Model and Predictive Uncertainty: A Foundation for Smooth Ambiguity Preferences*

Tommaso Denti[†] Luciano Pomatto[‡]

January 2020

Abstract

Smooth ambiguity preferences (Klibanoff, Marinacci, and Mukerji, 2005) describe a decision maker who evaluates each act f according to the twofold expectation

$$V(f) = \int_{\mathcal{P}} \phi \left(\int_{\Omega} u(f)\, \mathrm{d}p \right) \mathrm{d}\mu(p)$$

defined by a utility function u, an ambiguity index ϕ, and a belief μ over a set $\mathcal{P}$ of probabilities. We provide an axiomatic foundation for the representation, taking as a primitive a preference over standard Anscombe-Aumann acts. We study a special case where $\mathcal{P}$ is a subjective statistical model that is point identified, i.e. the decision maker believes that the true law $p \in \mathcal{P}$ can be recovered empirically. Our main axiom is a joint weakening of Savage's sure-thing principle and Anscombe-Aumann's mixture independence. In addition, we show that the parameters of the representation can be uniquely recovered from preferences, thereby making operational the separation between ambiguity attitude and perception, an hallmark feature of the smooth ambiguity representation.

*We are grateful to Simone Cerreia-Vioglio, Federico Echenique, Massimo Marinacci, Francesca Molinari, Kota Saito, Kyoungwon Seo, Jörg Stoye, and, especially, Fabio Maccheroni, for many comments and suggestions. We would like to thank Università Bocconi for its hospitality.

[†]Cornell University.

[‡]Caltech.

1 Introduction

Smooth ambiguity preferences, introduced by Klibanoff, Marinacci, and Mukerji (2005), have received great attention in economics and decision theory. Under these preferences, an act $f\colon \Omega \to X$ mapping states of the world to outcomes is ranked according to the representation

$$V(f) = \int_{\mathcal{P}} \phi\left(\int_{\Omega} u(f)\,\mathrm{d}p\right)\,\mathrm{d}\mu(p). \tag{1}$$

Acts are first evaluated by their expected utility with respect to each probability measure p in a set $\mathcal{P}$. These expectations are then averaged by means of a belief μ over probabilities and an increasing transformation ϕ. When the support of μ is not a singleton, the decision maker entertains multiple probabilistic scenarios. If in addition ϕ is not linear, then preferences can express ambiguity aversion or seeking, and can accommodate behavior that could not otherwise be modelled under subjective expected utility.

Smooth ambiguity preferences have seen a wide range of economic applications. They have also been the subject of a well-known debate, as attested by the exchange between Epstein (2010) and Klibanoff, Marinacci, and Mukerji (2012). The debate concerns the preferences' interpretation and behavioral foundations, and has cast doubts on whether the elements of the representation can be recovered from choice data.

In this paper we provide an axiomatic foundation for a class of smooth ambiguity preferences that admits an explicit statistical interpretation. Taking as a primitive a preference relation over Anscombe-Aumann acts, we show that smooth ambiguity preferences can be characterized by relating two tenets of Bayesian reasoning, the Anscombe-Aumann independence axiom and Savage's sure-thing principle; our main axiom is a joint weakening of these two principles. In addition, we show that the elements of the representation (1) can be uniquely recovered from preferences.

We can distinguish between two possible interpretations of smooth ambiguity preferences. In one view, the probability μ measures the agent's degree of confidence over different subjective beliefs. The motivating idea is that a person might be unable to deem an event A as being more or less likely than another event B, but nevertheless might have higher confidence in "A being more likely than B" than in "B being more likely than A." Such second-order beliefs are problematic, because it is difficult to envision what evidence could be used to elicit them. They also open the door to an infinite regress problem: there seems to be no clear reason for an agent to entertain second-order beliefs, but not third and higher-order beliefs as well (see, e.g., the discussions in Savage, 1954, p. 58; Marschak et al., 1975).

We adopt an alternative interpretation, already suggested by Klibanoff, Marinacci, and Mukerji (2005). According to this interpretation, the domain $\mathcal{P}$ is a subjective statistical model adopted by the agent as a guide for making decisions, and each measure $p \in \mathcal{P}$ corresponds to a possible law governing the states. The belief μ is a prior over the "true" law, by analogy with the framework of Bayesian statistics. Under this view, ambiguity is generated by uncertainty about the correct law of nature p, rather than by inability to express decisive first-order beliefs. Eliciting the prior μ amounts to observing the agent's bets on what is the true p.

The statistical interpretation of smooth preferences has become standard in applications and theoretical work (for a survey, see Marinacci, 2015). To formalize it, we adopt a general formulation introduced by Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013). We ask $\mathcal{P}$ to satisfy what is perhaps the single most fundamental assumption in statistical modeling, that of being identifiable. We say that a set $\mathcal{P}$ of probabilities over states is *identifiable* if there is a function $k\colon \Omega \to \mathcal{P}$, mapping observable states to probability models, such that for all $p \in \mathcal{P}$

$$p(\{\omega : k(\omega) = p\}) = 1.$$

In the mind of the decision maker, the quantity k will reveal, almost surely, the true law governing the state. The value taken by k is the missing information that generates ambiguity.

With different methods, terminologies, and motivations, a number of recent papers have made important progress in providing foundations for identifiable smooth preferences. A natural special case is one where $\mathcal{P}$ consists of the ergodic measures derived from a given transformation of the state space. This is the subject of Al-Najjar and De Castro (2014), who characterize identifiable smooth preferences in ergodic environments, as well as more general preferences. Klibanoff, Mukerji, Seo, and Stanca (2021) consider preferences that are invariant with respect to a permutation of the states, in the spirit of exchangeability. The question of characterizing general identifiable preferences was first addressed by Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013), who axiomatize identifiable preferences in an augmented Anscombe-Aumann framework where $\mathcal{P}$ is a primitive of the analysis.

Each of these papers involve nontrivial conceptual and technical innovations. However, the question of providing a foundation for identifiable smooth preferences that is purely behavioral—that is, solely in terms of preferences over acts—has remained open. Here we aim to fill this gap. The key difficulty is inferring the subjective statistical model $\mathcal{P}$ from preferences, rather than assuming that such a structure (or, alternatively, a notion of symmetry across states) is given from the outset.

The question is important for a number of reasons. A long-standing difficulty with smooth preferences is to understand how they differ behaviorally from other models of choice under ambiguity. This point has been raised, for example, by Epstein (2010), who writes: "*[...] because of its problematic foundations, the behavioral content of the model and how it differs from multiple priors, for example, are not clear.*" The seminal contributions of Schmeidler (1989) and Gilboa and Schmeidler (1989) characterize Choquet and maxmin expected utility with precise weakenings of the independence axiom. No equivalent result is known for smooth preferences.

A more practical challenge concerns the uniqueness of the elements ϕ, $\mathcal{P}$, and μ of the representation. In applications, a key reason for adopting smooth preferences is their separation between ambiguity perception, represented by $\mathcal{P}$ and μ, and ambiguity attitude, which is represented by the function ϕ and is meant to be a personal trait that is portable across decision problems. The status of such separation, however, is dubious if the mathematical objects are not uniquely pinned down by choice data. A behavioral foundation is thus necessary to clarify the meaning of the representation.

Finally, providing a behavioral foundation is required in order to give empirical content to the assumption that the decision maker reasons according to *some* statistical model $\mathcal{P}$. In many cases, what constitutes the appropriate model for a phenomenon of interest is a subjective matter. Indeed there is no shortage of examples where decision makers and analysts disagree not only in their beliefs, but also in the scientific or statistical models they deem relevant. In such situations, an analyst does typically not have access to the model the decision maker has in mind, and hence a method for eliciting such information is required.

In this paper we provide necessary and sufficient conditions for a preference over Anscombe-Aumann acts to admit an identifiable smooth representation. We also show that all elements of the representation are uniquely determined from preferences: in particular, the prior μ and the domain $\mathcal{P}$ are unique, up to null events.

In arriving at these results, the key step is to determine the exact behavioral counterpart of the identifying statistic k. We show that the σ-algebra generated by k, i.e. the missing information generating ambiguity, is in fact equal to the collection of events that satisfy the sure-thing principle, an object defined purely in terms of the agent's preferences. This collection of events was introduced by Gul and Pesendorfer (2014) to study a different class of preferences. A crucial result they discovered, and that we use in an essential way in this paper, is that under mild assumptions such a collection is a σ-algebra.

An additional contribution of our work is to give operational meaning to the subjective statistical model $\mathcal{P}$. From the decision maker's preference relation $\succsim$ over acts, we derive a new subrelation $\succsim_{\text{st}}$ defined by $f \succsim_{\text{st}} g$ if f is preferred to g conditional on every event that satisfies the sure-thing principle. This preference is incomplete, and we show it characterizes $\mathcal{P}$: given two acts f and g, $\int_\Omega u(f)\,\mathrm{d}p \geq \int_\Omega u(g)\,\mathrm{d}p$ holds for every $p \in \mathcal{P}$ if and only if $f \succsim_{\text{st}} g$.

Relaxing identifiability leads to a new class of preferences, which we fully characterize. We study decision makers for which the true law governing the states cannot be exactly recovered from observations: Given the realized state ω, they can infer a set $K(\omega)$ of possible laws, but remain agnostic about what law within the set generated the data. Related ideas have been explored by the econometric literature on *partial identification*

(Manski, 1989; Tamer, 2010; Molinari, 2019). We show that these ideas can be captured by the criterion

$$V(f) = \int_{\mathcal{C}} \phi\Big(\min_{p \in C} \int_{\Omega} u(f)\,\mathrm{d}p\Big)\,\mathrm{d}\mu(C) \tag{2}$$

where $\mathcal{C}$ is a collection of convex sets of probability measures, and μ is a prior over $\mathcal{C}$. Smooth-ambiguity preferences can be seen as a special case where each set C is a singleton, while the maxmin preferences of Gilboa and Schmeidler (1989) are a special case where the collection $\mathcal{C}$ consists of a single set.

Before moving on to our formal analysis, we discuss two alternative axiomatic approaches to smooth ambiguity preferences. Klibanoff, Marinacci, and Mukerji (2005) provide an axiomatic foundation for the smooth ambiguity representation by studying preferences over second-order acts. These are acts whose outcomes depend on the correct probability p over the states. As discussed in their paper, a key difficulty is that choices over second-order acts are, in general, not directly observable. An alternative approach is taken in Seo (2009), who considers Anscombe and Aumann's original framework with two stages of objective randomization. An important feature of Seo's representation theorem is that it imposes no restrictions on the domain $\mathcal{P}$. At the same time, in his approach decision makers can display sensitivity to ambiguity only if they fail to reduce objective compound lotteries.[1] By contrast, the primitive of our analysis is a preference relation over what are by now standard Anscombe-Aumann acts. This puts identifiable smooth ambiguity preferences on the same ground of the other main classes of ambiguity preferences, whose leading characterizations are consistent with reduction of compound lotteries.

Identifiable smooth-ambiguity preferences are based on a formal distinction between uncertainty about events and uncertainty about the odds that govern them. This long-standing idea is critical in many fields. Wald (1950) distinguishes between uncertainty about the sample realization and uncertainty about the parameter generating the sample. In robust mechanism design, Bergemann and Morris (2005) make a distinction between uncertainty about what signals players will observe, and uncertainty about the underlying information structure. In macroeconomics, Hansen and Sargent (2008) distinguish between uncertainty within a model and about the correct model.

2 Preliminary definitions

We consider a set Ω of *states of the world*, a σ-algebra $\mathcal{F}$ of subsets of Ω called *events*, and a set X of *consequences*. We assume that X is a convex subset of a Hausdorff topological vector space, endowed with the Borel σ-algebra. This is the case in the classic setting of Anscombe and Aumann (1963) where X is the set of simple lotteries on a fixed set of

[1] Related models of second-order expected utility have been studied by Segal (1987), Davis and Pate-Cornell (1994), Nau (2006), and Ergin and Gul (2009).

prizes. We also assume that $(\Omega, \mathcal{F})$ is a standard Borel measurable space, i.e. there exists a Polish topology on Ω such that $\mathcal{F}$ is the corresponding Borel σ-algebra, an assumption that covers most measurable spaces used in applications.

An *act* is a measurable function $f\colon \Omega \to X$. We consider the domain $\mathfrak{F}$ of acts f for which there exists a finite set $Y \subseteq X$ such that f takes values in the convex hull of Y (i.e., the image $f(\Omega)$ is included in a polytope). In particular, $\mathfrak{F}$ contains all acts whose range is finite. Our main object of study is a binary relation $\succsim$ over $\mathfrak{F}$ that represents the preferences of the decision marker. We denote by $\sim$ and $\succ$ the symmetric and asymmetric parts of $\succsim$, respectively.

We write x for the constant act f such that $f(\omega) = x$ for all $\omega \in \Omega$. Given $f, g \in \mathfrak{F}$ and $\alpha \in [0,1]$, we denote by $\alpha f + (1-\alpha)g$ the act in $\mathfrak{F}$ that takes value $\alpha f(\omega) + (1-\alpha)g(\omega)$ in state ω. Given acts f and g and event A, fAg is the act that coincides with f on A and with g on A^c.

An event A is *null* if $fAh \sim gAh$ for all $f, g, h \in \mathfrak{F}$. Two events A, B are *equivalent up to a null event* if their symmetric difference $A \triangle B$ is null. Two σ-algebras $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ are *equivalent up to null events* if for every $A \in \mathcal{G}$ there is a $B \in \mathcal{H}$ such that $A \triangle B$ is null, and for every $B \in \mathcal{H}$ there is a $A \in \mathcal{G}$ such that $A \triangle B$ is null.

We denote by $\Delta(\mathcal{F})$, or simply Δ, the space of countably additive probability measures on $(\Omega, \mathcal{F})$. Given $p \in \Delta$, the symbol E_p denotes the corresponding expectation operator. Two σ-algebras $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ are *p-equivalent* if for every $A \in \mathcal{G}$ there is a $B \in \mathcal{H}$ such that $p(A \triangle B) = 0$, and for every $B \in \mathcal{H}$ there is a $A \in \mathcal{G}$ such that $p(A \triangle B) = 0$.

We endow Δ with the weak* topology[2] and the corresponding Borel σ-algebra. This is the σ-algebra Σ generated by the functions $p \mapsto p(A)$ for $A \in \mathcal{F}$. Given a nonempty set $\mathcal{P} \subseteq \Delta$, let $\Sigma_{\mathcal{P}} = \{S \cap \mathcal{P} : S \in \Sigma\}$ be the relative σ-algebra. A *prior* on $\mathcal{P}$ is a countably additive probability measure μ on $(\Sigma_{\mathcal{P}}, \mathcal{P})$. To each prior μ we associate the *predictive* probability $\pi_\mu \in \Delta$ defined as

$$\pi_\mu(A) = \int_{\mathcal{P}} p(A)\, \mathrm{d}\mu(p).$$

for every event $A \in \mathcal{F}$.

3 Identifiable smooth representation

We begin with the formal definition of smooth ambiguity representation:[3]

Definition 1. A tuple $(u, \phi, \mathcal{P}, \mu)$ is a *smooth ambiguity representation* of a preference relation $\succsim$ if $u\colon X \to \mathbb{R}$ is a non-constant affine function, $\phi\colon u(X) \to \mathbb{R}$ a strictly increasing

[2] Recall that in this topology a net (p_α) in Δ converges to p if and only if $p_\alpha(A) \to p(A)$ for all $A \in \mathcal{F}$.

[3] Klibanoff, Marinacci, and Mukerji (2005) take as a primitive a preference relation over Savage acts defined over $\Omega \times [0,1]$, where $[0,1]$ is endowed with the Lebesgue measure and plays the role of a randomization device. Definition 1 translates their representation to the Anscombe-Aumann setting.

continuous function, $\mathcal{P} \subseteq \Delta$ a nonempty set, and μ a non-atomic prior on $\mathcal{P}$, such that

$$f \succsim g \quad \iff \quad \int_{\mathcal{P}} \phi\left(\int_{\Omega} u(f)\,\mathrm{d}p\right)\mathrm{d}\mu(p) \geq \int_{\mathcal{P}} \phi\left(\int_{\Omega} u(g)\,\mathrm{d}p\right)\mathrm{d}\mu(p)$$

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

We interpret each $p \in \mathcal{P}$ as a possible law, or probabilistic model, governing the state. The domain $\mathcal{P}$ can then be seen as a subjective statistical model. The agent's degree of confidence over different laws is expressed by a prior μ, while the functions u and ϕ reflect risk and ambiguity attitude, respectively.

We focus on representations where the set $\mathcal{P}$ is, at least in principle, identifiable from observations:

Definition 2. A nonempty set $\mathcal{P} \subseteq \Delta$ is *identifiable* if there exists a measurable function $k\colon \Omega \to \mathcal{P}$, or *kernel*, such that for all $p \in \mathcal{P}$

$$p(\{\omega : k(\omega) = p\}) = 1.$$

A smooth representation $(u, \phi, \mathcal{P}, \mu)$ is *identifiable* if the set $\mathcal{P}$ is identifiable.

The condition of identifiability makes concrete the interpretation of $\mathcal{P}$ as a statistical model. In statistical terms, Definition 2 amounts to the common assumption of $\mathcal{P}$ being point-identified: there exists a function k of the state that enables the decision maker to infer the true law p, almost surely.[4]

By varying the class $\mathcal{P}$ we obtain a number of canonical examples.

Example 1. The state space $\Omega = S^\infty$ is the product of infinitely many copies of a finite set S. The statistical model $\mathcal{P}$ is the set of i.i.d. probability distributions, represented as $\Delta(S)$. By the strong law of large numbers, the collection $\mathcal{P}$ is identified by a kernel $k\colon \Omega \to \Delta(S)$ where $k(\omega, s)$ is the limiting empirical frequency of the outcome s along the sequence $\omega = (\omega_1, \omega_2, \ldots)$ of realizations, whenever it is well-defined.

The logic in the previous example extends to any environment $\mathcal{P}$ for which appropriate laws of large numbers can be applied to recover the true law from empirical frequencies. A common example from macroeconomics is an economy where the state of fundamentals, consisting of aggregate and idiosyncratic shocks, follow a stochastic process p that is stationary and ergodic (e.g., a moving-average process or an autoregressive process without unit root). Another example is a portfolio selection problem with uncertainty about expected returns, variance, and covariances (see, e.g., Garlappi, Uppal, and Wang, 2006).

[4] Lehmann and Casella (2006) and Lewbel (2019) describe point identification in statistics and econometrics. Our definition of identifiability agrees with mainstream econometric usage in most standard settings (see, e.g., Example 1).

Definition 2, however, is not tied to the interpretation of probability models as empirical frequencies, nor it is limited to environments characterized by repetitions. More broadly, identifiability formalizes the common view that ambiguity is due to lack of information. The σ-algebra generated by the identifying kernel k, which we denote by $\sigma(k)$, is the information that would allow the decision maker to resolve their uncertainty about the correct probabilistic model.

Example 2. Consider a parametric statistical model $\mathcal{P} \subseteq \Delta(\Omega)$, defined by n measurable functions $\theta_i \colon \Omega \to \mathbb{R}$ and a measurable map $\varphi \colon \mathbb{R}^n \to \mathcal{P}$ such that

$$\mathcal{P} = \{\varphi\left(\theta_1(\omega), \ldots, \theta_n(\omega)\right) : \omega \in \Omega\}.$$

The unknown law p governing the states is a function of n parameters $\theta_1, \ldots, \theta_n$, according to a functional form that is described by the map φ. The σ-algebra generated by the functions $\theta_i \colon \Omega \to \mathbb{R}$ is the missing information that generates ambiguity

The parametric domain $\mathcal{P}$ is by construction identifiable and can describe environments without obvious repetitions or symmetries. For a concrete illustration, consider a forecaster who is uncertain about the outcome of the next presidential election. In this environment, a state ω might specify the outcome of the election, the voter turnout, the incumbent approval rate, the stock market index, among other factors. The parameters $\theta_1, \ldots, \theta_n$ can be interpreted as those factors the forecaster deems relevant for producing a precise probabilistic assessment. The function φ reflects her subjective judgment of how the different factors interact with one another and with the outcome of the election. For instance, the forecaster might believe that a higher voter turnout is negatively correlated with the job approval rate, and positively correlated with the event that a Democrat will be elected.

As an another example, consider a physician choosing between different treatments for a given patient. A state ω represents a vector of patient's characteristics the physician is uncertain about, such as their biological response to different drugs, or whether they have a history of substance abuse. The parameters $\theta_1, \ldots, \theta_n$ summarize the information that, if known, would allow the physician to formulate a probabilistic assessment and confidently prescribe a specific treatment. Different physicians might disagree about what constitutes relevant information. A physician might divide patients into a small number of "types" defined by a restricted number of summary statistics, while another physician might adopt a finer parametrization. There may be also disagreement on the relative importance of the different types, as reflected by the subjective nature of the function φ.

In the classic Ellsberg thought experiment, the missing information is the composition of the urns. The formulation of the Ellsberg setting in the next example was introduced by Klibanoff, Marinacci, and Mukerji (2012) in their reply to Epstein (2010).

Example 3. A ball is drawn from an urn that contains red, blue, and yellow balls. The composition of the urn is unknown, but is verifiable ex post. A state of the world $\omega = (c, \gamma)$ specifies the color of the extracted ball $c \in \{r, b, y\}$ and the composition of the urn $\gamma \in \Delta(\{r, b, y\})$. The set of probabilistic laws $\mathcal{P} = \{p_\gamma\}$ is indexed by the composition γ, and each p_γ assigns probability 1 to the event $\{r, b, y\} \times \{\gamma\}$.

The composition of the urn is the missing information that generates ambiguity. The obvious identifying kernel $k \colon \Omega \to \Delta$ is given by $k((c, \gamma), \omega) = p_\gamma$ and simply reports the composition of the urn.

As the previous example suggests, the set $\mathcal{P}$ is identifiable whenever each $p \in \mathcal{P}$ can be seen as the realization of a random variable that is unknown to the agent at the time of the decision, but that is verifiable at a future date.

3.1 Predictive representation

The interpretation of ambiguity as lack of information is emphasized by the following alternative representation for smooth-identifiable preferences:

Definition 3. A tuple $(u, \phi, \mathcal{G}, \pi)$ is a *predictive representation* of a preference relation $\succsim$ if $u \colon X \to \mathbb{R}$ is a non-constant affine function, $\phi \colon u(X) \to \mathbb{R}$ a strictly increasing continuous function, $\mathcal{G} \subseteq \mathcal{F}$ a σ-algebra, and $\pi \in \Delta$ a probability measure non-atomic on $\mathcal{G}$ such that

$$f \succsim g \quad \iff \quad E_\pi\Big[\phi\Big(E_\pi[u(f)|\mathcal{G}]\Big)\Big] \geq E_\pi\Big[\phi\Big(E_\pi[u(g)|\mathcal{G}]\Big)\Big]$$

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

In this representation, the agent is able to form a unique probability assessment π but is not confident about such a prediction. The sub σ-algebra $\mathcal{G}$ represents the additional information the agent would need in order to arrive at a reliable probability assessment. Given knowledge of $\mathcal{G}$, acts would be ranked according to their conditional expected utility $E_\pi[u(f)|\mathcal{G}]$. As shown by the next result, the predictive and the smooth-identifiable representations characterize the same class of preferences.

Proposition 1. *(i). If $\succsim$ admits an identifiable representation $(u, \phi, \mathcal{P}, \mu)$, then it admits a predictive representation $(u, \phi, \sigma(k), \pi_\mu)$ where k is a kernel that identifies $\mathcal{P}$.*

(ii). If $\succsim$ admits a predictive representation $(u, \phi, \mathcal{G}, \pi)$, then it admits an identifiable representation $(u, \phi, \mathcal{P}, \mu)$ where $\pi_\mu = \pi$ and $\mathcal{G}$ is equivalent to $\sigma(k)$ up to null events.

By relating the probability π to the measure π_μ induced by the prior μ, the proposition reinforces the interpretation of π as a predictive probability. The result ties together the σ-algebras $\mathcal{G}$ and $\sigma(k)$ as missing information. Particular instances of the predictive representation have already appeared in the literature:

Example 4. *(Second-Order Expected Utility).* Let $(u, \phi, \mathcal{G}, \pi)$ be a predictive representation where $\mathcal{G} = \mathcal{F}$. Then

$$f \succsim g \quad \Longleftrightarrow \quad E_\pi [\phi(u(f))] \geq E_\pi [\phi(u(g))].$$

This criterion for decision making under ambiguity was introduced by Neilson (2010). A special case are the *multiplier preferences* of Hansen and Sargent (2001), as shown by Strzalecki (2011).

Example 5. *(Source-Dependent Preferences).* Two sources of uncertainty a and b are represented by probability spaces $(\Omega_a, \mathcal{F}_a, \pi_a)$ and $(\Omega_b, \mathcal{F}_b, \pi_b)$. A state of the world $\omega = (\omega_a, \omega_b)$ specifies a realization for each source, and $\mathcal{F} = \mathcal{F}_a \times \mathcal{F}_b$ is the product σ-algebra. Nau (2006) and Ergin and Gul (2009) study preferences where acts are evaluated separately along each source. An important special case of their analysis is the representation

$$V(f) = \int_{\Omega_b} \phi\Big(\int_{\Omega_a} u(f(\omega_a, \omega_b))\, \mathrm{d}\pi_a(\omega_a) \Big)\, \mathrm{d}\pi_b(\omega_b).$$

This corresponds to a predictive representation with product measure $\pi = \pi_a \times \pi_b$ and sub σ-algebra $\mathcal{G} = \{\Omega_a \times B : B \in \mathcal{F}_b\}$.

4 Axioms

We begin by imposing three elementary assumptions on $\succsim$. In addition to completeness and transitivity, we require $\succsim$ to be monotone and to satisfy a standard continuity condition. In what follows, we call a sequence (f_n) of acts *bounded* if there exists a finite set $Y \subseteq X$ such that each f_n takes values in the convex hull of Y.

Axiom 1. *The preference $\succsim$ is complete, transitive, and nontrivial.*

Axiom 2. *If $f(\omega) \succ g(\omega)$ for all ω, then $f \succ g$.*

Axiom 3. *If (f_n) and (g_n) are bounded sequences that converge pointwise to f and g, respectively, and $f_n \succsim g_n$ for every n, then $f \succsim g$.*

It is a crucial insight due to Ellsberg (1961) that departures from Savage's sure-thing principle are key manifestations of ambiguity. We say that an event A *satisfies the sure-thing principle* if, for all $f, g, h, h' \in \mathfrak{F}$, the following conditions are satisfied:

(i). If $fAh \succsim gAh$, then $fAh' \succsim gAh'$.

(ii). If $hAf \succsim hAg$, then $h'Af \succsim h'Ag$.

In words, an event A satisfies the sure-thing principle if both A and its complement satisfy Savage's P2 axiom. We denote by $\mathcal{F}_{\text{st}}$ the family of all such events. The properties of $\mathcal{F}_{\text{st}}$ were first studied by Gul and Pesendorfer (2014) under the name of *ideal* events.

Following Ghirardato, Maccheroni, and Marinacci (2004), we say that an act f is *unambiguously preferred* to g if $f \succsim g$ and the ranking is preserved across mixtures:

$$f \succsim^* g \quad \text{if} \quad \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h \quad \text{for all} \quad \alpha \in [0,1], h \in \mathfrak{F}.$$

The relation $\succsim^*$ isolates those choices that cannot be reversed by mixing with a common act h. A key decision-theoretic insight, due to Schmeidler (1989), is that such preference reversals are characteristic of an agent who perceives ambiguity, as mixing with h may allow to hedge against the uncertainty connected with f and g.

We can now state our main axiom. For every non-null event $A \in \mathcal{F}_{\text{st}}$, we define the conditional preference relation $\succsim_A$ by $f \succsim_A g$ if $fAh \succsim gAh$ for some h. Since A satisfies the sure-thing principle, $\succsim_A$ is well defined and the choice of h is inessential.

