p. for } R \text{ given } \mathcal{G}\}$.

Step 1: $\mathcal{C}$ is measure convex and $\mathcal{P} \subseteq \mathcal{C}$.

Proof of the Step.

Since $(\Omega,\mathcal{F},\mathcal{P})$ is a Dynkin space, it is immediate to see that $\mathcal{P} \subseteq \mathcal{C}$. Let $\mu \in \Delta^{\sigma}(\mathcal{C},\mathcal{A}_{\mathcal{C}})$. Since $\mathcal{C} \subseteq \Delta^{\sigma}(\Omega,\mathcal{F},\mathcal{P})$, if $A \in \mathcal{F}$ and $P(A) = 0$ for all $P \in \mathcal{P}$ then

$$\bar{\mu}(A) = \int_{\mathcal{C}} R(A)\, d\mu(R) = \int_{\mathcal{C}} 0 d\mu(R) = 0$$

and $\bar{\mu} \in \Delta^{\sigma}(\Omega,\mathcal{F},\mathcal{P})$. Moreover, for each $A \in \mathcal{F}$ and $B \in \mathcal{G}$,

$$\begin{aligned}
\bar{\mu}(A\cap B) &= \int_{\mathcal{C}} R(A\cap B)\, d\mu(R) = \int_{\mathcal{C}} \left(\int_B p(A,\omega)\, dR(\omega)\right) d\mu(R) \\
&= \int_{\mathcal{C}} \langle p(A,\cdot)\, 1_B(\cdot), R\rangle\, d\mu(R) = \bar{\mu}\left(p(A,\cdot)\, 1_B(\cdot)\right) \\
&= \int_{\Omega} p(A,\cdot)\, 1_B(\cdot)\, d\bar{\mu} = \int_B p(A,\cdot)\, d\bar{\mu}_{\mathcal{G}},
\end{aligned}$$

that is, $p(A,\cdot) \in \bar{\mu}[A|\mathcal{G}]$. Summing up, $\bar{\mu} \in \Delta^{\sigma}(\Omega,\mathcal{F},\mathcal{P})$ and p is a r.c.p. for $\bar{\mu}$ given $\mathcal{G}$, for all $\mu \in \Delta^{\sigma}(\mathcal{C},\mathcal{A}_{\mathcal{C}})$, proving that $\bar{\mu} \in \mathcal{C}$ and that $\mathcal{C}$ is a measure convex set. □

Step 2: $\mathcal{C} = \left\{\int_{\Omega} p(\cdot,\omega)\, dQ(\omega) : Q \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})\right\}$.

Proof of the Step.

Let $R \in \mathcal{C}$ and let p be the r.c.p. for R given $\mathcal{G}$ (for R and $\mathcal{P}$). By assumption, we have that $R \in \Delta^{\sigma}(\Omega,\mathcal{F},\mathcal{P})$ and that for each $A \in \mathcal{F}$

$$R(A) = R(A\cap\Omega) = \int_{\Omega} p(A,\omega)\, dR_{\mathcal{G}}(\omega) \tag{29}$$

where, clearly, $R_{\mathcal{G}} \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})$. This implies that $R \in \left\{\int_{\Omega} p(\cdot,\omega)\, dQ(\omega) : Q \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})\right\}$.

Viceversa, let $R = \int_{\Omega} p(\cdot,\omega)\, d\tilde{Q}(\omega)$ for some $\tilde{Q} \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})$. By point 2. of Proposition 19, we have that

$$R(A) = \int_{\Omega} p(A,\omega)\, d\tilde{Q}(\omega) = \bar{\mu}_{\tilde{Q}^W}(A) \qquad \forall A \in \mathcal{F}.$$

By point 1. of Proposition 19, we have that $R = \bar{\mu}_{\tilde{Q}^W} \in \Delta^{\sigma}(\Omega,\mathcal{F},\mathcal{P})$. By point 4. of Proposition 19, we can conclude that p is a regular conditional probability for it, proving that $R \in \mathcal{C}$. □

Hence, by Proposition 19, we can conclude that

$$\begin{aligned}
\mathcal{P} &= \left\{\bar{\mu}_{(P_{\mathcal{G}})^W} : P \in \mathcal{P}\right\} \subseteq \left\{\bar{\mu}_{Q^W} : Q \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})\right\} \\
&= \left\{\int_{\Omega} p(\cdot,\omega)\, dQ(\omega) : Q \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})\right\} = \mathcal{C} = \left\{\bar{\mu}_{Q^W} : Q \in \Delta^{\sigma}(\Omega,\mathcal{G},\mathcal{P})\right\} \\
&\subseteq \left\{\bar{\mu} : \mu \in \Delta^{\sigma}\left(\mathcal{S}(\mathcal{P}),\mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)\right\} \subseteq \left\{\bar{\mu} : \mu \in \Delta^{\sigma}(\mathcal{P},\mathcal{A}_{\mathcal{P}})\right\} \subseteq \operatorname{mco}(\mathcal{P})
\end{aligned}$$

where the first equality follows by (26), the first inclusion is immediate since $P_{\mathcal{G}} \in \Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})$ for all $P \in \mathcal{P}$, the second equality follows by (28), the third equality follows by Step 2, the second inclusion follows by point 5. of Proposition 19, the third inclusion is obvious, and the last inclusion follows by the definition of $\operatorname{mco}(\mathcal{P})$.

Finally, by Step 1 we have that $\mathcal{C}$ is measure convex and it contains $\mathcal{P}$. Given the previous chain of inclusions and equalities and since $\mathcal{P} \subseteq \mathcal{C}$, this implies that

$$\begin{aligned}\operatorname{mco}(\mathcal{P}) \subseteq \mathcal{C} = \left\{\bar{\mu}_{Q^W} : Q \in \Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})\right\} &= \left\{\int_{\Omega} p(\cdot, \omega)\, dQ(\omega) : Q \in \Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})\right\} = \mathcal{C} \\ \subseteq \left\{\bar{\mu} : \mu \in \Delta^{\sigma}\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)\right\} &= \left\{\bar{\mu} : \mu \in \Delta^{\sigma}(\mathcal{P}, \mathcal{A}_{\mathcal{P}})\right\} \subseteq \operatorname{mco}(\mathcal{P}),\end{aligned}$$

proving the first part of the proposition.

1. Consider the restriction map $R \mapsto R_{\mathcal{G}}$ from $\operatorname{mco}(\mathcal{P}) = \mathcal{C}$ to $\Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})$. Clearly, this mapping is well defined and affine. Injectivity follows from the marginal-conditional decomposition (29). Indeed, if $R, \tilde{R} \in \mathcal{C}$ then

$$\begin{aligned}R_{\mathcal{G}} = \tilde{R}_{\mathcal{G}} &\Rightarrow \int_{\Omega} p(A, \cdot)\, dR_{\mathcal{G}} = \int_{\Omega} p(A, \cdot)\, d\tilde{R}_{\mathcal{G}}, \quad \forall A \in \mathcal{F} \\ &\Rightarrow R(A) = \tilde{R}(A) \qquad \forall A \in \mathcal{F}.\end{aligned}$$

Finally, by point 3. of Proposition 19, for each $Q \in \Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})$, $\left(\bar{\mu}_{Q^W}\right)_{\mathcal{G}} = Q$ and by the first part of the theorem $\bar{\mu}_{Q^W} \in \operatorname{mco}(\mathcal{P})$.

2. By point 5. of Proposition 19, for each $Q \in \Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})$, $\mu_{Q^W} = Q^W \circ p_{|W}^{-1}$ is a measure on $\mathcal{P}$ such that $\mu_{Q^W}(\mathcal{S}(\mathcal{P})) = 1$. Since $\mathcal{S}(\mathcal{P}) \in \mathcal{A}_{\mathcal{P}}$, without loss of generality, we can assume that $\mu_{Q^W} \in \Delta^{\sigma}\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$, by considering $\mu_{Q^W | \mathcal{A}_{\mathcal{P}} \cap \mathcal{S}(\mathcal{P})}$. By Fact 4, for each $Q, \tilde{Q} \in \Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})$ and $\alpha \in [0,1]$, we have that $\left(\alpha Q + (1-\alpha)\tilde{Q}\right)^W = \alpha Q^W + (1-\alpha)\tilde{Q}^W$. Then, it follows that for each $\Gamma \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$

$$\begin{aligned}\mu_{\left(\alpha Q + (1-\alpha)\tilde{Q}\right)^W}(\Gamma) &= \left(\alpha Q + (1-\alpha)\tilde{Q}\right)^W \left(p_{|W}^{-1}(\Gamma)\right) \\ &= \alpha Q^W \left(p_{|W}^{-1}(\Gamma)\right) + (1-\alpha)\tilde{Q}^W \left(p_{|W}^{-1}(\Gamma)\right) \\ &= \alpha \mu_{Q^W}(\Gamma) + (1-\alpha)\mu_{\tilde{Q}^W}(\Gamma),\end{aligned}$$

that is, $Q \mapsto \mu_{Q^W}$ is affine. Injectivity follows from point 3. of Proposition 19. Indeed, for each $Q, \tilde{Q} \in \Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})$,

$$\mu_{Q^W} = \mu_{\tilde{Q}^W} \Rightarrow Q = \left(\bar{\mu}_{Q^W}\right)_{\mathcal{G}} = \left(\bar{\mu}_{\tilde{Q}^W}\right)_{\mathcal{G}} = \tilde{Q}.$$

Finally, if $\lambda \in \Delta^{\sigma}\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$ then $\bar{\lambda} \in \operatorname{mco}(\mathcal{P})$. By the initial part of the proof and point 1., $\bar{\lambda} \in \Delta^{\sigma}(\Omega, \mathcal{F}, \mathcal{P})$ and $Q = \bar{\lambda}_{\mathcal{G}} \in \Delta^{\sigma}(\Omega, \mathcal{G}, \mathcal{P})$ and $\bar{\lambda} = \bar{\mu}_{Q^W}$. It follows that for each $\Gamma \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$

$$\begin{aligned}\mu_{\left(\bar{\lambda}_{\mathcal{G}}\right)^W}(\Gamma) &= \left(\bar{\lambda}_{\mathcal{G}}\right)^W \left(p_{|W}^{-1}(\Gamma)\right) = Q^W \left(p_{|W}^{-1}(\Gamma)\right) \\ &= \bar{\mu}_{Q^W}\left(p_{|W}^{-1}(\Gamma)\right) = \bar{\lambda}\left(p_{|W}^{-1}(\Gamma)\right) = \int_{\mathcal{S}(\mathcal{P})} P\left(p_{|W}^{-1}(\Gamma)\right) d\lambda(P) \\ &= \int_{\mathcal{S}(\mathcal{P})} 1_{\Gamma}(P)\, d\lambda(P) = \lambda(\Gamma),\end{aligned}$$

that is, $\lambda = \mu_{\left(\bar{\lambda}_{\mathcal{G}}\right)^W}$. We prove the last equality below. Indeed, we will show that $P\left(p_{|W}^{-1}(\Gamma)\right) = 1_{\Gamma}(P)$ for all $P \in \mathcal{S}(\mathcal{P})$.

If P in $\mathcal{S}(\mathcal{P})$ belongs to Γ then

$$\begin{aligned}1 = P(\{\omega \in \Omega : p^{\omega} = P\}) &= P(\{\omega \in W : p^{\omega} = P\}) \\ &\leq P(\{\omega \in W : p^{\omega} \in \Gamma\}) = P\left(p_{|W}^{-1}(\Gamma)\right) \leq 1.\end{aligned}$$

If P in $\mathcal{S}(\mathcal{P})$ does not belong to Γ then $\{\omega \in W : p^\omega \in \Gamma\} \subseteq \{\omega \in W : p^\omega \neq P\}$ and

$$\begin{aligned} 0 = P(\{\omega \in \Omega : p^\omega \neq P\}) &= P(\{\omega \in W : p^\omega \neq P\}) \\ \geq P(\{\omega \in W : p^\omega \in \Gamma\}) &= P\left(p_{|W}^{-1}(\Gamma)\right) \geq 0, \end{aligned}$$

as wanted.

3. Let $R \in \Delta^\sigma(\Omega, \mathcal{F})$ be such that there exists a net $\{R_\alpha\}_{\alpha\in A}$ in $\mathrm{mco}(\mathcal{P})$ such that

$$\lim_\alpha \langle f, R_\alpha\rangle = \langle f, R\rangle \qquad \forall f \in B(\Omega, \mathcal{F}).$$

It is immediate to see that $R \in \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$. Moreover, for each $A \in \mathcal{F}$ and $B \in \mathcal{G}$,

$$\begin{aligned} R(A \cap B) &= \lim_\alpha R_\alpha(A \cap B) = \lim_\alpha \int_\Omega p(A, \cdot)\, 1_B(\cdot)\, dR_\alpha \\ &= \lim_\alpha \langle p(A, \cdot)\, 1_B(\cdot), R_\alpha\rangle = \langle p(A, \cdot)\, 1_B(\cdot), R\rangle \\ &= \int_\Omega p(A, \cdot)\, 1_B(\cdot)\, dR, \end{aligned}$$

that is, p is a r.c.p. given $\mathcal{G}$ for R. By the initial part of the proof, it follows that $R \in \mathrm{mco}(\mathcal{P})$. ■

With the symbol $\leftrightarrow$, we denote affine bijections. In light of Proposition 18, we can conclude that

$$\begin{array}{ccccc} \mathrm{mco}(\mathcal{P}) & \leftrightarrow & \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P}) & \leftrightarrow & \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right) \\ R & \rightarrowtail & R_{\mathcal{G}} & \rightarrowtail & \mu_{(R_{\mathcal{G}})^W} \\ \bar{\lambda} & \leftarrowtail & \bar{\lambda}_{\mathcal{G}} & \leftarrowtail & \lambda \\ \mathcal{P} \ni P & \rightarrowtail & P_{\mathcal{G}} & \rightarrowtail & \mu_P \end{array} \tag{30}$$

and we just write μ_R for $\mu_{(R_{\mathcal{G}})^W}$ if $R \in \mathrm{mco}(\mathcal{P})$.[34] By (30):

Corollary 20 *For each $R \in \mathrm{mco}(\mathcal{P})$, $\bar{\mu}_R = R$ and μ_R is the unique element of $\Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$ having R as its barycenter.*

C Quasiconcave Duality Theory

In this appendix, we present the basic notions of quasiconcave duality theory. We denote by (S, Σ) a measurable space. In particular, (S, Σ) will be either $(\Omega, \mathcal{F})$ or $\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. We denote by $B_0(S, \Sigma)$ and $B(S, \Sigma)$, respectively, the set of all simple and measurable real valued functions on S and the set of all bounded and measurable functions on S. We denote by $\mathcal{L}(\mathbb{R} \times \Delta(S, \Sigma))$ the class of functions $G : \mathbb{R} \times \Delta(S, \Sigma) \to (-\infty, \infty]$ such that

[34] Indeed, for each $B \in \mathcal{G}$,

$$R(B) = R(B \cap W) = (R_{\mathcal{G}\cap W})_W(B),$$

that is, $(R_{\mathcal{G}\cap W})_W = R_{\mathcal{G}}$, $(R_{\mathcal{G}})^W = R_{\mathcal{G}\cap W}$, and $\mu_{(R_{\mathcal{G}})^W} = \mu_{R_{\mathcal{G}\cap W}}$. In other words, for each $\Gamma \in \mathcal{A}_{\mathcal{P}}$

$$\mu_{(R_{\mathcal{G}})^W}(\Gamma) = R_{\mathcal{G}\cap W}\left(p_{|W}^{-1}(\Gamma)\right) = R(\{\omega \in W : p^\omega \in \Gamma\}).$$

Moreover, $\{\omega \in \Omega : p^\omega \in \Gamma\} = \{\omega \in W : p^\omega \in \Gamma\} \cup \{\omega \in W^c : p^\omega \in \Gamma\}$, might not belong to $\mathcal{F}$, but it must belong to the R completion of $\mathcal{F}$ for all $R \in \mathrm{mco}(\mathcal{P})$ ($\{\omega \in W : p^\omega \in \Gamma\} \in \mathcal{G} \cap W \subseteq \mathcal{F}$, $P(W^c) = 0$ for all $P \in \mathcal{P}$, and $R \in \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$). Thus we can write

$$\mu_R(\Gamma) := \mu_{(R_{\mathcal{G}})^W}(\Gamma) = R(\{\omega \in W : p^\omega \in \Gamma\}) = R(\{\omega \in \Omega : p^\omega \in \Gamma\}), \quad \forall R \in \mathrm{mco}(\mathcal{P}), \Gamma \in \mathcal{A}_{\mathcal{P}}.$$

In other words, modulo completion $\mu_R(\Gamma) = R(\{\omega \in \Omega : p^\omega \in \Gamma\})$.

- G is quasiconvex and lower semicontinuous;
- $G(\cdot, P)$ is increasing for all $P \in \Delta(S, \Sigma)$;
- $t = \min_{P \in \Delta(S,\Sigma)} G(t, P)$ for all $t \in \mathbb{R}$.

We denote by $\mathrm{dom}_{\Delta} G$ the set $\{P \in \Delta(S, \Sigma) : G(t, P) < \infty$ for some $t \in \mathbb{R}\}$.

Theorem 21 *Let Y be either $B_0(S, \Sigma)$ or $B(S, \Sigma)$. A functional $I : Y \to \mathbb{R}$ is normalized, monotone, lower semicontinuous, and quasiconcave if and only if there exists a unique function $G \in \mathcal{L}(\mathbb{R} \times \Delta(S, \Sigma))$ such that*

$$I(\varphi) = \min_{P \in \Delta(S,\Sigma)} G\left(\int_S \varphi dP, P\right) \qquad \forall \varphi \in Y. \tag{31}$$

Moreover,

1. $G(t, P) = \sup\left\{I(\varphi) : \int_S \varphi dP \leq t \text{ and } \varphi \in Y\right\} = \left\{I(\varphi) : \int_S \varphi dP \leq t \text{ and } \varphi \in B_0(S, \Sigma)\right\}$ *for all* (t, P) *in* $\mathbb{R} \times \Delta(S, \Sigma)$.

2. *I is inner continuous on $B_0(S, \Sigma)$ if and only if*

$$G(t, P) = \infty \qquad \forall (t, P) \notin \mathbb{R} \times \Delta^{\sigma}(S, \Sigma).$$

3. *I is translation invariant if and only if there exists a convex and lower semicontinuous function c : $\Delta(S, \Sigma) \to [0, \infty]$ such that*

$$\min_{P \in \Delta(S,\Sigma)} c(P) = 0 \text{ and } G(t, P) = t + c(P) \qquad \forall (t, P) \in \mathbb{R} \times \Delta(S, \Sigma).$$

Proof. We start by observing that Y, in both cases, is an M-space with unit. By [11, Theorem 3 and Lemma 20] and since I is real valued, if $I : Y \to \mathbb{R}$ is a normalized, monotone, lower semicontinuous, quasiconcave, real valued functional then there exists a unique function $G : \mathbb{R} \times \Delta(S, \Sigma) \to (-\infty, \infty]$ such that
- $G(\cdot, P)$ is increasing for all $P \in \Delta(S, \Sigma)$;
- $\lim_{t \to \infty} G(t, P) = \lim_{t \to \infty} G(t, P')$ for all $P, P' \in \Delta(S, \Sigma)$;
- G is quasiconvex and lower semicontinuous;
- $t = \min_{P \in \Delta(\Omega,\mathcal{F})} G(t, P)$ for all $t \in \mathbb{R}$;
- $I(\varphi) = \min_{P \in \Delta(\Omega,\mathcal{F})} G\left(\int_\Omega \varphi dP, P\right)$ for all $\varphi \in Y$.

Moreover, G satisfies the equation in 1. This proves necessity. Viceversa, consider $G \in \mathcal{L}(\mathbb{R} \times \Delta(S, \Sigma))$ and a functional $I : Y \to [-\infty, \infty]$ such that (31) holds. Notice that the condition $t = \min_{P \in \Delta(\Omega,\mathcal{F})} G(t, P)$ for all $t \in \mathbb{R}$ implies that $\lim_{t \to \infty} G(t, P) = \infty = \lim_{t \to \infty} G(t, P')$ for all $P, P' \in \Delta(S, \Sigma)$. By [11, Theorem 3] and the proof of [11, Lemma 20], we have that I is a normalized, monotone, lower semicontinuous, and quasiconcave functional. In particular, I is real valued.

1. It follows from [11, Theorem 3].
2. It follows from [10, Theorem 54].
3. It follows from [11, Theorem 9]. ■

D Dynkin Functionals

If S is a nonempty set then we say that a subset $\mathcal{L} \subseteq B(S)$ is a Stone vector lattice if and only if $\mathcal{L}$ is a vector subspace of $B(S)$, a lattice,[35] and $1_S \in \mathcal{L}$. We endow $\mathcal{L}$ with the supnorm. Given a sequence $\{\xi_n\}_{n \in \mathbb{N}}$ and an element ξ in $B(S)$, we write $\xi_n \to \xi$ if and only if $\{\xi_n\}_{n \in \mathbb{N}}$ converges uniformly to ξ. On

[35] That is, for each $\xi_1, \xi_2 \in \mathcal{L}$ we have that $\xi_1 \vee \xi_2, \xi_1 \wedge \xi_2 \in \mathcal{L}$.

the other hand, we write that $\xi_n \uparrow \xi$ (resp., $\xi_n \downarrow \xi$) if and only if $\xi_{n+1} \geq \xi_n$ (resp., $\xi_n \geq \xi_{n+1}$) for all $n \in \mathbb{N}$ and $\lim_n \xi_n(s) = \xi(s)$ for all $s \in S$. Given a function $\varphi \in B(\Omega, \mathcal{F})$, we denote by $\langle \varphi, \cdot \rangle_{|\mathcal{S}(\mathcal{P})} = \langle \varphi, \cdot \rangle$ the functional on $\mathcal{S}(\mathcal{P})$ such that $P \mapsto \langle \varphi, P \rangle_{|\mathcal{S}(\mathcal{P})} = \int \varphi dP$ for all $P \in \mathcal{S}(\mathcal{P})$. Sometimes, with a small abuse of notation, we denote by $\langle \varphi, \cdot \rangle$ the functional $\mathcal{P}$ such that $P \mapsto \int \varphi dP$ for all $P \in \mathcal{P}$. It will be clear from the context what is the domain for the functional $\langle \varphi, \cdot \rangle$.

