Dynamic Consistency and Rectangularity for the Smooth Ambiguity Model*

Andrei Savochkin[†] Alexander Shklyaev[‡] Alexey Galatenko[§]

December 11, 2023

Abstract

We study the Smooth Ambiguity decision criterion in the dynamic setting to understand when it can satisfy the Dynamic Consistency and Consequentialism properties. These properties allow one to rewrite the decision criterion recursively and solve for optimal decisions by Dynamic Programming. Our result characterizes the possibility of having these properties through a condition that resembles Epstein and Schneider's (2003) rectangularity condition for the maxmin model. At the same time, we show that Dynamic Consistency and Consequentialism can be achieved for Smooth Ambiguity preferences in a narrower set of scenarios than one would hope for.

1 Introduction

1.1 Motivation and our contribution

In the past decades of research on decisions under uncertainty, the most progress has been associated with the idea of ambiguity.[1] The majority of the theoretical literature studies ambiguity and decision criteria in the static setting, assuming that uncertainty resolves at once, without any provisions for partial information. By contrast, a large fraction of applied research, by necessity, works with choices that are sequential, and in environments in which some information arrives in every period.

In the dynamic setting (with or without ambiguity), there are two natural properties, Dynamic Consistency and Consequentialism, that a modeler frequently wants to impose on

*We thank Pierpaolo Battigalli, Enrico Diecidue, David Dillenberger, Adam Dominiak, Ani Guerdjikova, Massimo Marinacci, Stefania Minardi, Camillo Padoa-Schioppa, Lorenzo Stanca, Peter Wakker, seminar audiences and the audience of D-TEA 2023 for helpful comments. Declarations of interest: none.

†Bocconi University, andrei.savochkin@unibocconi.it

‡Moscow State University, alexander.shklyaev@math.msu.ru

§Moscow State University, agalat@msu.ru

[1] The latter refers to situations in which, due to scarce information or her perception of the environment, a decision maker is unable to settle on a single probability distribution that describes the uncertainty that she is facing.

the agent's decision criterion. We will discuss these properties later in the paper, but, at the intuitive level, they indeed look quite appealing: Dynamic Consistency, in particular, implies that the decision maker is willing to carry out a plan of action that was optimal at the ex ante stage. This rules out various traits from the realm of Behavioral Economics, such as issues with self-control, unstable preferences, and many others. In turn, Consequentialism postulates the irrelevance of the decision maker's payoffs that would obtain under contingencies that are no longer possible. Beyond that, Dynamic Consistency and Consequentialism are also highly desirable from a purely practical perspective: for instance, they are fundamental in any setup that deals with the value of information and are indispensable for rewriting and solving models through Dynamic Programming.

As puzzling as it may be, modeling ambiguity-sensitive choices in the dynamic setting typically runs into difficulties with satisfying Dynamic Consistency and Consequentialism. Gilboa and Schmeidler (1993) gave an early indication that satisfying those properties may not be so easy. Later, Al-Najjar and Weinstein (2009) argued that the difficulties in working with ambiguity in the dynamic setting seem to have a fundamental nature. For the model that is most frequently used in both theory and applications—the maxmin model of Gilboa and Schmeidler (1989)—there is a partial solution for the conundrum. To have a positive result, Epstein and Schneider (2003) made an additional framework assumption that the timing and stucture of information disclosure is fixed—that is, the agent knows in advance the type of information that will arrive in each future time period. Then, they obtained a result that shows that the maxmin model behaves well in the dynamic setting and satisfies the desirable properties if and only if the model uses so-called *rectangular sets of probabilities*.[2] However, it has not been known whether a conceptually similar solution exists for the second most popular model of choice under ambiguity, the Smooth Ambiguity model of Klibanoff, Marinacci and Mukerji (2005).[3] In this paper, we seek to fill this gap.

We study properties of the Smooth Ambiguity model in the dynamic setting with the goal of understanding when Dynamic Consistency and Consequentialism can hold simultaneously with respect to a fixed two-period structure of information disclosure. Our main contribution is a theorem that provides "if and only if" conditions for that. It demonstrates that, in fact, Dynamic Consistency and Consequentialism hold in a set of scenarios that are narrower than one would have hoped for. First, they hold trivially if there is no ambiguity in the first or the second period. Besides those trivial cases, we show that they can hold *only* if preferences have the constant absolute ambiguity aversion property. Under constant absolute ambiguity aversion, the theorem states a condition for the decision maker's beliefs that delivers Dynamic Consistency and Consequentialism. Although our condition for the Smooth Ambiguity model is much more complicated than Epstein and Schneider's (2003) rectangularity condition for the maxmin model, there is still a conceptual similarity between them. Besides theoretical interest, our condition gives practical guidance to writing tractable models that use Smooth Ambiguity preferences in a dynamic environment.

The rest of the paper is organized as follows. In order to better define our research question, we continue this Introduction with an informal presentation of the Smooth Ambiguity

[2]The precursor of rectangular sets of probabilities was introduced by Sarin and Wakker (1998) under the name *reduced families of probability measures*.

[3]See the discussion of the literature in Section 3.

Date: 2023-12-11 14:40:02 Revision: e8dda11

model and the properties of Dynamic Consistency and Consequentialism. In Section 2, we formally set the stage, state our result, and illustrate it with a concrete example of Smooth Ambiguity preferences that have the desired properties. Section 3 discusses what can be learned from our result, alternative approaches, and related literature.

1.2 Informal presentation of the problem

Uncertainty To give the gist of the difficulties surrounding ambiguity in the dynamic setting and to outline our results, let us assume that uncertainty is represented by the state space Ω in which each element represents one possible resolution of the uncertainty. For the purpose of making decisions, each available action can be described by a mapping from states into outcomes from some set of outcomes X, where the latter can represent monetary outcomes, levels of the decision maker's satisfaction, social outcomes, or anything else. Without losing much of the content, we can also assume that the decision maker chooses among real-valued random variables defined on this state space. Her choice is assumed to be captured by a preference relation—a reflexive, transitive, and complete binary relation that compares mappings from Ω to X.

Dynamic Consistency and Consequentialism Now, we introduce the basic principles of Dynamic Consistency and Consequentialism that are essential for describing the objectives of our research and our results. Suppose that the state space Ω is partitioned into two events, E_1 and E_2 and, to simplify the illustration, assume that these events, as well as the entire Ω, are finite: $E_1 = \{\omega_1, \ldots, \omega_k\}$ and $E_2 = \{\omega_{k+1}, \ldots, \omega_n\}$. We are interested in the relationship between the decision maker's choices at two stages—the ex ante stage at which no information is revealed to her yet, and the interim stage at which she has been informed whether the true state belongs to E_1 or E_2. Consider the situation in which the decision maker's choice matters only if E_1 occurs; in the case of E_2, she obtains outcomes $h(\omega_{k+1}), \ldots, h(\omega_n)$ in states $\omega_{k+1}, \ldots, \omega_n$ regardless of her choice. If we think of the decision maker and Nature making moves sequentially, one can compare the two decision trees shown on Figure 1 that differ in the order of moves.

Dynamic Consistency postulates that, in the first tree, the decision maker should realize that her choice between the left branch and the right branch matters only if E_1 occurs. Hence, in the first tree, her choice should be the same as her choice in the second tree after she has learned that E_1 has occurred: she should exhibit strict preference for the left branch in the first tree if and only if she exhibits strict preference for the left branch in the second tree; she should exhibit strict preference for the right branch in the first tree if and only if she exhibits strict preference for the right branch in the second tree; and she should exhibit indifference between the two branches in the first tree if and only if she exhibits indifference in the second tree.

Next, consider a different version of the second tree in which the vector of outcomes $(h(\omega_{k+1}), \ldots, h(\omega_n))$ for states $\omega_{k+1}, \ldots, \omega_n$ is replaced with some other vector of outcomes $(\tilde{h}(\omega_{k+1}), \ldots, \tilde{h}(\omega_n))$, keeping the rest of the tree unchanged. Consequentialism postulates that the decision maker's choice between the left and the right branches should be the same for the two versions of the second tree—by the time the decision maker is given an

Date: 2023-12-11 14:40:02 Revision: e8dda11

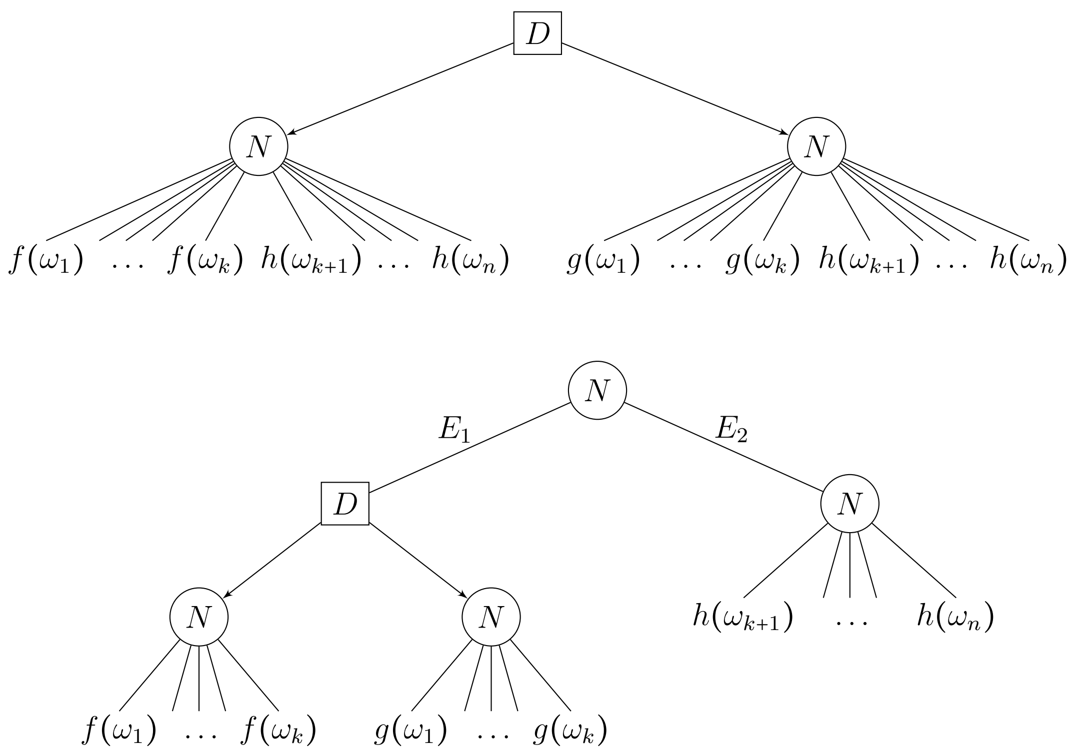

Figure 1: Decision trees with ex ante and conditional decisions

The nodes are marked by D and N to reflect whether it is the decision maker or Nature, respectively, that makes a move at this node.

opportunity to choose, the payoffs in the states that belong to E_2 should become irrelevant ("inconsequential").

More generally, a decision maker is said to satisfy Dynamic Consistency and Consequentialism with respect to a partitioning (E_1, E_2) of Ω if, for any two mappings f and g from Ω to X that coincide on E_2, her preference between them at the ex ante stage are the same as her preference conditional on learning that E_1 has occurred, and is independent of the common payoff vector in the states that belong to E_2; and similarly with E_1 and E_2 interchanged.

As has been argued in the literature, the principles of Dynamic Consistency and Consequentialism are quite reasonable normatively, as well as desirable from the applied perspective. For instance, violations of Dynamic Consistency imply negative value of information—not only does the decision maker refrain from acquiring information that is relevant for her decisions, but she must also be willing to pay a positive amount of money in order not to get informed. Violations of Consequentialism are also unappealing. They look as if the decision maker is influenced by immaterial details of the presentation of decision problems and, thus, falls victim to the "framing effect." Applied analysis frequently relies on Dynamic Consistency and Consequentialism implicitly because these properties always hold if the environment presumes no ambiguity, and decisions follow the canonical expected utility criterion—a mapping f is preferred over g whenever

$$\int_\Omega u(f(\omega))\,\mu(d\omega) \geq \int_\Omega u(g(\omega))\,\mu(d\omega) \tag{1}$$

(for some fixed probability measure μ on Ω and utility index $u : X \to \mathbb{R}$), with conditional probabilities $\mu(\cdot \mid E_i)$ replacing μ for decisions made after learning an event E_i.