Axiom 4. *If $f \succsim_A g$ for all non-null $A \in \mathcal{F}_{\text{st}}$, then $f \succsim^* g$.*

The axiom relates mixture independence to the sure-thing principle. Recall that under subjective expected utility, a preference $f \succsim g$ implies the unambiguous ranking $f \succsim^* g$. Axiom 4 is more permissive: the conclusion that f is unambiguously preferred to g is reached under the premise that f is preferred to g conditional on every event that satisfies the sure-thing principle.

The two final axioms correspond to Savage's postulates P4 and P6, but applied to events that satisfy the sure-thing principle, as in Gul and Pesendorfer (2014). Because the meaning of these conditions is well understood, we do not elaborate further on them.

Axiom 5. *If $A, B \in \mathcal{F}_{\text{st}}$ and $x, y, z, w \in X$ are such that $x \succ y$ and $w \succ z$, then*

$$xAy \succ xBy \quad \Rightarrow \quad wAz \succ wBz.$$

Axiom 6. *For all acts f, g, h that are $\mathcal{F}_{\text{st}}$-measurable, if $f \succ g$ then there is a partition $A_1, \ldots, A_n$ of events in $\mathcal{F}_{\text{st}}$ such that $hA_if \succ g$ and $f \succ hA_ig$ for all i.*

4.1 Discussion

We now discuss more in detail the interpretation of Axiom 4, our main axiom. As is well known, the unambiguous preference relation $\succsim^*$ admits the representation

$$f \succsim^* g \quad \iff \quad \int_\Omega u(f)\,\mathrm{d}\pi \geq \int_\Omega u(g)\,\mathrm{d}\pi \quad \text{for all } \pi \in C^*. \tag{3}$$

where u is an affine utility function, and C^* is a set of probabilities over $(\Omega, \mathcal{F})$.[5] When the set C^* is not a singleton, the agent is unable or unwilling to formulate a single probabilistic

[5] See Bewley (2002), Ghirardato, Maccheroni, and Marinacci (2004) and Ghirardato and Siniscalchi (2012). In this paper, the von Neumann-Morgenstern independence axiom for constant acts is an immediate implication of Axioms 1-4.

assessment π under which to evaluate acts according to expected utility. We call *predictive uncertainty* the indeterminacy described by the multiplicity of probabilities in C^*.

Axiom 4 describes a rationale for the unambiguous ranking of two acts. A common strategy for simplifying complex decision problems consists in first isolating a set of hypotheses, and then drawing conclusions by evaluating the available options conditional on each hypothesis. We formalize this form of case-by-case reasoning by interpreting each $A \in \mathcal{F}_{\text{st}}$ as a different hypothesis about the state of the world entertained by the decision maker. What makes this interpretation suggestive is the defining feature of the collection $\mathcal{F}_{\text{st}}$. That is, the fact that each event $A \in \mathcal{F}_{\text{st}}$ and its complement (i.e., the alternative hypothesis) induce a well-defined conditional preference consistent with $\succsim$.

Under this interpretation, Axiom 4 states that uncertainty about the correct hypothesis $A \in \mathcal{F}_{\text{st}}$ is a key determinant of predictive uncertainty: if contingent on every event $A \in \mathcal{F}_{\text{st}}$ the act f is preferred to g, then, according to the axiom, predictive uncertainty plays no role in ranking the two acts.

The axiom allows to describe the behavioral implications of identifiability, within the class of preferences that admit a smooth representation. In Section 8.1 we present a preference relation that admits a smooth representation (Definition 1) but does not satisfy Axiom 4. As our representation theorem will show, this means that the preference does not admit a smooth-identifiable representation.

5 Representation theorem

Theorem 1. *A preference relation $\succsim$ satisfies Axioms 1-6 if and only if it admits an identifiable smooth representation $(u, \phi, \mu, \mathcal{P})$.*

The theorem provides a behavioral foundation for the class of identifiable smooth representations. By Proposition 1 the preference $\succsim$ admits an identifiable representation $(u, \phi, \mu, \mathcal{P})$ if and only if it admits a predictive representation $(u, \phi, \mathcal{G}, \pi)$. Thus Axioms 1-6 also characterize the predictive representation. Next we describe the uniqueness properties of the representations.

Theorem 2. *Two identifiable representations $(u_1, \phi_1, \mathcal{P}_1, \mu_1)$ and $(u_2, \phi_2, \mathcal{P}_2, \mu_2)$ of the same preference $\succsim$ are related by the following conditions:*

(i). *There are $a, c > 0$ and $b, d \in \mathbb{R}$ such that $u_2(x) = au_1(x) + b$ and $\phi_2(t) = c\phi_1(\frac{t-b}{a}) + d$ for all $x \in X$ and $t \in u_2(X)$.*

(ii). *$\pi_{\mu_1} = \pi_{\mu_2}$ and, provided that ϕ_1 is not affine, $\mu_1(\mathcal{P}_1 \cap S) = \mu_2(\mathcal{P}_2 \cap S)$ for all $S \in \Sigma$.*

If $(u_1, \phi_1, \mathcal{G}_1, \pi_1)$ and $(u_2, \phi_2, \mathcal{G}_2, \pi_2)$ are two predictive representations of $\succsim$, then (i) above holds, $\pi_1 = \pi_2$, and, provided that ϕ_1 is not affine, $\mathcal{G}_1$ and $\mathcal{G}_2$ are equivalent up to null events.

The agent's risk attitude, ambiguity attitude, and ambiguity perception are uniquely determined from their preferences: the utility function u and the ambiguity index ϕ are determined up to positive affine transformations, and the prior μ is unique. An obvious exception is the case in which ϕ is affine. If the agent is ambiguity neutral, then their perception of ambiguity is inessential and their behavior can reveal only the predictive probability π_μ. In this case, the relation $\succsim$ reduces to a subjective expected utility preference and the uniqueness of π_μ follows from Savage's theorem.

Analogous uniqueness properties hold for the predictive representation. The predictive measure π is unique and, provided that ϕ is not affine, the σ-algebra $\mathcal{G}$ is unique up to null events.

6 Model uncertainty

A key step in our analysis is the study of a new relation over acts derived from the agent's preferences. We define a relation $\succsim_{\text{st}}$ over acts by

$$f \succsim_{\text{st}} g \quad \text{if} \quad f \succsim_A g \quad \text{for all non-null} \quad A \in \mathcal{F}_{\text{st}}.$$

In words, $f \succsim_{\text{st}} g$ if f is preferred to g conditional on each event that satisfies the sure-thing principle. Following the discussion in Section 4.1, this reflects the idea that f is preferred to g conditional on each hypothesis $A \in \mathcal{F}_{\text{st}}$ entertained by the decision maker about the state of the world.

The next result is a representation theorem for $\succsim_{\text{st}}$.

Proposition 2. *Let $\succsim$ admit identifiable representation $(u, \phi, \mathcal{P}, \mu)$ and predictive representation $(u, \phi, \mathcal{G}, \pi)$. If ϕ is not affine, then the following are equivalent:*

(i) $f \succsim_{\text{st}} g$,

(ii) $\int_\Omega u(f)\,\mathrm{d}p \geq \int_\Omega u(g)\,\mathrm{d}p$ *for μ-almost all $p \in \mathcal{P}$,*

(iii) $E_\pi[u(f)|\mathcal{G}] \geq E_\pi[u(g)|\mathcal{G}]$.

The preference relation $\succsim_{\text{st}}$ describes a robust ranking over acts that is based on the set of probabilistic models p the agent considers plausible. The equivalence between (i) and (ii) shows that $f \succsim_{\text{st}} g$ holds exactly when model uncertainty does not affect the ranking of the two acts, since f is deemed better than g under each probabilistic model $p \in \mathcal{P}$, almost surely. The equivalence between (i) and (iii) is the natural counterpart for the predictive representation: the preference $f \succsim_{\text{st}} g$ holds when the missing information $\mathcal{G}$ does not affect the ranking of the two acts.[6]

[6]Proposition 2 assumes that ϕ is not affine. Otherwise, the agent is ambiguity neutral, $\succsim$ reduces to a subjective expected utility preference, all events satisfy the sure-thing principle, and $f \succsim_{\text{st}} g$ if and only if $f(\omega) \succsim g(\omega)$ for all $\omega \in \Omega$, up to a null event.

The next proposition shows that $\mathcal{F}_{\text{st}}$ can be interpreted as the missing information that generates ambiguity. In this context, we recall a result of Gul and Pesendorfer (2014): under broad conditions on $\succsim$ that are satisfied in this paper, the collection of events $\mathcal{F}_{\text{st}}$ is a σ-algebra.[7]

Proposition 3. *Let $\succsim$ admit identifiable representation $(u, \phi, \mathcal{P}, \mu)$ and predictive representation $(u, \phi, \mathcal{G}, \pi)$. If ϕ is not affine and k is a kernel that identifies $\mathcal{P}$, then the σ-algebras $\mathcal{F}_{\text{st}}$, $\sigma(k)$, and $\mathcal{G}$ are all equivalent up to null events.*

For a smooth identifiable representation, the collection of events that satisfy the sure-thing principle coincides, up to null events, with the σ-algebra generated by a kernel k that identifies $\mathcal{P}$. In the representation, knowledge of the value taken by k resolves the decision maker's uncertainty about the correct law $p \in \mathcal{P}$ governing the state. Hence, $\sigma(k)$ can be seen as the missing information that generates ambiguity. The proposition shows that $\mathcal{F}_{\text{st}}$ stands for the behavioral counterpart of this information. An analogous result holds for the predictive representation where $\mathcal{F}_{\text{st}}$ and $\mathcal{G}$ are equivalent up to null events.

7 Partial identification

In many contexts, statistical models are point-identified only under non-trivial assumptions over the data generating process. For this reason, in recent years the study of models that are not point-identified has received growing attention, starting with the work of Manski (1989) and in the subsequent literature on partial identification (Tamer, 2010). Partially identified models have been applied to the study of treatment effects, random choice, entry games, auctions, and network formation (Molinari, 2019, reviews the literature on partial identification).

A common theme in this body of work is that the assumptions necessary to deliver identification may vary in their plausibility. While for some there might be an established consensus, others might be tentative, leading to a tradeoff between identifiability and more robust conclusions. Lack of identifiability can also arise from the intrinsic inability of a theory to provide point predictions. A classic example is the multiplicity of equilibria in game theoretic models.

In this section we study partial identification from an axiomatic and behavioral perspective. Our focus is not on studying identifiability as an assumption for statistical inference, but rather as a principle adopted by a decision maker for reasoning about uncertainty. We introduce a new choice criterion for decision makers whose subjective statistical model $\mathcal{P}$ is not point identified, and which we characterize by a weakening of Axiom 4. Our characterization shows that the subjective statistical model $\mathcal{P}$ adopted by the decision maker is only partially identified when preferences display a novel type of hedging against ambiguity.

[7] See Lemma 16 in the appendix for a precise statement of this result.

7.1 Motivating example

The next example, in the spirit of Manski (1989), is a simple illustration of a statistical model that is partially identified.

Example 6. An entrepreneur is uncertain about the probability of success of her investment. The probability is correlated with the entrepreneur's personal characteristics, such as experience, skills, and cognitive abilities. For simplicity, we summarize these characteristics as a binary variable, that can be either "high" or "low."

The entrepreneur operates in a large market populated by infinitely many agents. The entrepreneur knows that her type is high, and knows the fraction $\alpha \in [0,1]$ of agents in the market that are of high type as well. For each agent i, the entrepreneur can observe ex post the outcome $y_i \in \{0,1\}$ of the investment, where $y_i = 1$ denotes a success. Personal characteristics of other agents might be difficult to infer or measure. We assume therefore that types are private information.

The conditional probabilities of success are

$$\theta_h = \text{Prob}(y_i = 1|\text{high}) \quad \text{and} \quad \theta_l = \text{Prob}(y_i = 1|\text{low}).$$

The pair $\theta = (\theta_h, \theta_l)$ is the unknown parameter of interest. For the entrepreneur, the probability of success equals θ_h, while for every other agent i the unconditional probability of success is

$$\overline{\theta} = \alpha\theta_h + (1-\alpha)\theta_l.$$

The state space is $\Omega = \{0,1\} \times \{0,1\}^\infty$. Each state $\omega = (y_e, y_1, y_2, \ldots)$ reports the entrepreneur's outcome y_e together with the realized sequence of outcomes $(y_1, y_2, \ldots)$ in the market. The statistical model $\mathcal{P} = \{p_\theta\}$ is indexed by $\theta = (\theta_h, \theta_l)$. Each $p_\theta \in \mathcal{P}$ is a probability measure on Ω that is independent across agents, and such that the probability of $y_e = 1$ is θ_h and the probability of $y_i = 1$ is $\overline{\theta}$ for all other agents i.

Upon observing a state ω, let $\ell(\omega)$ be the frequency of success in the economy. From this knowledge, every agent can infer the unconditional probability $\ell(\omega) = \overline{\theta}$. In turn, the entrepreneur can infer that the true law belongs to the set

$$K(\omega) = \{p_\theta \in \mathcal{P} : \overline{\theta} = \ell(\omega)\},$$

a compact convex set of probability measures. By knowing α, and upon inferring $\overline{\theta}$, the entrepreneur can conclude that the true value θ_h belongs to the interval

$$[0,1] \cap \left[\frac{1}{\alpha}\overline{\theta} - \frac{1-\alpha}{\alpha}, \frac{1}{\alpha}\overline{\theta}\right]$$

Additional uncertainty about α can enlarge this interval. Conversely, a decision maker who is willing to make further assumptions on the relation between θ_h and θ_l would be able to pin down a smaller set of potential true laws.

In the example, the set $\mathcal{P}$ is partially identifiable. Knowledge of the state ω does not allow to determine exactly the true law p governing the state, but pins down a set $K(\omega)$ of potential distributions. Following the literature in econometrics, we refer to the set $K(\omega)$ as the *identified set* at ω.

7.2 Set-identifiable models

In this section, we introduce a criterion of decision making for agents whose subjective statistical model $\mathcal{P}$ may not be point identifiable. We begin by describing the broader class of statistical models that we consider.

Let $\mathscr{C}$ be the collection of nonempty subsets of Δ that are convex and weak* compact. Each bounded measurable function $\xi\colon \Omega \to \mathbb{R}$ defines a *support function* mapping each $C \in \mathscr{C}$ to $\min_{p \in C} E_p[\xi]$. We endow $\mathscr{C}$ with the σ-algebra generated by all support functions; we denote it by $\mathfrak{S}$.

Definition 4. A set $\mathcal{P} \subseteq \Delta$ has *compact and convex identified sets* if there exists a measurable function $K\colon \Omega \to \mathscr{C}$ such that

(i). For all $\omega \in \Omega$, $K(\omega) \subseteq \mathcal{P}$;

(ii). For all $p \in \mathcal{P}$, $p(\{\omega : p \in K(\omega)\}) = 1$;

(iii). For all $\omega \in \Omega$ and $p \in K(\omega)$, $p(\{\omega' : K(\omega') = K(\omega)\}) = 1$.

The set-valued kernel K associates to each realization ω a set $K(\omega)$ of probability laws the decision maker deems compatible with the observed evidence. We refer to $K(\omega)$ as the *identified set* at ω. When (i)-(iii) are satisfied, we say that K *set-identifies* $\mathcal{P}$. We focus on the case in which the case in which K takes convex and compact values, as in Example 6.[8]

Assumption (i) is the natural requirement that candidate laws belong to the statistical model $\mathcal{P}$ adopted by the decision maker. Assumption (ii) guarantees that the true law p is correctly identified as a possible law governing the states. Condition (iii) guarantees that every law $p \in K(\omega)$ is equal to its conditional probability $p(\cdot|K(\omega))$. This means that all the information that led to the identified set $K(\omega)$ is already included in the description of p. By requiring (ii) and (iii) to hold exactly, rather than only approximately, we abstract from additional difficulties such as sampling or measurement errors, which occur with limited data but vanish asymptotically.

For a concrete illustration, observe that (i)-(iii) are satisfied by the set-valued kernel of Example 6. In addition, if $\mathcal{P}$ is identifiable as in Definition 2, then the identifying kernel $k\colon \Omega \to \mathcal{P}$ can be seen as set-valued kernel $K\colon \Omega \to \mathscr{C}$ such that $K(\omega) = \{k(\omega)\}$.

[8]As we show in the appendix (Lemma 7), measurability of the set-valued kernel K implies that the events appearing in (ii) and (iii) are measurable with respect to $\mathcal{F}$, and thus the conditions are well defined.

7.3 Set-identifiable smooth representation

In what follows, it will be without loss of generality to focus directly on the collection $\{K(\omega) : \omega \in \Omega\}$ of identified sets, without keeping track of the underlying statistical model $\mathcal{P}$ and of the identifying kernel K. Identified sets constitute a collection of convex and compact subsets of $\Delta(\Omega)$, as we define next.

Definition 5. A collection $\mathcal{C} \subseteq \mathscr{C}$ is a *collection of identified sets* if there is a set $\mathcal{P} \subseteq \Delta$ and a measurable function $K : \Omega \to \mathscr{C}$ that set-identifies $\mathcal{P}$, such that

$$\mathcal{C} = \{K(\omega) : \omega \in \Omega\}. \tag{4}$$

We can now present our generalization of identifiable smooth-ambiguity preferences. In the next definition, given a subcollection $\mathcal{C}$ of $\mathscr{C}$, we endow $\mathcal{C}$ with the relative σ-algebra $\mathfrak{S}_{\mathcal{C}} = \{\mathcal{S} \cap \mathcal{C} : \mathcal{S} \in \mathfrak{S}\}$, and call a *prior on* $\mathcal{C}$ a probability measure on $(\mathcal{C}, \mathfrak{S}_{\mathcal{C}})$.

Definition 6. A tuple $(u, \phi, \mathcal{C}, \mu)$ is a *set-identifiable smooth representation* of a preference $\succsim$ if $u : X \to \mathbb{R}$ is a non-constant affine function, $\phi : u(X) \to \mathbb{R}$ a strictly increasing continuous function, $\mathcal{C} \subseteq \mathscr{C}$ a collection of identified sets, and μ a nonatomic prior on $\mathcal{C}$, such that

$$f \succsim g \iff \int_{\mathcal{C}} \phi\Big(\min_{p \in C} \int_{\Omega} u(f)\,\mathrm{d}p\Big)\,\mathrm{d}\mu(C) \geq \int_{\mathcal{C}} \phi\Big(\min_{p \in C} \int_{\Omega} u(f)\,\mathrm{d}p\Big)\,\mathrm{d}\mu(C).$$

An act f is first evaluated by the maxmin criterion with respect to each identified set $C \in \mathcal{C}$. These values are then averaged by means of a prior μ over identified sets and an increasing transformation ϕ. The point-identifiable smooth representation corresponds to the special case where $\mathcal{C} = \{\{p\} : p \in \mathcal{P}\}$ for some identifiable $\mathcal{P}$.

The smooth representation $(u, \phi, \mathcal{C}, \mu)$ distinguishes between two layers of ambiguity. It distinguishes between ambiguity about knowables, and ambiguity that cannot be resolved. The identified set C is the value taken by the underlying kernel K, hence knowable. A different layer of ambiguity comes from uncertainty about the true law p within C. Uncertainty about the identified set is described by the prior μ, and the attitude towards this uncertainty is represented by the function ϕ. Attitude towards uncertainty within the identified set C is captured by the maxmin evaluation, as in Gilboa and Schmeidler (1989).

Maxmin principles, which have a long history in statistical decision theory, have received renewed attention in the econometrics of partial identification. For instance, maxmin principles have been applied to define robust point estimators for partially identified parameters (e.g., Manski, 2007, Song, 2014, Giacomini, Kitagawa, and Uhlig, 2019). Our representation reflects a similar attitude towards uncertainty.

The role played the two layers of uncertainty can be illustrated in the context of Example 6.

Example 6. *(Continued).* The collection $\mathcal{C}$ is the set of values taken by the identified set $K(\omega) = \{p_\theta : \overline{\theta} = \ell(\omega)\}$, and μ is the decision maker's belief on the distribution of K. In particular, the collection $\mathcal{C} = \{C_{\overline{\theta}}\}$ can be parametrized by the value taken by the unconditional probability of success $\overline{\theta}$.

Consider an act f that depends only on the average success rate $\ell(\omega) = \overline{\theta}$ in the market. For example, a bet on an aggregate market index. Because $\overline{\theta}$ is identifiable, f is evaluated according to the expectation

$$\int_{\mathcal{C}} \phi(f(\overline{\theta}))\, \mathrm{d}\mu(C_{\overline{\theta}}),$$

as in the standard smooth representation.

Now consider an act g that pays x_1 or x_0 depending on whether the entrepreneur's investment succeeds or fails. As discussed above, knowledge of $\overline{\theta}$ would allow the entrepreneur to conclude that θ_h belongs to the interval

$$J(\overline{\theta}) = [0, 1] \cap \left[\frac{1}{\alpha}\overline{\theta} - \frac{1-\alpha}{\alpha}, \frac{1}{\alpha}\overline{\theta}\right].$$

Hence, the act g is evaluated as

$$\int_{\mathcal{C}} \phi\left(\min_{\theta_h \in J(\overline{\theta})} \theta_h u(x_1) + (1-\theta_h)u(x_0)\right) \mathrm{d}\mu(C_{\overline{\theta}}).$$

The curvature of ϕ describes the decision maker's sensitivity to uncertainty about the knowable quantity $\overline{\theta}$. The inability to identify a single probability measure leads to an additional layer of ambiguity, captured by the minimization within the set $J(\overline{\theta})$.

7.4 Axioms and representation theorem

In this section we provide an axiomatic foundation of set-identifiable smooth representations. The foundation relies on the following weakening of Axiom 4:

Axiom 7. *If $f \succsim_{\mathrm{st}} g$ then $\alpha f + (1-\alpha)x \succsim \alpha g + (1-\alpha)x$ for all $\alpha \in [0,1]$ and $x \in X$.*

Axiom 8. *If $f \sim_{\mathrm{st}} g$ then $\alpha f + (1-\alpha)g \succsim f$ for all $\alpha \in [0,1]$.*

As discussed in Section 4.1, each event $A \in \mathcal{F}_{\mathrm{st}}$ can be seen as a different hypothesis about the state of the world entertained by the decision maker. Axioms 7 and 8 characterize an agent who is not confident that these hypotheses are an exhaustive description of the decision problem at hand. Even if f is preferred to g contingent on every event $A \in \mathcal{F}_{\mathrm{st}}$, it is still possible that the ranking of f and g remains ambiguous, and that the agent sees value in hedging. Axiom 8 reveals aversion to this residual ambiguity, following a logic analogous to Schmeidler (1989). Axiom 7 rules out hedging opportunities when mixing with a common constant act. The axiom plays a role similar to the certainty independence axiom in Gilboa and Schmeidler (1989).

The next theorem shows that Axioms 7 and 8 are exactly the weakening of Axiom 4 that characterizes set-identifiable smooth preferences.

Theorem 3. *A preference relation* $\succsim$ *satisfies Axioms 1-3 and 5-8 if and only if it admits a set-identifiable smooth representation* $(u, \phi, \mathcal{C}, \mu)$*.*

Next we describe the uniqueness properties of the representation:

Theorem 4. *Two set-identifiable representations* $(u_1, \phi_1, \mathcal{C}_1, \mu_1)$ *and* $(u_2, \phi_2, \mathcal{C}_2, \mu_2)$ *of the same preference* $\succsim$ *are related by the following conditions:*

(i). *There are* $a, c > 0$ *and* $b, d \in \mathbb{R}$ *such that* $u_2(x) = au_1(x) + b$ *and* $\phi_2(t) = c\phi_1(\frac{t-b}{a}) + d$ *for all* $x \in X$ *and* $t \in u_2(X)$*.*

(ii). *If* ϕ_1 *is not affine, then* $\mu_1(\mathcal{S} \cap \mathcal{C}_1) = \mu_2(\mathcal{S} \cap \mathcal{C}_2)$ *for all* $\mathcal{S} \in \mathfrak{S}$*.*

The functions u and ϕ are uniquely determined from the preference relation, up to positive affine transformation. In the case where ϕ is not affine, then the prior μ is unique.

8 Discussion

8.1 Axioms and non-identifiable smooth representations

The following example presents a preference relation that admits a smooth representation, but is incompatible with Axiom 4.

Example 7. Let Ω contain at least three distinct elements ω_1, ω_2, and ω_3. The decision maker is confident that the event $\Omega_1 = \{\omega_2, \omega_3\}^c$ has probability $\alpha \in (0, 1)$. She is also confident about the relative likelihood of the states in Ω_1, as described by a measure $q \in \Delta$ that assigns probability one to Ω_1. She is unsure, however, about the relative likelihood $\beta \in (0, 1)$ of the two remaining states ω_2 and ω_3. Overall, her uncertainty is described by a domain $\mathcal{P} = \{p_\beta : \beta \in (0, 1)\}$ such that

$$p_\beta = \alpha q + (1 - \alpha)(\beta\delta_{\omega_2} + (1 - \beta)\delta_{\omega_3})$$

where δ_ω is the Dirac measure concentrated on state ω. The domain $\mathcal{P}$ is not identifiable, as any event having probability 1 under a law p_β must also have probability 1 under any other law $p_{\beta'}$. So, no kernel can satisfy Definition 2. The domain $\mathcal{P}$ represents a situation where the decision maker does not believe her ambiguity about β can ever be resolved, even in the idealized framework of this paper where the state of the world can be observed without error. When Ω consists of exactly three states $\{\omega_1, \omega_2, \omega_3\}$, the example can be seen as a three-color Ellsberg urn whose composition is not verifiable (unlike Example 3).

For simplicity, assume that $X = \mathbb{R}$. Let $\succsim$ admit a smooth representation $(u, \phi, \mathcal{P}, \mu)$ such that u is the identity function, $\phi(x) = -e^{-x}$ for all $x \in \mathbb{R}$, and μ is uniform. The resulting preference is represented by the functional

$$V(f) = \int_0^1 \phi\left(\alpha E_q[f] + (1 - \alpha)(\beta f(\omega_2) + (1 - \beta)f(\omega_3))\right) \mathrm{d}\beta. \tag{5}$$

It can be shown that the events that satisfy the sure-thing principle are

$$\mathcal{F}_{\text{st}} = \{A \in \mathcal{F} : \{\omega_2, \omega_3\} \subseteq A \text{ or } \{\omega_2, \omega_3\} \subseteq A^c\}.$$

The claim is based on the following intuition. Consider an event $A \in \mathcal{F}$ such that either $\omega_2 \in A$ or $\omega_3 \in A$, but not both. As in the familiar Ellsberg paradox, ambiguity about the relative likelihood of ω_2 and ω_3 makes the ranking of an act fAh sensitive to the choice h. Thus, provided that $q(A) > 0$, the event A does not satisfy $P2$. Conversely, being ϕ exponential, it can be shown that A satisfies P2 whenever $\{\omega_2, \omega_3\} \subseteq A$. Indeed, evaluating the act fAh, the term $E_q[f \cdot 1_A]$ can be factored out of the integral in (5).

Axiom 4 does not hold. For instance, let f be a bet on ω_2 defined as $f(\omega_2) = 1$ and $f(\omega) = 0$ for $\omega \neq \omega_2$, and let g be a bet on ω_3 defined as $g(\omega_3) = 1$ and $g(\omega) = 0$ for $\omega \neq \omega_3$. On one hand, $f \sim_{\text{st}} g$. On the other hand, because ϕ is strictly concave,

$$V(f) = V(g) = \int_0^1 \phi\left((1-\alpha)\beta\right) \mathrm{d}\beta < \phi\left(\frac{1-\alpha}{2}\right) = V\left(\frac{1}{2}g + \frac{1}{2}f\right).$$

Thus the unambiguous ranking $f \sim^* g$ is not satisfied, and therefore Axiom 4 is violated.

8.2 Proof sketch

We now describe, rather informally, the main arguments used in the proofs of Theorems 1 and 3. In the proof of Theorem 1, sufficiency of the axioms is established according to the following steps.