Lemma 22 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space. The sets*

$$L_0 = \left\{ \langle \psi, \cdot \rangle_{|\mathcal{S}(\mathcal{P})} : \psi \in B_0(\Omega, \mathcal{G}) \right\}$$
$$\text{and } L = \left\{ \langle \psi, \cdot \rangle_{|\mathcal{S}(\mathcal{P})} : \psi \in B(\Omega, \mathcal{G}) \right\} = \left\{ \langle \varphi^\star, \cdot \rangle_{|\mathcal{S}(\mathcal{P})} : \varphi \in B(\Omega, \mathcal{F}) \right\} = \left\{ \langle \varphi, \cdot \rangle_{|\mathcal{S}(\mathcal{P})} : \varphi \in B(\Omega, \mathcal{F}) \right\}$$

are Stone vector lattices in $B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. Moreover, L_0 is supnorm dense in L.

Proof. The chain of equalities contained in the statement follows from the fact that for each $\varphi \in B(\Omega, \mathcal{F})$ we have that $\langle \varphi^\star, \cdot \rangle_{|\mathcal{S}(\mathcal{P})} = \langle \varphi, \cdot \rangle_{|\mathcal{S}(\mathcal{P})}$, moreover, $\varphi^\star$ is an element of $B(\Omega, \mathcal{G})$ for each $\varphi \in B(\Omega, \mathcal{F})$ and $B(\Omega, \mathcal{G})$ is contained in $B(\Omega, \mathcal{F})$. Since $B_0(\Omega, \mathcal{G})$ and $B(\Omega, \mathcal{G})$ are vector spaces, it is immediate to see that L_0 and L are vector spaces as well.

For each $\psi \in B(\Omega, \mathcal{G})$,

$$|\langle \psi, P \rangle| = \left| \int_\Omega \psi dP \right| \leq \int_\Omega |\psi| \, dP \leq \|\psi\| \qquad \forall P \in \mathcal{S}(\mathcal{P}).$$

This implies that L_0 and L are vector subspaces of $B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. Moreover, since $\psi = 1_\Omega$ belongs to $B_0(\Omega, \mathcal{G})$, we have that $1_{\mathcal{S}(\mathcal{P})}$ belongs to L_0 and L.

Finally, consider $\langle \psi_1, \cdot \rangle, \langle \psi_2, \cdot \rangle \in L_0$ with $\psi_1, \psi_2 \in B_0(\Omega, \mathcal{G})$. Recall that $P \in \mathcal{S}(\mathcal{P})$ if and only if $P \in \mathcal{P}$ and $P(\mathcal{G}) = \{0, 1\}$. Then, there exist a partition $\{E_j\}_{j=1}^h \subseteq \mathcal{G}$ and two sets $\left\{\alpha_j^1\right\}_{j=1}^h, \left\{\alpha_j^2\right\}_{j=1}^h \subseteq \mathbb{R}$ such that

$$\psi_1 = \sum_{j=1}^h \alpha_j^1 1_{E_j} \text{ and } \psi_2 = \sum_{j=1}^h \alpha_j^2 1_{E_j}.$$

Notice that for each $P \in \mathcal{S}(\mathcal{P})$ there exists a unique E_{j_P} such that $P(E_{j_P}) = 1$. This implies that $\int_\Omega \psi_i dP = \alpha_{j_P}^i$ for $i \in \{1, 2\}$ and for all $P \in \mathcal{S}(\mathcal{P})$. Therefore, it follows that

$$(\langle \psi_1, \cdot \rangle \vee \langle \psi_2, \cdot \rangle)(P) = \langle \psi_1, P \rangle \vee \langle \psi_2, P \rangle = \alpha_{j_P}^1 \vee \alpha_{j_P}^2 = \int_\Omega (\psi_1 \vee \psi_2) \, dP = \langle \psi_1 \vee \psi_2, P \rangle \tag{32}$$

for all $P \in \mathcal{S}(\mathcal{P})$. Since $\psi_1 \vee \psi_2 \in B_0(\Omega, \mathcal{G})$, it follows that $\langle \psi_1, \cdot \rangle \vee \langle \psi_2, \cdot \rangle \in L_0$. If $\langle \psi_1, \cdot \rangle, \langle \psi_2, \cdot \rangle \in L$ with $\psi_1, \psi_2 \in B(\Omega, \mathcal{G})$ then there exist two sequences, $\{\psi_{1,n}\}_{n \in \mathbb{N}}$ and $\{\psi_{2,n}\}_{n \in \mathbb{N}}$, in $B_0(\Omega, \mathcal{G})$, such that $\psi_{i,n} \to \psi_i$ for $i \in \{1, 2\}$. Therefore, by (32) and the continuity of the lattice operations (that is, $\psi_{1,n} \vee \psi_{2,n} \to \psi_1 \vee \psi_2$) it follows that

$$\begin{aligned} (\langle \psi_1, \cdot \rangle \vee \langle \psi_2, \cdot \rangle)(P) &= \langle \psi_1, P \rangle \vee \langle \psi_2, P \rangle = \lim_n \left(\langle \psi_{1,n}, P \rangle \vee \langle \psi_{2,n}, P \rangle \right) \\ &= \lim_n \langle \psi_{1,n} \vee \psi_{2,n}, P \rangle = \langle \psi_1 \vee \psi_2, P \rangle \end{aligned} \tag{33}$$

for all $P \in \mathcal{S}(\mathcal{P})$. Since $\psi_1 \vee \psi_2 \in B(\Omega, \mathcal{G})$, it follows that $\langle \psi_1, \cdot \rangle \vee \langle \psi_2, \cdot \rangle \in L$. By the same argument, $\langle \psi_1, \cdot \rangle \wedge \langle \psi_2, \cdot \rangle \in L_0$ (resp., L). Thus, L_0 and L are Stone vector lattices.

Finally, since $B_0(\Omega, \mathcal{G})$ is supnorm dense in $B(\Omega, \mathcal{G})$, if $\langle \psi, \cdot \rangle \in L$ with $\psi \in B(\Omega, \mathcal{G})$ then there exists a sequence $\{\psi_n\}_{n \in \mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$ such that $\psi_n \to \psi$. It follows that for each $n \in \mathbb{N}$ and for each $P \in \mathcal{S}(\mathcal{P})$

$$|\langle \psi, P \rangle - \langle \psi_n, P \rangle| = |\langle \psi - \psi_n, P \rangle| \leq \|\psi - \psi_n\|.$$

This implies that $\sup_{P\in\mathcal{S}(\mathcal{P})}|\langle\psi,P\rangle-\langle\psi_n,P\rangle|\leq\|\psi-\psi_n\|\to 0$. By definition of L_0, we have that $\{\langle\psi_n,\cdot\rangle\}_{n\in\mathbb{N}}$ is a sequence in L_0, proving that L_0 is supnorm dense in L. ■

By similar arguments, we obtain the following lemma:

Lemma 23 *Let $(\Omega,\mathcal{F},\mathcal{P})$ be a Dynkin space. The set*

$$L_{00}=\left\{\langle\psi,\cdot\rangle_{|\mathcal{S}(\mathcal{P})}:\psi\in B_0(\Omega,\mathcal{F})\right\}$$

is a vector subspace of $B\left(\mathcal{S}(\mathcal{P}),\mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$ and $1_{\mathcal{S}(\mathcal{P})}\in L_{00}$. Moreover, L_{00} is supnorm dense in L.

Lemma 24 *Let $(\Omega,\mathcal{F},\mathcal{P})$ be a Dynkin space.*

(a) If $\xi_1,\xi_2\in L_0$ are such that $\xi_1\geq\xi_2$ then there exist φ_1 and φ_2 in $B_0(\Omega,\mathcal{G})$ such that $\varphi_1\geq\varphi_2$ and

$$\xi_i(P)=\langle\varphi_i,P\rangle\qquad\forall P\in\mathcal{S}(\mathcal{P}),\forall i\in\{1,2\}. \tag{34}$$

In particular, if $\xi_2=0$ then we can take $\varphi_2=0$.

(b) Given $\{\xi_n\}_{n\in\mathbb{N}}\subseteq L_0$ and $\xi\in L$, if $\xi_n\downarrow\xi$ (resp., $\uparrow$) then there exist a sequence $\{\varphi_n\}_{n\in\mathbb{N}}\subseteq B_0(\Omega,\mathcal{G})$ and $\varphi\in B(\Omega,\mathcal{G})$ such that $\varphi_n\downarrow\varphi$ (resp., $\uparrow$) and $\xi=\langle\varphi,\cdot\rangle$ and $\xi_n=\langle\varphi_n,\cdot\rangle$ for all $n\in\mathbb{N}$.

(c) If $\xi\in L^+$ then there exists $\varphi\in B^+(\Omega,\mathcal{G})$ such that $\xi=\langle\varphi,\cdot\rangle$.

(d) If $\xi_1,\xi_2\in L$ are such that $\xi_1\geq\xi_2$ then there exist φ_1 and φ_2 in $B(\Omega,\mathcal{G})$ such that $\varphi_1\geq\varphi_2$ and

$$\xi_i(P)=\langle\varphi_i,P\rangle\qquad\forall P\in\mathcal{S}(\mathcal{P}),\forall i\in\{1,2\}. \tag{35}$$

(e) Given $\{\xi_n\}_{n\in\mathbb{N}}\subseteq L$ and $\xi\in L$, $\xi_n\downarrow\xi$ (resp., $\uparrow$) if and only if there exist a sequence $\{\varphi_n\}_{n\in\mathbb{N}}\subseteq B(\Omega,\mathcal{G})$ and $\varphi\in B(\Omega,\mathcal{G})$ such that $\varphi_n\downarrow\varphi$ (resp., $\uparrow$) and $\xi=\langle\varphi,\cdot\rangle$ and $\xi_n=\langle\varphi_n,\cdot\rangle$ for all $n\in\mathbb{N}$.

(f) $L=B\left(\mathcal{S}(\mathcal{P}),\mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$.

(g) If $\varphi\in B(\Omega,\mathcal{G})$ and $v:\mathbb{R}\to\mathbb{R}$ is continuous then $v\circ\langle\varphi,\cdot\rangle=\langle v\circ\varphi,\cdot\rangle$.

(h) Given ξ_1 and ξ_2 in L, ξ_1 and ξ_2 are comonotonic if and only if there exist φ_1 and φ_2 in $B(\Omega,\mathcal{G})$ that are comonotonic and such that (35) holds. In particular, ξ_1 and ξ_2 are in L_0 if and only if φ_1 and φ_2 can be chosen to be in $B_0(\Omega,\mathcal{G})$.

(i) If $A\in\mathcal{G}$ then $\langle 1_A,\cdot\rangle\in B_0\left(\mathcal{S}(\mathcal{P}),\mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. In particular, $L_0\subseteq B_0\left(\mathcal{S}(\mathcal{P}),\mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$.

(l) If $C\in\mathcal{A}_{\mathcal{S}(\mathcal{P})}$ then there exists $A\in\mathcal{G}$ such that $\langle 1_A,\cdot\rangle=1_C$. In particular, $B_0\left(\mathcal{S}(\mathcal{P}),\mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)\subseteq L_0$.

(m) $L_0=B_0\left(\mathcal{S}(\mathcal{P}),\mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$.

(n) Given $\{C_n\}_{n\in\mathbb{N}}\subseteq\mathcal{A}_{\mathcal{S}(\mathcal{P})}$, $C_n\downarrow\emptyset$ if and only if there exists a sequence $\{E_n\}_{n\in\mathbb{N}}\subseteq\mathcal{G}$ such that $E_n\downarrow\emptyset$ and $1_{C_n}=\langle 1_{E_n},\cdot\rangle$ for all $n\in\mathbb{N}$.

Proof. (a) Consider $\xi_1,\xi_2\in L_0$ such that $\xi_1(P)\geq\xi_2(P)$ for all $P\in\mathcal{S}(\mathcal{P})$. Moreover, let $\psi_1,\psi_2\in B_0(\Omega,\mathcal{G})$ be such that

$$\xi_i=\langle\psi_i,\cdot\rangle\qquad\forall i\in\{1,2\}.$$

Then, there exist a partition $\{E_j\}_{j=1}^h \subseteq \mathcal{G}$ and two sets $\{\alpha_j\}_{j=1}^h, \{\beta_j\}_{j=1}^h \subseteq \mathbb{R}$ such that

$$\psi_1 = \sum_{j=1}^h \alpha_j 1_{E_j} \text{ and } \psi_2 = \sum_{j=1}^h \beta_j 1_{E_j}.$$

Set $J = \{j = 1, ..., h : P(E_j) \neq 0 \text{ for some } P \in \mathcal{P}\}$, $E = \bigcup_{j \in J} E_j$, and $\varphi_i = \psi_i 1_E$ for $i \in \{1, 2\}$. It is immediate to see that $\langle \varphi_i, \cdot \rangle = \langle \psi_i, \cdot \rangle = \xi_i$ for $i \in \{1, 2\}$.[36] For each $\omega \in \Omega$ there exists a unique $j \in \{1, ..., h\}$ such that $\omega \in E_j$. We have two cases:

1. $j \in J$. By definition of J, there exists $P_j \in \mathcal{P}$ such that $P_j(E_j) > 0$. By Theorem 17, it follows that

$$0 < P_j(E_j) = \int_{\mathcal{S}(\mathcal{P})} P(E_j) \, d\mu_{P_j}(P).$$

where $P(E_j) \in \{0, 1\}$ for all $P \in \mathcal{S}(\mathcal{P})$. This implies that there exists R_j in $\mathcal{S}(\mathcal{P})$ such that $R_j(E_j) = 1$. It follows that $R_j(E_k) = 0$ if $k \neq j$. Finally, we have that

$$\varphi_1(\omega) = \langle \varphi_1, R_j \rangle = \xi_1(R_j) \geq \xi_2(R_j) = \langle \varphi_2, R_j \rangle = \varphi_2(\omega).$$

2. $j \notin J$. Then, $1_E(\omega) = 0$ and $\varphi_1(\omega) = 0 \geq 0 = \varphi_2(\omega)$.

Summing up, $\langle \varphi_i, \cdot \rangle = \xi_i$ for $i \in \{1, 2\}$ and $\varphi_1 \geq \varphi_2$.

(b) Consider $\{\xi_n\}_{n \in \mathbb{N}} \subseteq L_0$ and $\xi \in L$ such that $\xi_n \downarrow \xi$. Moreover, assume that $\xi \geq 0$. By definition of L_0 and L, it follows that there exist $\{\psi_n\}_{n \in \mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$ and $\psi \in B(\Omega, \mathcal{G})$ such that $\xi = \langle \psi, \cdot \rangle$ and $\xi_n = \langle \psi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. By point (a), without loss of generality, we can assume that $\psi_n \geq 0$ for all $n \in \mathbb{N}$. Then, define the sequence $\{\varphi_n\}_{n \in \mathbb{N}}$ by

$$\varphi_n = \inf_{k \leq n} \psi_k \geq 0 \qquad \forall n \in \mathbb{N}.$$

Notice that $\{\varphi_n\}_{n \in \mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$ and $\varphi_n \downarrow \varphi = \inf_n \varphi_n = \inf_n \psi_n$ where φ is well defined, it belongs to $B(\Omega, \mathcal{G})$, and $\varphi \geq 0$. By (33), each element P in $\mathcal{S}(\mathcal{P})$ induces a lattice homomorphism on $B(\Omega, \mathcal{G})$. It follows that

$$\langle \varphi_n, P \rangle = \left\langle \inf_{k \leq n} \psi_k, P \right\rangle = \inf_{k \leq n} \langle \psi_k, P \rangle = \inf_{k \leq n} \xi_k(P) = \xi_n(P) = \langle \psi_n, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}), \forall n \in \mathbb{N}.$$

By the Dominated Convergence Theorem, this implies that

$$\xi(P) = \lim_n \xi_n(P) = \lim_n \langle \psi_n, P \rangle = \lim_n \langle \varphi_n, P \rangle = \langle \varphi, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}).$$

Consider $\{\xi_n\}_{n \in \mathbb{N}} \subseteq L_0$ and $\xi \in L$ such that $\xi_n \downarrow \xi$. There exists a constant $k \in \mathbb{R}$ such that $\xi_n + k1_{\mathcal{S}(\mathcal{P})} \downarrow \xi + k1_{\mathcal{S}(\mathcal{P})} \geq 0$. Since L_0 and L are Stone vector lattices and by the previous part of the proof, there exist a sequence $\{\psi_n\}_{n \in \mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$ and $\psi \in B(\Omega, \mathcal{G})$ such that $\psi_n \downarrow \psi$ where $\xi + k1_{\mathcal{S}(\mathcal{P})} = \langle \psi, \cdot \rangle$ and $\xi_n + k1_{\mathcal{S}(\mathcal{P})} = \langle \psi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. If we set $\varphi = \psi - k1_\Omega$ and $\varphi_n = \psi_n - k1_\Omega$ for all $n \in \mathbb{N}$ then the statement follows. Analogous considerations hold for the increasing case.

(c) Consider $\xi \geq 0$ in L. By definition, there exists $\psi \in B(\Omega, \mathcal{G})$ such that $\xi = \langle \psi, \cdot \rangle$. It follows that there exists a sequence $\{\psi_n\}_{n \in \mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$ such that $\psi_n \downarrow \psi$. Define $\{\xi_n\}_{n \in \mathbb{N}}$ by $\xi_n = \langle \psi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. By the Dominated Convergence Theorem, $\xi_n \downarrow \xi \geq 0$. By point (a) and since $\xi_n \geq 0$ for all $n \in \mathbb{N}$, there exists $\{\bar{\psi}_n\}_{n \in \mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$ such that $\xi_n = \langle \bar{\psi}_n, \cdot \rangle$ and $\bar{\psi}_n \geq 0$ for all $n \in \mathbb{N}$. Then, define the sequence $\{\varphi_n\}_{n \in \mathbb{N}}$ by

$$\varphi_n = \inf_{k \leq n} \bar{\psi}_k \geq 0 \qquad \forall n \in \mathbb{N}.$$

[36] Notice that if $\xi_2 = 0$ then ψ_2 can be chosen to be equal to 0. In turn, this implies that $\varphi_2 = 0$.

Notice that $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq B_0(\Omega,\mathcal{G})$ and $\varphi_n \downarrow \varphi = \inf_n \varphi_n = \inf_n \bar{\psi}_n$ where φ is well defined, it belongs to $B(\Omega,\mathcal{G})$, and $\varphi \geq 0$. By the same arguments used to prove point (b), it follows that $\xi = \langle \varphi, \cdot \rangle$, proving the statement.

(d) Let $\xi_1, \xi_2 \in L$ be such that $\xi_1 \geq \xi_2$. Then, $\xi_1 = \xi_2 + (\xi_1 - \xi_2)$. By point (c), there exists ψ in $B^+(\Omega,\mathcal{G})$ such that $\xi_1 - \xi_2 = \langle \psi, \cdot \rangle$. Consider $\varphi \in B(\Omega,\mathcal{G})$ such that $\xi_2 = \langle \varphi, \cdot \rangle$. It follows that $\xi_1 = \langle \varphi + \psi, \cdot \rangle$ and clearly $\varphi + \psi \geq \varphi$.

(e) We first prove necessity. Consider $\{\xi_n\}_{n\in\mathbb{N}} \subseteq L$ and $\xi \in L$ such that $\xi_n \downarrow \xi$. Moreover, assume that $\xi \geq 0$. By definition of L, it follows that there exist $\{\psi_n\}_{n\in\mathbb{N}} \subseteq B(\Omega,\mathcal{G})$ and $\psi \in B(\Omega,\mathcal{G})$ such that $\xi = \langle \psi, \cdot \rangle$ and $\xi_n = \langle \psi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. By point (c), without loss of generality, we can assume that $\psi_n \geq 0$ for all $n \in \mathbb{N}$. Define the sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ by

$$\varphi_n = \inf_{k\leq n} \psi_k \geq 0 \qquad \forall n \in \mathbb{N}.$$

Notice that $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq B(\Omega,\mathcal{G})$ and $\varphi_n \downarrow \varphi = \inf_n \varphi_n = \inf_n \psi_n$ where φ is well defined, it belongs to $B(\Omega,\mathcal{G})$, and $\varphi \geq 0$. By (33), each element P in $\mathcal{S}(\mathcal{P})$ induces a lattice homomorphism on $B(\Omega,\mathcal{G})$. It follows that

$$\langle \varphi_n, P \rangle = \left\langle \inf_{k\leq n} \psi_k, P \right\rangle = \inf_{k\leq n} \langle \psi_k, P \rangle = \inf_{k\leq n} \xi_k(P) = \xi_n(P) = \langle \psi_n, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}), \forall n \in \mathbb{N}.$$

By the Dominated Convergence Theorem, this implies that

$$\xi(P) = \lim_n \xi_n(P) = \lim_n \langle \psi_n, P \rangle = \lim_n \langle \varphi_n, P \rangle = \langle \varphi, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}).$$

Consider $\{\xi_n\}_{n\in\mathbb{N}} \subseteq L$ and $\xi \in L$ such that $\xi_n \downarrow \xi$. There exists a constant $k \in \mathbb{R}$ such that $\xi_n + k1_{\mathcal{S}(\mathcal{P})} \downarrow \xi + k1_{\mathcal{S}(\mathcal{P})} \geq 0$. Since L is a Stone vector lattice and by the previous part of the proof, there exist a sequence $\{\psi_n\}_{n\in\mathbb{N}} \subseteq B(\Omega,\mathcal{G})$ and $\psi \in B(\Omega,\mathcal{G})$ such that $\psi_n \downarrow \psi$ and $\xi + k1_{\mathcal{S}(\mathcal{P})} = \langle \psi, \cdot \rangle$ and $\xi_n + k1_{\mathcal{S}(\mathcal{P})} = \langle \psi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. If we set $\varphi = \psi - k1_\Omega$ and $\varphi_n = \psi_n - k1_\Omega$ for all $n \in \mathbb{N}$ then the statement follows. Analogous considerations hold for the increasing case.

Sufficiency follows by the Dominated Convergence Theorem.