Restrictions on preferences imposed by Dynamic Consistency and Consequentialism As mentioned earlier, Dynamic Consistency and Consequentialism present surprising difficulties for modeling ambiguity in settings with gradual disclosure of information. Without going into a general discussion, we are interested in specific restrictions that these two principles imply for one of the most frequently used models of ambiguity-sensitive preferences—the Smooth Ambiguity model.

Epstein and Schneider's (2003) seminal paper solved a similar problem for a different model of preferences—the maxmin model of Gilboa and Schmeidler (1989). In that model, a mapping f is preferred over g whenever it yields a higher utility value computed as

$$V(f) = \min_{\mu \in M} \int_\Omega u(f(\omega))\,\mu(d\omega) \tag{2}$$

for some fixed convex and compact set M of probability measures and some utility index u. The key insight from Epstein and Schneider's (2003) analysis goes as follows. Suppose that the decision maker's ex ante preferences $\succsim$ have the maxmin form with the utility index u and a set M of probability measures on Ω; her preferences $\succsim_1$ conditional on the event E_1 have the maxmin form with the utility index u and a set M_1 of probability measures on E_1; and her preferences $\succsim_2$ conditional on the event E_2 have the maxmin form with the utility index u and a set M_2 of probability measures on E_2. Then, her system of preferences satisfies

Dynamic Consistency and Consequentialism with respect to the partitioning (E_1, E_2) if and only if the set M has a rectangular form:

$$M = \left\{\alpha\mu_1 + (1-\alpha)\mu_2 \mid \alpha \in [\underline{\alpha}, \overline{\alpha}], \mu_1 \in M_1, \mu_2 \in M_2\right\}.$$

In words, the set of probability measures M for the ex ante preference relation is obtained by combining three components *independently*: probability measures conditional on E_1; probability measures conditional on E_2; and the weight that is given to E_1 versus E_2.[4] Clearly, this structure of the set of probability measures is very particular, and many sets do not have the rectangular form. Nevertheless, this result gives a practical answer to the question about what the modeler can do to achieve Dynamic Consistency and Consequentialism.

The other very popular model of ambiguity-sensitive preferences is the Smooth Ambiguity model of Klibanoff, Marinacci and Mukerji (2005). In that model, a mapping f is evaluated as

$$V(f) = \int_\Delta \varphi\left(\int_\Omega u(f(\omega))\,\mu(d\omega)\right)\pi(d\mu),$$

where Δ is the set of probability measures on Ω; π is a probability measure defined on Δ (the second-order probability measure); $u : X \to \mathbb{R}$ is a utility index; and $\varphi : \mathbb{R} \to \mathbb{R}$ is a second-order utility. If the function φ is assumed to be smooth, then, unlike in the maxmin model, this model's evaluation functional is smooth, which is often useful in applications.

With this background, our research question can be formulated as follows: under what conditions on the parameters π and φ do smooth ambiguity preferences satisfy Dynamic Consistency and Consequentialism?

2 Formal analysis

2.1 Setup

Let Ω be the state space that we assume to be finite, and let Δ be the space of probability measures on $(\Omega, 2^\Omega)$. We identify Δ with a subset of $\mathbb{R}^{|\Omega|}$ and endow it with the Euclidean metric and the Borel σ-algebra.

For a set Z, let $\mathcal{F}(Z)$ denote the set of functions from Ω to Z; traditionally, they are called *acts*. Our objects of interest are preference relations—reflexive, transitive, and complete binary relations—over acts. For any preference relation $\succsim$, its asymmetric part is denoted by $\succ$ and its symmetric part is denoted by $\sim$.

A functional $V : \mathcal{F}(Z) \to \mathbb{R}$ is said to *represent* a preference relation $\succsim$ if $f \succsim g$ holds if and only if $V(f) \geq V(g)$ for any $f, g \in \mathcal{F}(Z)$. We are interested in preference relations that belong to the class of the Smooth Ambiguity preferences; that is, they are such that there exist a nonconstant function $u : Z \to \mathbb{R}$, a continuous and weakly increasing function

[4] For practical purposes, it is also useful to know that, in the rectangular case, it is sufficient to know only the ex ante set of probability measures M; and the sets of probability measures M_1 and M_2 for conditional preferences can be computed as $M_1 = \{\mu(\cdot \mid E_1) \mid \mu \in M\}$ and $M_2 = \{\mu(\cdot \mid E_2) \mid \mu \in M\}$, while $[\underline{\alpha}, \overline{\alpha}] = \{\mu(E_1) \mid \mu \in M\}$.

$\varphi : \operatorname{range} u \to \mathbb{R}$, and a probability measure π on Δ such that the preference relation admits a representation by the functional

$$V(f) = \int_\Delta \varphi\left(\int_\Omega u(f(\omega))\,\mu(d\omega)\right)\pi(d\mu). \tag{3}$$

Regarding the function u, it is common to assume that u is convex-ranged. The outcome space Z and the function u do not play any substantial role in our results. Thus, to simplify the exposition, we let $I := \operatorname{range} u$ and work with preferences defined on $\mathcal{F}(I)$ that admit a representation by

$$V(f) = \int_\Delta \varphi\left(\int_\Omega f(\omega)\,\mu(d\omega)\right)\pi(d\mu). \tag{4}$$

Next, we want to formally introduce the properties of Dynamic Consistency and Consequentialism. Following the convention, for any $f, h \in \mathcal{F}(I)$ and any $E \subseteq \Omega$, we write $f\,E\,h$ for an element of $\mathcal{F}(I)$ such that $(f\,E\,h)(\omega) = f(\omega)$ if $\omega \in E$ and $(f\,E\,h)(\omega) = h(\omega)$ if $\omega \notin E$. We say that a state $\omega \in \Omega$ is *null* if, for any two functions $f, h \in F(I)$, $f\,\{\omega\}\,h \sim h$; otherwise, the state ω is *non-null.* Similarly, an event $E \subseteq \Omega$ is null if, for any two functions $f, h \in F(I)$, $f\,E\,h \sim h$; otherwise, it is non-null.

Suppose that Ω is partitioned as $\Omega = E_1 \sqcup E_2$, where E_1 and E_2 are $\succsim$-non-null. We say that $\succsim$ satisfies *Dynamic Consistency and Consequentialism* with respect to $\{E_1, E_2\}$ if

$$f\,E_i\,h \succsim g\,E_i\,h \quad \Leftrightarrow \quad f\,E_i\,h' \succsim g\,E_i\,h' \qquad \text{for all } f, g, h, h' \in \mathcal{F}(I) \text{ and } i = 1, 2.$$

As elaborated in the Introduction, we are interested in two properties: 1) consistency between ex ante decisions and decisions conditional on events E_1 and E_2, and 2) independence of conditional decisions from payoffs in unrealized states. The above definition imposes both of them jointly, and also allows us to state our assumptions succinctly by bypassing notation for conditional preferences.

We note that the combination of these two properties is related to the Savagean Sure Thing Principle (Axiom P2): if Dynamic Consistency and Consequentialism hold universally—for all possible partitionings—then the Sure Thing Principle obtains. The Sure Thing Principle is known to be very restrictive; by contrast, we are interested in the implications of Dynamic Consistency and Consequentialism in application to one given partitioning.

2.2 The result

Let $\succsim$ be a preference relation on $\mathcal{F}(I)$ for some nondegenerate open interval $I \subseteq \mathbb{R}$. Fix a partitioning $\{E_1, E_2\}$ of Ω and, for $i = 1, 2$, let Δ_i be the space of probability measures on E_i, Σ_i be the Borel σ-algebra on Δ_i, and $\mathcal{F}_i(I)$ be the set of functions from E_i to I.

We will impose the following condition on the probability measure π on (Δ, Σ).

Condition 1. $\int_\Delta \mu(E_1)^p \mathbb{1}_{\mu(E_1)>0}\,\pi(d\mu) < \infty$ *and* $\int_\Delta \mu(E_2)^p \mathbb{1}_{\mu(E_2)>0}\,\pi(d\mu) < \infty$ *for all* $p \in \mathbb{R}$.

Our main result can be stated as follows.

Date: 2023-12-11 14:40:02 Revision: e8dda11

Theorem 1. *Let $\succsim$ admit Representation (4) for some Borel probability measure π on Δ.*

Suppose that $\varphi : I \to \mathbb{R}$ is three times continuously differentiable with $\varphi' > 0$ and $\varphi'' < 0$. Suppose, also, that E_1 and E_2 each has at least two non-null states, and π satisfies Condition 1. If $\succsim$ jointly satisfies Dynamic Consistency and Consequentialism with respect to $\{E_1, E_2\}$, then one of the following must hold.

(i) There exist $\alpha \in (0,1)$ and probability measures π_1 and π_2 on (Δ_1, Σ_1) and (Δ_2, Σ_2), respectively, such that $\succsim$ admits a representation via

$$V(f) = \alpha \int_{\Delta_1} \varphi\left(\int_{E_1} f(\omega)\, \mu_1(d\omega) \right) \pi_1(d\mu_1) + (1-\alpha) \int_{\Delta_2} \varphi\left(\int_{E_2} f(\omega)\, \mu_2(d\omega) \right) \pi_2(d\mu_2).$$

(ii) There exist a Borel probability measure $\tilde{\pi}$ on $[0,1]$ and probability measures $\mu_1 \in \Delta_1$ and $\mu_2 \in \Delta_2$ such that $\succsim$ admits a representation via

$$V(f) = \int_{[0,1]} \varphi\left(\alpha \int_{E_1} f(\omega)\, \mu_1(d\omega) + (1-\alpha) \int_{E_2} f(\omega)\, \mu_2(d\omega) \right) \tilde{\pi}(d\alpha).$$

(iii) There exist $\gamma > 0$, $k > 0$, and $c \in \mathbb{R}$ such that $\varphi(z) = -ke^{-\gamma z} + c$, and $\succsim$ admits a representation via

$$V(f) = -\int_{[0,1]} v_1(f, \alpha)\, v_2(f, 1-\alpha)\, \tilde{\pi}(d\alpha), \tag{5}$$

where $\tilde{\pi}$ is a Borel probability measure on $[0,1]$ and, for $i = 1,2$, $v_i : \mathcal{F}(I) \times [0,1] \to \mathbb{R}$ are defined as $v_i(f, 0) = 1$ and

$$v_i(f, \alpha) = \int_{\Delta_i} e^{-\gamma\alpha \int_{E_i} f(\omega)\, \mu(d\omega)}\, \pi_{i,\alpha}(d\mu) \quad \text{for } \alpha \in (0,1], \tag{6}$$

where $\pi_{i,\alpha}$ for $\alpha \in (0,1]$ are probability measures on (Δ_i, Σ_i) derived as conditionals,

$$\pi_{i,\alpha}(S) = \pi\big(\{\mu \in \Delta : \mu(\cdot \mid E_i) \in S\} \,\big|\, \mu(E_i) = \alpha \big) \quad \text{for any } S \in \Sigma_i;$$

moreover, it must be that

$$v_i(f, \alpha) \geq v_i(g, \alpha) \Leftrightarrow v_i(f, \alpha') \geq v_i(g, \alpha') \qquad \text{for all } f, g \in \mathcal{F}(I) \tag{7}$$

holds for $\tilde{\pi}$-almost all $\alpha, \alpha' \in (0,1]$.

Conversely, suppose that φ is strictly increasing. Suppose that $\succsim$ admits the representation of Case (i) for some $\alpha \in (0,1)$ and some probability measures π_1 and π_2 on (Δ_1, Σ_1) and (Δ_2, Σ_2), respectively; or admits the representation of Case (ii) for some Borel probability measure $\tilde{\pi}$ on $[0,1]$ such that $\tilde{\pi}(\{0,1\}) < 1$ and probability measures μ_1 and μ_2 from Δ_1 and Δ_2, respectively; or admits the representation of Case (iii) for some $\gamma > 0$, probability measure $\tilde{\pi}$ on $[0,1]$ such that $\tilde{\pi}(\{0,1\}) < 1$, and, for $i = 1,2$ and $\alpha \in (0,1]$, probability measures $\pi_{i,\alpha}$ on (Δ_i, Σ_i) such that, for v_i defined by (6) and $v_i(f, 0) = 1$, equivalence (7) holds for $\tilde{\pi}$-almost all $\alpha, \alpha' \in (0,1]$. Then, $\succsim$ satisfies Dynamic Consistency and Consequentialism.

