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

Tommaso Denti† Luciano Pomatto‡

September 24, 2021

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 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 the editor and the referees for their comments and suggestions. We also thank Simone Cerreia-Vioglio, Federico Echenique, Massimo Marinacci, Francesca Molinari, Kota Saito, Kyoungwon Seo, Jörg Stoye, and, especially, Fabio Maccheroni. 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.

Beyond analogies with statistical modeling, identifiable smooth preferences formalize the common view that ambiguity is due to lack of information. Our initial finding is that they admit the alternative representation

$$V(f) = E_\pi\Big[\phi\Big(E_\pi[u(f)|\mathcal{T}]\Big)\Big] \tag{2}$$

where $\mathcal{T}$ is a sub-σ algebra of events and π is a probability measure over states. In this representation, the decision maker expresses a single predictive assessment π, but she is not confident about it. The sub-σ algebra $\mathcal{T}$ represents the additional information that would make her sure about her predictions. Both $\mathcal{T}$ and π are purely subjective and make no reference to any agreed-upon statistical notion of "true" law of nature.

With different methods, terminologies, and motivations, a number of recent papers have made important progress in providing foundations for identifiable smooth preferences. An important 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 representations (1) and (2) are uniquely determined from preferences: in particular, the prior μ and the domain $\mathcal{P}$ are unique, up to null events; the σ-algebra $\mathcal{T}$ and the predictive π are also unique, up to null events.

In arriving at these results, the key step is to determine the exact behavioral counterpart of the missing information generating ambiguity—in the smooth representation, it is the σ-algebra generated by the identifying kernel k; in the predictive representation, it is the sub σ-algebra $\mathcal{T}$. We show that these abstract objects are in fact equal to the collection of events that satisfy the sure-thing principle, a property 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 $\succsim_{\text{stp}}$ defined by $f \succsim_{\text{stp}} g$ 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 $f \succsim_{\text{stp}} g$.

The literature proposes 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. In their analysis, the decision maker's preferences over second-order acts are assumed to have a subjective expected utility representation. Epstein (2010) highlights a problematic aspect of this assumption: intuitively, ambiguity averse behavior should be more pronounced in the case of second-order acts rather than first-order acts. Our analysis bypasses these complications: what is ambiguous—and what is not—is completely determined by the decision maker's preferences.

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 $\mathcal{S}$ 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 prizes. We also assume that $(\Omega, \mathcal{S})$ is a standard Borel measurable space, i.e. there exists a Polish topology on Ω such that $\mathcal{S}$ 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

[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).

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 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 $f, g, h \in \mathfrak{F}$. Similarly, two σ-algebras $\mathcal{T}_1, \mathcal{T}_2 \subseteq \mathcal{S}$ are *equivalent up to null events* if for every $A \in \mathcal{T}_1$ there is a $B \in \mathcal{T}_2$ such that $A \triangle B$ is null, and for every $B \in \mathcal{T}_2$ there is a $A \in \mathcal{T}_1$ such that $A \triangle B$ is null.

We denote by $\Delta(\mathcal{S})$, or simply Δ, the space of countably additive probability measures on $(\Omega, \mathcal{S})$. Given $p \in \Delta$, the symbol E_p denotes the corresponding expectation operator.

We endow Δ with the weak* topology and the corresponding Borel σ-algebra.[2] This is the σ-algebra Σ generated by the functions $p \mapsto p(A)$ for $A \in \mathcal{S}$. Given a nonempty set $\mathcal{P} \subseteq \Delta$, let $\Sigma_{\mathcal{P}} = \{B \cap \mathcal{P} : B \in \Sigma\}$ 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 by $\pi_\mu(A) = \int_{\mathcal{P}} p(A) \, \mathrm{d}\mu(p)$.

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

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

[2] In the weak* topology, a net (p_α) in Δ converges to p if and only if $p_\alpha(A) \to p(A)$ for all $A \in \mathcal{S}$.

[3] Klibanoff, Marinacci, and Mukerji (2005) consider 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.

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 attitudes, 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, or *kernel*, $k \colon \Omega \to \mathcal{P}$, 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.

Our notion of identifiability agrees with mainstream econometric usage in most standard settings; see Section H of the online appendix for a formal comparison. The identifiable smooth representation was introduced by Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013) using a different terminology; for further discussion, see Section I of the online appendix.

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).

Definition 2, however, is not tied to the interpretation of probability models as empirical frequencies, nor it is limited to environments characterized by repetitions. Next is the classic Ellsberg thought experiment, in the formulation of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013, p. 980).

Example 2. 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 $\{r, b, y\} \times \{\gamma\}$.

Ambiguity is generated by uncertainty about the composition of the urn. 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. In the last example, the set $\mathcal{P}$ has a simple structure, and the kernel k is given. In more complex decision problems, there may be no obvious choice of a statistical model or identifying kernel. We illustrate this point in the context of policy making under uncertainty:

Example 3. A new virus is discovered and a policy maker is pondering whether or not it will develop into an epidemic. The spread of the virus will depend on a number of uncertain factors, such as the virus reproduction number, its infectious period, and its mode of transmission. Epidemiological models aggregate these factors in different ways, corresponding to different statistical models that the policy maker may adopt to guide their decision. Given the novelty of the virus, there is substantial disagreement among epidemiologists, and the adoption of a particular statistical model is a judgement call.[4]

To represent this problem, we define a state as a tuple $\omega = (e, x_1, \ldots, x_n)$, where $e \in \{0, 1\}$ describes whether or not an epidemic will occur, and each $x_i \in \mathbb{R}$ is a potential factor affecting the spread of the virus. An epidemiological model describes a

[4]See, e.g., Best and Boice (2021) for COVID-19.

set of relevant factors $I \subseteq \{1, \dots, n\}$ and the probability of an epidemic $\varphi(x_I) \in [0, 1]$ as a function of $x_I = (x_i)_{i \in I}$. We can then define a corresponding kernel k, where $k(e, x_1, \dots, x_n)$ is the measure in $\Delta(\Omega)$ that assigns probability $\varphi(x_I)$ to the state $(1, x_1, \dots, x_n)$ and the remaining probability to the state $(0, x_1, \dots, x_n)$. The σ-algebra generated by the factors in I describes the ambiguity perceived by the policy maker.

The literature on model averaging in policy evaluation provide additional examples of this nature (see, e.g., Brock, Durlauf, and West, 2007).

3.1 Predictive representation

Identifiable smooth preferences admit an alternative representation which, going beyond analogies with statistical modeling, formalizes the common view that ambiguity is due to lack of information. The alternative representation, which below we show is equivalent to that of Definition 2, is defined as follows:

Definition 3. A tuple $(u, \phi, \mathcal{T}, \pi)$ 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, $\mathcal{T} \subseteq \mathcal{S}$ a σ-algebra, and $\pi \in \Delta$ a probability measure non-atomic on $\mathcal{T}$ such that

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

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 $\mathcal{T}$ represents the additional information the agent would need in order to arrive at a reliable probability assessment. Given knowledge of $\mathcal{T}$, acts would be ranked according to their conditional expected utility $E_\pi[u(f)|\mathcal{T}]$. The idea that ambiguity aversion is a reaction to lack of information has been advanced many times in the literature, typically as an informal motivation.[5] What is new, in this representation, is that the information that generates ambiguity is explicitly part of the model and, in addition, that such information is elicited from preferences.

[5]An exception is Gajdos, Hayashi, Tallon, and Vergnaud (2008), where information is instead part of the model.

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* $(u, \phi, \mathcal{T}, \pi)$*, then it admits an identifiable representation* $(u, \phi, \mathcal{P}, \mu)$ *where* $\pi_\mu = \pi$ *and* $\mathcal{T}$ *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 also ties together the σ-algebras $\mathcal{T}$ and $\sigma(k)$.

The statistical interpretation of identifiable smooth preferences describes ambiguity as uncertainty about a true law. Proposition 1 provides a different interpretation for the same class of preferences: it is the lack of information—i.e., the σ-algebra $\mathcal{T}$—that generates ambiguity. The case where the events in $\mathcal{T}$ consists of long-run frequencies or other "objective" sources of information is a special case, as highlighted in Examples 1-3. de Finetti (1977) provides another vivid example: an individual may be "perplexed at having to assess the probability of victory (or defeat, or a draw) of a given team in a given match by stating a single number" but nevertheless be "willing to give assessments conditional upon certain contingencies (presence of star player, terrain conditions, bonus).[6]"

Particular instances of the predictive representation have already appeared in the literature:

Example 4. Let $(u, \phi, \mathcal{T}, \pi)$ be a predictive representation where $\mathcal{S} = \mathcal{T}$. Then

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

[6]De Finetti describes the underlying thought process as follows: "Given a situation well-defined as to the outside aspects which I regard as more significant, I, on the basis of my own more or less great knowledge of football and of the calibre of the two teams, am willing to express my opinion by assessing the probabilities of the three possible results. However, as for other contingencies, I feel incompetent and I would not venture an opinion. As for the prospects of recovery, or suspension, of that player I lack any grounds for an opinion; likewise, I feel unable to venture weather forecasts for the day of the match or the immediately preceding ones, lacking any knowledge of meteorology; and I have no crystal ball to fore see whether the rumours about the promised bonus will be confirmed. Therefore I am willing to assess probabilities $Prob(A|H_i)$ conditional on the various hypotheses H_i, but as for the $Prob(H_i)$, and hence $Prob(A)$, I cannot commit myself."

