The B.E. Journal of Theoretical Economics

Advances

Volume 9, *Issue* 1	2009	*Article* 37

Updating Ambiguity Averse Preferences

Eran Hanany* Peter Klibanoff†

*Tel Aviv University, hananye@post.tau.ac.il

†Northwestern University, peterk@kellogg.northwestern.edu

Recommended Citation

Eran Hanany and Peter Klibanoff (2009) "Updating Ambiguity Averse Preferences," *The B.E. Journal of Theoretical Economics*: Vol. 9: Iss. 1 (Advances), Article 37.

Available at: http://www.bepress.com/bejte/vol9/iss1/art37

Updating Ambiguity Averse Preferences*

Eran Hanany and Peter Klibanoff

Abstract

Dynamic consistency leads to Bayesian updating under expected utility. We ask what it implies for the updating of more general preferences. In this paper, we characterize dynamically consistent update rules for preference models satisfying ambiguity aversion. This characterization extends to regret-based models as well. As applications of our general result, we characterize dynamically consistent updating for two important models of ambiguity averse preferences: the ambiguity averse smooth ambiguity preferences (Klibanoff, Marinacci and Mukerji [Econometrica 73 2005, pp. 1849-1892]) and the variational preferences (Maccheroni, Marinacci and Rustichini [Econometrica 74 2006, pp. 1447-1498]). The latter includes max-min expected utility (Gilboa and Schmeidler [Journal of Mathematical Economics 18 1989, pp. 141-153]) and the multiplier preferences of Hansen and Sargent [American Economic Review 91(2) 2001, pp. 60-66] as special cases. For smooth ambiguity preferences, we also identify a simple rule that is shown to be the unique dynamically consistent rule among a large class of rules that may be expressed as reweightings of the Bayes' rule.

KEYWORDS: updating, dynamic consistency, ambiguity, regret, Ellsberg, Bayesian, consequentialism, smooth ambiguity

*We thank Nabil Al-Najjar, Sandeep Baliga, Eddie Dekel, Jurgen Eichberger, Larry Epstein, Raphaël Giraud, Simon Grant, Ian Jewitt, Massimo Marinacci, Sujoy Mukerji, Bob Nau, Ben Polak, Chris Shannon, Marciano Siniscalchi, Tomasz Strzalecki, Kane Sweeney, Tim Van Zandt and especially Fabio Maccheroni for helpful comments and discussion. We also thank the co-editor, Federico Echenique, and two anonymous referees for their comments. Additional thanks to seminar audiences at INSEAD, HEC, Johns Hopkins, Northwestern and Technion and at FUR XII, RUD '07, Fall '07 Midwest Economic Theory, ESNASM '08, the 3^{rd} Israeli Game Theory Conference and the 3^{rd} World Congress of the Game Theory Society. This research was partially supported by Grant No. 2006264 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel.

1 Introduction

A central question facing any theory of decision making under uncertainty is how preferences are updated to incorporate new information. Since updated preferences govern future choices, it is important to know how they relate to information contingent choices made ex-ante. *Dynamic consistency* is the requirement that ex-ante contingent choices are respected by updated preferences. This consistency is implicit in the standard way of thinking about a dynamic choice problem as equivalent to a single ex-ante choice to which one is committed, and is thus ubiquitous in economic modeling. We formally define dynamic consistency in Section 2.2.

Under subjective expected utility, updating preferences by applying Bayes' rule to the subjective probability is the standard way to update. Why is this so? Dynamic consistency is the primary justification for Bayesian updating. Not only does Bayesian updating imply dynamic consistency, but, if updating consists of specifying a conditional probability measure for each (non-null) event, dynamic consistency implies these conditional measures must be the Bayesian updates.[1] The requirement that updating consists of specifying a conditional probability measure ensures *closure* for expected utility preferences – i.e., that each such preference remains an expected utility preference after updating.

Since dynamic consistency and closure lead to a well-established theory of updating for expected utility, it is important to ask what they imply for the updating of more general preferences. Closure for a set of preferences means that each member of that set remains in the set after updating. In earlier work (Hanany and Klibanoff [2007]), we began to address this issue by identifying and characterizing the first dynamically consistent update rules satisfying closure for the maxmin expected utility (MEU) model of decision-making under ambiguity (Gilboa and Schmeidler [1989]).[2] In this paper, we are able to substantially generalize the approach taken there and characterize dynamically consistent update rules for essentially all continuous, monotonic preferences that are ambiguity averse in that they satisfy the uncertainty aversion axiom introduced in Schmeidler [1989] (i.e., have convex upper contour sets in utility space). Schmeidler's axiom (stated in Section 2.2) is very commonly used to describe ambiguity aversion in the literature and MEU is but one of many mod-

[1] See e.g., Proposition 6 in Hanany and Klibanoff [2007] or Proposition 3.1 in Section 2.2.

[2] Epstein and Schneider [2003] had previously investigated updating of MEU preferences using a stronger notion of dynamic consistency (see our discussion of recursive dynamic consistency in Section 6) and, as a result, to avoid a collapse to expected utility had to restrict attention to updating given a limited set of events (the so-called rectangular events).

els satisfying this property.[3] We consider several applications of our general result in some detail, characterizing dynamically consistent update rules satisfying closure for important specific models of ambiguity averse preferences: the smooth ambiguity preferences (Klibanoff, Marinacci and Mukerji [2005], henceforth KMM),[4] the variational preferences (Maccheroni, Marinacci and Rustichini [2006a], henceforth MMR) and regret-based models of ambiguity aversion, such as minimax regret with multiple priors (Hayashi [2008], Stoye [2008b]). We propose the first dynamically consistent update rules satisfying closure for these models.

What is a main difference in the implications of dynamic consistency when updating ambiguity averse preferences rather than standard preferences? Consider Ellsberg's three-color example, in which bets are made over the color of a ball drawn randomly from an urn with 90 balls, of which 30 are black (B) and the remaining 60 are somehow divided between red (R) and yellow (Y). Taking triples $(u_B, u_R, u_Y) \in \mathbb{R}^3$ to represent utility acts, i.e. state (color) contingent utility payoffs, the typical preference $(1, 0, 0) \succ (0, 1, 0)$ (betting on black rather than red) and $(0, 1, 1) \succ (1, 0, 1)$ (betting against black rather than against red) is inconsistent with expected utility but is consistent (and is one of the primary motivations for) many models of preferences under ambiguity. Let us introduce dynamics by supposing that after the ball is drawn from the urn, the decision maker (DM) is informed whether or not the ball is yellow.[5] The DM is allowed to condition her choice of bets on this information. Notice that this conditioning opportunity does not expand the feasible set of utility acts compared to the original problem of choosing between betting on B or on R – only the utility payoffs $(1, 0, 0)$ or $(0, 1, 0)$ (and their convex combinations, if randomization is considered) are achievable. The same applies to the choice between betting on the event $\{B, Y\}$ or on $\{R, Y\}$ – even with conditioning allowed, only convex combinations of $(1, 0, 1)$ or $(0, 1, 1)$ are feasible. Ex-ante, then, the static and dynamic choice problems are the same. A DM with the typical Ellsberg preferences is dynamically consistent in this problem if $(1, 0, 0)$ is chosen over $(0, 1, 0)$ when the dynamic problem is played out (i.e., conditional on the event $\{B, R\}$, betting on B is chosen over betting on R)

[3] For alternative notions of ambiguity aversion, see Epstein [1999] and Ghirardato and Marinacci [2002].

[4] See also Ergin and Gul [2009], Nau [2006], Neilson [2009], and Seo [2009] for related preference analyses and Hansen [2007] for an application to macroeconomic risk. While smooth ambiguity preferences may display a full range of ambiguity attitudes, we focus exclusively on the ambiguity averse subset of these preferences.

[5] Such dynamic extensions of Ellsberg have been considered before, e.g., Epstein and Schneider [2003], Hanany and Klibanoff [2007].

and $(0,1,1)$ is chosen over $(1,0,1)$ when the dynamic problem is played out (i.e., conditional on the event $\{B,R\}$, betting on $\{R,Y\}$ is chosen over betting on $\{B,Y\}$). Dynamic consistency creates a problem for the usual methods of updating preferences. To see this, observe that a choice of $(1,0,0)$ over $(0,1,0)$ conditional on $\{B,R\}$ from a feasible set consisting of this pair would directly conflict with a choice of $(0,1,1)$ over $(1,0,1)$ conditional on $\{B,R\}$ from a feasible set consisting of this latter pair. The conflict results because, when restricted to the event $\{B,R\}$, each pair involves exactly the choice of $(1,0)$ versus $(0,1)$. Therefore, to achieve dynamic consistency of the prototypical ambiguity averse preferences while ruling out updated preferences that give weight to unrealized events, updating must depend on the ex-ante contingent plan and/or feasible set of acts.

Why is dynamic consistency important enough to justify the additional complications it necessitates under ambiguity aversion? The fundamental reason is that dynamically consistent updating results in higher (ex-ante) welfare than any other form of updating. If a DM could choose an update rule, it would always be optimal to choose a dynamically consistent one. This is a strong justification for investigating dynamically consistent updating – such update rules are precisely what a DM should use. In sum, characterizing dynamically consistent update rules is a way to characterize optimal update rules.

There are two other approaches to modeling ambiguity averse preferences in dynamic settings. One approach deals with recursive extensions (e.g., Epstein and Schneider [2003], MMR [2006b], KMM [2009], Hayashi [2009]), while the other posits dynamic inconsistency and adopts (either explicitly or implicitly) assumptions, such as backward induction (e.g., Siniscalchi [2009]) or naive ignorance of the inconsistency, to pin down behavior. We compare them with our approach in some detail in Section 6, but want to mention here that in both of these approaches updating is independent of the ex-ante contingent plan and feasible set of acts, thus these approaches cannot deliver dynamic consistency for many events.

The fact that dynamically consistent updating of preferences that are not expected utility requires dependence on more than just the conditioning event was recognized by Machina [1989] and McClennen [1990]. Such updating is referred to as non-consequentialist. Machina [1989] proposed a dynamically consistent update rule in the context of preferences over objective lotteries, and this was later adapted by Machina and Schmeidler [1992] to satisfy dynamic consistency and closure for probabilistically sophisticated preferences over acts.[6]

[6] They show, in this latter context, that the rule is equivalent to updating beliefs by Bayes'

Capturing ambiguity averse behavior, however, requires models that go beyond probabilistic sophistication. We would like to find, for a variety of such models, dynamically consistent update rules that satisfy closure. In an early contribution in this direction, Eichberger and Grant [1997] succeed in characterizing a class of preferences quadratic in probabilities that allows for ambiguity aversion and is closed under Machina-Schmeidler updating. Why closure? Economists often choose to work with specific models because they are tractable and parsimoniously capture desired behavioral features (frequently expressed through axioms on preferences). A theory of updating will be most helpful when it allows the use of the same type of model throughout the problem under consideration. At a normative level, if properties of a model have strong appeal before updating, it would seem odd for this appeal to disappear after updating. Thus, desired properties imposed on the ex-ante preferences should also be imposed on the updated preferences, as this has the advantage of delivering the most relevant and interesting theory.

Unfortunately, the Machina-Schmeidler update rule fails closure for any set of preferences that includes non-probabilistically sophisticated members, as long as the preferences satisfy essentially the Savage [1954] axioms without the Sure-Thing Principle (Savage's P2) (see Epstein and Le Breton [1993]). Similarly, one can show that this rule is not closed for smooth ambiguity preferences or for variational preferences (both of which violate Savage's P4 as well as P2). In this paper, we successfully develop and characterize dynamically consistent update rules satisfying closure, for models ranging from imposing only ambiguity aversion, continuity and monotonicity to those requiring particular families of ambiguity averse preferences such as smooth ambiguity preferences, variational preferences, MEU preferences or multiplier preferences.

Why are these models of particular interest? The smooth ambiguity model simultaneously allows: separation of ambiguity attitude from perception of ambiguity; flexibility and non-constancy in ambiguity attitude; subjective and flexible perception of which events are ambiguous; the tractability of smooth preferences; and expected utility as a special case for any given ambiguity attitude. Furthermore, its modeling of ambiguity attitude allows tools and insights from the usual treatment of risk attitude under expected utility to be imported. The variational preferences model is an elegant and highly flexible model most notable for including and relating widely used models such as the MEU model and the robust-control/multiplier preferences of Hansen and

rule while simultaneously updating risk preferences in a specific, non-consequentialist way. See also Segal [1997] and Wakker [1997] who investigate alternative non-consequentialist update rules.

Sargent [2001] as special cases. The multiplier preferences of Hansen and Sargent [2001] model ambiguity or model uncertainty using relative entropy with respect to a reference subjective probability measure and have proven tractable in applications.

In addition to specializing our characterization of dynamically consistent updating for these models, we provide some more results of interest for each. Among rules satisfying closure for smooth ambiguity preferences, a rule we construct is shown to be the unique dynamically consistent rule that is a reweighting of Bayes' rule. We also show that it has other good properties, for example, it is commutative, in the sense that only the total information available and not the order in which it arrives is important for updating. Among rules satisfying closure for variational preferences, we show that a large set of dynamically consistent update rules imply that multiplier preferences should be updated by applying Bayes' rule to the reference measure.

What allows us to characterize dynamic consistency for such a broad class of preferences? The key is characterizing global (conditional) optimality of the ex-ante optimum. For convex feasible sets and preferences that have convex upper contour sets, this comes down to the existence of a measure in the intersection of two sets: the set of measures corresponding to hyperplanes supporting the conditional indifference curve at the ex-ante optimum and the set of measures supporting the relevant feasible set at the ex-ante optimum.[7]

1.1 An illustration of our approach: the smooth rule

To give some understanding of our results and the type of argument that underlies them, we briefly describe a dynamically consistent update rule proposed in this paper satisfying closure for ambiguity averse smooth ambiguity preferences and apply it to the dynamic Ellsberg example. Consider preferences having the following representation: beliefs are represented by a finite support probability measure μ over the set Δ of all probability measures π over a state space S, risk attitudes are represented by a von Neumann-Morgenstern expected utility function $u : X \to \mathbb{R}$ over a set of lotteries X, and ambiguity attitudes are represented by a strictly increasing, concave and differentiable function $\phi : u(X) \to \mathbb{R}$, such that for all acts f and h, $f \succsim h \iff \mathbb{E}_\mu \phi(\mathbb{E}_\pi u \circ f) \geq \mathbb{E}_\mu \phi(\mathbb{E}_\pi u \circ h)$, where $\mathbb{E}$ is the expectation operator. Think of an update rule defined by updating μ to a new belief, denoted $\mu_{E,g}$,

[7] In Hanany and Klibanoff [2007] on updating MEU, the former set was available "directly" in the form of the updated set of measures in the representation of conditional preferences. The generalization from this specific form is one of the key theoretical insights of the present paper and what enables the vastly expanded scope of application.

given any non-null event $E \subseteq S$ and act g that is unconditionally optimal within the feasible set of acts available to the DM. Imagine such a rule leading to conditional preferences represented by $\mathbb{E}_{\mu_{E,g}}\phi\left(\mathbb{E}_{\pi_E} u \circ f\right)$ where π_E denotes the Bayesian update of π given E. To understand when such a rule is dynamically consistent, consider the ex-ante optimization problem that g solves over some convex feasible set of acts B:

$$\max_{f \in B} \mathbb{E}_\mu \phi\left(\mathbb{E}_\pi u \circ f\right).$$

Since ϕ is concave and differentiable and the feasible set is convex and given an interiority condition, a necessary and sufficient condition for the optimality of g is that the utility gradient of the objective function when $f = g$ be normal to the feasible set at $u \circ g$. To satisfy dynamic consistency (essentially to remain optimal on E within the feasible set given that one can no longer change outcomes on E^c), g must also solve the conditional optimization problem:

$$\max_{f \in B \text{ s.t. } f = g \text{ on } E^c} \mathbb{E}_{\mu_{E,g}} \phi\left(\mathbb{E}_{\pi_E}\left(u \circ f\right)\right).$$

Again, a necessary and sufficient condition for g to solve this problem is that, on E, the utility gradient of the objective function when $f = g$ be normal to the set $u \circ \{f \in B \text{ s.t. } f = g \text{ on } E^c\}$ at $u \circ g$.

From the above, one can show that it is necessary and sufficient for dynamic consistency that the utility gradients of the unconditional and conditional objective functions are proportional on E when evaluated at $f = g$. The gradient of the unconditional objective function with respect to $u \circ f$ is

$$\mathbb{E}_\mu \phi'\left(\mathbb{E}_\pi\left(u \circ f\right)\right)\pi.$$

The gradient of the conditional objective function with respect to $u \circ f$ for $s \in E$ is

$$\mathbb{E}_{\mu_{E,g}} \phi'\left(\mathbb{E}_{\pi_E}\left(u \circ f\right)\right)\pi_E.$$

Thus such a rule is dynamically consistent if and only if

$$\mathbb{E}_{\mu_{E,g}} \phi'\left(\mathbb{E}_{\pi_E}\left(u \circ g\right)\right)\pi_E \propto \mathbb{E}_\mu \phi'\left(\mathbb{E}_\pi\left(u \circ g\right)\right)\pi \text{ restricted to } s \in E.$$

This argument is informative because it may be modified to work for more general models of preferences and more general notions of gradient. The important aspect of the gradient at the optimum is that it is normal to a hyperplane separating a feasible set from an upper contour set. Our most general characterization, applying to quasiconcave (ambiguity averse) preference models, exploits this by working directly at the level of separating hyperplanes and their normals.