(f) By Lemma 22, L is a Stone vector lattice in $B(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})})$ and $\mathcal{A}_{\mathcal{S}(\mathcal{P})}$ is the σ-algebra generated by L. By using point (e), we show that L is closed under bounded monotone convergence. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a bounded sequence in L such that $\xi_n \downarrow$. Define $\xi \in B(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})})$ to be the pointwise limit of $\{\xi_n\}_{n\in\mathbb{N}}$. Since $\{\xi_n\}_{n\in\mathbb{N}}$ is bounded and converging, it follows that ξ is well defined. Next, we show that $\xi \in L$. Assume that $\xi \geq 0$. By definition of L, it follows that there exists a sequence $\{\psi_n\}_{n\in\mathbb{N}} \subseteq B(\Omega,\mathcal{G})$ such that $\xi_n = \langle \psi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. By point (c), without loss of generality, we can assume that $\psi_n \geq 0$ for all $n \in \mathbb{N}$. Define now the sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ by

$$\varphi_n = \inf_{k\leq n} \psi_k \geq 0 \qquad \forall n \in \mathbb{N}.$$

Notice that $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq B(\Omega,\mathcal{G})$ and $\varphi_n \downarrow \varphi = \inf_n \varphi_n = \inf_n \psi_n$ where φ is well defined, it belongs to $B(\Omega,\mathcal{G})$, and $\varphi \geq 0$. By (33), each element P in $\mathcal{S}(\mathcal{P})$ induces a lattice homomorphism on $B(\Omega,\mathcal{G})$. It follows that

$$\langle \varphi_n, P \rangle = \left\langle \inf_{k\leq n} \psi_k, P \right\rangle = \inf_{k\leq n} \langle \psi_k, P \rangle = \inf_{k\leq n} \xi_k(P) = \xi_n(P) = \langle \psi_n, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}), \forall n \in \mathbb{N}.$$

Next, by the Dominated Convergence Theorem, observe that

$$\xi(P) = \lim_n \xi_n(P) = \lim_n \langle \psi_n, P \rangle = \lim_n \langle \varphi_n, P \rangle = \langle \varphi, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}).$$

Thus, $\xi = \langle \varphi, \cdot \rangle \in L^+$.

In general, if $\xi \not\geq 0$ then there exists a constant $k \in \mathbb{R}$ such that $\xi_n + k1_{\mathcal{S}(\mathcal{P})} \downarrow \xi + k1_{\mathcal{S}(\mathcal{P})} \geq 0$. Since L is a Stone vector lattice, $\left\{\xi_n + k1_{\mathcal{S}(\mathcal{P})}\right\}_{n\in\mathbb{N}} \subseteq L^+$. By the previous part of the proof, it follows that $\xi + k1_{\mathcal{S}(\mathcal{P})} \in L^+$, proving that $\xi \in L$. This shows that L is closed under bounded monotone (from above) pointwise convergence. A similar argument shows that L is closed under bounded monotone from below pointwise convergence. By [17, Theorem 22.3], we have that

$$\left\{\xi \in B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right) : \xi \in L\right\} = L \supseteq B\left(\mathcal{S}(\mathcal{P}), \sigma(L)\right) = B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right).$$

By Lemma 22, the statement follows.

(g) Let $\varphi = \sum_{j=1}^{h} \alpha_j 1_{A_j} \in B_0(\Omega, \mathcal{G})$ where $\{\alpha_j\}_{j=1}^h \subseteq \mathbb{R}$ and $\{A_j\}_{j=1}^h$ is a partition of Ω in $\mathcal{G}$. Each $P \in \mathcal{S}(\mathcal{P})$ is $\{0,1\}$-valued on $\mathcal{G}$. Therefore, for each $P \in \mathcal{S}(\mathcal{P})$ there exists a unique A_{j_P} in $\mathcal{G}$ such that $P(A_{j_P}) = 1$. This implies that

$$v(\langle \varphi, P \rangle) = v(\alpha_{j_P}) = \langle v \circ \varphi, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}),$$

proving the statement for $\varphi \in B_0(\Omega, \mathcal{G})$. If $\varphi \in B(\Omega, \mathcal{G})$ then there exists a sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ that converges uniformly to φ. Since v is continuous, it follows that the sequence $\{v \circ \varphi_n\}_{n\in\mathbb{N}}$ converges uniformly to $v \circ \varphi$. Since v is continuous and by the previous part of the proof, it follows that

$$v(\langle \varphi, P \rangle) = \lim_n v(\langle \varphi_n, P \rangle) = \lim_n \langle v \circ \varphi_n, P \rangle = \langle v \circ \varphi, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}),$$

proving the statement.

(h) Consider ξ_1 and ξ_2 in L and assume they are comonotonic. By [18, Proposition 4.5], there exist two monotone and continuous functions $v_1, v_2 : \mathbb{R} \to \mathbb{R}$ such that $\xi_i = v_i(\xi_1 + \xi_2)$ for $i \in \{1,2\}$. By definition of L, there exists $\varphi \in B(\Omega, \mathcal{G})$ such that $\xi_1 + \xi_2 = \langle \varphi, \cdot \rangle$. Define $\varphi_i = v_i \circ \varphi$ for $i \in \{1,2\}$. Since v_1 and v_2 are monotone and continuous and by [18, Proposition 4.5], it follows that φ_1 and φ_2 are comonotonic elements of $B(\Omega, \mathcal{G})$. By point (g), it follows that

$$\xi_i(P) = v_i((\xi_1 + \xi_2)(P)) = v_i(\langle \varphi, P \rangle) = \langle v_i \circ \varphi, P \rangle = \langle \varphi_i, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}), \forall i \in \{1,2\}.$$

Finally, observe that if ξ_1 and ξ_2 were elements of L_0 then φ could be chosen to be an element of $B_0(\Omega, \mathcal{G})$ and so $v_i \circ \varphi = \varphi_i \in B_0(\Omega, \mathcal{G})$ for $i \in \{1,2\}$. Viceversa, if φ_1 and φ_2 are comonotonic elements of $B(\Omega, \mathcal{G})$ then there exist two monotone and continuous functions $v_1, v_2 : \mathbb{R} \to \mathbb{R}$ such that $\varphi_i = v_i(\varphi_1 + \varphi_2)$ for $i \in \{1,2\}$. Consider $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1,2\}$. By point (g), it follows that

$$\xi_i(P) = \langle \varphi_i, P \rangle = \langle v_i \circ (\varphi_1 + \varphi_2), P \rangle = v_i(\langle \varphi_1 + \varphi_2, P \rangle) = v_i((\xi_1 + \xi_2)(P)) \qquad \forall P \in \mathcal{S}(\mathcal{P}), \forall i \in \{1,2\}.$$

By [18, Proposition 4.5], it follows that $\xi_1 = \langle \varphi_1, \cdot \rangle$ and $\xi_2 = \langle \varphi_2, \cdot \rangle$ are comonotonic. By definition of L_0, observe that if φ_1 and φ_2 were elements of $B_0(\Omega, \mathcal{G})$ then ξ_1 and ξ_2 would be elements of L_0.

(i) Let $A \in \mathcal{G}$. Since $P \in \mathcal{S}(\mathcal{P})$ if and only if $P(A) \in \{0,1\}$ for all $A \in \mathcal{G}$, it follows that $\langle 1_A, \cdot \rangle$ is a $\{0,1\}$-valued function. Since $\langle 1_A, \cdot \rangle \in L_0$, $\langle 1_A, \cdot \rangle$ is $\mathcal{A}_{\mathcal{S}(\mathcal{P})}$-measurable, proving the statement. In particular, if $\varphi \in B_0(\Omega, \mathcal{G})$ then $\varphi = \sum_{j=1}^h \alpha_j 1_{A_j}$ where $\{\alpha_j\}_{j=1}^h \subseteq \mathbb{R}$ and $\{A_j\}_{j=1}^h$ is a partition of Ω in $\mathcal{G}$. In other words, φ is a linear combination of $\mathcal{G}$-measurable indicator functions. Thus, $\langle \varphi, \cdot \rangle = \sum_{j=1}^h \alpha_j \langle 1_{A_j}, \cdot \rangle$ is a linear combination of $\mathcal{A}_{\mathcal{S}(\mathcal{P})}$-measurable indicator functions. That is, $L_0 \subseteq B_0\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$.

(l) Consider $C \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$. By point (f), $1_C \in L$. By definition, there exists $\psi \in B(\Omega, \mathcal{G})$ such that $\langle \psi, \cdot \rangle = 1_C$. Without loss of generality, we can assume that $0 \le \psi \le 1$.[37] Then, there exists a sequence $\{\psi_n\}_{n\in\mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$ such that $\psi_n \uparrow \psi$ uniformly and $0 \le \psi_n \le \psi$ for all $n \in \mathbb{N}$. Define $\xi_n = \langle \psi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. It follows that $0 \le \xi_n \uparrow 1_C$ uniformly. By construction, for each $n \in \mathbb{N}$ there exist a partition $\{E_{n,j}\}_{j=1}^{h_n}$ of Ω in $\mathcal{G}$ and a set $\{\alpha_{n,j}\}_{j=1}^{h_n} \subseteq \mathbb{R}$ such that

$$\psi_n = \sum_{j=1}^{h_n} \alpha_{n,j} 1_{E_{n,j}}.$$

For each $n \in \mathbb{N}$ set $J_n = \{j = 1, ..., h_n : P(E_{n,j}) \neq 0 \text{ for some } P \in \mathcal{P}\}$, $E = \bigcap_{n\in\mathbb{N}} \bigcup_{j\in J_n} E_{n,j}$, and set $\varphi_n = \psi_n 1_E$. Notice that $0 \le \varphi_n \le \psi_n \le \psi \le 1$ and $\xi_n = \langle \varphi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. Moreover, we have that $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$, $\varphi_n \uparrow \varphi = \psi 1_E \in B(\Omega, \mathcal{G})$, and $\langle \varphi, \cdot \rangle = \langle \psi, \cdot \rangle = 1_C$.[38] Finally, we show that φ is $\{0,1\}$-valued, that is, $\varphi = 1_A$ for some $A \in \mathcal{G}$. By contradiction, assume that there exists $\bar{\omega} \in \Omega$ such that $\varphi(\bar{\omega}) = \alpha \in (0,1)$. Since $\varphi_n \uparrow \varphi$, there exist $\varepsilon > 0$ and $n_1 \in \mathbb{N}$ such that $1 - \varepsilon > \varphi_n(\bar{\omega}) > \varepsilon$ for all $n \ge n_1$. By construction, it follows that for each $n \ge n_1$ there exists $P_n \in \mathcal{S}(\mathcal{P})$ such that $\langle \varphi_n, P_n \rangle = \varphi_n(\bar{\omega}) \in (\varepsilon, 1-\varepsilon)$. Since ξ_n converges uniformly to 1_C, there exists $n_2 \in \mathbb{N}$ such that $|\xi_n(P) - 1_C(P)| < \frac{\varepsilon}{2}$ for all $n \ge n_2$ and all $P \in \mathcal{S}(\mathcal{P})$. We can conclude that for each $n \ge \max\{n_1, n_2\}$:

- If $P_n \in C$ then $\frac{\varepsilon}{2} > |\xi_n(P_n) - 1| = 1 - \xi_n(P_n) = 1 - \langle \varphi_n, P_n \rangle$, that is, $\langle \varphi_n, P_n \rangle > 1 - \frac{\varepsilon}{2}$;
- If $P_n \notin C$ then $\frac{\varepsilon}{2} > |\xi_n(P_n)| = \xi_n(P_n) = \langle \varphi_n, P_n \rangle$, that is, $\langle \varphi_n, P_n \rangle < \frac{\varepsilon}{2}$,

a contradiction, since $\langle \varphi_n, P_n \rangle = \varphi_n(\bar{\omega}) \in (\varepsilon, 1-\varepsilon)$ for all $n \ge \max\{n_1, n_2\}$. Finally, since L_0 is a vector space and indicator functions of $\mathcal{A}_{\mathcal{S}(\mathcal{P})}$ belong to L_0, it follows that $B_0\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right) \subseteq L_0$.

(m) By points (i) and (l), the statement clearly follows.

---

[37] Consider $\psi \in B(\Omega, \mathcal{G})$ such that $\langle \psi, \cdot \rangle = 1_C$. By assumption, we have that $\langle \psi, P \rangle \in [0,1]$ for all $P \in \mathcal{S}(\mathcal{P})$. Define $E_1 = \{\omega \in \Omega : \psi(\omega) > 1\}$ and $E_2 = \{\omega \in \Omega : \psi(\omega) < 0\}$. It is immediate to see that $E_1, E_2 \in \mathcal{G}$. By contradiction, assume that there exists $\bar{P} \in \mathcal{S}(\mathcal{P})$ such that $\bar{P}(E_1) \neq 0$. Recall that $P(\mathcal{G}) = \{0,1\}$ for all $P \in \mathcal{S}(\mathcal{P})$. It follows that $\bar{P}(E_1) = 1$. Notice that

$$E_1 = \bigcup_n \left\{\omega \in \Omega : \psi(\omega) \ge 1 + \frac{1}{n}\right\}$$

and

$$\left\{\omega \in \Omega : \psi(\omega) \ge 1 + \frac{1}{n}\right\} \in \mathcal{G} \qquad \forall n \in \mathbb{N}.$$

Since $\bar{P} \in \mathcal{S}(\mathcal{P})$, it follows that there exists $\bar{n}$ such that

$$\bar{P}\left(\left\{\omega \in \Omega : \psi(\omega) \ge 1 + \frac{1}{\bar{n}}\right\}\right) = 1.$$

We could then conclude that $1 \ge \langle \psi, \bar{P} \rangle \ge 1 + \frac{1}{\bar{n}} > 1$, a contradiction. A similar argument shows that $P(E_2) = 0$ for all $P \in \mathcal{S}(\mathcal{P})$. It is then immediate to see that if we define $\bar{\psi} = \psi 1_{E_1^c \cap E_2^c}$ then $0 \le \bar{\psi} \le 1$ and

$$\langle \bar{\psi}, \cdot \rangle = 1_C.$$

[38] In fact, $E \in \mathcal{G}$ and $E^c = \bigcup_{n\in\mathbb{N}} \left(\bigcup_{j\in J_n} E_{n,j}\right)^c = \bigcup_{n\in\mathbb{N}} \bigcup_{j\notin J_n} E_{n,j}$. It follows that for each $n \in \mathbb{N}$ and each $P \in \mathcal{P}$,

$$P\left(\left(\bigcup_{j\in J_n} E_{n,j}\right)^c\right) = P\left(\bigcup_{j\notin J_n} E_{n,j}\right) = \sum_{j\notin J_n} P(E_{n,j}) = 0.$$

Thus, $P(E^c) = 0$ for all $P \in \mathcal{P}$.

(n) We first prove necessity. Consider $\{C_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}_{\mathcal{S}(\mathcal{P})}$ such that $C_n \downarrow \emptyset$. By point (l), it follows that there exists a sequence $\{B_n\}_{n\in\mathbb{N}} \subseteq \mathcal{G}$ such that $1_{C_n} = \langle 1_{B_n}, \cdot \rangle$ for all $n \in \mathbb{N}$. Define $\{\psi_n\}_{n\in\mathbb{N}}$ by

$$\psi_n = \inf_{k\le n} 1_{B_k} \ge 0 \qquad \forall n \in \mathbb{N}.$$

Notice that $\{\psi_n\}_{n\in\mathbb{N}}$ is a decreasing sequence. Furthermore, ψ_n is $\{0,1\}$-valued for all $n \in \mathbb{N}$ and $\psi_n \downarrow \psi = \inf_n \psi_n = \inf_n 1_{B_n}$. It is immediate to see that ψ is well defined, $\{0,1\}$-valued, and $\mathcal{G}$-measurable. By (33), each element P in $\mathcal{S}(\mathcal{P})$ induces a lattice homomorphism on $B(\Omega, \mathcal{G})$. Since $C_n \downarrow \emptyset$, it follows that

$$\langle \psi_n, P \rangle = \left\langle \inf_{k\le n} 1_{B_k}, P \right\rangle = \inf_{k\le n} \langle 1_{B_k}, P \rangle = \inf_{k\le n} 1_{C_k}(P) = 1_{C_n}(P) \qquad \forall P \in \mathcal{S}(\mathcal{P}), \forall n \in \mathbb{N}.$$

Next, by the Dominated Convergence Theorem and since $C_n \downarrow \emptyset$, observe that

$$0 = \lim_n 1_{C_n}(P) = \lim_n \langle \psi_n, P \rangle = \langle \psi, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}). \tag{36}$$

Since $\psi \in B(\Omega, \mathcal{G})$ and ψ is $\{0,1\}$-valued, there exists a set $E \in \mathcal{G}$ such that $\psi = 1_E$. By (36), it follows that $P(E) = 0$ for all $P \in \mathcal{S}(\mathcal{P})$. Define $\{\varphi_n\}_{n\in\mathbb{N}}$ by $\varphi_n = \psi_n 1_{E^c}$ for all $n \in \mathbb{N}$. It follows that

$$1_{C_n}(P) = \langle \psi_n, P \rangle = \langle \varphi_n, P \rangle \qquad \forall P \in \mathcal{S}(\mathcal{P}), \forall n \in \mathbb{N}. \tag{37}$$

Since $\{\psi_n\}_{n\in\mathbb{N}}$ is a decreasing sequence of $\mathcal{G}$-measurable and $\{0,1\}$-valued functions, it follows that $\{\varphi_n\}_{n\in\mathbb{N}}$ is a decreasing sequence of $\mathcal{G}$-measurable and $\{0,1\}$-valued functions. Moreover, we have that $\varphi_n \downarrow \psi 1_{E^c} = 0$. Since each φ_n is $\mathcal{G}$-measurable and $\{0,1\}$-valued for all $n \in \mathbb{N}$, for each $n \in \mathbb{N}$ there exists a set $E_n \in \mathcal{G}$ such that $\varphi_n = 1_{E_n}$. By (37), it follows that $1_{C_n} = \langle 1_{E_n}, \cdot \rangle$ for all $n \in \mathbb{N}$. Since $\varphi_n \downarrow 0$, it follows that $E_n \downarrow \emptyset$, proving necessity.

Sufficiency follows by the Dominated Convergence Theorem. ■

Proposition 25 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space. $I : B(\Omega, \mathcal{G}) \to \mathbb{R}$ is a normalized functional such that*

$$\int_\Omega \varphi_1 dP \ge \int_\Omega \varphi_2 dP \qquad \forall P \in \mathcal{P} \Rightarrow I(\varphi_1) \ge I(\varphi_2) \tag{38}$$

if and only if there exists a normalized and monotone functional $\breve{I} : L \to \mathbb{R}$ such that

$$I(\varphi) = \breve{I}(\langle \varphi, \cdot \rangle) \qquad \forall \varphi \in B(\Omega, \mathcal{G}). \tag{39}$$

Moreover, $\breve{I}$ is unique and

(1) $\breve{I}$ is translation invariant if and only if I is translation invariant;

(2) $\breve{I}$ is lower/upper semicontinuous if and only if I is lower/upper semicontinuous;

(3) $\breve{I}$ is concave if and only if I is concave;

(4) $\breve{I}$ is quasiconcave if and only if I is quasiconcave;

(5) $\breve{I}$ is supermodular if and only if I is supermodular;

(6) $\breve{I}$ is comonotonic additive if and only if I is comonotonic additive;

(7) If I is quasiconcave and lower semicontinuous, $\breve{I}$ is inner continuous on L_0 if and only if I is inner continuous on $B_0(\Omega, \mathcal{G})$;

(8) $\breve{I}$ is inner/outer continuous if and only if I is inner/outer continuous.

Proof.

In order to prove the main statement, we proceed by steps. We start by proving necessity. Define $\breve{I}: L \to \mathbb{R}$ to be such that for each $\xi \in L$

$$\breve{I}(\xi) = I(\varphi) \text{ where } \varphi \in B(\Omega, \mathcal{G}) \text{ and } \xi(P) = \int_\Omega \varphi dP \text{ for all } P \in \mathcal{S}(\mathcal{P}). \tag{40}$$

Step 1: Let $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$.

$$\begin{aligned} \int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP \qquad \forall P \in \mathcal{P} \\ \Leftrightarrow \int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP \qquad \forall P \in \mathcal{S}(\mathcal{P}) \\ \Leftrightarrow \int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP \qquad \forall P \in \operatorname{mco}(\mathcal{P}) \end{aligned}$$

Proof of the Step.

Consider $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$. Since $\mathcal{S}(\mathcal{P}) \subseteq \mathcal{P}$, it is immediate to see that if $\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP$ for all $P \in \mathcal{P}$ then $\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP$ for all $P \in \mathcal{S}(\mathcal{P})$. Consider $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$. If φ_1 and φ_2 are such that

$$\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP \qquad \forall P \in \mathcal{S}(\mathcal{P})$$

then for each $\mu \in \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$

$$\bar{\mu}(\varphi_1) = \int_{\mathcal{S}(\mathcal{P})} \langle \varphi_1, P \rangle \, d\mu(P) \geq \int_{\mathcal{S}(\mathcal{P})} \langle \varphi_2, P \rangle \, d\mu(P) = \bar{\mu}(\varphi_2). \tag{41}$$

By Proposition 18, we have that $\operatorname{mco}(\mathcal{P}) = \left\{\bar{\mu} : \mu \in \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)\right\}$. By (41), this implies that

$$\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP \qquad \forall P \in \operatorname{mco}(\mathcal{P}).$$

Finally, since $\mathcal{P} \subseteq \operatorname{mco}(\mathcal{P})$, it is immediate to see that if $\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP$ for all $P \in \operatorname{mco}(\mathcal{P})$ then $\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP$ for all $P \in \mathcal{P}$, proving the statement. □

Step 2: $\breve{I}$ is well defined.

Proof of the Step.

Consider $\xi \in L$ and assume that there exist φ_1 and φ_2 in $B(\Omega, \mathcal{G})$ such that

$$\int_\Omega \varphi_1 dP = \xi(P) = \int_\Omega \varphi_2 dP \qquad \forall P \in \mathcal{S}(\mathcal{P}).$$

By Step 1, this implies that

$$\int_\Omega \varphi_1 dP = \xi(P) = \int_\Omega \varphi_2 dP \qquad \forall P \in \mathcal{P}.$$

By (38), it follows that $I(\varphi_1) = I(\varphi_2)$, proving the statement. □

Step 3: I is monotone.

Proof of the Step.

Consider $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ such that $\varphi_1 \geq \varphi_2$. Since $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$, this implies that

$$\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP \qquad \forall P \in \mathcal{P}.$$

By (38), it follows that $I(\varphi_1) \geq I(\varphi_2)$. □

Step 4: $\breve{I}$ is monotone.