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

Example 5. Two sources of uncertainty a and b are represented by probability spaces $(\Omega_a, \mathcal{S}_a, \pi_a)$ and $(\Omega_b, \mathcal{S}_b, \pi_b)$. A state of the world $\omega = (\omega_a, \omega_b)$ specifies a realization for each source, and $\mathcal{S} = \mathcal{S}_a \times \mathcal{S}_b$ 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

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

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

The predictive representation allows to generalize source-dependent preferences along two natural directions: by dispensing with the assumption of stochastic independence, and by making the classification between different sources of uncertainty subjective. We illustrate these features in a simple optimal portfolio problem.

Example 6. An investor must choose how to allocate their wealth w across n assets. Let $r_i(\omega) \in \mathbb{R}$ be the gross return of asset $i = 1, \ldots, n$ in state $\omega \in \Omega$. After normalizing w to 1, a portfolio $x \in \mathbb{R}^n$ leads to the monetary payoff $(x \cdot r)(\omega) = \sum_{i=1}^n x_i r_i(\omega)$.

We consider an investor who is endowed with a belief $\pi \in \Delta(\Omega)$ and who sees the returns of the assets in a subset $I \subseteq \{1, \ldots, n\}$ as ambiguous. In the corresponding predictive representation, a portfolio x is evaluated according to

$$E_\pi\Big[\phi\Big(E_\pi[u(x \cdot r)|\mathcal{T}]\Big)\Big],$$

where $\mathcal{T}$ is the σ-algebra generated by ambiguous returns $(r_i)_{i \in I}$. Thus, conditional on the returns of the ambiguous assets, the portfolio is ranked according to its expected utility $E_\pi[u(x \cdot r)|\mathcal{T}]$. In line with the standard interpretation of smooth preferences, the function ϕ describes the investor's attitude towards uncertainty about the returns of the ambiguous assets. The representation of Ergin and Gul (2009) can be obtained as the special case where the set I is exogenous, rather than a feature of the preference, and where the two vectors of returns $(r_i)_{i \in I}$ and $(r_i)_{i \notin I}$ are stochastically independent of each other.

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 postulate P2. We denote by $\mathcal{S}_{\text{stp}}$ the family of all such events. The properties of $\mathcal{S}_{\text{stp}}$ 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 $A \in \mathcal{S}_{\text{stp}}$, 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 $A \in \mathcal{S}_{\mathrm{stp}}$, 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 only 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 $A, B \in \mathcal{S}_{\mathrm{stp}}$ 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 $\mathcal{S}_{\mathrm{stp}}$-measurable, if $f \succ g$ then there is a partition $A_1, \ldots, A_n$ of events in $\mathcal{S}_{\mathrm{stp}}$ such that $hA_if \succ g$ and $f \succ hA_ig$ for all i.*

We now discuss more in detail the interpretation of Axiom 4, our main assumption. The axiom ties hedging opportunities to the collection of events that satisfy the sure-thing principle. To better see the behavioral content of this assumption, consider a decision maker who is indifferent between two acts f and g, but who exhibits a preference reversal $f \succsim_A g$ and $f \prec_{A^c} g$, where A is an event that satisfies the sure-thing principle. Because the conditional preferences $\succsim_A$ and $\prec_{A^c}$ are well defined, the decision maker expresses an unequivocal preference for f when A happens, and for g when A^c happens.

Under Axiom 4, uncertainty about whether A will or will not occur generates a scope for hedging. Intuitively, the act $\frac{1}{2}f + \frac{1}{2}g$, being midway between f and g, is less exposed to uncertainty about the event A. This may result in the ranking $\frac{1}{2}f + \frac{1}{2}g \succ g$, a violation of independence, being that $f \sim g$.

The axiom can also be interpreted as capturing an elementary form of statistical reasoning, even without reference to probabilities or other representations of uncertainty. A common strategy for simplifying complex decision problems consists in first isolating a set of relevant hypotheses, and then drawing conclusions by evaluating the available options conditional on each hypothesis. For example, the quality of a test is evaluated by considering the probability of error conditional on each possible hypothesis (e.g., type I and type II error for binary hypotheses). This type of reasoning hinges on making statements conditional on an hypothesis being true or false. Since the events that satisfy the sure-thing principle are precisely those events A for which the decision maker's preferences conditional on A and A^c are well-defined, it is suggestive to interpret the collection $\mathcal{S}_{\text{stp}}$ as the set of hypotheses the decision maker has in mind. Under this interpretation, Axiom 4 states that uncertainty about the correct hypothesis $A \in \mathcal{S}_{\text{stp}}$ is what generates ambiguity: if contingent on every hypothesis the act f is preferred to g, then, according to the axiom, the ranking between the two acts is unambiguous.

5 Representation theorem and uniqueness

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 identifiable smooth representation. By Proposition 1 the preference $\succsim$ admits an identifiable representation $(u, \phi, \mu, \mathcal{P})$ if and only if it admits a predictive representation $(u, \phi, \mathcal{T}, \pi)$. Thus Axioms 1-6 characterize both representations. Next we describe their uniqueness properties.

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

6 Model and predictive 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 $\succsim_{\text{stp}}$ over acts by

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

In words, $f \succsim_{\text{stp}} g$ if f is preferred to g conditional on each event that satisfies the sure-thing principle. Following the discussion in Section 4, this reflects the idea that f is preferred to g conditional on each hypothesis $A \in \mathcal{S}_{\text{stp}}$ entertained by the decision maker about the states of the world. The next result is a representation theorem for the relation $\succsim_{\text{stp}}$.

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

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

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

The preference relation $\succsim_{\text{stp}}$ 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 $f \succsim_{\text{stp}} g$ 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 $f \succsim_{\text{stp}} g$ holds when the information $\mathcal{T}$ does not affect the ranking of the two acts.[7]

The next proposition shows that $\mathcal{S}_{\text{stp}}$ 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 $\mathcal{S}_{\text{stp}}$ is a σ-algebra.[8]

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

For an identifiable smooth 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. The proposition shows that $\mathcal{S}_{\text{stp}}$ stands for the behavioral counterpart of this information. An analogous result holds for the predictive representation where $\mathcal{S}_{\text{stp}}$ and $\mathcal{T}$ are equivalent up to null events.

As is well known, the unambiguous preference relation $\succsim^*$ admits the representation

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

where u is an affine utility function, and C^* is a set of probabilities over $(\Omega, \mathcal{S})$.[9] When the set C^* is not a singleton, the agent formulates a range of probabilistic assessments π under which to evaluate acts according to expected utility. We call *predictive uncertainty* the indeterminacy described by the multiplicity of probabilities in C^*.

We can rephrase Axiom 4 as follows:

$$f \succsim_{\text{stp}} g \quad \Longrightarrow \quad f \succsim^* g.$$

[7]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 $f \succsim_{\text{stp}} g$ if and only if $f(\omega) \succsim g(\omega)$ for all $\omega \in \Omega$, up to a null event.

[8]See Lemma 9 in the appendix for a precise statement of this result.

[9]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.

In view of Proposition 2, the axiom illustrates the following principle: if model uncertainty does not affect the ranking of f and g, then predictive uncertainty should not either.

6.1 Perceived ambiguity and ambiguity attitude

The separation between the ambiguity perceived by the decision maker, represented by $\mathcal{P}$ and μ, and their attitude towards ambiguity, represented by ϕ, is a central feature of the smooth ambiguity representation. This separation mimics the distinction of tastes from beliefs in Subjective Expected Utility, one of the most appealing features of Savage's theory. Such clear distinction is absent in many other classes of preferences.

In this section we characterize how the identifiable smooth representation captures changes in perceived ambiguity or in ambiguity aversion. We model comparisons in ambiguity perception in the following way. Let $\succsim_1$ and $\succsim_2$ be two preferences that admit identifiable smooth representations $(u_1, \phi_1, \mathcal{P}_1, \mu_1)$ and $(u_2, \phi_2, \mathcal{P}_2, \mu_2)$, respectively. We say the relation $\succsim_1$ *perceives more ambiguity* than $\succsim_2$ if

$$f \succsim^1_{\text{stp}} g \quad \Longrightarrow \quad f \succsim^2_{\text{stp}} g.$$

According to this definition, if the first decision maker prefers f to g conditional on every event that satisfies the sure-thing principle, then the same is true for the second decision maker. As already discussed, the relation $\succsim_{\text{stp}}$ is a robust ranking that singles out those pairs of acts that are ranked conditional on every hypothesis entertained by the decision maker. Thus, the decision maker who perceives more ambiguity formulates a richer sets of hypothesis, and as a consequence, fewer acts can be ranked in a robust way.

The next proposition characterizes this comparative notion:

Proposition 4. *Let ϕ_1 and ϕ_2 be not affine. The following conditions are equivalent:*

(i). $\succsim_1$ *perceives more ambiguity than* $\succsim_2$.