Cases (i) and (ii) cover the situations of degenerate ambiguity. In Case (i), there is no ambiguity about the probabilities of events E_1 and E_2. Indeed, for any $x, y \in I$, act $f = x\, E_1\, y$ that is constant on E_1 and E_2 is valued as $V(x\, E_1\, y) = \alpha\varphi(x) + (1-\alpha)\varphi(y)$, which is an Expected Utility evaluation (1), in which event E_1 has the probability of α. In Case (ii), there is no ambiguity after the decision maker learns E_1 or E_2. Indeed, for any two acts $f\, E_1\, h$ and $g\, E_1\, h$,

$$V(f\, E_1\, h) \geq V(g\, E_1\, h) \quad \Leftrightarrow \quad \int_{E_1} f(\omega)\, \mu_1(d\omega) \geq \int_{E_1} g(\omega)\, \mu_1(d\omega)$$

because the function

$$z \mapsto \int_{[0,1]} \varphi\Big(\alpha z + (1-\alpha) \int_{E_2} f(\omega)\, \mu_2(d\omega)\Big)\, \tilde{\pi}(d\alpha)$$

is strictly increasing in z (under our assumption that E_1 is non-null). Therefore, after learning E_1, the decision maker compares acts by the expected utility criterion. The same reasoning applies to E_2.

Case (iii) is the main case of the theorem. Representation (5) is still a Smooth Ambiguity representation—it can be seen by plugging the expressions for v_1 and v_2 there, collecting the integrals on the outside, and observing that a product of exponentials is the exponential of a sum. However, in the right-hand side of (5), we see a certain separation of the evaluation of f on E_1 and on E_2.[5] What is important is that Case (iii) has the *independent combination* feature due to Condition (7): the evaluation of f on E_1 via integrals with respect to various $\mu_1 \in \Delta_1$ and the evaluation of f on E_2 via integrals with respect to various $\mu_2 \in \Delta_2$ are combined when the weight α varies, but the ranking of acts f by $v_1(f, \alpha)$, as well as by $v_2(f, 1-\alpha)$, are *independent* of the value of α. From this perspective, our Case (iii) resembles the rectangularity condition of Epstein and Schneider (2003) that we recalled in the Introduction.

We illustrate the substance of Case (iii) with the following example.

Example 1. *Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $E_1 = \{\omega_1, \omega_2\}$, $E_2 = \{\omega_3, \omega_4\}$, and $I = \mathbb{R}$. To specify the representation of a preference relation, suppose that an act f takes values x_1, x_2, x_3, x_4 in states $\omega_1, \omega_2, \omega_3, \omega_4$, respectively, and let*

$$\begin{aligned} V(f) = &-\frac{1}{16} e^{-\frac{1}{9}x_1 - \frac{2}{9}x_2 - \frac{2}{9}x_3 - \frac{4}{9}x_4} - \frac{1}{16} e^{-\frac{2}{9}x_1 - \frac{1}{9}x_2 - \frac{2}{9}x_3 - \frac{4}{9}x_4} - \frac{1}{8} e^{-\frac{1}{9}x_1 - \frac{2}{9}x_2 - \frac{1}{3}x_3 - \frac{1}{3}x_4} - \frac{1}{8} e^{-\frac{2}{9}x_1 - \frac{1}{9}x_2 - \frac{1}{3}x_3 - \frac{1}{3}x_4} \\ &- \frac{1}{16} e^{-\frac{1}{9}x_1 - \frac{2}{9}x_2 - \frac{4}{9}x_3 - \frac{2}{9}x_4} - \frac{1}{16} e^{-\frac{2}{9}x_1 - \frac{1}{9}x_2 - \frac{4}{9}x_3 - \frac{2}{9}x_4} - \frac{1}{16} e^{-\frac{2}{9}x_1 - \frac{4}{9}x_2 - \frac{1}{9}x_3 - \frac{2}{9}x_4} - \frac{1}{8} e^{-\frac{1}{3}x_1 - \frac{1}{3}x_2 - \frac{1}{9}x_3 - \frac{2}{9}x_4} \\ &- \frac{1}{16} e^{-\frac{4}{9}x_1 - \frac{2}{9}x_2 - \frac{1}{9}x_3 - \frac{2}{9}x_4} - \frac{1}{16} e^{-\frac{2}{9}x_1 - \frac{4}{9}x_2 - \frac{2}{9}x_3 - \frac{1}{9}x_4} - \frac{1}{8} e^{-\frac{1}{3}x_1 - \frac{1}{3}x_2 - \frac{2}{9}x_3 - \frac{1}{9}x_4} - \frac{1}{16} e^{-\frac{4}{9}x_1 - \frac{2}{9}x_2 - \frac{2}{9}x_3 - \frac{1}{9}x_4} \end{aligned}$$

Observe that this functional captures a nondegenerate case: the probability of event E_1 can be equal to $\frac{1}{3}$ (and this happens with probability $\frac{1}{2}$), and can also be equal to $\frac{2}{3}$ (and this

[5] In fact, the mappings $f \mapsto (-v_1(f, \alpha))$ and $f \mapsto (-v_2(f, \alpha))$ can be thought of as representing the decision maker's preferences conditional on E_1 and E_2, respectively.

Date: 2023-12-11 14:40:02 Revision: e8dda11

also happens with probability $\frac{1}{2}$). For this preference relation, the functions $v_1(\cdot, \alpha)$ from the statement of the theorem can be computed as

$$v_1(f, \tfrac{1}{3}) = \frac{1}{2} e^{-\frac{2}{9}x_1 - \frac{1}{9}x_2} + \frac{1}{2} e^{-\frac{1}{9}x_1 - \frac{2}{9}x_2},$$
$$v_1(f, \tfrac{2}{3}) = \frac{1}{4} e^{-\frac{4}{9}x_1 - \frac{2}{9}x_2} + \frac{1}{2} e^{-\frac{1}{3}x_1 - \frac{1}{3}x_2} + \frac{1}{4} e^{-\frac{2}{9}x_1 - \frac{4}{9}x_2}.$$

They are distinct functions on $\mathbb{R}^4$; however, they rank vectors identically. Indeed, it can be seen that $v_1(f, \frac{2}{3}) = \big(v_1(f, \frac{1}{3})\big)^2$ for all $f \in \mathbb{R}^4$.

3 Discussion of the result and related literature

Issues related to ambiguity in a dynamic setting

When ambiguity is brought to the multi-period setting, one of the main topics in the literature has been the question of updating—that is, a procedure for transforming an ex ante preference relation of a particular form into a conditional preference relation that takes into account the fact that some event has occurred. As Gilboa and Schmeidler (1993) and Eichberger and Kelsey (1996) noted quite early, proposing a reasonable updating procedure for popular models of ambiguity-sensitive preferences (namely, maxmin and Choquet Expected Utility) is a very challenging task.

Al-Najjar and Weinstein (2009) discuss at a more general level the challenges that the dynamic setting presents to modeling ambiguity. They argue that there seems to be a genuine tension between the entire concept of ambiguity and a number of natural requirements for dynamic choice. This provocative paper prompted a thoughtful discussion in the literature. Siniscalchi (2009) provides a succinct summary:

> [T]he modeller interested in ambiguity can essentially pick *any two out of three ingredients*: full generality in the representation of ambiguity attitudes, Consequentialism, or Dynamic Consistency. She *cannot* have all three (however, she can have a little bit of each).

It should also be noted that Dynamic Consistency without Consequentialism can always be achieved.

Gumen and Savochkin (2013) focus on a different aspect of updating: a modeler typically imposes, implicitly or explicitly, another requirement beyond those mentioned above—the requirement that updated preferences have the same "form." For example, if the ex ante preferences are assumed to have the maxmin form (2), modelers wish the conditional preferences also to be maxmin, and similarly for other types of preferences (variational preferences, confidence function preferences, multiplier preferences, and so on). This requirement further constraints the modeler, and these constraints sometimes can be quite strong. Continuing with the maxmin example, if one takes such a preference relation and derives conditional preferences that satisfy Dynamic Consistency, then these preferences will not only typically fail Consequentialism, but will also fail to be maxmin. Gumen and Savochkin's (2013) general result shows the boundaries of what can be achieved for preferences that have certain invariance properties. For maxmin and variational preferences, this invariance emerges from weaker versions of the Independence axiom that define those classes of preferences. That result is not applicable to the Smooth Ambiguity model because it is not defined by postulating a weaker version of the Independence axiom.

In an earlier paper, Hanany and Klibanoff (2007) propose a way to address many of the difficulties mentioned above, including the difficulty of maintaining the same form of preferences (what they call "closure"). Their solution is to revise what we require from preferences. In particular, they not only discard Consequentialism,[6] but also substantially weaken the Dynamic Consistency requirement by focusing only on the best act in the set of feasible ones. To illustrate the gist of their idea, fix a decision maker and consider a situation in which she is asked ex ante to rank three acts: $r\,E_1\,h$; $f\,E_1\,h$; and $g\,E_1\,h$. Suppose, furthermore, that she declares that $r\,E_1\,h$ is preferred to $f\,E_1\,h$; $f\,E_1\,h$ is preferred to $g\,E_1\,h$; and, by transitivity, $r\,E_1\,h$ is preferred to $g\,E_1\,h$. Hanany and Klibanoff's (2007) model requires that preferences conditional on the event E_1 only respect the ranking of $r\,E_1\,h$ above $f\,E_1\,h$ and the ranking of $r\,E_1\,h$ above $g\,E_1\,h$, while the conditional ranking between $f\,E_1\,h$ and $g\,E_1\,h$ is not restricted—according to the model, the decision maker may declare that the conditional ranking of $g\,E_1\,h$ is higher than that of $f\,E_1\,h$. By allowing this, Hanany and Klibanoff (2007) achieve a lot in resolving various modeling difficulties. Starting with an ex ante preference relation that is maxmin, they provide a way to have conditional preferences that are maxmin as well, and relate probability measures that are used to compute the ex ante and conditional values of acts. Subsequently, Hanany and Klibanoff (2009) generalize the idea beyond the maxmin preferences to cover the Smooth Ambiguity model.

The literature discussed above is concerned mainly with the universal concept of updating preferences to account for arbitrary events about which the decision maker may get informed. As mentioned in the introduction, Epstein and Schneider (2003) pursue a practical route of studying updating given a fixed timing and structure of the arriving information. In the context of the maxmin model, they derive conditions that ensure Dynamic Consistency and Consequentialism and that the maxmin form is preserved for conditional preferences. Subsequently, this approach has been applied to other models of preferences (e.g., to variational by Maccheroni, Marinacci and Rustichini, 2006). In the present paper, we also follow Epstein and Schneider's (2003) route and contribute to the literature by studying the conditions under which Dynamic Consistency and Consequentialism can hold for the Smooth Ambiguity model.

Smooth Ambiguity preferences and "intuitive" updating

One intuitively appealing way to update a Smooth Ambiguity preference relation is to replace each probability distribution in the support of π with the corresponding conditional probability. Put differently, if the decision maker's ex ante preferences have the Smooth Ambiguity representation (3) with some utility index u, second-order utility φ, and a second-order

[6]Their solution does not satisfy Consequentialism because the preference relation conditional on learning an event E_1 is allowed to depend on the outcomes outside of E_1 that were feasible at the ex ante stage.

probability measure π, then one can assume that, conditional on learning an event E_1, the decision maker compares acts by using the evaluation functional $\hat{V}$ defined as

$$\hat{V}(f) = \int_{\Delta_1} \varphi\Big(\int_\Omega u(f(\omega))\,\hat{\mu}(d\omega)\Big)\,\hat{\pi}(d\hat{\mu}), \tag{8}$$

where the probability measure $\hat{\pi}$ on Δ_1 is defined as

$$\hat{\pi}(\hat{\mu}) = \pi\big(\{\mu \in \Delta : \mu(\cdot \mid E_1) = \hat{\mu}\}\big). \tag{9}$$

As our Theorem 1 suggests, this way of updating preferences generally leads to dynamic inconsistencies, in the sense that the equivalence

$$V(f\,E_1\,h) \geq V(g\,E_1\,h) \quad \Leftrightarrow \quad \hat{V}(f\,E_1\,h) \geq \hat{V}(g\,E_1\,h) \qquad \text{for all } f,g,h \in \mathcal{F}(Z)$$

will typically fail, unless the ex ante preference relation falls into the first two cases of the theorem.

