Online Appendix

H Identifiability in statistics and econometrics

Our definition of identifiability agrees with mainstream statistic and econometric usage in most standard settings. A standard example is random sampling, where the state space Ω is a product of infinitely many copies of a measurable set S, and a state $\omega = (\omega_1, \omega_2, \ldots)$ corresponds to an infinite sample realization. A parameter $\theta \in \Theta$ determines the i.i.d. probability distribution $p_\theta \in \Delta(\Omega)$ of the sample. The statistical model is parametrized and given by $\mathcal{P} = \{p_\theta : \theta \in \Theta\}$.

In the statistics and econometrics tradition (see, e.g., Lehmann and Casella, 2006, p. 24), a statistical model $\mathcal{P} = \{p_\theta : \theta \in \Theta\}$ is *identifiable* if different parameter values induce different sample distributions:

$$\text{if} \quad \theta \neq \theta' \quad \text{then} \quad p_\theta \neq p_{\theta'}. \tag{14}$$

A related property is the existence of a consistent estimator (see, e.g., Lehmann and Casella, 2006, p. 54). Let $\mathcal{S}_n \subseteq \mathcal{S}$ be the σ-algebra generated by the first n sample realizations. A statistical model $\mathcal{P} = \{p_\theta : \theta \in \Theta\}$ admits a *consistent estimator* if, for every $n = 1, 2, \ldots$, there is a $\mathcal{F}_n$-measurable function $k_n : \Omega \to \Theta$ such that, for all $\theta \in \Theta$ and $\epsilon > 0$,

$$\lim_{n\to\infty} p_\theta \left(\{\omega : |k_n(\omega) - \theta| \geq \epsilon\}\right) = 0. \tag{15}$$

It is easy to see that identifiability is a necessary condition for the existence of a consistent estimator. Under some regularity conditions, identifiability is also a sufficient condition for the existence of a consistent estimator. For example, if S is a finite set and the statistical model is identifiable, then the sequence (k_n) can be constructed using the sequence of empirical distribution functions (see, e.g., LeCam and Schwartz, 1960, for more general results on the existence of consistent estimators).

In our paper, the state space may not feature obvious symmetries or repetitions, nor the statistical model have a natural parametrization. Therefore, we find it convenient to translate (15) rather than (14). In our paper, a statistical model $\mathcal{P} \subseteq \Delta(\Omega)$ is identifiable if there exists a $\mathcal{S}$-measurable function $k : \Omega \to \mathcal{P}$ such that, for all $p \in \mathcal{P}$,

$$p(\{\omega : k(\omega) = p\}) = 1. \tag{16}$$

Essentially, the kernel k is a consistent estimator for the statistical model $\mathcal{P}$.

Barring technical differences, (14), (15), and (16) all reflect the same idea: there is some information that allows the decision maker to resolve their uncertainty about the true law governing the state of the world.

I Identifiability and Dynkin spaces

In the paper we follow Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013) for a general formulation of identifiable smooth preferences. The exposition is slightly different: to put restrictions on $\mathcal{P}$, they adopt the formalism of Dynkin spaces, while we use the notion of identifying kernel. In this section we verify that the two approaches are equivalent.

Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013) provide the following definition of Dynkin space (Dynkin, 1978).

Definition 5. Let $\mathcal{P} \subseteq \Delta$ be a nonempty set. The triple $(\Omega, \mathcal{S}, \mathcal{P})$ is a *Dynkin space* if there are a σ-algebra $\mathcal{T} \subseteq \mathcal{S}$, a set $W \in \mathcal{S}$, and a measurable function $k : \Omega \to \Delta$ such that

(i). for every $p \in \mathcal{P}$, the kernel k is a regular conditional probability of p given $\mathcal{T}$;

(ii). $p(W) = 1$ for all $p \in \mathcal{P}$ and $k(W) \subseteq \mathcal{P}$.

Among other results, they study smooth-ambiguity preferences $(u, \phi, \mathcal{S}(\mathcal{P}), \mu)$ where $(\Omega, \mathcal{S}, \mathcal{P})$ is a Dynkin space and $\mathcal{S}(\mathcal{P})$ is the set of strong extreme points of $\mathcal{P}$.