Proof of the Step.

Consider $\xi_1, \xi_2 \in L$ such that $\xi_1 \geq \xi_2$. By point (d) of Lemma 24, there exist $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ such that $\varphi_1 \geq \varphi_2$ and

$$\xi_i(P) = \int_\Omega \varphi_i dP \qquad \forall P \in \mathcal{S}(\mathcal{P}), \forall i \in \{1, 2\}.$$

By definition of $\breve{I}$ and since I is monotone, this implies that $\breve{I}(\xi_1) = I(\varphi_1) \geq I(\varphi_2) = \breve{I}(\xi_2)$, proving the statement. □

Step 5: $\breve{I}$ is normalized.

Proof of the Step.

Consider $k \in \mathbb{R}$. Then, it follows that $k1_{\mathcal{S}(\mathcal{P})} = \langle k1_\Omega, \cdot \rangle$. By definition of $\breve{I}$ and since I is normalized,

$$\breve{I}\left(k1_{\mathcal{S}(\mathcal{P})}\right) = I(k1_\Omega) = k.$$

Since k was arbitrarily chosen, the statement follows. □

Step 6: If $\breve{I}$ is a normalized and monotone functional on L such that (39) *holds then I is a normalized functional on $B(\Omega, \mathcal{G})$. Moreover, if $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ then*

$$\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP \qquad \forall P \in \mathcal{P} \Rightarrow I(\varphi_1) \geq I(\varphi_2). \tag{42}$$

Proof of the Step.

Fix $k \in \mathbb{R}$. By (39) and since $\breve{I}$ is normalized, it follows that

$$I(k1_\Omega) = \breve{I}(\langle k1_\Omega, \cdot \rangle) = \breve{I}\left(k1_{\mathcal{S}(\mathcal{P})}\right) = k.$$

Since k was arbitrarily chosen, it follows that I is normalized. Next, consider $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ such that $\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP$ for all $P \in \mathcal{P}$. It follows that $\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP$ for all $P \in \mathcal{S}(\mathcal{P})$. Define $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1, 2\}$. It is immediate to see that $\xi_1, \xi_2 \in L$ and $\xi_1 \geq \xi_2$. By (39) and since $\breve{I}$ is monotone, it follows that

$$I(\varphi_1) = \breve{I}(\langle \varphi_1, \cdot \rangle) = \breve{I}(\xi_1) \geq \breve{I}(\xi_2) = \breve{I}(\langle \varphi_2, \cdot \rangle) = I(\varphi_2),$$

proving that I satisfies (42). □

Step 7: $\breve{I}$ is unique.

Proof of the Step.

Assume that there exist two functionals over L, $\breve{I}_1$ and $\breve{I}_2$, such that

$$\breve{I}_1(\langle \varphi, \cdot \rangle) = I(\varphi) = \breve{I}_2(\langle \varphi, \cdot \rangle) \qquad \forall \varphi \in B(\Omega, \mathcal{G}).$$

Consider $\xi \in L$. By definition of L, there exists $\varphi \in B(\Omega, \mathcal{G})$ such that $\xi = \langle \varphi, \cdot \rangle$. It follows that

$$\breve{I}_1(\xi) = \breve{I}_1(\langle \varphi, \cdot \rangle) = I(\varphi) = \breve{I}_2(\langle \varphi, \cdot \rangle) = \breve{I}_2(\xi) \qquad \forall \xi \in L,$$

proving the statement. □

Steps 1 to 6 prove the main statement. Step 7 proves the uniqueness of $\breve{I}$.

(1) $\breve{I}$ is translation invariant if and only if I is translation invariant.

Proof.

We first prove sufficiency. Fix $\xi \in L$ and $k \in \mathbb{R}$. Consider $\varphi \in B(\Omega, \mathcal{G})$ such that $\xi = \langle \varphi, \cdot \rangle$. It follows that $\left(\xi + k1_{\mathcal{S}(\mathcal{P})}\right)(P) = \int_\Omega (\varphi + k1_\Omega)\, dP$ for all $P \in \mathcal{S}(\mathcal{P})$. By (39) and since I is translation invariant, we have that

$$\breve{I}\left(\xi + k1_{\mathcal{S}(\mathcal{P})}\right) = I(\varphi + k1_\Omega) = I(\varphi) + k = \breve{I}(\xi) + k.$$

Since ξ and k were arbitrarily chosen, sufficiency follows. Viceversa, fix $\varphi \in B(\Omega, \mathcal{G})$ and $k \in \mathbb{R}$. Define $\xi \in L$ by $\xi = \langle \varphi, \cdot \rangle$. It follows that $\xi + k1_{\mathcal{S}(\mathcal{P})} = \langle \varphi + k1_\Omega, \cdot \rangle$. By (39) and since $\breve{I}$ is translation invariant, we have that

$$I(\varphi + k1_\Omega) = \breve{I}\left(\xi + k1_{\mathcal{S}(\mathcal{P})}\right) = \breve{I}(\xi) + k = I(\varphi) + k.$$

Since φ and k were arbitrarily chosen, necessity follows.

(2) $\breve{I}$ is lower/upper semicontinuous if and only if I is lower/upper semicontinuous.

Proof.

We first prove a claim.

Claim: If I is lower (resp., upper) semicontinuous then the set $\left\{\alpha \in [0,1] : \breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) \leq c\right\}$ (resp., $\left\{\alpha \in [0,1] : \breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) \geq c\right\}$) is closed for all $\xi_1, \xi_2 \in L$ and $c \in \mathbb{R}$.

Proof of the Claim.

Fix $\xi_1, \xi_2 \in L$ and $c \in \mathbb{R}$. Set

$$\begin{aligned} L_{\xi_1,\xi_2,c} &= \left\{\alpha \in [0,1] : \breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) \leq c\right\} \\ \text{and } U_{\xi_1,\xi_2,c} &= \left\{\alpha \in [0,1] : \breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) \geq c\right\}. \end{aligned}$$

By point (f) of Lemma 24, we have that $L = B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. This implies that $\breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2)$ is well defined for all $\alpha \in [0,1]$. If $L_{\xi_1,\xi_2,c}$ (resp., $U_{\xi_1,\xi_2,c}$) is empty then $L_{\xi_1,\xi_2,c}$ (resp., $U_{\xi_1,\xi_2,c}$) is closed. Otherwise, consider $\{\alpha_n\}_{n\in\mathbb{N}} \subseteq L_{\xi_1,\xi_2,c}$ (resp., $\{\alpha_n\}_{n\in\mathbb{N}} \subseteq U_{\xi_1,\xi_2,c}$) such that $\alpha_n \to \alpha$. Define $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ to be such that $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1,2\}$. It follows that

$$\begin{aligned} \alpha_n\xi_1 + (1-\alpha_n)\xi_2 &= \langle \alpha_n\varphi_1 + (1-\alpha_n)\varphi_2, \cdot \rangle \qquad \forall n \in \mathbb{N} \\ \text{and } \alpha\xi_1 + (1-\alpha)\xi_2 &= \langle \alpha\varphi_1 + (1-\alpha)\varphi_2, \cdot \rangle. \end{aligned}$$

It is immediate to see that the sequence $\{\alpha_n\varphi_1 + (1-\alpha_n)\varphi_2\}_{n\in\mathbb{N}}$ converges uniformly to $\alpha\varphi_1 + (1-\alpha)\varphi_2$. By (39) and since I is lower (resp., upper) semicontinuous on $B(\Omega, \mathcal{G})$, this implies that

$$\begin{aligned} \breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) &= I(\alpha\varphi_1 + (1-\alpha)\varphi_2) \leq \liminf_n I(\alpha_n\varphi_1 + (1-\alpha_n)\varphi_2) \\ &= \liminf_n \breve{I}(\alpha_n\xi_1 + (1-\alpha_n)\xi_2) \leq c \\ &(resp., \\ \breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) &= I(\alpha\varphi_1 + (1-\alpha)\varphi_2) \geq \limsup_n I(\alpha_n\varphi_1 + (1-\alpha_n)\varphi_2) \\ &= \limsup_n \breve{I}(\alpha_n\xi_1 + (1-\alpha_n)\xi_2) \geq c), \end{aligned}$$

proving that $L_{\xi_1,\xi_2,c}$ (resp., $U_{\xi_1,\xi_2,c}$) is closed. Since ξ_1, ξ_2, and c were arbitrarily chosen, the statement follows. □

We now prove sufficiency. By the previous claim, if I is lower (resp., upper) semicontinuous then the set $L_{\xi_1,\xi_2,c}$ (resp., $U_{\xi_1,\xi_2,c}$) is closed for all $\xi_1, \xi_2 \in L$ and $c \in \mathbb{R}$. By [10, Lemma 46] and since $L = B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$ and $\breve{I}$ is monotone, this implies that $\breve{I}$ is lower (resp., upper) semicontinuous. Viceversa, consider $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq B(\Omega, \mathcal{G})$ and $\varphi \in B(\Omega, \mathcal{G})$ such that $\varphi_n \to \varphi$. Define $\xi \in L$ and $\{\xi_n\}_{n\in\mathbb{N}} \subseteq L$ such that

$\xi = \langle \varphi, \cdot \rangle$ and $\xi_n = \langle \varphi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. It is immediate to see that $\xi_n \to \xi$. By (39) and since $\breve{I}$ is lower (resp., upper) semicontinuous, this implies that

$$I(\varphi) = \breve{I}(\xi) \leq \liminf_n \breve{I}(\xi_n) = \liminf_n I(\varphi_n)$$
$$(resp.,$$
$$I(\varphi) = \breve{I}(\xi) \geq \limsup_n \breve{I}(\xi_n) = \limsup_n I(\varphi_n)),$$

proving necessity.

(3) $\breve{I}$ is concave if and only if I is concave.

Proof.

We first prove sufficiency. Pick $\xi_1, \xi_2 \in L$ and $\alpha \in (0,1)$. By assumption, there exist $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ such that $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1,2\}$. It follows that $\alpha\xi_1 + (1-\alpha)\xi_2 = \langle \alpha\varphi_1 + (1-\alpha)\varphi_2, \cdot \rangle$. By (39) and since I is concave, we have that

$$\breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) = I(\alpha\varphi_1 + (1-\alpha)\varphi_2) \geq \alpha I(\varphi_1) + (1-\alpha) I(\varphi_2) = \alpha\breve{I}(\xi_1) + (1-\alpha)\breve{I}(\xi_2),$$

proving that $\breve{I}$ is concave. Viceversa, pick $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ and $\alpha \in (0,1)$. Define $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1,2\}$. It follows that $\alpha\xi_1 + (1-\alpha)\xi_2 = \langle \alpha\varphi_1 + (1-\alpha)\varphi_2, \cdot \rangle$. By (39) and since $\breve{I}$ is concave, we have that

$$I(\alpha\varphi_1 + (1-\alpha)\varphi_2) = \breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) \geq \alpha\breve{I}(\xi_1) + (1-\alpha)\breve{I}(\xi_2) = \alpha I(\varphi_1) + (1-\alpha) I(\varphi_2),$$

proving that I is concave.

(4) $\breve{I}$ is quasiconcave if and only if I is quasiconcave.

Proof.

We first prove sufficiency. Pick $\xi_1, \xi_2 \in L$ and $\alpha \in (0,1)$. By assumption, there exist $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ such that $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1,2\}$. It follows that $\alpha\xi_1 + (1-\alpha)\xi_2 = \langle \alpha\varphi_1 + (1-\alpha)\varphi_2, \cdot \rangle$. By (39) and since I is quasiconcave, we have that

$$\breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) = I(\alpha\varphi_1 + (1-\alpha)\varphi_2) \geq \min\{I(\varphi_1), I(\varphi_2)\} = \min\left\{\breve{I}(\xi_1), \breve{I}(\xi_2)\right\},$$

proving that $\breve{I}$ is quasiconcave. Viceversa, pick $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ and $\alpha \in (0,1)$. Define $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1,2\}$. It follows that $\alpha\xi_1 + (1-\alpha)\xi_2 = \langle \alpha\varphi_1 + (1-\alpha)\varphi_2, \cdot \rangle$. By (39) and since $\breve{I}$ is quasiconcave, we have that

$$I(\alpha\varphi_1 + (1-\alpha)\varphi_2) = \breve{I}(\alpha\xi_1 + (1-\alpha)\xi_2) \geq \min\left\{\breve{I}(\xi_1), \breve{I}(\xi_2)\right\} = \min\{I(\varphi_1), I(\varphi_2)\},$$

proving that I is quasiconcave.

(5) $\breve{I}$ is supermodular if and only if I is supermodular.

Proof.

We first prove sufficiency. Pick $\xi_1, \xi_2 \in L$. By assumption, there exist $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ such that $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1,2\}$. By (33), each element P in $\mathcal{S}(\mathcal{P})$ induces a lattice homomorphism on $B(\Omega, \mathcal{G})$. It follows that $\xi_1 \vee \xi_2 = \langle \varphi_1 \vee \varphi_2, \cdot \rangle$ and that $\xi_1 \wedge \xi_2 = \langle \varphi_1 \wedge \varphi_2, \cdot \rangle$. By (39) and since I is supermodular, we have that

$$\breve{I}(\xi_1 \vee \xi_2) + \breve{I}(\xi_1 \wedge \xi_2) = I(\varphi_1 \vee \varphi_2) + I(\varphi_1 \wedge \varphi_2) \geq I(\varphi_1) + I(\varphi_2) = \breve{I}(\xi_1) + \breve{I}(\xi_2),$$

proving that $\breve{I}$ is supermodular. Viceversa, pick $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$. Define $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1,2\}$. By the same argument used before, it follows that $\xi_1 \vee \xi_2 = \langle \varphi_1 \vee \varphi_2, \cdot \rangle$ and that $\xi_1 \wedge \xi_2 = \langle \varphi_1 \wedge \varphi_2, \cdot \rangle$. By (39) and since $\breve{I}$ is supermodular, we have that

$$I(\varphi_1 \vee \varphi_2) + I(\varphi_1 \wedge \varphi_2) = \breve{I}(\xi_1 \vee \xi_2) + \breve{I}(\xi_1 \wedge \xi_2) \geq \breve{I}(\xi_1) + \breve{I}(\xi_2) = I(\varphi_1) + I(\varphi_2),$$

proving that I is supermodular.

(6) $\breve{I}$ is comonotonic additive if and only if I is comonotonic additive;

Proof of the Step.

We first prove sufficiency. Consider $\xi_1, \xi_2 \in L$ comonotonic. By point (h) of Lemma 24, there exist $\varphi_1, \varphi_2 \in B(\Omega, \mathcal{G})$ such that φ_1 and φ_2 are comonotonic and $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i = 1, 2$. It is immediate to see that $\xi_1 + \xi_2 = \langle \varphi_1 + \varphi_2, \cdot \rangle$. By (39) and since I is comonotonic additive, this implies that

$$\breve{I}(\xi_1 + \xi_2) = I(\varphi_1 + \varphi_2) = I(\varphi_1) + I(\varphi_2) = \breve{I}(\xi_1) + \breve{I}(\xi_2),$$

proving that $\breve{I}$ is comonotonic additive. Viceversa, pick φ_1 and φ_2 in $B(\Omega, \mathcal{G})$ that are further comonotonic. Define $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1, 2\}$. By point (h) of Lemma 24, it follows that ξ_1 and ξ_2 are comonotonic. Moreover, we have that $\xi_1 + \xi_2 = \langle \varphi_1 + \varphi_2, \cdot \rangle$. By (39) and since $\breve{I}$ is comonotonic additive, we have that

$$I(\varphi_1 + \varphi_2) = \breve{I}(\xi_1 + \xi_2) = \breve{I}(\xi_1) + \breve{I}(\xi_2) = I(\varphi_1) + I(\varphi_2),$$

proving that I is comonotonic additive.

(7) If I is quasiconcave and lower semicontinuous, $\breve{I}$ is inner continuous on L_0 if and only if I is inner continuous on $B_0(\Omega, \mathcal{G})$.

Proof.

We first prove sufficiency. Consider $\xi_1, \xi_2 \in L_0$ such that $\breve{I}(\xi_1) > \breve{I}(\xi_2)$, $\{C_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}_{\mathcal{S}(\mathcal{P})}$ such that $C_n \downarrow \emptyset$, and $k \in \mathbb{R}$. By definition of L_0, it follows that there exist $\varphi_1, \varphi_2 \in B_0(\Omega, \mathcal{G})$ such that $\xi_i = \langle \varphi_i, \cdot \rangle$ for $i \in \{1, 2\}$. By point (n) of Lemma 24, there exists a sequence $\{E_n\}_{n \in \mathbb{N}} \subseteq \mathcal{G}$ such that $E_n \downarrow \emptyset$ and $1_{C_n} = \langle 1_{E_n}, \cdot \rangle$ for all $n \in \mathbb{N}$. It is immediate to see that

$$k 1_{C_n} + \xi_1 1_{C_n^c} = \left\langle k 1_{E_n} + \varphi_1 1_{E_n^c}, \cdot \right\rangle \qquad \forall n \in \mathbb{N}.$$ [39]

By (39), we have that $I(\varphi_1) = \breve{I}(\xi_1) > \breve{I}(\xi_2) = I(\varphi_2)$. By [10, Theorem 54], (39), and since I is inner continuous on $B_0(\Omega, \mathcal{G})$, there exists $n \in \mathbb{N}$ such that

$$\breve{I}\left(k 1_{C_n} + \xi_1 1_{C_n^c}\right) = I\left(k 1_{E_n} + \varphi_1 1_{E_n^c}\right) > I(\varphi_2) = \breve{I}(\xi_2).$$

By point (2) and (4), it follows that $\breve{I}$ is lower semicontinuous and quasiconcave, other than being monotone and normalized. By point (m) of Lemma 24 and [10, Theorem 54], this implies that $\breve{I}$ is inner continuous. Viceversa, pick $\{\varphi_n\}_{n \in \mathbb{N}} \subseteq B_0(\Omega, \mathcal{G})$ such that $\varphi_n \uparrow \varphi \in B_0(\Omega, \mathcal{G})$. Define $\xi = \langle \varphi, \cdot \rangle$ and $\xi_n = \langle \varphi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. By point (e) of Lemma 24, it follows that $\xi_n \uparrow \xi$ where $\{\xi_n\}_{n \in \mathbb{N}} \subseteq L_0$ and $\xi \in L_0$. By (39) and since $\breve{I}$ is inner continuous on L_0, we have that

$$\lim_n I(\varphi_n) = \lim_n \breve{I}(\xi_n) = \breve{I}(\xi) = I(\varphi),$$

proving that I is inner continuous.

---

[39] Observe that for each $n \in \mathbb{N}$

$$\xi_1 = \langle \varphi_1, \cdot \rangle = \langle \varphi_1 1_{E_n}, \cdot \rangle + \left\langle \varphi_1 1_{E_n^c}, \cdot \right\rangle.$$

If $P \notin C_n$ then $0 = 1_{C_n}(P) = P(E_n)$. Thus, $\left(k 1_{C_n} + \xi_1 1_{C_n^c}\right)(P) = \xi_1(P) 1_{C_n^c}(P)$. It follows that

$$\langle k 1_{E_n}, P \rangle = 0 \text{ and } \left\langle k 1_{E_n} + \varphi_1 1_{E_n^c}, P \right\rangle = \left\langle \varphi_1 1_{E_n^c}, P \right\rangle = \langle \varphi_1, P \rangle = \xi_1(P) = \xi_1(P) 1_{C_n^c}(P).$$

Viceversa, if $P \in C_n$ then $1 = 1_{C_n}(P) = P(E_n)$. Thus, $\left(k 1_{C_n} + \xi_1 1_{C_n^c}\right)(P) = k$. It follows that

$$\left\langle k 1_{E_n} + \varphi_1 1_{E_n^c}, P \right\rangle = k.$$

(8) $\check{I}$ is inner/outer continuous if and only if I is inner/outer continuous.

Proof.

Consider $\{\xi_n\}_{n\in\mathbb{N}} \subseteq L$ such that $\xi_n \uparrow \xi$ (resp., $\xi_n \downarrow \xi$). By point (e) of Lemma 24, it follows that there exists a sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ and φ in $B(\Omega, \mathcal{G})$ such that $\varphi_n \uparrow \varphi$ (resp., $\varphi_n \downarrow \varphi$) and such that $\xi = \langle \varphi, \cdot \rangle$ and $\xi_n = \langle \varphi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. By (39) and since I is inner (resp., outer) continuous, we have that

$$\lim_n \check{I}(\xi_n) = \lim_n I(\varphi_n) = I(\varphi) = \check{I}(\xi),$$

proving that $\check{I}$ is inner (resp., outer) continuous. Viceversa, pick $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq B(\Omega, \mathcal{G})$ such that $\varphi_n \uparrow \varphi \in B(\Omega, \mathcal{G})$ (resp., $\varphi_n \downarrow \varphi$). Define $\xi = \langle \varphi, \cdot \rangle$ and $\xi_n = \langle \varphi_n, \cdot \rangle$ for all $n \in \mathbb{N}$. By point (e) of Lemma 24, it follows that $\xi_n \uparrow \xi$ (resp., $\xi_n \downarrow \xi$) where $\{\xi_n\}_{n\in\mathbb{N}} \subseteq L$ and $\xi \in L$. By (39) and since $\check{I}$ is inner (resp., outer) continuous, we have that

$$\lim_n I(\varphi_n) = \lim_n \check{I}(\xi_n) = \check{I}(\xi) = I(\varphi),$$

proving that I is inner (resp., outer) continuous. ■

Let $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ be a normalized, monotone, and supnorm continuous functional. Define the binary relation $\succsim_I$ on $B(\Omega, \mathcal{F})$ by

$$\varphi \succsim_I \psi \Leftrightarrow I(\lambda\varphi + (1-\lambda)\eta) \geq I(\lambda\psi + (1-\lambda)\eta) \qquad \forall \lambda \in [0,1], \forall \eta \in B(\Omega, \mathcal{F}).$$

By the same arguments contained in [24] or [25], there exists a unique closed and convex subset $\mathcal{C}(I) \subseteq \Delta(\Omega, \mathcal{F})$ such that

$$\varphi \succsim_I \psi \Leftrightarrow \int_\Omega \varphi dP \geq \int_\Omega \psi dP \qquad \forall P \in \mathcal{C}(I).$$

Similarly, given a set $\mathcal{C} \subseteq \Delta(\Omega, \mathcal{F})$, we define the binary relation $\succsim_{\mathcal{C}}$ on $B(\Omega, \mathcal{F})$ by

$$\varphi \succsim_{\mathcal{C}} \psi \Leftrightarrow \int_\Omega \varphi dP \geq \int_\Omega \psi dP \qquad \forall P \in \mathcal{C}.$$

Observe that $\succsim_{\mathcal{C}}$ is a conic, monotone, and continuous preorder.