To illustrate this failure, we give an example that is built upon the well-known three-color version of Ellsberg's (1961) paradox.

Example 2. *Suppose that the state space is $\Omega = \{R,B,Y\}$, where the states represent possible colors of a ball drawn from an urn, and the decision maker's π has the following form: when the space Δ is thought of as a subset of $\mathbb{R}^3$, let the support of π be the linear segment connecting $(1/3,2/3,0)$ and $(1/3,0,2/3)$, and let π be the uniform measure on that segment. For concreteness, assume that $Z = \mathbb{R}$, $u(x) \equiv x$, and $\varphi(x) = \ln(1+x)$. Then, the decision maker evaluates any act f by*

$$V(f) = \int_0^1 \ln\Big(1 + \frac{1}{3}f(R) + \frac{2}{3}pf(B) + \frac{2}{3}(1-p)f(Y)\Big)\,dp.$$

In turn, the updating procedure (8)–(9) conditional on the event $E_1 = \{R,B\}$ leads to evaluation of an act f by

$$\hat{V}(f) = \int_0^1 \ln\Big(1 + \frac{1}{1+2p}f(R) + \frac{2p}{1+2p}pf(B)\Big)\,dp.$$

Consider acts f, g, and h that are "bets" that pay \$100 in events $\{B,Y\}$, $\{R,Y\}$, and $\{Y\}$, respectively, and nothing otherwise. That is, written as vectors from $\mathbb{R}^3$, $f = (0,100,100)$, $g = (100,0,100)$, and $h = (0,0,100)$.[7] *Then, it can be computed that $V(f\,E_1\,h) \approx 4.21 > 4.17 \approx V(g\,E_1\,h)$ but $\hat{V}(f\,E_1\,h) \approx 3.69 < 3.98 \approx \hat{V}(g\,E_1\,h)$. This means that if the decision maker compares two acts, $f\,E_1\,h$ and $g\,E_1\,h$, that differ only on E_1, she finds that $f\,E_1\,h$ is better than $g\,E_1\,h$; however, after E_1 has occurred, she reverts her choice and picks $g\,E_1\,h$.*

As is well known, these types of preference reversals have a startling implication for the decision maker's willingness to pay for information. For a standard Bayesian decision maker and in a single player environment, information always has nonnegative value. Dynamic inconsistencies introduced by the updating procedure (8)–(9) change the way information is

[7] In our example, in fact, $f\,E_1\,h = f$ and $g\,E_1\,h = g$.

Date: 2023-12-11 14:40:02 Revision: e8dda11

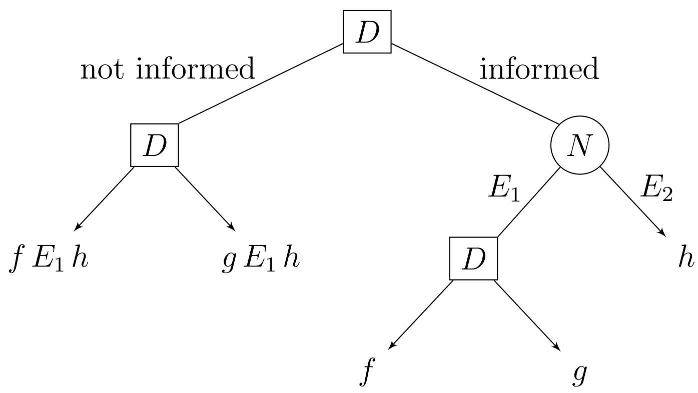

Figure 2: Decision tree with the choice of being informed

To make the picture more compact, we represent the final stage of the resolution of uncertainty by writing just the name of the act instead of drawing a subtree with branches to individual outcomes such as $f(\omega)$ for all $\omega \in E_1$. Choice among decision trees under ambiguity and dynamic inconsistencies is studied in depth by Siniscalchi (2011).

valued. In the above example, the decision maker strictly prefers not to get informed about whether E_1 occurs. The scenario of choosing whether or not to get informed is depicted in Figure 2. When standing at the top node, the decision maker knows what she will choose in each subtree. In the left one, she is uninformed and chooses $f\,E_1\,h$, as calculated earlier. In the right subtree she is informed and chooses $g\,E_1\,h$. At the top node, her preferences over acts are represented by V, and, hence, she prefers to be in the left subtree that leads to her eventually having the (ex ante) better option $f\,E_1\,h$. Moreover, she is willing to pay a small but positive amount of money to be in the left subtree and stay uninformed.

Hanany and Klibanoff's (2009) updating rule

The question of updating in application to the Smooth Ambiguity model is the sole focus of Hanany and Klibanoff (2009)'s (2009) analysis. As mentioned earlier, they follow the general idea of Hanany and Klibanoff (2007) and impose a much weaker requirement on the consistency of ex ante and conditional preferences than the dynamic consistency property that is commonly used in the literature (and is also adopted in this paper). The essence of their consistency condition with respect to a fixed event E_1 is as follows. Let E_2 be E_1's complement in Ω, and suppose that B is a nonempty set of feasible acts, which is assumed to be convex and compact, and $g \in B$. Furthermore, suppose that, for any $f \in B$ such that $f|_{E_2} = g|_{E_2}$, we have $g \succsim f$—that is, among all acts that coincide with g outside of E_1, g is the best. Then, Hanany and Klibanoff allow the conditional preferences after learning E_1 to depend on B and g and denote them by $\succsim_{E_1,g,B}$. Their weaker consistency condition requires $g \succsim_{E_1,g,B} f$ for all $f \in B$, but otherwise does not restrict the way $\succsim_{E_1,g,B}$ ranks all other acts.

Returning to Smooth Ambiguity preferences, suppose that the conditional preference relation is also required to have the following Smooth Ambiguity form

$$V_{E_1}(f) = \int_{\Delta} \varphi\left(\int_{E_1} f(\omega)\, \mu(d\omega \mid E_1) \right) \pi_{E_1}(d\mu),$$

for some probability measure π_{E_1} on Δ, where the second-order utility function φ is the same as in the Smooth Ambiguity representation of the ex ante preferences. Then, Hanany and Klibanoff (2009) show that, with minor technical qualifications, their consistency condition between ex ante and conditional preferences holds if and only if

$$\frac{\int_{\Delta} \varphi'\left(\int_{E_1} g(\omega)\, \mu(d\omega \mid E_1) \right) \mu(\omega^* \mid E_1)\, \pi_{E_1}(d\mu)}{\int_{\Delta} \varphi'\left(\int_{E_1} g(\omega)\, \mu(d\omega \mid E_1) \right) \pi_{E_1}(d\mu)} = \frac{\int_{\Delta} \varphi'\left(\int_{\Omega} g(\omega)\, \mu(d\omega) \right) \mu(\omega^*)\, \pi(d\mu)}{\int_{\Delta} \varphi'\left(\int_{\Omega} g(\omega)\, \mu(d\omega) \right) \mu(E_1)\, \pi(d\mu)} \tag{10}$$

holds for all $\omega^* \in E_1$. Subsequently, they advocate one specific updating rule (i.e., the rule to compute π_{E_1}) that factors in the derivative of φ at the expected value of g. We illustrate the relationship between our result and their updating rule with the following example.

Example 3. *As in Example 1, we let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $E_1 = \{\omega_1, \omega_2\}$, $E_2 = \{\omega_3, \omega_4\}$, and $I = \mathbb{R}$, and enumerate the values of an act f in the four states as x_1, x_2, x_3, x_4. We define a preference relation $\succsim$ through its utility representation as*

$$V(f) = -\tilde{\pi}_0 v_2(f,1) - \tilde{\pi}_1 v_1(f, \tfrac{1}{3})\, v_2(f, \tfrac{2}{3}) - \tilde{\pi}_2 v_1(f, \tfrac{2}{3})\, v_2(f, \tfrac{1}{3}) - \tilde{\pi}_3 v_1(f,1),$$

where $\tilde{\pi}_k \in (0,1)$ for $k = 0, \ldots, 3$ are arbitrary numbers such that $\tilde{\pi}_0 + \tilde{\pi}_1 + \tilde{\pi}_2 + \tilde{\pi}_3 = 1$, and the functions v_1 and v_2 are defined as

$$\begin{aligned}
v_1(f, \tfrac{1}{3}) &= \frac{1}{2} e^{-\frac{2}{9}\gamma x_1 - \frac{1}{9}\gamma x_2} + \frac{1}{2} e^{-\frac{1}{9}\gamma x_1 - \frac{2}{9}\gamma x_2}, \\
v_1(f, \tfrac{2}{3}) &= \frac{1}{4} e^{-\frac{4}{9}\gamma x_1 - \frac{2}{9}\gamma x_2} + \frac{1}{2} e^{-\frac{1}{3}\gamma x_1 - \frac{1}{3}\gamma x_2} + \frac{1}{2} e^{-\frac{2}{9}\gamma x_1 - \frac{4}{9}\gamma x_2}, \\
v_1(f, 1) &= \frac{1}{8} e^{-\frac{2}{3}\gamma x_1 - \frac{1}{3}\gamma x_2} + \frac{3}{8} e^{-\frac{5}{9}\gamma x_1 - \frac{4}{9}\gamma x_2} + \frac{3}{8} e^{-\frac{4}{9}\gamma x_1 - \frac{5}{9}\gamma x_2} + \frac{1}{8} e^{-\frac{1}{3}\gamma x_1 - \frac{2}{3}\gamma x_2}, \\
v_2(f, \tfrac{1}{3}) &= \frac{1}{2} e^{-\frac{2}{9}\gamma x_3 - \frac{1}{9}\gamma x_4} + \frac{1}{2} e^{-\frac{1}{9}\gamma x_3 - \frac{2}{9}\gamma x_4}, \\
v_2(f, \tfrac{2}{3}) &= \frac{1}{4} e^{-\frac{4}{9}\gamma x_3 - \frac{2}{9}\gamma x_4} + \frac{1}{2} e^{-\frac{1}{3}\gamma x_3 - \frac{1}{3}\gamma x_4} + \frac{1}{2} e^{-\frac{2}{9}\gamma x_3 - \frac{4}{9}\gamma x_4}, \\
v_2(f, 1) &= \frac{1}{8} e^{-\frac{2}{3}\gamma x_3 - \frac{1}{3}\gamma x_4} + \frac{3}{8} e^{-\frac{5}{9}\gamma x_3 - \frac{4}{9}\gamma x_4} + \frac{3}{8} e^{-\frac{4}{9}\gamma x_3 - \frac{5}{9}\gamma x_4} + \frac{1}{8} e^{-\frac{1}{3}\gamma x_3 - \frac{2}{3}\gamma x_4},
\end{aligned}$$

where $\gamma > 0$ is a parameter. After substituting v_1 and v_2 into the definition of V and expanding, it can be seen that the preference relation has a Smooth Ambiguity representation with the second-order utility function $\varphi(x) = -e^{-\gamma x}$. As in Example 1, note that $v_1(f, \frac{2}{3}) = \left(v_1(f, \frac{1}{3})\right)^2$ and $v_1(f,1) = \left(v_1(f, \frac{1}{3})\right)^3$, and similar identities hold for v_2. These identities establish that the conditions of Case (iii) of Theorem 1 hold.