An example of a dynamically consistent rule we propose in the smooth ambiguity setting is the *smooth rule* defined by setting

$$\mu_{E,g}(\pi) = \frac{\mu(\pi)\,\pi(E)\,\frac{\phi'(\mathbb{E}_{\pi}(u\circ g))}{\phi'\left(\mathbb{E}_{\pi_E}(u\circ g)\right)}}{\sum_{\hat{\pi}\in\Delta}\mu(\hat{\pi})\,\hat{\pi}(E)\,\frac{\phi'(\mathbb{E}_{\hat{\pi}}(u\circ g))}{\phi'\left(\mathbb{E}_{\hat{\pi}_E}(u\circ g)\right)}} \text{ if } \pi(E) > 0,$$

(and equals 0 otherwise). Notice that when the smooth rule is used to define $\mu_{E,g}$,

$$\begin{aligned} & \mathbb{E}_{\mu_{E,g}}\phi'(\mathbb{E}_{\pi_E}(u\circ g))\,\pi_E \\ \propto\ & \mathbb{E}_{\mu}\phi'(\mathbb{E}_{\pi_E}(u\circ g))\,\frac{\phi'(\mathbb{E}_{\pi}(u\circ g))}{\phi'(\mathbb{E}_{\pi_E}(u\circ g))}\pi(E)\pi_E \\ =\ & \mathbb{E}_{\mu}\phi'(\mathbb{E}_{\pi}(u\circ g))\,\pi \text{ for } s\in E \end{aligned}$$

verifying dynamic consistency. We will see that this rule has many nice properties.

A numerical example of the smooth rule serves to show how it generates dynamically consistent preferences in the dynamic Ellsberg example described above. Recall that the relevant state space is $S = \{B, R, Y\}$ corresponding to the three colors that might be drawn from the urn. We assume that the DM has smooth ambiguity preferences with μ a half-half distribution over the distributions $\pi^{\theta} \equiv (\frac{1}{3}, \frac{2}{3}\theta, \frac{2}{3}(1-\theta))$ for $\theta \in \left\{\frac{1}{3}, \frac{2}{3}\right\}$. This is consistent with the fact that the chance of drawing a black ball is known to be one-third and the ex-ante symmetry of the situation. Lotteries are monetary and evaluated by their expected value (i.e., u is the identity) and $\phi(x) = -e^{-\alpha x}$ where $\alpha > 0$ so that the DM displays constant absolute ambiguity aversion with coefficient α (see KMM [2005]). We can verify, using Jensen's inequality, that these preferences display the modal Ellsberg choices: $(1,0,0) \succ (0,1,0)$ and $(0,1,1) \succ (1,0,1)$. Now consider updating on the event that the ball is not yellow so that $E = \{B, R\}$. According to the smooth rule, conditional preferences are represented by

$$V_{E,g}(u\circ f) = \phi[\mathbb{E}_{\pi_E^{1/3}}(u\circ f)]\mu_{E,g}(\pi^{1/3}) + \phi[\mathbb{E}_{\pi_E^{2/3}}(u\circ f)]\mu_{E,g}(\pi^{2/3})$$

where $\mu_{E,g}(\pi^{\theta}) = \dfrac{\frac{\phi'\left(\mathbb{E}_{\pi^{\theta}}(u\circ g)\right)}{\phi'\left(\mathbb{E}_{\pi_E^{\theta}}(u\circ g)\right)}\mu(\pi)\pi^{\theta}(E)}{\sum_{\hat{\theta}\in\left\{\frac{1}{3},\frac{2}{3}\right\}}\frac{\phi'\left(\mathbb{E}_{\pi^{\hat{\theta}}}(u\circ g)\right)}{\phi'\left(\mathbb{E}_{\pi_E^{\hat{\theta}}}(u\circ g)\right)}\mu(\pi)\pi^{\hat{\theta}}(E)}$ and $\pi_E^{\theta} = (\frac{1}{1+2\theta}, \frac{2\theta}{1+2\theta}, 0)$ for $\theta \in \left\{\frac{1}{3}, \frac{2}{3}\right\}$.

For the feasible set $B^1 = co\{(1,0,0),(0,1,0)\}$ corresponding to the first Ellsberg choice pair, $u \circ g = (1,0,0)$.[8] Applying the smooth rule, $\mu_{E,(1,0,0)}(\pi^\theta) = \frac{(3+6\theta)e^{\alpha(\frac{1}{1+2\theta})}}{5e^{\frac{3\alpha}{5}}+7e^{\frac{3\alpha}{7}}}$. Similarly, for the feasible set $B^2 = co\{(1,0,1),(0,1,1)\}$ corresponding to the second Ellsberg choice pair, $u \circ g = (0,1,1)$ and $\mu_{E,(0,1,1)}(\pi^\theta) = \frac{(3+6\theta)e^{\alpha(\frac{2\theta}{1+2\theta})}}{5e^{\frac{2\alpha}{5}}+7e^{\frac{4\alpha}{7}}}$. Notice that when $u \circ g = (1,0,0)$ the updated belief puts more weight on the measure giving higher conditional probability of black than it does when $u \circ g = (0,1,1)$. If, for example, $\alpha = 1$, then $\mu_{E,(1,0,0)}(\pi^{\frac{1}{3}}) \approx 0.4588$, while $\mu_{E,(0,1,1)}(\pi^{\frac{1}{3}}) \approx 0.3757$. What is an intuition for why these weights are different? When facing problem B^1, preferring $(1,0,0)$ reveals that the DM must in some sense be assigning more prior weight to the scenario $\theta = \frac{1}{3}$ where red has a low probability relative to black. Upon updating, this will mean that such scenarios continue to get more weight than they would under symmetric prior weighting. When facing problem B^2, however, preferring $(0,1,1)$ reveals that the DM must in the same sense be assigning less prior weight to the scenario $\theta = \frac{1}{3}$. After updating then, this scenario will continue to get less weight than under symmetric prior weighting. Further calculation shows that with smooth rule updating, as dynamic consistency requires, $(1,0,0)$ is conditionally optimal within B^1 and $(0,1,1)$ is conditionally optimal within B^2.

By way of contrast, consider standard Bayesian updating in this example, given by $\mu_E(\pi^\theta) = \frac{1+2\theta}{4}$ for $\theta \in \left\{\frac{1}{3}, \frac{2}{3}\right\}$, so that $\mu_E(\pi^{\frac{1}{3}}) = \frac{5}{12} \approx 0.4167$. In comparing $(1,0,0)$ and $(0,1,0)$, observe that the distribution of conditional expected utilities for $(1,0,0)$, ($\frac{3}{5}$ w.prob. $\frac{5}{12}$; $\frac{3}{7}$ w.prob. $\frac{7}{12}$), is a mean-preserving spread of the distribution ($\frac{2}{5}$ w.prob. $\frac{5}{12}$; $\frac{4}{7}$ w.prob. $\frac{7}{12}$) for $(0,1,0)$. Since ϕ is strictly concave, this implies $(0,1,0) \succ_E (1,0,0)$, and thus Bayesian updating violates dynamic consistency for such a DM.

What happens if we use the Machina-Schmeidler update rule in this example? In our context, their rule is equivalent to saying that if ex-ante preferences are represented by the functional $V(u \circ f)$ for all acts f, then updated preferences after learning E, given that g was unconditionally optimal within the feasible set of acts B, are represented by $V(u \circ f_E g)$ for all acts f where $f_E g$ denotes the act equal to f on E and g on E^c. In the example, $V(u \circ f) \equiv \mathbb{E}_\mu \phi(\mathbb{E}_\pi u \circ f)$ and thus the updated preferences are represented by $V(u \circ f_E g) \equiv \mathbb{E}_\mu \phi(\mathbb{E}_\pi u \circ f_E g)$. Substituting yields

$$-\frac{1}{2}e^{-\alpha(\frac{1}{3}f(B)+\frac{2}{9}f(R))} - \frac{1}{2}e^{-\alpha(\frac{1}{3}f(B)+\frac{4}{9}f(R))}$$

[8] We use co to denote the convex hull operator. We take the convex hull of the available acts to reflect the fact that the DM may randomize among them.

when $g = (1,0,0)$ and

$$-\frac{1}{2}e^{-\alpha(\frac{1}{3}f(B)+\frac{2}{9}f(R)+\frac{4}{9})}-\frac{1}{2}e^{-\alpha(\frac{1}{3}f(B)+\frac{4}{9}f(R)+\frac{2}{9})}$$

when $g = (0,1,1)$. One can show these updated preferences are *not* smooth ambiguity preferences over acts.[9] Therefore the Machina-Schmeidler rule, although dynamically consistent, does not allow one to talk about updating beliefs μ, nor does it produce updated preferences that satisfy the properties of the smooth ambiguity model (something that the ex-ante preferences do satisfy), so it cannot even be described as updating a combination of tastes, ϕ, and beliefs, μ.

The rest of the paper is organized as follows. Section 2 contains an exposition of the framework and notation, formally defines dynamic consistency and presents our main result characterizing dynamic consistency in ambiguity averse preference models. This result is applied in Section 3 to characterize dynamically consistent update rules satisfying closure for the smooth ambiguity model. We then derive a novel update rule, the smooth rule mentioned above, which is shown to be the unique dynamically consistent rule among a large class of rules that may be expressed as reweightings of Bayes' rule. We describe additional desirable properties of this update rule, including invariance to the order in which information arrives (commutativity) and a strict version of dynamic consistency. We then show several impossibility results when strengthening dynamic consistency. Section 4 applies the general characterization result to the representation developed in Cerreia-Vioglio, Maccheroni, Marinacci and Montrucchio [2008] (henceforth CMMM) and to rules satisfying closure for variational preferences. We also construct some consistent update rules for variational preferences. The result on dynamically consistent updating of multiplier preferences is here as well. Section 5 explains how our results extend to updating regret-based models, including minimax regret with multiple priors. Related literature, including recursive approaches and dynamic inconsistency, is discussed in Section 6. The final section contains a brief conclusion. All proofs not appearing in the main text are collected in Appendix A.

[9] Not even if restricted to acts of the form $f_E g$.

2 Dynamic consistency and updating ambiguity averse preferences

2.1 Setting and notation

Consider an Anscombe-Aumann framework [1963], where X is the set of all simple (i.e., finite-support) lotteries over a set of consequences Z, S is a finite set of states of nature and $\mathcal{A}$ is the set of all acts, i.e., functions $f : S \to X$.[10] Abusing notation, $x \in X$ is also used to denote the constant act for which $\forall s \in S$, $f(s) = x$.

Let $\mathcal{B}$ denote the set of all non-empty subsets of acts $B \subseteq \mathcal{A}$ such that B is convex (with respect to the usual Anscombe-Aumann mixtures) and compact (according to the norm taking the maximum over states and Euclidean distance on lotteries in X). Elements of $\mathcal{B}$ are considered feasible sets and their convexity could be justified, for example, by randomization over acts.[11] Compactness is needed to ensure the existence of optimal acts.

Let $\mathcal{P}$ denote the set of preference relations (i.e., complete and transitive binary relations) $\succsim$ on the acts $\mathcal{A}$ satisfying the following quite standard assumptions: preferences (i) are non-degenerate (i.e., $f \succ h$ for some acts f, h), (ii) are mixture continuous (i.e., the sets $\{\alpha \in [0,1] \mid \alpha f + (1-\alpha)i \succsim h\}$ and $\{\alpha \in [0,1] \mid h \succsim \alpha f + (1-\alpha)i\}$ are closed), (iii) are weakly monotonic (i.e., $f(s) \succsim h(s)$ for all $s \in S$ implies $f \succsim h$) and (iv) when restricted to constant acts, obey the von Neumann-Morgenstern expected utility axioms.

Given $V : \mathbb{R}^S \to \mathbb{R}$ such that (i) V is non-constant and continuous and (ii) $a \geq b$ implies $V(a) \geq V(b)$ for all $a, b \in \mathbb{R}^S$, and given a non-constant von Neumann-Morgenstern utility function $u : X \to \mathbb{R}$, the functional $V(u \circ \cdot) : \mathcal{A} \to \mathbb{R}$ represents a unique preference in $\mathcal{P}$. Without loss of generality, fix a particular normalization of u throughout the analysis. Consider a pair (V, u) where V and u satisfy these conditions. Let Ψ denote the set of all such pairs. Each element of Ψ is thus associated with a unique preference $\succsim \in \mathcal{P}$. Conversely, given any preference in $\mathcal{P}$ there exists a unique u and at least one V such that $(V, u) \in \Psi$ represents it.[12] For any $\succsim \in \mathcal{P}$, it will be useful to identify the set of interior acts, $\text{int}(\mathcal{A}) \equiv \{f \in \mathcal{A} \mid u \circ f$ is in the interior of

[10] The lottery structure, *per se*, is not essential. Extension of all of our analysis to the case where X is a general convex subset of a vector space is straightforward.

[11] The important implication of convexity with respect to mixtures will be that, if $u : X \to \mathbb{R}$ is a von Neumann-Morgenstern utility function, feasible sets have a convex image in utility space, i.e., $\{u \circ f \mid f \in B\}$ is convex. Alternatively, one could assume the latter directly.

[12] Existence may be shown along the lines of Lemma 67 in the Appendix of CMMM [2008].

$u \circ \mathcal{A}\}$. These are the acts that do not give a best or a worst lottery in any state. If utility is unbounded, then all acts are interior. Sometimes we refer to utility acts, by which we mean elements of $u \circ \mathcal{A}$.

For an event $E \subseteq S$, let $\Delta(E)$ denote the set of all probability measures on 2^S giving weight 0 to E^c (the complement of E in S). Let $\Delta \equiv \Delta(S)$. For any $q \in \Delta$ with $q(E) > 0$, we denote by $q_E \in \Delta(E)$ the measure obtained through Bayesian conditioning of q on E (i.e., for $F \subseteq S$, $q_E(F) = \frac{q(E \cap F)}{q(E)}$). For $E \subseteq S$ and $f, h \in \mathcal{A}$, we use $f_E h$ to denote the act equal to f on E and h on E^c. Say that an event E is Savage-null [1954] according to $\succsim$ if, for all acts $f, h, i \in \mathcal{A}$, $h_E f \sim i_E f$. Say that an event E is non-null according to $\succsim$ if, for all acts $f, h, i \in \mathcal{A}$, $h(s) \succ i(s)$ for all $s \in E$ implies $h_E f \succ i_E f$.[13]

Denote by $\mathcal{Q}$ the set of all quadruples $(\succsim, E, g, B)$ where $\succsim \in \mathcal{P}$, E is a non-null event, and $g \in \text{int}(\mathcal{A})$[14] is an act optimal according to $\succsim$ within a feasible set $B \in \mathcal{B}$ (i.e., $g \succsim f$ for all $f \in B$). $\mathcal{Q}$ is the largest domain that we consider in defining an update rule. It will also be useful to be able to restrict an update rule to domains that are strict subsets of $\mathcal{Q}$, for example if we want to consider update rules that apply only to a specific model of preferences, only to specific events or only to feasible sets with unique utility optima. To this end, say that a non-empty $\mathcal{D} \subseteq \mathcal{Q}$ is a *domain* if $(\succsim, E, g, B) \in \mathcal{D}$ implies $(\succsim, E, g', B') \in \mathcal{D}$ for each g', B' such that $(\succsim, E, g', B') \in \mathcal{Q}$ and $u \circ g'$ is the unique maximizer of V over $u \circ B'$. An update rule takes a quadruple and produces a representation of preferences. These preferences are conditional in the sense that they make E^c a Savage-null event. In this paper, we restrict attention to update rules that preserve risk preferences (by leaving u unchanged). Formally:

Definition 2.1 *An* update rule *is a function U mapping elements of a domain $\mathcal{D}$ to a codomain consisting of pairs $(V_{E,g,B}, u) \in \Psi$ such that, after updating, E^c is Savage-null and E remains non-null.*

We use $\succsim_{E,g,B}$ to denote the preferences represented by $V_{E,g,B}(u \circ \cdot)$. Such a conditional preference is viewed as governing choice upon the realization of

[13] This notion of non-null excludes some events that are not Savage-null – under expected utility the two concepts agree, but they need not more generally. Throughout the paper we will restrict attention to updating on non-null events. This stronger, yet still quite mild, sense of "events that matter" is useful in avoiding extraneous complications related to what should be required of update rules if considering events sometimes given positive weight and sometimes not.

[14] The interiority assumption is mild and allows the simplification of a number of characterizations provided below by avoiding the multiplicity of supporting hyperplanes that can occur on the boundaries (if any) of the utility act space.

the conditioning event E.[15] Recall from the Introduction that, generally, allowing dependence on g and/or B is necessary for dynamic consistency under ambiguity aversion. In some specific applications, we may be able to maintain consistency while imposing independence from either g or B (but not both). For example, when considering update rules for smooth ambiguity preferences (in Section 3) we limit attention to rules depending on B only through g. Another example, as we explicitly remark after Theorem 2.1 in the next subsection, is that all of our results hold when one limits attention to feasible sets where $u \circ g$ is unique, an environment where g is unimportant given B.