Lemma 26 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space. If $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ is normalized, monotone, supnorm continuous, and such that $\mathcal{C}(I) = cl(co(\mathcal{C}))$ where $\mathcal{C} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$ then the following conditions are equivalent:*

(i) $\mathcal{C} \subseteq \mathrm{mco}(\mathcal{P})$;

(ii) $\varphi \succsim_{\mathcal{P}} \psi$ implies $\varphi \succsim_{\mathcal{C}} \psi$

(iii) $\mathcal{C} \subseteq \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$ and $\varphi \sim_{\mathcal{C}} \varphi^\star$ for all $\varphi \in B(\Omega, \mathcal{F})$;

(iv) $\mathcal{C} \subseteq \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$ and if $\varphi, \psi \in B(\Omega, \mathcal{F})$ then $\varphi^\star \geq \psi^\star$ $\mathcal{P}$-a.s. implies $\varphi \succsim_{\mathcal{C}} \psi$;

(v) $\varphi \succsim_{\mathcal{S}(\mathcal{P})} \psi$ implies $\varphi \succsim_{\mathcal{C}} \psi$;

(vi) $\varphi \succsim_{\mathcal{S}(\mathcal{P})} \psi$ implies $I(\varphi) \geq I(\psi)$;

(vii) $\varphi \succsim_{\mathcal{P}} \psi$ implies $I(\varphi) \geq I(\psi)$.

Proof. Before starting observe that, since $\mathcal{C}(I) = cl(co(\mathcal{C}))$,

$$\varphi \succsim_{\mathcal{C}(I)} \psi \Leftrightarrow \varphi \succsim_{\mathcal{C}} \psi \qquad \forall \varphi, \psi \in B(\Omega, \mathcal{F}).$$

(i) implies (v). Let $\varphi, \psi \in B(\Omega, \mathcal{F})$. By definition of $\succsim_{\mathcal{S}(\mathcal{P})}$, if $\varphi \succsim_{\mathcal{S}(\mathcal{P})} \psi$ then

$$\int_\Omega \varphi dP \geq \int_\Omega \psi dP \qquad \forall P \in \mathcal{S}(\mathcal{P}).$$

This implies that for each $\mu \in \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$

$$\bar{\mu}(\varphi) = \int_{\mathcal{S}(\mathcal{P})} \langle \varphi, P \rangle \, d\mu(P) \geq \int_{\mathcal{S}(\mathcal{P})} \langle \psi, P \rangle \, d\mu(P) = \bar{\mu}(\psi). \tag{43}$$

By Proposition 18, we have that $\text{mco}(\mathcal{P}) = \left\{\bar{\mu} : \mu \in \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)\right\}$. By (43), this implies that

$$\int_\Omega \varphi dP \geq \int_\Omega \psi dP \qquad \forall P \in \text{mco}(\mathcal{P}).$$

By (i), we have that $\mathcal{C} \subseteq \text{mco}(\mathcal{P})$, thus, $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{C}$, proving the statement.

(v) implies (ii). Since $\mathcal{S}(\mathcal{P}) \subseteq \mathcal{P}$, the statement is obvious.

(ii) implies (iii). For each $\varphi \in B(\Omega, \mathcal{F})$ we have that $\int_\Omega \varphi dP = \int_\Omega \varphi^\star dP$ for all $P \in \mathcal{P}$. By definition of $\succsim_{\mathcal{P}}$, this is equivalent to say that $\varphi \sim_{\mathcal{P}} \varphi^\star$. By (ii), this implies that $\varphi \sim_{\mathcal{C}} \varphi^\star$. Next, observe that if $P(A) = 0$ for all $P \in \mathcal{P}$ then $1_A \sim_{\mathcal{P}} 0$. By (ii), it follows that $1_A \sim_{\mathcal{C}} 0$. By definition of $\succsim_{\mathcal{C}}$, we can conclude that $\mathcal{C} \subseteq \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$.

(iii) implies (iv). By assumption, we have that $\mathcal{C} \subseteq \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$. Next, since $\mathcal{C} \subseteq \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$, if $\varphi^\star \geq \psi^\star$ $\mathcal{P}$-a.s. then $\varphi^\star \succsim_{\mathcal{C}} \psi^\star$. By (iii) and since $\succsim_{\mathcal{C}}$ is a preorder, it follows that

$$\varphi \sim_{\mathcal{C}} \varphi^\star \succsim_{\mathcal{C}} \psi^\star \sim_{\mathcal{C}} \psi,$$

that is, $\varphi \succsim_{\mathcal{C}} \psi$.

(iv) implies (i). If $\varphi \succsim_{\text{mco}(\mathcal{P})} \psi$ then

$$\int_\Omega \varphi dP \geq \int_\Omega \psi dP \qquad \forall P \in \text{mco}(\mathcal{P}). \tag{44}$$

Since $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space, consider W where $W \in \mathcal{F}$ is such that $P(W) = 1$ for all $P \in \mathcal{P}$ and $p(\cdot, \omega) \in \mathcal{P}$ for all $\omega \in W$. Since $\mathcal{P} \subseteq \text{mco}(\mathcal{P})$ and by (44), it follows that if $\omega \in W$ then

$$\int_\Omega \varphi dp^\omega \geq \int_\Omega \psi dp^\omega,$$

that is, $\varphi^\star \geq \psi^\star$ $\mathcal{P}$-a.s.. By assumption, this implies that $\varphi \succsim_{\mathcal{C}} \psi$. By [24, Proposition A.1.] and since $\text{mco}(\mathcal{P})$ is convex, it follows that $\mathcal{C} \subseteq cl(co(\mathcal{C})) \subseteq cl(\text{mco}(\mathcal{P}))$. By assumption $\mathcal{C} \subseteq \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$, this implies that

$$\mathcal{C} \subseteq cl(\text{mco}(\mathcal{P})) \cap \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P}).$$

We conclude the proof by showing that $cl(\text{mco}(\mathcal{P})) \cap \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P}) \subseteq \text{mco}(\mathcal{P})$. If P belongs to $cl(\text{mco}(\mathcal{P})) \cap \Delta^\sigma(\Omega, \mathcal{F}, \mathcal{P})$ then there exists a net $\{P_\alpha\}_{\alpha \in A} \in \text{mco}(\mathcal{P})$ such that $\langle \varphi, P_\alpha \rangle \to \langle \varphi, P \rangle$ for all $\varphi \in B(\Omega, \mathcal{F})$. By Proposition 18, we have that p is a r.c.p. for $\text{mco}(\mathcal{P})$. For each $A \in \mathcal{F}$ and each $B \in \mathcal{G}$, we have that

$$\begin{aligned}\int_B p(A, \cdot) \, dP &= \int_\Omega p(A, \cdot) \, 1_B(\cdot) \, dP = \lim_\alpha \int_\Omega p(A, \cdot) \, 1_B(\cdot) \, dP_\alpha \\ &= \lim_\alpha P_\alpha(A \cap B) = P(A \cap B).\end{aligned}$$

Since $P \in \Delta^{\sigma}(\Omega, \mathcal{F}, \mathcal{P})$, it follows that p is a r.c.p. for P. By Proposition 18, we conclude that $P \in \text{mco}(\mathcal{P})$.

We thus have proved the equivalence between points (i), (ii), (iii), (iv), and (v).

(v) implies (vi). Since $\mathcal{C}(I) = cl(co(\mathcal{C}))$, the statement follows immediately.

(vi) implies (vii). Since $\mathcal{S}(\mathcal{P}) \subseteq \mathcal{P}$, clearly the statement follows.

(vii) implies (v). By definition of $\succsim_{\mathcal{P}}$, we have that if $\varphi \succsim_{\mathcal{P}} \psi$ then

$$\lambda\varphi + (1-\lambda)\eta \succsim_{\mathcal{P}} \lambda\psi + (1-\lambda)\eta \qquad \forall \lambda \in [0,1], \forall \eta \in B(\Omega, \mathcal{F}). \tag{45}$$

By (vii) and (45), it follows that if $\varphi \succsim_{\mathcal{P}} \psi$ then $I(\lambda\varphi + (1-\lambda)\eta) \geq I(\lambda\psi + (1-\lambda)\eta)$ for all $\lambda \in [0,1]$ and $\eta \in B(\Omega, \mathcal{F})$, that is, $\varphi \succsim_{\mathcal{C}(I)} \psi$ and so $\varphi \succsim_{\mathcal{C}} \psi$. In other words, (vii) implies (ii). Given the previous part of the proof, we have that (ii) implies (v), proving the statement. ■

Lemma 27 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and let $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ be a normalized, monotone, quasi-concave, and lower semicontinuous functional. If $G \in \mathcal{L}(\mathbb{R} \times \Delta(\Omega, \mathcal{F}))$ is such that*

$$I(\varphi) = \min_{P \in \Delta(\Omega, \mathcal{F})} G\left(\int_{\Omega} \varphi dP, P\right) \qquad \forall \varphi \in B(\Omega, \mathcal{F})$$

and $\mathcal{C} = \text{dom}_{\Delta} G$ then the following conditions are equivalent:

(i) $\varphi \succsim_{\mathcal{S}(\mathcal{P})} \psi$ implies $\varphi \succsim_{\mathcal{C}} \psi$;

(ii) $\varphi \succsim_{\mathcal{S}(\mathcal{P})} \psi$ implies $I(\varphi) \geq I(\psi)$;

(iii) $\varphi \succsim_{\mathcal{P}} \psi$ implies $I(\varphi) \geq I(\psi)$;

(iv) if $\varphi, \psi \in B_0(\Omega, \mathcal{F})$, $\varphi \succsim_{\mathcal{P}} \psi$ implies $I(\varphi) \geq I(\psi)$.

Proof. First, observe that if $\varphi \succsim_{\mathcal{C}} \psi$ then $I(\varphi) \geq I(\psi)$. Indeed, consider $\varphi, \psi \in B(\Omega, \mathcal{F})$ such that $\varphi \succsim_{\mathcal{C}} \psi$. Since G is increasing in the first component and by definition of $\succsim_{\mathcal{C}}$, it follows that

$$G\left(\int_{\Omega} \varphi dP, P\right) \geq G\left(\int_{\Omega} \psi dP, P\right) \qquad \forall P \in \mathcal{C}.$$

By definition of $\text{dom}_{\Delta} G$, this implies that

$$I(\varphi) = \min_{P \in \mathcal{C}} G\left(\int_{\Omega} \varphi dP, P\right) \geq \min_{P \in \mathcal{C}} G\left(\int_{\Omega} \psi dP, P\right) = I(\psi).$$

Finally, by Theorem 21 (see also [11]), we have that

$$G(t, P) = \sup\left\{I(\varphi) : \int_{\Omega} \varphi dP \leq t\right\} \qquad \forall (t, P) \in \mathbb{R} \times \Delta(\Omega, \mathcal{F}).$$

(i) implies (ii). Since if $\varphi \succsim_{\mathcal{C}} \psi$ then $I(\varphi) \geq I(\psi)$, the statement follows immediately.

(ii) implies (iii). Since $\mathcal{S}(\mathcal{P}) \subseteq \mathcal{P}$, clearly, the statement follows.

(iii) implies (iv). Since $B_0(\Omega, \mathcal{F}) \subseteq B(\Omega, \mathcal{F})$, the statement is obvious.

(iv) implies (i). We first show that $\Delta(\Omega, \mathcal{F}) \backslash cl(\text{mco}(\mathcal{P})) \subseteq \Delta(\Omega, \mathcal{F}) \backslash \text{dom}_{\Delta} G$. Since $\text{mco}(\mathcal{P})$ is convex, it follows that $cl(\text{mco}(\mathcal{P}))$ is convex and closed. If $\Delta(\Omega, \mathcal{F}) \backslash cl(\text{mco}(\mathcal{P})) = \emptyset$ then the inclusion is true.

Otherwise, pick $\bar{P} \in \Delta(\Omega, \mathcal{F}) \setminus cl(\text{mco}(\mathcal{P}))$. By an usual separation argument, we have that there exist $\bar{\varphi} \in B_0(\Omega, \mathcal{F})$ and $\varepsilon > 0$ such that

$$\int_\Omega \bar{\varphi} d\bar{P} < 0 < \varepsilon \leq \int_\Omega \bar{\varphi} dP \qquad \forall P \in \text{mco}(\mathcal{P}).$$

This implies that for each $s, t \in \mathbb{R}$ there exists $\bar{n} \in \mathbb{N}$ such that

$$\int_\Omega \bar{n}\bar{\varphi} d\bar{P} \leq t \text{ and } \int_\Omega s 1_\Omega dP = s \leq \int_\Omega \bar{n}\bar{\varphi} dP \qquad \forall P \in \text{mco}(\mathcal{P}).$$

Since I is normalized, $\mathcal{P} \subseteq \text{mco}(\mathcal{P})$, $\bar{n}\bar{\varphi}, s1_\Omega \in B_0(\Omega, \mathcal{F})$, and by (iv), we have that

$$\int_\Omega \bar{n}\bar{\varphi} d\bar{P} \leq t \text{ and } s \leq I(\bar{n}\bar{\varphi}).$$

Since $G(t, \bar{P}) = \sup\{I(\varphi) : \int_\Omega \varphi d\bar{P} \leq t\}$, this implies that $G(t, \bar{P}) \geq s$ for all $s, t \in \mathbb{R}$. We can conclude that $G(t, \bar{P}) = \infty$ for all $t \in \mathbb{R}$. This proves that $\bar{P} \in \Delta(\Omega, \mathcal{F}) \setminus \text{dom}_\Delta G$. Since $\Delta(\Omega, \mathcal{F}) \setminus cl(\text{mco}(\mathcal{P})) \subseteq \Delta(\Omega, \mathcal{F}) \setminus \text{dom}_\Delta G$, we have that $\text{dom}_\Delta G = \mathcal{C} \subseteq cl(\text{mco}(\mathcal{P}))$. Next, consider $\varphi, \psi \in B(\Omega, \mathcal{F})$ such that $\varphi \succsim_{\mathcal{S}(\mathcal{P})} \psi$. By the same arguments contained in Step 1 of the proof of Proposition 25 and since $\mathcal{C} \subseteq cl(\text{mco}(\mathcal{P}))$, it follows that

$$\varphi \succsim_{\mathcal{S}(\mathcal{P})} \psi \Rightarrow \varphi \succsim_{\text{mco}(\mathcal{P})} \psi \Rightarrow \varphi \succsim_{cl(\text{mco}(\mathcal{P}))} \psi \Rightarrow \varphi \succsim_{\mathcal{C}} \psi,$$

proving the statement. ∎

E Proofs

In this appendix, we prove the main statements of the paper. Before starting, we introduce a new piece of notation. Given a Dynkin space $(\Omega, \mathcal{F}, \mathcal{P})$, for each $P \in \mathcal{P}$ and for each $h \in B_0(X)$ we use indifferently the notation $\int_\Omega h dP$ and P_h, that is,

$$P_h = \sum_{x \in X} P(\{\omega \in \Omega : f(\omega) = x\}) x = \int_\Omega h dP.$$

Lemma 28 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$. If $\succsim$ satisfies the Basic Conditions, Consistency, and Risk Independence then $\succsim$ satisfies Monotonicity.*

Proof. By [34] and since $\succsim$ satisfies the Basic Conditions and Risk Independence, there exists an affine function $u : X \to \mathbb{R}$ that represents $\succsim$ restricted to X. Consider $f, g \in B_0(X)$ such that $f(\omega) \succsim g(\omega)$ for all $\omega \in \Omega$. We have that $u(f), u(g) \in B_0(\Omega, \mathcal{F})$. Since u represents $\succsim$ over X, it follows that $u(f) \geq u(g)$. Next, recall that for each $P \in \mathcal{P}$

$$P_f = \sum_{x \in X} P(\{\omega \in \Omega : f(\omega) = x\}) x.$$

Since u is affine and f takes just finitely many values, this implies that for each $P \in \mathcal{P}$

$$\begin{aligned} u(P_f) &= u\left(\sum_{x \in X} P(\{\omega \in \Omega : f(\omega) = x\}) x\right) \\ &= \sum_{x \in X} P(\{\omega \in \Omega : f(\omega) = x\}) u(x) \\ &= \int_\Omega u(f) dP. \end{aligned}$$

Similarly, we have that $u(P_g) = \int_\Omega u(g)\,dP$ for all $P \in \mathcal{P}$. Since $u(f) \geq u(g)$ and $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$, it follows that

$$u(P_f) = \int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP = u(P_g) \qquad \forall P \in \mathcal{P}.$$

Since u represents $\succsim$ on X, this implies that $P_f \succsim P_g$ for all $P \in \mathcal{P}$. Since $\succsim$ satisfies Consistency, we can conclude that $f \succsim g$, proving the statement. ■

Remark 29 *If $(\Omega, \mathcal{F}, \mathcal{P})$ is a Dynkin space and $\succsim$ a binary relation on $B_0(X)$ that satisfies the Basic Conditions, Consistency, and Risk Independence then it follows that there exist x and y in X such that $x \succ y$. Moreover, observe that, in the previous proof, we could dispense with completeness of $\succsim$ on $B_0(X)$ and just require completeness of $\succsim$ over constant acts. In other words, if $\succsim$ satisfies the Weak Basic Conditions, Consistency, and Risk Independence then it satisfies Monotonicity and, in particular, it is reflexive.*

Lemma 30 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$ that satisfies the Basic Conditions, Consistency, and Risk Independence. If $u : X \to \mathbb{R}$, $I : B_0(\Omega, \mathcal{F}, u(X)) \to \mathbb{R}$, and $J : B(\Omega, \mathcal{G}, u(X)) \to \mathbb{R}$ are such that*

1. *u is nonconstant and affine;*
2. *I is normalized, monotone, and continuous;*
3. *J is normalized, monotone, and continuous;*
4. *if $\varphi, \psi \in B_0(\Omega, \mathcal{F}, u(X))$ then $\int_\Omega \varphi dP \geq \int \psi dP$ for all $P \in \mathcal{P}$ implies $I(\varphi) \geq I(\psi)$;*
5. *if $f, g \in B_0(X)$ then $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$;*
6. *if $f, g \in B_0(X)$ and f and g are $\mathcal{G}$-measurable then $f \succsim g$ if and only if $J(u(f)) \geq J(u(g))$;*

then, $I(\varphi) = J(\varphi)$ for all $\varphi \in B_0(\Omega, \mathcal{G}, u(X))$ and $I(\varphi) = J(\varphi^\star)$ for all $\varphi \in B_0(\Omega, \mathcal{F}, u(X))$.

Proof. Since u is nonconstant and affine, notice that $u(X)$ is an interval with nonempty interior. By Lemma 28 and since $\succsim$ satisfies the Basic Conditions, Consistency, and Risk Independence, it follows that $\succsim$ satisfies Monotonicity. This implies that for each $f \in B_0(X)$ there exists $x_f \in X$ such that $f \sim x_f$. If $\varphi \in B_0(\Omega, \mathcal{G}, u(X))$ then there exists f in $B_0(X)$, which is further $\mathcal{G}$-measurable, such that $\varphi = u(f)$. Since I and J are normalized and, once composed with u, they represent $\succsim$, it follows that

$$I(\varphi) = I(u(f)) = I(u(x_f) 1_\Omega) = u(x_f) = J(u(x_f) 1_\Omega) = J(u(f)) = J(\varphi),$$

proving the first part of the statement. If $\varphi \in B_0(\Omega, \mathcal{F}, u(X))$ then we have that $\varphi^\star \in B(\Omega, \mathcal{G}, u(X))$ and that there exists two sequences, $\{\varphi_n\}_{n\in\mathbb{N}}$ and $\{\psi_n\}_{n\in\mathbb{N}}$, in $B_0(\Omega, \mathcal{G}, u(X))$ such that $\varphi_n \downarrow \varphi^\star$ and $\psi_n \uparrow \varphi^\star$ where the convergence is uniform. Moreover, since $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$, observe that for each $n \in \mathbb{N}$

$$\int_\Omega \varphi_n dP \geq \int_\Omega \varphi^\star dP = \int_\Omega \varphi dP = \int_\Omega \varphi^\star dP \geq \int_\Omega \psi_n dP \qquad \forall P \in \mathcal{P}.$$

By the previous part of the proof and point 4., this implies that

$$J(\varphi_n) = I(\varphi_n) \geq I(\varphi) \geq I(\psi_n) = J(\psi_n) \qquad \forall n \in \mathbb{N}.$$

Since J is continuous, passing to the limit, we obtain that

$$J(\varphi^\star) \geq I(\varphi) \geq J(\varphi^\star),$$

proving the last part of the statement. ■

Lemma 31 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$. $\succsim$ satisfies the Basic Conditions, Consistency, and Comonotonic Independence on $\mathcal{G}$-measurable acts if and only if there exist a nonconstant affine function $u : X \to \mathbb{R}$ and a normalized, monotone, translation invariant, and positively homogeneous functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$, which is further comonotonic additive on $B(\Omega, \mathcal{G})$ and such that*

(a) $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ implies $I(\varphi) \geq I(\psi)$;

(b) $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$.