Next, we consider a preference relation conditional on E_1 that is consistent (in the sense of Dynamic Consistency and Consequentialism) with the ex ante relation $\succsim$. This conditional preference relation is unique, but it admits multiple Smooth Ambiguity representations. For *instance, we can take* $V_{E_1}(f) = -v_1(f,1)$ *as its representation. As can be seen from the specification of* v_1 *stated above, the second-order probability measure* $\pi_{E_1} \in \Delta_1$ *that arises from the expression for* $(-v_1(\cdot,1))$ *has four measures* $\mu \in \Delta$ *in its support and assigns weights to them as follows:*

$$\begin{aligned} \mu &= (\tfrac{2}{3}, \tfrac{1}{3}, 0, 0) \quad \textit{gets } \tfrac{1}{8}, \\ \mu &= (\tfrac{5}{9}, \tfrac{4}{9}, 0, 0) \quad \textit{gets } \tfrac{3}{8}, \\ \mu &= (\tfrac{4}{9}, \tfrac{5}{9}, 0, 0) \quad \textit{gets } \tfrac{3}{8}, \\ \mu &= (\tfrac{1}{3}, \tfrac{2}{3}, 0, 0) \quad \textit{gets } \tfrac{1}{8}. \end{aligned}$$ [8]

Our assumption of Dynamic Consistency and Consequentialism implies Hanany and Klibanoff's (2007) consistency condition. It can be verified directly that our probability measure π_{E_1} *satisfies their characterizing equation stated above in (10) for any* $\omega^* \in E_1$ *and any act* g. *However, Hanany and Klibanoff's 2009 specific updating rule (the "smooth rule") produces a different conditional preference relation. In their relation, the probability measure* $\pi_{E_1,g}$ *has* 16 *elements in its support and assigns weights to them in a more complicated manner. For instance, if* $g = (a,0,0,0)$ *for some* $a \in \mathbb{R}$, *then their smooth rule assigns to* $\mu = (\frac{1}{9}, \frac{2}{9}, \frac{2}{9}, \frac{4}{9})$ *the weight*

$$\pi_{E_1,g}\big((\tfrac{1}{9}, \tfrac{2}{9}, \tfrac{2}{9}, \tfrac{4}{9})\big) = \frac{e^{\frac{2}{9}\gamma a}\tilde{\pi}_1}{\Big(4e^{\frac{2}{9}\gamma a} + 4e^{\frac{4}{9}\gamma a}\Big)\tilde{\pi}_1 + \Big(4e^{\frac{1}{9}\gamma a} + 8e^{\frac{1}{6}\gamma a} + 4e^{\frac{2}{9}\gamma a}\Big)\tilde{\pi}_2 + 24\tilde{\pi}_3}.$$

Hanany and Klibanoff's (2009) smooth rule provides a way to define conditional preferences that is applicable to any smooth second-order utility function φ *and any ex ante probability measure* π. *However, this universality comes at a cost: the probability measure* $\pi_{E_1,g}$ *in the representation of their conditional preferences depends on the maximal element* g *in the choice set and parameters of the utility function (in the current case,* γ*).*

Recursive preferences

Klibanoff, Marinacci and Mukerji (2009) follow one aspect of Epstein and Schneider's (2003) ideas and bring Smooth Ambiguity to the multi-period setting using a recursive procedure. Assuming two periods and uncertainty that is partially resolved by informing which event of the partition $\{E_1, E_2\}$ has been realized (see, e.g., the second decision tree in Figure 1 on p. 4), the recursive Klibanoff et al.'s (2009) procedure prescribes that acts be evaluated by the functional $V^* : \mathcal{F}(Z) \to \mathbb{R}$ defined as

$$V^*(f) = \int_{[0,1]} \varphi\Big(\alpha\varphi^{-1}\big(V_1^*(f)\big) + (1-\alpha)\varphi^{-1}\big(V_2^*(f)\big)\Big)\tilde{\pi}(d\alpha), \tag{11}$$

for some Borel probability measure $\tilde{\pi}$ on $[0,1]$, where

$$V_i^*(f) = \int_{\Delta_i} \varphi\left(\int_{E_i} f(\omega)\,\mu_i(d\omega)\right)\pi_i(d\mu_i) \quad \text{for } i = 1,2$$

[8] Probability measures from Δ in the table are represented as a vector of numbers, $(\mu(\omega_1), \mu(\omega_2), \mu(\omega_3), \mu(\omega_4))$.

represent conditional evaluations of acts after learning the event E_i ($i = 1, 2$), and where π_1 and π_2 are probability measures on Δ_1 and Δ_2, respectively. Put differently, the decision maker's ex ante preferences are represented by the functional

$$V^*(f\,E_1\,h) = \int_{[0,1]} \varphi\Bigg(\alpha\varphi^{-1}\left[\int_{\Delta_1} \varphi\left(\int_{E_1} f(\omega)\,\mu_1(d\omega)\right)\pi_1(d\mu_1)\right] + \\ (1-\alpha)\varphi^{-1}\left[\int_{\Delta_2} \varphi\left(\int_{E_2} h(\omega)\,\mu_2(d\omega)\right)\pi_2(d\mu_2)\right]\Bigg)\tilde{\pi}(d\alpha). \quad (12)$$

If E_1 and E_2 are non-null, then the decision maker's preferences conditional on learning E_1, as well as on learning E_2, have the Smooth Ambiguity structure. In addition, if we restrict attention to acts of the form $x\,E_1\,y$ in which x and y are constant on E_1 and E_2, respectively—the so-called one-step-ahead preferences—then, the evaluation functional V^* also has the Smooth Ambiguity form. It can be easily seen that the preferences with the recursive representation satisfy Dynamic Consistency and Consequentialism.

Looking at this construction, however, it is not very clear whether the recursive preferences defined by (11) or (12) generally fall within the class of Smooth Ambiguity preferences, as defined by Klibanoff et al. (2005) and recalled earlier in this paper.[9] Having the same form of preferences—in our case, Smooth Ambiguity—at the ex ante and subsequent stages may be very desirable both for normative and practical reasons. Therefore, it would be useful to know when representations (3) and (12) may be in agreement and induce the same ranking of acts. Note that Epstein and Schneider (2003) fully resolve a similar question for the maxmin model. Indeed, if the set of probability measures in the representation of ex ante maxmin preferences is rectangular, then the preferences satisfy Dynamic Consistency and Consequentialism and also admit a recursive maxmin representation; moreover, the representations for conditional preferences are easy to compute. Conversely, if an ex ante preference relation is recursive maxmin, it is always maxmin as well, and its maxmin representation uses a set of probability measures that is rectangular.

Our Theorem 1 helps to answer the above-stated question for the Smooth Ambiguity model. If an ex ante preference relation admits both the Smooth Ambiguity and recursive forms, then it must satisfy Dynamic Consistency and Consequentialism. Hence, our theorem applies (subject to the conditions of nontriviality, differentiability, and the conditions on the prior including the existence of moments). Then, the two forms can agree only in one of the three cases. The first case of Theorem 1 captures a situation in which the preference relation admits both representations, and $\tilde{\pi}$ in the recursive representation is a degenerate measure that is concentrated on a single point α. The second case captures a situation in which the probability measures π_1 and π_2 in the recursive representation are degenerate. With the exception of these two cases, agreement between the Smooth Ambiguity and Recursive Smooth Ambiguity forms requires φ be an exponential function. We illustrate such an agreement in the exponential case with the following example.

Example 4. *Consider the setup and the ex ante preference relation from Example 3. It can be seen that the functional $V(f)$ from its representation can be computed through recursive*

[9]That is, if we look at all acts without restricting them to be constant on E_1 and E_2, their ranking according to (12) may well be inconsistent with the ranking according to (3) for any choice of the probability distribution π in the latter representation.

Date: 2023-12-11 14:40:02 Revision: e8dda11

Representation (11) as

$$V(f) = V^*(f) = -\sum_{k=0}^{3} \tilde{\pi}_k \, exp\left(\frac{k}{3} \ln v_1(f,1) + \frac{3-k}{3} \ln v_2(f,1)\right)$$
$$= -\sum_{k=0}^{3} \tilde{\pi}_k \, exp\left(-\gamma \frac{k}{3}\big(-1/\gamma \ln v_1(f,1)\big) - \gamma \frac{3-k}{3}\big(-1/\gamma \ln v_2(f,1)\big)\right).$$

In this representation, we have $V_i^*(f) = -v_i(f,1)$ for $i = 1,2$.

Appendix

A A functional equation

We establish Theorem 1 by reducing the problem to solving a functional equation

$$\mathbf{E}\big[\varphi(z + x\,X + y(x,z)\,Y + h\,(1 - X - Y))\big] = \mathbf{E}\big[\varphi(z + h\,(1 - X - Y))\big] \tag{13}$$

that must hold for all values of z, x, and h in some open connected set, where X and Y are random variables; $\varphi(z)$ is an unknown function; and $y(x,z)$ is another unknown function such that $y(0,z) = 0$ and that does not depend on h.

In this Appendix, we solve this functional equation under certain assumptions on the function φ and the distribution of X and Y.

Our starting assumptions are as follows.

Assumption 1. *X and Y are random variables that are jointly distributed in the triangle $T \in \mathbb{R}^2$ with the vertices $(0,0)$, $(1,0)$, $(0,1)$.*

Assumption 2. *φ is a real function that is defined and three times continuously differentiable in a neighborhood O_{z_0} of some point $z_0 \in \mathbb{R}$, and $\varphi'(z) > 0$ and $\varphi''(z) < 0$ for all $z \in O_{z_0}$.*

Assumption 3. *The joint distribution of X and Y has the following properties:*

- $\mathbf{P}(X = 0) < 1$ *and* $\mathbf{P}(Y = 0) < 1$*;*
- $\mathbf{P}(0 < X + Y < 1) > 0$*; and*
- *for any $c \in \mathbb{R}_{++}$,* $\mathbf{P}(X = cY) < 1$.

Lemma 2. *Let g be a function defined as*

$$g(x,y,z) = \mathbf{E}[\varphi(z + x\,X + y\,Y)] - \varphi(z). \tag{14}$$

Then, under Assumptions 1–3,

- *g is defined and is three times continuously differentiable in some neighborhood of the point $(0,0,z_0)$; and*
- *there exists a unique function $y(x,z)$ defined in a neighborhood of $(0,z_0)$ such that $g(x,y(x,z),z) = 0$ for all (x,z) in a neighborhood of $(0,z_0)$ and $y(0,z_0) = 0$; moreover, y is three times continuously differentiable in that neighborhood.*

Proof. First, we verify that the function g is three times continuously differentiable in some neighborhood of the point $(0,0,z_0)$.

Let $[a,b]$ be an interval in $\mathbb{R}$ such that $z_0 \in (a,b) \subset I$. Since the variables $X, Y \in [0,1]$, we can choose a neighborhood $O(0,0,z_0)$ of the point $(0,0,z_0)$ such that $a < z + xX + yY < b$ for all $(x,y,z) \in O(0,0,z_0)$. Then, for all $(x,y,z), (x_1,y_1,z_1) \in O(0,0,z_0)$, we have

$$\frac{\mathbf{E}\varphi(z_1 + x_1 X + y_1 Y) - \mathbf{E}\varphi(z + xX + yY)}{\sqrt{(z_1 - z)^2 + (x_1 - x)^2 + (y_1 - y)^2}} = \mathbf{E}\frac{\varphi(z_1 + x_1 X + y_1 Y) - \varphi(z + xX + yY)}{\sqrt{(z_1 - z)^2 + (x_1 - x)^2 + (y_1 - y)^2}}. \tag{15}$$

The value under the expectation sign in (15) is bounded from above by

$$\frac{|z_1 - z + (x_1 - x)X + (y_1 - y)Y|}{\sqrt{(z_1 - z)^2 + (x_1 - x)^2 + (y_1 - y)^2}} \max_{t \in [a,b]} |\varphi'(t)| \le 3 \max_{t \in [a,b]} |\varphi'(t)|.$$

Hence, by the Dominated Convergence Theorem, we can pass to the limit under the expectation sign as $(x_1, y_1, z_1) \to (x,y,z)$. The proof that $g(x,y,z)$ has continuous second and third derivatives in a neighborhood of $(0,0,z_0)$ is similar.