(ii). u_2 *is a positive affine transformation of* u_1 *and there exists a measurable function* $M\colon \mathcal{P}_1 \to \Delta(\mathcal{P}_2)$ *such that for every event* $A \in \mathcal{S}$ *and* μ_2*-almost every* $p_2 \in \mathcal{P}_2$,

$$p_2(A) = \int_{\mathcal{P}_1} p_1(A) \,\mathrm{d}M(p_1|p_2).$$

Both conditions imply $\sigma(k_2) \subseteq \sigma(k_1)$, π_{μ_2}-almost surely.[10]

An increase in ambiguity perception translates into a contraction of the set of probability models: if agent 1 perceives more ambiguity than agent 2, then each probability law $p_2 \in \mathcal{P}_2$ can be obtained by a weighted average across laws in the set $\mathcal{P}_1$.

A simple instance of this is when $\mathcal{P}_2$ is a subset of $\mathcal{P}_1$, i.e., when the first decision maker has in mind a larger set of probability models. The proposition extends this intuition by allowing each p_2 to be a mixture of models in $\mathcal{P}_1$. For example, this allows for the possibility that for some law $p \in \mathcal{P}_2$, the first decision maker considers a richer set of distributions $p(\cdot|H_1), \ldots, p(\cdot|H_n)$, where $H_1, \ldots, H_n$ are conditional events that the first decision maker, unlike the second, deems relevant for resolving uncertainty. This naturally arises if, in the predictive representations of their preferences, the two decision makers share the same predictive probability $\pi = \pi_1 = \pi_2$, but more information is needed to resolve the ambiguity of the first decision maker, that is, $\mathcal{T}_2 \subseteq \mathcal{T}_1$. More generally, Proposition 4 also shows that an increase in the ambiguity perceived by the decision maker always translates into a larger σ-algebra $\sigma(k_i)$ of events that resolve ambiguity. In particular, up to null events, it implies a larger collection of events that satisfy the sure-thing principle.

For modeling an increase in ambiguity aversion we employ the standard notion of comparative ambiguity aversion introduced by Ghirardato and Marinacci (2002), according to which $\succsim_1$ is *more ambiguity averse* than $\succsim_2$ if

$$f \succsim_1 x \quad \Longrightarrow \quad f \succsim_2 x.$$

Intuitively, the evaluation of constant acts is unambiguous because the outcome is independent of the state. Thus, decision makers who are more ambiguity averse should choose constant acts more often.

The next proposition characterizes the agents' degree of ambiguity aversion in terms of the relative concavity of ϕ_1 and ϕ_2. As it is well known, if $\succsim_1$ is more ambiguity averse than $\succsim_2$, then the decision makers have the same risk preferences, i.e., u_1 is a positive affine transformation of u_2.[11] To simplify the exposition, for the next result we assume that $u_1 = u_2$. To state the result, let $\mathcal{S}^i_{\text{stp}}$ be the collection of

[10] That is, for every $A_2 \in \sigma(k_2)$ there exists $A_1 \in \sigma(k_1)$ such that $\pi_{\mu_2}(A_1 \triangle A_2) = 0$.
[11] See, e.g., Ghirardato, Maccheroni, and Marinacci (2004, Corollary B.3)

events that satisfy the sure-thing principle according to agent i.

Proposition 5. *Let ϕ_1 and ϕ_2 be continuously differentiable and not affine, and let $u_1 = u_2$. If $\mathcal{S}^1_{\text{stp}} = \mathcal{S}^2_{\text{stp}}$, then the following conditions are equivalent:*

(i). The preference $\succsim_1$ is more ambiguity averse than $\succsim_2$.

(ii). The function $\phi_1 \circ \phi_2^{-1}$ is concave and $\pi_{\mu_1} = \pi_{\mu_2}$.

Klibanoff, Marinacci, and Mukerji (2005) derive an analogous characterization for general smooth preferences (Klibanoff, Marinacci, and Mukerji, 2005, Theorem 2), but under the stronger hypothesis that $\mu_1 = \mu_2$, and hence that $\mathcal{P}_1 = \mathcal{P}_2$. This implies that in their analysis agents are comparable in their ambiguity attitude only if they share the same ambiguity attitude.

In contrast, focusing on identifiable smooth preferences, we show the same conclusion holds under the hypothesis that the agents agree on the events that satisfy the sure-thing principle. By Proposition 3 this is equivalent to saying that the decision makers agree on what information can resolve ambiguity. This assumption is fully behavioral and distinctive of smooth identifiable preferences. In addition, as we show in Example 7 in the online appendix, it cannot be further weakened.

7 Discussion

7.1 On the behavioral foundations of smooth ambiguity

Our analysis contributes to the recent debate on the behavioral foundations of smooth ambiguity preferences. Much of the debate has centered around the role played by second-order acts in the axiomatization of Klibanoff, Marinacci, and Mukerji (2005). These are acts of the form $f^2 \colon \mathcal{P} \to X$, whose outcome $x = f^2(p)$ depends on the true probability distribution $p \in \mathcal{P}$. Their setup takes choices over second-order acts as part of the primitives.

In his critique to their approach, Epstein (2010) presents a variation on Ellsberg's (1961) thought experiment where the decision maker is asked to bet both on the color extracted from an urn and its composition. Both variables are ambiguous, and it is intuitive that a decision maker could exhibit Ellsberg-like behavior with respect to both bets. However, whether or not smooth preferences can accommodate this behavior depends on how the problem is formalized. In Epstein's formulation, a bet on the composition of the urn is a second-order act, and hence ambiguity aversion is not possible. To the contrary, Klibanoff, Marinacci, and Mukerji (2012) argue that the state space should be richer, and should be taken to contain all information that is relevant for the decision maker, including the composition of the urn. Unlike Epstein's formulation, this approach makes a bet on the composition a first-order act.

Our axiomatization can help to clarify some aspects of this debate. Epstein's critique starts from the observation that smooth preferences are ambiguity neutral over second-order acts. This may seem counterintuitive: how can an agent who finds it difficult to assess the likelihood of an event have no difficulties in assessing the probability of a second-order event?

In our analysis, which involves only Anscombe-Aumann acts, the object that is closest to a second-order act is an act that is measurable with respect to the σ-algebra $\mathcal{S}_{\text{stp}}$. Over such acts, the decision maker is ambiguity neutral in the sense of being consistent with Savage's axioms. But since it is up to the decision maker to decide which events belong to $\mathcal{S}_{\text{stp}}$, this assumption seems less extreme than the assumption of ambiguity neutrality over second-order acts that is critiqued by Epstein (2010).

One can draw an analogy with the interpretation of maxmin preferences. In a maxmin representation where the set C of probabilities is exogenously given, a common criticism of the minmax criterion is its extreme pessimism. On the contrary, Gilboa and Schmeidler (1989) show that when C is elicited from preferences, minmax preferences do not express any obvious form of extreme pessimism. In a similar way, once we allow for $\mathcal{P}$ to be endogeneous, our axiomatization shows that smooth-preferences do not presume any obvious or extreme form of ambiguity neutrality.

7.2 Scope of identifiable smooth preferences

The literature has introduced several alternative models for decision making under ambiguity. A natural question, then, is what criteria can we use to select among models? The question is particularly relevant for the smooth model, given the difficulties in its interpretation and problematic foundations. Our analysis contributes to this question by providing an axiomatic foundation for the (identifiable) smooth model that is cast within the same Anscombe-Aumann framework shared by the other classic axiomatizations of Choquet and Maxmin expected utility, and other models.

Our axiomatic analysis helps understanding some potential limitations of (identifiable) smooth preferences, and the contexts where they can be more appropriate.

As highlighted among others by Gilboa and Marinacci (2013), "smooth preferences have the disadvantage of imposing non-trivial epistemological demands on the decision maker" since "the smooth model requires the specification of a prior over probability models." Our main axiom translates this observation in the concrete realm of preferences over acts. To satisfy Axiom 4, the decision maker first needs to identify a number of contingencies that would resolve ambiguity (i.e. the sure-thing events), then check the ranking over acts contingency by contingency, and finally, in the case of a unanimous ranking, be sure not violate the Anscombe-Aumann independence axiom.

It is not difficult to imagine situations where this cognitive task can be difficult, if not impossible. For example, a policy maker asked to make a decision on an unfamiliar, and complex issue, such as climate change, may have simply no idea of what facts could be useful for understanding the phenomenon. In contexts such these, other models, such as max-min or Choquet expected utility, may be less cognitively demanding.

An interesting question concerns identifiability for general ambiguity preferences; we discuss some of the issues involved in Section J of the online appendix.

Appendix

The appendix is organized as follows: In Section A we set up 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 1-6 admits a predictive representation, as defined in the main text. We study the uniqueness properties of the predictive representation as well as provide characterizations for $\mathcal{S}_{\text{stp}}$ and $\succsim_{\text{stp}}$. The analysis in Section D is applied in Sections E-G to prove the results stated in the main text. Proposition 1 follows from standard arguments on regular conditional probabilities; for completeness, a formal proof is in the online appendix.