Definition 6. Let $\mathcal{P} \subseteq \Delta$ be a nonempty set. An element $p \in \mathcal{P}$ is a *strong extreme point* of $\mathcal{P}$ if, for every prior μ on $\mathcal{P}$, $\pi_\mu = p$ implies $\mu(\{p\}) = 1$.

The next result shows that the class of identifiable smooth preferences we study in this paper coincides with the class of smooth preferences they consider.

Proposition 9. *A nonempty set $\mathcal{P} \subseteq \Delta$ is identifiable if and only if $\mathcal{P}$ is the set of strong extreme points of a Dynkin space.*

The proof of the proposition relies on the following characterization of the strong extreme points of a Dynkin space, which is due to Dynkin (1978) (see also Theorem 17 of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio, 2013).

Lemma 22. *If* $(\Omega, \mathcal{S}, \mathcal{P})$ *is a Dynkin space, then*

$$\mathcal{S}(\mathcal{P}) = \{p \in \mathcal{P} : p(\{\omega : k(\omega) = p\}) = 1\}.$$

Proof of Proposition 9. If $(\Omega, \mathcal{S}, \mathcal{P})$ is a Dynkin space, then by Lemma 22 the set $\mathcal{S}(\mathcal{P})$ is identifiable. Conversely, suppose that $\mathcal{P}$ is identifiable. Let $k : \Omega \to \mathcal{P}$ be a kernel that witnesses the identifiability of $\mathcal{P}$. Let $\mathcal{T} \subseteq \mathcal{S}$ be given by

$$\mathcal{T} = \{A : p(A) \in \{0, 1\} \text{ for all } p \in \mathcal{P}\}.$$

Define $W = \Omega$. By Lemma 23, for every $p \in \mathcal{P}$, the kernel k is a regular conditional probability of p given $\mathcal{T}$. Moreover, trivially $p(W) = 1$ for every $p \in \mathcal{P}$ and $k(W) \subseteq \mathcal{P}$. Thus $(\Omega, \mathcal{S}, \mathcal{P})$ is Dynkin space. By Lemma 22 we conclude that $\mathcal{P} = \mathcal{S}(\mathcal{P})$. □

J Identifiability for general ambiguity preferences

An interesting question concerns identifiability for general ambiguity preferences. Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013) and Al-Najjar and De Castro (2014) characterize the identifiable versions of general ambiguity preferences when the statistical model is objectively given or is based on exogenous exchangeability assumptions. In our paper the statistical model is fully subjective but we focus on smooth ambiguity preferences. A natural question is how to extend our analysis to more general preferences.

It is our impression that making progress on such a question is challenging, and that it requires non-trivial technical innovations. What is certain is that some of the results that we establish for identifiable smooth preferences cannot be obtained for more general preferences. For example, our uniqueness result, which shows that $\mathcal{P}$ and μ can be recovered uniquely from choice behavior, does not hold for maxmin preferences. To illustrate, consider the identifiable version of maxmin preferences. Following Cerreia-Vioglio et al. (2013, Section 4.1) and Al-Najjar and De Castro (2014, Section 4.2), in such model an act f is evaluated according to the criterion

$$V(f) = \min_{\mu \in \mathcal{M}} \int_{\mathcal{P}} \int_{\Omega} u(f) \, \mathrm{d}p \, \mathrm{d}\mu(p) \tag{17}$$

where $\mathcal{P} \subseteq \Delta$ is an identifiable set of probability measures, and $\mathcal{M} \subseteq \Delta(\mathcal{P})$ is a compact convex set of priors on $\mathcal{P}$. Suppose that as in this paper, and unlike Cerreia-Vioglio et al. (2013) and Al-Najjar and De Castro (2014), that the only primitive is a preference relation over Anscombe-Aumann acts. It is then impossible to uniquely recover $\mathcal{P}$ and $\mathcal{M}$.