2.2 Dynamically consistent update rules

Dynamic consistency of one form or another has often been put forward as a rationality criterion and thus, from a normative point of view, it is important to identify rules that satisfy some version of this property. Moreover, from the normative point of view, optimal acts are the most important for dynamic consistency to be satisfied on because those are the acts that are chosen. Under dynamic consistency, optimal contingent plans remain optimal according to updated preferences. Therefore dynamically consistent updating necessarily maximizes ex-ante welfare among all update rules. Another normative argument in favor of dynamic consistency is the dominated choice argument described in the context of an example in Section 6.2. Dynamic consistency also makes it easier to describe an individual planning ahead and to make welfare statements in dynamic models because it is only under dynamic consistency that ex-ante and ex-post preferences agree on optimal feasible choices. Our axiom requires that, for each feasible set of acts, B, the update rule guarantees, for any preference and any non-null conditioning event E, if g is chosen within B, it remains optimal conditionally. Formally,

Axiom 2.1 ***DC*** *(Dynamic Consistency). For any* $(\succsim, E, g, B) \in \mathcal{D}$*, if* $f \in B$ *with* $f = g$ *on* E^c*, then* $g \succsim_{E,g,B} f$*.*

This formalization of dynamic consistency appeared in Hanany and Klibanoff [2007] and is related to ideas and concepts in earlier literature such as Machina [1989], McClennen [1990], Segal [1997] and Grant, Kajii and Polak [2000].[16] Recall from the Introduction that dependence of conditional preference on g

[15] For a discussion of the observability of conditional preferences see Appendix A.2 of Hanany and Klibanoff [2007].

[16] For a more detailed discussion of the relation to other definitions, see Hanany and Klibanoff [2007] and the discussion of recursive approaches in Section 6.1 of this paper.

and B is needed, in general, to attain dynamic consistency under ambiguity aversion. With ambiguity aversion, the weighting of states supporting the choice of g from B is typically different than that supporting the choice of some g' optimal in B'. It is quite natural, then, that g and/or B also affect the manner in which new information changes the DM's view of these uncertainties through updating.

Observe that conditional optimality of g is checked against all $f \in B$ such that f and g agree on E^c. Why check conditional optimality only against these acts? Dynamic consistency is relevant only *ceteris paribus*, i.e., when exactly the same consequences occur on E^c. To make clear why this is reasonable, consider an environment where the DM has a fixed budget to allocate across bets on various events. It would be nonsensical to require that the ex-ante optimal allocation of bets remained better than placing all of one's bets on the realized event. This justifies the restriction of the conditional comparisons to acts that agree on E^c.

Is there a way to describe the set of dynamically consistent update rules for ambiguity averse preferences? We give a characterization of dynamic consistency that applies to update rules that produce representations of preferences satisfying Schmeidler's [1989] uncertainty aversion axiom. Schmeidler's axiom says that $f \succsim h$ implies $\alpha f + (1-\alpha)h \succsim h$ for $\alpha \in [0,1]$, and is equivalent in our setting to the quasiconcavity of V (i.e., preferences are convex in utility space). All smooth ambiguity preferences that are ambiguity averse in the sense of KMM [2005] satisfy this condition, as do all variational preferences, and the general uncertainty averse preferences studied by CMMM [2008], so this includes a very large and interesting class of models.[17] Formally, then, we restrict attention to update rules satisfying the following property that we have verbally described above:

(i) (uncertainty/ambiguity aversion) all $V_{E,g,B}$ in the codomain of the update rule are quasiconcave.

Letting $\mathcal{Y}$ denote the family of all update rules satisfying the above property, $\mathcal{Y}$ will be the largest family for which we describe the rules that are dynamically consistent.

[17] What about preferences that are not quasiconcave (i.e., that violate Schmeidler's [1989] axiom)? The methods we describe will still apply if one is willing to update these so that the conditional preferences satisfy Schmeidler's axiom. However, if one wants to allow non-quasiconcavity after updating, while the existence of dynamically consistent rules is easy to show, the general characterization of such rules would suffer from all of the substantial difficulties in characterizing global maxima of non-quasiconcave problems, and is therefore beyond the scope of this paper.

The following definitions are key in characterizing conditional optimality and thus dynamic consistency. We define two sets of probability measures. The first set will correspond to hyperplanes supporting the conditional indifference curve at a given point. The second will correspond to hyperplanes supporting the feasible set of acts at a given point.

Definition 2.2 *For a quasiconcave* $V : \mathbb{R}^S \to \mathbb{R}$*, act* $h \in \mathcal{A}$ *and event* $E \subseteq S$*, define the* measures supporting the conditional indifference curve at h on E *to be*

$$T_{E,h}(V) = \left\{ \begin{array}{c} q \in \Delta(E) \mid \int (u \circ f) dq > \int (u \circ h) dq \\ \text{for all } f \in \mathcal{A} \text{ such that } V(u \circ f) > V(u \circ h) \end{array} \right\}.$$

If we think of V as representing conditional preferences over utility acts given the event E, then the measures in $T_{E,h}$ are normals to hyperplanes supporting the conditional indifference curve containing $u \circ h$. The modifier "conditional" is important – if V is such that E^c matters, then $T_{E,h}(V)$ will typically be empty. Thus the name "measures supporting the *conditional* indifference curve at h on E."

Definition 2.3 *For an act* $h \in \mathcal{A}$*, event* $E \subseteq S$ *and feasible set* $B \in \mathcal{B}$*, define the* measures supporting the conditional optimality of h in B *to be* $Q_E^{E,h,B}$ *where*

$$Q^{E,h,B} = \left\{ \begin{array}{c} q \in \Delta \mid q(E) > 0 \text{ and} \int (u \circ h) dq \geq \int (u \circ f) dq \\ \text{for all } f \in B \text{ with } f = h \text{ on } E^c \end{array} \right\},$$

and $Q_E^{E,h,B}$ *is given by the Bayesian updates on* E *of measures in* $Q^{E,h,B}$*.*

There are two reasons why calling these sets "measures supporting the conditional optimality of h" makes sense. The first is obvious: if we consider a conditional expected utility preference with measure $q_E \in Q_E^{E,h,B}$, then according to such a preference, h will be conditionally optimal in the set $\{f \in B \mid f = h \text{ on } E^c\}$. Similarly, if h can't be conditionally optimal because it is dominated on E by an element of $\{f \in B \mid f = h \text{ on } E^c\}$ then $Q_E^{E,h,B}$ is empty. The second reason is deeper: as we will show, the existence of a measure in $Q_E^{E,h,B}$ that is a measure supporting the conditional indifference curve at h on E is equivalent to the conditional optimality of h.

We can now state our characterization of dynamically consistent updating.

Theorem 2.1 $U \in \mathcal{Y}$ *is dynamically consistent if and only if* $T_{E,g}(V_{E,g,B}) \cap Q_E^{E,g,B} \neq \emptyset$ *for all* $(\succsim, E, g, B) \in \mathcal{D}$*.*

This theorem provides a test for dynamic consistency of an update rule for ambiguity averse preferences. First, given an event E, a feasible set B, and an act g unconditionally optimal within B, one can calculate $Q_E^{E,g,B}$. If the image in utility space of the set of feasible acts agreeing with g on E^c is smooth at $u \circ g$, $Q_E^{E,g,B}$ will be a singleton and can be found by differentiation. Second, apply the candidate update rule to produce a representation of the updated preferences, $V_{E,g,B}$. Third, calculate $T_{E,g}(V_{E,g,B})$. If, for example, $V_{E,g,B}$ is smooth at $u \circ g$ this is no more complicated than usual differentiation. Finally, see if the two sets intersect.

The intuition is quite familiar from convex optimization: g will be conditionally optimal within $\{f \in B \mid f = g \text{ on } E^c\}$ if and only if there exists a hyperplane containing $u \circ g$ that separates $u \circ \{f \in B \mid f = g \text{ on } E^c\}$ from the utility acts conditionally better than $u \circ g$. Each hyperplane in utility space may be uniquely associated with a probability measure on the state space. The proof in Appendix A shows that the measures in $T_{E,g}(V_{E,g,B}) \cap Q_E^{E,g,B}$ exactly correspond to such separating hyperplanes.

Remark 2.1 Can we say anything special about update rules that do not depend on g (given E and B)? Using the characterization in Theorem 2.1, we can show that such dynamically consistent update rules generally exist. However, they may be very restrictive in the ambiguity they permit after updating. For an illustration of this in the context of MEU, see Example 1 in Hanany and Klibanoff [2007]. As an alternative, less restrictive, approach to independence from g, one could limit attention to feasible sets having a unique optimum. The characterization in Theorem 2.1 remains true when restricted to these sets. However, for later, more specialized results in the paper to go through, a slightly larger collection of feasible sets is needed. In this regard, note that one domain $\mathcal{D}$ to which Theorem 2.1 applies limits attention to feasible sets for which $u \circ g$ is uniquely optimal in $u \circ B$ (i.e., there may be multiple optima in B, but each generates the same utility act). If we take this domain and replace "$f = g$ on E^c" with "$f(s) \sim g(s)$ on E^c" in the definitions of dynamic consistency and of $Q_E^{E,h,B}$, we arrive at an application of Theorem 2.1 where the choice of g is unimportant. Specifically, given B allowed by this domain, further dependence of the update rule on g would never affect dynamic consistency. With this domain restriction and the above-mentioned replacement, all of the subsequent results in the paper and their proofs (with appropriate replacements of $=$ with $\sim$ if necessary) remain true.
What about update rules that depend on B only through g? In this case, kinks in preferences can cause difficulties in achieving dynamic consistency (see e.g., Proposition 7 in Hanany and Klibanoff [2007]). In Section 3, we explore such

rules in detail for the (non-kinked) smooth ambiguity model.

Above, we suggested that calculation of $T_{E,g}(V_{E,g,B})$ amounted to differentiation if $V_{E,g,B}$ is smooth at $u \circ g$. In fact, the set $T_{E,g}(V_{E,g,B})$ is closely related to a general differential notion developed in Greenberg and Pierskalla [1973] for use in quasiconvex analysis. The definition below is adapted from their paper and, up to normalization, simplifies, for monotone and quasiconcave functions, to the usual gradient at points of continuous differentiability whenever the gradient is non-zero.[18]

Definition 2.4 *Given a function* $I : \mathbb{R}^S \to \mathbb{R}$*, its* Greenberg-Pierskalla superdifferential at a *is*

$$\partial^{\star} I(a) = \left\{ \begin{array}{c} r : 2^S \to \mathbb{R} \mid r \text{ is bounded and additive and} \\ \left\{ b \in \mathbb{R}^S : I(b) > I(a) \right\} \subseteq \left\{ b \in \mathbb{R}^S : \int b dr > \int a dr \right\} \end{array} \right\}.$$

The next result says that we can replace $T_{E,g}(V_{E,g,B})$ in our characterization by the Greenberg-Pierskalla superdifferential of $V_{E,g,B}$ at $u \circ g$ (or, since it is the only relevant part, only the conditional probability measures in this superdifferential):

Proposition 2.1 $U \in \mathcal{Y}$ *is dynamically consistent if and only if*

$$\partial^{\star} V_{E,g,B}(u \circ g) \cap Q_E^{E,g,B} \neq \emptyset \text{ for all } (\succsim, E, g, B) \in \mathcal{D}.$$

This result is helpful in that it facilitates the use of gradients when available. In the next section, we apply Proposition 2.1 to characterize dynamically consistent update rules satisfying closure for ambiguity averse smooth ambiguity preferences. In the subsequent section, we apply our results to do the same for the CMMM [2008] representation and variational preferences. Recall from the Introduction, when we say that an update rule satisfies closure for a set of preferences, we mean that both unconditional and conditional preferences are restricted to that set. For example, if one wishes to restrict ex-ante preferences to be variational preferences (i.e., to obey the axioms in MMR [2006a]) then it will be most desirable to have an update rule that generates only variational preferences. Similarly, if one thinks it appropriate or useful to further restrict to multiplier preferences (i.e., to obey the axioms in Strzalecki [2008]), then an update rule defined on and generating only multiplier preferences will be of interest. No dynamically consistent update rules have been previously proposed satisfying closure for either smooth ambiguity or variational preferences.

[18] Since we consider updating only on non-null events, points with zero gradients do not play a role in our analysis. We are grateful to Fabio Maccheroni for bringing the Greenberg-Pierskalla superdifferential to our attention and his substantial help and suggestions regarding the proposition on this and its proof.

3 Dynamically consistent updating of smooth ambiguity preferences

In this section we investigate dynamically consistent update rules satisfying closure for ambiguity averse smooth ambiguity preferences. Let $\mathcal{P}^{SM}$ denote the set of smooth ambiguity preference relations over $\mathcal{A}$ (KMM [2005]).[19] For any preference $\succsim \in \mathcal{P}^{SM}$, there exists a countably additive probability measure, μ, over the set Δ, a von Neumann-Morgenstern expected utility function, $u : X \to \mathbb{R}$ and a strictly increasing function $\phi : u(X) \to \mathbb{R}$, such that $\forall f, h \in \mathcal{A}$, $f \succsim h \iff \mathbb{E}_\mu \phi(\mathbb{E}_\pi u \circ f) \geq \mathbb{E}_\mu \phi(\mathbb{E}_\pi u \circ h)$, where $\mathbb{E}$ is the expectation operator. We will assume that ϕ is differentiable and concave. According to KMM, concavity of ϕ reflects ambiguity aversion. If μ is degenerate, we adopt the convention that ϕ is the identity (with a degenerate μ, ϕ is irrelevant for $\succsim$). For simplicity, we will also assume that the support of μ is finite. Let $\text{supp}(\mu)$ denote this support.

Say that $(V, u) \in \Psi$ is an ambiguity averse smooth ambiguity representation whenever $V(a) = \mathbb{E}_\mu \phi(\mathbb{E}_\pi a)$ for all $a \in \mathbb{R}^S$. Such a V is completely determined by specifying μ and ϕ. Let Ψ^{SM} denote the set of all such (V, u). Each element of Ψ^{SM} is thus associated with a preference $\succsim \in \mathcal{P}^{SM}$. To update ambiguity averse smooth ambiguity preferences we consider rules defined on an appropriate subset of quadruples $(\succsim, E, g, B)$ – specifically, $\mathcal{D}^{SM}$ is the set of all elements of $\mathcal{Q}$ such that $\succsim \in \mathcal{P}^{SM}$. Note that the only events in $\mathcal{D}^{SM}$ are those for which $\sum_{\pi \in \text{supp}(\mu)} \mu(\pi) \pi(E) > 0$ as they are the non-null events in the sense defined in Section 2.1.

We consider update rules $\mathcal{U} \subseteq \mathcal{Y}$ satisfying:

(i) (closure for ambiguity averse smooth ambiguity preferences) the domain is $\mathcal{D}^{SM}$ and the codomain is Ψ^{SM},

(ii) (preservation of ambiguity attitude) $\phi_{E,g,B} = \phi$, and

(iii) (independence from feasible sets) $V_{E,g,B_1} = V_{E,g,B_2}$ for all $(\succsim, E, g, B_1), (\succsim, E, g, B_2) \in \mathcal{D}^{SM}$.

Property (i) reflects the scope of this particular application of our results – we want to update within the class of ambiguity averse smooth ambiguity

[19] The framework in Klibanoff, Marinacci and Mukerji [2005] is not precisely an Anscombe-Aumann framework as the acts there need not have lottery consequences. However their state space S is assumed to be a product space with an ordinate that is $[0, 1]$ and is treated as a randomizing device. Thus, their theoretical development could be easily adapted to the setting here. See Seo [2009] for an alternative axiomatization and model.

preferences. Since this class of preferences separates tastes (risk and ambiguity attitudes) from beliefs, property (ii) extends the common notion that new information should affect beliefs but not affect tastes to encompass ambiguity attitudes in addition to the previously assumed preservation of risk attitudes. Property (iii) is assumed mainly because update rules satisfying it are in a natural sense simpler than those requiring dependence on B through a channel other than g, and because we can – we shall see that even with this restriction interesting dynamically consistent update rules exist for this class of preferences. The fact that we are able to summarize dependence on the feasible set through g stems from the smoothness of the preferences under consideration.[20]

For these preferences, a given $\succsim$, and even a given V, may be represented by multiple (ϕ, μ) pairs. In KMM, (ϕ, μ) are pinned down by considering more preference information than simply $\succsim$. When we characterize dynamically consistent update rules in this section, we fix a function selecting a particular (ϕ, μ) representation for each $\succsim$ that is to be updated. Note that one can imagine the selection function being based on information outside the present framework, for example the preferences over second-order acts used and discussed in KMM to pin down the characteristics of ϕ and μ. Given such a selection function, a formula for updating (ϕ, μ) determines a well-defined update rule. Since update rules in $\mathcal{U}$ are required to satisfy properties (ii) and (iii) above, for the remainder of this section, an update rule in $\mathcal{U}$ will be identified with a mapping taking μ to an updated measure $\mu_{E,g}$, with the selection rule implicitly fixed in the background.