A Preliminaries

Notation. For every σ-algebra $\mathcal{T} \subseteq \mathcal{S}$ and nonempty interval $U \subseteq \mathbb{R}$, we denote by $B(\mathcal{T}, U)$ the space of $\mathcal{T}$-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 $B_0(\mathcal{T}, U) \subseteq B(\mathcal{T}, U)$ the subspace of functions taking finitely many values, and let $B_b(\mathcal{T}, U) \subseteq B(\mathcal{T}, U)$ be the set of all $\xi \in B(\mathcal{T}, U)$ for which there exist $a, b \in U$ that satisfy $a \geq \xi \geq b$. A sequence (ξ_n) in $B_b(\mathcal{T}, U)$ is *bounded* if there are $a, b \in U$ such that $a \geq \xi_n \geq b$ for all n.

Let $q \in \Delta(\mathcal{T})$ be a probability measure. We denote by $L_\infty(\mathcal{T}, q)$ the space of equivalence classes of real-valued, $\mathcal{T}$-measurable, and almost-surely bounded functions. Given $\xi \in B_b(\mathcal{T})$ we denote by $[\xi] \in L_\infty(\mathcal{T}, q)$ the corresponding equivalence class. We refer to an element $\zeta \in [\xi]$ of the equivalence class as a *representative* of $[\xi]$. We denote by

$$L_\infty(\mathcal{T}, q, U) = \{[\xi] : \xi \in B_b(\mathcal{T}, U)\}$$

the set of equivalence classes induced from functions in $B_b(\mathcal{T}, U)$. For an increasing function $\phi\colon U \to \mathbb{R}$ and $\xi \in B_b(\mathcal{T}, U)$, we denote by $\phi([\xi])$ the equivalence class $[\phi(\xi)] \in L_\infty(\mathcal{T}, q)$.

Given a probability measure $\pi \in \Delta$ that agrees with q on $\mathcal{T}$, we denote by

$$E_\pi[\xi|\mathcal{T}] \in L_\infty(\mathcal{T}, q, U)$$

the conditional expectation of $\xi \in B_b(\mathcal{S}, U)$ with respect to $\mathcal{T}$.

Equivalent σ-algebras Let $\pi \in \Delta(\mathcal{S})$ be a probability measure on $\mathcal{S}$. Two events $A, B \in \mathcal{S}$ are *π-equivalent* if $\pi(A \triangle B) = 0$. Two σ-algebras $\mathcal{T}_1, \mathcal{T}_2 \subseteq \mathcal{S}$ are *π-equivalent* if every $A \in \mathcal{T}_1$ has a π-equivalent $B \in \mathcal{T}_2$, and vice versa every $B \in \mathcal{T}_2$ has a π-equivalent $A \in \mathcal{T}_1$. Two functions $\xi, \zeta \in B_b(\mathcal{S}, U)$ are *π-equivalent* if they are equal π-almost surely. The next lemma describes some basic properties of equivalent σ-algebras. We omit the simple proof.

Lemma 1. *If $\mathcal{T}_1$ and $\mathcal{T}_2$ are π-equivalent, then the following conditions are satisfied:*

(i). *For every $\xi \in B_b(\mathcal{T}_1, U)$ there is a π-equivalent $\zeta \in B_b(\mathcal{T}_2, U)$.*

(ii). *For each $\xi \in B_b(\mathcal{S}, U)$, every $\zeta \in E_\pi[\xi|\mathcal{T}_1]$ has a π-equivalent $\psi \in E_\pi[\xi|\mathcal{T}_2]$.*

Pexider functional equation Let $I, J \subseteq \mathbb{R}$ be non-empty open intervals of the real line. Let $I + J = \{s + t : s \in I, t \in J\}$. Let $\phi\colon I + J \to \mathbb{R}$ be a measurable function that is not constant on every sub-interval of positive length. The following result on the Pexider equation is due to Aczél (2005, Theorem 2 and its corollary).

Lemma 2. *Suppose there are $\alpha\colon I \to \mathbb{R}$, $\beta\colon I \to \mathbb{R}$, and $\gamma : J \to \mathbb{R}$ such that*

$$\phi(s+t) = \alpha(s) + \beta(s)\gamma(t) \quad \forall s \in I, t \in J.$$

Then ϕ is either affine or exponential, that is, only two cases can arise:

(i). *There are $\alpha, \beta \in \mathbb{R}$ with $\beta \neq 0$ such that $\phi(t) = \alpha + \beta t$ for all $t \in I + J$.*

(ii). *There are $\alpha, \beta, \gamma \in \mathbb{R}$ with $\beta\gamma \neq 0$ such that $\phi(t) = \alpha + \beta e^{\gamma t}$ for all $t \in I + J$.*

B Decomposable operators

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

Definition 4. An operator $T\colon B_b(\mathcal{S}, U) \to L_\infty(\mathcal{T}, q, U)$ is:

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

The next technical results derive some basic properties of decomposable operators (proofs in the online appendix).

Lemma 3. *If T is decomposable, then for every partition $A_1, \ldots, A_n$ of Ω in events that are $\mathcal{T}$-measurable, and every $\xi_1, \ldots, \xi_n$ in $B_b(\mathcal{S}, U)$*

$$T\left(\sum_{i=1}^{n} \xi_i \cdot 1_{A_i}\right) = \sum_{i=1}^{n} T(\xi_i) \cdot [1_{A_i}] \tag{3}$$

Lemma 4. *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 $T\xi_n \to T\xi$.*

Lemma 5. *If T is monotone, decomposable, normalized, and σ-order continuous, then it is projective.*

An operator $T\colon B_b(\mathcal{S}, U) \to L_\infty(\mathcal{T}, q, U)$ is *affine* if for all $\alpha \in [0,1]$ and $\xi, \zeta \in B_b(\mathcal{S}, U)$, it satisfies $T(\alpha\xi + (1-\alpha)\zeta) = \alpha T(\xi) + (1-\alpha)T(\zeta)$.

Theorem 3. *An operator $T\colon B_b(\mathcal{S}, U) \to L_\infty(\mathcal{T}, q, U)$ is monotone, decomposable, normalized, σ-order continuous, and affine if and only if there is a probability measure $\pi \in \Delta(\mathcal{S})$ that extends q and satisfies for all $\xi \in B_b(\mathcal{S}, U)$*

$$T\xi = E_\pi[\xi|\mathcal{T}]. \tag{4}$$

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

The next lemmas follows from standard arguments; for completeness, a formal proof is in the online appendix.

Lemma 6. *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 7, then there exists a non-constant affine function $u\colon X \to \mathbb{R}$ representing $\succsim$ on X.*

Lemma 7. *For every σ-algebra $\mathcal{T} \subseteq \mathcal{S}$ and affine function $u\colon X \to \mathbb{R}$,*

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

For a preference relation $\succsim$ that satisfies Axioms 1-3 and 7, Lemma 6(i) and 6(ii) imply that for every $A \in \mathcal{S}_{\text{stp}}$ 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 8. *Assume Axioms 1-3 and 7 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 $A \in \mathcal{S}_{\text{stp}}$ is not null, then $x \succ y$ implies $x \succ_A y$.*

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

Lemma 9. *If Axioms 1-3, 5, and 7 are satisfied, then $\mathcal{S}_{\text{stp}}$ 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.

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

Theorem 4. *If Axioms 1-3, 5-6, and 7 are satisfied, then there are*

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

(ii). *a nonatomic probability measure $q \in \Delta(\mathcal{S}_{\text{stp}})$,*

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

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

$$T\colon B_b(\mathcal{S}, u(X)) \to L_\infty(\mathcal{S}_{\text{stp}}, q, u(X)),$$

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

$$\begin{aligned} f \succsim g &\iff \int_\Omega \phi(Tu(f))\,\mathrm{d}q \geq \int_\Omega \phi(Tu(g))\,\mathrm{d}q, \\ f \succsim_{\text{stp}} g &\iff Tu(f) \geq Tu(g). \end{aligned}$$

C.1 Proof of Theorem 4

The proof of the result is divided in lemmas. For the remaining of this section, we assume that Axioms 1-3, 5-6, and 7 are satisfied. By Lemma 9 the collection of events $\mathcal{S}_{\text{stp}}$ is a σ-algebra.

Lemma 10. *There exist a non-atomic probability measure $q \in \Delta(\mathcal{S}_{\text{stp}})$ and a continuous strictly increasing function $\phi\colon u(X) \to \mathbb{R}$ such that for all $\mathcal{S}_{\text{stp}}$-measurable acts f and g*

$$f \succsim g \iff \int_\Omega \phi(u(f))\,\mathrm{d}q \geq \int_\Omega \phi(u(g))\,\mathrm{d}q \tag{5}$$

Proof. First we show that (5) holds for simple acts. Let $\succsim_0$ be the restriction of $\succsim$ to the acts that are simple and $\mathcal{S}_{\text{stp}}$-measurable. Observe that $\succsim_0$ satisfy Savage's P1-P6: Axiom 1 implies P1 and P5; P2 holds by definition of $\mathcal{S}_{\text{stp}}$; P3 follows from Lemmas 6(i) and 8(iii); Axiom 5 is P4; Axiom 6 is P6.[12] In addition, $\succsim_0$ satisfies risk independence (Axiom 7), mixture continuity (Lemma 6(ii)), and monotone continuity is implied by Lemma 8(i). By Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2012, Proposition 3) there exist a non-atomic probability measure $q \in \Delta(\mathcal{S}_{\text{stp}})$ and a continuous strictly increasing function $\phi\colon u(X) \to \mathbb{R}$ such that (5) holds for all f and g that are $\mathcal{S}_{\text{stp}}$-measurable and simple.