Moreover,

1. *u is unique up to an affine transformation and I is unique.*
2. *I is quasiconcave if and only if $\succsim$ satisfies Uncertainty Aversion.*
3. *I is inner and outer continuous on $B(\Omega, \mathcal{G})$ if and only if $\succsim$ satisfies Monotone Continuity.*

Proof. We first prove necessity. Since $\succsim$ satisfies the Basic Conditions and Comonotonic Independence on $\mathcal{G}$-measurable acts, $\succsim$ satisfies Risk Independence *à la* Herstein and Milnor. By [34], it follows that there exists an affine function $u : X \to \mathbb{R}$ that represents $\succsim$ restricted to X. This implies that $\succsim$ satisfies Risk Independence and, in particular, that u is nonconstant. Moreover, u is cardinally unique. Since u is cardinally unique, without loss of generality, we assume that $u(X) \supseteq [-1, 1]$. Since $\succsim$ further satisfies Consistency and by Lemma 28, $\succsim$ satisfies Monotonicity. By [9] and since $\succsim$ satisfies the Basic Conditions, Monotonicity, and Risk Independence, it follows that there exists a normalized, monotone, and continuous functional $\hat{I} : B_0(\Omega, \mathcal{F}, u(X)) \to \mathbb{R}$ such that for each f and g in $B_0(X)$ we have that

$$f \succsim g \Leftrightarrow \hat{I}(u(f)) \geq \hat{I}(u(g)). \tag{46}$$

If $\varphi, \psi \in B_0(\Omega, \mathcal{F}, u(X))$ are such that $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ then there exist f and g in $B_0(X)$ such that $\varphi = u(f), \psi = u(g)$, and $\int_\Omega u(f) dP \geq \int_\Omega u(g) dP$ for all $P \in \mathcal{P}$. By definition of P_f and P_g and since u is affine, it follows that

$$u(P_f) = \int_\Omega u(f) dP \geq \int_\Omega u(g) dP = u(P_g) \qquad \forall P \in \mathcal{P}.$$

Since u represents $\succsim$ over X, this implies that $P_f \succsim P_g$ for all $P \in \mathcal{P}$. Since $\succsim$ satisfies Consistency and $\hat{I}$ satisfies (46), it follows that

$$\hat{I}(\varphi) = \hat{I}(u(f)) \geq \hat{I}(u(g)) = \hat{I}(\psi).$$

Next, recall that $\succsim$, restricted to $\mathcal{G}$-measurable acts, satisfies the Basic Conditions, Monotonicity, and Comonotonic Independence. By [47], it follows that there exists a capacity $\rho : \mathcal{G} \to [0, 1]$ such that

$$f \succsim g \Leftrightarrow \int_\Omega u(f) d\rho \geq \int_\Omega u(g) d\rho, \tag{47}$$

where the integrals are in the Choquet sense. If we define $J : B(\Omega, \mathcal{G}) \to \mathbb{R}$ by

$$J(\varphi) = \int_\Omega \varphi d\rho \qquad \forall \varphi \in B(\Omega, \mathcal{G})$$

then J is a normalized, monotone, and continuous functional. Moreover, J is comonotonic additive and such that

$$f \succsim g \Leftrightarrow J(u(f)) d\rho \geq J(u(g)).$$

By Lemma 30, it follows that $\hat{I}(\varphi) = J(\varphi)$ for all $\varphi \in B_0(\Omega, \mathcal{G}, u(X))$ and $\hat{I}(\varphi) = J(\varphi^\star)$ for all $\varphi \in B_0(\Omega, \mathcal{F}, u(X))$. Define the functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ by

$$\varphi \mapsto \int_\Omega \varphi^\star d\rho = J(\varphi^\star).$$

Since I is the composition of J with the linear and continuous operator $^\star : B(\Omega, \mathcal{F}) \to B(\Omega, \mathcal{G})$, it follows that I is a normalized, monotone, translation invariant, and positively homogeneous functional. In particular, it follows that I is Lipschitz continuous. By the previous part of the proof and by definition of I, we have that for each $\varphi \in B_0(\Omega, \mathcal{G}, u(X))$

$$J(\varphi) = \hat{I}(\varphi) = J(\varphi^\star) = I(\varphi). \tag{48}$$

Since $u(X) \supseteq [-1, 1]$, if $\varphi \in B_0(\Omega, \mathcal{G})$ then there exist $\lambda \in (0, \infty)$ such that $\bar{\varphi} = \lambda\varphi \in B_0(\Omega, \mathcal{G}, u(X))$. By (48) and since both I and J are positively homogeneous, it follows that

$$\lambda J(\varphi) = J(\lambda\varphi) = J(\bar{\varphi}) = I(\bar{\varphi}) = I(\lambda\varphi) = \lambda I(\varphi) \Rightarrow J(\varphi) = I(\varphi),$$

proving that I coincides to J on $B_0(\Omega, \mathcal{G})$. Since both I and J are continuous on $B(\Omega, \mathcal{G})$, this implies that I and J coincide on $B(\Omega, \mathcal{G})$. Therefore, it follows that I is comonotonic additive on $B(\Omega, \mathcal{G})$. Since $u(X) \supseteq [-1, 1]$, if $\varphi, \psi \in B_0(\Omega, \mathcal{F})$ then there exist $\lambda \in (0, \infty)$ such that $\bar{\varphi} = \lambda\varphi$ and $\bar{\psi} = \lambda\psi$ are both elements of $B_0(\Omega, \mathcal{F}, u(X))$. If, furthermore, $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ then $\int_\Omega \bar{\varphi} dP \geq \int_\Omega \bar{\psi} dP$ for all $P \in \mathcal{P}$. By the previous part of the proof, it follows that

$$\begin{aligned}
\lambda I(\varphi) = I(\lambda\varphi) = I(\bar{\varphi}) &= J(\bar{\varphi}^\star) \\
&= \hat{I}(\bar{\varphi}) \\
&\geq \hat{I}(\bar{\psi}) \\
&= J(\bar{\psi}^\star) = I(\bar{\psi}) = I(\lambda\psi) = \lambda I(\psi).
\end{aligned}$$

It follows that $I(\varphi) \geq I(\psi)$. On the other hand, if $\varphi, \psi \in B(\Omega, \mathcal{F})$ are such that $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ then there exist two sequences, $\{\varphi_n\}_{n\in\mathbb{N}}$ and $\{\psi_n\}_{n\in\mathbb{N}}$, in $B_0(\Omega, \mathcal{F})$ such that $\varphi_n \downarrow \varphi$ and $\psi_n \uparrow \psi$ and the convergence is uniform. Since $\mathcal{P} \subseteq \Delta^\sigma(\Omega, \mathcal{F})$, it follows that $\int_\Omega \varphi_n dP \geq \int_\Omega \psi_n dP$ for all $P \in \mathcal{P}$ and for all $n \in \mathbb{N}$. By the previous part of the proof, it follows that $I(\varphi_n) \geq I(\psi_n)$ for all $n \in \mathbb{N}$. Since I is continuous and passing to the limit, it follows that $I(\varphi) \geq I(\psi)$, proving that I satisfies (a). Finally, by (46) and since $\hat{I}(\varphi) = J(\varphi^\star) = I(\varphi)$ for all $\varphi \in B_0(\Omega, \mathcal{F}, u(X))$, it follows that

$$f \succsim g \Leftrightarrow \hat{I}(u(f)) \geq \hat{I}(u(g)) \Leftrightarrow I(u(f)) \geq I(u(g)),$$

proving that I satisfies (b), and thus proving necessity. Sufficiency follows from routine arguments.

1. It follows from [24, Lemma 1] and the Lipschitz continuity of I.
2. By [10, Lemma 60], if $\succsim$ further satisfies Uncertainty Aversion then

$$f \succsim g \quad \Rightarrow \quad \alpha f + (1-\alpha) g \succsim g \qquad \forall \alpha \in (0, 1).$$

By the previous part of the proof, recall that $u(X) \supseteq [-1, 1]$. If $\varphi, \psi \in B_0(\Omega, \mathcal{F})$ and $\alpha \in (0, 1)$ then there exists $\lambda \in (0, \infty)$ such that $\bar{\varphi} = \lambda\varphi$ and $\bar{\psi} = \lambda\psi$ belong to $B_0(\Omega, \mathcal{F}, u(X))$. Thus, there exist f and g in $B_0(X)$ such that $\bar{\varphi} = u(f)$ and $\bar{\psi} = u(g)$. Without loss of generality, assume that $f \succsim g$. Since u is affine and I is positively homogeneous and satisfies (b), it follows that

$$\begin{aligned}
\lambda I(\alpha\varphi + (1-\alpha)\psi) = I(\alpha\bar{\varphi} + (1-\alpha)\bar{\psi}) \geq \min\{I(\bar{\varphi}), I(\bar{\psi})\} = \lambda \min\{I(\varphi), I(\psi)\} \\
\Rightarrow I(\alpha\varphi + (1-\alpha)\psi) \geq \min\{I(\varphi), I(\psi)\},
\end{aligned}$$

proving the quasiconcavity of I on $B_0(\Omega, \mathcal{F})$. By the continuity of I, it follows that I is quasiconcave on $B(\Omega, \mathcal{F})$, proving sufficiency. Necessity is obvious.

3. We next prove sufficiency. By [46] and since I is a normalized, monotone, and comonotonic additive functional on $B(\Omega, \mathcal{G})$, there exists a unique capacity $\rho : \mathcal{G} \to [0,1]$ such that $I(\varphi) = \int_\Omega \varphi d\rho$ for all $\varphi \in B(\Omega, \mathcal{G})$. Consider $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathcal{G}$ and $E \in \mathcal{G}$ such that $E_n \downarrow E$ (resp., $E_n \uparrow E$). It is immediate to see that $\lim_n \rho(E_n)$ is well defined, it belongs to $[0,1]$, and that $\lim_n \rho(E_n) \geq \rho(E)$ (resp., $\leq \rho(E)$). By contradiction, assume that $\lim_n \rho(E_n) > \rho(E)$ (resp., $< \rho(E)$). Since, without loss of generality, we can assume that $u(X) \supseteq [-1,1]$, consider $x, y, z \in X$ such that $u(x) = 1$, $u(z) = \lim_{n\in\mathbb{N}} \rho(E_n)$, and $u(y) = 0$. Moreover, define $F_n = E_n \backslash E$ (resp., $E \backslash E_n$) for all $n \in \mathbb{N}$, $f = z$ and $g = xEy$ (resp., $f = xEy$ and $g = z$). It is immediate to see that $u(g) = 1_E$ and $u(xF_n g) = 1_{E_n}$ (resp., $u(yF_n f) = 1_{E_n}$) for all $n \in \mathbb{N}$. By working hypothesis, this implies that $f \succ g$ (resp., $f \succ g$). Since I satisfies (b) and $\succsim$ satisfies Monotone Continuity, there exists $N \in \mathbb{N}$ such that $f \succ xF_N g$ (resp., $yF_N f \succ g$), that is,

$$\lim_n \rho(E_n) = u(z) = I(u(z) 1_\Omega) > I(u(xF_N g)) = I(1_{E_N}) = \rho(E_N) \geq \lim_n \rho(E_n)$$
$$(resp.,\ \lim_n \rho(E_n) = u(z) = I(u(z) 1_\Omega) < I(u(yF_N f)) = I(1_{E_N}) = \rho(E_N) \leq \lim_n \rho(E_n)),$$

a contradiction. It follows that ρ is continuous. It is well known (see, e.g., [13]) that the continuity of ρ implies that I is inner and outer continuous on $B(\Omega, \mathcal{G})$.

Necessity follows by observing that, since I satisfies point (a), inner and outer monotone continuity of I on $B(\Omega, \mathcal{G})$ implies inner and outer monotone continuity of I on $B(\Omega, \mathcal{F})$. By a routine argument, this latter fact implies that $\succsim$ satisfies Monotone Continuity. ■

Lemma 32 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$. $\succsim$ satisfies the Basic Conditions, Consistency, Weak Certainty Independence, Uncertainty Aversion, and Unboundedness if and only if there exist an unbounded affine function $u : X \to \mathbb{R}$ and a normalized, monotone, translation invariant, and concave functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ such that*

(a) $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ implies $I(\varphi) \geq I(\psi)$;

(b) $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$.

Moreover,

1. *u is unique up to an affine transformation and, given u, I is unique.*
2. *I is inner continuous if $\succsim$ satisfies Monotone Continuity.*

Proof. We first prove necessity. Since $\succsim$ satisfies the Basic Conditions and Weak Certainty Independence, $\succsim$ satisfies Risk Independence *à la* Herstein and Milnor. By [34], it follows that there exists an affine function $u : X \to \mathbb{R}$ that represents $\succsim$ restricted to X. This implies that $\succsim$ satisfies Risk Independence. Since $\succsim$ satisfies Unboundedness, it is easy to check that $u(X)$ is unbounded. Furthermore, u is unique up to an affine transformation. Since u is unique up to an affine transformation, without loss of generality, we can assume that u is such that $0 \in \operatorname{int}(u(X))$. Since $\succsim$ further satisfies Consistency and by Lemma 28, $\succsim$ satisfies Monotonicity. Since $\succsim$ further satisfies Uncertainty Aversion and by [40, Lemma 25 and Lemma 28] and the proof of [40, Theorem 3], it follows that there exists a normalized and concave niveloid $I : B_0(\Omega, \mathcal{F}, u(X)) \to \mathbb{R}$ such that $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$. By [40, pag. 1476], it follows that I has a unique extension to $B_0(\Omega, \mathcal{F})$ which is normalized and concave as well. With a small abuse of notation, we denote this extension by I. By [40, Lemma 25], it follows that $I : B_0(\Omega, \mathcal{F}) \to \mathbb{R}$ is a normalized, monotone, translation invariant, and concave functional.

Before proceeding with the proof, observe that

$$B_0(\Omega,\mathcal{F},u(X)) = \{u(f) : f \in B_0(X)\}. \tag{49}$$

Consider $\varphi,\psi \in B_0(\Omega,\mathcal{F})$. Assume that $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$. Since $u(X)$ is unbounded, there exists $k \in \mathbb{R}$ such that $\bar{\varphi} = \varphi + k1_\Omega$ and $\bar{\psi} = \psi + k1_\Omega$ belong to $B_0(\Omega,\mathcal{F},u(X))$. Since $\mathcal{P} \subseteq \Delta^\sigma(\Omega,\mathcal{F})$, we have that $\int_\Omega \bar{\varphi} dP \geq \int_\Omega \bar{\psi} dP$ for all $P \in \mathcal{P}$. By (49), it follows that there exist f and g in $B_0(X)$ such that $\bar{\varphi} = u(f)$ and $\bar{\psi} = u(g)$. This implies that $\int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP$ for all $P \in \mathcal{P}$. By definition of P_f and P_g and since u is affine, we have that

$$u(P_f) = \int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP = u(P_g) \qquad \forall P \in \mathcal{P}.$$

Since u represents $\succsim$ over X, this implies that $P_f \succsim P_g$ for all $P \in \mathcal{P}$. Since $\succsim$ satisfies Consistency and I is translation invariant and satisfies (b), it follows that

$$I(\varphi) + k = I(\bar{\varphi}) = I(u(f)) \geq I(u(g)) = I\left(\bar{\psi}\right) = I(\psi) + k \Rightarrow I(\varphi) \geq I(\psi). \tag{50}$$

Since I is normalized, monotone, and translation invariant, observe that I is Lipschitz continuous. This implies that I admits a unique continuous extension to $B(\Omega,\mathcal{F})$. It is routine to check that this extension is itself a normalized, monotone, translation invariant, and concave functional over $B(\Omega,\mathcal{F})$. Again with a small abuse of notation, we denote this extension by I. Moreover, I clearly satisfies (b). Finally, consider $\varphi,\psi \in B(\Omega,\mathcal{F})$ such that $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$. It follows that there exist two sequences, $\{\varphi_n\}_{n\in\mathbb{N}}$ and $\{\psi_n\}_{n\in\mathbb{N}}$, in $B_0(\Omega,\mathcal{F})$ such that $\varphi_n \downarrow \varphi$ and $\psi_n \uparrow \psi$ and the convergence is uniform. Since $\mathcal{P} \subseteq \Delta^\sigma(\Omega,\mathcal{F})$, we have that $\int_\Omega \varphi_n dP \geq \int_\Omega \psi_n dP$ for all $P \in \mathcal{P}$ and for all $n \in \mathbb{N}$. By (50), this implies that $I(\varphi_n) \geq I(\psi_n)$ for all $n \in \mathbb{N}$. Since I is continuous, we can conclude that $I(\varphi) \geq I(\psi)$, proving that I satisfies (a) as well, thus, proving necessity. Sufficiency follows from routine arguments.

1. By the proof of [40, Lemma 28], it follows that u is unique up to an affine transformation. Next, fix u and consider two normalized, monotone, and translation invariant functions, I_1 and I_2, satisfying (b). Consider $\varphi \in B_0(\Omega,\mathcal{F})$. Since $u(X)$ is unbounded, it follows that there exists $k \in \mathbb{R}$ such that $\bar{\varphi} = \varphi + k1_\Omega$ belongs to $B_0(\Omega,\mathcal{F},u(X))$. This implies that there exists $f \in B_0(X)$ such that $\bar{\varphi} = u(f)$. Moreover, given the assumptions, we have that for each $f \in B_0(X)$ there exists $x_f \in X$ such that $f \sim x_f$. Since I_1 and I_2 are normalized, translation invariant, and they both satisfy (b), we have that

$$\begin{aligned} I_1(\varphi) + k = I_1(\bar{\varphi}) &= I_1(u(f)) \\ &= I_1(u(x_f)1_\Omega) \\ &= u(x_f) \\ &= I_2(u(x_f)1_\Omega) = I_2(u(f)) = I_2(\bar{\varphi}) = I_2(\varphi) + k. \end{aligned}$$

We can conclude that I_1 and I_2 coincide on $B_0(\Omega,\mathcal{F})$. Since both functionals are Lipschitz continuous functionals and $B_0(\Omega,\mathcal{F})$ is dense in $B(\Omega,\mathcal{F})$, the statement follows.

2. Consider $\varphi,\psi \in B_0(\Omega,\mathcal{F})$ such that $I(\varphi) > I(\psi)$, $k \in \mathbb{R}$, and $\{E_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ such that $E_n \downarrow \emptyset$. Since $u(X)$ is unbounded, it follows that there exists $h \in \mathbb{R}$ such that $\bar{\varphi} = \varphi + h1_\Omega$, $\bar{\psi} = \psi + h1_\Omega$, and $(k+h)1_\Omega$ belong to $B_0(\Omega,\mathcal{F},u(X))$. Since I is translation invariant, notice that $I(\bar{\varphi}) > I\left(\bar{\psi}\right)$. Moreover, we have that there exist $f,g \in B_0(X)$ and $x \in X$ such that $\bar{\varphi} = u(f)$, $\bar{\psi} = u(g)$, and $u(x) = k+h$. Since I satisfies (b), it follows that

$$f \succ g \text{ and } u(xE_nf) = (k+h)1_{E_n} + \bar{\varphi}1_{E_n^c} \qquad \forall n \in \mathbb{N}.$$

Since $\succsim$ satisfies Monotone Continuity, it follows that there exists an $N \in \mathbb{N}$ such that $xE_N f \succ g$. Since I is translation invariant and it satisfies (b), we can conclude that

$$I\left(k1_{E_N} + \varphi 1_{E_N^c}\right) + h = I\left((k+h)\,1_{E_N} + \bar{\varphi}1_{E_N^c}\right) = I\left(u\left(xE_N f\right)\right) > I\left(u\left(g\right)\right) = I\left(\bar{\psi}\right) = I\left(\psi\right) + h$$
$$\Rightarrow I\left(k1_{E_N} + \varphi 1_{E_N^c}\right) > I\left(\psi\right).$$

By Theorem 21 and [10, Theorem 54], it follows that I is inner continuous on $B_0\left(\Omega, \mathcal{F}\right)$. Since I is translation invariant and by using the same techniques contained in [13, Lemma 15], it follows that I is inner continuous on $B\left(\Omega, \mathcal{F}\right)$. ■

Lemma 33 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$. $\succsim$ satisfies the Basic Conditions, Consistency, Risk Independence, Uncertainty Aversion, and Full Unboundedness if and only if there exist an onto affine function $u : X \to \mathbb{R}$ and a normalized, monotone, lower semicontinuous, and quasiconcave functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ such that*

(a) $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ implies $I(\varphi) \geq I(\psi)$;

(b) $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$;

(c) I is continuous on $B_0(\Omega, \mathcal{F})$.

Moreover,

1. *u is unique up to an affine transformation and, given u, I is unique.*
2. *I is inner continuous if $\succsim$ satisfies Monotone Continuity.*

Proof. We first prove necessity. By [34] and since $\succsim$ satisfies the Basic Conditions and Risk Independence, it follows that there exists an affine function $u : X \to \mathbb{R}$ that represents $\succsim$ restricted to X. Furthermore, u is unique up to an affine transformation. Since $\succsim$ further satisfies Consistency and by Lemma 28, $\succsim$ satisfies Monotonicity. Since $\succsim$ satisfies Full Unboundedness, it is easy to check that $u(X) = \mathbb{R}$. It follows that

$$B_0\left(\Omega, \mathcal{F}\right) = \left\{u\left(f\right) : f \in B_0\left(X\right)\right\}. \tag{51}$$

By [10, Lemma 61], there exists a unique normalized, monotone, continuous, and quasiconcave functional $I : B_0(\Omega, \mathcal{F}) \to \mathbb{R}$ such that $f \succsim g \Leftrightarrow I(u(g)) \geq I(u(f))$.

Consider $\varphi, \psi \in B_0(\Omega, \mathcal{F})$. Assume that $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$. By (51), it follows that there exist f and g in $B_0(X)$ such that $\varphi = u(f)$ and $\psi = u(g)$. This implies that $\int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP$ for all $P \in \mathcal{P}$. By definition of P_f and P_g and since u is affine, we have that

$$u\left(P_f\right) = \int_\Omega u\left(f\right)dP \geq \int_\Omega u\left(g\right)dP = u\left(P_g\right) \qquad \forall P \in \mathcal{P}.$$

Since u represents $\succsim$ over X, this implies that $P_f \succsim P_g$ for all $P \in \mathcal{P}$. Since $\succsim$ satisfies Consistency and I represents $\succsim$, once composed with u, it follows that

$$I\left(\varphi\right) = I\left(u\left(f\right)\right) \geq I\left(u\left(g\right)\right) = I\left(\psi\right). \tag{52}$$

By Theorem 21, it follows that there exists a unique $G \in \mathcal{L}\left(\mathbb{R} \times \Delta\left(\Omega, \mathcal{F}\right)\right)$ such that

$$I\left(\varphi\right) = \min_{P \in \Delta(\Omega, \mathcal{F})} G\left(\int_\Omega \varphi dP, P\right) \qquad \forall \varphi \in B_0\left(\Omega, \mathcal{F}\right). \tag{53}$$

By Theorem 21, if we define $J : B(\Omega, \mathcal{F}) \to \mathbb{R}$ by

$$J(\varphi) = \min_{P \in \Delta(\Omega, \mathcal{F})} G\left(\int_\Omega \varphi dP, P\right) \qquad \forall \varphi \in B(\Omega, \mathcal{F})$$

then we have that J is a normalized, monotone, lower semicontinuous, and quasiconcave extension of I to $B(\Omega, \mathcal{F})$. With a small abuse of notation, we also call the extension I. By the previous part of the proof, it follows that I satisfies all the requirements of the statement as well as points (b) and (c). Moreover, I satisfies point (a) when φ and ψ are elements of $B_0(\Omega, \mathcal{F})$. This implies that I satisfies point (iv) of Lemma 27. Thus, we can conclude that I satisfies point (iii) of Lemma 27, which is point (a).