Given $V(a) = \mathbb{E}_{\mu}\phi(\mathbb{E}_{\pi}a)$, an updated representation $V_{E,g}(a)$ could be written as $\mathbb{E}_{\breve{\mu}_{E,g}}\phi(\mathbb{E}_{\pi}a)$, where $\breve{\mu}_{E,g}$ satisfies $\breve{\mu}_{E,g}(\Delta(E)) = 1$. Notice that such a $V_{E,g}$ may also be written as $\mathbb{E}_{\mu_{E,g}}\phi(\mathbb{E}_{\pi_E}a)$ where $\sum_{\left\{\hat{\pi}\in\text{supp}(\mu_{E,g})\mid\hat{\pi}_E=\pi\right\}}\mu_{E,g}(\hat{\pi}) = \breve{\mu}_{E,g}(\pi)$ for all $\pi \in \Delta(E)$. We find it convenient to work with such representations. Any $\mu_{E,g}$ determines a unique $\breve{\mu}_{E,g}$ in this way. Thus, to specify an update rule in $\mathcal{U}$ we can and will specify conditional measures $\mu_{E,g}$ rather than $\breve{\mu}_{E,g}$. Of course this relationship is not one-to-one – all $\mu_{E,g}$ corresponding to the same $\breve{\mu}_{E,g}$ identify the same update rule in $\mathcal{U}$ (i.e., result in the same $V_{E,g}$).

An example of an update rule in $\mathcal{U}$ is Bayes' rule (i.e., Bayesian updating of beliefs). Given $(\succsim, E, g, B) \in \mathcal{D}^{SM}$, applying Bayes' rule produces $\mu_{E,g}$ where

$$\mu_{E,g}(\pi) = \frac{\mu(\pi)\,\pi(E)}{\sum\limits_{\hat{\pi}\in\text{supp}(\mu)}\mu(\hat{\pi})\,\hat{\pi}(E)}.$$

[20] In contrast, for models with kinks, (iii) would lead to the non-existence of dynamically consistent updates. See Hanany and Klibanoff [2007] for proof in the MEU case.

Dynamic consistency is intimately connected with Bayesian updating under expected utility. In fact, the following result shows that dynamic consistency justifies adopting Bayes' rule when preferences are expected utility.

Proposition 3.1 *Any update rule in* $\mathcal{U}$ *with domain restricted so that* $\succsim$ *are expected utility (i.e.,* ϕ *is affine) and that satisfies* ***DC*** *must produce the same conditional preferences as Bayes' rule.*

What does dynamic consistency say about updating smooth ambiguity preferences beyond expected utility? We can use Proposition 2.1 to characterize the dynamically consistent rules in $\mathcal{U}$. We will also show that this characterization is not vacuous – such rules exist. Unfortunately, it is no longer true that Bayesian updating is dynamically consistent for this larger class of preferences. In fact, as the dynamic Ellsberg example in the Introduction demonstrated, no rule that depends on only $\succsim$ and E can be dynamically consistent.[21]

The following result completely characterizes the update rules in $\mathcal{U}$ satisfying **DC**.[22]

Theorem 3.1 $U \in \mathcal{U}$ *satisfies* ***DC*** *if and only if*

$$\frac{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi_E(s)]}{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))]} = \frac{\mathbb{E}_{\mu}[\phi'(\mathbb{E}_{\pi}(u \circ g))\pi(s)]}{\mathbb{E}_{\mu}[\phi'(\mathbb{E}_{\pi}(u \circ g))\pi(E)]} \tag{3.1}$$

for all $s \in E$.

The sketch of the proof is to use differentiation and the unconditional optimality and interiority of g to show that for any $U \in \mathcal{U}$, $\frac{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi_E(s)]}{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))]}$ for $s \in S$ is the unique element of $\partial^{\star} V_{E,g,B}\,(u \circ g) \cap \Delta(E)$ and that $\frac{\mathbb{E}_{\mu}[\phi'(\mathbb{E}_{\pi}(u \circ g))\pi(s)]}{\mathbb{E}_{\mu}[\phi'(\mathbb{E}_{\pi}(u \circ g))\pi(E)]}$ for $s \in E$ and 0 for $s \in E^c$ is an element of $Q_E^{E,g,B}$. Interiority and unconditional optimality of g implies that for some feasible set B, $Q_E^{E,g,B}$ is a singleton. Proposition 2.1 and the fact that rules in $\mathcal{U}$ are independent of B then delivers the result.

Another perspective on Equation 3.1 can be gained by thinking of normalized gradients (in utility space) as local probability measures. Since preferences

[21] Even if we expand the set of update rules $\mathcal{U}$ to allow arbitrary updating of the ambiguity attitude function ϕ by dropping properties (ii) and (iii) above, one can show that there is still no dynamically consistent rule that involves updating μ by Bayes' rule.

[22] This result may be easily extended to the larger set of rules obtained by dropping properties (ii) and (iii) used to define $\mathcal{U}$ – simply add a subscript for B to $\mu_{E,g}$ and an E, g, B subscript for ϕ on the left-hand side of (3.1).

are smooth, they are locally expected utility and normalizing the gradient of the unconditional representation, $\mathbb{E}_\mu[\phi(\mathbb{E}_\pi(u \circ g))]$, gives the local probability measure at $u \circ g$. Similarly, normalizing the gradient of the conditional representation, $\mathbb{E}_{\mu_{E,g}}[\phi(\mathbb{E}_{\pi_E}(u \circ g))]$, gives a local conditional probability measure. Equation 3.1 reveals that dynamically consistent updating is equivalent to the statement that, at $u \circ g$, the local conditional probability measure is exactly the Bayesian update of the local probability measure – in this sense there is still a connection between Bayes' rule and dynamic consistency. When preferences are expected utility, updating the local measure at $u \circ g$ according to Bayes' rule is accomplished by updating the overall measure according to Bayes' rule, since the overall measure is also the local measure for all acts. For smooth ambiguity preferences more generally, updating the local measure at $u \circ g$ according to Bayes' rule no longer corresponds to updating the overall measure using Bayes' rule. In the next section, we show how this can be done through a specific reweighting of the prior belief μ.

3.1 An attractive update rule: the smooth rule

For a very large class of rules in $\mathcal{U}$ that are particularly appealing, in that they can be expressed as reweightings of the prior μ, we show using Theorem 3.1 that dynamic consistency selects a unique reweighting. We call this class of rules reweighted Bayesian because Bayes' rule corresponds to the special case where the prior is reweighted by (any positive multiple of) the likelihood. We will also show that dynamically consistent reweighting has further desirable properties, such as invariance to the order in which information is presented ("commutativity") and obeying a strict version of dynamic consistency.

Formally a reweighted Bayesian rule is the following:

Definition 3.1 *An update rule in* $\mathcal{U}$ *is* reweighted Bayesian (**RB**) *if* $\mu_{E,g}(\pi) = \frac{\alpha(\pi,\phi,u,g,E)\mu(\pi)\pi(E)}{\sum_{\hat{\pi}\in\Delta}\alpha(\hat{\pi},\phi,u,g,E)\mu(\hat{\pi})\hat{\pi}(E)}$ *for some real-valued function* α *satisfying* $\alpha(\pi,\phi,u,g,E) > 0$ *if* $\pi(E) > 0$.

The formula generating $\mu_{E,g}$ from μ will be referred to as a *reweighting*. In analogy with the above definition, reweightings of the form given there will be said to be ***RB*** *reweightings*. There are several things worth noticing in this definition. First, the weights, α, are independent of μ. This expresses the fact that the rule should be the "same" for all beliefs. Second, the value of α when $\pi(E) = 0$ is clearly irrelevant, so attention may be restricted to values on $\{\pi \in \Delta \mid \pi(E) > 0\}$. Third, Bayesian updating corresponds to the special case where α is constant in π, so that these rules include Bayes' rule. Fourth, the positivity restriction on α is needed to ensure that $\mu_{E,g}(\pi)$ is a well-defined probability measure for all μ. To see this, note that if $\alpha(\hat{\pi}, \phi, u, g, E) = 0$, taking $\mu(\hat{\pi}) = 1$ does not yield a well-defined $\mu_{E,g}$. Note also that if, for given (ϕ, u, g, E), some values of α were positive and some negative, $\mu_{E,g}$ would be either ill-defined or not a probability measure for some μ's. It would be fine for all values of α to be negative, but this produces no new rules, so we rule it out. Finally, all reweighted Bayesian rules preserve ambiguity in the sense that any π with $\pi(E) > 0$ that is given positive weight by μ is also given positive weight by the updated measure $\mu_{E,g}$.

The next result shows that dynamic consistency identifies a unique **RB** reweighting. We will refer to this novel updating procedure as the *smooth rule*, and it is the reweighting generated by setting

$$\alpha(\pi, \phi, u, g, E) = \begin{cases} \frac{\phi'(\mathbb{E}_{\pi}(u \circ g))}{\phi'\left(\mathbb{E}_{\pi_E}(u \circ g)\right)} \text{ if } \pi(E) > 0 \\ 0 \text{ otherwise} \end{cases}.$$

If ϕ is affine (ambiguity neutrality according to KMM), ϕ' is constant and the smooth rule collapses to Bayes' rule and preferences collapse to expected utility. The intersection with Bayes' rule is even larger than this, however. For example, whenever the act g is a constant act, our rule collapses to Bayes' rule. In general, the departure from Bayes' rule depends both on the ambiguity attitude, as reflected in ϕ (and in particular, in ratios of derivatives of ϕ, a quantity preserved under positive affine transformations) and on conditional and unconditional valuations of the unconditionally optimal act g. Observe that since ϕ is concave, relative to Bayes' rule the smooth rule overweights measures π that result in higher conditional valuations of g relative to unconditional valuations of g. Therefore, updating using this reweighting favors the unconditionally chosen act, making it easier to satisfy dynamic consistency. That an update rule generated by fixing any selection function and applying the reweighting given by the smooth rule is dynamically consistent follows directly from substituting into the equation for dynamic consistency established in Theorem 3.1. The contribution of Theorem 3.2 is in showing that the smooth rule is the unique **RB** reweighting for which this is always true – for any other **RB** reweighting, at least one update rule generated by combining the reweighting with a selection function will violate **DC**.

Formally:

Definition 3.2 *An* ***RB*** *reweighting* satisfies **DC** *if each* ***RB*** *update rule generated by combining a selection function with the* ***RB*** *reweighting satisfies* ***DC.***

Theorem 3.2 *The smooth rule is the unique* ***RB*** *reweighting to satisfy* ***DC****.*

To understand why the smooth rule is the only dynamically consistent **RB** reweighting, note that given Theorem 3.1, the question of uniqueness becomes whether a distinct **RB** reweighting can always satisfy Equation 3.1. Since the weights α do not depend on μ and the selection rule is arbitrary, it is enough to show uniqueness for a single μ. Using an event E containing at least two states and a μ with two-point support (and where the two points have distinct conditionals on E), one can show that Equation 3.1 generates a system of linear equations having a unique solution.

Given this result, from a normative point of view, dynamic consistency is a justification for adopting the smooth rule instead of Bayes' rule. A more psychological interpretation of this update rule is that it reflects an ex-post "rationalization" effect or a way to make current beliefs consonant with ex-ante optimal choices or plans.

The smooth rule has a number of nice properties beyond the dynamic consistency that selects it. First, because the smooth rule is an **RB** reweighting, it preserves ambiguity – the only measures in the support of μ that are eliminated by updating are the measures that assigned zero weight to the realized event, E. This distinguishes the smooth rule from, for example, the dynamically consistent rule described by setting $\mu_{E,g}$ equal to the degenerate measure on $\pi^*_g \equiv \frac{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi]}{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))]}$. This degenerate rule removes all ambiguity as soon as any information is learned. Given that ambiguity and reaction to it is the main feature of interest, such a degenerate rule is undesirable.

Second, the smooth rule satisfies *commutativity.* Commutativity means that the order of information received does not affect updating. If two sequences of events have the same intersection, ceteris paribus, beliefs following those sequences must be identical. Gilboa and Schmeidler [1993] advocate commutativity as an important property for an update rule to satisfy.

Definition 3.3 *A formula for updating μ satisfies* commutativity *if, for any $(\succsim, E, g, B) \in \mathcal{D}^{SM}$ and any non-null event F such that $E \cap F$ is non-null, applying the formula on E then F yields the same as updating on $E \cap F$ (i.e., $\left(\mu_{E,g,B}\right)_{F,g,\{f \in B | f=g \text{ on } E^c\}} = \mu_{E \cap F,g,B}$).*

Proposition 3.2 *The smooth rule satisfies commutativity.*

Third, when ϕ is strictly concave, the smooth rule satisfies a desirable strengthening of dynamic consistency. One way in which **DC** is weak is that it requires only weak conditional preference of g over f. Therefore, it is compatible with the axiom, for example, to unconditionally have $g \succ f$ for some

$f = g$ on E^c while conditionally $g \sim_{E,g,B} f$. In such a circumstance, it is true that the DM is willing to continue with g, but this is only weakly so. It turns out that the smooth rule satisfies a strengthening of **DC** that rules out such shifts from strict preference to indifference, as well as similar shifts from indifference to strict preference. Ruling out the latter shifts is a robustness requirement for dynamic consistency – just because an indifference was broken in favor of g unconditionally, why should it necessarily continue to be broken in favor of g conditional on E?

Formally, the stronger **DC** is:

Axiom 3.1 ***Strict DC***. *For any* $(\succsim, E, g, B) \in \mathcal{D}$*, if* $f \in B$ *with* $f = g$ *on* E^c*, then* $g \succ$ *(resp.* $\sim$*)*f *implies* $g \succ_{E,g,B}$ *(resp.* $\sim_{E,g,B}$*)*f*.*

Definition 3.4 *A real-valued function* τ *is* strictly concave *if, for all* $x, y \in$ *Domain* (τ)*,* $x \neq y$ *implies* $\tau(\alpha x + (1-\alpha)y) > \alpha\tau(x) + (1-\alpha)\tau(y)$ *for all* $\alpha \in (0, 1)$*.*

Proposition 3.3 *Assume* ϕ *is strictly concave. Updating using the smooth rule satisfies* ***Strict DC****.*

One may wonder if it might be appropriate to strengthen **DC** even further. The next section shows that, at least for two directions suggested by some other consistency concepts in the literature, further strengthening results in the impossibility of dynamically consistent updating.

3.2 The impossibility of stronger consistency when updating smooth ambiguity preferences

Recall that **DC** requires only conditional optimality of g among those feasible acts agreeing with g on E^c. Adding **Strict DC** extends the checking of conditional optimality to any $f \sim g$ and equal to g on E^c. Why not go well beyond preserving the optimality of g and, fixing g, check that the ordering of *all* feasible acts agreeing with g on E^c is preserved under conditional preference $\succsim_{E,g,B}$? The following axiom does exactly this.

Axiom 3.2 ***DC1*** *For any* $(\succsim, E, g, B) \in \mathcal{D}$*, if* $f, h \in B$ *with* $f = h = g$ *on* E^c *and* $f \succsim h$ *then* $f \succsim_{E,g,B} h$*.*

Requirements implying this appear in a number of places in the literature (e.g., Machina [1989], Machina and Schmeidler [1992], Epstein and Le Breton

[1993] and Ghirardato [2002]). Is such a stronger axiom desirable? This is debatable. At least two arguments that might be used to support **DC** don't seem to extend support to the additional requirements of **DC1**. First, the verbal essence of dynamic consistency involves preventing reversals, which will only ever have the opportunity to occur when they involve ex-ante optimal acts. As Machina [1989] writes (pp. 1636-7) "... behavior... will be *dynamically inconsistent*, in the sense that ... actual choice upon arriving at the decision node would differ from ... planned choice for that node." Second, many normative arguments in support of dynamic consistency, such as arguments showing how lack of consistency may lead to payoff-dominated outcomes (see e.g., Machina [1989], McClennen [1990], Seidenfeld [2004], and Segal [1997]), require only the conditional optimality of g. Most importantly for our purposes, we will show below that no update rules in $\mathcal{U}$ can satisfy **DC1**.

Epstein and Schneider [2003], when discussing differences between recursive multiple priors and the robust control model of Hansen and Sargent [2001] point out that the robust control model satisfies a version of dynamic consistency that checks only optimality of g. Aside from minor differences in the framework, the following is that condition:

Axiom 3.3 ***DC2*** *For any* $(\succsim, E, g, B) \in \mathcal{D}$*, if* $f \in \mathcal{A}$ *with* $f = g$ *on* E^c*, then* $g \succsim f$ *implies* $g \succsim_{E,g,B} f$*.*

The only difference from **DC** is that comparisons with g are not restricted to acts in the feasible set. Why restrict comparisons of g to feasible acts? Again we point out that the essence of dynamic consistency involves reversals, which are only relevant if they involve ex-ante feasible acts. Moreover, if we impose **DC2**, we will show that impossibility of consistent updating results.

Proposition 3.4 *No update rule in* $\mathcal{U}$ *satisfies, for all selection functions,* ***DC1*** *or* ***DC2****.*

Remark 3.1 We note that the same proof used for Proposition 3.4 suffices for further results. In particular, the restriction to rules in $\mathcal{U}$ is overly strong – the same impossibility holds even if the update rule is allowed to depend on the feasible set B. Moreover, one could replace unconditional and conditional weak preference with indifference in the statement of **DC2** and also yield impossibility. Furthermore, by exploiting the property of smooth ambiguity preferences that marginal rates of utility substitution between two states are constant around the constant acts, we can show an even stronger impossibility holds for **DC1**: fixing any selection function, no update rule in $\mathcal{U}$ satisfies **DC1**. We do not know whether such a strengthening holds for **DC2**.