Step 1. Consider a relation $\succsim$ that satisfies Axioms 1-6. The first four axioms imply that when restricted over X, the relation $\succsim$ satisfies the von Neumann-Morgenstern independence axiom. By standard arguments, there exists an affine utility function $u\colon X \to \mathbb{R}$ representing $\succsim$ on X, and any two acts satisfy $f \sim g$ whenever $u(f) = u(g)$. For expositional simplicity, we assume here that $X = \mathbb{R}$ and u is the identity. We can therefore identify the set of acts with the set $B(\mathcal{F})$ of real-valued, bounded, and $\mathcal{F}$-measurable functions.

A result due to Gul and Pesendorfer (2014) guarantees that the collection $\mathcal{F}_{\text{st}}$ is a σ-algebra. The Savage postulates are satisfied on $\mathcal{F}_{\text{st}}$-measurable acts, and there exist therefore a strictly increasing function $\phi\colon \mathbb{R} \to \mathbb{R}$ and a non-atomic probability measure $q : \mathcal{F}_{\text{st}} \to [0, 1]$ such that

$$f \succsim g \iff \int_\Omega \phi(f)\, \mathrm{d}q \geq \int_\Omega \phi(g)\, \mathrm{d}q \tag{6}$$

for all $\mathcal{F}_{\text{st}}$-measurable acts f and g.

Step 2. We next show that for every act $f \in B(\mathcal{F})$ there exists a $\mathcal{F}_{\text{st}}$-measurable act $\hat{f}$ that satisfies $f \sim_{\text{st}} \hat{f}$. To this end, define $V\colon B(\mathcal{F}) \to \mathbb{R}$ as $V(f) = \phi(c(f))$, where $c(f)$ is the certainty equivalent of f. The functional V represents $\succsim$.

Now given an act f, fix x such that $f(\omega) \succ x$ for every state ω, and consider the set function $q_f \colon \mathcal{F}_{\text{st}} \to \mathbb{R}$ given by $q_f(A) = V(fAx) - V(x)$. The crucial observation is that q_f is a positive measure, absolutely continuous with respect to q. By applying the Radon-Nikodym theorem, we can find a $\mathcal{F}_{\text{st}}$-measurable act $\hat{f}$ such that $\hat{f} \sim_{\text{st}} f$. See Lemma 19 in the appendix for more details.

Step 3. Two acts $\hat{f}$ and $\hat{f}'$ that are $\mathcal{F}_{\text{st}}$-measurable and satisfy $f \sim_{\text{st}} \hat{f} \sim_{\text{st}} \hat{f}'$ are equal q-almost surely. We can thus define an operator

$$T \colon B(\mathcal{F}) \to L_\infty(\mathcal{F}_{\text{st}}, q)$$

mapping each act f to the set $T(f)$ of $\mathcal{F}_{\text{st}}$-measurable acts $\hat{f}$ that satisfy $f \sim_{\text{st}} \hat{f}$. This set forms an indifference class in the L_∞ space defined by q. Moreover, the functional

$$V(f) = \int_\Omega \phi(T(f)) \, \mathrm{d}q$$

represents $\succsim$.

Step 4. We establish a number of properties for the operator T. We show T is monotone, normalized (i.e. $T(x) = x$ for every constant x), and is *decomposable*, that is $T(1_A f) = 1_A T(f)$ for every f and every event A in $\mathcal{F}_{\text{st}}$. In addition it represents $\succsim_{\text{st}}$, in the sense that $f \succsim_{\text{st}} g$ if and only if $T(f) \geq T(g)$ almost surely. Decomposable operators are discussed by Section B in the appendix. As we show, they are connected to the notion of rectangular sets of probability measures in Epstein and Schneider (2003). Related classes of operators, and their representations, play an important role in the theory of dynamic risk measures (Föllmer and Schied, 2011, Chapter 11).

Step 5. In the last step we study the implications of Axiom 4. Given the previous steps, we show that Axiom 4 holds if and only if the operator T is affine. Affinity of the operator implies, as we establish, the existence of a probability measure π extending q from $\mathcal{F}_{\text{st}}$ to the original σ-algebra $\mathcal{F}$, such that

$$T(f) = E_\pi[f|\mathcal{F}_{\text{st}}].$$

This leads to the predictive representation $V(f) = E_\pi[\phi(E_\pi[f|\mathcal{F}_{\text{st}}])]$. By Proposition 1, we conclude that $\succsim$ admits an identifiable smooth representation.

In the proof of Theorem 3 we show how weakening Axiom 4 leads to a broader class of preferences. Under Axioms 7 and 8, steps 1-4 above can be replicated without changes, but the operator T satisfies properties weaker than affinity. In particular, there exists a compact convex set $\Pi \subseteq \Delta$ of probability measures extending q such that T is the generalized conditional expectation operator

$$T(f) = \operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[f|\mathcal{F}_{\text{st}}].$$

Applied to V, this leads to a generalized predictive representation

$$V(f) = E_q\Big[\phi\Big(\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[f|\mathcal{F}_{\text{st}}]\Big)\Big]. \tag{7}$$

In the appendix we study in details the properties of this representation (Section D). The main result is an equivalence between (7) and the set-identifiable representation of Theorem 3. This equivalence is nontrivial, and is a based on a novel notion of regular conditional probability for set-valued kernels (Section B.3).

8.3 Relation with Epstein and Seo (2015)

Epstein and Seo (2015) provides a different decision-theoretic perspective on partial identification. In their setting, each state ω is defined by a sequence of draws $s_1, s_2, \ldots$ from a set S. The decision maker believes draws could be correlated, and contemplates different theories about the correlation structure. Each theory provides a potentially incomplete description of the data-generating process. A theory is represented by a set $\mathcal{P}_\theta \subseteq \Delta(\Omega)$ of probability measures, indexed by a structural parameter $\theta \in \Theta$. Uncertainty about the parameter is quantified by a prior μ.

Epstein and Seo (2015) consider settings in which every $\mathcal{P}_\theta$ corresponds to the core of a capacity ν_θ^∞ on Ω. The capacity ν_θ^∞ is the (suitably defined) i.i.d. product of a belief function ν_θ on S. An act f is evaluated according to the double integral

$$V(f) = \int_\Theta \left(\int_\Omega f \,\mathrm{d}\nu_\theta^\infty \right) \mathrm{d}\mu(\theta)$$

where the inner integral $\int_\Omega f \,\mathrm{d}\nu_\theta^\infty$ is a Choquet integral.

In Section 7 we develop a different decision-theoretic approach to partial identification. In particular, we borrow from the econometrics literature the notion of identified set, which is central to our analysis and does not have an immediate counterpart in Epstein and Seo (2015).[9] As an example, the collection $\{\mathcal{P}_\theta\}$ is not a collection of identified sets in the sense of our Definition 5. As Epstein and Seo (2015) point out, even after observing the entire sequence of draws $(s_1, s_2, \ldots)$ the decision maker may not be able to determine with certainty which theory to adopt, and this contradicts condition (iii) of our Definition 4.

Appendix

The appendix is organized as follows: In Section A we introduce the necessary notation and preliminary results; Section B introduces the notion of decomposable operator, which is then used in Section C to provide a representation for a preference relation that satisfies Axioms 1-3 and 5-6; Starting from this baseline representation, in Section D we show that a preference relation that satisfies Axioms 1-3 and 5-8 can be represented by a

[9] See also the discussion in Epstein, Kaido, and Seo (2016, p. 1805).

generalization of the predictive representation defined in the main text. We study the uniqueness properties of this representation as well as provide characterizations for $\mathcal{F}_{\text{st}}$ and $\succsim_{\text{st}}$; The analysis in Section D is applied in Sections E-**??** to prove the results stated in the main text.

A Preliminaries

A.1 Notation

For every σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ and nonempty interval $U \subseteq \mathbb{R}$, we denote by $B(\mathcal{G}, U)$ the space of $\mathcal{G}$-measurable bounded functions $\xi \colon \Omega \to \mathbb{R}$ taking values in U. As usual, we identify $a \in U$ with the constant function taking value a. We denote by $B_0(\mathcal{G}, U) \subseteq B(\mathcal{G}, U)$ the subspace of functions taking finitely many values, and let $B_b(\mathcal{G}, U) \subseteq B(\mathcal{G}, U)$ be the set of all $\xi \in B(\mathcal{G}, U)$ for which there exist $a, b \in U$ that satisfy $a \geq \xi \geq b$. A sequence (ξ_n) in $B_b(\mathcal{G}, U)$ is *bounded* if there are $a, b \in U$ such that $a \geq \xi_n \geq b$ for all n.

Let $q \in \Delta(\mathcal{G})$ be a probability measure. We denote by $L_\infty(\mathcal{G}, q)$ the space of equivalence classes of real-valued, $\mathcal{G}$-measurable, and almost-surely bounded functions. Given $\xi \in B_b(\mathcal{G})$ we denote by $[\xi] \in L_\infty(\mathcal{G}, q)$ the corresponding equivalence class. We refer to an element $\zeta \in [\xi]$ of the equivalence class as a *representative* of $[\xi]$. We denote by $L_\infty(\mathcal{G}, q, U) = \{[\xi] : \xi \in B_b(\mathcal{G}, U)\}$ the set of equivalence classes induced from functions in $B_b(\mathcal{G}, U)$. Given an increasing function $\phi \colon U \to \mathbb{R}$ and $\xi \in B_b(\mathcal{G}, U)$, we denote by $\phi([\xi])$ the equivalence class $[\phi(\xi)] \in L_\infty(\mathcal{G}, q)$.

Let $\Pi \subseteq \Delta$ be a set of probability measures that agree with q on $\mathcal{G}$. For every $\pi \in \Pi$, we denote by $E_\pi[\xi|\mathcal{G}] \in L_\infty(\mathcal{G}, q, U)$ the conditional expectation of $\xi \in B_b(\mathcal{F}, U)$ with respect to $\mathcal{G}$. We denote by $\operatorname{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}]$ and $\operatorname{ess\,sup}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}]$ the (essential) supremum and infimum, respectively, of $\{E_\pi[\xi|\mathcal{G}] : \pi \in \Pi\} \subseteq L_\infty(\mathcal{G}, q, U)$. We recall that $L_\infty(\mathcal{G}, q)$ has the countable sup property (see, e.g., Aliprantis and Border, 2006, page 326). This implies that for every $\xi \in B_b(\mathcal{F}, U)$ there exists a countable set $\{\pi_1, \pi_2, \ldots\} \subseteq \Pi$ such that

$$\operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}] = \operatorname*{ess\,inf}_{n} E_{\pi_n}[\xi|\mathcal{G}] \quad \text{and} \quad \operatorname*{ess\,sup}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}] = \operatorname*{ess\,sup}_{n} E_{\pi_n}[\xi|\mathcal{G}].$$

Since the set is countable, $\operatorname{ess\,inf}_n E_{\pi_n}[\xi|\mathcal{G}]$ is equal to the equivalence class $[\inf_n \zeta_n]$, where $\zeta_n \in E_{\pi_n}[\xi|\mathcal{G}]$ for every n. The same is true for $\operatorname{ess\,sup}_n E_{\pi_n}[\xi|\mathcal{G}]$.

A.2 Equivalent σ-algebras

Let $\Pi \subseteq \Delta(\mathcal{F})$ be a set of probability measures on $\mathcal{F}$. Two events $A, B \in \mathcal{F}$ are Π-*equivalent* if $\pi(A \triangle B) = 0$ for all $\pi \in \Pi$. Two σ-algebras $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ are Π-*equivalent* if every $G \in \mathcal{G}$ has a Π-equivalent $H \in \mathcal{H}$, and vice versa every $H \in \mathcal{H}$ has a Π-equivalent $G \in \mathcal{G}$. Two functions $\xi, \zeta \in B_b(\mathcal{F}, U)$ are Π-*equivalent* if they are equal π-almost surely for all $\pi \in \Pi$. The next lemma describes some basic properties of equivalent σ-algebras. We omit the simple proof.

Lemma 1. *If $\mathcal{G}$ and $\mathcal{H}$ are Π-equivalent, then the following conditions are satisfied:*

(i). *For every $\xi \in B_b(\mathcal{G}, U)$ there is a Π-equivalent $\zeta \in B_b(\mathcal{H}, U)$.*

(ii). *For each $\xi \in B_b(\mathcal{F}, U)$ and $\pi \in \Pi$, every $\zeta \in E_\pi[\xi|\mathcal{G}]$ has a Π-equivalent $\psi \in E_\pi[\xi|\mathcal{H}]$.*

(iii). *If all $\pi \in \Pi$ agree on $\mathcal{G}$, then they agree on $\mathcal{H}$.*

(iv). *If all $\pi \in \Pi$ agree on $\mathcal{G}$ and $\mathcal{H}$, then for each $\xi \in B_b(\mathcal{F}, U)$ every two representatives of $\operatorname{ess\,inf}_{\pi\in\Pi} E_\pi[\xi|\mathcal{G}]$ and $\operatorname{ess\,inf}_{\pi\in\Pi} E_\pi[\xi|\mathcal{H}]$ are Π-equivalent.*

A.3 Weak and weak* compactness

Let $ba(\mathcal{F})$ be the space of finitely-additive measures on $\mathcal{F}$ of bounded variation, endowed with the variation norm. The *weak topology* on $ba(\mathcal{F})$ is the weakest topology that makes continuous every bounded linear functional on $(ba(\mathcal{F}), \|\cdot\|_v)$. The *weak* topology* on $ba(\mathcal{F})$ is the weakest topology that makes continuous every bounded linear functional of the form $p \mapsto \int_\Omega \xi \,\mathrm{d}p$ for $\xi \in B(\mathcal{F}, \mathbb{R})$. We denote by $ba_1^+(\mathcal{F})$, the space of finitely additive probabilities. It is weak* compact.

The following result relates weak and weak* compactness for sets of σ-additive measures. It is proved in Maccheroni and Marinacci (2001). Let $ca(\mathcal{F}) \subseteq ba(\mathcal{F})$ be the space of σ-additive measures and $ca^+(\mathcal{F}) \subseteq ca(\mathcal{F})$ the subset of positive measures.

Lemma 2. *For a set $\Pi \subseteq ca(\mathcal{F})$, the following statements are equivalent:*

(i). Π *is weak* compact.*

(ii). Π *is weakly compact.*

(iii). Π *is bounded, weakly closed, and there exists $\lambda \in ca^+(\mathcal{F})$ such that for every $\epsilon > 0$ there exists $\delta > 0$ such that $\lambda(A) < \delta$ implies $\pi(A) < \epsilon$ for all $\pi \in \Pi$.*

(iv). Π *is bounded, weakly closed, and $A_n \downarrow \varnothing$ implies $\sup_{\pi\in\Pi} \pi(A_n) \to 0$.*

In addition, if $\Pi \in \mathscr{C}$, then λ in (iii) can be chosen to belong to Π.

We refer to a measure λ that satisfies (iii) as a *control measure* for Π. Let $ca(\mathcal{F}, \lambda) \subseteq ba(\mathcal{F})$ be the space of σ-additive measures that are absolutely continuous with respect to $\lambda \in ca^+(\mathcal{F})$. The Banach space $(ca(\mathcal{F}, \lambda), \|\cdot\|_v)$ is isometrically isomorphic to $L_1(\mathcal{F}, \lambda)$ (Aliprantis and Border, 2006, Theorem 13.19). Observe that λ is a control measure for Π if and only if $\Pi \subseteq ca(\mathcal{F}, \lambda)$ and the Radon-Nikodym derivatives $\{\frac{\mathrm{d}\pi}{\mathrm{d}\lambda} : \pi \in \Pi\} \subseteq L_1(\mathcal{F}, \lambda)$ are uniformly integrable. That is, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\lambda(A) < \delta$ implies $\int_A |\frac{\mathrm{d}\pi}{\mathrm{d}\lambda}| \,\mathrm{d}\lambda < \varepsilon$ for every $\pi \in \Pi$.

Lemma 3. *If $\Pi \subseteq ca(\mathcal{F})$ is weak* compact, then it is weak* separable.*

Proof. Let $\lambda \in ca^+(\mathcal{F})$ be a control measure for Π. Being $(\Omega, \mathcal{F})$ standard Borel, $\mathcal{F}$ is countably generated, and hence the space $L_1(\mathcal{F}, \lambda)$ is separable (Brezis, 2010, Theorem 4.13). Thus the Banach space $(ca(\mathcal{F}, \lambda), \|\cdot\|_v)$ is separable as well. This implies that Π is $\|\cdot\|_v$-separable (Aliprantis and Border, 2006, Corollary 3.5). Since the topology corresponding to $\|\cdot\|_v$ is stronger than the weak* topology, Π is weak*-separable. □

The following characterization of uniform integrability is due to de la Vallée Poussin (see, e.g., Diestel, 1991, Theorem 2).

Lemma 4. *For $\lambda \in ca^+(\mathcal{F})$ and $\Pi \subseteq ca(\mathcal{F}, \lambda)$, the Radon-Nikodym derivatives $\{\frac{\mathrm{d}\pi}{\mathrm{d}\lambda} : \pi \in \Pi\} \subseteq L_1(\mathcal{F}, \lambda)$ are uniformly integrable if and only if there is convex even function $\psi : \mathbb{R} \to \mathbb{R}$ such that $\psi(0) = 0$, $\lim_{t\to\infty} \frac{\psi(t)}{t} = \infty$, and*

$$\sup_{\pi\in\Pi} \int_\Omega \psi\left(\left|\frac{\mathrm{d}\pi}{\mathrm{d}\lambda}\right|\right) \mathrm{d}\lambda < \infty.$$

A.4 Support functions

For every $C \in \mathscr{C}$, we denote by $\sigma_C \colon B(\mathcal{F}, \mathbb{R}) \to \mathbb{R}$ the induced support functional $\sigma_C(\xi) = \min_{\pi\in C} E_\pi[\xi]$. By a standard separation argument, two sets in $C_1, C_2 \in \mathcal{C}$ satisfy $\sigma_{C_1} = \sigma_{C_2}$ if and only if $C_1 = C_2$. Moreover, it is immediate to verify that given an interval $U \subseteq \mathbb{R}$ of positive length, $\sigma_{C_1} = \sigma_{C_2}$ holds if and only if $\sigma_{C_1}(\xi) = \sigma_{C_2}(\xi)$ for all $\xi \in B_b(\mathcal{F}, U)$.

Lemma 5. *If (ξ_n) is bounded sequence in $B(\mathcal{F}, \mathbb{R})$ converging pointwise to ξ, then $\sigma_C(\xi_n) \to \sigma_C(\xi)$ for all $C \in \mathscr{C}$.*

Proof. It is enough to show that $E_p[\xi_n] \to E_p[\xi]$ uniformly in $p \in C$. Fix $\varepsilon > 0$ and for every n define the event $A_n = \{\omega : |\xi_m(\omega) - \xi(\omega)| \geq \varepsilon \text{ for some } m \geq n\}$. Since (ξ_n) converges pointwise to ξ, we have $A_n \downarrow \varnothing$. Moreover

$$\max_{p\in C} \left|\int_\Omega \xi_n - \xi \, \mathrm{d}p\right| \leq \sup_{\omega,n} |\xi_n(\omega) - \xi(\omega)| \cdot \max_{p\in C} p(A_n) + \varepsilon.$$

The converge of $\max_{p\in C} p(A_n)$ to zero follows from C being weak* compact (Lemma 2). Since ε is arbitrary we conclude that $E_p[\xi_n] \to E_p[\xi]$ uniformly in $p \in C$. □

Lemma 6. *There exist a countable set $\{\xi_1, \xi_2, \ldots\} \subseteq B(\mathcal{F}, \mathbb{R})$ such that for all $C_1, C_2 \in \mathscr{C}$, if $\sigma_{C_1}(\xi_n) \leq \sigma_{C_2}(\xi_n)$ for all n, then $C_1 \subseteq C_2$.*

Proof. Assume first Ω is uncountable. Being $(\Omega, \mathcal{F})$ standard Borel, by the Borel Isomorphism Theorem it is Borel isomorphic to $([0,1], \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra on $[0,1]$. Let Ξ be the space of continuous function $[0,1]$ and denote by $\Delta([0,1])$ the space of Borel probability measures $[0,1]$. We denote by $\hat{p}$ a generic element of $\Delta([0,1])$. Let $\hat{\tau}$ be the topology on $\Delta([0,1])$ induced by the maps $\hat{p} \mapsto E_{\hat{p}}[\hat{\xi}]$ where $\hat{\xi} \in \Xi$. As is well known, under this topology $\Delta([0,1])$ is compact.

Denote by τ the topology on $\Delta(\Omega, \mathcal{F})$ induced by the maps $p \mapsto E_p[\xi]$ where $\xi \in B(\mathcal{F}, \mathbb{R})$ (i.e. the weak* topology we defined in the main text). Recall that $\mathscr{C}$ is the collection of nonempty τ-compact convex subsets of $\Delta(\Omega, \mathcal{F})$.

Let $\vartheta\colon \Omega \to [0,1]$ be a Borel isomorphism. For every $p \in \Delta(\Omega, \mathcal{F})$ denote by $p \circ \vartheta^{-1}$ the pushforward of p under ϑ. For every $C \in \mathscr{C}$ define the set $C \circ \vartheta^{-1} = \{p \circ \vartheta^{-1} : p \in C\}$. It follows from convexity of C that $C \circ \vartheta^{-1}$ is convex. Because C is τ-compact, the set $C \circ \vartheta^{-1}$ is $\hat{\tau}$-closed, hence $\hat{\tau}$-compact (being that $\Delta([0,1])$ is $\hat{\tau}$-compact). To see this, let (p_α) be a net in C and suppose that $(p_\alpha \circ \vartheta^{-1})$ has $\hat{\tau}$-limit $\hat{p}$. Being that C is τ-compact, we can assume without loss of generality that (p_α) has τ-limit $p \in C$. For every $\hat{\xi} \in \hat{\Xi}$

$$\int_{[0,1]} \hat{\xi} \, \mathrm{d}\hat{p} = \lim_\alpha \int_{[0,1]} \hat{\xi} \, \mathrm{d}(p_\alpha \circ \vartheta^{-1}) = \lim_\alpha \int_\Omega \hat{\xi}(\vartheta) \, \mathrm{d}p_\alpha = \int_\Omega \hat{\xi}(\vartheta) \, \mathrm{d}p = \int_{[0,1]} \hat{\xi} \, \mathrm{d}(p \circ \vartheta^{-1}).$$

Thus $\hat{p} = p \circ \vartheta^{-1}$, which implies that $C \circ \vartheta^{-1}$ is $\hat{\tau}$-closed.

Let $\{\hat{\xi}_1, \hat{\xi}_2, \ldots\}$ be a countable supnorm-dense subset of Ξ. For every $C_1, C_2 \in \mathscr{C}$

$$\min_{\hat{p} \in C_1 \circ \vartheta^{-1}} \int_{\hat{\Omega}} \hat{\xi} \, \mathrm{d}\hat{p} = \min_{\hat{p} \in C_2 \circ \vartheta^{-1}} \int_{\hat{\Omega}} \hat{\xi} \, \mathrm{d}\hat{p} \quad \forall \hat{\xi} \in \hat{\Xi}$$

holds if and only if

$$\min_{\hat{p} \in C_1 \circ \vartheta^{-1}} \int_{\hat{\Omega}} \hat{\xi}_n \, \mathrm{d}\hat{p} = \min_{\hat{p} \in C_2 \circ \vartheta^{-1}} \int_{\hat{\Omega}} \hat{\xi}_n \, \mathrm{d}\hat{p} \quad \forall n.$$

Moreover, by a standard application of the hyperplane separating theorem,

$$\min_{\hat{p} \in C_1 \circ \vartheta^{-1}} \int_{\hat{\Omega}} \hat{\xi} \, \mathrm{d}\hat{p} \leq \min_{\hat{p} \in C_2 \circ \vartheta^{-1}} \int_{\hat{\Omega}} \hat{\xi} \, \mathrm{d}\hat{p} \quad \forall \hat{\xi} \in \hat{\Xi} \iff C_1 \circ \vartheta^{-1} \subseteq C_2 \circ \vartheta^{-1}. \tag{8}$$

For every n define $\xi_n = \hat{\xi}_n \circ \vartheta$. If $\sigma_{C_1}(\xi_n) \leq \sigma_{C_2}(\xi_n)$ for every n then (8) implies $C_1 \circ \vartheta^{-1} \subseteq C_2 \circ \vartheta^{-1}$. Because ϑ is injective, then so is the the map $p \mapsto p \circ \vartheta^{-1}$. Indeed, suppose $p(A) \neq q(A)$ for some event A. Then $\vartheta(A) \in \mathcal{B}$ and $(p \circ \vartheta^{-1})(\vartheta(A)) = p(A) \neq q(A) = (p \circ \vartheta^{-1})(\vartheta(B))$. It follows that $C_1 \subseteq C_2$.

The case where Ω is countable follows a similar proof. □

We conclude this section by deriving a basic property of the σ-algebra $\mathfrak{S}$ on $\mathscr{C}$.

Lemma 7. *For every $C \in \mathscr{C}$, the collections $\{D \in \mathscr{C} : C \subseteq D\}$ and $\{D \in \mathscr{C} : C \supseteq D\}$ are $\mathfrak{S}$-measurable.*

Proof. Let $\{\xi_1, \xi_2, \ldots\} \subseteq B(\mathcal{F}, \mathbb{R})$ as in Lemma 6. We have that $\{D \in \mathscr{C} : C \subseteq D\}$ equals $\bigcap_n \{D \in \mathscr{C} : \sigma_C(\xi_n) \leq \sigma_D(\xi_n)\}$. Each $\{D \in \mathscr{C} : \sigma_C(\xi_n) \leq \sigma_D(\xi_n)\}$ is measurable by definition of $\mathfrak{S}$. Being $\mathfrak{S}$ closed under countable intersections, it follows that $\{D \in \mathscr{C} : C \subseteq D\}$ belongs to $\mathfrak{S}$. A similar argument proves the measurability of $\{D \in \mathscr{C} : C \supseteq D\}$. □

A.5 Gilboa-Schmeidler's Theorem

Let $U \subseteq \mathbb{R}$ be an interval of positive length. A functional $I\colon B_b(\mathcal{F}, U) \to \mathbb{R}$ is: *monotone* if $\xi \geq \zeta$ implies $I(\xi) \geq I(\zeta)$; *normalized* if $I(a) = a$ for all $a \in U$; *constant-affine* if $I(\alpha\xi + (1-\alpha)a) = \alpha I(\xi) + (1-\alpha)a$ for all $\xi \in B_b(\mathcal{F}, U)$, $\alpha \in [0,1]$ and $a \in U$; *pointwise-continuous* if $I(\xi_n) \to I(\xi)$ whenever (ξ_n) is a bounded sequence that converges pointwise to ξ; *concave* if $I(\alpha\xi + (1-\alpha)\zeta) \geq \alpha I(\xi) + (1-\alpha)I(\zeta)$ for all $\alpha \in [0,1]$; and *affine* if $I(\alpha\xi + (1-\alpha)\zeta) = \alpha I(\xi) + (1-\alpha)I(\zeta)$ for all $\alpha \in [0,1]$.

The next result follows, up to minor details, from Gilboa and Schmeidler (1989) and Lemma 5. See also Chateauneuf, Maccheroni, Marinacci, and Tallon (2005).

Theorem 5. *A functional $I\colon B_b(\mathcal{F}, U) \to \mathbb{R}$ is monotone, normalized, constant-affine, concave, and monotone continuous if and only if there exists a set $C \in \mathscr{C}$ such that $I(\xi) = \sigma_C(\xi)$ for all $\xi \in B_b(\mathcal{F}, U)$. It is additionally affine if and only if C is a singleton. Moreover, two sets $C_1, C_2 \in \mathscr{C}$ satisfy $I = \sigma_{C_1} = \sigma_{C_2}$ if and only if $C_1 = C_2$.*

B Decomposable operators

Throughout this section, $U \subseteq \mathbb{R}$ is an interval of positive length, $\mathcal{G} \subseteq \mathcal{F}$ a σ-algebra, and q a measure in $\Delta(\mathcal{G})$.