The underlying intuition is simple. As is well known, the Gilboa-Schmeidler representation does not distinguish between ambiguity perception and attitude. But at the same time the set $\mathcal{P}$ is supposed to reflect only ambiguity perception. This tension leads to an impossibility result: We can rewrite any representation such as (17) as

$$V(f) = \min_{\mu \in \mathcal{M}} \int_{\Omega} u(f)\, \mathrm{d}\pi_{\mu}$$

where $\pi_{\mu} = \int_{\mathcal{P}} p\, \mathrm{d}\mu(p)$ is the predictive probability associated to the prior μ. Thus (17) can be rewritten with the alternative parametrization

$$V(f) = \min_{\mu \in \mathcal{M}'} \int_{\mathcal{P}'} \int_{\Omega} u(f)\, \mathrm{d}p\, \mathrm{d}\mu(p) \tag{18}$$

where $\mathcal{P}' = \{\delta_{\omega} : \omega \in \Omega\}$ is the set of the Dirac probability measures over states of the world, and $\mathcal{M}'$ is obtained by pushing forward the elements of $\{\pi_{\mu} : \mu \in \mathcal{M}\}$ under the map $\omega \to \delta_{\omega}$. While $\mathcal{P}$ and $\mathcal{P}'$ could be very different from each other, both sets are identifiable and represent the same preference relation. This disappointing conclusion shows that, once we go beyond smooth preferences, identifiability is no longer a sufficient assumption for uniquely recovering the components of the representation.[13] Some separation between ambiguity attitude and perception seems an additional necessary ingredient.

K Comparative statics: an example

The next example shows that the assumption $\mathcal{S}^1_{\text{stp}} = \mathcal{S}^2_{\text{stp}}$ in Proposition 5 cannot be weakened.

Example 7. Let $(u_i, \phi_i, \mathcal{T}_i, \pi_i)$ be the predictive representation of $\succsim_i$. Suppose that $u_1 = u_2 = u$, $\phi_1 = \phi_2 = \phi$, $\mathcal{T}_2 \subseteq \mathcal{T}_1$, and $\pi_1 = \pi_2 = \pi$: the predictive representations of $\succsim_1$ and $\succsim_2$ are the same expect for $\mathcal{T}_2$ being a sub-σ-algebra of $\mathcal{T}_1$. Thus the agent corresponding to $\succsim_1$ needs more information to resolve their ambiguity.

[13]This observation is quite probably known, but we could not find it anywhere in the literature.

If ϕ is concave, then $\succsim_1$ is more ambiguity averse than $\succsim_2$. This follows from

$$E_\pi\Big[\phi\Big(E_\pi[u(f)|\mathcal{T}_2]\Big)\Big] = E_\pi\Big[\phi\Big(E_\pi[E_\pi[u(f)|\mathcal{T}_1]|\mathcal{T}_2]\Big)\Big] \geq E_\pi\Big[\phi\Big(E_\pi[u(f)|\mathcal{T}_1]\Big)\Big].$$

Thus, provided that ϕ is concave, $\succsim_1$ is more ambiguity averse than $\succsim_2$, despite the fact that $\phi_1 = \phi_2$. If ϕ is not affine, then $\mathcal{T}^1$ and $\mathcal{T}^2$ are equivalent to $\mathcal{S}^1_{\text{stp}}$ and $\mathcal{S}^2_{\text{stp}}$ up to null events. Therefore $\mathcal{S}^2_{\text{stp}}$ is a sub-σ-algebra $\mathcal{S}^1_{\text{stp}}$ up to null events. If the inclusion is strict, then the hypothesis of Proposition 5 is not satisfied.

L Omitted proofs

L.1 Proof of Lemma 3

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

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

Summing over i and subtracting $T(a)$ yields (3).

L.2 Proof of Lemma 4

The operator satisfies

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

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

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

The desired result follows.

L.3 Proof of Lemma 5

Let $\xi = \sum_{i=1}^{n} a_i 1_{A_i}$ where $A_1, \ldots, A_n$ is a $\mathcal{T}$-measurable partition and $a_1, \ldots, a_n \in U$. By applying Lemma 3 and the fact that T is normalized, we obtain

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

The general case where ξ is not simple now follows by Lemma 4 (being $B_0(\mathcal{S}, U)$ dense in $B_b(\mathcal{S}, U)$ with respect to the supnorm).

L.4 Proof of Theorem 3

Necessity is easy to verify. Turning to sufficiency, suppose T is monotone, decomposable, normalized, σ-order continuous, and affine. Define the functional $I \colon B_b(\mathcal{S}, U) \to \mathbb{R}$ by $I(\xi) = E_q[T\xi]$.