1. By [10, Lemma 61], it follows that u is unique up to an affine transformation. Moreover, by the same result and given u, I is unique over $B_0(\Omega, \mathcal{F})$. As a consequence of Theorem 21, I admits a unique normalized, monotone, lower semicontinuous, and quasiconcave extension to $B(\Omega, \mathcal{F})$, proving the uniqueness of I.

2. If $\succsim$ further satisfies Monotone Continuity then it is immediate to check that I, restricted to $B_0(\Omega, \mathcal{F})$ and represented as in (53), satisfies the conditions of [10, Theorem 54]. By [10, Theorem 54], it follows that I is inner continuous on $B_0(\Omega, \mathcal{F})$ and that there exists a probability measure R such that $G(t, P') = \infty$ for all $t \in \mathbb{R}$ and for all $P' \notin \{P \in \Delta^\sigma(\Omega, \mathcal{F}) : P << R\}$. In light of this observation and by using the same arguments of [10, Theorem 54], it follows that I is inner continuous on $B(\Omega, \mathcal{F})$. ■

Lemma 34 *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$ that satisfies the Basic Conditions, Consistency, and Risk Independence. If $u : X \to \mathbb{R}$, $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$, $\phi : u(X) \to \mathbb{R}$, and $I : B_0(\Omega, \mathcal{F}, u(X)) \to \mathbb{R}$ are such that*

1. *u is nonconstant and affine;*
2. *I is normalized, monotone, and continuous;*
3. *ϕ is strictly increasing and continuous;*
4. *if $\varphi_1, \varphi_2 \in B_0(\Omega, \mathcal{F}, u(X))$ then $\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP$ for all $P \in \mathcal{P}$ implies $I(\varphi_1) \geq I(\varphi_2)$;*
5. *if $f, g \in B_0(X)$ then $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$;*
6. *if $f, g \in B_0(X)$ and f and g are $\mathcal{G}$-measurable then*

$$f \succsim g \Leftrightarrow \phi^{-1}\left(\int_\Omega \phi(u(f)) dQ\right) \geq \phi^{-1}\left(\int_\Omega \phi(u(g)) dQ\right).$$

then, there exists $\mu \in \Delta^\sigma(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})})$ such that

$$I(\varphi) = \phi^{-1}\left(\int_{\mathcal{S}(\mathcal{P})} \phi\left(\int_\Omega \varphi(\omega) dP(\omega)\right) d\mu(P)\right) \qquad \forall \varphi \in B_0(\Omega, \mathcal{F}, u(X)). \tag{54}$$

Proof. Define $J : B(\Omega, \mathcal{G}, u(X)) \to \mathbb{R}$ by

$$J(\psi) = \phi^{-1}\left(\int_\Omega \phi(\psi) dQ\right) \qquad \forall \psi \in B(\Omega, \mathcal{G}, u(X)).$$

Since ϕ is strictly increasing and continuous, it is immediate to see that J is well defined, normalized, monotone, and continuous. Moreover, by point 6., it follows that for each f and g, which are further $\mathcal{G}$-measurable, we have that

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

By Lemma 30, it follows that $I(\varphi) = J(\varphi^\star)$ for all $\varphi \in B_0(\Omega, \mathcal{F}, u(X))$. Pick $\varphi \in B_0(\Omega, \mathcal{F}, u(X))$. It follows that

$$\begin{aligned} I(\varphi) &= J(\varphi^\star) = \phi^{-1}\left(\int_\Omega \phi(\varphi^\star)\, dQ\right) = \phi^{-1}\left(\int_W \phi(\varphi^\star)\, dQ^W\right) \\ &= \phi^{-1}\left(\int_{\mathcal{P}} \phi(\langle\varphi, \cdot\rangle)\, d\mu_{Q^W}\right) = \phi^{-1}\left(\int_{\mathcal{S}(\mathcal{P})} \phi(\langle\varphi, \cdot\rangle)\, d\mu_{Q^W}\right) \\ &= \phi^{-1}\left(\int_{\mathcal{S}(\mathcal{P})} \phi\left(\int_\Omega \varphi(\omega)\, dP(\omega)\right) d\mu_{Q^W}(P)\right) \end{aligned}$$

where the first equality follows from the previous part of the proof, the second follows by definition of J, the third equality follows from the fact that $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$ and by Fact 4, the fourth equality follows from the Change of Variables Theorem (see [1, Theorem 13.46]), the fifth equality follows from point 5. of Proposition 19, and the sixth equality is a trivial rewriting. Define $\mu : \mathcal{A}_{\mathcal{S}(\mathcal{P})} \to [0,1]$ by

$$\mu(\Gamma) = \mu_{Q^W}(\Gamma) \qquad \forall \Gamma \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}.$$

Since $\mu_{Q^W} \in \Delta^\sigma(\mathcal{P}, \mathcal{A}_{\mathcal{P}})$, $\mathcal{S}(\mathcal{P}) \in \mathcal{A}_{\mathcal{P}}$, and $\mu_{Q^W}(\mathcal{S}(\mathcal{P})) = 1$, it follows that μ is well defined and $\mu \in \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. By the previous part of the proof and since φ was arbitrarily chosen, it follows that μ satisfies (54). ■

Proof of Theorem 4. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$.

(i) implies (ii). By Lemma 32, if $\succsim$ satisfies the Basic Conditions, Consistency, Weak Certainty Independence, Uncertainty Aversion, and Unboundedness then there exist an unbounded affine function $u : X \to \mathbb{R}$ and a normalized, monotone, translation invariant, and concave functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ such that
(a) $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ implies $I(\varphi) \geq I(\psi)$;
(b) $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$.

By (b), we have that if we define $V : B_0(X) \to \mathbb{R}$ by $V(f) = I(u(f))$ for all $f \in B_0(X)$ then V represents $\succsim$.

Next, observe that if we define $\breve{I} : L \to \mathbb{R}$ as in Proposition 25. It follows that $\breve{I}$ is a normalized, monotone, translation invariant, and concave functional on L such that $I(\varphi) = \breve{I}(\langle\varphi, \cdot\rangle)$ for all $\varphi \in B(\Omega, \mathcal{G})$. Moreover, since I satisfies (a), we have that

$$I(\varphi) = I(\varphi^\star) = \breve{I}(\langle\varphi^\star, \cdot\rangle) = \breve{I}(\langle\varphi, \cdot\rangle) \qquad \forall \varphi \in B(\Omega, \mathcal{F}). \tag{55}$$

By point (f) of Lemma 24, $L = B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. Given the properties of $\breve{I}$ and by Theorem 21 (see also [40, Lemma 26]), there exists a unique, grounded, lower semicontinuous, and convex function $\gamma : \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right) \to [0, \infty]$ such that

$$\breve{I}(\xi) = \min_{\mu \in \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)} \left\{\int_{\mathcal{S}(\mathcal{P})} \xi d\mu + \gamma(\mu)\right\} \qquad \forall \xi \in L. \tag{56}$$

By the previous part of the proof and (55), we can conclude that

$$\begin{aligned} V(f) = I(u(f)) = \breve{I}(\langle u(f), \cdot\rangle) &= \min_{\mu \in \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)} \left\{\int_{\mathcal{S}(\mathcal{P})} \langle u(f), \cdot\rangle\, d\mu + \gamma(\mu)\right\} \\ &= \min_{\mu \in \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)} \left\{\int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega u(f(\omega))\, dP(\omega)\right) d\mu(P) + \gamma(\mu)\right\} \qquad \forall f \in B_0(X). \end{aligned}$$

(ii) implies (i). It is routine.

1. Uniqueness follows from point 1. of Lemma 32 and the uniqueness of $\breve{I}$ and its representation.

2. Sufficiency is routine. Viceversa, by point 2. of Lemma 32, if $\succsim$ further satisfies Monotone Continuity then I is inner continuous. By point (8) of Proposition 25, this implies that $\breve{I}$ is inner continuous. By Theorem 21, this implies that $\gamma(\mu) = \infty$ for all $\mu \notin \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. ■

Proof of Theorem 5. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$.

(i) implies (ii). By Lemma 33, if $\succsim$ satisfies the Basic Conditions, Consistency, Risk Independence, Uncertainty Aversion, and Full Unboundedness then there exist an onto affine function $u : X \to \mathbb{R}$ and a normalized, monotone, lower semicontinuous and quasiconcave functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$ such that:
(a) $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ implies $I(\varphi) \geq I(\psi)$;
(b) $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$;
(c) I is continuous on $B_0(\Omega, \mathcal{F})$.

By (b), we have that if we define $V : B_0(X) \to \mathbb{R}$ by $V(f) = I(u(f))$ for all $f \in B_0(X)$ then V represents $\succsim$.

Next, observe that if we define $\breve{I} : L \to \mathbb{R}$ as in Proposition 25. It follows that $\breve{I}$ is a normalized, monotone, lower semicontinuous, and quasiconcave functional on L such that $I(\varphi) = \breve{I}(\langle \varphi, \cdot \rangle)$ for all $\varphi \in B(\Omega, \mathcal{G})$. Moreover, since I satisfies (a), we have that

$$I(\varphi) = I(\varphi^\star) = \breve{I}(\langle \varphi^\star, \cdot \rangle) = \breve{I}(\langle \varphi, \cdot \rangle) \qquad \forall \varphi \in B(\Omega, \mathcal{F}). \tag{57}$$

By point (f) of Lemma 24, $L = B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. Given the properties of $\breve{I}$ and by Theorem 21 (see also [11]), there exists a unique function $G \in \mathcal{L}\left(\mathbb{R} \times \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)\right)$ such that

$$\breve{I}(\xi) = \min_{\mu \in \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)} G\left(\int_{\mathcal{S}(\mathcal{P})} \xi d\mu, \mu\right) \qquad \forall \xi \in L. \tag{58}$$

By the previous part of the proof and (57), we can conclude that

$$\begin{aligned} V(f) = I(u(f)) = \breve{I}(\langle u(f), \cdot \rangle) &= \min_{\mu \in \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)} G\left(\int_{\mathcal{S}(\mathcal{P})} \langle u(f), \cdot \rangle d\mu, \mu\right) \\ &= \min_{\mu \in \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)} G\left(\int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega u(f(\omega))\, dP(\omega)\right) d\mu(P), \mu\right) \qquad \forall f \in B_0(X). \end{aligned}$$

We are left to show that G is linearly continuous. By (57) and (58), notice that for each $\varphi \in B_0(\Omega, \mathcal{F}) = B_0(\mathbb{R})$ we have that

$$\min_{\mu \in \Delta\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)} G\left(\int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega \varphi(\omega)\, dP(\omega)\right) d\mu(P), \mu\right) = \breve{I}(\langle \varphi, \cdot \rangle) = I(\varphi).$$

By point (c), I is continuous on $B_0(\Omega, \mathcal{F})$, proving the linear continuity of G.

(ii) implies (i). It is routine.

1. Uniqueness follows from point 1. of Lemma 33 and the uniqueness of $\breve{I}$ and its representation.

2. Sufficiency is routine. Viceversa, by point 2. of Lemma 33, if $\succsim$ further satisfies Monotone Continuity then I is inner continuous. By point (8) of Proposition 25, this implies that $\breve{I}$ is inner continuous. By Theorem 21, this implies that $G(t, \mu) = \infty$ for all $(t, \mu) \notin \mathbb{R} \times \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. ■

Proof of Theorem 6. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$.

(i) implies (ii). By Lemma 28 and since $\succsim$ satisfies the Basic Conditions, Consistency, and Risk Independence, it follows that $\succsim$ satisfies Monotonicity. By [9], this implies that there exist a nonconstant affine function $u : X \to \mathbb{R}$ and a normalized, monotone, and continuous functional $I : B_0(\Omega, \mathcal{F}, u(X)) \to \mathbb{R}$ such that for each f and g in $B_0(X)$

$$f \succsim g \Leftrightarrow I(u(f)) \geq I(u(g)). \tag{59}$$

On the other hand, by Savage [45] (see also [12, Proposition 3]) and since $\succsim$ satisfies the Basic Conditions, Consistency, Risk Independence, Monotone Continuity, and its restriction to $\mathcal{G}$-measurable acts satisfies P2-P6 then there exist a nonatomic probability measure $Q \in \Delta^\sigma(\Omega, \mathcal{G})$, a nonconstant affine function $\hat{u} : X \to \mathbb{R}$, and a strictly increasing and continuous function $\phi : \hat{u}(X) \to \mathbb{R}$ such that for each f and g in $B_0(X)$ that are further $\mathcal{G}$-measurable

$$f \succsim g \Leftrightarrow \phi^{-1}\left(\int_\Omega \phi(\hat{u}(f))\, dQ\right) \geq \phi^{-1}\left(\int_\Omega \phi(\hat{u}(g))\, dQ\right). \tag{60}$$

Moreover, Q is unique, u is cardinally unique, and ϕ is cardinally unique given $\hat{u}$. By [9] and [12, Proposition 3], without loss of generality, we can assume that $\hat{u} = u$, that $0, 1 \in int(u(X))$, and that $\phi(1) - 1 = 0 = \phi(0)$. Consider $\varphi_1, \varphi_2 \in B_0(\Omega, \mathcal{F}, u(X))$ such that $\int_\Omega \varphi_1 dP \geq \int_\Omega \varphi_2 dP$ for all $P \in \mathcal{P}$. It follows that there exist f_1 and f_2 in $B_0(X)$ such that

$$\varphi_i = u(f_i) \text{ for } i \in \{1, 2\} \text{ and } u(P_{f_1}) = \int_\Omega u(f_1)\, dP \geq \int_\Omega u(f_2)\, dP = u(P_{f_2}) \qquad \forall P \in \mathcal{P}.$$

By (59) and since $\succsim$ satisfies Consistency, it follows that $I(\varphi_1) = I(u(f_1)) \geq I(u(f_2)) = I(\varphi_2)$, that is, I satisfies point 4. of Lemma 34. Similarly, consider $A \in \mathcal{G}$ such that $P(A) = 0$ for all $P \in \mathcal{P}$. Since $0, 1 \in int(u(X))$, it follows that $1_A, 1_\emptyset \in B_0(\Omega, \mathcal{G}, u(X))$ and that $\int_\Omega 1_\emptyset dP = 0 = P(A) = \int_\Omega 1_A dP$ for all $P \in \mathcal{P}$. Define $x, y \in X$ to be such that $u(x) = 1$ and $u(y) = 0$. By (60) and since $\phi(1) - 1 = 0 = \phi(0)$ and $\succsim$ satisfies Consistency, it follows that

$$\phi^{-1}(Q(A)) = \phi^{-1}\left(\int_\Omega \phi(1_A)\, dQ\right) = \phi^{-1}\left(\int_\Omega \phi(u(xAy))\, dQ\right) = \phi^{-1}\left(\int_\Omega \phi(u(y))\, dQ\right) = 0,$$

that is, $Q(A) = 0$ and $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$. By Lemma 34, it follows that there exists $\mu \in \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$ such that

$$I(\varphi) = \phi^{-1}\left(\int_{\mathcal{S}(\mathcal{P})} \phi\left(\int_\Omega \varphi(\omega)\, dP(\omega)\right) d\mu(P)\right) \qquad \forall \varphi \in B_0(\Omega, \mathcal{F}, u(X)). \tag{61}$$

We next show that μ is nonatomic. Consider $C \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$ and $A \in \mathcal{G}$ such that $\langle 1_A, \cdot \rangle = 1_C$ and $\mu(C) > 0$. By (61), it follows that

$$0 < \phi^{-1}(\mu(C)) = \phi^{-1}\left(\int_{\mathcal{S}(\mathcal{P})} \phi\left(\int_\Omega 1_A(\omega)\, dP(\omega)\right) d\mu(P)\right) = I(1_A) = \phi^{-1}(Q(A)) \tag{62}$$

where the last equality is consequence of Lemma 30.[40] Next, consider $C_1 \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$ such that $\mu(C_1) > 0$. By point (l) of Lemma 24, there exists $A_1 \in \mathcal{G}$ such that $\langle 1_{A_1}, \cdot \rangle = 1_{C_1}$. By (62), it follows that $Q(A_1) > 0$. Since Q is nonatomic, it follows that there exists $A_2 \in \mathcal{G}$ such that $A_2 \subseteq A_1$ and $0 < Q(A_2) < Q(A_1)$. By the proof of point (i) of Lemma 24, we have that $1_{C_1} = \langle 1_{A_1}, \cdot \rangle \geq \langle 1_{A_2}, \cdot \rangle = 1_{C_2}$ where $C_2 \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$, that is, $C_2 \subseteq C_1$. By (62), we can conclude that

$$0 < \mu(C_2) = Q(A_2) < Q(A_1) = \mu(C_1),$$

thus proving that μ is nonatomic.

[40] $J : B(\Omega, \mathcal{G}, u(X)) \to \mathbb{R}$ is defined by

$$\psi \mapsto \phi^{-1}\left(\int_\Omega \phi(\psi)\, dQ\right).$$

Finally, by (61), notice that the function $V : B_0(X) \to \mathbb{R}$, defined by

$$V(f) = \int_{\mathcal{S}(\mathcal{P})} \phi\left(\int_\Omega u(f(\omega))\, dP(\omega)\right) d\mu(P) \qquad \forall f \in B_0(X),$$

is such that $V = \phi \circ I \circ u$. Since ϕ is strictly increasing and by (59), it follows that V represents $\succsim$.

(ii) implies (i). Assume that there exist a nonatomic probability measure $\mu \in \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$, a nonconstant affine function $u : X \to \mathbb{R}$, and a strictly increasing and continuous function $\phi : u(X) \to \mathbb{R}$ such that

$$V(f) = \int_{\mathcal{S}(\mathcal{P})} \phi\left(\int_\Omega u(f(\omega))\, dP(\omega)\right) d\mu(P) \qquad \forall f \in B_0(X)$$

represents $\succsim$. Define $I : B_0(\Omega, \mathcal{F}, u(X)) \to \mathbb{R}$ by

$$I(\varphi) = \phi^{-1}\left(\int_{\mathcal{S}(\mathcal{P})} \phi\left(\int_\Omega \varphi(\omega)\, dP(\omega)\right) d\mu(P)\right) \qquad \forall \varphi \in B_0(\Omega, \mathcal{F}, u(X)).$$

Since ϕ is strictly increasing and continuous and $\mu \in \Delta^\sigma\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$, it is immediate to see that I is normalized, monotone, continuous, and inner/outer continuous. Since ϕ is strictly increasing and continuous, it follows that the function $V' = \phi^{-1} \circ V = I \circ u$ represents $\succsim$. By the arguments contained in [9], it follows that $\succsim$ satisfies the Basic Conditions and Risk Independence. It is immediate to see that $\succsim$ satisfies Consistency and Monotone Continuity as well. Next, define $Q = \bar{\mu}_\mathcal{G}$ where $\bar{\mu}$ is the barycenter of μ. Notice that $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$. We next show that Q is nonatomic. Consider $A \in \mathcal{G}$ and $C \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$ such that $\langle 1_A, \cdot \rangle = 1_C$. It follows that

$$\mu(C) = \int_{\mathcal{S}(\mathcal{P})} 1_C d\mu = \int_{\mathcal{S}(\mathcal{P})} \langle 1_A, \cdot \rangle\, d\mu = \int_\Omega 1_A d\bar{\mu} = \int_\Omega 1_A dQ = Q(A). \tag{63}$$

Consider $A_1 \in \mathcal{G}$ such that $Q(A_1) > 0$. By the proof of point (i) of Lemma 24, we have that $\langle 1_{A_1}, \cdot \rangle = 1_{C_1}$ where $C_1 \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$. By (63), it follows that $\mu(C_1) > 0$. Since μ is nonatomic, it follows that there exists $C_2 \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$ such that $C_2 \subseteq C_1$ and $0 < \mu(C_2) < \mu(C_1)$. By point (l) of Lemma 24, there exists $A_2 \in \mathcal{G}$ such that $\langle 1_{A_2}, \cdot \rangle = 1_{C_2}$. Define $A_3 = A_1 \cap A_2 \in \mathcal{G}$. Since $C_2 \subseteq C_1$ and $A_1, A_2 \in \mathcal{G}$, it follows that

$$1_{C_2} = 1_{C_1} \wedge 1_{C_2} = \langle 1_{A_1}, \cdot \rangle \wedge \langle 1_{A_2}, \cdot \rangle = \langle 1_{A_1} \wedge 1_{A_2}, \cdot \rangle = \langle 1_{A_3}, \cdot \rangle.$$

It follows that $A_3 \subseteq A_1$. By (63), we thus have that $0 < Q(A_3) = \mu(C_2) < \mu(C_1) = Q(A_1)$, proving that Q is nonatomic. Next, observe that for each $\psi \in B_0(\Omega, \mathcal{G}, u(X))$ we have that

$$\begin{aligned} I(\psi) &= \phi^{-1}\left(\int_{\mathcal{S}(\mathcal{P})} \phi(\langle \psi, \cdot \rangle)\, d\mu\right) = \phi^{-1}\left(\int_{\mathcal{S}(\mathcal{P})} \phi(\langle \psi, \cdot \rangle)\, d\mu_{Q^W}\right) = \phi^{-1}\left(\int_{\mathcal{P}} \phi(\langle \psi, \cdot \rangle)\, d\mu_{Q^W}\right) \\ &= \phi^{-1}\left(\int_W \phi(\psi^\star)\, dQ^W\right) = \phi^{-1}\left(\int_\Omega \phi(\psi^\star)\, dQ\right) = \phi^{-1}\left(\int_\Omega \phi(\psi)\, dQ\right). \end{aligned}$$

where the first equality follows by definition of I, the second equality follows by (30), the third equality follows from point 5. of Proposition 19, the fourth equality follows from the Change of Variables Theorem (see [1, Theorem 13.46]), the fifth equality follows from the fact that $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$ and by Fact 4, and the sixth equality follows from the fact that $Q \in \Delta^\sigma(\Omega, \mathcal{G}, \mathcal{P})$ and $\psi = \psi^\star$ $\mathcal{P}$-a.s., since $\psi \in B(\Omega, \mathcal{G})$. This implies that $V'(f) = \phi^{-1}(V(f)) = I(u(f)) = \phi^{-1}\left(\int_\Omega \phi(u(f))\, dQ\right)$ for all $f \in B_0(X)$ that are further $\mathcal{G}$-measurable. By [12, Proposition 3] and since Q is a nonatomic probability measure, $u : X \to \mathbb{R}$ is nonconstant and affine, $\phi : u(X) \to \mathbb{R}$ is strictly increasing and continuous, and V' represents $\succsim$, it follows that the restriction of $\succsim$ to $\mathcal{G}$-measurable acts satisfies P2-P6.