Now we extend the result to acts that are not simple. Let f be a $\mathcal{S}_{\text{stp}}$-measurable act. From Lemma 7 we have $u(f) \in B_b(\mathcal{S}_{\text{stp}}, u(X))$. Thus we can find a sequence (ξ_n) in $B_0(\mathcal{S}_{\text{stp}}, u(X))$ that converges uniformly to $u(f)$. Applying Lemma 7 again we can find a sequence (f_n) of simple $\mathcal{S}_{\text{stp}}$-measurable acts such that $u(f_n) = \xi_n$ for all n. It follows from Lemma 8(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 (5) 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 (5) holds for all $\mathcal{S}_{\text{stp}}$-measurable acts. $\square$

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

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

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

Lemma 10 shows that V represents $\succsim$ on $\mathfrak{F}$. Moreover $V(f) = \int_\Omega \phi(u(f))\,\mathrm{d}q$ for all $\mathcal{S}_{\text{stp}}$-measurable acts f. The next lemmas establish key properties of $\succsim_{\text{stp}}$.

Lemma 11. *For all $f, g \in \mathfrak{F}$, the following conditions are satisfied.*

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

(ii) *$u(f) \geq u(g)$ implies $f \succsim_{\text{stp}} g$.*

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

Proof. (i). If $f \succsim_{\text{stp}} g$, then for all $B \in \mathcal{S}_{\text{stp}}$ we have $A \cap B \in \mathcal{S}_{\text{stp}}$ 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 $B \in \mathcal{S}_{\text{stp}}$, then $(fAh)Bh' \succsim (gAh)Bh'$ for every $h' \in \mathfrak{F}$, which implies $fAh \succsim_{\text{stp}} gAh$. The other implication is obvious. (ii). It follows from Lemma 6(i).

(iii). "If." Let $A \in \mathcal{S}_{\text{stp}}$ be the event where $g(\omega) \succ f(\omega)$. Because $q(A) = 0$, it follows from Lemma 8(iii) that A is null. Thus $fAg \sim_{\text{stp}} g$. Moreover $f \succsim_{\text{stp}} fAg$ by (ii) above. We conclude that $f \succsim_{\text{stp}} g$. "Only if." Fix $x \in X$. For all $A \in \mathcal{S}_{\text{stp}}$ 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. $\square$

Lemma 12. *For every $f \in \mathfrak{F}$ there exists a $\mathcal{S}_{\text{stp}}$-measurable act $\hat{f}$ such that $f \sim_{\text{stp}} \hat{f}$.*

Proof. Fix $x \in X$ such that $f(\omega) \succsim x$ for all ω. Let $q_f \colon \mathcal{S}_{\text{stp}} \to \mathbb{R}$ 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 6(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 $\mathcal{S}_{\text{stp}}$. 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 $\mathcal{S}_{\text{stp}}$ such that $A_n \downarrow \varnothing$. The sequence (fA_nx) is bounded and converges pointwise to x. By Lemma 8(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 8(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 $\mathcal{S}_{\text{stp}}$-measurable function $\xi\colon \Omega \to \mathbb{R}_+$ such that for all $A \in \mathcal{S}_{\text{stp}}$, $q_f(A) = \int_A \xi \,\mathrm{d}q$. Let $y \in X$ such that $y \succsim f(\omega)$ for all ω. For all $A \in \mathcal{S}_{\text{stp}}$ we have by Lemma 6(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 $\zeta = \phi^{-1}(\xi + \phi(u(x))) \in B_b(\mathcal{S}_{\text{stp}}, u(X))$. By Lemma 7 there is a $\mathcal{S}_{\text{stp}}$-measurable act $\hat{f}$ such that $u(\hat{f}) = \zeta$. For all $A \in \mathcal{S}_{\text{stp}}$

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

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

We define the operator $T\colon B_b(\mathcal{S}, u(X)f) \to L_\infty(\mathcal{S}_{\text{stp}}, q, u(X))$ by $Tu(f) = [u(\hat{f})]$, where $\hat{f}$ is a $\mathcal{S}_{\text{stp}}$-measurable act that satisfies $\hat{f} \sim_{\text{stp}} f$. By Lemmas 7, 11(iii), and 12 the operator is well defined. In addition, $f \succsim_{\text{stp}} g$ if and only if $Tu(f) \geq Tu(g)$. Moreover, since $\hat{f} \sim_{\text{stp}} f$ 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 4.

Lemma 13. *T is monotone, normalized, decomposable, and σ-order continuous.*

Proof. Monotonicity follows from Lemmas 11(ii) and 11(iii). Normalization is obvious. Decomposability follows from Lemma 11(i): Given f let $\hat{f}$ be $\mathcal{S}_{\text{stp}}$-measurable and such that $\hat{f} \sim_{\text{stp}} f$. Lemma 11(i) implies that for every $A \in \mathcal{S}_{\text{stp}}$ 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 8(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 $L_1(\mathcal{T}, q)$:

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

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 $T\xi_n \uparrow T\xi$. $\square$

D Predictive representation

D.1 Properties of the representation

We first characterize the collection of events that satisfy the sure-thing principle for a preference relation that admits a predictive representation.

Proposition 6. *If $\succsim$ admits a predictive representation $(u, \phi, \mathcal{T}, \pi)$, then $\mathcal{T} \subseteq \mathcal{S}_{\text{stp}}$. If, in addition, ϕ is not affine, then $\mathcal{S}_{\text{stp}}$ and $\mathcal{T}$ are π-equivalent.*

The proof of the result is divided in lemmas. Assume $\succsim$ admits a predictive representation $(u, \phi, \mathcal{T}, \pi)$. Let q be the restriction of π on $\mathcal{T}$, and let

$$Tu(f) = E_\pi[u(f)|\mathcal{T}].$$

By Lemma 7 and Theorem 3, the operator $T\colon B_b(\mathcal{S}, u(X)) \to L_\infty(\mathcal{T}, q, u(X))$ is monotone, decomposable, normalized, σ-order continuous, and affine. We define

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

Without loss of generality, assume that

$$\inf u(X) < 0 = \phi(0) \qquad \text{and} \qquad \sup u(X) > 1 = \phi(1). \tag{6}$$

Lemma 14. *If $A \in \mathcal{S}$ is π-equivalent to a $B \in \mathcal{T}$, then $A \in \mathcal{S}_{\mathrm{stp}}$. In particular, $\mathcal{T} \subseteq \mathcal{S}_{\mathrm{stp}}$.*

Proof. Let $A \in \mathcal{S}$ and $B \in \mathcal{T}$ be π-equivalent. For all acts f and h, we have $E_\pi[u(fAh)|\mathcal{T}] = E_\pi[u(fBh)|\mathcal{T}]$, which implies

$$\begin{aligned} T(u(f)\cdot 1_A + u(h)\cdot 1_{A^c}) &= T(u(f)\cdot 1_B + u(h)\cdot 1_{B^c}) \\ &= T(u(f))\cdot[1_B] + T(u(h))\cdot[1_{B^c}]. \end{aligned}$$

We deduce that $fAh \succsim gAh$ if only if $\int_B \phi(Tu(f))\,\mathrm{d}q \geq \int_B \phi(Tu(g))\,\mathrm{d}q$ if and only if $fAh' \succsim gAh'$. The same argument applies also to A^c, being π-equivalent to B^c. It follows that $A \in \mathcal{S}_{\mathrm{stp}}$. $\square$

Lemma 15. *Let $A \in \mathcal{S}_{\mathrm{stp}}$ and $B, C \in \mathcal{T}$. If $T1_A = [1_B]$ and $T1_{A^c} = [1_C]$, then A is π-equivalent to B.*

Proof. From $T1_A = [1_B]$ it follows that

$$E_\pi[1_A \cdot 1_{B^c}] = E_\pi[1_B \cdot 1_{B^c}] = 0.$$

Since, in addition, $T1_{A^c} = [1_C]$, we obtain that

$$[1_{B^c}] = 1 - [1_B] = 1 - T1_A = T1_{A^c} = [1_C].$$

Thus we have

$$E_\pi[1_{A^c} \cdot 1_B] = E_\pi[1_C \cdot 1_B] = E_\pi[1_{B^c} \cdot 1_B] = 0.$$

We conclude that A is π-equivalent to B. $\square$

The next lemma concludes the proof of Proposition 6.