Definition 7. An operator $T\colon B_b(\mathcal{F}, U) \to L_\infty(\mathcal{G}, q, U)$ is:

- *monotone* if $\xi \geq \zeta$ implies $T\xi \geq T\zeta$,
- *decomposable* if for all $\xi \in B_b(\mathcal{F}, U)$, $A \in \mathcal{G}$, and $a \in U$

$$T(\xi \cdot 1_A + a \cdot 1_{A^c}) = T(\xi) \cdot [1_A] + T(a) \cdot [1_{A^c}],$$

- *normalized* if $T(a) = [a]$ for all $a \in U$,
- *σ-order continuous* if $\xi_n \downarrow \xi$ implies $T\xi_n \downarrow T\xi$ and $\xi_n \uparrow \xi$ implies $T\xi_n \uparrow T\xi$,
- *projective* if $T(\xi) = [\xi]$ for all $\xi \in B_b(\mathcal{G}, U)$.

The next lemmas derive some basic properties of decomposable operators.

Lemma 8. *If T is decomposable, then for every partition $A_1, \ldots, A_n$ of Ω in events that are $\mathcal{G}$-measurable, and every $\xi_1, \ldots, \xi_n$ in $B_b(\mathcal{F}, U)$*

$$T\left(\sum_{i=1}^{n} \xi_i \cdot 1_{A_i}\right) = \sum_{i=1}^{n} T(\xi_i) \cdot [1_{A_i}] \tag{9}$$

Proof. Let $\zeta = \sum_{i=1}^n \xi_i \cdot 1_{A_i}$. Trivially $T(\zeta) = \sum_{i=1}^n T(\zeta) \cdot [1_{A_i}]$. Now fix $a \in U$. Using the fact that T is decomposable, for every i we obtain

$$\begin{aligned} T(\zeta) \cdot [1_{A_i}] + T(a) \cdot [1_{A_i^c}] &= T\left(\zeta \cdot 1_{A_i} + a \cdot 1_{A_i^c}\right) \\ &= T\left(\xi_i \cdot 1_{A_i} + a \cdot 1_{A_i^c}\right) = T(\xi_i) \cdot [1_{A_i}] + T(a) \cdot [1_{A_i^c}]. \end{aligned}$$

Summing over i and subtracting $T(a)$ yields (9). □

Lemma 9. *Assume T is monotone and σ-order continuous. If (ξ_n) is a bounded sequence such that $\xi_n \to \xi$ pointwise, then q-almost surely $T\xi_n \to T\xi$.*

Proof. The operator satisfies

$$T(\xi) = T\left(\lim_{n\to\infty} \sup_{m\geq n} \xi_m\right) = \lim_{n\to\infty} T\left(\sup_{m\geq n} \xi_m\right) \geq \limsup_{n\to\infty} T(\xi_n)$$

where the second equality follows σ-order continuity, and the inequality follows from monotonicity. Similarly, T satisfies

$$T(\xi) = T\left(\lim_{n\to\infty} \inf_{m\geq n} \xi_m\right) = \lim_{n\to\infty} T\left(\inf_{m\geq n} \xi_m\right) \leq \liminf_{n\to\infty} T(\xi_n).$$

The desired result follows. □

Lemma 10. *If T is monotone, decomposable, normalized, and σ-order continuous, then it is projective.*

Proof. Let $\xi = \sum_{i=1}^n a_i 1_{A_i}$ where $A_1, \ldots, A_n$ is a $\mathcal{G}$-measurable partition and $a_1, \ldots, a_n \in U$. By applying Lemma 8 and the fact that T is normalized, we obtain

$$T(\xi) = \sum_{i=1}^n T(a_i) \cdot [1_{A_i}] = \sum_{i=1}^n [a_i] \cdot [1_{A_i}] = [\xi].$$

The general case where ξ is not simple now follows by Lemma 9 (being $B_0(\mathcal{F}, U)$ dense in $B_b(\mathcal{F}, U)$ with respect to the supnorm). □

B.1 Affine decomposable operators

An operator $T\colon B_b(\mathcal{F}, U) \to L_\infty(\mathcal{G}, q, U)$ is *affine* if for all $\alpha \in [0,1]$ and $\xi, \zeta \in B_b(\mathcal{F}, U)$, it satisfies $T(\alpha\xi + (1-\alpha)\zeta) = \alpha T(\xi) + (1-\alpha)T(\zeta)$. The next result provides a representation for affine decomposable operators.

Theorem 6. *An operator $T\colon B_b(\mathcal{F}, U) \to L_\infty(\mathcal{G}, q, U)$ is monotone, decomposable, normalized, σ-order continuous, and affine if and only if there is a probability measure $\pi \in \Delta(\mathcal{F})$ that extends q and satisfies for all $\xi \in B_b(\mathcal{F}, U)$*

$$T\xi = E_\pi[\xi|\mathcal{G}]. \tag{10}$$

Proof. Necessity is easy to verify. Turning to sufficiency, suppose T is monotone, decomposable, normalized, σ-order continuous, and affine. Define the functional $I\colon B_b(\mathcal{F},U) \to \mathbb{R}$ by $I(\xi) = E_q[T\xi]$. It is immediate to verify that I satisfies the following properties described in Section A.5: it is normalized, monotone, and affine. Lemma 9 implies I is pointwise continuous. It therefore follows from Theorem 5 that there exists $\pi \in \Delta$ such that $I(\xi) = E_\pi[\xi]$. By Lemma 10 T is projective, hence for every $\xi \in \mathcal{B}_b(\mathcal{G},U)$ we have $E_\pi[\xi] = I(\xi) = E_q[T\xi] = E_q[\xi]$. This implies π agrees with q on $\mathcal{G}$. For all $A \in \mathcal{G}$

$$\int_A E_\pi[\xi|\mathcal{G}]\,\mathrm{d}q + aq(A^c) = I(\xi \cdot 1_A + a \cdot 1_{A^c}) = E_q[T(\xi \cdot 1_A + a \cdot 1_{A^c})] = \int_A T\xi\,\mathrm{d}q + aq(A^c).$$

where the last equality follows from the fact that T is decomposable. □

B.2 Decomposable operators and rectangular sets of measures

We now turn our attention to more general decomposable operators.

Definition 8. An operator $T\colon B_b(\mathcal{F},U) \to L_\infty(\mathcal{G},q,U)$ is

- *constant-affine* if for all $\alpha \in [0,1]$, $\xi \in B_b(\mathcal{F},U)$, and $a \in U$

$$T(\alpha\xi + (1-\alpha)a) = \alpha T(\xi) + (1-\alpha)[a],$$

- *concave* if for all $\alpha \in [0,1]$ and $\xi,\zeta \in B_b(\mathcal{F},U)$

$$T(\alpha\xi + (1-\alpha)\zeta) \geq \alpha T(\xi) + (1-\alpha)T(\zeta),$$

The following definition is adapted from Epstein and Schneider (2003, Definition 3.1). For every $A \in \mathcal{F}$ and every $\pi \in \Pi$ we denote by $\pi(\cdot|A) \in \Delta$ the corresponding conditional probability, with the convention that $\pi(\cdot|A) = q$ if $\pi(A) = 0$.

Definition 9. A set $\Pi \in \mathscr{C}$ is *q-rectangular* if all $\pi \in \Pi$ agree with q on $\mathcal{G}$, and if for every $\mathcal{G}$-measurable finite partition $A_1,\dots,A_n$ of Ω and every $\pi_1,\dots,\pi_n \in \Pi$

$$\sum_{i=1}^{n} q(A_i)\pi_i(\cdot|A_i) \in \Pi.$$

Regular sets of measures satisfy a generalization of the law of iterated expectations:

Lemma 11. *Let $\Pi \in \mathscr{C}$ be a set of measures that agree with q on $\mathcal{G}$. Then Π is q-rectangular if and only if*

$$\min_{\pi\in\Pi} \int_\Omega \xi\,\mathrm{d}\pi = \int_\Omega \operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[\xi|\mathcal{G}]\,\mathrm{d}q \quad \forall \xi \in B(\mathcal{F},U). \tag{11}$$

Proof. Let Π be q-rectangular. Fix $\xi \in B(\mathcal{F}, U)$ and let $\{\pi_1, \pi_2, \ldots\} \subseteq \Pi$ be a countable subset such that

$$\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[\xi|\mathcal{G}] = \operatorname*{ess\,inf}_{n} E_{\pi_n}[\xi|\mathcal{G}] = \left[\inf_n \zeta_n\right]$$

where each ζ_n belongs to $B_b(\mathcal{G}, U)$ and is a version of $E_{\pi_n}[\xi|\mathcal{G}]$. Now let $\varepsilon > 0$ and for every $i = 1, 2, \ldots$ let $G_i \in \mathcal{G}$ be the set of states ω where $\zeta_i(\omega) < \inf_n \zeta_n(\omega) + \varepsilon$. Define $A_1 = G_1$ and inductively, for every n, $A_n = G_n \setminus \bigcup_{i=1}^{n-1} G_i$. The events $\{A_1, A_2, \ldots\}$ form a countable partition of Ω in $\mathcal{G}$-measurable events.

Define, for every n, $B_n = \Omega \setminus \bigcup_{i=1}^{n} A_i$. Notice that $q(B_n) \downarrow 0$ as $n \to \infty$. Because Π is q-rectangular, each finite partition $\{A_1, \ldots, A_n, B_n\}$ defines a measure $\pi^n \in \Pi$ as

$$\pi^n = \sum_{i=1}^{n} \pi_i(\cdot|A_i) q(A_i) + \pi_{n+1}(\cdot|B_n) q(B_n).$$

Because $\pi^n(\cdot|A_i) = \pi_i(\cdot|A_i)$ for each $i = 1, \ldots, n$, π^n satisfies

$$\int_{A_i} E_{\pi^n}[\xi|\mathcal{G}]\,\mathrm{d}q = \int_{A_i} E_{\pi_i}[\xi|\mathcal{G}]\,\mathrm{d}q \quad i = 1, \ldots, n$$

and thus

$$\int_{A_i} E_{\pi^n}[\xi|\mathcal{G}]\,\mathrm{d}q \le \int_{A_i} \operatorname*{ess\,inf}_{\pi} E[\xi|\mathcal{G}]\,\mathrm{d}q + \varepsilon q(A_i) \quad i = 1, \ldots, n.$$

by summing over i and using the fact that ξ is bounded, we can choose n large enough such that

$$\min_{\pi\in\Pi} \int_\Omega \xi\,\mathrm{d}\pi \le \int_\Omega E_{\pi^n}[\xi|\mathcal{G}]\,\mathrm{d}q \le \int_\Omega \operatorname*{ess\,inf}_{\pi} E[\xi|\mathcal{G}]\,\mathrm{d}q + 2\varepsilon.$$

Because ε is arbitrary, we conclude that (11) holds. In the opposite direction, assume (11). Let $G_1, \ldots, G_n$ be a $\mathcal{G}$-measurable partition of Ω. Define

$$\Pi^* = \sum_{i=1}^{n} q(G_i)\{\pi(\cdot|G_i) : \pi \in \Pi\}.$$

The set Π^* is weak*-compact. It is also convex: since any two π and π' in Π agree on $\mathcal{G}$, they satisfy $(\alpha\pi + (1-\alpha)\pi')(\cdot|G_i) = \alpha\pi(\cdot|G_i) + (1-\alpha)\pi'(\cdot|G_i)$ for every i such that $\pi(G_i) > 0$ and $\pi'(G_i) > 0$ and every $\alpha \in [0, 1]$. The convexity of Π^* now follows from that of Π. For every $\pi^* \in \Pi$,

$$\int_\Omega \xi\,\mathrm{d}\pi^* \ge \sum_{i=1}^{n} q(G_i) \min_{\pi\in\Pi} \int_{G_i} E_\pi[\xi|\mathcal{G}]\,\mathrm{d}q \ge \int_\Omega \operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[\xi|\mathcal{G}]\,\mathrm{d}q.$$

It therefore follows from (11) that $\min_{\pi^*\in\Pi^*} \int_\Omega \xi\,\mathrm{d}\pi^* \ge \min_{\pi\in\Pi} \int_\Omega \xi\,\mathrm{d}\pi$ but because $\Pi \subseteq \Pi^*$, the converse inequality holds as well. We conclude that Π and Π^* satisfy $\sigma_\Pi(\xi) = \sigma_{\Pi^*}(\xi)$ for every ξ, and thus $\Pi = \Pi^*$. Hence Π is q-rectangular. □

Theorem 7. *For $T \colon B_b(\mathcal{F}, U) \to L_\infty(\mathcal{G}, q, U)$, the following statements are equivalent:*

(i). *The operator is monotone, decomposable, normalized, σ-order continuous, constant-affine, and concave.*

(ii). *There is a q-rectangular $\Pi \in \mathscr{C}$ such that for all $\xi \in B_b(\mathcal{F}, U)$*

$$T\xi = \operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}]. \tag{12}$$

Proof. "Necessity." It is immediate to verify that T, as defined in (12), is monotone and normalized. It is decomposable, since for all $A \in \mathcal{G}$, $\xi \in B_b(\mathcal{F}, U)$ and $a \in U$,

$$\begin{aligned}\operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi \cdot 1_A + a \cdot 1_{A^c}|\mathcal{G}] &= \operatorname*{ess\,inf}_{\pi \in \Pi} \left(E_\pi[\xi|\mathcal{G}] \cdot [1_A] + a \cdot [1_{A^c}]\right) \\ &= \left(\operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}]\right) \cdot [1_A] + a \cdot [1_{A^c}].\end{aligned}$$

That T is constant-affine and concave follows from the affinity of conditional expectation operator. To check σ-order continuity, let (ξ_n) be a sequence in $B_b(\mathcal{F}, U)$ such that $\xi_n \uparrow \xi$ (a similar argument applies to $\xi_n \downarrow \xi$). By Lemma 5 and the compactness of Π we have $\min_{\pi \in \Pi} E_\pi[\xi_n] \to \min_{\pi \in \Pi} E_\pi[\xi]$. Because Π is q-rectangular, Lemma 11 implies that $T\xi_n \to T\xi$ in $L_1(\mathcal{G}, q)$:

$$\int_\Omega |T\xi - T\xi_n| \,\mathrm{d}q = \int_\Omega T\xi \,\mathrm{d}q - \int_\Omega T\xi_n \,\mathrm{d}q = \min_{\pi \in \Pi} E_\pi[\xi] - \min_{\pi \in \Pi} E_\pi[\xi_n] \to 0.$$

We can therefore extract a subsequence (ξ_{n_m}) such that $T\xi_{n_m} \uparrow T\xi$ (Aliprantis and Border, 2006, Theorems 13.38 and 13.39). Because the whole sequence $(T\xi_n)$ is monotone, we conclude that $T\xi_n \uparrow T\xi$ as desired.

"Sufficiency." We define the functional $I\colon B_b(\mathcal{F}, U) \to \mathbb{R}$ by $I(\xi) = E_q[T\xi]$. It can be verified that I is monotone, normalized, constant-affine, and concave. Lemma 9 implies I is pointwise continuous. It therefore follows from Theorem 5 that there exists a set $\Pi \in \mathscr{C}$ such that $I(\xi) = \sigma_\Pi(\xi)$. Lemma 10 implies T is projective. Hence $\min_{\pi \in \Pi} E_\pi[\xi] = I(\xi) = E_q[T\xi] = E_q[\xi]$ for every $\xi \in B_b(\mathcal{G}, U)$. It turn, this implies $\min_{\pi \in \Pi} E_\pi[\xi] = E_q[\xi]$ for every $\xi \in B_b(\mathcal{G}, \mathbb{R})$. It follows from a standard separation argument that all $\pi \in \Pi$ agree with q on $\mathcal{G}$. It remains to show that (12) holds. By Lemma 11 this immediately implies that Π is q-rectangular. For all $\pi \in \Pi$ we have

$$\int_\Omega \operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}] \,\mathrm{d}q \le I(\xi) \le \int_\Omega E_\pi[\xi|\mathcal{G}] \,\mathrm{d}q.$$

Because T is decomposable, for all $a \in U$ and $A \in \mathcal{G}$

$$\int_\Omega T(\xi \cdot 1_A + a \cdot 1_{A^c}) \,\mathrm{d}q = \int_A T(\xi) \,\mathrm{d}q + aq(A^c)$$

Thus for all $A \in \mathcal{G}$ and $\pi \in \Pi$

$$\int_A \operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}] \,\mathrm{d}q \le \int_A T(\xi) \,\mathrm{d}q \le \int_A E_\pi[\xi|\mathcal{G}] \,\mathrm{d}q.$$

We conclude that (12) holds. □

B.3 Regular conditional probabilities for set-valued kernels

Being $(\Omega, \mathcal{F})$ standard Borel, each $\pi \in \Delta$ admits a regular conditional probability with respect to $\mathcal{G}$. The notion of regular conditional probability can be extended to compact convex sets of measures.

Definition 10. Let $\Pi \in \mathscr{C}$ be a set of measures that agree with q on $\mathcal{G}$. A $(\mathcal{G}, \mathfrak{S})$-measurable function $K \colon \Omega \to \mathscr{C}$ is a *regular conditional probability* of Π with respect to $\mathcal{G}$ if for all $G \in \mathcal{G}$ and $\xi \in B(\mathcal{F}, \mathbb{R})$

$$\min_{\pi \in \Pi} \int_G \xi \, \mathrm{d}\pi = \int_G \min_{p \in K(\omega)} E_p[\xi] \, \mathrm{d}q(\omega).$$

If it exists, the regular conditional probability of Π given $\mathcal{G}$ is essentially unique. Formally, if K and K' satisfy the conditions of Definition 10, then for all $\xi \in B(\mathcal{F}, \mathbb{R})$ and for q-almost all ω

$$\min_{p \in K(\omega)} E_p[\xi] = \min_{p \in K'(\omega)} E_p[\xi].$$

By Lemma 6 this implies that K and K' are equal q-almost surely. The next theorem shows that q-rectangularity is equivalent to existence of a regular conditional probability.

Theorem 8. *If Π and K satisfy the conditions of Definition 10, then Π is q-rectangular and for all $\xi \in B(\mathcal{F}, \mathbb{R})$*

$$\min_{p \in K(\cdot)} E_p[\xi] \in \operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi | \mathcal{G}].$$

Conversely, if $\Pi \in \mathscr{C}$ is q-rectangular, then it admits a regular conditional probability $K : \Omega \to \mathscr{C}$ with respect to $\mathcal{G}$ such that for q-almost all ω

$$\min_{p \in K(\omega)} p(\{\omega' : K(\omega') = K(\omega)\}) = 1.$$

B.4 Proof of Theorem 8

Necessity. Let $K \colon \Omega \to \mathscr{C}$ be a regular conditional probability of Π given $\mathcal{G}$. On one hand, for every $\xi \in B(\mathcal{F}, \mathbb{R})$

$$\int_G \operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi | \mathcal{G}] \, \mathrm{d}q \leq \min_{\pi \in \Pi} \int_G \xi \, \mathrm{d}\pi = \int_G \min_{p \in K(\omega)} E_p[\xi] \, \mathrm{d}q(\omega) \quad \forall G \in \mathcal{G}.$$

But for every $\pi \in \Pi$ and every $G \in \mathcal{G}$, $\int_G E_\pi[\xi|\mathcal{G}] \, \mathrm{d}q \geq \int_G \min_{p \in K(\omega)} E_p[\xi] \, \mathrm{d}q(\omega)$. We conclude that for all $\xi \in B(\mathcal{F}, \mathbb{R})$, $\min_{p \in K(\cdot)} E_p[\xi] \in \operatorname{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}]$, which by Lemma 9 yields that Π is q-rectangular.

Sufficiency. Assume Π is q-rectangular. Let $\lambda \in \Pi$ be a control measure for Π (see Lemma 2). Let $\{\pi_1, \pi_2, \ldots\}$ be a countable weak*-dense subset of Π (see Lemma 3). For every n, let $\xi_n \colon \Omega \to \mathbb{R}_+$ be a version of the Radon-Nikodym derivative of π_n with respect to λ.

Lemma 12. *There is a kernel $k_\lambda \colon \Omega \to \Delta$ that satisfies the following properties:*

(i). *k_λ is a regular conditional probability of λ given $\mathcal{G}$.*

(ii). *For every n, the kernel $k_n \colon \Omega \to \Delta$ defined by*

$$k_n(\omega, A) = \int_A \xi_n \, \mathrm{d}k_\lambda(\omega) \quad \forall A \in \mathcal{F}$$

is a regular conditional probability of π_n given $\mathcal{G}$. In particular, each $k_n(\omega)$ is absolutely continuous with respect to $k_\lambda(\omega)$, and ξ_n is a version of the corresponding Radon-Nikodym derivative.

(iii). *For every ω the set $\{\xi_1, \xi_2, \ldots\}$ is uniformly integrable with respect to $k_\lambda(\omega)$.*

Proof. Let $k \colon \Omega \to \Delta$ be a regular conditional probability of λ given $\mathcal{G}$. For every n, $A \in \mathcal{F}$, and $G \in \mathcal{G}$

$$\int_G \left(\int_A \xi_n \, \mathrm{d}k(\omega) \right) \mathrm{d}q(\omega) = \int_G E_\lambda[\xi_n \cdot 1_A | \mathcal{G}] \, \mathrm{d}\lambda = \int_{G \cap A} \xi_n \, \mathrm{d}\lambda(\omega) = \pi_n(G \cap A) \tag{13}$$

where the first equality uses the fact that q and λ agree on $\mathcal{G}$. In particular, by choosing $A = \Omega$ and varying G, we obtain that $\int_\Omega \xi_n \, \mathrm{d}k(\omega) = 1$ for every n and q-almost all ω. As a result, the kernel $k' \colon \Omega \to \Delta$ defined by

$$k'(\omega) = \begin{cases} k(\omega) & \text{if } \int_\Omega \xi_n \, \mathrm{d}k(\omega) = 1 \text{ for all } n, \\ \lambda & \text{otherwise,} \end{cases}$$

satisfies (i). Moreover, by substituting k with k' in (13) we obtain that k' satisfies (ii). Since λ is a control measure for Π, by Lemma 4 we can find a convex even function $\psi : \mathbb{R} \to \mathbb{R}$ such that $\psi(0) = 0$, $\lim_{t \to \infty} \frac{\psi(t)}{t} = \infty$, and

$$\sup_{\pi \in \Pi} \int_\Omega \psi \left(\frac{\mathrm{d}\pi}{\mathrm{d}\lambda} \right) \mathrm{d}\lambda < \infty.$$

For every n, take $\zeta_n \colon \Omega \to [0, \infty]$ given by $\zeta_n(\omega) = \int_\Omega \psi(\xi_n) \, \mathrm{d}k'(\omega)$ and define

$$k_\lambda(\omega) = \begin{cases} k'(\omega) & \text{if } \sup_n \zeta_n(\omega) < \infty, \\ \lambda & \text{otherwise.} \end{cases}$$

In the last part of the proof we show that $\sup_n \zeta_n(\omega) < \infty$ for q-almost all ω. This implies that k_λ continues to satisfy (i) and (ii). By Lemma 4 its also implies that k_λ satisfies (iii) as desired. For every n and $t > 0$ consider the event $G(n,t) \in \mathcal{G}$ given by

$$G(n,t) = \left\{ \omega : \max_{i=1,\ldots,n} \zeta_i(\omega) > t \right\}.$$

The event can be partitioned into the events $G_1(n,t), \ldots, G_n(n,t)$ defined by

$$G_i(n,t) = \{\omega : \zeta_i(\omega) > t \text{ and } \zeta_1(\omega), \ldots, \zeta_{i-1}(\omega) \le t\}.$$

Because Π is q-rectangular, we can find $\pi \in \Pi$ such that for all $A \in \mathcal{F}$,

$$\pi\left(A \cap G_i(n,t)\right) = \pi_i\left(A \cap G_i(n,t)\right) \quad i = 1, \ldots, n.$$

Thus, for all $A \in \mathcal{F}$,

$$\pi(A \cap G(n,t)) = \sum_{i=1}^{n} \pi_i\left(A \cap G_i(n,t)\right) = \sum_{i=1}^{n} \int_{A \cap G_i(n,t)} \xi_i \, \mathrm{d}\lambda = \int_A \left(\sum_{i=1}^{n} 1_{G_i(n,t)} \xi_i \right) \mathrm{d}\lambda.$$

This shows that $\sum_{i=1}^{n} 1_{G_i(n,t)} \xi_i$ is a version of

$$\frac{\mathrm{d}\pi}{\mathrm{d}\lambda} \cdot [1_{G(n,t)}].$$

Using the fact that $\psi(0) = 0$ we obtain that

$$\int_\Omega \psi\left(\frac{\mathrm{d}\pi}{\mathrm{d}\lambda}\right) \mathrm{d}\lambda \ge \int_{G(n,t)} \psi\left(\frac{\mathrm{d}\pi}{\mathrm{d}\lambda}\right) \mathrm{d}\lambda = \sum_{i=1}^{n} \int_{G_i(n,t)} \zeta_i \, \mathrm{d}q \ge t q(G(n,t)).$$

By varying n and t, we conclude that

$$0 = \lim_{t\to\infty} \frac{1}{t} \sup_{\pi \in \Pi} \int_\Omega \psi\left(\frac{\mathrm{d}\pi}{\mathrm{d}\lambda}\right) \mathrm{d}\lambda \ge \lim_{t\to\infty} \lim_{n\to\infty} q(G(n,t)) = q\left(\{\omega : \sup_n \zeta_n(\omega) = \infty\} \right) \square$$

Take k_λ and $\{k_1, k_2, \ldots\}$ as in Lemma 12. Parts (ii) and (iii) of the lemma show that for every ω, the derivatives

$$\left\{ \frac{\mathrm{d}k_n(\omega)}{\mathrm{d}k_\lambda(\omega)} : n = 1, 2, \ldots \right\} \subseteq L_1(\mathcal{F}, k_\lambda(\omega))$$

form a uniformly integrable set. So, $k_\lambda(\omega)$ is a control measure for $\{k_1(\omega), k_2(\omega), \ldots\}$.

For every $\omega \in \Omega$, let $K(\omega) \subseteq ba_1^+(\mathcal{F})$ be the closed convex hull of $\{k_1(\omega), k_2(\omega), \ldots\}$ in the weak* topology. Since $ba_1^+(\mathcal{F})$ is weak* compact then so is $K(\omega)$. In addition, because $k_\lambda(\omega)$ is a control measure, then $K(\omega)$ is a subset of $\Delta(\mathcal{F})$.[10]

[10] Let $\pi \in ba_1^+(\mathcal{F})$ be the weak* limit of a net that take values in the convex hull of $\{k_1(\omega), k_2(\omega), \ldots\}$. Given $\varepsilon > 0$ let $\delta > 0$ be such that for every $A \in \mathcal{F}$, $k_\lambda(\omega, A) < \delta$ implies $k_n(\omega, A) < \varepsilon$ for every n. Since weak* convergence is equivalent to eventwise convergence, it is then immediate that $k_n(\omega, A) < \varepsilon$ for every n implies $\pi(A) < \varepsilon$ as well. Now given a sequence (A_n) in $\mathcal{F}$, if $A_n \downarrow \emptyset$ then $k_\lambda(\omega, A_n) \downarrow 0$ and hence $\pi(A_n) \downarrow 0$. Therefore $\pi \in \Delta$.

We now show that the function $K : \Omega \to \mathscr{C}$ is $\mathcal{G}\backslash\mathfrak{S}$-measurable. This is equivalent to the statement that for every $\xi \in B(\mathcal{F}, \mathbb{R})$ the function $\omega \mapsto \sigma_{K(\omega)}(\xi)$ is $\mathcal{G}$-measurable. For every such ξ and every $\omega \in \Omega$, since $\{k_1(\omega), k_2(\omega), \dots\}$ is weak* dense in $K(\omega)$ we obtain

$$\sigma_{K(\omega)}(\xi) = \inf_n \int_\Omega \xi \, \mathrm{d}k_n(\omega).$$

Thus $\omega \mapsto \sigma_{K(\omega)}(\xi)$ is $\mathcal{G}$-measurable as claimed. That K is a regular conditional probability of Π given $\mathcal{G}$ now follows from Π being q-rectangular: for all $G \in \mathcal{G}$

$$\begin{aligned} \int_G \sigma_K(\xi) \, \mathrm{d}q &= \int_G \inf_n \int_\Omega \xi \, \mathrm{d}k_n(\omega) \, \mathrm{d}q \leq \inf_n \int_G \xi \, \mathrm{d}\pi_n \\ &= \min_{\pi \in \Pi} \int_G \xi \, \mathrm{d}\pi = \int_G \operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}] \, \mathrm{d}q \\ &\leq \int_G \inf_n E_{\pi_n}[\xi|\mathcal{G}] \, \mathrm{d}q = \int_G \sigma_K(\xi) \, \mathrm{d}q. \end{aligned}$$

We conclude that K is a regular conditional probability of Π given $\mathcal{G}$.