1. Uniqueness is routine to check.

2. By [12, Proposition 3], if $\succsim$ satisfies Uncertainty Aversion then ϕ is concave. Viceversa, if ϕ is concave then it is immediate to see that $\succsim$ satisfies Uncertainty Aversion. ■

Proof of Theorem 7. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a Dynkin space and $\succsim$ a binary relation on $B_0(X)$.

(i) implies (ii). By Lemma 31, if $\succsim$ satisfies the Basic Conditions, Consistency, and Comonotonic Independence on $\mathcal{G}$-measurable acts then there exist a nonconstant affine function $u : X \to \mathbb{R}$ and a normalized, monotone, translation invariant, and positively homogeneous functional $I : B(\Omega, \mathcal{F}) \to \mathbb{R}$, which is further comonotonic additive on $B(\Omega, \mathcal{G})$ and such that:
(a) $\int_\Omega \varphi dP \geq \int_\Omega \psi dP$ for all $P \in \mathcal{P}$ implies $I(\varphi) \geq I(\psi)$;
(b) $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$.

By (b), we have that if we define $V : B_0(X) \to \mathbb{R}$ by $V(f) = I(u(f))$ for all $f \in B_0(X)$ then V represents $\succsim$.

Next, observe that if we define $\breve{I} : L \to \mathbb{R}$ as in Proposition 25. It follows that $\breve{I}$ is a normalized, monotone, and comonotonic additive functional on L such that $I(\varphi) = \breve{I}(\langle \varphi, \cdot \rangle)$ for all $\varphi \in B(\Omega, \mathcal{G})$. Moreover, since I satisfies (a) we have that

$$I(\varphi) = I(\varphi^\star) = \breve{I}(\langle \varphi^\star, \cdot \rangle) = \breve{I}(\langle \varphi, \cdot \rangle) \qquad \forall \varphi \in B(\Omega, \mathcal{F}). \tag{64}$$

By point (f) of Lemma 24, $L = B\left(\mathcal{S}(\mathcal{P}), \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right)$. Given the properties of $\breve{I}$ and by [46], there exists a unique capacity $\nu : \mathcal{A}_{\mathcal{S}(\mathcal{P})} \to [0,1]$ such that

$$\breve{I}(\xi) = \int_{\mathcal{S}(\mathcal{P})} \xi d\nu \qquad \forall \xi \in L. \tag{65}$$

By the previous part of the proof and (64), we can conclude that

$$V(f) = I(u(f)) = \breve{I}(\langle u(f), \cdot \rangle) = \int_{\mathcal{S}(\mathcal{P})} \langle u(f), \cdot \rangle d\nu = \int_{\mathcal{S}(\mathcal{P})} \left(\int_\Omega u(f(\omega)) dP(\omega) \right) d\nu(P) \qquad \forall f \in B_0(X),$$

proving the implication.

(ii) implies (i). It is routine.

1. Uniqueness follows from point 1. of Lemma 31 and the uniqueness of $\breve{I}$ and its representation.

2. Necessity is trivial. Viceversa, by point 2. of Lemma 31, if $\succsim$ further satisfies Uncertainty Aversion then I is quasiconcave. By point (4) of Proposition 25, this implies that $\breve{I}$ is quasiconcave. By [46] and (65), it follows that ν is convex.

3. Necessity is trivial. Viceversa, by point 3. of Lemma 31, if $\succsim$ further satisfies Monotone Continuity then I is inner and outer continuous on $B(\Omega, \mathcal{G})$. By point (8) of Proposition 25, this implies that $\breve{I}$ is inner and outer continuous. Given (65), it follows that ν is inner and outer continuous, that is, continuous. ■

Proof of Theorem 8. (i) implies (ii). By Remark 29 and since $\succsim$ satisfies the Weak Basic Conditions, Consistency, and Independence, it follows that $\succsim$ satisfies Monotonicity and it is reflexive. By [26, Theorem 1] and since $\succsim$ is reflexive and satisfies the Weak Basic Conditions, Monotonicity, and Independence, there exist a nonconstant affine function $u : X \to \mathbb{R}$ and a nonempty, closed, and convex set $C \subseteq \Delta(\Omega, \mathcal{F})$ such that

$$f \succsim g \Leftrightarrow \int_\Omega u(f) dP \geq \int_\Omega u(g) dP \qquad \forall P \in C. \tag{66}$$

Moreover, u is unique up to an affine transformation and it represents $\succsim$ restricted to X. Without loss of generality, we can assume that $1, 0 \in int(u(X))$. Define $x, z \in X$ and the sequence $\{y_k\}_{k\in\mathbb{N}} \subseteq X$ to be such that $u(x) = 1$, $u(z) = 0$, and $u(y_k) = 1/k$ for all $k \in \mathbb{N}$. Next, consider a sequence of events $\{E_n\}_{n\in\mathbb{N}}$ in $\mathcal{F}$ such that $E_n \downarrow \emptyset$. It follows that $u(xE_nz) = 1_{E_n}$ for all $n \in \mathbb{N}$. Since $\succsim$ satisfies Binary Monotone Continuity and $y_k \succ z$ for all $k \in \mathbb{N}$, it follows that for each $k \in \mathbb{N}$ there exists $N_k \in \mathbb{N}$ such that $y_k \succsim xE_{N_k}z$. By (66), this implies that

$$\frac{1}{k} = u(y_k) \geq \int_\Omega u(xE_{N_k}z)\, dP = P(E_{N_k}) \geq \lim_n P(E_n) \geq 0 \qquad \forall P \in C. \tag{67}$$

Passing to the limit in (67), it follows that $\lim_n P(E_n) = 0$ for all $P \in C$. Since $\{E_n\}_{n\in\mathbb{N}}$ was arbitrarily chosen, it follows that $C \subseteq \Delta^\sigma(\Omega, \mathcal{F})$. By Proposition 18 and (66), and since $\succsim$ satisfies Consistency, u is affine, and it represents $\succsim$ restricted to X, it follows that for each f and g in $B_0(X)$

$$\begin{aligned}
\int_\Omega u(f)\, dP \geq \int_\Omega u(g)\, dP \quad \forall P \in \text{mco}(\mathcal{P}) &\Leftrightarrow \int_\Omega u(f)\, dP \geq \int_\Omega u(g)\, dP \quad \forall P \in \mathcal{P} \\
&\Leftrightarrow \int_\Omega f dP \succsim \int_\Omega g dP \qquad \forall P \in \mathcal{P} \\
&\Rightarrow \int_\Omega u(f)\, dP \geq \int_\Omega u(g)\, dP \qquad \forall P \in C.
\end{aligned}$$

By [24, Proposition A.1] and since C is closed and convex and $\text{mco}(\mathcal{P})$ is convex, it follows that $C \subseteq cl(\text{mco}(\mathcal{P}))$. Recall that $\text{mco}(\mathcal{P}), C \subseteq \Delta^\sigma(\Omega, \mathcal{F})$. By point 3. of Proposition 18, this implies that

$$C = C \cap \Delta^\sigma(\Omega, \mathcal{F}) \subseteq cl(\text{mco}(\mathcal{P})) \cap \Delta^\sigma(\Omega, \mathcal{F}) = \text{mco}(\mathcal{P}).$$

Define $\mathcal{M} = \left\{\mu \in \Delta^\sigma\left(\mathcal{S}_\mathcal{P}, \mathcal{A}_{\mathcal{S}(\mathcal{P})}\right) : \bar{\mu} = P \text{ for some } P \text{ in } C\right\}$. By Corollary 20, it follows that $\mathcal{M}$ is well defined. By (30) and since C is nonempty, closed (compact), and convex, it follows that $\mathcal{M}$ shares the same properties.[41] In light of (66) and given $\mathcal{M}$, it is immediate to check that $\mathcal{M}$ represents $\succsim$ as in (21).

(ii) implies (i). It is routine.

Finally, uniqueness follows from routine arguments. ■

Proof of Proposition 9. Since $\succsim$ satisfies the Basic Conditions and Risk Independence, observe that either under (i) or under (ii) $\succsim$ satisfies Monotonicity. By [9, Proposition 1 and Proposition 2] and since $\succsim$ satisfies the Basic Conditions, Monotonicity, and Risk Independence, it follows that there exist a nonconstant and affine function $u : X \to \mathbb{R}$ and a closed and convex set $C \subseteq \Delta(\Omega, \mathcal{F})$ such that

$$f \succsim^* g \Leftrightarrow \int_\Omega u(f)\, dP \geq \int_\Omega u(g)\, dP \qquad \forall P \in C. \tag{68}$$

[41] Notice tha the map from C to $\mathcal{M}$ defined by

$$P \mapsto \mu_P \qquad \forall P \in C$$

is a well defined affine homeomorphism when both the domain and the target space are endowed with the relative topologies induced by the respective weak* topologies. Recall that $C \subseteq \text{mco}(\mathcal{P})$. By Corollary 20 and definition, the above map is well defined and bijective. By (30) and its discussion, it is affine. Regarding continuity, consider a net $\{P_\alpha\}_{\alpha\in A} \subseteq C$ such that $P_\alpha \to \bar{P} \in C$. Consider $D \in \mathcal{A}_{\mathcal{S}(\mathcal{P})}$. By point (l) of Lemma 24, it follows that there exists $A \in \mathcal{G}$ such that $1_D = \langle 1_A, \cdot \rangle$. This implies that

$$\begin{aligned}
\mu_{P_\alpha}(D) &= \int_{\mathcal{S}(\mathcal{P})} 1_D d\mu_{P_\alpha} = \int_{\mathcal{S}(\mathcal{P})} \langle 1_A, \cdot \rangle\, d\mu_{P_\alpha} = P_\alpha(A) \to \bar{P}(A) \\
&= \int_{\mathcal{S}(\mathcal{P})} \langle 1_A, \cdot \rangle\, d\mu_{\bar{P}} = \int_{\mathcal{S}(\mathcal{P})} 1_D d\mu_{\bar{P}} = \mu_{\bar{P}}(D).
\end{aligned}$$

Since D was arbitrarily chosen, the continuity of the above map follows. By [1, Theorem 2.36] and since C is compact and $\mathcal{M}$ is an Hausdorff topological space, it follows that the above map is a homeomorphism.

Moreover, u is unique up to an affine transformation and it represents $\succsim$ restricted to X while C is independent of the choice of u. Without loss of generality, we can assume that $1, 0 \in int\,(u(X))$. Define $x, z \in X$ and the sequence $\{y_k\}_{k\in\mathbb{N}} \subseteq X$ to be such that $u(x) = 1$, $u(z) = 0$, and $u(y_k) = 1/k$ for all $k \in \mathbb{N}$. Next, consider a sequence of events $\{E_n\}_{n\in\mathbb{N}}$ in $\mathcal{F}$ such that $E_n \downarrow \emptyset$. It follows that $u(xE_n z) = 1_{E_n}$ for all $n \in \mathbb{N}$. Since $\succsim^*$ satisfies Binary Monotone Continuity and $y_k \succ^* z$ for all $k \in \mathbb{N}$, it follows that for each $k \in \mathbb{N}$ there exists $N_k \in \mathbb{N}$ such that $y_k \succsim^* xE_{N_k}z$. By (68), this implies that

$$\frac{1}{k} = u(y_k) \geq \int_\Omega u(xE_{N_k}z)\,dP = P(E_{N_k}) \geq \lim_n P(E_n) \geq 0 \qquad \forall P \in C. \tag{69}$$

Passing to the limit in (69), it follows that $\lim_n P(E_n) = 0$ for all $P \in C$. Since $\{E_n\}_{n\in\mathbb{N}}$ was arbitrarily chosen, it follows that $C \subseteq \Delta^\sigma(\Omega, \mathcal{F})$.

(i) implies (ii). By Lemma 28 and since $\succsim$ satisfies the Basic Conditions and Risk Independence, if $\succsim$ satisfies Consistency then it satisfies Monotonicity. By the initial and common part of the proof and since $\succsim$ satisfies Consistency, it follows that for $f, g \in B_0(X)$

$$\int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP \quad \forall P \in \operatorname{mco}(\mathcal{P}) \quad \Rightarrow \int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP \quad \forall P \in \mathcal{P} \quad \Rightarrow \quad f \succsim g.$$

By definition of $\succsim^*$ and the initial and common part of the proof, it is immediate to see that this implies that

$$\int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP \quad \forall P \in \operatorname{mco}(\mathcal{P}) \quad \Rightarrow \int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP \quad \forall P \in C.$$

By [24, Proposition A.1] and since C is closed and convex and $\operatorname{mco}(\mathcal{P})$ is convex, it follows that $C \subseteq cl\,(\operatorname{mco}(\mathcal{P}))$. Recall that $\operatorname{mco}(\mathcal{P}), C \subseteq \Delta^\sigma(\Omega, \mathcal{F})$. By point 3. of Proposition 18, this implies that

$$C = C \cap \Delta^\sigma(\Omega, \mathcal{F}) \subseteq cl\,(\operatorname{mco}(\mathcal{P})) \cap \Delta^\sigma(\Omega, \mathcal{F}) = \operatorname{mco}(\mathcal{P}).$$

However, it is immediate to see that the set of invariant measures is measure convex, thus $\mathcal{P} = \operatorname{mco}(\mathcal{P})$. Finally, if $f \in B_0(X)$ then $u(f) \in B_0(\Omega, \mathcal{F})$. It follows that there exists a partition $\{E_i\}_{i=1}^n \subseteq \mathcal{F}$ and a collection $\{\alpha_i\}_{i=1}^n$ such that $u(f) = \sum_{i=1}^n \alpha_i 1_{E_i}$. If $\pi \in G$ then we have that $u(f \circ \pi) = \sum_{i=1}^n \alpha_i 1_{\pi^{-1}(E_i)}$. Since $\mathcal{P}$ is the set of invariant measures and $C \subseteq \mathcal{P}$, we have that $\int_\Omega u(f)\,dP = \int_\Omega u(f \circ \pi)\,dP$ for all $P \in C$, that is, $f \circ \pi \sim^* f$, proving the statement.

(ii) implies (i). Consider $E \in \mathcal{F}$ and $\pi \in G$. By the initial and common part of the proof, recall that, without loss of generality, $1, 0 \in int\,(u(X))$. This implies that there exists $f \in B_0(X)$ such that $u(f) = 1_E$. By (68) and since $\succsim$ satisfies Unambiguous Symmetry, we have that

$$f \circ \pi \sim^* f \text{ and } P\left(\pi^{-1}(E)\right) = \int_\Omega u(f \circ \pi)\,dP = \int_\Omega u(f)\,dP = P(E) \qquad \forall P \in C.$$

Since $C \subseteq \Delta^\sigma(\Omega, \mathcal{F})$ and E and π were arbitrarily chosen, it follows that $C \subseteq \mathcal{P}$. Since $\succsim^*$ is a subrelation of $\succsim$, u represents $\succsim$ restricted to X, and u is affine, it follows that for each f and g in $B_0(X)$

$$\begin{aligned} \int_\Omega f dP \succsim \int_\Omega g dP \quad \forall P \in \mathcal{P} \Leftrightarrow \int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP \quad \forall P \in \mathcal{P} \\ \Rightarrow \int_\Omega u(f)\,dP \geq \int_\Omega u(g)\,dP \quad \forall P \in C \Rightarrow f \succsim^* g \Rightarrow f \succsim g, \end{aligned}$$

proving that $\succsim$ satisfies Consistency. ■

References

[1] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis*, 3rd ed., Springer Verlag, Berlin, 2006.

[2] N. I. Al-Najjar and L. De Castro, Parametric representation of preferences, mimeo, October 2010.

[3] M. Amarante, Foundations of neo-Bayesian statistics, *Journal of Economic Theory*, 144, 2146–2173, 2009.

[4] K. Arrow, *Essays in the Theory of Risk-Bearing*, North-Holland, Amsterdam, 1970.

[5] H. Becker and A. S. Kechris, Borel actions of polish groups, *Bulletin of the American Mathematical Society*, 28, 1993.

[6] J. O. Berger, The robust Bayesian viewpoint, in *Robustness of Bayesian analysis*, (J. B. Kadane ed.), North-Holland, Amsterdam, 1984.

[7] T. Bewley, Knightian decision theory: part I, *Decisions in Economics and Finance*, 25, 79–110, 2002.

[8] P. Billingsley, *Probability and measure*, 3rd ed., John Wiley & Sons, New York, 1995.

[9] S. Cerreia-Vioglio, P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi, Rational preferences under ambiguity, Collegio Carlo Alberto Notebook 169, 2010.

[10] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Uncertainty Averse Preferences, *Journal of Economic Theory*, forthcoming.

[11] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Complete monotone quasiconcave duality, *Mathematics of Operations Research*, forthcoming.

[12] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Probabilistic sophistication and uncertainty aversion, mimeo, 2009.

[13] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Signed integral representations of comonotonic additive functionals, mimeo, 2010.

[14] E. Danan and A. Ziegelmeyer, Are preferences complete? An experimental measurement of indecisiveness under risk, mimeo, 2006.

[15] L. De Castro and N. I. Al-Najjar, A subjective foundation of objective probability, mimeo, July 2009.

[16] B. de Finetti, La prevision: ses lois logiques, ses sources subjectives, Annales de l'Institute Henri Poincaré, 7, 1–68, 1937.

[17] C. Dellacherie and P. A. Meyer, *Probabilités et potentiel: Chapitres I à IV*, Hermann, Paris, 1975.

[18] D. Denneberg, *Non-additive measure and integral*, Kluwer, Dordrecht, 1994.

[19] E. B. Dynkin, Sufficient statistics and extreme points, *The Annals of Probability*, 6, 705–730, 1978.

[20] K. Eliaz and P. Ortoleva, Uncertain probabilities vs. uncertain outcomes: An experimental study of attitudes towards ambiguity, mimeo, 2010.

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

[22] L. Epstein and K. Seo, Symmetry of evidence without evidence of symmetry, *Theoretical Economics*, 5, 313–368, 2010.

[23] L. Epstein and K. Seo, A de Finetti theorem for capacities: ambiguity about correlation, mimeo, 2010.

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

[25] P. Ghirardato and M. Siniscalchi, A more robust definition of multiple priors, Carlo Alberto Notebook 100, 2010.

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

[27] I. Gilboa and M. Marinacci, Ambiguity and the Bayesian paradigm, mimeo, 2010.

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

[29] R. Giraud, Objective imprecise probabilistic information, second order beliefs and ambiguity aversion: an axiomatization, *ISIPTA'05 electronic proceedings*, Pittsburgh, 2005.

[30] R. M. Gray, *Probability, random processes, and ergodic properties*, 2nd ed. Springer, New York, 2009.

[31] P. D. Grünwald and A. P. Dawid, Game theory, maximum entropy, minimum discrepancy and robust Bayesian decision theory, *The Annals of Statistics*, 32, 1367–1433, 2004.

[32] J. Y. Halpern, *Reasoning about uncertainty*, The MIT Press, Boston, 2003.

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

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

[35] P. J. Huber and E. M. Ronchetti, *Robust statistics*, 2nd ed., Wiley, New York, 2009.

[36] P. J. Huber and J. Strassen, Minimax tests and the Neyman-Pearson Lemma for capacities, *The Annals of Statistics*, 1, 251–263, 1973.

[37] P. Klibanoff, M. Marinacci, and S. Mukerji, A smooth model of decision making under ambiguity, *Econometrica*, 73, 1849–1892, 2005.

[38] P. Klibanoff, S. Mukerji, and K. Seo, Relevance and symmetry, mimeo, September 2010.

[39] I. Kopylov, Choice deferral and ambiguity aversion, *Theoretical Economics*, 4, 199–225, 2009.

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

[41] R. Nau, Uncertainty aversion with second-order utilities and probabilities, *Management Science*, 52, 136–145, 2004.

[42] K. Nehring, Imprecise probabilistic beliefs as a context for decision making under ambiguity, *Journal of Economic Theory*, 144, 1054–1091, 2009.

[43] W. S. Neilson, A simplified axiomatic approach to ambiguity aversion, *Journal of Risk and Uncertainty*, 41, 113–124, 2010.

[44] K. R. Parthasarathy, Projective unitary antiunitary representations of locally compact groups, *Communications in Mathematical Physics*, 15, 305–328, 1969.

[45] L. J. Savage, *The foundations of statistics*, Wiley, New York, 1954.

[46] D. Schmeidler, Integral representation without additivity, *Proceedings of the American Mathematical Society*, 97, 255–261, 1986.

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

[48] K. Seo, Ambiguity and second-order belief, *Econometrica*, 77, 1575–1605, 2009.

[49] G. Shafer, Belief functions and parametric models, *Journal of the Royal Statistical Society: Series B*, 44, 322–352, 1982.

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

[51] V. S. Varadarajan, Groups of automorphisms of Borel spaces, *Transactions of the American Mathematical Society*, 109, 191–220, 1963.

[52] C. Villegas, On qualitative probability σ-algebras, *The Annals of Mathematical Statistics*, 35, 1787-1796, 1964.

[53] A. Wald, Basic ideas of a general theory of statistical decision rules, in *Proceedings of the International Congress of Mathematicians*, Harvard University Press, Cambridge, 1952.

[54] L. A. Wasserman and J. B. Kadane, Bayes' theorem for Choquet capacities, *The Annals of Statistics*, 18, 1328–1339, 1990.

[55] G. Winkler, Choquet order and simplices with applications in probabilistic models, Lecture Notes in Mathematics, 1145, Springer, Berlin, 1985.