Note that $g(0,0,z) = 0$ for any z in the neighborhood. We compute

$$\frac{\partial}{\partial y} g(0,0,z) = \mathbf{E}[\varphi'(z) Y] = \varphi'(z)\mathbf{E}Y.$$

Observe that $\mathbf{E}Y > 0$ (since $Y \ge 0$ and Y is not a.s. zero) and $\varphi'(z) \ne 0$ by assumption, so the computed derivative is non-zero. Hence, one can apply the implicit function theorem to obtain a unique function $y(x,z)$ such that $y(0,z_0) = 0$ and

$$g(x, y(x,z), z) = 0$$

for all (x,z) from some neighborhood of $(0,z_0)$; moreover, y is continuously differentiable, and

$$\frac{\partial}{\partial x} y(x,z) = -\frac{\mathbf{E}[X\varphi'(z + xX + yY)]}{\mathbf{E}[Y\varphi'(z + xX + yY)]}.$$

The proof that $y(x,z)$ has continuous second and third derivatives in a neighborhood of $(0,z_0)$ is similar. Note that $g(0,0,z) = 0$ and, hence, $y(0,z) = 0$ for all z from some neighborhood of z_0. ☐

Lemma 3. *Suppose that Eq. (13) holds for (x,h,z) in some open connected set containing the point $(0,0,z_0)$. Under Assumptions 1–3, it can be only if, in a neighborhood of z_0, φ has one of the following forms*

- $\varphi(z) = k\ln(z + b) + c$ *for some* $k > 0$ *and* $b, c \in \mathbb{R}$;
- $\varphi(z) = k(z + b)^\gamma + c$ *for some* $k, \gamma, b, c \in \mathbb{R}$ *such that* $k\gamma(1-\gamma) > 0$;
- $\varphi(z) = k\exp(\gamma z) + c$ *for some* $k < 0$, $\gamma < 0$, $c \in \mathbb{R}$.

Proof. Step 1. We differentiate the identity (13) by x when (x,h,z) belongs to a neighborhood of $(0,0,z)$:

$$\mathbf{E}[\varphi'(z + x\,X + y(x,z)\,Y + h\,U)(X + Y y'_x(x,z))] = 0, \tag{16}$$

where $U \equiv 1 - X - Y$. By setting $x = h = 0$, we obtain the equality

$$\mathbf{E}[X + Y y'_x(0,z)] = 0.$$

Thus, $y'_x(0,z)$ does not depend on z.

Step 2. Repeated differentiation of (16) by x produces the identity

$$\mathbf{E}\big[\varphi''(z + xX + yY + hU)(X + Y y'_x)^2\big] + \mathbf{E}[\varphi'(z + xX + yY + hU)]\, y''_{x,x} = 0. \tag{17}$$

By setting $x = h = 0$, we obtain

$$\mathbf{E}[\varphi''(z)(X + Y y'_x(0,z))^2] + \mathbf{E}[\varphi'(z)Y]\, y''_{x,x}(0,z) = 0$$

Thus, it must be that

$$\varphi''(z)\mathbf{E}[(X + Y y'_x(0,z))^2] = -\varphi'(z) y''_{x,x}(0,z)\mathbf{E}Y. \tag{18}$$

Note that $\mathbf{E}[(X + Y y'_x(0,z))^2] > 0$ since $X \neq y'_x(0,z)Y$ with a positive probability by Assumption 3. Differentiating (17) by h and setting $x = h = 0$, we also obtain

$$\varphi'''(z)\mathbf{E}[(X + Y y'_x(0,z))^2 U] + \varphi''(z) y''_{x,x}(0,z)\mathbf{E}[UY] = 0. \tag{19}$$

Step 3. Suppose that $\mathbf{E}[UY] = 0$. Then, $UY = 0$ a.s., and

$$\varphi'''(z)\mathbf{E}\left[(X + Y y'_x(0,z))^2 U\right] = 0. \tag{20}$$

Suppose that

$$\mathbf{E}\left[(X + Y y'_x(0,z))^2 U\right] = 0.$$

Expanding the left-hand side and using the fact that $UY = 0$, a.s., we obtain that $\mathbf{E}[X^2 U] = 0$, and, hence, $XU = 0$ a.s. It follows that $\mathbf{P}(X = Y = 0) + \mathbf{P}(X + Y = 1) = 1$, a contradiction to Assumption 3. Then, to satisfy (20), it must be that $\varphi'''(z) = 0$ for every z in some neighborhood of z_0, and, hence, φ is a quadratic function.

Step 4. Suppose that $\mathbf{E}\left[(X + y'_x(0,z)Y)^2 U\right] = 0$. Then, $(X + y'_x(0,z)Y)^2 U = 0$ a.s. and

$$\varphi''(z) y''_{x,x}(0,z)\mathbf{E}[UY] = 0.$$

The case $\mathbf{E}[UY] = 0$ has been considered in the previous step. If $y''_x(0,z) = 0$ for every z in some neighborhood of z_0, then it follows from (18) that $\varphi''(z) = 0$. The equality $\varphi''(z) = 0$ contradicts Assumption 2.

Step 5. In the remaining case, we combine (18) and (19) to obtain

$$c_1 \frac{\varphi''(z)}{\varphi'(z)} = -y''_{x,x}(0,z) \quad \text{and} \quad c_2 \frac{\varphi'''(z)}{\varphi''(z)} = -y''_{x,x}(0,z),$$

where

$$c_1 := \frac{\mathbf{E}\left[(X + Y y'_x(0,z))^2\right]}{\mathbf{E}[Y]}, \quad c_2 := \frac{\mathbf{E}\left[(X + Y y'_x(0,z))^2 U\right]}{\mathbf{E}[UY]}.$$

Date: 2023-12-11 14:40:02 Revision: e8dda11

It was noted in Step 1 that $y'_x(0,z)$ does not depend on z; therefore, c_1 and c_2 are some positive constants.

Since $\varphi'(z) > 0$ and $\varphi''(z) < 0$, we have

$$c_1(\ln \varphi'(z))' = c_2(\ln(-\varphi''(z)))'.$$

As a result, for some constant c_3, it must be that

$$\frac{c_1}{c_2} \ln \varphi'(z) = \ln(-\varphi''(z)) + c_3$$

in some neighborhood of z_0. Taking the exponential, we obtain $\varphi''(z) = \widetilde{c}\varphi'(z)^\alpha$, where $\widetilde{c} = -e^{-c_3}$ and $\alpha = \frac{c_1}{c_2}$. Since $\varphi'(z) \neq 0$, we can rewrite it as

$$\frac{\varphi''(z)}{\varphi'(z)^\alpha} = \widetilde{c}.$$

1. If $\alpha = 1$, then
$$\varphi(z) = d_1 \exp(\widetilde{c}z) + d_2$$
in some neighborhood of the point z_0, where d_1, d_2 are some constants.

2. If $\alpha = 2$, then
$$\varphi(z) = d_1 + \ln(d_2 z + d_3)$$
in some neighborhood of the point z_0, where d_1, d_2, d_3 are some constants.

3. If $\alpha \notin \{1, 2\}$, then
$$(\varphi'(z))^{1-\alpha}/(1-\alpha) = \widetilde{c}z + d_1$$
and, in turn,
$$\varphi(z) = d_2(d_3 z + d_4)^{d_5} + d_6,$$
in some neighborhood of the point z_0, where d_1 and d_6 are some constants and
$$d_2 = \frac{1}{(2-\alpha)\widetilde{c}}, \quad d_3 = (1-\alpha)\widetilde{c}, \quad d_4 = (1-\alpha)d_1, \quad d_5 = \frac{2-\alpha}{1-\alpha}.$$

□

Lemma 4. *Under Assumptions 1–3, $\varphi(z) = k(z+b)^\gamma + c$ for any $k < 0$, $b, c \in \mathbb{R}$ and $\gamma \in \mathbb{N}$ is not a solution of Eq. (13).*

Proof. Observe, first, that $\gamma \neq 1$ due to Assumption 2.

Note that if a function φ satisfies Eq. (13), then the function

$$\widetilde{\varphi}(z) = d\varphi(z-b) - c,$$

where b, c, d are arbitrary constants, also satisfies the equation. Hence, we can assume without loss of generality that $b = 0$, $c = 0$, and $k = -1$.

Observe that it cannot be that $z_0 = 0$, because otherwise either $\varphi'(z)$ changes sign in any neighborhood of z_0 (if γ is even), or $\varphi''(z)$ changes sign in any neighborhood of z_0 (if γ is odd), which contradicts Assumption 2.

As follows from our equation (13),

$$\mathbf{E}\left[(z_0 + xX + y(x)Y + hU)^\gamma\right] = \mathbf{E}\left[(z_0 + hU)^\gamma\right], \tag{21}$$

where, with a slight abuse of notation, we write $y(x)$ for $y(x, z_0)$ and, as earlier, $U \equiv 1 - X - Y$. Differentiating the above equation by h and setting $h = 0$, we obtain

$$\mathbf{E}\left[(z_0 + xX + y(x)Y)^{\gamma-n}U^n\right] = \mathbf{E}\left[z_0^{\gamma-n}U^n\right]$$

for all $n \in \{1, \ldots, \gamma\}$. If we let $n = \gamma - 1$, we obtain

$$x\,\mathbf{E}\left[XU^{\gamma-1}\right] + y(x)\,\mathbf{E}\left[YU^{\gamma-1}\right] = 0.$$

Thus, $y(x) = -cx$ for some $c \in \mathbb{R}$.

Substituting the latter result into (21) and setting $h = 0$, we have

$$\mathbf{E}\left[(z_0 + x(X - cY))^\gamma\right] = z_0^\gamma$$

for all x in a neighborhood of zero. Differentiating it by x and setting $x = 0$, we obtain

$$\mathbf{E}\left[z_0^{\gamma-k}(X - cY)^n\right] = 0$$

for $n = 1, 2$. For $n = 2$, this contradicts Assumption 3. ☐

Assumption 4. *The joint distribution of X and Y is such that $\mathbf{E}[(1 - X - Y)^p \mathbb{1}_{X+Y<1}] < \infty$ for all $p \in \mathbb{R}$.*

Lemma 5. *Under Assumptions 1–4, $\varphi(z) = k(z+b)^\gamma + c$ for any $k, b, c \in \mathbb{R}$ and $\gamma \in \mathbb{R}\backslash\mathbb{N}$ such that $k\gamma(1-\gamma) > 0$ is not a solution of Eq. (13).*

Proof. Step 1. As in the proof of Lemma 4, we can assume without loss of generality that $b = c = 0$. Since z_0 cannot be zero, we can also assume without loss of generality that $z_0 = 1$. Therefore, we have an equation

$$\mathbf{E}\left[(1 + xX + y(x)Y + hU)^\gamma\right] = \mathbf{E}\left[(1 + hU)^\gamma\right] \tag{22}$$

that holds for all (x, h) from some neighborhood of $(0, 0)$.

Step 2. Consider the expansion of the function $z \mapsto (z + hu)^\gamma$ for $z > 0$ in a Taylor series:

$$(z + hu)^\gamma = \sum_{n=0}^{\infty} a_n h^n u^n z^{\gamma-n}, \tag{23}$$

where a_n are the coefficients for which we have $a_n \neq 0$ for all $n \in \mathbb{N}$. The remainder of the series can be expressed as

$$(z + hu)^\gamma - \sum_{n=0}^{N-1} a_n h^n u^n z^{\gamma-n} = \frac{h^N u^N}{N!}\,\frac{d^N(x^\gamma)}{dx^N}\bigg|_{x=\xi}$$

for some ξ between z and $z+hu$. If (z,h) belongs to a sufficiently small neighborhood of the point $(1,0)$ and $u \in [0,1]$, then, due to the h^N term, the remainder goes to zero uniformly over (z,h,u) when $N \to \infty$; that is, the series in the right-hand side of (23) converges uniformly.

We use this series in (22) with $z = 1 + xX + y(x)Y$ and $u = U$ in the left-hand side, and $z = 1$ and $u = U$ in the right-hand side, to obtain

$$\sum_{n=0}^{\infty} a_n h^n \mathbf{E}\left[(1 + xX + y(x)Y)^{\gamma-n} U^n\right] = \sum_{n=0}^{\infty} a_n h^n \mathbf{E}\left[U^n\right],$$

where summation and expectation have been interchanged due to uniform convergence. We equate the coefficients of h^n to obtain

$$\mathbf{E}\left[(1 + xX + y(x)Y)^{\gamma-n} U^n\right] = \mathbf{E}\left[U^n\right] \tag{24}$$

that holds for all $n \in \mathbb{N}$.