Define the set $\Omega_0 \subseteq \Omega$ by

$$\Omega_0 = \left\{ \omega : \min_{p \in K(\omega)} p(\{\omega' : K(\omega') = K(\omega)\}) = 1 \right\}.$$

It remains to show that $q(\Omega_0) = 0$. To this end, let $\mathcal{A}$ be a countable collection of events generating $\mathcal{F}$. Given $A \in \mathcal{F}$, k_n and k_m, denote by $\Omega(n, m, A)$ the set of $\omega \in \Omega$ that satisfy

$$\int_\Omega (k_m(\omega', A) - k_m(\omega, A))^2 \, \mathrm{d}k_n(\omega, \omega') = 0 \tag{14}$$

The set $\Omega(n, m, A)$ belongs to $\mathcal{G}$. We now show it satisfies $q(\Omega_0) = 0$. It is enough to show that (14) holds for q-almost all ω. Condition (14) is equivalent to

$$\int_\Omega k_m(\omega', A)^2 \, \mathrm{d}k_n(\omega, \omega') + k_m(\omega, A)^2 = 2k_m(\omega, A) \int_\Omega k_m(\omega', A) \, \mathrm{d}k_n(\omega, \omega').$$

Because $k_m(\cdot, A)$ and $k_m(\cdot, A)^2$ are $\mathcal{G}$-measurable, for q-almost all ω

$$\begin{aligned} \int_\Omega k_m(\omega', A) \, \mathrm{d}k_n(\omega, \omega') &= k_m(\omega, A), \\ \int_\Omega k_m(\omega', A)^2 \, \mathrm{d}k_n(\omega, \omega') &= k_m(\omega, A)^2. \end{aligned}$$

Hence $q(\Omega(n, m, A)) = 0$. Now let $\Omega_1 = \bigcap_{n,m} \bigcap_{A \in \mathcal{A}} \Omega(n, m, A)$. Then $q(\Omega_1) = 1$, and for every $\omega \in \Omega_1$ and every n,

$$\begin{aligned} 1 &= k_n(\omega) \left(\{\omega' : \forall m, \forall A \in \mathcal{A}, k_m(\omega', A) = k_m(\omega, A)\} \right) \\ &= k_n(\omega)(\{\omega' : \forall m, k_m(\omega') = k_m(\omega)\}) = k_n(\omega)(\{\omega' : K(\omega') = K(\omega)\}). \end{aligned}$$

Because this is true for every $k_n(\omega)$ it follows that every $p \in K(\omega)$ in the closure of $\{k_1(\omega), k_2(\omega), \dots\}$ satisfies $p(\{\omega' : K(\omega') = K(\omega)\}) = 1$.

C Baseline representation under Axioms 1-3 and 5-6

We begin by studying some preliminary implications of our basic axioms. For the moment we consider binary relations that satisfy Axioms 1-3, as well as the von Neumann-Morgenstern independence axiom on X:

Axiom 9. *For all $x, y, z \in X$ and $\alpha \in [0,1]$, if $x \succsim y$ then $x + (1-\alpha)z \succsim \alpha y + (1-\alpha)z$.*

Lemma 13. *If $\succsim$ satisfies Axioms 1-3, then the following conditions hold:*

(i) *If $f(\omega) \succsim g(\omega)$ for all ω, then $f \succsim g$.*

(ii) *For all acts f, g, h the sets $\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succsim h\}$ and $\{\alpha \in [0,1] : h \succsim \alpha f + (1-\alpha)g\}$ are closed.*

(iii) *If in addition $\succsim$ satisfies Axiom 9, then there exists a non-constant affine function $u\colon X \to \mathbb{R}$ representing $\succsim$ on X.*

Proof. (i). Let $x \succ y$ and define $f_n = \frac{1}{n}x + \left(1 - \frac{1}{n}\right) f$ and $g_n = \frac{1}{n}y + \left(1 - \frac{1}{n}\right) g$. Axioms 1 and 2 imply $f_n \succ g_n$ for every n. The two sequences are bounded and converge pointwise to f and g, respectively. It follows from Axiom 3 that $f \succsim g$.

(ii). It follows from Axiom 3.

(iii). The claim is an application of the mixture space theorem (Herstein and Milnor, 1953) together with (ii) and Axioms 1 and 9. □

Lemma 14. *For every σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ and affine function $u\colon X \to \mathbb{R}$,*

$$B_b(\mathcal{G}, u(X)) = \{u(f) : f \in \mathfrak{F} \text{ and } f \text{ is } \mathcal{G}\text{-measurable}\}.$$

Proof. Let $f \in \mathfrak{F}$ be $\mathcal{G}$-measurable and let $Y \subseteq X$ be a polytope such that $f(\Omega) \subseteq Y$. The set Y is compact and u (being affine) is continuous on Y (Aliprantis and Border, 2006, Theorem 5.21). Thus $u(f)$ is $\mathcal{G}$-measurable and $\min u(Y) \leq u(f) \leq \max u(Y)$. It follows that $u(f)$ belongs to $B_b(\mathcal{G}, u(X))$. In the opposite direction, let $\xi \in B_b(\mathcal{G}, u(X))$ and $u(x) \geq \xi \geq u(y)$ for some $x, y \in X$. If $u(x) = u(y)$, take $f = x$. If instead $u(x) > u(y)$, take $\zeta \equiv \frac{\xi - u(y)}{u(x) - u(y)}$ and $f \equiv \zeta x + (1-\zeta)y$. The function f belongs to $\mathfrak{F}$ and $u(f) = \xi$. □

For a preference relation $\succsim$ that satisfies Axioms 1-3 and 9, Lemma 13(i) and 13(ii) imply that for every $A \in \mathcal{F}_{\text{st}}$ and $f \in \mathfrak{F}$ there exists an outcome $c(f|A) \in X$ such that $c(f|A) \sim_A f$. If $A = \Omega$, we simply write $c(f)$ instead of $c(f|\Omega)$.

Lemma 15. *Assume Axioms 1-3 and 9 are satisfied. For every affine function $u\colon X \to \mathbb{R}$ representing $\succsim$ on X, the following conditions hold:*

(i) *If (f_n) is bounded and $f_n \to f$ pointwise, then $u(c(f_n)) \to u(c(f))$.*

(ii) *If (f_n) is bounded and $u(f_n) \to u(f)$ pointwise, then $u(c(f_n)) \to u(c(f))$.*

(iii) *If Axiom 5 holds and $A \in \mathcal{F}_{\mathrm{st}}$ is not null, then $x \succ y$ implies $x \succ_A y$.*

Proof. (i). Choose $x, y \in X$ such that $x \succsim f_n(\omega) \succsim y$ for all n and ω. By Lemma 13(i) we have $x \succsim f_n \succsim y$ for all n. By Axiom 3 this implies that $x \succsim f \succsim y$ as well. If $x \sim y$, then $u(c(f_n)) = u(x) = u(c(f))$ for all n. Assume therefore that $x \succ y$. By Lemma 13(ii) we can choose $\alpha_n \in [0,1]$ and $\alpha \in [0,1]$ such that $f_n \sim \alpha_n x + (1-\alpha_n)y$ and $f \sim \alpha x + (1-\alpha)y$. Possibly passing to a subsequence, we can assume without loss of generality that $\alpha_n \to \beta$ for some $\beta \in [0,1]$. It follows from Axiom 3 that $f \sim \beta x + (1-\beta)y$, i.e., $u(c(f)) = \beta u(x) + (1-\beta)u(y)$, which in turn implies $\alpha = \beta$. Thus

$$u(c(f_n)) = \alpha_n u(x) + (1-\alpha_n)u(y) \longrightarrow \alpha u(x) + (1-\alpha)u(y) = u(c(f)).$$

(ii). Choose $x, y \in X$ such that $x \succsim f_n(\omega) \succsim y$ for all n and ω. By Axiom 3 this implies that $x \succsim f(\omega) \succsim y$ for all ω as well. Take $\xi_n \in B(\mathcal{F}, [0,1])$ and $\xi \in B(\mathcal{F}, [0,1])$ such that $u(f_n) = \xi_n u(x) + (1-\xi_n)u(y)$ and $u(f) = \xi u(x) + (1-\xi)u(y)$. Define $g_n = \xi_n x + (1-\xi_n)y$ and $g = \xi x + (1-\xi)y$. Observe that $u(f_n) = u(g_n)$ and $u(f) = u(g)$: it follows from Lemma 13(i) that $u(c(f_n)) = u(c(g_n))$ and $u(c(f)) = u(c(g))$. In addition, $u(f_n) \to u(f)$ pointwise implies $g_n \to g$ pointwise. The desired result then follows from (i) above.

(iii). Being A not null, there are f, g such that $f \succ_A g$. Take $w, z \in X$ such that $w \succsim f(\omega)$ and $g(\omega) \succsim z$ for all ω. By Lemma 13(i) we have $w \succ_A z$, that is, $w \succ zAw$. It follows from Axiom 5 that $x \succ yAx$, that is, $x \succ_A y$. □

The next lemma is due to Gul and Pesendorfer (2014, Lemma B2).

Lemma 16. *If Axioms 1-3, 5, and 9 are satisfied, then $\mathcal{F}_{\mathrm{st}}$ is a σ-algebra.*

Up to minor details, the result follows by replicating the proof in Gul and Pesendorfer (2014). A self-contained proof is available from the authors upon request.

C.1 Representation

The next theorem introduces a representation of the agent's preferences in terms of decomposable operators.

Theorem 9. *If Axioms 1-3, 5-6, and 9 are satisfied, then there are*

(i). *a non-constant affine function $u\colon X \to \mathbb{R}$,*

(ii). *a nonatomic probability measure $q \in \Delta(\mathcal{F}_{\mathrm{st}})$,*

(iii). *a continuous strictly increasing function $\phi\colon u(X) \to \mathbb{R}$, and*

(iv). *a monotone, normalized, decomposable, σ-order continuous operator*

$$T\colon B_b(\mathcal{F}, u(X)) \to L_\infty(\mathcal{F}_{\mathrm{st}}, q, u(X)),$$

such that for all $f, g \in \mathfrak{F}$

$$f \succsim g \iff \int_\Omega \phi(Tu(f))\,\mathrm{d}q \geq \int_\Omega \phi(Tu(g))\,\mathrm{d}q,$$
$$f \succsim_{\mathrm{st}} g \iff Tu(f) \geq Tu(g).$$

C.2 Proof of Theorem 9

The proof of the result is divided in lemmas. For the remaining of this section, we assume that Axioms 1-3, 5-6, and 9 are satisfied. By Lemma 16 the collection of events $\mathcal{F}_{\mathrm{st}}$ is a σ-algebra.

Lemma 17. *There exist a non-atomic probability measure* $q \in \Delta(\mathcal{F}_{\mathrm{st}})$ *and a continuous strictly increasing function* $\phi\colon u(X) \to \mathbb{R}$ *such that for all* $\mathcal{F}_{st}$*-measurable acts* f *and* g

$$f \succsim g \iff \int_\Omega \phi(u(f))\,\mathrm{d}q \geq \int_\Omega \phi(u(g))\,\mathrm{d}q \tag{15}$$

Proof. First we show that (15) holds for simple acts. Let $\succsim_0$ be the restriction of $\succsim$ to the acts that are simple and $\mathcal{F}_{\mathrm{st}}$-measurable. Observe that $\succsim_0$ satisfy Savage's P1-P6: Axiom 1 implies P1 and P5; P2 holds by definition of $\mathcal{F}_{\mathrm{st}}$; P3 follows from Lemmas 13(i) and 15(iii); Axiom 5 is P4; Axiom 6 is P6.[11] In addition, $\succsim_0$ satisfies risk independence (Axiom 9), mixture continuity (Lemma 13(ii)), and monotone continuity is implied by Lemma 15(i). By Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2012, Proposition 3) there exist a non-atomic probability measure $q \in \Delta(\mathcal{F}_{\mathrm{st}})$ and a continuous strictly increasing function $\phi\colon u(X) \to \mathbb{R}$ such that (15) holds for all f and g that are $\mathcal{F}_{\mathrm{st}}$-measurable and simple.

Now we extend the result to acts that are not simple. Let f be a $\mathcal{F}_{\mathrm{st}}$-measurable act. From Lemma 14 we have $u(f) \in B_b(\mathcal{F}_{\mathrm{st}}, u(X))$. Thus we can find a sequence (ξ_n) in $B_0(\mathcal{F}_{\mathrm{st}}, u(X))$ that converges uniformly to $u(f)$. Applying Lemma 14 again we can find a sequence (f_n) of simple $\mathcal{F}_{\mathrm{st}}$-measurable acts such that $u(f_n) = \xi_n$ for all n. It follows from Lemma 15(ii) and continuity of ϕ that $\phi(u(c(f_n))) \to \phi(u(c(f)))$. In addition, by the continuity of ϕ,

$$\lim_n \int_\Omega \phi(u(f_n))\,\mathrm{d}q = \int_\Omega \phi(u(f))\,\mathrm{d}q.$$

Because (15) holds for all simple acts, $\int_\Omega \phi(u(f_n))\,\mathrm{d}q = \phi(u(c(f_n)))$ for all n. We deduce that $\int_\Omega \phi(u(f))\,\mathrm{d}q = \phi(u(c(f)))$. It follows that (15) holds for all $\mathcal{F}_{\mathrm{st}}$-measurable acts. $\square$

We define the functional $V : \mathfrak{F} \to \mathbb{R}$ by

$$V(f) = \phi(u(c(f))).$$

Lemma 17 shows that V represents $\succsim$ on $\mathfrak{F}$. Moreover $V(f) = \int_\Omega \phi(u(f))\,\mathrm{d}q$ for all $\mathcal{F}_{\mathrm{st}}$-measurable acts f. The next lemmas establish key properties of $\succsim_{\mathrm{st}}$.

[11]See, e.g., Gilboa (2009, Section 10) for a textbook reference on Savage's theorem.

Lemma 18. *For all $f, g \in \mathfrak{F}$, the following conditions are satisfied.*

(i) *$f \succsim_{\mathrm{st}} g$ if and only if $fAh \succsim_{\mathrm{st}} gAh$ for all $A \in \mathcal{F}_{\mathrm{st}}$ and $h \in \mathfrak{F}$.*

(ii) *$u(f) \geq u(g)$ implies $f \succsim_{\mathrm{st}} g$.*

(iii) *If f and g are $\mathcal{F}_{\mathrm{st}}$-measurable, $f \succsim_{\mathrm{st}} g$ if and only if q-almost surely $u(f) \geq u(g)$.*

Proof. (i). If $f \succsim_{\mathrm{st}} g$, then for all $B \in \mathcal{F}_{\mathrm{st}}$ we have $A \cap B \in \mathcal{F}_{\mathrm{st}}$ and therefore, for every $h \in \mathfrak{F}$,

$$(fAh)Bh = f(A \cap B)h \succsim g(A \cap B)h = (gAh)Bh.$$

Now, since $B \in \mathcal{F}_{\mathrm{st}}$, then $(fAh)Bh' \succsim (gAh)Bh'$ for every $h' \in \mathfrak{F}$, which implies $fAh \succsim_{\mathrm{st}} gAh$. The other implication is obvious. (ii). It follows from Lemma 13(i).

(iii). "If." Let $A \in \mathcal{F}_{\mathrm{st}}$ be the event where $g(\omega) \succ f(\omega)$. Because $q(A) = 0$, it follows from Lemma 15(iii) that A is null. Thus $fAg \sim_{\mathrm{st}} g$. Moreover $f \succsim_{\mathrm{st}} fAg$ by (ii) above. We conclude that $f \succsim_{\mathrm{st}} g$. "Only if." Fix $x \in X$. For all $A \in \mathcal{F}_{\mathrm{st}}$ we have $fAx \succsim gAx$, that is,

$$\int_A \phi(u(f)) \,\mathrm{d}q \geq \int_A \phi(u(g)) \,\mathrm{d}q.$$

Thus q-almost surely $\phi(u(f)) \geq \phi(u(g))$ and hence $u(f) \geq u(g)$, being ϕ strictly increasing. □

Lemma 19. *For every $f \in \mathfrak{F}$ there exists a $\mathcal{F}_{\mathrm{st}}$-measurable act $\hat{f}$ such that $f \sim_{\mathrm{st}} \hat{f}$.*

Proof. Fix $x \in X$ such that $f(\omega) \succsim x$ for all ω. Let $q_f : \mathcal{F}_{\mathrm{st}} \to \mathbb{R}$ be defined by $q_f(A) = V(fAx) - V(x)$. The set function q_f is a σ-additive measure. Indeed, observe first that $q_f(\varnothing) = 0$. Second, we have that q_f is monotone: $A \subseteq B$ implies $fBx \succsim fAx$ by Lemma 13(i), which in turn implies $q_f(A) \leq q_f(B)$. To see that q_f is finitely additive, let A and B be disjoint element of $\mathcal{F}_{\mathrm{st}}$. Observe that

$$f(A \cup B)x = fAfBx \sim c(A|f)AfBx \sim c(A|f)Ac(B|f)Bx.$$

Define $g = c(A|f)Ac(B|f)Bx$. Then

$$\begin{aligned} q_f(A \cup B) &= V(g) - V(x) \\ &= \phi(u(c(A|f)))q(A) + \phi(u(c(B|f)))q(B) - \phi(u(x))(q(A) + q(B)) \\ &= V(c(A|f)Ax) - V(x) + V(c(B|f)Bx) - V(x) = q_f(A) + q_f(B). \end{aligned}$$

Finally, let (A_n) be a sequence in $\mathcal{F}_{\mathrm{st}}$ such that $A_n \downarrow \varnothing$. The sequence (fA_nx) is bounded and converges pointwise to x. By Lemma 15(i) $u(c(fA_nx)) \to u(x)$. It follows from continuity of ϕ that $q_f(A_n) \to 0$. We conclude that q_f is a σ-additive measure.

If $q(A) = 0$, it follows from Lemma 15(iii) that A is null, and therefore $q_f(A) = 0$. Thus q_f is absolutely continuous with respect to q and we can apply the Radon-Nikodym theorem to find a $\mathcal{F}_{\text{st}}$-measurable function $\xi\colon \Omega \to \mathbb{R}_+$ such that for all $A \in \mathcal{F}_{\text{st}}$, $q_f(A) = \int_A \xi \,\mathrm{d}q$. Let $y \in X$ such that $y \succsim f(\omega)$ for all ω. For all $A \in \mathcal{F}_{\text{st}}$ we have by Lemma 13(i) that $yAx \succsim fAx$, which means that

$$\phi(u(y))q(A) \geq \int_A \xi + \phi(u(x)) \,\mathrm{d}q.$$

Thus q-almost surely $\phi(u(y)) \geq \xi + \phi(u(x)) \geq \phi(u(x))$. Possibly passing to another version of the Radon-Nikodym derivative, we can assume without loss of generality that $\phi(u(y)) \geq \xi + \phi(u(x)) \geq \phi(u(x))$ everywhere. Because $u(X)$ is convex, the interval $[u(x), u(y)]$ is included by $u(X)$. Because ϕ is continuous and strictly increasing, $\phi([u(x), u(y)]) = [\phi(u(x)), \phi(u(y))]$. In addition the inverse function ϕ^{-1} is measurable, being strictly increasing. Thus we can define $\zeta = \phi^{-1}(\xi + \phi(u(x))) \in B_b(\mathcal{F}_{\text{st}}, u(X))$. By Lemma 14 there is a $\mathcal{F}_{\text{st}}$-measurable act $\hat{f}$ such that $u(\hat{f}) = \zeta$. For all $A \in \mathcal{F}_{\text{st}}$

$$V(\hat{f}Ax) = \int_A \xi \,\mathrm{d}q + \phi(u(x)) = V(fAx).$$

We conclude that $\hat{f} \sim_{\text{st}} f$. □

We define the operator $T\colon B_b(\mathcal{F}, X) \to L_\infty(\mathcal{F}_{\text{st}}, q, u(X))$ by $Tu(f) = [u(\hat{f})]$, where $\hat{f}$ is a $\mathcal{F}_{\text{st}}$-measurable act that satisfies $\hat{f} \sim_{\text{st}} f$. By Lemmas 14, 18(iii), and 19 the operator is well defined. In addition, $f \succsim_{\text{st}} g$ if and only if $Tu(f) \geq Tu(g)$. Moreover, since $\hat{f} \sim_{\text{st}} f$ implies $\hat{f} \sim f$, we obtain the representation

$$V(f) = \int_\Omega \phi(Tu(f)) \,\mathrm{d}q.$$

The next lemma concludes the proof of Theorem 9.

Lemma 20. *T is monotone, normalized, decomposable, and σ-order continuous.*

Proof. Monotonicity follows from Lemmas 18(ii) and 18(iii). Normalization is obvious. Decomposability follows from Lemma 18(i): Given f let $\hat{f}$ be $\mathcal{F}_{\text{st}}$-measurable and such that $\hat{f} \sim_{\text{st}} f$. Lemma 18(i) implies that for every $A \in \mathcal{F}_{\text{st}}$ and $x \in X$

$$T(u(f)1_A + u(x)1_{A^c}) = T(u(fAx)) = [u(\hat{f}Ax)] = Tu(f) \cdot [1_A] + [u(x)][1_{A^c}].$$

It remains to show T is σ-order continuous. Suppose $u(f_n) = \xi_n \uparrow \xi = u(f)$ (a similar argument applies to $\xi_n \downarrow \xi$). Lemma 15(ii) and continuity of ϕ imply that $V(f_n) \to V(f)$. Because, T is monotonic and ϕ is strictly increasing, $\phi(T\xi_n) \leq \phi(T\xi_{n+1}) \leq \phi(T\xi)$ for all n. Thus $\phi(T\xi_n) \to \phi(T\xi)$ in $L_1(\mathcal{G}, q)$:

$$\int_\Omega |\phi(T\xi) - \phi(T\xi_n)| \,\mathrm{d}q = \int_\Omega \phi(T\xi) \,\mathrm{d}q - \int_\Omega \phi(T\xi_n) \,\mathrm{d}q = V(f) - V(f_n) \to 0.$$

We can therefore extract a subsequence (ξ_{n_m}) such that q-almost surely $\phi(T\xi_{n_m}) \to \phi(T\xi)$ (Aliprantis and Border, 2006, Theorems 13.38 and 13.39). Monotonicity of the sequence allows us to conclude that $\phi(T\xi_n) \uparrow \phi(T\xi)$. The sequence $(T\xi_n)$ is monotonic as well. Because ϕ is strictly increasing, we conclude that $T\xi_n \uparrow T\xi$. □

D Multiple predictive representation

Definition 11. A tuple $(u, \phi, \mathcal{G}, q, \Pi)$ is a *multiple predictive representation* of $\succsim$ if $u\colon X \to \mathbb{R}$ is a non-constant affine function, $\phi\colon u(X) \to \mathbb{R}$ is a strictly increasing continuous function, $\mathcal{G} \subseteq \mathcal{F}$ is a σ-algebra, $q \in \Delta(\mathcal{G})$ is a nonatomic probability measure, and $\Pi \subseteq \Delta$ is a q-rectangular weak*-compact convex set such that for all $f, g \in \mathfrak{F}$

$$f \succsim g \iff E_q\Big[\phi\Big(\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(f)|\mathcal{G}]\Big)\Big] \geq E_q\Big[\phi\Big(\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(g)|\mathcal{G}]\Big)\Big].$$

The definition extends the predictive representation introduced in the main text. If Π is a singleton, then Definition 11 reduces to Definition 3. In what follows, to shorten notation we may write $(u, \phi, \mathcal{G}, \Pi)$ instead of $(u, \phi, \mathcal{G}, q, \Pi)$.

We first characterize the collection of events that satisfy the sure-thing principle for a preference relation that admits a multiple predictive representation. The proof is presented in Section D.1 below.

Lemma 21. *If $\succsim$ admits a multiple predictive representation $(u, \phi, \mathcal{G}, \Pi)$, then $\mathcal{G} \subseteq \mathcal{F}_{\text{st}}$. If, in addition, ϕ is not affine, then $\mathcal{F}_{\text{st}}$ and $\mathcal{G}$ are Π-equivalent.*

It follows a representation result for $\succsim_{\text{st}}$:

Proposition 4. *If $\succsim$ is represented by $(u, \phi, \mathcal{G}, \Pi)$, then $f \succsim_{\text{st}} g$ implies*

$$\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(f)|\mathcal{G}] \geq \operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(g)|\mathcal{G}]. \tag{16}$$

If in addition ϕ is not affine, then $f \succsim_{\text{st}} g$ if and only if (16) holds.

Proof. First observe that $f \succsim_G g$ for all $G \in \mathcal{G}$ is equivalent to

$$\int_G \phi\left(\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(f)|\mathcal{G}]\right) \mathrm{d}q \geq \int_G \phi\left(\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(g)|\mathcal{G}]\right) \mathrm{d}q \quad \forall G \in \mathcal{G},$$

which in turn is equivalent to (16), being ϕ strictly increasing. By Lemma 21 we have $\mathcal{G} \subseteq \mathcal{F}_{\text{st}}$. Thus $f \succsim_{\text{st}} g$ implies (16). If in addition ϕ is not affine, then by Lemma 21 $\mathcal{G}$ and $\mathcal{F}_{\text{st}}$ are Π-equivalent. If $A \in \mathcal{F}_{\text{st}}$ is Π-equivalent to $G \in \mathcal{G}$, then $u(fAh)$ and $u(fGh)$ are equal π-almost surely for all $\pi \in \Pi$ and every third act h, which implies

$$\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(fAh)|\mathcal{G}] = \operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(fGh)|\mathcal{G}] \quad \forall h \in \mathfrak{F}.$$

We deduce that $f \succsim_{\text{st}} g$ if and only if (16) holds. ∎

The next result is a representation theorem for $\succsim$ (see Section D.2 for the proof).

Theorem 10. *A preference $\succsim$ satisfies Axioms 1-3 and 5-8 if and only if it admits a multiple predictive representation.*

The next results describe the uniqueness properties of the representation.

Proposition 5. *If $\succsim$ admit a multiple predictive representation $(u, \phi, \mathcal{G}, \Pi)$ and $\mathcal{H} \subseteq \mathcal{F}$ is a σ-algebra Π-equivalent to $\mathcal{G}$, then $\succsim$ admits a multiple predictive representation $(u, \phi, \mathcal{H}, \Pi)$.*

Proof. By Lemma 1 all $\pi \in \Pi$ agree on $\mathcal{H}$. It remains to show that their common restriction on $\mathcal{H}$ is nonatomic. Fix any $\pi \in \Pi$ and let $H \in \mathcal{H}$ such that $\pi(H) > 0$. Take $G \in \mathcal{G}$ that is π-equivalent to H. Because π is nonatomic on $\mathcal{G}$, there exists $G' \subseteq G$ such that $0 < \pi(G') < \pi(G) = \pi(H)$. Let $H' \in \mathcal{H}$ be π-equivalent to G'. Then G' is also π-equivalent to $H \cap H'$. Thus $\pi(H \cap H') = \pi(G') \in (0, \pi(H))$. Thus π is nonatomic on $\mathcal{H}$. □

Theorem 11. *Two multiple predictive representations $(u_1, \phi_1, \mathcal{G}_1, \Pi_1)$ and $(u_2, \phi_2, \mathcal{G}_2, \Pi_2)$ of the same preference $\succsim$ are related by the following conditions:*

(i). *There are $a, c \in \mathbb{R}$ and $b, d > 0$ such that $u_2(x) = au_1(x) + b$ and $\phi_2(t) = c\phi_1(\frac{t-b}{a}) + d$ for all $x \in X$ and $t \in u_2(X)$.*

(ii). *$\Pi_1 = \Pi_2$ and, provided that ϕ_1 is not affine, $\mathcal{G}_1$ and $\mathcal{G}_2$ are Π_1-equivalent.*

The proof of Theorem 11 is presented in Section D.3. We conclude by characterizing the null events.