Lemma 16. *If there is $A \in \mathcal{S}_{\text{stp}}$ such that $T1_A \neq [1_B]$ for all $B \in \mathcal{T}$, then ϕ is affine.*

Proof. Let $\rho \in B_b(\mathcal{T}, [0,1])$ be a representative of the equivalence class $T1_A$. Since $[\rho] \neq [1_B]$ for all $B \in \mathcal{T}$, we have that

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

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 $t_* < \bar{t} < t^*$. Define $s_*, s^* \in [-\infty, \infty]$ by

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

By (6), $s_* < 0$ and $s^* > 1$. Let

$$\psi, \psi' \in B_b(\mathcal{T}, (s_* t_*, s^* t_*)) \quad \text{and} \quad \varphi \in B_b(\mathcal{T}, s_*(1-t^*), s^*(1-t^*)).$$

Then, for all $\omega \in G_{t_*,t^*}$,

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

Hence $\psi = \rho\xi$, $\psi' = \rho\xi'$, and $\varphi = (1-\rho)\zeta$ for some $\xi, \xi', \zeta \in B_b(\mathcal{T}, u(X))$. Thus

$$T(\xi \cdot 1_A + \zeta \cdot 1_{A^c}) = [\psi] + [\varphi] \quad \text{and} \quad T(\xi' \cdot 1_A + \zeta \cdot 1_{A^c}) = [\psi'] + [\varphi].$$

Because A satisfies the sure-thing principle, we obtain that

$$\begin{gathered}\int_\Omega \phi\left(\psi\right)\,\mathrm{d}q(\cdot|G_{t_*,t^*}) \geq \int_\Omega \phi\left(\psi'\right)\,\mathrm{d}q(\cdot|G_{t_*,t^*}) \\ \Longleftrightarrow \\ \int_\Omega \phi\left(\psi+\varphi\right)\,\mathrm{d}q(\cdot|G_{t_*,t^*}) \geq \int_\Omega \phi\left(\psi'+\varphi\right)\,\mathrm{d}q(\cdot|G_{t_*,t^*}).\end{gathered} \tag{7}$$

where $q(\cdot|G_{t_*,t^*})$ is the conditional probability given the event G_{t_*,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 $s \in (s_*(1-t^*), s^*(1-t^*))$ there are $\alpha(s) \in \mathbb{R}$ and $\beta(s) > 0$ such that for all $t \in (s_*t_*, s^*t_*)$

$$\phi(s+t) = \alpha(s) + \beta(s)\phi(t). \tag{8}$$

Moreover, ϕ is strictly increasing. Thus, by Lemma 2, ϕ is either affine or exponential on the interval

$$(s_*(t_*+1-t^*), s^*(t_*+1-t^*)).$$

By taking t_* and t^* arbitrarily close to zero and one, we deduce that ϕ is either affine or exponential on the interval (s_*, s^*). By continuity, ϕ is either affine or exponential on its whole domain $u(X)$.

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 $\xi \in B_b(\mathcal{T}, (s_*t_*, s^*t_*))$ 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 (7) 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. $\square$

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

Proposition 7. *If $\succsim$ is represented by $(u, \phi, \mathcal{T}, \pi)$, then $f \succsim_{\text{stp}} g$ implies*

$$E_\pi[u(f)|\mathcal{T}] \geq E_\pi[u(g)|\mathcal{T}]. \tag{9}$$

If in addition ϕ is not affine, then $f \succsim_{\text{stp}} g$ if and only if (9) holds.

Proof. Let q be the restriction of π on $\mathcal{T}$. First observe that $f \succsim_A g$ for all $A \in \mathcal{T}$ is equivalent to

$$\int_A \phi\left(E_\pi[u(f)|\mathcal{T}]\right) \mathrm{d}q \geq \int_A \phi\left(E_\pi[u(g)|\mathcal{T}]\right) \mathrm{d}q, \qquad \forall A \in \mathcal{T},$$

which in turn is equivalent to (9), being ϕ strictly increasing. By Proposition 6 we have $\mathcal{T} \subseteq \mathcal{S}_{\text{stp}}$. Thus $f \succsim_{\text{stp}} g$ implies (9). If in addition ϕ is not affine, then $\mathcal{T}$ and $\mathcal{S}_{\text{stp}}$ are π-equivalent by Proposition 6. If $A \in \mathcal{S}_{\text{stp}}$ is π-equivalent to $B \in \mathcal{T}$, then $u(fAh)$ and $u(fBh)$ are equal π-almost surely for every third act h, which implies

$$E_\pi[u(fAh)|\mathcal{T}] = E_\pi[u(fBh)|\mathcal{T}], \qquad \forall h \in \mathfrak{F}.$$

We deduce that $f \succsim_{\text{stp}} g$ if and only if (9) holds. $\square$

We next characterize the null events induced by the representation.

Lemma 17. *Let $\succsim$ admit a predictive representation $(u, \phi, \mathcal{T}, \pi)$. An event $A \in \mathcal{S}$ is null if and only if $\pi(A) = 0$.*

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

$$\phi(u(x)) = E_\pi\Big[\phi\Big(E_\pi[u(yAx)|\mathcal{T}]\Big)\Big] = E_\pi\Big[\phi\Big(u(y)\pi(A|\mathcal{T}) + u(x)\pi(A^c|\mathcal{T})\Big)\Big].$$

Being ϕ strictly increasing, $\pi(A|\mathcal{T}) = [0]$, which in turn implies that $\pi(A) = 0$.

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

D.2 Representation theorem and uniqueness

The next result is a representation theorem for $\succsim$.

Theorem 5. *A preference $\succsim$ satisfies Axioms 1-6 if and only if it admits a predictive representation.*

Proof. We first establish the sufficiency of the axioms for the representation. Assume axioms 1-6 are satisfied. Note that Axiom 7 is satisfied as well: by Lemma 6(i) if $x \succsim y$, then $x \succsim_{\text{stp}} y$, 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 4. By Theorem 3, to conclude the proof of sufficiency it is enough to show that T is affine.

To this end, we first show that

$$f \succsim_{\text{stp}} g \Rightarrow \alpha f + (1-\alpha)h \succsim_{\text{stp}} \alpha f + (1-\alpha)h \text{ for all } \alpha \in [0,1], h \in \mathfrak{F}. \tag{10}$$

By Lemma 11(i) we have $fAh \succsim_{\text{stp}} gAh$ for all $A \in \mathcal{S}_{\text{stp}}$ and $h \in \mathfrak{F}$. By Axiom 4

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

thus, since $A \in \mathcal{S}_{\text{stp}}$, we have $(\alpha f + (1-\alpha)h)Ah' \succsim (\alpha g + (1-\alpha)h)Ah'$ for all $h' \in \mathfrak{F}$. Hence (10) follows.

Now recall that T represents $\succsim_{\text{stp}}$. Hence for $\hat{f}$ such that $Tu(f) = [u(\hat{f})]$ and $\hat{g}$ such that $Tu(g) = [u(\hat{g})]$, (10) implies that for all $\alpha \in [0,1]$

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

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 7 that T is affine.

We now turn to the proof of necessity. Assume $\succsim$ admits a predictive representation $(u, \phi, \mathcal{T}, \pi)$. Let q be the restriction of π on $\mathcal{T}$, and let

$$Tu(f) = E_\pi[u(f)|\mathcal{T}].$$

By Lemma 7 and Theorem 3 the operator $T\colon B_b(\mathcal{S}, u(X)) \to L_\infty(\mathcal{T}, q, u(X))$ is monotone, decomposable, normalized, σ-order continuous, and affine.

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 $E_\pi[u(f)|\mathcal{T}] > E_\pi[u(g)|\mathcal{T}]$ and 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 4 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.

Let $f \succsim_{\text{stp}} g$. By Proposition 6 we have $E_\pi[u(f)|\mathcal{T}] \geq E_\pi[u(g)|\mathcal{T}]$. This implies for all $\alpha \in [0,1]$ and $h \in \mathfrak{F}$

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

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

If ϕ is affine, then

$$f \succsim g \quad \iff \quad E_\pi[u(f)] \geq E_\pi[u(g)],$$

which implies that Savage's P4 holds for all events in $\mathcal{S}$. Suppose now that ϕ is not affine and let $A, B \in \mathcal{S}_{\text{stp}}$. By Proposition 6 there are events $C, D \in \mathcal{T}$ such that $[1_C] = \pi(A|\mathcal{T})$ and $[1_D] = \pi(B|\mathcal{T})$. Thus for all $x, y \in X$ such that $x > y$

$$xAy \succsim xBy \quad \iff \quad q(C) \geq q(D).$$

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 $\mathcal{T}$-measurable events. By Proposition 6 each A_i satisfies the sure-thing principle. Because T is decomposable, by Lemma 3

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

A similar condition holds for $V(hA_ig)$. Since q is nonatomic, for every $\varepsilon > 0$ we can choose $A_1, \dots, A_n$ 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. $\square$

The next results describe the uniqueness properties of the representation.