Step 3. For all $n \in \mathbb{N}$, we can replace expectation with conditional expectation in (24) to obtain

$$\mathbf{E}\left[(1 + xX + y(x)Y)^{\gamma-n} U^n \,|\, U > 0\right] = \mathbf{E}\left[U^n \,|\, U > 0\right]$$

that holds for all $n \in \mathbb{N}$. Differentiating it by x, we obtain

$$\mathbf{E}\left[(X + y'(x)Y)(1 + xX + y(x)Y)^{\gamma-n-1} U^n \,\middle|\, U > 0\right] = 0$$

for all $n \in \mathbb{N}$. We rewrite it as

$$\mathbf{E}\left[(X + y'(x)Y)(1 + xX + y(x)Y)^{\gamma-1} \left(\frac{U}{1 + xX + y(x)Y}\right)^n \,\middle|\, U > 0\right] = 0$$

for $n \in \mathbb{N}$. Then,

$$\mathbf{E}\left[(X + y'(x)Y)(1 + xX + y(x)Y)^{\gamma-1}\, P\!\left(\frac{U}{1 + xX + y(x)Y}\right) \,\middle|\, U > 0\right] = 0 \tag{25}$$

holds for every polynomial P such that $P(0) = 0$.

Step 4. Consider the function $r : [0,1] \to \mathbb{R}$ defined as $r(t) = t^{\gamma-1}$ for $t > 0$ and $r(0) = 0$. By Assumption 4, this function is integrable on $[0,1]$ under the distribution of U conditional on $U > 0$, that is, under the Borel measure $\mathbf{P}_U$ defined as $\mathbf{P}_U(A) = \mathbf{P}(U \in A \,|\, U > 0)$. Hence, for the Borel measure $\mathbf{Q}$ defined as

$$\mathbf{Q}(A) = \mathbf{P}\left(\frac{U}{1 + xX + y(x)Y} \in A \,\middle|\, U > 0\right),$$

we also have that $\int_{[0,1]} |r(t)|\, \mathbf{Q}(dt) < \infty$. Then, the function $r(t)$ can be approximated in $L^1(\mathbf{Q})$ by a continuous integrable function that takes the value of 0 at 0, which, in turn, can be approximated by polynomials that take the value of 0 at 0. It follows from (25) that

$$\mathbf{E}\left[(X + y'(x)Y)(1 + xX + y(x)Y)^{\gamma-1}\, r\!\left(\frac{U}{1 + xX + y(x)Y}\right) \,\middle|\, U > 0\right] = 0,$$

which can be simplified to

$$\mathbf{E}\left[(X + y'(x)Y)\,U^{\gamma-1}\,\middle|\,U > 0\right] = 0.$$

Step 5. The latter equation can be solved for $y'(x)$ to observe that, in fact, $y'(x)$ does not depend on x; in turn, $y''(x) = 0$. Using these findings in Eq. (18) from the proof of Lemma 3 that follows from our main equation (13), we arrive at a contradiction to the assumptions that $\mathbf{P}(X = cY) < 1$ for any $c \in \mathbb{R}_+$ and $\varphi'' < 0$. ☐

Lemma 6. *Under Assumptions 1–4, $\varphi(z) = k\ln(z + b) + c$ for any $k > 0$ and $b, c \in \mathbb{R}$ is not a solution of Eq. (13).*

Proof. The proof is similar to the one of Lemma 5, with the only change that Step 2 considers a Tailor series for the logarithmic function, and the rest of the proof goes through with $\gamma = 0$. ☐

Theorem 7. *Suppose that (13) holds for all (x, h, z) in some neighborhood of $(0, 0, z_0)$ and some continuous function $y(x, z)$ with $y(0, z_0) = 0$, and Assumptions 1–4 hold. Then, in some neighborhood of z_0, it must be that*

$$\varphi(z) = k\exp(\gamma z) + c, \tag{26}$$

where $k < 0$, $\gamma < 0$, and c are some constants, and the function $y(x, z)$ must not depend on z.

Proof. Follows from Lemmas 2–6. ☐

B Proof of Theorem 1

Lemma 8. *Suppose that P is a Borel probability measure on $\mathbb{R}$ with compact support, $u, v \in L^1(P)$, and*

$$\int_{\mathbb{R}} u(t)\,e^{-ht}\,P(dt) = \int_{\mathbb{R}} v(t)\,e^{-ht}\,P(dt)$$

holds for all h in a neighborhood of some point h_0. Then, it must be that $u(t) = v(t)$ for all t, P-a.s.

Proof. Let μ_1 and μ_2 be Borel measures on $\mathbb{R}$ defined as

$$\mu_1(A) = \int_A u(t)\,P(dt), \qquad \mu_2(A) = \int_A v(t)\,P(dt).$$

With a change of variables in the Lebesgue integral, we have that

$$\int_{\mathbb{R}} e^{-hs}\,\mu_1(ds) = \int_{\mathbb{R}} e^{-hs}\,\mu_2(ds);$$

that is, the Laplace transforms of the measures μ_1 and μ_2 coincide in some neighborhood of zero. Both measures have compact support, and, hence, their Laplace transforms are infinitely differentiable. This gives the equalities

$$\int_{\mathbb{R}} s^k\,\mu_1(ds) = \int_{\mathbb{R}} s^k\,\mu_2(ds)$$

that must hold for all $k \in \mathbb{Z}_+$. A finite measure with compact support is uniquely identified by its moments (Hausdorff moment problem, see, e.g. Akhiezer, 1965, Theorem 2.6.4). Hence, the measures μ_1 and μ_2 coincide; in turn, their Radon-Nikodym derivatives with respect to the measure P must coincide, and we have that $u(t) = v(t)$ for all t, P-a.s. □

Proof of Theorem 1. *The direct part.* Suppose that $\gtrsim$ admits Representation (4), with φ and π as described in the statement of the theorem, and satisfies Dynamic Consistency and Consequentialism with respect to the partitioning $\{E_1, E_2\}$. We will proceed in steps.

Step 1. Suppose, first, that $\mu(E_1) \in \{0,1\}$ (π-almost surely). Note that this implies that $\mu(E_2) \in \{0,1\}$ (π-almost surely). Let $\alpha := \pi(\{\mu(E_1) = 1\})$ and define probability measures π_1 and π_2 on (Δ_1, Σ_1) and (Δ_2, Σ_2), respectively, through the conditionals of π: $\pi_i(S) := \pi\big(\{\mu|_{2^{E_i}} \in S\} \mid \{\mu(E_i) = 1\}\big)$ for $i = 1,2$, where, for any $\mu \in \Delta$, $\mu|_{2^{E_i}}$ is the restriction of μ to 2^{E_i}. Then, the outer integral in Representation (4) can be split into two integrals—the integral over the set of measures μ for which $\mu(E_1) = 1$ and the integral over the set for which $\mu(E_2) = 1$. We have obtained Case (i) from the statement of the theorem.

Step 2. Suppose, second, that, for any non-null E_{11} and E_{12} that form a partitioning of E_1, there is π-almost sure proportionality between $\mu(E_{11})$ and $\mu(E_{12})$; and similarly for E_2, for any non-null E_{21} and E_{22} that form its partitioning, there is π-almost sure proportionality between $\mu(E_{21})$ and $\mu(E_{22})$. For a set $A \subseteq E_1$, let $c(A)$ denote the proportionality coefficient: $\mu(A) = c(A)\mu(E_1 \backslash A)$, π-a.s., and similarly for subsets of E_2.

We define the function $\nu_1 : 2^{E_1} \to [0,1]$ as follows. For any $A \subseteq E_1$, we let $\nu_1(A) = 0$ if A is null; $\nu_1(A) = 1$ if $E_1 \backslash A$ is null; and $\nu_1(A) = c(A)/(1 + c(A))$ otherwise. In all of the three cases, we have: for any $A \subseteq E_1$, $(1 - \nu_1(A))\mu(A) = \nu_1(E)(\mu(E_1) - \mu(A))$ holds for all $\mu \in \Delta$, π-a.s. Then, arithmetically, for any $A \subseteq E_1$, $\mu(A) = \nu_1(A)\mu(E_1)$ holds for all $\mu \in \Delta$, π-a.s. Considering this equality for a fixed $\mu \in \Delta$ such that $\mu(E_1) > 0$, we observe that ν_1 is an additive set-function. Similarly, we define $\nu_2 : 2^{E_2} \to [0,1]$ to obtain that, for any $A \subseteq E_2$, $\mu(A) = \nu_2(A)\mu(E_2)$ holds for all $\mu \in \Delta$, π-a.s., and ν_2 is additive. Now we rewrite Representation (4) as

$$
\begin{aligned}
V(f) &= \int_\Delta \varphi\left(\int_{E_1} f(\omega)\,\mu(d\omega) + \int_{E_2} f(\omega)\,\mu(d\omega)\right)\pi(d\mu) \\
&= \int_\Delta \varphi\left(\mu(E_1)\int_{E_1} f(\omega)\,\nu_1(d\omega) + \mu(E_2)\int_{E_2} f(\omega)\,\nu_2(d\omega)\right)\pi(d\mu) \\
&= \int_{[0,1]} \varphi\left(\alpha\int_{E_1} f(\omega)\,\nu_1(d\omega) + (1-\alpha)\int_{E_2} f(\omega)\,\nu_2(d\omega)\right)\tilde{\pi}(d\alpha),
\end{aligned}
$$

where $\tilde{\pi}$ is a Borel probability measure on $[0,1]$ defined as $\tilde{\pi}(S) = \pi(\{\mu(E_1) \in S\})$. We have obtained Case (ii) from the statement of the theorem.

Step 3. Suppose that the assumptions of Steps 1–2 do not hold. In particular, there exist non-null E_{11} and E_{12} such that $E_1 = E_{11} \sqcup E_{12}$, and there is no π-almost sure proportionality between $\mu(E_{11})$ and $\mu(E_{12})$.

Step 3a. Fix an arbitrary $z_0 \in I$, and consider acts f of the form $f = (z + x)\,E_{11}\,(z + y)\,E_{12}\,(z + h)$ for all $(x, y, h, z) \in (-\varepsilon, \varepsilon)^3 \times (z_0 - \varepsilon, z_0 + \varepsilon)$, where $\varepsilon > 0$ is chosen so that

$f \in \mathcal{F}(I)$. For $(h,z) \in (-\varepsilon,\varepsilon) \times (z_0 - \varepsilon, z_0 + \varepsilon)$, let $L_{h,z} \subset \mathbb{R}^2$ be defined as

$$L_{h,z} = \Big\{(x,y) \in (-\varepsilon,\varepsilon)^2 : (z+x)\, E_{11}\, (z+y)\, E_{12}\, (z+h) \sim z\, E_1\, (z+h)\Big\},$$

where $\sim$ is the symmetric part of $\gtrsim$. As follows from Dynamic Consistency and Consequentialism, $L_{h,z}$, in fact, do not depend on h. Using Representation (4), it implies that

$$\int_\Delta \varphi\big(z + x\mu(E_{11}) + y\mu(E_{12}) + h\mu(E_2)\big)\, \pi(d\mu) = \int_\Delta \varphi\big(z + h\mu(E_2)\big)\, \pi(d\mu)$$

holds for all $(x,y) \in L_z$, all $h \in (-\varepsilon,\varepsilon)$, and all $z \in (z_0 - \varepsilon, z_0 + \varepsilon)$. We use Lemma 2 to obtain a function $y(x,z)$ and functional equation (13), where $\mu(E_{11})$ plays the role of the random variable X, and $\mu(E_{12})$ plays the role of the random variable Y.