It is immediate that I is normalized (i.e. $I(c) = c$ for all $c \in \mathbb{R}$), monotone (if $\xi \geq \zeta$ then $I(\xi) \geq I(\zeta)$), and affine. Lemma 4 and the σ-additivity of q imply that I is pointwise continuous: if (ξ_n) is a bounded sequence such that $\xi_n \to \xi$ pointwise, then $I(\xi_n) \to I(\xi)$. By a standard application of the Riesz representation theorem, there exists $\pi \in \Delta(\mathcal{S})$ such that $I(\xi) = E_\pi[\xi]$. By Lemma 5 the operator T is projective, hence for every $\xi \in B_b(\mathcal{T}, U)$ we have $E_\pi[\xi] = I(\xi) = E_q[T\xi] = E_q[\xi]$. This implies π agrees with q on $\mathcal{T}$. For all $A \in \mathcal{T}$

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

where the last equality follows from T being decomposable. We conclude that $T\xi = E_\pi[\xi|\mathcal{T}]$, as desired.

L.5 Proof of Lemma 6

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

(ii). It follows from Axiom 3.

(iii). The claim is an application of the mixture space theorem (Herstein and Milnor, 1953) together with (ii) and Axioms 1 and 7.

L.6 Proof of Lemma 7

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

L.7 Proof of Lemma 8

(i). Choose $x, y \in X$ such that $x \succsim f_n(\omega) \succsim y$ for all n and ω. By Lemma 6(i) we have $x \succsim f_n \succsim y$ for all n. By Axiom 3 this implies that $x \succsim f \succsim y$ as well. If $x \sim y$, then $u(c(f_n)) = u(x) = u(c(f))$ for all n. Assume therefore that $x \succ y$. By Lemma 6(ii) we can choose $\alpha_n \in [0,1]$ and $\alpha \in [0,1]$ such that $f_n \sim \alpha_n x + (1 - \alpha_n) y$ and $f \sim \alpha x + (1 - \alpha) y$. Possibly passing to a subsequence, we can assume without loss of generality that $\alpha_n \to \beta$ for some $\beta \in [0,1]$. It follows from Axiom 3 that $f \sim \beta x + (1 - \beta) y$, i.e., $u(c(f)) = \beta u(x) + (1 - \beta) u(y)$, which in turn implies $\alpha = \beta$. Thus

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

(ii). Choose $x, y \in X$ such that $x \succsim f_n(\omega) \succsim y$ for all n and ω. By Axiom 3 this implies that $x \succsim f(\omega) \succsim y$ for all ω as well. Take $\xi_n \in B_b(\mathcal{S}, [0,1])$ and $\xi \in B_b(\mathcal{S}, [0,1])$ such that $u(f_n) = \xi_n u(x) + (1 - \xi_n) u(y)$ and $u(f) = \xi u(x) + (1 - \xi) u(y)$. Define $g_n = \xi_n x + (1 - \xi_n) y$ and $g = \xi x + (1 - \xi) y$. Observe that $u(f_n) = u(g_n)$ and $u(f) = u(g)$: it follows from Lemma 6(i) that $u(c(f_n)) = u(c(g_n))$ and $u(c(f)) = u(c(g))$. In addition, $u(f_n) \to u(f)$ pointwise implies $g_n \to g$ pointwise. The desired result then follows from (i) above.

(iii). Being A not null, there are f, g such that $f \succ_A g$. Take $w, z \in X$ such that $w \succsim f(\omega)$ and $g(\omega) \succsim z$ for all ω. By Lemma 6(i) we have $w \succ_A z$, that is, $w \succ zAw$. It follows from Axiom 5 that $x \succ yAx$, that is, $x \succ_A y$.

L.8 Proof of Proposition 1

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

$$\mathcal{T}_{\mathcal{P}} = \{A \in \mathcal{S} : p(A) \in \{0, 1\} \text{ for all } p \in \mathcal{P}\}. \tag{19}$$

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

Lemma 23. *Let $\mathcal{P} \subseteq \Delta$. A kernel $k \colon \Omega \to \mathcal{P}$ identifies $\mathcal{P}$ if and only if it witnesses the sufficiency of $\mathcal{T}_{\mathcal{P}}$ for $\mathcal{P}$.*

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

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

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

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

Lemma 23 can be used to relate our definition of identifiability to the notion of *Dynkin space* (Dynkin, 1978; Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio, 2013). Some of the results that appear in this section are already discussed in the original paper by Dyknin and in Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013). See, in particular, their Appendix B.