4 Dynamically consistent updating of uncertainty averse and variational preferences

Recent work by CMMM [2008] explores the representation of a class of preferences they call uncertainty averse. This class corresponds to what we wrote down as the largest set of preferences we consider when discussing the output of dynamically consistent update rules. They show that these preferences may be represented by the functional $\inf_{p\in\Delta} G^{\star}\left(\int (u\circ f)\,dp, p\right)$, where $G^{\star}$ defined as follows:

Definition 4.1 *$G^{\star} : u(X)\times\Delta \to (-\infty,\infty]$ is given by*

$$G^{\star}(t,p) = \sup_{f\in\mathcal{A}}\left\{u\left(x^{f}\right) : \int (u\circ f)\,dp \leq t\right\},$$

where $x^{f}\in X$ satisfies $x^{f}\sim f$.

A natural question is how our results can be read in terms of this representation. We find that the measures in $T_{E,g}\left(V_{E,g,B}\right)$ are exactly those measures that minimize $G^{\star}_{E,g,B}\left(\int (u\circ g)\,dp, p\right)$ (where $G^{\star}_{E,g,B}$ denotes the $G^{\star}$ derived from $\succsim_{E,g,B}$). This leads to the following alternative characterization of dynamically consistent updating:[23]

Proposition 4.1 *$U\in\mathcal{Y}$ is dynamically consistent if and only if*

$$\arg\min_{p\in\Delta} G^{\star}_{E,g,B}\left(\int (u\circ g)\,dp, p\right)\cap Q_{E}^{E,g,B}\neq\emptyset$$

for all $(\succsim, E, g, B)\in\mathcal{D}$.

This result is helpful in analyzing dynamically consistent updating for the variational preference model characterized by MMR [2006a]. A variational preference over Anscombe-Aumann acts has the following concave representation:

$$\min_{p\in\Delta}\left(\int (u\circ f)\,dp + c(p)\right)$$

where $u : X\to\mathbb{R}$ is a nonconstant affine function and $c:\Delta\to[0,\infty]$ is grounded (i.e., has infimum zero), convex and lower semicontinuous (in the

[23] Again we thank Fabio Maccheroni for his help and suggestions concerning this result and its proof.

weak* topology). Thus, variational preferences correspond to $G^{\star}(t,p) = t + c(p)$. We will assume that u is unbounded (either above or below), as then c is unique given u (MMR [2006a], Proposition 6). They refer to c as the ambiguity index. Notice that $(V,u) \in \Psi$ is a variational representation whenever $V(a) = \min_{p\in\Delta}\left(\int adp + c(p)\right)$ for all $a \in \mathbb{R}^S$. Such a V is completely determined by specifying c. Let Ψ^{VR} denote the set of all such (V,u) and let $\mathcal{P}^{VR}$ denote the set of variational preference relations over $\mathcal{A}$. To update variational preferences we consider rules defined on an appropriate subset of quadruples $(\succsim, E, g, B)$ – specifically, $\mathcal{D}^{VR}$ is the set of all elements of $\mathcal{Q}$ such that $\succsim \in \mathcal{P}^{VR}$. The events E for which $p(E) = 0$ implies $c(p) = +\infty$ for all $p \in \Delta$ are the only events we consider conditioning on for variational preferences, as they are the non-null events in the sense defined in Section 2.1. Formally, restrict attention to the set of update rules $\mathcal{W} \subseteq \mathcal{Y}$ satisfying:

(i) (closure for variational preferences) the domain is contained in $\mathcal{D}^{VR}$ and the codomain is contained in Ψ^{VR}.

Property (i) reflects the scope of this application – we want to update within the class of variational preferences.

The following result completely characterizes the update rules in $\mathcal{W}$ that are dynamically consistent:

Corollary 4.1 *$U \in \mathcal{W}$ is dynamically consistent if and only if*

$$\arg\min_{p\in\Delta}\left(\int (u\circ g)\, dp + c_{E,g,B}(p)\right) \cap Q_E^{E,g,B} \neq \emptyset$$

for all $(\succsim, E, g, B) \in \mathcal{D} \subseteq \mathcal{D}^{VR}$.

This result does for variational preferences what Theorem 3.1 does for ambiguity averse smooth ambiguity preferences. Note that MEU preferences are a special case of variational preferences. MEU preferences with set of measures C correspond to variational preferences with ambiguity index

$$c(p) = \begin{cases} 0 & \text{if } p \in C \\ +\infty & \text{if } p \notin C \end{cases}.$$

Proposition 4.1 and Corollary 4.1 are thus strict generalizations of Proposition 1 in Hanany and Klibanoff [2007], which characterized dynamically consistent update rules for MEU preferences. For examples and characterizations of such rules, we refer the reader to that paper.

Although Corollary 4.1 is a complete characterization and is quite useful for checking if any given update rule for variational preferences is dynamically consistent, it is not as useful for explicitly constructing dynamically consistent update rules. To aid in this task, we present next an explicit family of dynamically consistent update rules for variational preferences.

Notation 4.1 *Given $p \in \Delta(E)$ and $q \in \Delta$, let $p \otimes^E q$ be the measure in Δ for which $p \otimes^E q(F) = q(E)p(F) + q(F \cap E^c)$ for all events F. Note that if $q(E^c) > 0$, we can write this as $p \otimes^E q(F) = q(E)\, p_E(F) + q(E^c)\, q_{E^c}(F)$ for all events F.*

In $p \otimes^E q$ the choice of p determines the probabilities conditional on E while q determines all other probabilities. The idea behind our family of update rules is to: (1) fix a probability measure r that is used to evaluate g unconditionally and supports the conditional optimality of g in B; (2) observe that if the updated ambiguity index, $c_{E,g,B}$, is to relate to the unconditional ambiguity index, c, we need a way to map measures conditional on E back to unconditional measures; (3) given that r was used to evaluate g unconditionally, a natural choice for this map is to treat any conditional measure $p \in \Delta(E)$ as if it were the unconditional measure $p \otimes^E r$; (4) since $c_{E,g,B}$ must be grounded to be part of a variational representation, ensure this without altering preferences by subtracting off the constant $\min_{q \in \Delta(E)} c\left(q \otimes^E r\right)$; (5) observe that since $\int (u \circ f)\, dp \otimes^E r = r(E) \int (u \circ f)\, dp + \int_{E^c} (u \circ f)\, dr$, the expected utility component of the contribution of p when evaluating an act f, $\int (u \circ f)\, dp$, is $\frac{1}{r(E)}$ times the expected utility component of the contribution of p as a part of $p \otimes^E r$ when evaluating f, so it is as if the utility function has been rescaled by the factor $\frac{1}{r(E)}$ when calculating the contribution of p; and (6) by the uniqueness properties of the variational representation, when the utility function is multiplicatively rescaled, the ambiguity index must also be rescaled by the same factor. This leads to the following set of update rules:

Definition 4.2

$$\mathcal{W}_0^{DC} = \left\{ \begin{array}{l} U \in \mathcal{W} \mid c_{E,g,B}(p) = \\ \quad \begin{cases} \frac{1}{r(E)} \left[c\left(p \otimes^E r\right) - \min_{q \in \Delta(E)} c\left(q \otimes^E r\right) \right] & \text{if } \ p \in \Delta(E) \\ +\infty & \text{if } \ p \notin \Delta(E) \end{cases} \\ \quad \textit{for some } r \in Q^{E,g,B} \cap \arg\min_{p \in \Delta} \left(\int (u \circ g)\, dp + c(p) \right). \end{array} \right\}.$$

The next result says that these rules exist and that all of these rules are dynamically consistent.

Proposition 4.2 $\emptyset \neq \mathcal{W}_0^{DC} \subseteq \mathcal{W}^{DC}$.

An important subset of variational preferences are smooth variational preferences (i.e., variational preferences that are everywhere differentiable). In addition to the tractability of smoothness, such preferences are rich in the sense that any variational preference may be approximated arbitrarily well by a smooth variational preference. Theorem 18 of MMR [2006a] shows that this smoothness is equivalent to the ambiguity index being essentially strictly convex (i.e., strictly convex on the domain of convex combinations of measures in Δ that are minimizers of the representing functional for at least one act). Assume this for the unconditional preference. In this case, since $u \circ g$ is interior in $u \circ \mathcal{A}$, $\arg\min_{p\in\Delta} \left(\int (u \circ g)\, dp + c(p)\right)$ is a singleton, meaning there is only one choice of r possible in $\mathcal{W}_0^{DC}$.

Corollary 4.2 *Restricted to smooth variational preferences, there is only one update rule in* $\mathcal{W}_0^{DC}$ *and this update rule does not depend on the feasible set* B *given* g, c *and* u.

Let's consider a prominent example of smooth variational preferences and see how it is updated according to our (now unique) rule.

Example 4.1 MMR point out that multiplier preferences (Hansen and Sargent [2001]) are a special case of variational preferences, where, for $\theta > 0$ and a reference probability $q \in \Delta$,

$$c(p) = \theta \sum_{s\in \text{supp}(q)} p(s) \ln \frac{p(s)}{q(s)} \text{ if } p << q \text{ (and } \infty \text{ otherwise).}$$

How does the rule in W_0^{DC} update these preferences? For $p \in \Delta(E)$,

$$c\left(p \otimes^E r\right) = \theta\Big(\sum_{s\in \text{supp}(q)\cap E} r(E)p(s) \ln \frac{r(E)p(s)}{q(s)} + \sum_{s\in \text{supp}(q)\cap E^c} r(s) \ln \frac{r(s)}{q(s)}\Big).$$

Observe that $\arg\min_{p\in\Delta(E)} c\left(p \otimes^E r\right) = \{q_E\}$. Thus,

$$\begin{aligned} c_{E,g,B}(p) &= \frac{1}{r(E)}\left[c\left(p \otimes^E r\right) - c\left((q_E) \otimes^E r\right)\right] \\ &= \theta\Big(\sum_{s\in \text{supp}(q)\cap E} p(s) \ln \frac{r(E)}{q(E)} \frac{p(s)}{q_E(s)}\Big) - \theta \ln \frac{r(E)}{q(E)} \\ &= \theta \sum_{s\in \text{supp}(q)\cap E} p(s) \ln \frac{p(s)}{q_E(s)}. \end{aligned}$$

So, our update rule says to update multiplier preferences simply by updating the reference measure q using Bayes' rule and otherwise leaving the ambiguity index unchanged. One may show that the same is true for any rule in $\mathcal{W}^{DC}$ that does not depend on the feasible set B and that preserves θ when updating multiplier preferences. This procedure seems quite natural, and its dynamic consistency makes sense in light of our Proposition 3.1 justifying Bayes' rule for expected utility and the result of Strzalecki [2008] that multiplier preferences satisfy the Savage [1954] axioms of subjective expected utility applied to acts mapping from S to X.

5 Updating minimax regret and other regret-based models

Thus far, we have considered models that generate a single complete and transitive binary relation over acts. In contrast, there is a literature in statistical decision theory and in economics (e.g., Chamberlain [2000], Bergemann and Schlag [2008]) that considers models of decision making under uncertainty that incorporate regret, and, as a result, cannot be represented by a single preference ordering. In particular, concepts like regret lead to different orderings when considering different feasible sets. In this section, we show that all of our results apply equally well to models incorporating regret, including for example, minimax regret with multiple priors (Hayashi [2008], Stoye [2008b]), a generalization of the classic minimax regret criterion (Savage [1951]).[24]

The key to extending our results to regret-based settings is recognizing that our characterizations of updating apply feasible set-by-feasible set, and thus are straightforward to adapt to models where preferences differ with the feasible set, but are standard ambiguity averse preferences *given* any fixed feasible set. Typically the only aspect of regret-based models that varies with the feasible set is the benchmark with respect to which regret is measured. This generates a different state-dependent utility function for each feasible set. To adapt our results, wherever we use the utility profile of an act, $u \circ f$, replace it with the "regret-adjusted" utility profile $\hat{u}_B(f)$ defined by $\hat{u}_B(f)(s) = u \circ f(s) - \max_{h \in B} u \circ h\,(s)$ for each $s \in S$. With this replacement, our results apply to any regret-based models, where, fixing the feasible set B, preferences over acts $f \in B$ are represented by $V\left(\hat{u}_B(f)\right)$ with V non-constant, continuous,

[24] There are also models of regret that generate intransitivity even when restricting to a fixed feasible set (for references see Stoye [2008a]). Our theory does not apply to these models.

weakly increasing and quasiconcave.

As an example, consider the model of minimax regret with multiple priors:

$$V(\hat{u}_B(f)) = -\max_{p \in C} \int \left[\max_{h \in B} u \circ h(s) - u \circ f(s) \right] dp = \min_{p \in C} \int \hat{u}_B(f) dp$$

with C a compact, convex subset of Δ. This is an MEU representation with regret-adjusted utility. Direct adaptation of Propositions 2.1 or 4.1 yields the following characterization of dynamically consistent rules for updating the set of priors C. A rule is dynamically consistent if and only if the updated sets of measures, $C_{E,g,B}$, satisfy

$$\hat{Q}_E^{E,g,B} \cap \arg \min_{q \in C_{E,g,B}} \int \hat{u}_B(g) dq \neq \emptyset,$$

where $\hat{Q}_E^{E,g,B}$ differs from $Q_E^{E,g,B}$ only in replacing $u \circ f$ and $u \circ g$ with $\hat{u}_B(f)$ and $\hat{u}_B(g)$ respectively. One can further apply the more explicit characterizations and examples of update rules for MEU provided in Hanany and Klibanoff [2007] and the algorithms for calculating these rules developed in Hanany, Klibanoff and Marom [2008]. See Hayashi [2009] for an alternative approach to dynamic extensions of the minimax regret model relaxing dynamic consistency.

6 Related literature

In some of the earliest relevant work, as mentioned in the Introduction, Machina [1989] and McClennen [1990] provide excellent and deep analyses of the problem of rational dynamic choice and advocate dropping consequentialism in order to maintain some type of consistency. McClennen proposed a theory of resolute choice, where it is assumed that conditional choices are in agreement with an unconditionally optimal plan, even when those conditional choices may conflict with some underlying conditional preference. He does not specify how this agreement is to be obtained. One way to view a dynamically consistent update rule is as a way of implementing McClennen's resolute choice while also preserving the property that conditional choices are based solely on conditional preferences. McClennen does not pursue this idea, and his definition of dynamic consistency is much too strong for this purpose (see Hanany and Klibanoff [2007]).

As explained in the Introduction, there are two approaches to modeling ambiguity averse preferences in dynamic settings that, unlike our approach, maintain consequentialism – using recursion on a limited set of events or adopting assumptions, such as consistent planning or naivete, that pin down behavior under dynamic inconsistency.[25] We compare these approaches with ours in the remainder of this section.

6.1 Recursive approaches

The recursive models most related to the preferences for which we examine updating are the recursive smooth ambiguity model of Klibanoff, Marinacci and Mukerji [2009], the recursive subset of the dynamic variational preferences model of Maccheroni, Marinacci and Rustichini [2006b] (which contains the recursive multiple priors model of Epstein and Schneider [2003] as a special case) and the model of regret with consistency to information arrival of Hayashi [2009]. These models satisfy what we refer to as recursive dynamic consistency for events in a given information filtration and benefit from the tractability delivered by recursion. However, the limitation to events in particular filtrations is no accident – these models are inherently incapable of satisfying **DC** for many events in the presence of ambiguity.

In our notation, recursive dynamic consistency is equivalent to the following condition:

Axiom 6.1 ***DC-R*** *(Recursive Dynamic Consistency). For any* $(\succsim, E, g, B) \in \mathcal{D}$*, if* $f, h \in \mathcal{A}$ *with* $f = h$ *on* E^c*,* $f \succsim h$ *if and only if* $f \succsim_{E,g,B} h$*.*

How does this relate to our **DC** condition? **DC-R** implies **DC** plus consequentialism. As mentioned earlier, consequentialism should be thought of as the requirement that updated preferences depend only on the ex-ante preference and the realized event, E, and make E^c a Savage-null event. Formally, we can write the following axiom:

Axiom 6.2 ***C*** *(Consequentialism). For any* $(\succsim, E, g_1, B_1)$*,* $(\succsim, E, g_2, B_2) \in \mathcal{D}$*,*
$\succsim_{E,g_1,B_1} = \succsim_{E,g_2,B_2}$ *and* $f, h, i \in \mathcal{A}$ *implies* $f_E h \sim_{E,g_1,B_1} f_E i$*.*

In fact, as stated in the next result, the straightforward proof of which is omitted, **DC-R** also implies the various strengthenings of **DC** mentioned in this paper.

[25] See also Ozdenoren and Peck [2008], who use extensive form games of conflict with nature to illustrate how varying the game the DM thinks she is playing is an alternative modeling strategy for generating variation in conditional choices in Ellsberg-like problems.

Proposition 6.1 *The update rules satisfying **DC-R** satisfy **DC**, **C**, **Strict DC**, **DC1** and **DC2**.*[26]

This result helps us understand why recursive models are simply not an option for updating on many events under ambiguity. For instance, from the dynamic Ellsberg example in the Introduction, we know that any update rule satisfying **C** must violate **DC** when applied to Ellsberg preferences upon observing the event $\{B, R\}$. By the proposition above, this implies every update rule violates **DC-R** when updating these preferences on $\{B, R\}$ and is therefore incompatible with recursion with respect to any filtration including the event $\{B, R\}$.