Step 3b. By Theorem 7, there exist $\gamma > 0$, $k > 0$ and $c \in \mathbb{R}$ such that $\varphi(z) = -ke^{-\gamma z} + c$ for all z in some neighborhood of z_0. Since positive affine transformations of φ change neither the preferences nor our functional equation, we may assume without loss of generality that $k = 1$ and $c = 0$. We claim that $\varphi(z) = -e^{-\gamma z}$ for all $z \in I$. Indeed, suppose that this equality fails at some z_1, and assume without loss of generality that $z_1 > z_0$. We let $z_2 = \sup\{z \geq z_0 : \varphi(z) = -e^{-\gamma z}\}$ and apply Theorem 7 to a neighborhood of z_2 to arrive at a contradiction.

Step 3c. We will continue using the interpretation of μ as a random vector whose probability law is π. In particular, for any $f \in \mathcal{F}(I)$ and $h \in I$, we can write the representation functional in (4) as

$$V(f) = \mathbf{E}\Big[-e^{-\gamma \int_{E_1} f(\omega)\,\mu(d\omega) - \gamma \int_{E_2} f(\omega)\,\mu(d\omega)}\Big]. \tag{27}$$

For any $f \in \mathcal{F}(I)$, let $v_1(f,\cdot)$ be defined according to the Doob-Dynkin lemma as a measurable function such that

$$\mathbf{E}\Big[e^{-\gamma \int_{E_1} f(\omega)\,\mu(d\omega)} \,\Big|\, \mu(E_1)\Big] = v_1(f, \mu(E_1)), \tag{28}$$

where the equality holds almost surely with respect to the probability law of $\mu(E_1)$. Let W_1 denote the random variable $\mu(E_1)$ and let P_{W_1} be its probability law, $P_{W_1}(S) = \pi(\{\mu(E_1) \in S\})$ for any Borel $S \subseteq \mathbb{R}$.

On $E_2 = \Omega \setminus E_1$, we define objects similarly. Note that $W_2 = 1 - W_1$, and, hence,

$$\mathbf{E}\Big[-e^{-\gamma \int_{E_2} f(\omega)\,\mu(d\omega)} \,\Big|\, \mu(E_1)\Big] = \mathbf{E}\Big[-e^{-\gamma \int_{E_2} f(\omega)\,\mu(d\omega)} \,\Big|\, \mu(E_2)\Big] = v_2(f, \mu(E_2)).$$

Using the law of iterated expectation and changing variables in the Lebesgue integral, we rewrite (27) as

$$V(f) = \mathbf{E}\Big[\mathbf{E}\Big[-e^{-\gamma \int_{E_1} f(\omega)\,\mu(d\omega) - \gamma \int_{E_2} f(\omega)\,\mu(d\omega)} \,\Big|\, \mu(E_1)\Big]\Big] =$$
$$\mathbf{E}\big[-v_1(f, W_1)\, v_2(f, W_2)\big] = -\int_{[0,1]} v_1(f,w)\, v_2(f, 1-w)\, P_{W_1}(dw)$$

to obtain representation (5). Note that the conditional expectation in (28) can be rewritten as an integral with respect to the conditional probability measure,

$$v_1(f,\alpha) = \int_\Omega e^{-\gamma \int_{E_1} f(\omega)\,\mu(d\omega)}\, \pi\big(d\mu \mid \mu(E_1) = \alpha\big),$$

and rearranged to obtain expression (6) for v_1. The expression for v_2 obtains similarly. This establishes the representation of Case (iii).

Step 3d. It remains to prove that $v_i(\cdot,\alpha)$ for $i=1,2$ rank acts in the same way when α varies in $(0,1]$.

Let $C_1 : \mathcal{F}(I) \to I$ be the conditional "certainty equivalent" functional on E_1: for any $f \in \mathcal{F}(I)$ and $h \in I$, $C_1(f,h)$ is a number such that

$$V(f\,E_1\,h) = \mathbf{E}\left[-e^{-\gamma\,C_1(f,h)\,\mu(E_1)-\gamma\,h\,(1-\mu(E_1))}\right]. \tag{29}$$

Note that $C_1(f,h)$ exists and is unique because of continuity and monotonicity of V, as well as the assumption that $\mu(E_1)>0$ with positive probability under π. Also, C_1 does not depend on h because of Dynamic Consistency and Consequentialism. We define C_2 similarly as the conditional certainty equivalent on E_2.

For all $h \in I$, we have that

$$V(f\,E_1\,h) = \int_{[0,1]} v_1(f,w)\,e^{-\gamma h(1-w)}\,P_{W_1}(dw).$$

Comparing this with (29), it follows from Lemma 8 that $v_1(f,W_1) = -e^{-\gamma\,C_1(f)\,W_1}$, P_{W_1}-a.s. Therefore, for all $f,g \in \mathcal{F}(I)$,

$$v_1(f,\alpha) \geq v_1(g,\alpha) \Leftrightarrow C_1(f) \geq C_1(g) \Leftrightarrow v_1(f,\alpha') \geq v_1(g,\alpha')$$

for P_{W_1}-a.s. $\alpha,\alpha' \in (0,1]$. The analysis for v_2 is similar. This completes the proof of the direct part of the theorem.

The converse part. Suppose that φ is strictly increasing.

(i) Suppose that $\succsim$ admits a representation via

$$V(f) = \alpha \int_{\Delta_1} \varphi\left(\int_{E_1} f(\omega)\,\mu_1(d\omega)\right)\pi_1(d\mu_1) + (1-\alpha)\int_{\Delta_2} \varphi\left(\int_{E_2} f(\omega)\,\mu_2(d\omega)\right)\pi_2(d\mu_2).$$

for some $\alpha \in (0,1)$ and some probability measures π_1 and π_2 on (Δ_1,Σ_1) and (Δ_2,Σ_2), respectively. Then, for arbitrary $f,g,h,h' \in \mathcal{F}(I)$,

$$V(f\,E_1\,h) \geq V(g\,E_1\,h) \Leftrightarrow \int_{\Delta_1} \varphi\left(\int_{E_1} f(\omega)\,\mu_1(d\omega)\right)\pi_1(d\mu_1) \geq \int_{\Delta_1} \varphi\left(\int_{E_1} g(\omega)\,\mu_1(d\omega)\right)\pi_1(d\mu_1) \Leftrightarrow V(f\,E_1\,h') \geq V(g\,E_1\,h').$$

A similar equivalence holds for E_2 and E_1 interchanged. Therefore, the preference relation satisfies Dynamic Consistency and Consequentialism.

(ii) Suppose that $\succsim$ admits a representation via

$$V(f) = \int_{[0,1]} \varphi\Big(\alpha \int_{E_1} f(\omega)\,\mu_1(d\omega) + (1-\alpha)\int_{E_2} f(\omega)\,\mu_2(d\omega)\Big)\,\tilde{\pi}(d\alpha).$$

for some Borel probability measure $\tilde{\pi}$ on $[0,1]$ such that $\tilde{\pi}\big(\{0,1\}\big) < 1$ and probability measures μ_1 and μ_2 from Δ_1 and Δ_2, respectively. Observe that the mappings $v_h : I \to \mathbb{R}$ for $h \in \mathcal{F}(I)$ defined as

$$v_h(t) = \int_{[0,1]} \varphi\Big(\alpha t + (1-\alpha)\int_{E_2} h(\omega)\,\mu_2(d\omega)\Big)\,\tilde{\pi}(d\alpha)$$

are strictly increasing. Then, for arbitrary $f, g, h, h' \in \mathcal{F}(I)$,

$$V(f\,E_1\,h) \geq V(g\,E_1\,h) \Leftrightarrow \int_{E_1} f(\omega)\,\mu_1(d\omega) \geq \int_{E_1} g(\omega)\,\mu_1(d\omega) \Leftrightarrow \\ V(f\,E_1\,h') \geq V(g\,E_1\,h').$$

A similar equivalence holds for E_2 and E_1 interchanged. Therefore, the preference relation satisfies Dynamic Consistency and Consequentialism.

(iii) Suppose that $\succsim$ admits a representation via

$$V(f) = -\int_{[0,1]} v_1(f,\alpha)\,v_2(f,1-\alpha)\,\tilde{\pi}(d\alpha),$$

for some $\gamma > 0$, probability measure $\tilde{\pi}$ on $[0,1]$ such that $\tilde{\pi}\big(\{0,1\}\big) < 1$, and probability measures $\pi_{i,\alpha}$ on (Δ_i, Σ_i) for $i = 1,2$ and $\alpha \in (0,1]$ such that, for $i = 1,2$ and v_i defined as $v_i(f,0) = 1$ and

$$v_i(f,\alpha) = \int_{\Delta_i} e^{-\gamma\alpha \int_{E_i} f(\omega)\,\mu(d\omega)}\,\pi_{i,\alpha}\big(d\mu\big) \quad \text{for } \alpha > 0,$$

the equivalence

$$v_i(f,\alpha) \geq v_i(g,\alpha) \Leftrightarrow v_i(f,\alpha') \geq v_i(g,\alpha') \qquad \forall_{f,g \in \mathcal{F}_i(I)}$$

holds for $\tilde{\pi}$-almost all $\alpha, \alpha' \in (0,1]$.

Fix arbitrary $f, g, h, h' \in \mathcal{F}(I)$. First, suppose that $v_1(f,\alpha) = v_1(g,\alpha)$ for $\tilde{\pi}$-almost all $\alpha, \alpha' \in (0,1]$. Then it is clear that $V(f\,E_1\,h) = V(g\,E_1\,h)$ and $V(f\,E_1\,h') = V(g\,E_1\,h')$ simultaneously. Second, suppose without loss of generality that $v_1(f,\alpha) > v_1(g,\alpha)$ for $\tilde{\pi}$-almost all $\alpha, \alpha' \in (0,1]$. Then it is clear that $V(f\,E_1\,h) < V(g\,E_1\,h)$ and $V(f\,E_1\,h') < V(g\,E_1\,h')$ simultaneously. The symmetric case with E_2 and E_1 interchanged is analogous. Therefore, the preference relation satisfies Dynamic Consistency and Consequentialism.

□

References

Akhiezer, Naum I., *The classical moment problem and some related questions in analysis*, Oliver and Boyd, 1965.

Al-Najjar, Nabil I. and Jonathan Weinstein, "The Ambiguity Aversion Literature: A Critical Assessment," *Economics and Philosophy*, November 2009, *25* (Special Issue 03), 249–284.

Eichberger, Jurgen and David Kelsey, "Uncertainty Aversion and Dynamic Consistency," *International Economic Review*, August 1996, *37* (3), 625–40.

Ellsberg, Daniel, "Risk, Ambiguity and the Savage Axioms," *The Quarterly Journal of Economics*, November 1961, *75* (4), 643–669.

Epstein, Larry G. and Martin Schneider, "Recursive multiple-priors," *Journal of Economic Theory*, November 2003, *113* (1), 1–31.

Gilboa, Itzhak and David Schmeidler, "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, April 1989, *18* (2), 141–153.

_ and _ , "Updating Ambiguous Beliefs," *Journal of Economic Theory*, February 1993, *59* (1), 33–49.

Gumen, Anna and Andrei Savochkin, "Dynamically stable preferences," *Journal of Economic Theory*, 2013, *148* (4), 1487–1508.

Hanany, Eran and Peter Klibanoff, "Updating preferences with multiple priors," *Theoretical Economics*, September 2007, *2* (3).

_ and _ , "Updating Ambiguity Averse Preferences," *The B.E. Journal of Theoretical Economics*, 2009, *9.*

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji, "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, November 2005, *73* (6), 1849–1892.

_ , _ , and _ , "Recursive smooth ambiguity preferences," *Journal of Economic Theory*, May 2009, *144* (3), 930–976.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini, "Dynamic variational preferences," *Journal of Economic Theory*, May 2006, *128* (1), 4–44.

Sarin, Rakesh and Peter P. Wakker, "Dynamic Choice and NonExpected Utility," *Journal of Risk and Uncertainty*, November 1998, *17* (2), 87–119.

Siniscalchi, Marciano, "Two out of Three Ain't Bad: A Comment on "The Ambiguity Aversion Literature: A Critical Assessment"," *Economics and Philosophy*, 2009, *25* (Special Issue 03), 335–356.

_ , "Dynamic choice under ambiguity," *Theoretical Economics*, September 2011, *6* (3), 379–421.