Lemma 24. *If a kernel $k \colon \Omega \to \mathcal{P}$ identifies $\mathcal{P} \subseteq \Delta$ and μ is a prior on $\mathcal{P}$, then (i) k is a regular conditional probability of π_μ given $\mathcal{T}_{\mathcal{P}}$, and (ii) $\sigma(k)$ and $\mathcal{T}_{\mathcal{P}}$ are π_μ-equivalent.*

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

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

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

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

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

Lemma 25. *If $\mathcal{P} \subseteq \Delta$ is identifiable, then (i) $\Sigma_{\mathcal{P}} = \Sigma^\star_{\mathcal{P}}$, and (ii) a prior μ on $\mathcal{P}$ is nonatomic if and only if π_μ is nonatomic on $\mathcal{T}_{\mathcal{P}}$.*

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

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

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

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

Lemma 26. *Let $\succsim$ admit an identifiable representation $(u, \phi, \mathcal{P}, \mu)$. Then it admits a predictive representation $(u, \phi, \mathcal{T}_{\mathcal{P}}, \pi_\mu)$.*

Proof. By Lemma 25 the measure π_μ is nonatomic on $\mathcal{T}_{\mathcal{P}}$. To conclude the proof it remains to show that for all $\xi \in B_b(\mathcal{S}, u(X))$

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

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

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

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

For an arbitrary $\mathcal{S}$-measurable ξ, Lemma 23 implies

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

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

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

The right-hand side is equal to $E_{\pi_\mu}\Big[\phi\Big(E_{\pi_\mu}[\xi|\mathcal{T}_{\mathcal{P}}]\Big)\Big]$, being k a regular conditional probability for π_μ (Lemma 24). □

Lemma 27. *Let $\succsim$ admit a predictive representation $(u, \phi, \mathcal{T}, \pi)$. Then it admits an identifiable representation $(u, \phi, \mathcal{P}, \mu)$ where $\pi_\mu = \pi$ and $\mathcal{T}_{\mathcal{P}}$ is π-equivalent to $\mathcal{T}$.*

Proof. Since $(\Omega, \mathcal{S})$ is standard Borel, π admits a regular conditional probability $k : \Omega \to \Delta$ with respect to $\mathcal{T}$. We define a prior μ on Σ as the pushforward of π under k. We now show that for each $A \in \mathcal{S}$ and for μ-almost all p

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

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

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

Hence, for μ-almost all p

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

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

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

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

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

The function $k \colon \Omega \to \mathcal{P}$ is $(\mathcal{T}, \Sigma_{\mathcal{P}})$-measurable and identifies $\mathcal{P}$. A simple change of variables shows that

$$E_\pi\Big[\phi\Big(E_\pi[u(f)|\mathcal{T}]\Big)\Big] = \int_\Omega \phi\left(\int_\Omega u(f(\omega'))k(\omega,\, \mathrm{d}\omega')\right) \mathrm{d}\pi(\omega) = \int_{\mathcal{P}} \phi\left(\int_\Omega u(f)\,\mathrm{d}p\right) \mathrm{d}\mu(p).$$

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

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

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

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

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

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

Hence $\mu(A^\star) + \mu((A^c)^\star) = 1$, and in particular $\mu(A^\star) = \pi(A)$. Lemma 25 shows $\Sigma_{\mathcal{P}} = \Sigma_{\mathcal{P}}^\star$. Thus there exists $B \in \mathcal{T}_{\mathcal{P}}$ such that $A^\star = B^\star$, and hence $(A^\star)^c = (B^\star)^c = (B^c)^\star$. Then

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

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

For a preference relation $\succsim$ that admits a predictive representation $(u, \phi, \mathcal{T}, \pi)$, an event $A \in \mathcal{S}$ is null if and only if $\pi(A) = 0$ (Lemma 17). Thus Proposition 1 follows from Lemmas 26 and 27, given that $\sigma(k)$ and $\mathcal{T}_{\mathcal{P}}$ are π_μ-equivalent (Lemma 24).