Lemma 22. *Let $\succsim$ admit a multiple predictive representation $(u, \phi, \mathcal{G}, \Pi)$. An event $A \in \mathcal{F}$ is null if and only if $\pi(A) = 0$ for all $\pi \in \Pi$.*

Proof. Let A be null. Take $x, y \in X$ such that $x \succ y$. From $xAy \sim y$ we obtain

$$\begin{aligned}\phi(u(x)) &= E_q\Big[\phi\Big(\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(yAx)|\mathcal{G}]\Big)\Big] \\ &= E_q\Big[\phi\Big(u(y)\operatorname*{ess\,sup}_{\pi\in\Pi}\pi(A|\mathcal{G}) + u(x)\operatorname*{ess\,inf}_{\pi\in\Pi}\pi(A^c|\mathcal{G})\Big)\Big].\end{aligned}$$

Being ϕ strictly increasing, $\operatorname{ess\,sup}_{\pi\in\Pi}\pi(A|\mathcal{G}) = [0]$. Thus $\pi(A|\mathcal{G}) = [0]$ for all $\pi \in \Pi$, which in turn implies that $\pi(A) = 0$ for all $\pi \in \Pi$.

Conversely, suppose that $\pi(A) = 0$ for all $\pi \in \Pi$. For every pair of acts f and h, we have $E_\pi[u(fAh)|\mathcal{G}] = E_\pi[u(h)|\mathcal{G}]$ for all $\pi \in \Pi$. Thus A is null. □

D.1 Proof of Lemma 21

The proof of the result is divided in lemmas. For the remaining of this section, we assume that $\succsim$ admit a multiple predictive representation $(u, \phi, \mathcal{G}, q, \Pi)$. Let

$$V(f) = \int_\Omega \phi(Tu(f))\,\mathrm{d}q.$$

and let $T\colon B_b(\mathcal{F}, u(X)) \to L_\infty(\mathcal{G}, q, U)$ be the operator defined by

$$Tu(f) = \operatorname*{ess\,inf}_{\pi\in\Pi} E[u(f)|\mathcal{G}].$$

By Lemma 14 and Theorem 7 the operator T is decomposable. Without loss of generality, assume $\inf u(X) < 0$, $\sup u(X) > 1$, $\phi(0) = 0$, and $\phi(1) = 1$.

Lemma 23. *If $A \in \mathcal{F}$ is Π-equivalent to a $G \in \mathcal{G}$, then $A \in \mathcal{F}_{\text{st}}$. In particular, $\mathcal{G} \subseteq \mathcal{F}_{\text{st}}$.*

Proof. Let $A \in \mathcal{F}$ and $G \in \mathcal{G}$ be Π-equivalent. For all acts f and h, we have $E_\pi[u(fAh)|\mathcal{G}] = E_\pi[u(fGh)|\mathcal{G}]$ for all $\pi \in \Pi$, which implies

$$T(u(f) \cdot 1_A + u(h) \cdot 1_{A^c}) = T(u(f) \cdot 1_G + u(h) \cdot 1_{G^c}).$$

Because in addition T is decomposable, by Lemma 8

$$T(u(f) \cdot 1_G + u(h) \cdot 1_{G^c}) = T(u(f)) \cdot [1_G] + T(u(h)) \cdot [1_{G^c}].$$

We deduce that $fAh \succsim gAh$ if only if $\int_G \phi(Tu(f))\,\mathrm{d}q \geq \int_G \phi(Tu(g))\,\mathrm{d}q$ if and only if $fAh' \succsim gAh'$. The same argument applies also to A^c, being Π-equivalent to G^c. It follows that $A \in \mathcal{F}_{\text{st}}$. □

Lemma 24. *Let $A \in \mathcal{F}_{\text{st}}$ and $G, H \in \mathcal{G}$. If $T1_A = [1_G]$ and $T1_{A^c} = [1_H]$, then A is Π-equivalent to G.*

Proof. For each π we denote by $\pi(A|\mathcal{G})$ the conditional expectation $E_\pi[1_A|\mathcal{G}]$. We first show that G and H^c are q-equivalent. For $\pi \in \Pi$ we have $\pi(A|\mathcal{G}) \geq T1_A = [1_G]$ and $\pi(A|\mathcal{G}) \leq 1 - T1_{A^c} = [1_{H^c}]$. It follows that $[1_{H^c}] \geq \pi(A|\mathcal{G}) \geq [1_G]$, which implies $q(G \cap H) = 0$.

Consider now the event $B = G^c \cap H^c$. Choose $x, y \in X$ such that $u(x) = 1$ and $u(y) = 0$. Since A satisfies the sure-thing principle

$$V(yAy) \geq V((xBy)Ay) \iff V(yA(xBy)) \geq V((xBy)A(xBy)),$$

moreover $V(yAy) = 0$, $V((xBy)A(xBy)) = q(B)$, and

$$\begin{aligned} V((xBy)Ay) &= \int_\Omega \phi(T1_{A\cap B})\,\mathrm{d}q = \int_B \phi(T1_{A\cap B})\,\mathrm{d}q \leq \int_B \phi(1_G)\,\mathrm{d}q = 0 \\ V(yA(xBy)) &= \int_\Omega \phi(T1_{A^c\cap B})\,\mathrm{d}q = \int_B \phi(T1_{A^c\cap B})\,\mathrm{d}q \leq \int_B \phi(1_H)\,\mathrm{d}q = 0. \end{aligned}$$

Thus $q(B) = 0$. We obtain that G and H^c are q-equivalent. Hence

$$\operatorname*{ess\,inf}_{\pi\in\Pi} \pi(A|\mathcal{G}) = [1_G] = [1_{H^c}] = \operatorname*{ess\,sup}_{\pi\in\Pi} \pi(A|\mathcal{G}).$$

It follows that $\pi(A|\mathcal{G}) = [1_G]$ for all π, which implies $E_\pi[1_A \cdot 1_{G^c}] = E_\pi[1_G \cdot 1_{G^c}] = 0$ and $E_\pi[1_{A^c} \cdot 1_G] = E_\pi[1_{G^c} \cdot 1_G] = 0$. We conclude that A is Π-equivalent to G. □

The next lemma concludes the proof of Lemma 21.

Lemma 25. *If there is $A \in \mathcal{F}_{\text{st}}$ such that $T1_A \neq [1_G]$ for all $G \in \mathcal{G}$, then ϕ is affine.*

Proof. Let $\rho \in B_b(\mathcal{G}, [0,1])$ be a representative $\rho \in T1_A$. Then

$$q(\{\omega : \rho(\omega) \in (0,1)\}) > 0.$$

For every $t_*, t^* \in (0,1)$ with $t_* < t^*$, we define the event

$$G_{t_*,t^*} = \{\omega \in \Omega : t_* \leq \rho(\omega) \leq t^*\}.$$

Because q is σ-additive, we can find $\bar{t} \in (0,1)$ such that for all t_*, t^* as above,

$$t_* < \bar{t} < t^* \quad \Rightarrow \quad q(G_{t_*,t^*}) > 0.$$

Indeed, if not, then for every $t \in (0,1)$ there is an interval $I_t \subseteq (0,1)$ of positive length such that $t \in I_t$ and $q(\{\omega : \rho(\omega) \in I_t\}) = 0$. But then the equality $(0,1) = \cup_{t \in \mathbb{Q} \cap (0,1)} I_t$ implies $q(\{\omega : \rho(\omega) \in (0,1)\}) = 0$ by σ-additivity of q. A contradiction.

Let $t_*, t^* \in (0,1)$ such that $t_* < \bar{t} < t^*$. We first observe that for all $\xi, \zeta \in B_b(\mathcal{G}, u(X))$ if $\xi \geq \zeta$ then

$$\operatorname*{ess\,inf}_{\pi} E_\pi[\xi \cdot 1_A + \zeta \cdot 1_{A^c} | \mathcal{G}] = \operatorname*{ess\,inf}_{\pi} ([\zeta] + E_\pi[1_A] \cdot ([\xi] - [\zeta])) = [\zeta] + [\rho] \cdot ([\xi] - [\zeta]).$$

Therefore, for all $\xi, \zeta \in B_b(\mathcal{G}, u(X))$

$$\xi \geq \zeta \quad \Rightarrow \quad T(\xi \cdot 1_A + \zeta \cdot 1_{A^c}) = [\rho] \cdot [\xi] + [1 - \rho] \cdot [\zeta].$$

Because A satisfies the sure-thing principle, for all $\xi, \xi', \zeta, \zeta' \in B_b(\mathcal{G}, u(X))$ such that

$$\omega \in G_{t_*,t^*} \quad \Rightarrow \quad \min\{\xi(\omega), \xi(\omega)'\} \geq \max\{\zeta(\omega), \zeta(\omega)'\}, \tag{17}$$

we have

$$\begin{gathered} \int_\Omega \phi\left(\rho\xi + (1-\rho)\zeta\right) \, \mathrm{d}q(\cdot | G_{t_*,t^*}) \geq \int_\Omega \phi\left(\rho\xi' + (1-\rho)\zeta\right) \, \mathrm{d}q(\cdot | G_{t_*,t^*}) \\ \Longleftrightarrow \\ \int_\Omega \phi\left(\rho\xi + (1-\rho)\zeta'\right) \, \mathrm{d}q(\cdot | G_{t_*,t^*}) \geq \int_\Omega \phi\left(\rho\xi' + (1-\rho)\zeta'\right) \, \mathrm{d}q(\cdot | G_{t_*,t^*}), \end{gathered} \tag{18}$$

where $q(\cdot | G_{t_*,t^*})$ is the conditional probability of q given G_{t_*,t^*}.

Define $s_*, s^* \in [-\infty, \infty]$ by

$$s_* = \inf u(X) \quad \text{and} \quad s^* = \sup u(X).$$

By assumption $s_* < 0$ and $s^* > 1$. Let $\psi \in B_b(\mathcal{G}, (s_* t_*, s^* t_*))$ and $\varphi \in B_b(\mathcal{G}, s_*(1 - t^*), s^*(1 - t^*))$. Then, for all $\omega \in G_{t_*,t^*}$,

$$\frac{\psi(\omega)}{\rho(\omega)} \in (s_*, s^*) \quad \text{and} \quad \frac{\varphi(\omega)}{1 - \rho(\omega)} \in (s_*, s^*).$$

Hence $\psi = \rho\xi$ and $\varphi = (1-\rho)\zeta$ for some $\xi, \zeta \in B_b(\mathcal{G}, u(X))$. Thus, by changing variables in (17) and (18), we obtain that for all $\psi, \psi' \in B_b(\mathcal{G}, (s_* t_*, s^* t_*))$ and $\varphi, \varphi' \in B_b(\mathcal{G}, (s_*(1-t^*), s^*(1-t^*)))$ such that

$$\omega \in G_{t_*,t^*} \quad \Rightarrow \quad \frac{\min\{\psi(\omega), \psi'(\omega)\}}{\rho(\omega)} \geq \frac{\max\{\varphi(\omega), \varphi'(\omega)\}}{1-\rho(\omega)} \tag{19}$$

we have

$$\begin{gathered} \int_\Omega \phi\left(\psi + \varphi\right) \mathrm{d}q(\cdot|G_{t_*,t^*}) \geq \int_\Omega \phi\left(\psi' + \varphi\right) \mathrm{d}q(\cdot|G_{t_*,t^*}) \\ \iff \\ \int_\Omega \phi\left(\psi + \varphi'\right) \mathrm{d}q(\cdot|G_{t_*,t^*}) \geq \int_\Omega \phi\left(\psi' + \varphi'\right) \mathrm{d}q(\cdot|G_{t_*,t^*}). \end{gathered} \tag{20}$$

Define the interval $I_{t_*,t^*} \subseteq u(X)$ by

$$I_{t_*,t^*} = \left(s_* \frac{1-t^*}{1-t_*}, s^* \frac{t_*}{t^*}\right).$$

Clearly $0 \in I_{t_*,t^*}$. For $s \in I_{t_*,t^*} \cap (-\infty, 0]$, (20) holds for all $\psi, \psi' \in B_b(\mathcal{G}, (st_*, s^* t_*))$ and $\varphi = b$ and $\varphi' = c$ with $b, c \in (s_*(1-t^*), s(1-t_*))$. Because q is nonatomic, $q(\cdot|G_{t_*,t^*})$ is nonatomic as well. Reasoning as in Strzalecki (2011, p. 67), by the uniqueness properties of the expected utility representation, for all $b, c \in (s_*(1-t^*), s(1-t_*))$ there are $\alpha(b,c) \in \mathbb{R}$ and $\beta(b,c) > 0$ such that for all $a \in (st_*, s^* t_*)$

$$\phi(a+b) = \alpha(b,c) + \beta(b,c)\phi(a+c). \tag{21}$$

For $s \in I_{t_*,t^*} \cap (0, +\infty)$ condition (20) holds for all $\psi, \psi' \in B_b(\mathcal{G}, (st^*, s^* t_*))$ and $\varphi = b$ and $\varphi' = c$ with $b, c \in (s_*(1-t^*), s(1-t^*))$. Hence, reasoning as above, we obtain that there are $\alpha(b,c) \in \mathbb{R}$ and $\beta(b,c) > 0$ such that (21) holds for all $a \in (st^*, s^* t_*)$.

Now we use (21) to show that ϕ is either affine or exponential. For every $s \in I_{t_*,t^*}$, define the interval $I_{t_*,t^*}(s) \subseteq u(X)$ by

$$I_{t_*,t^*}(s) = \begin{cases} (s_*(1-t^*) + st_*, s(1-t_*) + s^* t_*) & \text{if } s \leq 0, \\ (s_*(1-t^*) + st^*, s(1-t^*) + s^* t_*) & \text{otherwise.} \end{cases}$$

By Aczél (2005, Theorem 2 and its corollary), the function ϕ is affine or exponential on $I_{t_*,t^*}(s)$. If $I_{t_*,t^*}(s) \cap I_{t_*,t^*}(s') \neq \varnothing$, then ϕ is affine or exponential on $I_{t_*,t^*}(s) \cup I_{t_*,t^*}(s')$. Moreover, for $s \leq 0$, we have

$$s_*(1-t^*) + st_* \leq s \leq s(1-t_*) + s^* t_* \quad \iff \quad s_* \frac{1-t^*}{1-t_*} \leq s \leq s^*.$$

Thus $s \leq 0$ implies $s \in I_{t_*,t^*}(s)$. Similarly, for $s > 0$, we have

$$s_*(1-t^*) + st^* \leq s \leq s(1-t^*) + s^* t_* \quad \iff \quad s_* \leq s \leq \frac{t_*}{t^*} s^*.$$

Thus also $s > 0$ implies $s \in I_{t_*,t^*}(s)$. Overall, we obtain that

$$\bigcup_{s \in I_{t_*,t^*}} I_{t_*,t^*}(s) = I_{t_*,t^*}.$$

The function ϕ, therefore, is affine or exponential on I_{t_*,t^*}. This conclusion holds for all $t_*, t^* \in (0,1)$ such that $t_* < \bar{t} < t^*$. Thus ϕ is either affine or exponential on (s_*, s^*). Since ϕ is continuous, ϕ is either affine or exponential on $u(X)$.

It remains to show that ϕ is not exponential. Pick $t_*, t^* \in (0,1)$ such that $t_* < \bar{t} < t^*$. Let $\epsilon > 0$ be small enough so that $\epsilon < s^* t_*$ and $-\epsilon > s_*(1-t^*)$. Being $q(\cdot|G_{t_*,t^*})$ nonatomic, we can find $\xi \in B_b(\mathcal{G}, (s_* t_*, s^* t_*))$ such that $\xi = \epsilon$ with $q(\cdot|G_{t_*,t^*})$-probability $\frac{1}{2}$ and $\xi = 0$ with $q(\cdot|G_{t_*,t^*})$-probability $\frac{1}{2}$. Choose ξ', ζ, ζ' such that $\xi' = \frac{\epsilon}{2}$, $\zeta = 0$, and $\zeta' = -\xi$. It follows from (20) that

$$\frac{1}{2}\phi(\epsilon) + \frac{1}{2}\phi(0) \geq \phi\left(\frac{\epsilon}{2}\right) \quad \Longleftrightarrow \quad \phi(0) \geq \frac{1}{2}\phi\left(-\frac{\epsilon}{2}\right) + \frac{1}{2}\phi\left(\frac{\epsilon}{2}\right).$$

Thus ϕ is neither strictly convex nor strictly convex, which implies that ϕ is not exponential. $\square$

D.2 Proof of Theorem 10

Sufficiency. Assume axioms 1-3 and 5-8 are satisfied. Note that Axiom 9 is satisfied as well: by Lemma 13(i) if $x \succsim y$, then $x \succsim_{\text{st}} y$, which in turn implies $\alpha x + (1-\alpha)z \succsim \alpha y + (1-\alpha)z$ by Axiom 7. Thus we can pick u, ϕ, q, and T as in Theorem 9. By Theorem 7, to conclude the proof of sufficiency it is enough to show that T is constant-affine and concave. Without loss of generality, we assume that $[-1,1] \subseteq u(X)$.

Lemma 26. *For all acts f and g the following conditions are satisfied:*

(i) $f \succsim_{\text{st}} g$ *implies* $\alpha f + (1-\alpha)x \succsim_{\text{st}} \alpha f + (1-\alpha)x$ *for all* $\alpha \in [0,1]$ *and* $x \in X$.

(ii) $f \sim_{\text{st}} g$ *implies* $\alpha f + (1-\alpha)g \succsim_{\text{st}} f$ *for all* $\alpha \in [0,1]$.

Proof. (i). By Lemma 18(i) for all $A \in \mathcal{F}_{\text{st}}$ and h we have $fAh \succsim_{\text{st}} gAh$, and hence, by Axiom 7,

$$(\alpha f + (1-\alpha)x)A(\alpha h + (1-\alpha)x) \succsim (\alpha g + (1-\alpha)x)A(\alpha h + (1-\alpha)x).$$

which implies $(\alpha f + (1-\alpha)x)Ah \succsim (\alpha g + (1-\alpha)x)Ah$ given that $A \in \mathcal{F}_{\text{st}}$.

(ii). By Lemma 18(i) and Axiom 8, for all $A \in \mathcal{F}_{\text{st}}$ and h,

$$(\alpha f + (1-\alpha)g)Ah = \alpha fAh + (1-\alpha)gAh \succsim fAh.$$

The desired result follows. $\square$

Recall that T represents $\succsim_{\text{st}}$. For $\hat{f}$ such that $Tu(f) = [u(\hat{f})]$, Lemma 26(i) implies for all $\alpha \in [0,1]$ and $x \in X$, $\alpha f + (1-\alpha)x \sim_{\text{st}} \alpha \hat{f} + (1-\alpha)x$. Because T is projective (Lemma 10)

$$T(\alpha u(f) + (1-\alpha)u(x)) = T(\alpha u(\hat{f}) + (1-\alpha)u(x)) = \alpha T u(f) + (1-\alpha)[u(x)].$$

It follows from Lemma 14 that T is constant-affine. This implies T is positively homogeneous: for all $\alpha \geq 0$ and $\xi \in B_b(\mathcal{F}, u(X))$ such that $\alpha\xi(\Omega) \subseteq u(X)$

$$T(\alpha\xi) = \left[\alpha\hat{\xi}\right]$$

where $\hat{\xi}$ be a representative of $T\xi$ such that $\alpha\hat{\xi}(\Omega) \subseteq u(X)$.

Let $\xi \in B_b(\mathcal{F}, u(X))$, and let $\zeta \in B_0(\mathcal{F}_{\text{st}}, u(X))$ be of the form $\zeta = \sum_{i=1}^n a_i 1_{A_i}$ for some partition $A_1, \ldots, A_n$. Fix $\alpha \in [0,1]$ and define $\xi_i = \alpha\xi + (1-\alpha)a_i$. Then by applying the fact that T is constant-affine, together with Lemma 8, it follows that

$$T(\alpha\xi + (1-\alpha)\zeta) = T\left(\sum_{i=1}^n \xi_i \cdot 1_{A_i}\right) = \sum_{i=1}^n T(\xi_i)[1_{A_i}] = \sum_{i=1}^n \alpha T\xi + (1-\alpha)[a_i].$$

hence $T(\alpha\xi + (1-\alpha)\zeta) = \alpha T\xi + (1-\alpha)[\zeta]$.

Being $B_0(\mathcal{F}_{\text{st}}, u(X))$ supnorm-dense in $B_b(\mathcal{F}_{\text{st}}, u(X))$, we can apply Lemma 9 and extend the equation displayed above to all $\zeta \in B_b(\mathcal{F}_{\text{st}}, u(X))$. By positive homogeneity, for all $\xi \in B_b(\mathcal{F}, u(X))$ and $\zeta \in B_b(\mathcal{F}_{\text{st}}, u(X))$ such that $\xi + \zeta$ takes values in $u(X)$

$$T(\xi + \zeta) = 2T\left(\frac{1}{2}\xi + \frac{1}{2}\zeta\right) = \left[\hat{\xi} + \zeta\right] \tag{22}$$

where $\hat{\xi}$ is a representative of $T\xi$ such that $\hat{\xi} + \zeta$ takes values in $u(X)$.

To show that T is concave, pick $\alpha \in [0,1]$ and $\xi, \zeta \in B_b(\mathcal{F}, u(X))$ such that both ξ and ζ take values in $[-\frac{1}{3}, \frac{1}{3}]$ (the general case follows by positive homogeneity). Choose $\hat{\xi} \in T\xi$ and $\hat{\zeta} \in T\zeta$ such that $\hat{\xi}$ and $\hat{\zeta}$ take values in $[-\frac{1}{3}, \frac{1}{3}]$. By (22) we have

$$T\xi = \left[\hat{\zeta} + \hat{\xi} - \hat{\zeta}\right] = T\left(\zeta + \hat{\xi} - \hat{\zeta}\right).$$

We can then apply Lemma 14 and Lemma 26(ii) to obtain

$$T\left(\alpha\xi + (1-\alpha)\left(\zeta + \hat{\xi} - \hat{\zeta}\right)\right) \geq \alpha T(\xi) + (1-\alpha)T\left(\zeta + \hat{\xi} - \hat{\zeta}\right).$$

Applying (22) to both sides of the inequality yields $T(\alpha\xi + (1-\alpha)\zeta) \geq \alpha T(\xi) + (1-\alpha)T(\zeta)$.

Necessity. Assume $\succsim$ admits a multiple predictive representation $(u, \phi, \mathcal{G}, q, \Pi)$. Let

$$Tu(f) = \operatorname*{ess\,inf}_{\pi \in \Pi} E[u(f)|\mathcal{G}].$$

By Lemma 14 and Theorem 7 the operator $T\colon B_b(\mathcal{F}, u(X)) \to L_\infty(\mathcal{G}, q, u(X))$ is monotone, decomposable, normalized, σ-order continuous, constant-affine, and concave.

The preference relation $\succsim$ is obviously complete and transitive. Because u is not constant and ϕ is strictly increasing, it is also nontrivial: Axiom 1 is satisfied.

Assume $f(\omega) \succ g(\omega)$ for all ω. Because $E_\pi[u(f)|\mathcal{G}] > E_\pi[u(g)|\mathcal{G}]$ for all $\pi \in \Pi$, we have $Tu(f) \geq Tu(g)$. In addition, because Π is q-rectangular, Lemma 11 implies

$$E_q[Tu(f)] = \min_{\pi\in\Pi} E_\pi[u(f)] > \min_{\pi\in\Pi} E_\pi[u(g)] = E_q[Tu(g)].$$

Thus $Tu(f) > Tu(g)$ with positive q-probability. Since ϕ is strictly increasing, we deduce that $f \succ g$. So, Axiom 2 is satisfied.

Let (f_n) and (g_n) be bounded sequences such that $f_n \succsim g_n$ for every n. Suppose $f_n \to f$ and $g_n \to g$ pointwise. If $Y \subseteq X$ is a polytope, then Y is compact and u (being affine) is continuous on Y (Aliprantis and Border, 2006, Theorem 5.21). Thus the sequences $(u(f_n))$ and $(u(g_n))$ are bounded and converge pointwise to $u(f)$ and $u(g)$, respectively. By Lemma 9 and monotonicity of T, the sequences $(Tu(f_n))$ and $(Tu(g_n))$ are (essentially) bounded and converge q-almost surely to $Tu(f)$ and $Tu(g)$, respectively. Because ϕ is continuous and q is σ-additive, $E_q[\phi(Tu(f_n))] \to E_q[\phi(Tu(f))]$ and $E_q[\phi(Tu(g_n))] \to E_q[\phi(Tu(g))]$. We conclude that $f \succsim g$: Axiom 3 is satisfied.

If ϕ is affine, then by q-rectangularity of Π and Lemma 11,

$$f \succsim g \qquad \Longleftrightarrow \min_{\pi\in\Pi} E_\pi[u(f)] \geq \min_{\pi\in\Pi} E_\pi[u(g)],$$

which implies that Savage's P4 holds for all events in $\mathcal{F}$. Suppose now that ϕ is not affine and let $A, B \in \mathcal{F}_{\text{st}}$. By Lemma 21 there are events $G, H \in \mathcal{G}$ such that q-almost surely $1_G = \pi(A|\mathcal{G})$ and $1_H = \pi(B|\mathcal{G})$ for all $\pi \in \Pi$. Thus for all $x, y \in X$ such that $x > y$

$$xAy \succsim xBy \qquad \Longleftrightarrow \qquad q(G) \geq q(H).$$

It follows that Axiom 5 holds.

Let f, g, h such that $f \succ g$. Let $A_1, \ldots, A_n$ be a finite partition of $\mathcal{G}$-measurable events. By Lemma 21 each A_i satisfies the sure-thing principle. Because T is decomposable, by Lemma 8

$$\begin{aligned} V(hA_if) &= \int_\Omega \phi(Tu(hA_if))\,\mathrm{d}q = \int_\Omega \phi(Tu(h)\cdot[1_{A_i}] + Tu(f)\cdot[1_{A_i^c}])\,\mathrm{d}q \\ &= \int_{A_i} \phi(Tu(h))\,\mathrm{d}q + \int_{A_i^c} \phi(Tu(f))\,\mathrm{d}q. \end{aligned}$$

A similar condition holds for $V(hA_ig)$. Since q is nonatomic, for every $\varepsilon > 0$ we can choose $A_1, \ldots, A_n$ so that $\max_i |V(hA_if) - V(f)| \leq \varepsilon$ and $\max_i |V(hA_ig) - V(g)| \leq \varepsilon$. It follows that Axiom 6 holds.

Assume $f \succsim_{\text{st}} g$. By Lemma 21 we have $Tu(f) \geq Tu(g)$. Being T constant-affine, this implies $Tu(\alpha f + (1-\alpha)x) \geq Tu(\alpha g + (1-\alpha)x)$ for all $\alpha \in [0,1]$ and $x \in X$. It follows that $\alpha f + (1-\alpha)x \succsim \alpha g + (1-\alpha)x$. Thus, Axiom 7 holds.

Assume $f \sim_{\text{st}} g$. By Lemma 21 we have $Tu(f) = Tu(g)$. Being T concave, this implies $Tu(\alpha f + (1-\alpha)g) \geq Tu(f)$. It follows that $\alpha f + (1-\alpha)g \succsim f$. Hence, Axiom 8 holds.

D.3 Proof of Theorem 11

Since u_1 and u_2 both represent $\succsim$ on X, by the uniqueness properties of the expected utility representation, u_2 is a positive affine transformation of u_1. For the rest of the proof, we can assume without loss of generality that $u_1 = u_2 = u$.