Another way to understand this failure of recursive models to handle updating on many events under ambiguity is to recognize that recursion leads to a lack of reduction in terms of information. For example, Figure 6.1 displays the choice between betting on black or betting on red in the traditional 3-color Ellsberg problem under two possible information structures. Squares indicate choice nodes and circles indicate nodes where uncertainty is (partially) resolved. In the pair on the right, the DM simply chooses and then learns which color was drawn, as is typical in the literature on Ellsberg behavior. In the pair on the left, the DM chooses and then the same information is revealed, but in two stages. First, the DM is told whether or not the ball drawn was yellow, and then, if it was not yellow, whether it was black or red. Assume that under the information structure on the right, choices follow the usual Ellsberg pattern (in the figure, $b' \succ r'$.) This, by itself, is consistent with recursion because all information is revealed at one time. However, these choices could be reversed ($r \succ b$) under a recursive model and the information structure on the left.[27]

Notice that this change in information structure does not affect the feasible actions or payoffs at all. In both cases the DM is choosing between betting on black or red once and for all at the beginning. The reversal across the two pairs under recursion is due purely to non-indifference toward timing of information

[26] We can also say something about the implications of **DC** plus **C**. It can be shown that they imply that the ex-ante and conditional optima are the same for any feasible sets containing only acts that agree on E^c. We conjecture that under natural conditions this can be used to show that **DC** plus **C** is, in fact, equivalent to **DC-R**. We have proved this conjecture for the case where preferences are MEU. Therefore, at least in this case, we can be confident that consequentialism is the only difference between **DC-R** and **DC**.

[27] Moreover, if preference does not reverse in this example, then, in an example that differs only in replacing the 0 payoffs on E^c with payoffs of 1, recursion will necessarily force a reversal between the two information structures.

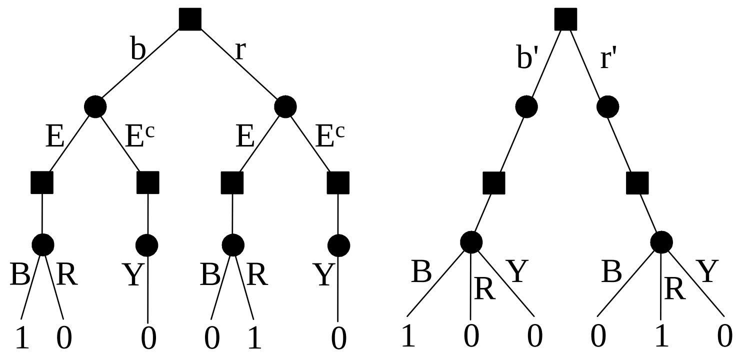

Figure 6.1: Betting on black vs. red with and without interim information revelation

even though that information has no instrumental value (and even when the actual time involved could be merely the time it takes to say "not yellow").

6.2 Dynamic inconsistency

Dynamic consistency is a central feature of our analysis. In contrast, there are approaches to dynamic decision making that take dynamic *inconsistency* of preferences over acts as given. In fact, several authors have explicitly suggested that a DM may wish to maintain consequentialism in the face of ambiguity, thus generating dynamic inconsistency (e.g., Eichberger and Grant [1997], Eichberger, Grant and Kelsey [2007]). Although not usually thought of in this way, this includes the vast literature (excepting Hanany and Klibanoff [2007]) exploring update rules for MEU such as full Bayesian updating (applying Bayes' rule to each measure in the set of measures)[28] and Maximum likelihood updating (applying Bayes' rule to only those measures assigning the largest probability to the observed event)[29], and for Choquet expected utility (Schmeidler [1989]), such as the Dempster-Shafer rule (Dempster [1968] and

[28] See Jaffray ([1992],[1994]), Fagin and Halpern [1990], Wasserman and Kadane [1990], and Walley [1991]. Sarin and Wakker [1998], Pires [2002], Siniscalchi [2009], Wang [2003] and Epstein and Schneider [2003] formally characterize this update rule using preference axioms in various settings.

[29] Explored in terms of preferences in Gilboa and Schmeidler [1993].

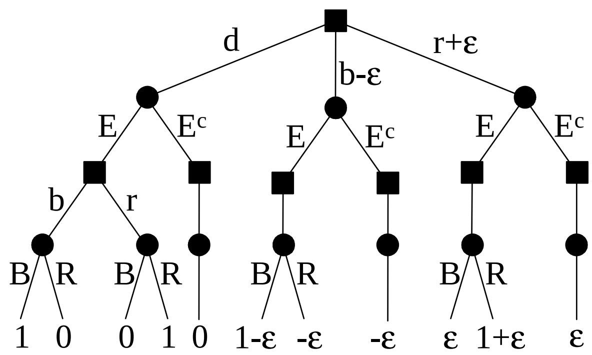

Figure 6.2: Dynamic consistency vs. inconsistency in a dynamic Ellsberg problem

Shafer [1976]).[30] All of these classic rules are dynamically inconsistent, as they satisfy consequentialism. Unlike preferences that are dynamically consistent, dynamically inconsistent preferences need to be coupled with assumptions about how this inconsistency is resolved before these preferences may be translated into behavior. The two most popular assumptions in the literature are *naivete* and *sophistication*. This distinction was made by Strotz [1955-6] and many that followed. A sophisticated DM correctly anticipates future conditional preferences and chooses by optimizing current preferences taking these future preferences as a constraint. Under the assumption of consequentialism, Siniscalchi [2009] shows how sophistication relates to consistent planning (a refinement of backward induction) in an environment rich enough to allow for ambiguity. A naive DM incorrectly anticipates that future conditional desires will be the same as currently desired plans for the future and therefore chooses assuming that ex-ante desired plans will be carried out. Figure 6.2 presents an example illustrating the distinction between these approaches and ours in the context of ambiguity. The setting of the example is as in the dynamic Ellsberg example presented in the Introduction. As there, the three states correspond to the three colors that might be drawn from the urn – black (B), red (R) and yellow (Y) – with the odds of drawing black known. The event E is the event that the drawn ball is not yellow. At the initial choice node, the DM must choose between d, $b - \varepsilon$ and $r + \varepsilon$. The choice d leads to exactly the situa-

[30] For preference characterizations of this rule see e.g., Gilboa and Schmeidler [1993], Wang [2003], and Nishimura and Ozaki [2003].

tion described in the dynamic Ellsberg problem where the DM must choose between betting on black or on red contingent on whether E or E^c is realized. The choice $b-\varepsilon$ leads to commitment to bet on black (still contingent, though trivially so, on whether E or E^c is realized) by removing the contingent option of betting on red, and requires payment of a fee of ε to do so. Similarly, the choice $r+\varepsilon$ leads to commitment to bet on red (again trivially contingent) by removing the contingent option of betting on black, and offers a payment of ε to do so. Assume that the DM unconditionally strictly prefers betting on black to betting on red (as is usual in the Ellsberg problem), that higher payoffs are desirable and that $\varepsilon > 0$ is small enough, the DM will have preferences with

$$(1,0,0) \succ (1-\varepsilon,-\varepsilon,-\varepsilon) \succ (\varepsilon,1+\varepsilon,\varepsilon) \succ (0,1,0)\,.$$

If the DM updates in a dynamically consistent way, then d will be chosen initially and will be followed by the choice of B if the event E is realized. These choices result in the payoff vector $(1,0,0)$, the ex-ante optimum. Suppose instead that the DM updates in a dynamically inconsistent way. For example, the DM might have the smooth ambiguity preferences used in the numerical example in the Introduction and use Bayes' rule to update μ. As was shown in that example, this would lead R to be chosen over B if the event E were realized. A naive DM would not anticipate this choice of R over B, and thus would choose d initially, planning to then choose B, generating $(1,0,0)$. However, what would actually be realized is d followed by R, generating $(0,1,0)$. Observe that naive dynamic inconsistency leads to a strictly dominated outcome – choosing $r+\varepsilon$ rather than d would generate the dominating payoffs $(\varepsilon,1+\varepsilon,\varepsilon)$.[31] A sophisticated DM, realizing correctly that the choice of d would lead to $(0,1,0)$ would instead choose $b-\varepsilon$, generating payoff vector $(1-\varepsilon,-\varepsilon,-\varepsilon)$. In effect, the sophisticated but dynamically inconsistent DM is willing to pay a fee to remove the option R contingent on the event E occurring. Thus, under dynamically inconsistent updating, the choice between adopting sophistication and naivete becomes relevant. In the example above, sophistication was superior to naivete from an ex-ante point of view. However, in other examples the DM may be ex-ante better off if future selves were naive than if they were sophisticated.[32] Thus there is not a clear ex-ante welfare justification for taking a sophisticated approach over a naive one or vice-versa. In contrast, there is a clear welfare comparison between dynamic consistency

[31] For previous papers arguing that naive dynamic inconsistency may lead to strictly dominated choices, see e.g., Green [1987], Machina [1989], and Segal [1997] to name just a few.

[32] O'Donoghue and Rabin's [1999] Example 2 demonstrates this point under certainty when dynamic inconsistency is generated by non-exponential discounting. Similar examples may be constructed using updating as the source of dynamic inconsistency.

and dynamic inconsistency. If a DM could choose an update rule at the ex-ante stage, it would always be optimal to choose a dynamically consistent one, such as the rules explored in this paper.

Note that while recursion (as noted in Figure 6.1) violates invariance (reduction) with respect to information timing (when holding the action timing and feasible payoffs fixed), Figure 6.2 shows sophistication plus dynamic inconsistency violates invariance (reduction) with respect to action timing (holding the information timing and feasible payoffs fixed). Our dynamically consistent updating satisfies both these invariances while violating invariance with respect to the ex-ante optimum induced by the decision problem (consequentialism).

7 Summary

In this paper, we have characterized dynamically consistent updating for general ambiguity averse (quasiconcave) preference models and proposed novel update rules for broad classes of preferences for which there was little prior understanding of how to update in a consistent way. For ambiguity averse smooth ambiguity preferences, we characterized consistent updating and proposed a rule, called the smooth rule, that has a number of attractive properties including dynamic consistency and invariance to the order in which information arrives (commutativity). The form of the rule is a "reweighting" of Bayes' rule, where the weights depend on the DM's ambiguity aversion and her unconditionally chosen act in a decision problem. The rule is the unique reweighting to be dynamically consistent. We also characterized dynamically consistent update rules for variational preferences, constructed such rules and applied them to multiplier preferences. Finally, we showed that our results on updating apply equally well to regret-based modifications of ambiguity averse models.

A Appendix: Proofs not in the main text

The following lemma is the key to proving Theorem 2.1.

Notation A.1 *For* $a \in \mathbb{R}^S$, $E \subseteq S$, *the restriction of* a *to* E *is denoted* $a_{|E}$.

Lemma A.1 *Fix a preference relation,* $\succsim_{E,g,B}$, *on* $\mathcal{A}$, *a von Neumann - Morgenstern utility* $u : X \to \mathbb{R}$ *representing* $\succsim_{E,g,B}$ *on constant acts, and an event* $E \subseteq S$. *Assume there exists a quasiconcave* $V_{E,g,B} : \mathbb{R}^S \to \mathbb{R}$ *such* *that (i)* $f \succsim_{E,g,B} h$ *if and only if* $V_{E,g,B}(u \circ f) \geq V_{E,g,B}(u \circ h)$*, (ii)* $a_{|E} = b_{|E}$ *implies* $V_{E,g,B}(a) = V_{E,g,B}(b)$*, and (iii)* $a(s) > b(s)$ *for all* $s \in E$ *implies* $V_{E,g,B}(a) > V_{E,g,B}(b)$*. Then for any* $h \in B$*, [*$h \succsim_{E,g,B} f$ *for all* $f \in B$ *with* $f = h$ *on* E^c*] is equivalent to* $T_{E,h}(V_{E,g,B}) \cap Q_E^{E,h,B} \neq \emptyset$*.*

Proof. ($h \in B$, [$h \succsim_{E,g,B} f$ for all $f \in B$ with $f = h$ on E^c] $\implies$ $T_{E,h}(V_{E,g,B}) \cap Q_E^{E,h,B} \neq \emptyset$) Let I be $\mathbb{R}^E$. Let $\succsim$ be a complete, transitive binary relation on I defined by

$$a_{|E} \succsim b_{|E} \text{ if and only if } V_{E,g,B}(a) \geq V_{E,g,B}(b).$$

Note that $\succsim$ is well-defined because of (ii). Let $\succ$ and $\approx$ be the asymmetric and symmetric parts of $\succsim$. Consider the sets $D_1 \equiv \{a \mid a \in I \text{ with } a \succ u \circ h_{|E}\}$ and $D_2 \equiv \{u \circ f_{|E} \mid f \in B \text{ with } f = h \text{ on } E^c\}$. Quasiconcavity of $V_{E,g,B}$ implies D_1 is convex, while B convex implies D_2 is convex. Conditional optimality of h implies $D_1 \cap D_2 = \emptyset$. D_1 is non-empty by (iii) and also has a non-empty interior. D_2 is non-empty since it contains $u \circ h_{|E}$. By a separating hyperplane theorem (e.g., Aliprantis and Border [1999], Thm. 5.50, p. 190), there must exist a hyperplane separating D_1 and D_2. Without loss of generality, such a hyperplane may be defined by $\{a \in I \mid \int a dr = \alpha\}$ for $r \in \Delta(E)$ restricted to 2^E and real α such that $\int a dr \geq \alpha \geq \int b dr$ for all $a \in D_1$ and $b \in D_2$. Since $u \circ h_{|E} \in D_2$, $\alpha \geq \int (u \circ h_{|E}) dr$. Suppose $\alpha > \int (u \circ h_{|E}) dr$. Then by the event-wise continuity of $\int (\cdot) dr$ and (iii), there would exist an $a \in D_1$ such that $\int a dr < \alpha$, a contradiction. Thus, $\alpha = \int (u \circ h_{|E}) dr$. Similarly, one can show that $a \in D_1$ implies $\int a dr > \int (u \circ h_{|E}) dr$. Therefore, $\int a dr > \int (u \circ h_{|E}) dr \geq \int b dr$ for all $a \in D_1$ and $b \in D_2$. Let $\hat{r}$ be the extension of r to 2^S obtained by assigning zero to all measurable events in E^c. By the definitions of $T_{E,h}(V_{E,g,B})$ and $Q_E^{E,h,B}$, $\hat{r} \in T_{E,h}(V_{E,g,B}) \cap Q_E^{E,h,B}$.
($h \in B$, $T_{E,h}(V_{E,g,B}) \cap Q_E^{E,h,B} \neq \emptyset \implies$ [$h \succsim_{E,g,B} f$ for all $f \in B$ with $f = h$ on E^c]) Let $\hat{r}$ be an element of $T_{E,h}(V_{E,g,B}) \cap Q_E^{E,h,B}$. Since $\hat{r} \in Q_E^{E,h,B}$, $\int (u \circ h) d\hat{r} \geq \int (u \circ f) d\hat{r}$ for all $f \in B$ with $f = h$ on E^c. Since $\hat{r} \in T_{E,h}(V_{E,g,B})$, $\int (u \circ f) d\hat{r} > \int (u \circ h) d\hat{r}$ for all f such that $V_{E,g,B}(u \circ f) > V_{E,g,B}(u \circ h)$. Thus $V_{E,g,B}(u \circ f) - V_{E,g,B}(u \circ h) \leq 0$ for all $f \in B$ with $f = h$ on E^c. Since $V_{E,g,B}(u \circ \cdot)$ represents $\succsim_{E,g,B}$ by (i), this implies h is conditionally optimal among all $f \in B$ with $f = h$ on E^c. ■

Proof of Theorem 2.1. Note that applying the above lemma where h is taken to be the unconditionally optimal act g and E is non-null yields $T_{E,g}(V_{E,g,B}) \cap Q_E^{E,g,B} \neq \emptyset$ as a characterization of **DC** for update rules in $\mathcal{Y}$. ■

The next lemma is the key to proving Proposition 2.1.