Proposition 8. *If $\succsim$ admits a predictive representation $(u, \phi, \mathcal{T}, \pi)$ and $\mathcal{U} \subseteq \mathcal{S}$ is a σ-algebra π-equivalent to $\mathcal{T}$, then $\succsim$ admits a predictive representation $(u, \phi, \mathcal{U}, \pi)$.*

Proof. It follows from Lemma 1 that

$$E_\pi\Big[\phi\Big(E_\pi[u(f)|\mathcal{T}]\Big)\Big] = E_\pi\Big[\phi\Big(E_\pi[u(f)|\mathcal{U}]\Big)\Big]$$

It remains to show that π is non-atomic on $\mathcal{U}$. Let $A \in \mathcal{U}$ such that $\pi(A) > 0$. Take $B \in \mathcal{T}$ that is π-equivalent to A. Because π is nonatomic on $\mathcal{T}$, there exists $B' \subseteq B$ such that $0 < \pi(B') < \pi(B) = \pi(A)$. Let $A' \in \mathcal{U}$ be π-equivalent to B'. Then B' is also π-equivalent to $A \cap A'$. Thus $\pi(A \cap A') = \pi(B') \in (0, \pi(A))$. We conclude that π is nonatomic on $\mathcal{U}$. $\square$

Theorem 6. *Two predictive representations $(u_1, \phi_1, \mathcal{T}_1, \pi_1)$ and $(u_2, \phi_2, \mathcal{T}_2, \pi_2)$ 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, $\mathcal{T}_1$ and $\mathcal{T}_2$ are π_1-equivalent.*

Proof. 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 Proposition 6 and Proposition 8 the preference $\succsim$ admits a predictive representation $(u, \phi_2, \mathcal{S}_{\text{stp}}, \pi_2)$. Moreover $\mathcal{T}_1 \subseteq \mathcal{S}_{\text{stp}}$ again by Proposition 6. Thus for all acts f and g that are $\mathcal{T}_1$-measurable

$$\int_\Omega \phi_1(u(f))) \, \mathrm{d}\pi_1 \geq \int_\Omega \phi_1(u(g))) \, \mathrm{d}\pi_1 \iff \int_\Omega \phi_2(u(f)) \, \mathrm{d}\pi_2 \geq \int_\Omega \phi_2(u(g)) \, \mathrm{d}\pi_2.$$

In particular, for all $A, B \in \mathcal{T}_1$, $\pi_1(A) \geq \pi_1(B)$ if and only if $\pi_2(A) \geq \pi_2(B)$. Because π_1 is nonatomic on $\mathcal{T}_1$, then by standard arguments we obtain that $\pi_1(A) = \pi_2(A)$ for all $A \in \mathcal{T}_1$. 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. In addition, for $i \in \{1, 2\}$,

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

By the uniqueness properties of the expected utility representation, $\pi_1 = \pi_2$.

Assume now that both ϕ_1 and ϕ_2 are not affine. By Proposition 6 and Proposition 8 the preference $\succsim$ admits the representations $(u, \phi_1, \mathcal{S}_{\text{stp}}, \pi_1)$ and $(u, \phi_2, \mathcal{S}_{\text{stp}}, \pi_2)$. Moreover $\mathcal{T}_1$ is π_1-equivalent to $\mathcal{S}_{\text{stp}}$ and $\mathcal{T}_2$ is π_2-equivalent to $\mathcal{S}_{\text{stp}}$. Let q_i be the restriction of π_i, $i = 1, 2$, on $\mathcal{S}_{\text{stp}}$. It is non-atomic. For all acts f and g that are $\mathcal{S}_{\text{stp}}$-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 7 that for all act f, $E_{\pi_1}[u(f)|\mathcal{S}_{\text{stp}}] = E_{\pi_2}[u(f)|\mathcal{S}_{\text{stp}}]$. Since $q_1 = q_2$, we obtain that $E_{\pi_1}[u(f)] = E_{\pi_2}[u(f)]$. By the uniqueness properties of the expected utility representation, $\pi_1 = \pi_2$. $\square$

E Proofs of the results in Sections 5 and 6

E.1 Proof of Theorem 1

The result follows immediately from Proposition 1 and Theorem 5.

E.2 Proof of Theorem 2

The uniqueness properties of the predictive representation follow from Theorem 6. 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 $\mu_1(B \cap \mathcal{P}_1) = \mu_2(B \cap \mathcal{P}_2)$ for all $B \in \Sigma$. By Lemma 24, 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 $A \in \mathcal{S}$ 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 $\mu_1(B \cap \mathcal{P}_1) = \mu_2(B \cap \mathcal{P}_2)$ for the set $B = \{p \in \Delta : p(A) \leq t\}$. Since sets of this form generate Σ, the desired result follows.

E.3 Proof of Proposition 2

The equivalence of (i) and (iii) follows from Proposition 7. Let k witness the identifiability of $\mathcal{P}$. By Lemma 27 the preference $\succsim$ admits a predictive representation $(u, \phi, \mathcal{T}_{\mathcal{P}}, \pi_\mu)$. By Lemma 24 the kernel k is a regular conditional probability of π_μ given $\mathcal{T}_{\mathcal{P}}$. Thus, being (i) and (iii) equivalent,

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

The event $A = \{\omega : \int_\Omega u(f) - u(g) \, \mathrm{d}k(\omega) \geq 0\}$ belongs to $\mathcal{T}_{\mathcal{P}}$. 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

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

E.4 Proof of Proposition 3

By Proposition 6 the σ-algebras $\mathcal{S}_{\text{stp}}$ and $\mathcal{T}$ 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 $\mathcal{T}$. From Lemma 17 it follows that $\mathcal{S}_{\text{stp}}$, $\sigma(k)$, and $\mathcal{T}$ are all equivalent up to null events.

F Proof of Proposition 4

That (i) follows from (ii) is an immediate consequence of Proposition 2. We now prove the converse implication. By Lemma 24, the two preferences admit predictive representations $(u_1, \phi_1, \mathcal{T}_1, \pi_1)$ and $(u_2, \phi_2, \mathcal{T}_2, \pi_2)$ where each $\mathcal{T}_i = \mathcal{T}_{\mathcal{P}_i}$ is the corresponding σ-algebra of zero-one events and $\pi_i = \pi_{\mu_i}$ is the barycenter of μ_i. Moreover, by Lemma 26 each k_i is a regular conditional probability of π_i, and $\mathcal{T}_i$ and $\sigma(k_i)$ are π_i-equivalent.

Assume $f \succsim^1_{\text{stp}} g$ implies $f \succsim^2_{\text{stp}} g$. This yields that $\succsim^1$ and $\succsim^2$ agree on X and thus, without loss of generality, that $u_1 = u_2 = u$ and $(-1, 1) \subseteq u(X)$. Proposition 2 shows (i) can be rewritten as

$$E_{\pi_1}[u(f)|\mathcal{T}_1] \geq E_{\pi_1}[u(g)|\mathcal{T}_1] \implies E_{\pi_2}[u(f)|\mathcal{T}_2] \geq E_{\pi_2}[u(g)|\mathcal{T}_2]. \tag{11}$$

By the linearity of the conditional expectation operator, (11) is equivalent to the the property that for all $\xi \in B(\mathcal{S})$,

$$E_{\pi_1}[\xi|\mathcal{T}_1] \geq 0 \implies E_{\pi_2}[\xi|\mathcal{T}_2] \geq 0. \tag{12}$$

A first implication of this property is that for all $A \in \mathcal{S}$, if $\pi_1(A) = 0$ then $E_{\pi_1}[1_A|\mathcal{T}_1] = 0$, and thus $E_{\pi_2}[1_A|\mathcal{T}_2] = 0$, which is equivalent to $\pi_2(A) = 0$. Thus, π_2 is absolutely continuous with respect to π_1. Given $\xi \in B(\mathcal{S})$, $E_{\pi_1}[\int_\Omega \xi \, \mathrm{d}k_1 - \xi|\mathcal{T}_1] = 0$ since k_1 is a regular conditional probability for π_1. Thus, by (12),

$$E_{\pi_2}[\xi|\mathcal{T}_2] = E_{\pi_2}\left[\int_\Omega \xi \, \mathrm{d}k_1 \middle| \mathcal{T}_2\right].$$

Taking $\xi = 1_A$, $A \in \mathcal{S}$, and using the fact that k_2 is a regular conditional probability for π_2 we obtain that, for π_2-almost all ω,

$$k_2(\omega, A) = \int_\Omega k_1(\omega', A) k_2(\omega, \mathrm{d}\omega'). \tag{13}$$

Because $\pi_2 = \pi_{\mu_2} = \int_{\mathcal{P}_2} p_2 \, \mathrm{d}\mu_2(p_2)$ and (13) holds π_2-almost surely, then it must hold p_2-almost surely for μ_2-almost every $p_2 \in \mathcal{P}_2$. But then, the identifiability of $\mathcal{P}_2$ yields that for μ_2-almost every p_2,

$$p_2(A) = \int_\Omega k_1(\omega, A)\,\mathrm{d}p_2(\omega).$$

Define $M\colon \Delta(\Omega) \to \Delta(\mathcal{P}_1)$ by letting $M(p)$ be the pushforward of $p \in \Delta$ under k_1. It is routine to prove that M is measurable, and so is its restriction on $\mathcal{P}_2$. It follows that for μ_2-almost every p_2, $p_2(A) = \int_{\mathcal{P}_1} p_1(A)\,\mathrm{d}M(p_1|p_2)$.