We first show that if ϕ_1 is affine, then ϕ_2 is affine as well. We prove the contrapositive statement. Suppose ϕ_2 is not affine. By Lemma 21 and Proposition 5 the preference $\succsim$ admits a multiple predictive representation $(u, \phi_2, \mathcal{F}_{\text{st}}, \Pi_2)$. Moreover $\mathcal{G}_1 \subseteq \mathcal{F}_{\text{st}}$ again by Lemma 21. Thus for all acts f and g that are $\mathcal{G}_1$-measurable

$$\int_\Omega \phi_1(u(f)))\,\mathrm{d}q_1 \geq \int_\Omega \phi_1(u(g)))\,\mathrm{d}q_1 \iff \int_\Omega \phi_2(u(f))\,\mathrm{d}q_2 \geq \int_\Omega \phi_2(u(g))\,\mathrm{d}q_2.$$

In particular, for all $A, B \in \mathcal{G}_1$, $q_1(A) \geq q_1(B)$ if and only if $q_2(A) \geq q_2(B)$. Because q_1 is nonatomic then by standard arguments we obtain $q_1 = q_2$. Hence, by the uniqueness properties of the expected utility representation, ϕ_1 must be a positive affine transformation of ϕ_2. We conclude as desired that ϕ_1 is not affine.

We have therefore two cases to consider: either both ϕ_1 and ϕ_2 are affine, or both ϕ_1 and ϕ_2 are not affine. Suppose first that both ϕ_1 and ϕ_2 are affine. Because the two are also strictly increasing, then ϕ_2 is an affine transformation of ϕ_1. For $i \in \{1,2\}$, the set Π_i is q_i-rectangular and therefore for all acts f and g we can apply Lemma 11 and the affinity of ϕ_i to conclude

$$\begin{aligned} f \succsim g &\iff \min_{\pi_i \in \Pi_i} E_{\pi_i}[\phi_i(u(f))] \geq \min_{\pi_i \in \Pi_i} E_{\pi_i}[\phi_i(u(g))] \\ &\iff \min_{\pi_i \in \Pi_i} E_{\pi_i}[u(f)] \geq \min_{\pi_i \in \Pi_i} E_{\pi_i}[u(g)]. \end{aligned}$$

Theorem 5 implies $\Pi_2 = \Pi_1$.

Assume now that both ϕ_1 and ϕ_2 are not affine. By Lemma 21 and Proposition 5 the preference $\succsim$ admits the representations $(u, \phi_1, \mathcal{F}_{\text{st}}, \Pi_1)$ and $(u, \phi_2, \mathcal{F}_{\text{st}}, \Pi_2)$. Moreover $\mathcal{G}_1$ is Π_1-equivalent to $\mathcal{F}_{\text{st}}$ and $\mathcal{G}_2$ is Π_2-equivalent to $\mathcal{F}_{\text{st}}$. Let q_i' be the common restriction of Π_i, $i = 1, 2$, on $\mathcal{F}_{\text{st}}$. It is non-atomic. For all acts f and g that are $\mathcal{F}_{\text{st}}$-measurable

$$\int_\Omega \phi_1(u(f))\,\mathrm{d}q_1' \geq \int_\Omega \phi_1(u(g))\,\mathrm{d}q_1' \iff \int_\Omega \phi_2(u(f))\,\mathrm{d}q_2' \geq \int_\Omega \phi_2(u(g))\,\mathrm{d}q_2'.$$

By the uniqueness properties of the subjective expected utility representation, $q_1' = q_2'$ and ϕ_2 is a positive affine transformation of ϕ_1. It follows from Proposition 4 that for all act f

$$\operatorname*{ess\,inf}_{\pi_1 \in \Pi_1} E_{\pi_1}[u(f)|\mathcal{F}_{\text{st}}] = \operatorname*{ess\,inf}_{\pi_2 \in \Pi_2} E_{\pi_2}[u(f)|\mathcal{F}_{\text{st}}].$$

Thus, because both Π_1 and Π_2 are q_1'-rectangular, Lemma 11 implies

$$\min_{\pi_1 \in \Pi_1} E_{\pi_1}[u(f)] = \min_{\pi_2 \in \Pi_2} E_{\pi_2}[u(f)].$$

Thus $\Pi_1 = \Pi_2$ by Theorem 5.

E Proof of Proposition 1

The proof of Proposition 1 is divided in lemmas. Given $\mathcal{P} \subseteq \Delta$, we denote by $\mathcal{G}_{\mathcal{P}}$ the collection of *zero-one* events:

$$\mathcal{G}_{\mathcal{P}} = \{A \in \mathcal{F} : p(A) \in \{0,1\} \text{ for all } p \in \mathcal{P}\}. \tag{23}$$

By Breiman, LeCam, and Schwartz (1964, Proposition 1), the collection $\mathcal{G}_{\mathcal{P}}$ is a σ-algebra. Given a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, we say that a kernel $k \colon \Omega \to \mathcal{P}$ *witnesses the sufficiency* of $\mathcal{G}$ for $\mathcal{P}$ if, for every $p \in \mathcal{P}$, k is a regular conditional probability of p with respect to $\mathcal{G}$.

Lemma 27. *Let $\mathcal{P} \subseteq \Delta$. A kernel $k \colon \Omega \to \mathcal{P}$ identifies $\mathcal{P}$ if and only if it witnesses the sufficiency of $\mathcal{G}_{\mathcal{P}}$ for $\mathcal{P}$.*

Proof. "If." Being $\mathcal{F}$ standard Borel, we can pick a countable algebra of events $\mathcal{A}$ that generates $\mathcal{F}$. Since k is $\mathcal{G}_{\mathcal{P}}$-measurable, for every $A \in \mathcal{F}$ and $p \in \mathcal{P}$ the events $\{\omega : k(\omega, A) > p(A)\}$ and $\{\omega : k(\omega, A) < p(A)\}$ have p-probability 0 or 1. From $p(A) = \int_{\Omega} k(\omega, A)\, \mathrm{d}p(\omega)$ it follows that $p(\{\omega : k(\omega, A) = p(A)\}) = 1$. Since $\mathcal{A}$ is countable and generates $\mathcal{F}$ we obtain $p(\{\omega : k(\omega) = p\}) = 1$.

"Only if." For every $A \in \mathcal{F}$, $t \in \mathbb{R}$, and $p \in \mathcal{P}$, the probability $p(\{\omega : k(\omega, A) \geq t\})$ equals 1 if $p(A) \geq t$ and 0 otherwise. Hence $\{\omega : k(\omega, A) \geq t\} \in \mathcal{G}_{\mathcal{P}}$. We deduce that k is $\mathcal{G}_{\mathcal{P}}$-measurable. Moreover, for all $A \in \mathcal{F}$ and $B \in \mathcal{G}_{\mathcal{P}}$

$$\int_B k(\omega, A) dp(\omega) = p(B) \int_{\Omega} p(A) dp(\omega) = p(A)p(B) = p(A \cap B).$$

where the last two equalities follow from $p(B)$ being in $\{0,1\}$. We conclude that k is a common regular conditional probability of all $p \in \mathcal{P}$ with respect to $\mathcal{G}_{\mathcal{P}}$. □

Lemma 27 can be used to relate our definition of identifiability to the notion of *Dynkin space* (Dynkin, 1978; Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio, 2013). Some of the results that appear in this section are already discussed in the original paper by Dyknin and in Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013). See, in particular, their Appendix B.

Lemma 28. *If a kernel $k \colon \Omega \to \mathcal{P}$ identifies $\mathcal{P} \subseteq \Delta$ and μ is a prior on $\mathcal{P}$, then (i) k is a regular conditional probability of π_μ given $\mathcal{G}_{\mathcal{P}}$, and (ii) $\sigma(k)$ and $\mathcal{G}_{\mathcal{P}}$ are π_μ-equivalent.*

Proof. (i). For all $A \in \mathcal{F}$ and $B \in \mathcal{G}_{\mathcal{P}}$,

$$\pi_\mu(A \cap B) = \int_{\mathcal{P}} p(A \cap B)\,\mathrm{d}\mu(p) = \int_{\mathcal{P}} \left(\int_{\Omega} 1_B k(\omega, A)\,\mathrm{d}p(\omega) \right) d\mu(p).$$

It follows that $\pi_\mu(A \cap B) = \int_B k(\omega, A)\,\mathrm{d}\pi_\mu(\omega)$. By varying A and B we conclude that k is a regular conditional probability of π_μ with respect to $\mathcal{G}_{\mathcal{P}}$.

(ii). By (i) the kernel k is a regular conditional probability of π_μ with respect to $\mathcal{G}_{\mathcal{P}}$. Thus each $A \in \mathcal{G}_{\mathcal{P}}$ is π_μ-equivalent to $B = \{\omega : k(A, \omega) = 1\} \in \sigma(k)$. Moreover, $\sigma(k) \subseteq \mathcal{G}_{\mathcal{P}}$. We conclude that $\sigma(k)$ and $\mathcal{G}_{\mathcal{P}}$ are π_μ-equivalent. □

Let $\mathcal{P} \subseteq \Delta$. For every $A \in \mathcal{F}$ we define $A^\star \in \Sigma_{\mathcal{P}}$ by $A^\star = \{p \in \mathcal{P} : p(A) = 1\}$. We also define the collection $\Sigma^\star_{\mathcal{P}} = \{A^\star : A \in \mathcal{G}_{\mathcal{P}}\} \subseteq \Sigma_{\mathcal{P}}$. It is a σ-algebra, as shown by Breiman, LeCam, and Schwartz (1964, Proposition 1).

Lemma 29. *If $\mathcal{P} \subseteq \Delta$ is identifiable, then (i) $\Sigma_{\mathcal{P}} = \Sigma^\star_{\mathcal{P}}$, and (ii) a prior μ on $\mathcal{P}$ is nonatomic if and only if π_μ is nonatomic on $\mathcal{G}_{\mathcal{P}}$.*

Proof. (i). Let k identify $\mathcal{P}$. For every $A \in \mathcal{F}$ and $t \in \mathbb{R}$ we have

$$\{p \in \mathcal{P} : p(A) \geq t\} = \{\omega : k(\omega, A) \geq t\}^\star.$$

Since k is $\mathcal{G}_{\mathcal{P}}$-measurable we have $\{p \in \mathcal{P} : p(A) \geq t\} \in \Sigma^\star_{\mathcal{P}}$. Since $\Sigma^\star_{\mathcal{P}} \subseteq \Sigma_{\mathcal{P}}$, and the sets of the form $\{p \in \mathcal{P} : p(A) \geq t\}$ generate $\Sigma_{\mathcal{P}}$, it follows that $\Sigma^\star_{\mathcal{P}} = \Sigma_{\mathcal{P}}$.

(ii). Observe that $\mu(A^\star) = \pi_\mu(A)$ for every $A \in \mathcal{G}_{\mathcal{P}}$. If μ is nonatomic, given $A \in \mathcal{G}_P$ and $\alpha \in [0,1]$, by (i) there is $B \in \mathcal{G}_{\mathcal{P}}$ such that $B^\star \subseteq A^\star$ and $\mu(B^\star) = \alpha\mu(A^\star)$. Because $B^\star \cap A^\star = (A \cap B)^\star$ then $\pi_\mu(A \cap B) = \alpha\pi_\mu(A)$. The proof that if π_μ is nonatomic then so is μ follows from an analogous argument. □

Lemma 30. *Let $\succsim$ admit an identifiable representation $(u, \phi, \mathcal{P}, \mu)$. Then it admits a predictive representation $(u, \phi, \mathcal{G}_{\mathcal{P}}, \pi_\mu)$.*

Proof. By Lemma 29 the measure π_μ is nonatomic on $\mathcal{G}_P$. To conclude the proof it remains to show that for all $\xi \in B_b(\mathcal{F}, u(X))$

$$\int_{\mathcal{P}} \phi\left(\int_{\Omega} \xi\,\mathrm{d}p \right)\,\mathrm{d}\mu(p) = E_{\pi_\mu}[\phi(E_{\pi_\mu}[\xi|\mathcal{G}_{\mathcal{P}}])].$$

Assume first that ξ is $\mathcal{G}_{\mathcal{P}}$-measurable. Each $p \in \mathcal{P}$ satisfies

$$p(\{\omega : \xi(\omega) = E_p[\xi]\}) = 1.$$

Hence $E_p[\phi(\xi)] = \phi(E_p[\xi])$ for all $p \in \mathcal{P}$, which implies $\int_\Delta \phi\left(\int_\Omega \xi\,\mathrm{d}p\right)\,\mathrm{d}\mu(p) = E_{\pi_\mu}[\phi(\xi)]$.

For an arbitrary $\mathcal{F}$-measurable ξ, Lemma 27 implies

$$\int_{\mathcal{P}} \phi\left(\int_{\Omega} \xi\,\mathrm{d}p \right)\,\mathrm{d}\mu(p) = \int_{\mathcal{P}} \phi\left(\int_{\Omega} \left(\int_{\Omega} \xi\,\mathrm{d}k(\omega) \right)\,\mathrm{d}p(\omega) \right)\,\mathrm{d}\mu(p)$$

where k identifies $\mathcal{P}$. The function $\omega \mapsto \int_\Omega \xi \,\mathrm{d}k(\omega)$ is $\mathcal{G}_\mathcal{P}$-measurable and therefore

$$\int_\mathcal{P} \phi\left(\int_\Omega \left(\int_\Omega \xi \,\mathrm{d}k(\omega)\right) \mathrm{d}p(\omega)\right) \mathrm{d}\mu(p) = \int_\Omega \phi\left(\int_\Omega \xi \,\mathrm{d}k(\omega)\right) \mathrm{d}\pi_\mu(\omega).$$

The right-hand side is equal to $E_{\pi_\mu}[\phi(E_{\pi_\mu}[\xi|\mathcal{G}_\mathcal{P}])]$, being k a regular conditional probability for π_μ (Lemma 28). □

Lemma 31. *Let $\succsim$ admit a predictive representation $(u, \phi, \mathcal{G}, \pi)$. Then it admits an identifiable representation $(u, \phi, \mathcal{P}, \mu)$ where $\pi_\mu = \pi$ and $\mathcal{G}_\mathcal{P}$ is π-equivalent to $\mathcal{G}$.*

Proof. Since $(\Omega, \mathcal{F})$ is standard Borel, π admits a regular conditional probability $k : \Omega \to \Delta$ with respect to $\mathcal{G}$. We now show that for each $A \in \mathcal{F}$ and for μ-almost all p

$$p(\{\omega : k(\omega, A) = p(A)\}) = 1.$$

Indeed, the functions $\omega \mapsto k(\omega, A)$ and $\omega \mapsto k(\omega, A)^2$ are $\mathcal{G}$-measurable, and therefore, by definition of regular conditional probability, for π-almost all ω

$$\int_\Omega k(\omega', A) k(\omega, \mathrm{d}\omega') = k(\omega, A) \quad \text{and} \quad \int_\Omega k(\omega', A)^2 k(\omega, \mathrm{d}\omega') = k(\omega, A)^2.$$

Hence, for μ-almost all p

$$\int_\Omega k(\omega, A)^2 \,\mathrm{d}p(\omega) + p(A)^2 = 2p(A) \int_\Omega k(\omega, A) \,\mathrm{d}p(\omega),$$

which is equivalent to $\int_\Omega (k(\omega, A) - p(A))^2 \,\mathrm{d}p(\omega) = 0$. The desired conclusion follows.

Being the state space standard Borel, we can find a countable collection $\mathcal{A}$ of events that generates $\mathcal{F}$. For μ-almost all p

$$p(\{\omega : k(\omega, A) = p(A) \text{ for all } A \in \mathcal{A}\}) = 1,$$

which implies that $p(\{\omega : k(\omega) = p\}) = 1$. Let $\mathcal{P} = \{p : p(\{\omega : k(\omega) = p\}) = 1\}$.

The function $k \colon \Omega \to \mathcal{P}$ is $(\mathcal{G}, \Sigma_\mathcal{P})$-measurable and identifies $\mathcal{P}$. We define a prior μ on Σ as the pushforward of π under k. A simple change of variables shows that

$$E_\pi\Big[\phi\Big(E_\pi[u(f)|\mathcal{G}]\Big)\Big] = \int_\Omega \phi\left(\int_\Omega u(f(\omega')) k(\omega, \mathrm{d}\omega')\right) \mathrm{d}\pi(\omega) = \int_\mathcal{P} \phi\left(\int_\Omega u(f) \,\mathrm{d}p\right) \mathrm{d}\mu(p).$$

By a similar reasoning, for every $A \in \mathcal{F}$

$$\pi_\mu(A) = \int_\mathcal{P} p(A) \,\mathrm{d}\mu(p) = \int_\Omega k(\omega, A) \,\mathrm{d}\pi(\omega) = \pi(A).$$

It remains to show μ is nonatomic. Let $A_1, \ldots, A_n$ be a partition of events in $\mathcal{G}$ that have equal π-probability. The sets $A_1^\star, \ldots, A_n^\star$ are pairwise disjoint, and satisfy

$$\mu(A_i^\star) = \pi(\{\omega : k(\omega, A_i) = 1\}) = \pi(A_i) = \frac{1}{n}.$$

It follows that μ is nonatomic. Hence, the tuple $(u, \phi, \mathcal{P}, \mu)$ is an identifiable representation. It remains to show $\mathcal{G}$ and $\mathcal{G}_{\mathcal{P}}$ are π-equivalent. If $A \in \mathcal{G}$ then

$$\mu(\{p : p(A) \in \{0,1\}\}) = \pi(\{\omega : k(\omega, A) \in \{0,1\}\}) = \pi(\{\omega : 1_A(\omega) \in \{0,1\}\}) = 1.$$

Hence $\mu(A^\star) + \mu((A^c)^\star) = 1$, and in particular $\mu(A^\star) = \pi(A)$. Lemma 29 shows $\Sigma_{\mathcal{P}} = \Sigma^\star_{\mathcal{P}}$. Thus there exists $B \in \mathcal{G}_{\mathcal{P}}$ such that $A^\star = B^\star$, and hence $(A^\star)^c = (B^\star)^c = (B^c)^\star$. Then

$$\pi(A) = \mu(A^\star) = \mu(B^\star) = \pi(B) \quad \text{and} \quad \pi(A^c) = \mu((A^c)^\star) = \mu((B^c)^\star) = \pi(B^c)$$

so $\pi(A \triangle B) = 0$. Conversely, if $A \in \mathcal{G}_{\mathcal{P}}$ then $k(\omega, A) \in \{0,1\}$ for every ω. This implies $\pi(A \triangle B) = 0$ for $B = \{\omega : k(\omega, A) = 1\} \in \mathcal{G}$. □

For a preference relation $\succsim$ that admits a predictive representation $(u, \phi, \mathcal{G}, \pi)$, an event $A \in \mathcal{F}$ is null if and only if $\pi(A) = 0$ (Lemma 22). Thus Proposition 1 follows from Lemmas 30 and 31, given that $\sigma(k)$ and $\mathcal{G}_{\mathcal{P}}$ are π_μ-equivalent (Lemma 28).

F Proofs of the results in Section 5

F.1 Proof of Theorem 1

We prove the equivalent statement (as implied by Proposition 1) that $\succsim$ satisfies Axioms 1-6 if and only if it admits a predictive representation.

Sufficiency. Assume axioms 1-6 are satisfied. Note that Axiom 9 is satisfied as well: by Lemma 13(i) if $x \succsim y$, then $x \succsim_{\text{st}} y$, which in turn implies $\alpha x + (1-\alpha)z \succsim \alpha y + (1-\alpha)z$ by Axiom 4. Thus we can pick u, ϕ, q, and T as in Theorem 9. By Theorem 7, to conclude the proof of sufficiency it is enough to show that T is affine. To this end, we first show that

$$f \succsim_{\text{st}} g \Rightarrow \alpha f + (1-\alpha)h \succsim_{\text{st}} \alpha f + (1-\alpha)h \text{ for all } \alpha \in [0,1], h \in \mathfrak{F}. \tag{24}$$

By Lemma 18(i) we have $fAh \succsim_{\text{st}} gAh$ for all $A \in \mathcal{F}_{\text{st}}$ and h. By Axiom 7

$$(\alpha f + (1-\alpha)h)A(\alpha h + (1-\alpha)h) \succsim (\alpha g + (1-\alpha)h)A(\alpha h + (1-\alpha)h),$$

thus, since $A \in \mathcal{F}_{\text{st}}$, we have $(\alpha f + (1-\alpha)h)Ah' \succsim (\alpha g + (1-\alpha)h)Ah'$ for all $h' \in \mathfrak{F}$. Hence (24) follows. Now recall T represents $\succsim_{\text{st}}$. Hence for $\hat{f}$ such that $Tu(f) = [u(\hat{f})]$ and $\hat{g}$ such that $Tu(g) = [u(\hat{g})]$, (24) implies that for all $\alpha \in [0,1]$

$$\alpha f + (1-\alpha)g \sim_{\text{st}} \alpha \hat{f} + (1-\alpha)g \sim_{\text{st}} \alpha \hat{f} + (1-\alpha)\hat{g}.$$

Thus, being u affine,

$$T(\alpha u(f) + (1-\alpha)u(g)) = [\alpha u(\hat{f}) + (1-\alpha)u(\hat{g})] = \alpha Tu(f) + (1-\alpha)Tu(g).$$

It follows from Lemma 14 that T is affine.

Necessity. Assume $\succsim$ admits a predictive representation $(u, \phi, \mathcal{G}, \pi)$. By Theorem 10, to conclude the proof of necessity it is enough to show that Axiom 4 is satisfied. Let $f \succsim_{\text{st}} g$. By Lemma 21 we have $E_\pi[u(f)|\mathcal{G}] \geq E_\pi[u(g)|\mathcal{G}]$. This implies for all $\alpha \in [0,1]$ and $h \in \mathfrak{F}$

$$\begin{aligned} E_\pi[u(\alpha f + (1-\alpha)h)|\mathcal{G}] &= \alpha E_\pi[u(f)|\mathcal{G}] + (1-\alpha)E_\pi[u(h)|\mathcal{G}] \\ &\geq \alpha E_\pi[u(g)|\mathcal{G}] + (1-\alpha)E_\pi[u(h)|\mathcal{G}] = E_\pi[u(\alpha g + (1-\alpha)h)|\mathcal{G}]. \end{aligned}$$

It follows that $\alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h$. Hence, Axiom 4 holds.

F.2 Proof of Theorem 2

The uniqueness properties of the predictive representation follow from Theorem 11. Consider now two identifiable representations $(u_1, \phi_1, \mathcal{P}_1, \mu_1)$ and $(u_2, \phi_2, \mathcal{P}_2, \mu_2)$. By Proposition 1, the preference $\succsim$ admits predictive representations $(u_1, \phi_1, \sigma(k_1), \pi_{\mu_1})$ and $(u_1, \phi_1, \sigma(k_2), \pi_{\mu_2})$. Thus u_2 is a positive affine transformation of u_1; normalizing the utility indexes, ϕ_2 is a positive affine transformation of ϕ_1; $\pi_{\mu_1} = \pi_{\mu_2}$; if ϕ_1 is not affine, then $\sigma(k_1)$ and $\sigma(k_2)$ are π_{μ_1}-equivalent.

It remains to show that, if ϕ_1 is not affine, than $\mu_1(S \cap \mathcal{P}_1) = \mu_2(S \cap \mathcal{P}_2)$ for all $S \in \Sigma$. By Lemma 28, for every $i \in \{1, 2\}$ the kernel k_i is a regular conditional probability of π_{μ_i} given $\sigma(k_i)$. Thus, if ϕ_1 is not affine, k_1 and k_2 are equal π_{μ_1}-almost surely, being that $\sigma(k_1)$ and $\sigma(k_2)$ are π_{μ_1}-equivalent (see Lemma 1). For all $A \in \mathcal{F}$ and $t \in [0,1]$, we obtain from the condition of identifiability that

$$\pi_{\mu_i}(\{\omega : k_i(\omega, A) \leq t\}) = \int_{\mathcal{P}} p(\{\omega : k_i(\omega, A) \leq t\})\,\mathrm{d}\mu_i(p) = \mu_i(\{p \in \mathcal{P} : p(A) \leq t\}).$$

Since $\pi_{\mu_1} = \pi_{\mu_2}$, it follows that $\mu_1(S \cap \mathcal{P}_1) = \mu_2(S \cap \mathcal{P}_2)$ for the set $S = \{p \in \Delta : p(A) \leq t\}$. Since sets of this form generate Σ, the desired result follows.

G Proofs of the results in Section 6

G.1 Proof of Proposition 2

The equivalence of (i) and (iii) follows from Proposition 4. Let k witness the identifiability of $\mathcal{P}$. By Lemma 31 the preference $\succsim$ admits a predictive representation $(u, \phi, \mathcal{G}_{\mathcal{P}}, \pi_\mu)$. By Lemma 28 the kernel k is a regular conditional probability of π_μ given $\mathcal{G}_{\mathcal{P}}$. Thus, being (i) and (iii) equivalent,

$$f \succsim_{\text{st}} g \iff \int_\Omega u(f)\,\mathrm{d}k(\omega) \geq \int_\Omega u(g)\,\mathrm{d}k(\omega) \text{ for } \pi_\mu\text{-almost all } \omega.$$

The event $A = \{\omega : \int_\Omega u(f) - u(g)\,\mathrm{d}k(\omega) \geq 0\}$ belongs to $\mathcal{G}_{\mathcal{P}}$. Thus $\pi_\mu(A) = 1$ if and only if $\mu(\{p : p(A) = 1\}) = 1$. Because each p satisfies $p(\{\omega : k(\omega) = p\}) = 1$, we obtain

$$f \succsim_{\text{st}} g \iff \int_\Omega u(f)\,\mathrm{d}p \geq \int_\Omega u(g)\,\mathrm{d}p \text{ for } \mu\text{-almost all } p.$$

G.2 Proof of Proposition 3

By Lemma 21 the σ-algebras $\mathcal{F}_{\mathrm{st}}$ and $\mathcal{G}$ are π-equivalent. By Proposition 1 the preference $\succsim$ admits a predictive representation $(u, \phi, \sigma(k), \pi_\mu)$. By Theorem 2 we obtain $\pi_\mu = \pi$ and $\sigma(k)$ is π-equivalent to $\mathcal{G}$. From Lemma 22 it follows that $\mathcal{F}_{\mathrm{st}}$, $\sigma(k)$, and $\mathcal{G}$ are all equivalent up to null events.

H Proofs of the results in Section 7

H.1 Preliminary results on identified sets

Let $\mathcal{C} \subseteq \mathscr{C}$ be a collection of identified sets and μ a prior on $\mathcal{C}$. Define $\mathcal{G}_{\mathcal{C}} \subseteq \mathcal{F}$ by

$$\mathcal{G}_{\mathcal{C}} = \left\{ A \in \mathcal{F} : \forall C \in \mathcal{C}, \min_{p \in C} p(A) = \max_{p \in C} p(A) \in \{0,1\} \right\}.$$

We can think of elements of $\mathcal{G}_{\mathcal{C}}$ as *zero-one* events, a generalization of (23).