Lemma A.2 *For any* $(\succsim, E, g, B) \in \mathcal{D}$ *and* $h \in \operatorname{int}(\mathcal{A})$,

$$T_{E,h}(V_{E,g,B}) = \partial^{\star} V_{E,g,B}(u \circ h) \cap \Delta(E).$$

Proof. Define $I : \mathbb{R}^S \to \mathbb{R}$ by

$$I(a) = \begin{cases} V_{E,g,B}(a) & \text{if} \quad a \in u(X)^S \\ -\infty & \text{otherwise.} \end{cases}$$

Fix $a = u \circ h$. We begin by proving $T_{E,h}(V_{E,g,B}) = \partial^{\star} I(a) \cap \Delta(E)$:

$$\begin{aligned} &\partial^{\star} I(a) \cap \Delta(E) \\ &= \left\{ p \in \Delta(E) : \left\{ b \in \mathbb{R}^S : I(b) > I(a) \right\} \subseteq \left\{ b \in \mathbb{R}^S : \int b dp > \int a dp \right\} \right\} \\ &= \left\{ p \in \Delta(E) : \left\{ b \in u(X)^S : I(b) > I(a) \right\} \subseteq \left\{ b \in \mathbb{R}^S : \int b dp > \int a dp \right\} \right\} \\ &= \left\{ \begin{array}{c} p \in \Delta(E) : \left\{ b \in u(X)^S : V_{E,g,B}(b) > V_{E,g,B}(a) \right\} \\ \subseteq \left\{ b \in u(X)^S : \int b dp > \int a dp \right\} \end{array} \right\} \\ &= \left\{ \begin{array}{c} p \in \Delta(E) : \int (u \circ f) dp > \int a dp \\ \text{for all } f \in \mathcal{A} \text{ such that } V_{E,g,B}(u \circ f) > V_{E,g,B}(a) \end{array} \right\} \\ &= T_{E,h}(V_{E,g,B}). \end{aligned}$$

To complete the proof, we show that $\partial^{\star} I(a) = \partial^{\star} V_{E,g,B}(a)$. Since $V_{E,g,B}(b) \geq I(b)$ for all $b \in \mathbb{R}^S$ and $V_{E,g,B}(a) = I(a)$, $\partial^{\star} V_{E,g,B}(a) \subseteq \partial^{\star} I(a)$. We now prove the opposite inclusion. Suppose $q \in \partial^{\star} I(a)$. If $q \notin \partial^{\star} V_{E,g,B}(a)$ then there exists a $b \in \mathbb{R}^S$ such that $V_{E,g,B}(b) > V_{E,g,B}(a)$ and $\int b dq \leq \int a dq$. Since E is non-null and $V_{E,g,B}$ is continuous, there exists an $\hat{a}$ in the interior of $u(X)^S$ such that $V_{E,g,B}(b) > V_{E,g,B}(\hat{a}) > V_{E,g,B}(a)$ and $\int b dq \leq \int a dq < \int \hat{a} dq$. Since $\hat{a}$ is in the interior of $u(X)^S$ there exists a $\lambda \in (0,1)$ such that $\lambda b + (1-\lambda)\hat{a}$ is in the interior of $u(X)^S$. By quasiconcavity, $V_{E,g,B}(\lambda b + (1-\lambda)\hat{a}) \geq V_{E,g,B}(\hat{a})$ while by linearity, $\int (\lambda b + (1-\lambda)\hat{a}) dq < \int \hat{a} dq$. Now by increasing $\lambda b + (1-\lambda)\hat{a}$ by a small enough amount on each state in E, we obtain by a $c \in u(X)^S$ such that $V_{E,g,B}(c) > V_{E,g,B}(\hat{a})$ and $\int c dq < \int \hat{a} dq$. This contradicts $q \in \partial^{\star} I(a)$. Therefore, $\partial^{\star} I(a) \subseteq \partial^{\star} V_{E,g,B}(a)$ and thus $\partial^{\star} I(a) = \partial^{\star} V_{E,g,B}(a)$ as desired. ■

Proof of Proposition 2.1. By Lemma A.2, we can replace $T_{E,g}(V_{E,g,B})$ by $\partial^{\star} V_{E,g,B}(u \circ g) \cap \Delta(E)$ in the statement of Theorem 2.1. Since $Q_E^{E,g,B} \subseteq \Delta(E)$, $\partial^{\star} V_{E,g,B}(u \circ g) \cap \Delta(E)$ may be further replaced by $\partial^{\star} V_{E,g,B}(u \circ g)$ without affecting the result. ■

The following lemma is used in proving the subsequent lemma that, in turn, is invoked in proving the next two results in the paper:

Lemma A.3 *Fix an interior act g and preference representation $(V, u) \in \Psi$. For any closed, convex set of utility acts, D, and interior utility act $a = u \circ g \in D$ such that $V(a) > V(b)$ for all $b \in D$, $b \neq a$, there exists a closed, convex set of acts, B, such that $g \in B$ and $u \circ B = D$.*

Proof. Fix any such D and a. There exist consequences $\overline{z}$ and $\underline{z}$ such that $V(u \circ \overline{z}) \geq V(b) \geq V(u \circ \underline{z})$ for all $b \in D$. Thus, each $b \in D$ can be associated with an act that in each state yields lotteries putting positive probability on only $\overline{z}$ and/or $\underline{z}$ and for each $b \in D$ there is a unique such act. Let $\hat{B}$ be the set of acts constructed in this way from D. Consider any $f, h \in \hat{B}$ and any $\lambda \in (0,1)$. Since $u \circ f$, $u \circ h \in D$, D is convex and u is affine, $u \circ (\lambda f + (1-\lambda)h) = \lambda u \circ f + (1-\lambda) u \circ h \in D$. Since $\lambda f + (1-\lambda)h$ in each state yields lotteries putting positive probability on only $\overline{z}$ and/or $\underline{z}$, this must be the unique such act associated with $\lambda u \circ f + (1-\lambda) u \circ h \in D$ and thus $\lambda f + (1-\lambda)h \in \hat{B}$. This proves $\hat{B}$ is convex. Let $B = co(\hat{B} \cup \{g\})$. B is closed and convex and contains g. Since $u \circ \hat{B} = D$ by construction and $u \circ g \in u \circ \hat{B}$, the fact that u is affine implies $u \circ B = D$. ■

Lemma A.4 *Given a smooth ambiguity preference, $\succsim$, with ϕ concave and differentiable, an interior act g and a non-null event E, there always exists a feasible set B such that (i) $g \in B$, (ii) g is optimal in B according to $\succsim$ and if $f \in B$ and $f \sim g$ then $u \circ f = u \circ g$, (iii) $Q_E^{E,g,B}$ (where $f(s) \sim g(s)$ on E^c is substituted for $f = g$ on E^c in the definition) is a singleton.*

Proof. If E is a singleton, then $B = \{g\}$ works. Otherwise, consider the indifference curve through g. Let D be the intersection of the set $u \circ \{f \in \mathcal{A} \mid f(s) \sim g(s) \text{ on } E^c\}$ and a closed ball of utility acts having $u \circ g$ on its boundary, tangent at $u \circ g$ to the unique (by differentiability of ϕ and interiority of $u \circ g$) hyperplane in utility space through $u \circ g$ supporting the set $u \circ \{f \in \mathcal{A} \mid f \succsim g\}$. D is convex since it is the intersection of two convex sets and it is also closed. By construction, the projection of D onto E, denoted $proj_E(D)$, has a unique supporting hyperplane at $proj_E(u \circ g)$. Thus, there is a unique $r \in \Delta(E)$ such that $\int (u \circ g) dr \geq \int b dr$ for all $b \in D$. By Lemma A.3, there exists a feasible set of acts B such that $g \in B$ and $u \circ B = D$. By construction of D, g is optimal in B and any other optimal act in B must have the same utility profile, so (ii) is satisfied. Finally, applying the definition of $Q_E^{E,g,B}$ (and replacing $f(s) \sim g(s)$ on E^c for $f = g$ on E^c in that definition), yields $Q_E^{E,g,B} = \{r\}$ proving (iii). ■

Proof of Proposition 3.1. For expected utility preferences, ϕ is affine and the only aspect of the belief relevant for preferences over acts is the reduced measure $v \equiv E_\mu \pi$. Applying Bayes' rule yields expected utility preferences with ϕ and u and updated reduced measure $v_{E,g} = E_{\mu_{E,g}} \pi_E = v_E$. Since the conditional representation is continuously differentiable, $\partial^\star V_{E,g,B}(u \circ g) \cap \Delta(E) = \{v_E\}$. Unconditional optimality of g implies $v_E \in Q_E^{E,g,B}$. As $u \circ g$ is interior, Proposition 2.1 then implies that Bayes' rule satisfies **DC** for such preferences. We prove it does so uniquely. Since the updated preferences may depend on only $(\succsim, E, g)$, and since $u \circ g$ is interior in $u \circ \mathcal{A}$, by Lemma A.4 we can without loss of generality fix a problem with a feasible set having optimum g for which $Q_E^{E,g,B}$ is a singleton. Then invoking Proposition 2.1 again, **DC** implies

$$v_{E,g}(s) = v_E(s) \text{ for all } s \in E.$$

Thus, updating must be Bayesian in the sense that the updated preferences over acts are identical to those generated by using Bayes' rule. ■

Proof of Theorem 3.1. Since $\mathbb{E}_{\mu_{E,g}} \phi(\mathbb{E}_{\pi_E} a)$ is finite, concave and differentiable at $a = u \circ g$ and $u \circ g$ is interior, the unique element of $\partial^\star V_{E,g,B}(u \circ g) \cap \Delta(E)$ may be found by taking the gradient of $\mathbb{E}_{\mu_{E,g}} \phi(\mathbb{E}_{\pi_E} a)$ at $a = u \circ g$ and normalizing. This yields $\left(\frac{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi_E(s)]}{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))]} \right)_{s \in S}$. Therefore, Proposition 2.1 says U satisfies **DC** if and only if $\frac{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi_E(s)]}{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))]} \in Q_E^{E,g,B}$. Since g is unconditionally optimal, an element of $Q^{E,g,B}$ may be obtained by differentiating $\mathbb{E}_\mu \phi(\mathbb{E}_\pi a)$ with respect to $a \in \mathbb{R}^S$, evaluating at $a = u \circ g$, and normalizing. This yields $\frac{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(s)]}{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))]}$ for all $s \in S$. Since $\frac{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(s)]}{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))]} \in Q^{E,g,B}$, its conditional on E, $\frac{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(s)]}{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(E)]}$ for $s \in E$ and 0 otherwise, must be an element of $Q_E^{E,g,B}$. As $u \circ g$ is interior in $u \circ \mathcal{A}$, Lemma A.4 implies there exists an appropriate B such that $Q_E^{E,g,B}$ is a singleton. For such a B, therefore, $\frac{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi_E(s)]}{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))]} \in Q_E^{E,g,B}$ if and only if $\frac{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi_E(s)]}{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))]} = \frac{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(s)]}{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(E)]}$ for all $s \in E$. As U does not depend on B except through g (since $U \in \mathcal{U}$), the proposition follows. ■

Proof of Theorem 3.2.

By Theorem 3.1, in order to satisfy **DC** it must be that

$$\frac{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi_E(s)]}{\mathbb{E}_{\mu_{E,g}}[\phi'(\mathbb{E}_{\pi_E}(u \circ g))]} = \frac{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(s)]}{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(E)]} \text{ for all } s \in E.$$

For any **RB** reweighting, substituting for $\mu_{E,g}$ yields

$$\frac{\mathbb{E}_\mu[\alpha(\pi, \phi, u, g, E)\, \phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi(s)]}{\mathbb{E}_\mu[\alpha(\pi, \phi, u, g, E)\, \phi'(\mathbb{E}_{\pi_E}(u \circ g))\pi(E)]} = \frac{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(s)]}{\mathbb{E}_\mu[\phi'(\mathbb{E}_\pi(u \circ g))\pi(E)]} \text{ for all } s \in E. \tag{A.1}$$

The smooth rule is defined by setting

$$\alpha(\pi,\phi,u,g,E)=\alpha^{SM}(\pi,\phi,u,g,E)\equiv\begin{cases}\frac{\phi'(\mathbb{E}_{\pi}(u\circ g))}{\phi'\left(\mathbb{E}_{\pi_E}(u\circ g)\right)} \text{ if } \pi(E)>0\\ 0 \text{ otherwise}\end{cases}.$$

Substituting this α into Equation A.1, it is immediate that the smooth rule satisfies **DC**.

To show that it is the unique such **RB** reweighting, fix ϕ, u, g and E. When E is a singleton, the value of α at E is irrelevant. Therefore, assume E contains at least two states. Observe that if $\alpha(\pi,\phi,u,g,E)=k\alpha^{SM}(\pi,\phi,u,g,E)$ for $k\neq 0$, this is the same as the smooth rule, since the multiplicative factor cancels out in the normalization. So, in order for an **RB** reweighting to be different than the smooth rule, there must exist $\pi^1\neq\pi^2$ with $\pi^i(E)>0, i=1,2$ and positive real numbers $k_1\neq k_2$ such that $\alpha(\pi^1,\phi,u,g,E)=k_1\alpha^{SM}(\pi^1,\phi,u,g,E)$ and $\alpha(\pi^2,\phi,u,g,E)=k_2\alpha^{SM}(\pi^2,\phi,u,g,E)$. Fix such π^1,π^2. There exists a π^3 with $\pi^3(E)>0$ such that π^3_E is different than at least one of π^1_E or π^2_E and a $k_3>0$ such that $\alpha(\pi^3,\phi,u,g,E)=k_3\alpha^{SM}(\pi^3,\phi,u,g,E)$. k_3 must differ from at least one of k_1 and k_2. Therefore, without loss of generality, assume $\pi^1_E\neq\pi^2_E$ (since we can substitute π^3 for one or the other if this is not the case). We will show that this implies **DC** is violated. Consider $\succsim\in\mathcal{P}^{SM}$ corresponding to a smooth ambiguity representation using u, ϕ and a μ having support $\{\pi^1,\pi^2\}$ and fix a selection function selecting that (ϕ,μ) for $\succsim$.
For such a μ, the left-hand side of (A.1) is equal to:

$$\frac{k_1\phi'(\mathbb{E}_{\pi^1}(u\circ g))\pi^1(s)\mu(\pi^1)+k_2\phi'(\mathbb{E}_{\pi^2}(u\circ g))\pi^2(s)\mu(\pi^2)}{k_1\phi'(\mathbb{E}_{\pi^1}(u\circ g))\pi^1(E)\mu(\pi^1)+k_2\phi'(\mathbb{E}_{\pi^2}(u\circ g))\pi^2(E)\mu(\pi^2)} \text{ for all } s\in E.$$

From (A.1), this must equal

$$\frac{\phi'(\mathbb{E}_{\pi^1}(u\circ g))\pi^1(s)\mu(\pi^1)+\phi'(\mathbb{E}_{\pi^2}(u\circ g))\pi^2(s)\mu(\pi^2)}{\phi'(\mathbb{E}_{\pi^1}(u\circ g))\pi^1(E)\mu(\pi^1)+\phi'(\mathbb{E}_{\pi^2}(u\circ g))\pi^2(E)\mu(\pi^2)} \text{ for all } s\in E$$

for **DC** to hold. Simplifying yields

$$\begin{aligned} & k_1\left[\phi'(\mathbb{E}_{\pi^1}(u\circ g))\mu\left(\pi^1\right)\phi'(\mathbb{E}_{\pi^2}(u\circ g))\mu\left(\pi^2\right)\right]\left[\pi^1(s)\pi^2(E)-\pi^2(s)\pi^1(E)\right]\\ = & k_2\left[\phi'(\mathbb{E}_{\pi^1}(u\circ g))\mu\left(\pi^1\right)\phi'(\mathbb{E}_{\pi^2}(u\circ g))\mu\left(\pi^2\right)\right]\left[\pi^1(s)\pi^2(E)-\pi^2(s)\pi^1(E)\right].\end{aligned}$$

Since $k_1\neq k_2$, the only case in which this can be true is if, for all $s\in E$,

$$\left[\phi'(\mathbb{E}_{\pi^1}(u\circ g))\mu\left(\pi^1\right)\phi'(\mathbb{E}_{\pi^2}(u\circ g))\mu\left(\pi^2\right)\right]\left[\pi^1(s)\pi^2(E)-\pi^2(s)\pi^1(E)\right]=0.$$

As $\phi' > 0$, this requires

$$\pi^1(s)\pi^2(E) - \pi^2(s)\pi^1(E) = 0 \text{ for all } s \in E,$$

meaning

$$\pi^1_E = \pi^2_E,$$

a contradiction. Therefore no other **RB** reweighting can satisfy dynamic consistency. ■

Proof of Proposition 3.2. According to the smooth rule, $\mu_{E\cap F,g} = \frac{\frac{\phi'(\mathbb{E}_\pi(u\circ g))}{\phi'(\mathbb{E}_{\pi_{E\cap F}}(u\circ g))}\mu(\pi)\pi(E\cap F)}{\sum_{\hat\pi\in\Delta}\frac{\phi'(\mathbb{E}_{\hat\pi}(u\circ g))}{\phi'(\mathbb{E}_{\hat\pi_{E\cap F}}(u\circ g))}\mu(\hat\pi)\hat\pi(E\cap F)}$. Similarly, updating on E then F using the smooth rule produces

$$\begin{aligned}
\left(\mu_{E,g}\right)_{F,g} &= \frac{\frac{\phi'(\mathbb{E}_{\pi_E}(u\circ g))}{\phi'(\mathbb{E}_{\pi_{E_F}}(u\circ g))}\mu_{E,g}(\pi)\,\pi_E(F)}{\sum_{\hat\pi\in\Delta}\frac{\phi'(\mathbb{E}_{\hat\pi_E}(u\circ g))}{\phi'(\mathbb{E}_{\hat\pi_{E_F}}(u\circ g))}\mu_{E,g}(\hat\pi)\,\hat\pi_E(F)} \\
&= \frac{\frac{\phi'(\mathbb{E}_{\pi_E}(u\circ g))}{\phi'(\mathbb{E}_{\pi_{E_F}}(u\circ g))}\frac{\phi'(\mathbb{E}_{\pi}(u\circ g))}{\phi'(\mathbb{E}_{\pi_E}(u\circ g))}\mu(\pi)\,\pi(E)\,\pi_E(F)}{\sum_{\hat\pi\in\Delta}\frac{\phi'(\mathbb{E}_{\hat\pi_E}(u\circ g))}{\phi'(\mathbb{E}_{\hat\pi_{E_F}}(u\circ g))}\frac{\phi'(\mathbb{E}_{\hat\pi}(u\circ g))}{\phi'(\mathbb{E}_{\hat\pi_E}(u\circ g))}\mu(\hat\pi)\,\hat\pi(E)\,\hat\pi_E(F)} \\
&= \frac{\frac{\phi'(\mathbb{E}_{\pi}(u\circ g))}{\phi'(\mathbb{E}_{\pi_{E_F}}(u\circ g))}\mu(\pi)\,\pi(E\cap F)}{\sum_{\hat\pi\in\Delta}\frac{\phi'(\mathbb{E}_{\hat\pi}(u\circ g))}{\phi'(\mathbb{E}_{\hat\pi_{E_F}}(u\circ g))}\mu(\hat\pi)\,\hat\pi(E\cap F)}.
\end{aligned}$$

Since $\pi_{E_F} = \begin{cases} \frac{\pi_E}{\pi_E(F)} & \text{if } s \in E\cap F \\ 0 & \text{otherwise} \end{cases} = \pi_{E\cap F}$, $\left(\mu_{E,g}\right)_{F,g} = \mu_{E\cap F,g}$. This proves commutativity. ■

Before proving Proposition 3.3, we show a quite useful lemma and corollary deriving implications of strict concavity of ϕ.