Let $\xi = 1_A$ with $A \in \mathcal{T}_2$, and let $\zeta = k_1(A|\omega)$. We have $1_A = E_{\pi_2}[\zeta|\mathcal{T}_2]$. Because $0 \le \zeta \le 1$, it follows that $E = \{\zeta = 1\} \in \mathcal{T}_1$ satisfies $\pi_2(A \triangle E) = 0$. So, $\mathcal{T}_2 \subseteq \mathcal{T}_1$, π_2-almost surely. By Lemma 24, each $\mathcal{T}_i$ is π_i-equivalent to $\sigma(k_i)$, and the fact that π_2 is absolutely continuous with respect to π_1 implies $\sigma(k_2)$ and $\mathcal{T}_2$ are π_1-equivalent. It follows that $\sigma(k_2) \subseteq \sigma(k_1)$, π_1-almost surely.

G Proof of Proposition 5

Set $u = u_1 = u_2$ and $\mathcal{S}_{\text{stp}} = \mathcal{S}^1_{\text{stp}} = \mathcal{S}^2_{\text{stp}}$. For each agent i, let $k_i : \Omega \to \Delta$ be a kernel that identifies $\mathcal{P}_i$. By Proposition 1, each preference $\succsim_i$ admits a predictive representation $(u, \phi_i, \sigma(k_i), \pi_{\mu_i})$. Being ϕ_i not affine, by Proposition 3 the σ-algebra $\sigma(k_i)$ and $\mathcal{S}_{\text{stp}}$ are π_{μ_i}-equivalent. Thus $\succsim_i$ admits a predictive representation $(u, \phi_i, \mathcal{S}_{\text{stp}}, \pi_{\mu_i})$. We define

$$V_i(f) = E_{\pi_{\mu_i}}\left[\phi_i\left(E_{\pi_{\mu_i}}[u(f)|\mathcal{S}_{\text{stp}}]\right)\right] \quad \text{and} \quad W_i(f) = \phi_i^{-1}\left(V_i(f)\right).$$

Begin ϕ_i strictly increasing, $f \succsim_i g$ if and only if $W_i(f) \ge W_i(g)$. In addition, $W_i(x) = u(x)$. The rest of the proof is organized in lemmas. The first lemma is the standard characterization of comparative ambiguity aversion in terms of certainty equivalents.

Lemma 18. *Condition (i) holds if and only if $W_1 \le W_2$.*

Proof. "If." Suppose $f \succsim_1 x$, i.e., $W_1(f) \ge W_1(x) = u(x)$. From $W_2 \ge W_1$ it follows that $W_2(f) \ge W_1(f)$, thus $W_2(f) \ge u(x) = W_2(x)$. We conclude that $f \succsim_2 x$.

"Only if." Take f and x such that $W_1(f) = W_1(x)$, i.e., $f \sim_1 x$. From (i) it follows that $f \succsim_2 x$, i.e., $W_2(f) \ge W_2(x)$. Since $W_1(x) = W_2(x)$, we deduce that $W_1(f) \le W_2(f)$. ☐

To prove the next result, we adapt an argument used in the proof of Klibanoff, Mukerji, and Seo (2014, Lemma C.1). See also Yaari (1969, Remark 1).

Lemma 19. *If (i) holds, then $\pi_{\mu_1} = \pi_{\mu_2}$.*

Proof. Let $A \in \mathcal{S}$. Take $x \in X$ such that $u(x)$ is in the interior of $u(X)$. Being $u(X)$ an interval, for every $t \in \mathbb{R}$ such that $u(x) + t \in u(X)$, we can find an outcome $y_t \in X$ such act $u(y_t) = u(x) + t$. Define $f_t = y_t A x$. Observe that

$$\begin{aligned}\lim_{t\to 0} \frac{W_i(f_t) - W_i(x)}{t} &= \lim_{t\to 0} \frac{\phi_i^{-1}\left(E_{\pi_{\mu_i}}\left[\phi_i\left(u(x) + t\pi_{\mu_i}(A|\mathcal{S}_{\mathrm{stp}})\right)\right]\right) - u(x)}{t} \\ &= \frac{E_{\pi_{\mu_i}}\left[\phi_i'(u(x))\pi_{\mu_i}(A|\mathcal{S}_{\mathrm{stp}})\right]}{\phi_i'(u(x))} = \pi_{\mu_i}(A)\end{aligned}$$

where ϕ_i' is the derivative of ϕ_i. In addition, by Lemma 18 we have

$$W_1(f_t) - W_1(x) = W_1(f_t) - u(x) \le W_2(f_t) - u(x) = W_2(f_t) - W_2(x).$$

Overall, we obtain

$$\begin{aligned}\pi_{\mu_1}(A) &= \lim_{t\to 0^+} \frac{W_1(f_t) - W_1(x)}{t} \le \lim_{t\to 0^+} \frac{W_2(f_t) - W_2(x)}{t} = \pi_{\mu_2}(A), \\ \pi_{\mu_1}(A) &= \lim_{t\to 0^-} \frac{W_1(f_t) - W_1(x)}{t} \ge \lim_{t\to 0^-} \frac{W_2(f_t) - W_2(x)}{t} = \pi_{\mu_2}(A).\end{aligned}$$

We conclude that $\pi_{\mu_1}(A) = \pi_{\mu_2}(A)$. $\square$

Lemma 20. *If (i) holds, then the function $\psi = \phi_1 \circ \phi_2^{-1}$ is concave.*

Proof. By Lemma 19 we can set $\pi = \pi_{\mu_1} = \pi_{\mu_2}$. Take $\alpha \in (0,1)$ and $x, y \in X$. Since π is non-atomic on $\mathcal{S}_{\mathrm{stp}}$, we can find an event $A \in \mathcal{S}_{\mathrm{stp}}$ such that $\pi_\mu(A) = \alpha$. Define $f = xAy$. It follows from Lemma 18 that

$$\begin{aligned}\alpha\psi(\phi_2(u(x))) + (1-\alpha)\psi(\phi_2(u(y))) &= \alpha\phi_1(u(x)) + (1-\alpha)\phi_1(u(y)) = \phi_1(W_1(f)) \\ &\le \phi_1(W_2(f)) = \psi(\alpha\phi_2(u(x)) + (1-\alpha)\phi_2(u(y))).\end{aligned}$$

We deduce that the function ψ is concave. $\square$

The last two lemmas show that (i) implies (ii). The next result concludes the proof of Proposition 5.

Lemma 21. *If (ii) holds, then (i) holds.*

Proof. Set $\pi = \pi_{\mu_1} = \pi_{\mu_2}$ and $\psi = \phi_1 \circ \phi_2^{-1}$. Assume $f \succsim_1 x$, that is,

$$E_\pi \left[\phi_1 \left(E_\pi[u(f)|\mathcal{S}_{\text{stp}}]\right)\right] \geq \phi_1(u(x)).$$

We can rewrite the inequality as

$$E_\pi \left[(\psi \circ \phi_2) \left(E_\pi[u(f)|\mathcal{S}_{\text{stp}}]\right)\right] \geq (\psi^{-1} \circ \phi_2)(u(x)).$$

Being ψ concave, by Jensen's inequality

$$E_\pi \left[\phi_2 \left(E_\pi[u(f)|\mathcal{S}_{\text{stp}}]\right)\right] \geq \phi_2(u(x)),$$

that is, $f \succsim_2 x$. $\square$

In section K in the online appendix we show that the assumption $\mathcal{S}^1_{\text{stp}} = \mathcal{S}^2_{\text{stp}}$ cannot be weakened.

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.

R. Best and J. Boice. Where the latest covid-19 models think we're headed—and why they disagree, 2021. URL `https://projects.fivethirtyeight.com/covid-forecasts/`.

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.

W. A. Brock, S. N. Durlauf, and K. D. West. Model uncertainty and policy evaluation: Some theory and empirics. *Journal of Econometrics*, 136:629–664, 2007.

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.

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.

B. de Finetti. Probabilities of probabilities: A real problem or a misunderstanding. pages 1–10. North-Holland, 1977.

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.

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

T. Gajdos, T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud. Attitude toward imprecise information. *Journal of Economic Theory*, 140(1):27–65, 2008.

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. Marinacci. Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102:251–289, 2002.

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.

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

I. Gilboa and M. Marinacci. Ambiguity and the bayesian paradigm. In *Advances in Economics and Econometrics: Tenth World Congress*, volume 1, page 179. Cambridge University Press, 2013.

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, and K. Seo. Perceived ambiguity and relevant measures. *Econometrica*, 82:1945–1978, 2014.

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

L. LeCam and L. Schwartz. A necessary and sufficient condition for the existence of consistent estimates. *The Annals of Mathematical Statistics*, 31(1):140–150, 1960.

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

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.

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.

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

A. Wald. *Statistical Decision Functions*. Wiley, 1950.

M. E. Yaari. Some remarks on measures of risk aversion and on their uses. *Journal of Economic Theory*, 1:315–329, 1969.