Lemma 32. *The collection $\mathcal{G}_{\mathcal{C}}$ is a σ-algebra.*

Proof. It is clear that $\Omega \in \mathcal{G}_{\mathcal{C}}$, and that $A \in \mathcal{G}_{\mathcal{C}}$ implies $A^c \in \mathcal{G}_{\mathcal{C}}$. Now let $A, B \in \mathcal{G}_{\mathcal{C}}$. If $\min_{p\in C} p(A) = 0$ or $\min_{p \in C} p(B) = 0$, then $p(A \cap B) = 0$ for all $p \in C$. Otherwise, $p(A \cap B) = 1$ for all $p \in C$. We deduce that $A \cap B \in \mathcal{G}_{\mathcal{C}}$. So, $\mathcal{G}_{\mathcal{C}}$ is an algebra. Let (A_n) be a sequence in $\mathcal{G}_{\mathcal{C}}$ such that $A_n \uparrow A$. From Lemma 5 it follows that $\min_{p\in C} p(A_n) \to \min_{p \in C} p(A)$ and $\max_{p\in C} p(A_n) \to \max_{p\in C} p(A)$. Thus $A \in \mathcal{G}_{\mathcal{C}}$. We conclude that $\mathcal{G}_{\mathcal{C}}$ is a σ-algebra. □

Let $\mathcal{P} \subseteq \Delta$ and let $K : \Omega \to \mathscr{C}$ be a measurable function that set-identifies $\mathcal{P}$ such that $\mathcal{C} = \{K(\omega) : \omega \in \Omega\}$. The set-valued kernel K is $(\mathcal{G}_{\mathcal{C}} \backslash \mathfrak{S}_{\mathcal{C}})$-measurable. Indeed, for every $C \in \mathcal{C}$, $p \in C$, $\xi \in B(\mathcal{F}, \mathbb{R})$, and $t \in \mathbb{R}$ we have

$$p(\{\omega : \sigma_{K(\omega)}(\xi) \le t\}) = p(\{\omega : \sigma_C(\xi) \le t\}) = \begin{cases} 1 & \text{if } \sigma_C(\xi) \le t, \\ 0 & \text{otherwise.} \end{cases}$$

Thus $\{\omega : \sigma_{K(\omega)}(\xi) \le t\} \in \mathcal{G}_{\mathcal{C}}$. By varying ξ and t we deduce that K is $\mathcal{G}_{\mathcal{C}}$-measurable.

Let $C \in \mathcal{C}$. By contruction, all $p \in C$ agree on $\mathcal{G}_{\mathcal{C}}$. Letting p_0 be the common restriction on $\mathcal{G}_{\mathcal{C}}$, we have that for every $\xi \in B(\mathcal{F}, \mathbb{R})$ and $G \in \mathcal{G}_{\mathcal{C}}$

$$\min_{p \in C} \int_G \xi \, \mathrm{d}p = p_0(G) \sigma_C(\xi) = p_0(G) \int_\Omega \sigma_K(\xi) \, \mathrm{d}p_0 = \int_G \sigma_K(\xi) \, \mathrm{d}p_0 \tag{25}$$

where we used the fact that $p_0(G) \in \{0,1\}$. We conclude in particular that K is a regular conditional probability of C given $\mathcal{G}_{\mathcal{C}}$ (cfr. Definition 10).

For every $A \in \mathcal{F}$, define $A^\star = \{C \in \mathcal{C} : \min_{p\in C} p(A) = 1\}$. Let $q_\mu : \mathcal{G}_{\mathcal{C}} \to [0,1]$ be given by $q_\mu(A) = \mu(A^\star)$.

Lemma 33. *The set function q_μ is a countably additive probability measure.*

Proof. Because $\varnothing^\star = \varnothing$ and $\Omega^\star = \mathcal{C}$, we have $q_\mu(\varnothing) = 0$ and $q_\mu(\Omega) = 1$. If A and B are disjoint, then $A^\star$ and $B^\star$ are disjoint. If in addition $A, B \in \mathcal{G}_C$ then $(A \cup B)^\star = A^\star \cup B^\star$. Thus $q_\mu(A \cup B) = q_\mu(A) + q_\mu(B)$. If $A_n \downarrow \varnothing$, then $\min_{p \in C} p(A_n) \to 0$ for all $C \in \mathcal{C}$ (Lemma 5). Thus $A_n^\star \downarrow \varnothing$, which implies $q_\mu(A_n) \downarrow 0$. We conclude that q_μ is a probability measure. □

The next lemma generalizes the notion of predictive probability.

Lemma 34. *There exists a unique $\Pi_\mu \in \mathscr{C}$ such that for all $\xi \in B(\mathcal{F}, \mathbb{R})$*

$$\sigma_{\Pi_\mu}(\xi) = \int_{\mathcal{C}} \sigma_C(\xi)\, \mathrm{d}\mu(C).$$

The set Π_μ satisfies the following properties:

(i). $\pi(G) = q_\mu(G)$ *for all* $G \in \mathcal{G}_{\mathcal{C}}$ *and* $\pi \in \Pi_\mu$.

(ii). K *is a regular conditional probability of* Π_μ *given* $\mathcal{G}_{\mathcal{C}}$.

Proof. Define the functional $I \colon B(\mathcal{F}, \mathbb{R}) \to \mathbb{R}$ by $I(\xi) = \int_{\mathcal{C}} \sigma_C(\xi)\, \mathrm{d}\mu(C)$. It can be easily verified that I is monotone, normalized, constant-affine and concave. Lemma 5 implies I is pointwise continuous, and thus by Theorem 5 there exists a unique $\Pi_\mu \in \mathcal{C}$ such that $I = \sigma_{\Pi_\mu}$. This implies that if $G \in \mathcal{G}_{\mathcal{C}}$ then $\min_{\pi \in \Pi_\mu} \pi(G) = 1 \cdot \mu(G^\star)$. Similarly, $\min_{\pi \in \Pi_\mu} \pi(G^c) = \mu((G^c)^\star) = \mu((G^\star)^c)$. It follows that $\min_{\pi \in \Pi_\mu} \pi(G) = \max_{\pi \in \Pi_\mu} \pi(G) = q_\mu(G)$.

To establish (ii), let $G \in \mathcal{G}_C$ and $\xi \in B(\mathcal{F}, \mathbb{R})$. It follows from (25) that $\sigma_C(\xi) = \sigma_C(1_G \cdot \sigma_K(\xi))$ and hence

$$\min_{\pi \in \Pi} \int_G \xi\, \mathrm{d}\pi = \int_{\mathcal{C}} \sigma_C(1_G \cdot \sigma_K(\xi))\, \mathrm{d}\mu(C) = \min_{\pi \in \Pi} \int_G \sigma_K(\xi)\, \mathrm{d}\pi = \int_G \sigma_K(\xi)\, \mathrm{d}q_\mu.$$

where the last equality follows from (i) and the $\mathcal{G}_{\mathcal{C}}$-measurability of $1_G \cdot \sigma_K(\xi)$. □

Let $\mathfrak{S}_{\mathcal{C}}^\star \subseteq \mathfrak{S}_{\mathcal{C}}$ be given by $\mathfrak{S}_{\mathcal{C}}^\star = \{A^\star : A \in \mathcal{G}_{\mathcal{C}}\}$. Given that $\mathcal{G}_{\mathcal{C}}$ is a σ-algebra, it is easy to check that $\mathfrak{S}_\mu$ is a σ-algebra as well. Indeed, $\emptyset^\star = \emptyset$, and $\Omega^\star = \mathcal{C}$ belong to $\mathfrak{S}_{\mathcal{C}}^\star$. Moreover if $A \in \mathcal{G}_{\mathcal{C}}$ then $(A^\star)^c = (A^c)^\star \in \mathfrak{S}_{\mathcal{C}}^\star$. Now let (A_i) be a sequence of pairwise disjoint events in $\mathcal{G}_{\mathcal{C}}$ and let $A = \bigcup_i A_i$. If $C \in A_i^\star$ for some i, then $C \in A^\star$. In the other direction, let $C \in A^\star$. If $C \in \mathcal{C} - A_i^\star = (A_i)_\star$ for every i, then, since each $p \in C$ is σ-additive and satisfies $p(A_i) = 0$ for every i, then $C \in A_\star$, a contradiction. Hence $A^\star = \bigcup_i A_i^\star \in \mathfrak{S}_{\mathcal{C}}^\star$.

Lemma 35. *The σ-algebras $\mathfrak{S}_{\mathcal{C}}^\star$ and $\mathfrak{S}_{\mathcal{C}}$ coincide. Moreover, μ is nonatomic if and only if q_μ is nonatomic.*

Proof. By definition $\mathfrak{S}^\star_{\mathcal{C}} \subseteq \mathfrak{S}_{\mathcal{C}}$. For every $\xi \in B(\mathcal{F}, \mathbb{R})$ and $t \in \mathbb{R}$ we have

$$\{C \in \mathcal{C} : \sigma_C(\xi) \le t\} = \left\{C \in \mathcal{C} : \min_{p \in C}(\{\omega \in \Omega : \sigma_{K(\omega)}(\xi) \le t\}) = 1\right\}.$$

Since K is $\mathcal{G}_{\mathcal{C}}$-measurable, we obtain $\{C \in \mathcal{C} : \sigma_C(\xi) \le t\} \in \mathfrak{S}^\star_{\mathcal{C}}$. Because the sets of the form $\{C \in \mathcal{C} : \sigma_C(\xi) \le t\}$ generate $\mathfrak{S}_{\mathcal{C}}$, we conclude that $\mathfrak{S}^\star_{\mathcal{C}} = \mathfrak{S}_{\mathcal{C}}$.

If μ is nonatomic, given $G \in \mathcal{G}_{\mathcal{C}}$ and $\alpha \in [0,1]$ it follows from $\mathfrak{S}^\star_{\mathcal{C}} = \mathfrak{S}_{\mathcal{C}}$ that we can find $B \in \mathcal{G}_{\mathcal{C}}$ such that $B^\star \subseteq A^\star$ and $\mu(B^\star) = \alpha\mu(A^\star)$. Because $B^\star \cap A^\star = (B \cap A)^\star$ then $q_\mu(B \cap A) = \alpha q_\mu(A)$. So, q_μ is nonatomic. The proof that if q_μ is nonatomic then so is μ follows from an analogous argument. □

Let $\mathcal{G}_\mu = \{A \in \mathcal{F} : \mu(A^\star) + \mu((A^c)^\star) = 1\}$. Being $A^\star$ and $(A^c)^\star$ disjoint, if $A \in \mathcal{G}_\mu$ then $\mu(A^\star), \mu((A^c)^\star) \in \{0,1\}$. We can think of elements of $\mathcal{G}_\mu$ as *weak zero-one events*. In a way similar to that in Lemma 32 it can be shown that $\mathcal{G}_\mu$ is a σ-algebra.

Lemma 36. *The σ-algebra $\mathcal{G}_{\mathcal{C}}$ and $\mathcal{G}_\mu$ are Π_μ-equivalent.*

Proof. By definition $\mathcal{G}_{\mathcal{C}} \subseteq \mathcal{G}_\mu$. For $A \in \mathcal{G}_\mu$ take $B = \{\omega : p(A) = 1 \text{ for all } p \in K(\omega)\}$. Since K is $(\mathcal{G}_{\mathcal{C}} \backslash \mathfrak{S}_C)$-measurable then $B \in \mathcal{G}_{\mathcal{C}}$. Moreover, for all $C \in \mathcal{C}$ such $\min_{p \in C} p(A) = 1$ or $\max_{p \in C} p(A) = 0$, we have

$$\begin{aligned}
p(A \triangle B) &= p\left(\left\{\omega \in A : \min_{p \in K(\omega)} p(A) < 1\right\}\right) + p\left(\left\{\omega \notin A : \min_{p \in K(\omega)} p(A) = 1\right\}\right) \\
&= p\left(\left\{\omega \in A : \min_{p \in C} p(A) < 1\right\}\right) + p\left(\left\{\omega \notin A : \min_{p \in C} p(A) = 1\right\}\right) = 0.
\end{aligned}$$

Thus $\pi(A \triangle B) \le \int_{\mathcal{C}} \max_{p \in C} p(A \triangle B)\, \mathrm{d}\mu(C) = 0$ for all $\pi \in \Pi_\mu$. □

H.2 Proof of Theorem 3

The result follows from Theorem 10 toghether with the next two Lemmas, which generalize Lemmas 30 and 31.

Lemma 37. *If $\succsim$ admits a multiple predictive representation $(u, \phi, \mathcal{G}, \Pi)$, then it admits a set-identifiable representation $(u, \phi, \mathcal{C}, \mu)$ where $\Pi_\mu = \Pi$ and $\mathcal{G}_{\mathcal{C}}$ is Π-equivalent to $\mathcal{G}$.*

Proof. Let $(u, \phi, \mathcal{G}, \Pi)$ be a multiple predictive representation for $\succsim$. Denote by q the common restrictions of all $\pi \in \Pi$ on $\mathcal{G}$. By Lemma 12 the set Π, being q-rectangular, admits a regular conditional probability $K\colon \Omega \to \mathscr{C}$ such that for q-almost all ω

$$\min_{p \in K(\omega)} p(\{\omega' : K(\omega') = K(\omega)\}) = 1. \tag{26}$$

In addition, for all $\xi \in B(\mathcal{F}, \mathbb{R})$

$$\sigma_K(\xi) \in \operatorname*{ess\,inf}_{\pi \in \Pi} E_\pi[\xi|\mathcal{G}]. \tag{27}$$

Let $\Omega_0 \in \mathcal{G}$ be the set of states that satisfy (26). Let $K_0 \colon \Omega_0 \to \mathcal{C}$ be the restriction of K to Ω_0. Take $\mathcal{C} = \{K_0(\omega) : \omega_0 \in \Omega_0\}$. By contruction $\mathcal{C}$ is a collection of identified sets. Let q_0 be restriction of q to $\mathcal{G} \cap \Omega_0 = \{G \cap \Omega_0 : G \in \mathcal{G}\}$. By (26) the measure q_0 has total mass one, hence it is a probability measure. Let $\mu : \mathfrak{S}_{\mathcal{C}} \to [0,1]$ be the pushforward of q_0 by K_0. By (27), a change of variables shows that for every act f

$$E_q\Big[\phi\Big(\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[u(f)|\mathcal{G}]\Big)\Big] = \int_{\mathcal{C}} \phi\Big(\min_{p\in C}\int_\Omega u(f)\,\mathrm{d}p\Big)\,\mathrm{d}\mu(C).$$

By a similar reasoning, for every $\xi \in B(\mathcal{F},\mathbb{R})$

$$E_q\Big[\operatorname*{ess\,inf}_{\pi\in\Pi} E_\pi[\xi|\mathcal{G}]\Big] = \int_{\mathcal{C}} \left(\min_{p\in C}\int_\Omega \xi\,\mathrm{d}p\right)\,\mathrm{d}\mu(C).$$

Because Π is q-rectangular, we deduce from Lemma 9 that $\Pi = \Pi_\mu$.

The σ-algebras $\mathcal{G}$ and $\mathcal{G}_{\mathcal{C}}$ are Π-equivalent. To see this, by Lemma 36 it is enough to verify that $\mathcal{G}$ and $\mathcal{G}_\mu$ are Π-equivalent. If $A \in \mathcal{G}$, then by (27) for q-almost all ω

$$\min_{p\in K(\omega)} p(A) = \max_{p\in K(\omega)} p(A) = 1_A(\omega).$$

By a change of variable this implies that $A \in \mathcal{G}_\mu$. Conversely, if $A \in \mathcal{G}_\mu$, then by a change of variables for q-almost all ω

$$\min_{p\in K(\omega)} p(A) = \max_{p\in K(\omega)} p(A) \in \{0,1\}.$$

Take $B = \{\omega : \min_{p\in K(\omega)} p(A) = 1\} \in \mathcal{G}_{\mathcal{C}}$ and $B' = \{\omega : \max_{p\in K(\omega)} p(A) = 0\} \in \mathcal{G}_{\mathcal{C}}$. We have $q(B \cup B') = 1$ and therefore by (26) for q-almost all ω

$$\begin{aligned}\max_{p\in K(\omega)} p(A\triangle B) &= \max_{p\in K(\omega)} p\left(\left\{\omega' \in A : \sigma_{K(\omega'))}(1_A) < 1\right\}\right) + p\left(\left\{\omega' \notin A : \sigma_{K(\omega'))}(1_A) = 1\right\}\right)\\ &= \max_{p\in K(\omega)} p(A)1_{B'}(\omega) + p(A^c)1_B(\omega) = 0.\end{aligned}$$

Being K a regular conditional probability of Π given $\mathcal{G}_{\mathcal{C}}$, we deduce that $\pi(A\triangle B) = 0$ for all $\pi \in \Pi$. Overall, $\mathcal{G}$ and $\mathcal{G}_\mu$ are Π-equivalent.

It reamins to show that μ is nonatomic. The probability measure q on $\mathcal{G}$ is nonatomic. Since $\mathcal{G}$ and $\mathcal{G}_{\mathcal{C}}$ are Π-equivalent and $\Pi = \Pi_\mu$, the probability measures q_μ on $\mathcal{G}_{\mathcal{C}}$ is nonatomic as well. It follows from Lemma 35 that μ is nonatomic. □

Lemma 38. *If $\succsim$ admits a set-identifiable representation $(u,\phi,\mathcal{C},\mu)$, then it admits a multiple predictive representation $(u,\phi,\mathcal{G}_{\mathcal{C}},\Pi_\mu)$.*

Proof. By Lemma 34 the set Π_μ is q_μ-rectangular, and by Lemma 35 the probability measure q_μ is nonatomic. It remains to show that for all $\xi \in B_b(\mathcal{F}, u(X))$

$$\int_{\mathcal{C}} \phi\left(\sigma_C(\xi)\right)\,\mathrm{d}\mu(C) = \int_\Omega \phi\left(\operatorname*{ess\,inf}_{\pi\in\Pi_\mu} E_\pi[\xi|\mathcal{G}_{\mathcal{C}}]\right)\,\mathrm{d}q_\mu.$$

Assume first that ξ is $\mathcal{G}_{\mathcal{C}}$-measurable. Let $\mathcal{P} \subseteq \Delta$ be a statistical model and $K : \Omega \to \mathscr{C}$ a measurable function that set-identifies $\mathcal{P}$ such that $\mathcal{C} = \{K(\omega) : \omega \in \Omega\}$. Fix $C \in \mathcal{C}$ and let q_C be the common restriction of all $p \in C$ to $\mathcal{G}_C$. The set C admits K as a regular conditional probability with respect to $\mathcal{G}_{\mathcal{C}}$. Thus, because $\sigma_K(\xi) \in \operatorname{ess\,inf}_{p \in C} E_p[\xi|\mathcal{G}]$ (Lemma 12), for q_C-almost all ω, $\sigma_{K(\omega)}(\xi) = \xi(\omega)$. Then, from the condition of set-identifiability,

$$1 = q_C(\{\omega : \sigma_{K(\omega)}(\xi) = \sigma_C(\xi)\}) = q_C(\{\omega : \xi(\omega) = E_{q_C}[\xi]\}).$$

We deduce that $\phi(\sigma_C(\xi)) = \phi(E_{q_C}[\xi]) = E_{q_C}[\phi(\xi)] = \sigma_C(\phi(\xi))$. We conclude that $\int_{\mathcal{C}} \phi(\sigma_C(\xi))\, \mathrm{d}\mu(C) = \int_\Omega \phi(\xi)\, \mathrm{d}q_\mu$. Consider now a general $\mathcal{F}$-measurable ξ. We have

$$\int_{\mathcal{C}} \phi(\sigma_C(\xi))\, \mathrm{d}\mu(C) = \int_{\mathcal{C}} \phi\left(\int_\Omega \sigma_{K(\omega)}(\xi)\, \mathrm{d}q_C(\omega)\right) \mathrm{d}\mu(C).$$

The function $\omega \mapsto \sigma_{K(\omega)}(\xi)$ is $\mathcal{G}_{\mathcal{C}}$-measurable, and therefore

$$\int_{\mathcal{C}} \phi\left(\int_\Omega \sigma_{K(\omega)}(\xi)\, \mathrm{d}q_C(\omega)\right) \mathrm{d}\mu(C) = \int_\Omega \phi(\sigma_{K(\omega)}(\xi))\, \mathrm{d}q_\mu(\omega).$$

The right-hand side is equal to $\int_\Omega \phi\left(\operatorname{ess\,inf}_{\pi \in \Pi_\mu} E_\pi[\xi|\mathcal{G}_{\mathcal{C}}]\right) \mathrm{d}q_\mu$, being K a regular conditional probability of Π_μ with respect to $\mathcal{G}_C$ (Lemma 34). □

H.3 Proof of Theorem 4

For completeness, we prove the following stronger result:

Lemma 39. *Two set-identifiable representations $(u_1, \phi_1, \mathcal{C}_1, \mu_1)$ and $(u_2, \phi_2, \mathcal{C}_2, \mu_2)$ of the same preference $\succsim$ are related by the following conditions:*

(i). *There are $a, c > 0$ and $b, d \in \mathbb{R}$ such that $u_2(x) = au_1(x) + b$ and $\phi_2(t) = c\phi_1(\frac{t-b}{a}) + d$ for all $x \in X$ and $t \in u_2(X)$.*

(ii). *$\Pi_{\mu_1} = \Pi_{\mu_2}$ and, provided that ϕ_1 is not affine, $\mu_1(\mathcal{S} \cap \mathcal{C}_1) = \mu_2(\mathcal{S} \cap \mathcal{C}_2)$ for all $\mathcal{S} \in \mathfrak{S}$.*

Proof. The preference $\succsim$ admits multiple predictive representations $(u_1, \phi_1, \mathcal{G}_{\mathcal{C}_1}, \Pi_{\mu_1})$ and $(u_2, \phi_2, \mathcal{G}_{\mathcal{C}_2}, \Pi_{\mu_2})$ by Lemma 38. By Theorem 11, the utility function u_2 is a positive affine transformation of u_1; normalizing the utility indexes, ϕ_2 is a positive affine transformation of ϕ_1; $\Pi_{\mu_1} = \Pi_{\mu_2}$; if ϕ_1 is not affine, then $\mathcal{G}_{\mathcal{C}_1}$ and $\mathcal{G}_{\mathcal{C}_2}$ are Π_{μ_1}-equivalent.

It remains to show that, if ϕ_1 is not affine, then $\mu_1(\mathcal{S} \cap \mathcal{C}_1) = \mu_2(\mathcal{S} \cap \mathcal{C}_2)$ for all $\mathcal{S} \in \mathfrak{S}$. It is enough to verify the result for $S = \{C \in \mathscr{C} : \sigma_C(\xi) \leq t\}$ for some $\xi \in B(\mathcal{F}, \mathbb{R})$ and $t \in \mathbb{R}$. For every $i \in \{1, 2\}$, let $K_i \colon \Omega \to \mathcal{C}_i$ witness the identifiability of $\mathcal{C}_i$. It follows from Lemma 34 that K_i is a regular conditional probability of Π_{μ_i} given $\mathcal{G}_{\mathcal{C}_i}$. Thus, if ϕ_1 is not affine, then K_1 and K_2 are equal π-almost surely for all $\pi \in \Pi_{\mu_1}$, being that $\mathcal{G}_{\mathcal{C}_1}$ and $\mathcal{G}_{\mathcal{C}_2}$ are Π_{μ_1}-equivalent. For all $\xi \in B(\mathcal{F}, \mathbb{R})$ and $t \in \mathbb{R}$

$$\begin{aligned}\mu_i\left(\{C_i \in \mathcal{C}_i : \sigma_{C_i}(\xi) \leq t\}\right) &= \int_{\mathcal{C}_i} \min_{p \in C_i} p\left(\{\omega : \sigma_{C_i}(\xi) \leq t\}\right) \mathrm{d}\mu_i(C_i) \\ &= \int_{\mathcal{C}_i} \min_{p \in C_i} p\left(\left\{\omega : \sigma_{K_i(\omega)}(\xi) \leq t\right\}\right) \mathrm{d}\mu_i(C_i) \\ &= \min_{\pi \in \Pi_{\mu_i}} \pi\left(\left\{\omega : \sigma_{K_i(\omega)}(\xi) \leq t\right\}\right).\end{aligned}$$

Since $\Pi_{\mu_1} = \Pi_{\mu_2}$, it follows that $\mu_1(\mathcal{S} \cap \mathcal{C}_1) = \mu_2(\mathcal{S} \cap \mathcal{C}_2)$ as desired. □

References

J. Aczél. Extension of a generalized Pexider equation. *Proceedings of the American Mathematical Society*, 133:3227–3233, 2005.

N. I. Al-Najjar and L. De Castro. Parametric representation of preferences. *Journal of Economic Theory*, 150:642–667, 2014.

C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis.* Springer, third edition, 2006.

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

D. Bergemann and S. Morris. Robust mechanism design. *Econometrica*, 73:1771–1813, 2005.

T. F. Bewley. Knightian decision theory. Part I. *Decisions in Economics and Finance*, 25: 79–110, 2002.

L. Breiman, L. LeCam, and L. Schwartz. Consistent estimates and zero-one sets. *Annals of Mathematical Statistics*, 35:157–161, 1964.

H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, 2010.

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

S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio. Ambiguity and robust statistics. *Journal of Economic Theory*, 148:974–1049, 2013.

A. Chateauneuf, F. Maccheroni, M. Marinacci, and J.-M. Tallon. Monotone continuous multiple priors. *Economic Theory*, 26:973–982, 2005.

D. B. Davis and M.-E. Pate-Cornell. A challenge to the compound lottery axiom: A two-stage normative structure and comparison to other theories. *Theory and Decision*, 37:267–309, 1994.

J. Diestel. Uniform integrability: an introduction. *Rendiconti dell'Istituto di Matematica dell'Università di Trieste*, 23:41–80, 1991.

E. B. Dynkin. Sufficient statistics and extreme points. *Annals of Probability*, 6:705–730, 1978.

D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75: 643–669, 1961.

L. G. Epstein. A paradox for the "smooth ambiguity" model of preference. *Econometrica*, 78:2085–2099, 2010.

L. G. Epstein and M. Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113:1–31, 2003.

L. G. Epstein and K. Seo. Exchangeable capacities, parameters and incomplete theories. *Journal of Economic Theory*, 157:879–917, 2015.

L. G. Epstein, H. Kaido, and K. Seo. Robust confidence regions for incomplete models. *Econometrica*, 84:1799–1838, 2016.

H. Ergin and F. Gul. A theory of subjective compound lotteries. *Journal of Economic Theory*, 144:899–929, 2009.

H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. de Gruyter, 2011.

L. Garlappi, R. Uppal, and T. Wang. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, 20:41–81, 2006.

P. Ghirardato and M. Siniscalchi. Ambiguity in the small and in the large. *Econometrica*, 80:2827–2847, 2012.

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

R. Giacomini, T. Kitagawa, and H. Uhlig. Estimation under ambiguity. 2019.

I. Gilboa. *Theory of Decision Under Uncertainty*. Cambridge University Press, 2009.

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

F. Gul and W. Pesendorfer. Expected uncertain utility theory. *Econometrica*, 82:1–39, 2014.

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

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

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

P. Klibanoff, M. Marinacci, and S. Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73:1849–1892, 2005.

P. Klibanoff, M. Marinacci, and S. Mukerji. On the smooth ambiguity model: A reply. *Econometrica*, 80:1303–1321, 2012.

P. Klibanoff, S. Mukerji, K. Seo, and L. Stanca. Foundations of ambiguity models under symmetry. *Journal of Economic Theory*, 2021. Forthcoming.

E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer, 2006.

A. Lewbel. The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature*, 7, 2019.

F. Maccheroni and M. Marinacci. A Heine-Borel theorem for $ba(\Sigma)$. *Rivista Internazionale di Scienze Economiche e Commerciali*, 48:353–362, 2001.

C. F. Manski. Anatomy of the selection problem. *Journal of Human resources*, 24:343–360, 1989.

C. F. Manski. Minimax-regret treatment choice with missing outcome data. *Journal of Econometrics*, 139:105–115, 2007.

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

J. Marschak et al. Personal probabilities of probabilities. *Theory and Decision*, 6:121–153, 1975.

F. Molinari. Econometrics with partial identification. Invited chapter for the *Handbook of Econometrics*, 2019.

R. F. Nau. Uncertainty aversion with second-order utilities and probabilities. *Management Science*, 52:136–145, 2006.

W. S. Neilson. A simplified axiomatic approach to ambiguity aversion. *Journal of Risk and Uncertainty*, 41:113–124, 2010.

L. J. Savage. *The Foundations of Statistics*. John Wiley & Sons, 1954.

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

U. Segal. The Ellsberg paradox and risk aversion: An anticipated utility approach. *International Economic Review*, 28:175–202, 1987.

K. Seo. Ambiguity and second-order belief. *Econometrica*, 77:1575–1605, 2009.

K. Song. Point decisions for interval–identified parameters. *Econometric Theory*, 30: 334–356, 2014.

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

E. Tamer. Partial identification in econometrics. *Annual Review of Economics*, 2:167–195, 2010.

A. Wald. *Statistical Decision Functions*. Wiley, 1950.