Lemma A.5 *Assume ϕ is strictly concave. If $f \in B$ and $f \sim g$, then $\mu(\{\pi \in \Delta(S) \mid \mathbb{E}_\pi u\circ f \neq \mathbb{E}_\pi u\circ g\}) = 0$.*

Proof of Lemma A.5. Let $b = u\circ g$ and $a = u\circ f$. We show that for any f such that $\mu(\{\pi\in\Delta \mid \mathbb{E}_\pi a \neq \mathbb{E}_\pi b\}) > 0$ and $\mathbb{E}_\mu\phi(\mathbb{E}_\pi a) = \mathbb{E}_\mu\phi(\mathbb{E}_\pi b)$, $\alpha f + (1-\alpha)g \succ g$ for all $\alpha\in(0,1)$. To see this, fix such an a. By strict concavity of ϕ and $\mu(\{\pi\in\Delta \mid \mathbb{E}_\pi a \neq \mathbb{E}_\pi b\}) > 0$, $\mathbb{E}_\mu\phi(\alpha\mathbb{E}_\pi a + (1-\alpha)\mathbb{E}_\pi b) >$

$\mathbb{E}_\mu\left(\alpha\phi\left(\mathbb{E}_\pi a\right)+(1-\alpha)\phi\left(\mathbb{E}_\pi b\right)\right)=\alpha\mathbb{E}_\mu\phi\left(\mathbb{E}_\pi a\right)+(1-\alpha)\mathbb{E}_\mu\phi\left(\mathbb{E}_\pi b\right)=\mathbb{E}_\mu\phi\left(\mathbb{E}_\pi b\right)$. Since g is optimal in B and B is convex, if $f\in B$ with $f\sim g$ then

$$\mu\left(\{\pi\in\Delta\mid\mathbb{E}_\pi u\circ f\neq\mathbb{E}_\pi u\circ g\}\right)=0.$$

■

Corollary A.1 *Assume ϕ is strictly concave and ν is absolutely continuous with respect to μ. If $f\in B$, $f\sim g$ and $f(s)\sim g(s)$ on E^c then $\nu\left(\{\pi\in\Delta(S)\backslash\Delta(E^c)\mid\mathbb{E}_{\pi_E}u\circ f\neq\mathbb{E}_{\pi_E}u\circ g\}\right)=0$.*

Proof of Corollary A.1. Lemma A.5 and absolute continuity imply $\nu\left(\{\pi\in\Delta(S)\mid\mathbb{E}_\pi u\circ f\neq\mathbb{E}_\pi u\circ g\}\right)=0$. $f(s)\sim g(s)$ on E^c implies $\mathbb{E}_\pi u\circ f\neq\mathbb{E}_\pi u\circ g\iff\mathbb{E}_{\pi_E}u\circ f\neq\mathbb{E}_{\pi_E}u\circ g$ for $\pi\in\Delta(S)\backslash\Delta(E^c)$. Thus, $\nu\left(\{\pi\in\Delta(S)\backslash\Delta(E^c)\mid\mathbb{E}_{\pi_E}u\circ f\neq\mathbb{E}_{\pi_E}u\circ g\}\right)=0$. ■

Proof of Proposition 3.3. By Theorem 3.2, the smooth rule satisfies **DC**. We first show that $g\sim f$ implies $g\sim_{E,g,B}f$. By strict concavity of ϕ and Lemma A.5, if g is optimal in B and $f\in B$ with $g\sim f$ then $\mu\left(\{\pi\in\Delta\mid\mathbb{E}_\pi u\circ f\neq\mathbb{E}_\pi u\circ g\}\right)=0$. By Corollary A.1, absolute continuity and $f(s)\sim g(s)$ on E^c imply $\mu_{E,g}\left(\{\pi\in\Delta(S)\backslash\Delta(E^c)\mid\mathbb{E}_{\pi_E}u\circ f\neq\mathbb{E}_{\pi_E}u\circ g\}\right)=0$. Thus, for all $f\in B$ with $f(s)\sim g(s)$ on E^c, $g\sim f$ implies $g\sim_{E,g,B}f$.

It remains to show that $g\succ f$ implies $g\succ_{E,g,B}f$. We prove this by contradiction. Suppose $f\in B$ with $f(s)\sim g(s)$ on E^c, $g\succ f$ and $g\sim_{E,g,B}f$ (since **DC** holds, the possibility $f\succ_{E,g,B}g$ is not relevant). Since the set $\{f\in B\mid f(s)\sim g(s)\text{ on }E^c\}$ is convex and ϕ is strictly concave, by the arguments used in the proof of Lemma A.5 applied to $\succsim_{E,g,B}$,

$$\mu_{E,g}\left(\{\pi\in\Delta(S)\backslash\Delta(E^c)\mid\mathbb{E}_{\pi_E}u\circ f\neq\mathbb{E}_{\pi_E}u\circ g\}\right)=0.$$

From the definition of the smooth rule, this implies

$$\mu\left(\{\pi\in\Delta(S)\backslash\Delta(E^c)\mid\mathbb{E}_\pi u\circ f\neq\mathbb{E}_\pi u\circ g\}\right)=0,$$

implying $g\sim f$, a contradiction. ■

Proof of Proposition 3.4. Fix $\succsim\in\mathcal{P}^{SM}$ with $S=\{1,2,3\}$, $E=\{1,2\}$, $Z=\mathbb{R}$, $u(z)=z$ for $z\in Z$, and smooth ambiguity representation with $\mu\{(0.01,0,0.99)\}=\mu\{(0,0.01,0.99)\}=\frac{1}{2}$, and $\phi(z)=\begin{cases}2(z-1)^{\frac{1}{2}} & \text{if } z\geq 2\\ z & \text{if } z\leq 2\end{cases}$. The key feature of this ϕ that we exploit in our argument is that it exhibits a decreasing coefficient of Arrow-Pratt relative risk aversion (corresponding to decreasing relative ambiguity aversion) in the range $z\geq 2$. Consider

$g = (1700, 1700, 0)$, $h = (1000, 1000, 0)$, $f_1 = (1700, 500, 0)$, $f_2 = (500, 1700, 0)$ and $B = co\{g, h, f_1, f_2\}$. Since g weakly dominates all acts in B, it is unconditionally optimal (and, in fact, uniquely so). Observe that $h \sim f_1 \sim f_2$ (since $6 = \phi(10) = \frac{1}{2}\phi(17) + \frac{1}{2}\phi(5) = 4 + 2 = 6$). Since E consists of only two states, any update rule $U \in \mathcal{U}$ delivers a probability measure, call it $\hat{p}$, over the conditional probability of state 1 (which we denote by α). We will show that either $f_1 \succ_{E,g,B} h$ or $f_2 \succ_{E,g,B} h$ and thus that **DC1** fails.

Observe that $f_1 \succ_{E,g,B} h$ if and only if

$$\begin{aligned} & \phi(1000) < \mathbb{E}_{\hat{p}}\phi(\alpha 1700 + (1-\alpha)500) \\ \Leftrightarrow \quad & 999^{\frac{1}{2}} < \mathbb{E}_{\hat{p}}(499 + 1200\alpha)^{\frac{1}{2}}. \end{aligned}$$

Similarly, $f_2 \succ_{E,g,B} h$ if and only if

$$\begin{aligned} & \phi(1000) < \mathbb{E}_{\hat{p}}\phi(\alpha 500 + (1-\alpha)1700) \\ \Leftrightarrow \quad & 999^{\frac{1}{2}} < \mathbb{E}_{\hat{p}}(1699 - 1200\alpha)^{\frac{1}{2}}. \end{aligned}$$

We will show that the sum of these two inequalities holds. This implies that at least one of the two inequalities must hold.

The sum of the two right-hand side expectations is

$$\mathbb{E}_{\hat{p}}\left[(499 + 1200\alpha)^{\frac{1}{2}} + (1699 - 1200\alpha)^{\frac{1}{2}}\right]$$

and subtracting twice $999^{\frac{1}{2}}$ yields

$$\mathbb{E}_{\hat{p}}\left[(499 + 1200\alpha)^{\frac{1}{2}} + (1699 - 1200\alpha)^{\frac{1}{2}} - 2 * 999^{\frac{1}{2}}\right].$$

Examining the integrand for a fixed $\alpha \in [0, 1]$ gives

$$(499 + 1200\alpha)^{\frac{1}{2}} + (1699 - 1200\alpha)^{\frac{1}{2}} - 2 * 999^{\frac{1}{2}}. \tag{A.2}$$

As a function of α, (A.2) is strictly concave, and attains a maximum at $\alpha = \frac{1}{2}$. Therefore, if (A.2) is positive at the boundary points $\alpha = 0$ and $\alpha = 1$ then it is positive for all $\alpha \in [0, 1]$. At $\alpha = 0$ and $\alpha = 1$, (A.2) becomes

$$1699^{\frac{1}{2}} + 499^{\frac{1}{2}} - 2 * 999^{\frac{1}{2}} > 0.3433 > 0.$$

Therefore, for any $\hat{p}$, either $f_1 \succ_{E,g,B} h$ or $f_2 \succ_{E,g,B} h$ and thus **DC1** fails.

To show that **DC2** must fail, consider the same setting as above, except that now $g = (1000, 1000, 0)$ and $B = \{g\}$. Since B is a singleton, g is trivially unconditionally optimal. Recall that $g \sim f_1 \sim f_2$. That either $f_1 \succ_{E,g,B} g$ or $f_2 \succ_{E,g,B} g$ and thus that **DC2** fails follows from the same calculations as above. ■

Proof of Proposition 4.1. Given Theorem 2.1, it suffices to show that $T_{E,g}(V_{E,g,B}) = \arg\min_{p\in\Delta} G^{\star}_{E,g,B}\left(\int (u\circ g)\,dp, p\right)$. Define $I : \mathbb{R}^S \to \mathbb{R}$ by

$$I(a) = \begin{cases} V_{E,g,B}(a) & \text{if} \quad a \in u(X)^S \\ -\infty & \text{otherwise.} \end{cases}$$

Fix $a = u \circ g$. One step in the proof of Lemma A.2 showed that $T_{E,g}(V_{E,g,B}) = \partial^{\star} I(a) \cap \Delta(E)$. Therefore, we are done if we can show $\arg\min_{p\in\Delta} G^{\star}_{E,g,B}\left(\int adp, p\right) = \partial^{\star} I(a) \cap \Delta(E)$. To prove this equality we show the double inclusion of the sets.

$\subseteq$) Suppose $q \in \arg\min_{p\in\Delta} G^{\star}_{E,g,B}\left(\int adp, p\right)$. By definition, $G^{\star}_{E,g,B}$ is increasing in its first argument. For all $f \in \mathcal{A}$ s.t. $\int (u\circ f)\,dq \leq \int adq$, $\inf_{p\in\Delta} G^{\star}_{E,g,B}\left(\int (u\circ f)\,dp, p\right) \leq G^{\star}_{E,g,B}\left(\int (u\circ f)\,dq, q\right) \leq G^{\star}_{E,g,B}\left(\int adq, q\right) = \inf_{p\in\Delta} G^{\star}_{E,g,B}\left(\int adp, p\right)$. Therefore, $\int (u\circ f)\,dq \leq \int adq$ implies $V_{E,g,B}(u\circ f) \leq V_{E,g,B}(a)$ and thus $I(u\circ f) \leq I(a)$. So,

$$\begin{aligned} & I(b) \leq I(a) \text{ for all } b \in \mathbb{R}^S \text{ s.t. } \int bdq \leq \int adq \\ & \Rightarrow \left\{b \in \mathbb{R}^S : \int bdq \leq \int adq\right\} \subseteq \left\{b \in \mathbb{R}^S : I(b) \leq I(a)\right\} \\ & \Rightarrow \left\{b \in \mathbb{R}^S : I(b) > I(a)\right\} \subseteq \left\{b \in \mathbb{R}^S : \int bdq > \int adq\right\} \end{aligned}$$

thus $q \in \partial^{\star} I(a)$. Clearly $q \in \Delta$. It remains to show $q \in \Delta(E)$. Since $h \in \operatorname{int}(\mathcal{A})$, there is an $\varepsilon > 0$ such that $(a + \varepsilon q(E^c))_E(a - \varepsilon q(E)) \in u(X)^S$. Observe $\int (a + \varepsilon q(E^c))_E(a - \varepsilon q(E))dq = \int adq$. Since E is non-null, if $q(E^c) > 0$, $V_{E,g,B}((a + \varepsilon q(E^c))_E(a - \varepsilon q(E))) > V_{E,g,B}(a)$ contradicting $q \in \partial^{\star} I(a)$. So $q \in \partial^{\star} I(a) \cap \Delta(E)$.

$\supseteq$) Suppose $q \in \partial^{\star} I(a) \cap \Delta(E)$. Then $q \in \Delta$ and

$$\begin{aligned} \inf_{p\in\Delta} G^{\star}_{E,g,B}\left(\int adp, p\right) &= \inf_{p\in\Delta} \sup_{f\in\mathcal{A}} \left\{u\left(x^f\right) : \int (u\circ f)\,dp \leq \int (u\circ h)dp\right\} \\ &\geq u\left(x^h\right) = G^{\star}_{E,g,B}\left(\int adq, q\right), \end{aligned}$$

where the inequality holds because h satisfies the expectation constraint for any p and the last equality uses the definition of $\partial^\star I(a)$ to rule out doing any better than h when the measure is q. This shows $q \in \arg\min_{p\in\Delta} G^\star_{E,g,B}\left(\int adp, p\right)$. ■

Proof of Corollary 4.1. By Proposition 4.1, a rule in $\mathcal{W}$ satisfies **DC** if and only if $Q_E^{E,g,B} \cap \arg\min_{p\in\Delta} G^\star_{E,g,B}\left(\int (u\circ g)\, dp, p\right) \neq \emptyset$. For variational preferences, one can show (using e.g., [2008] and Theorem 3 in [2006a]) that

$$G^\star_{E,g,B}\left(\int (u\circ g)\, dp, p\right) = \int (u\circ g)\, dp + c_{E,g,B}(p).$$

■

Proof of Proposition 4.2. By Proposition 4.1 and the form of $G^\star$ for variational preferences, optimality of g is equivalent to

$$Q^{E,g,B} \cap \arg\min_{p\in\Delta}\left(\int (u\circ g)\, dp + c(p)\right) \neq \emptyset.$$

Let $r \in Q^{E,g,B} \cap \arg\min_{p\in\Delta}\left(\int (u\circ g)\, dp + c(p)\right)$. Since $p(E) = 0$ implies $c(p) = \infty$, $r(E) > 0$. For any $p \in \Delta(E)$, $\int (u\circ g)\, dp \otimes^E r + c\left(p \otimes^E r\right) \geq \int (u\circ g)\, dr + c(r)$. Since $\int (u\circ g)\, dp \otimes^E r = r(E)\int (u\circ g)\, dp + \int_{E^c} (u\circ g)\, dr$, $r(E)\int (u\circ g)\, dp + c\left(p\otimes^E r\right) \geq r(E)\int (u\circ g)\, dr_E + c(r)$. Thus $\int (u\circ g)\, dp + \frac{1}{r(E)}c\left(p\otimes^E r\right) \geq \int (u\circ g)\, dr_E + \frac{1}{(E)}c\left(r_E \otimes^E r\right)$, so r_E is an element of $\arg\min_{p\in\Delta(E)}$ $\left(\int (u\circ g)\, dp + \frac{1}{r(E)}c\left(p\otimes^E r\right)\right)$. Since E^c Savage-null implies $c_{E,g,B}(p) = +\infty$ for $p \notin \Delta(E)$ (because u is unbounded either above or below) and since subtracting the constant, $\min_{q\in\Delta(E)} c\left(q\otimes^E r\right)$, does not affect the minimization, $r_E \in \arg\min_{p\in\Delta}(\int (u\circ g)\, dp + c_{E,g,B}(p))$. Since

$$r_E \in Q_E^{E,g,B} \cap \arg\min_{p\in\Delta}\left(\int (u\circ g)\, dp + c_{E,g,B}(p)\right),$$

g is conditionally optimal by Corollary 4.1. It remains to show that $c_{E,g,B}$ satisfies the conditions required of an ambiguity index in a variational representation. $c_{E,g,B}$ is non-negative and grounded by construction. It is straightforward to verify that $c_{E,g,B}$ inherits convexity and lower semi-continuity from c. ■

References

C.D. Aliprantis, K.C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Second Edition (1999), Springer-Verlag, Berlin.

F. Anscombe, R. Aumann, A Definition of Subjective Probability, *Annals of Mathematical Statistics* **34** (1963), 199-205.

D. Bergemann, K.H. Schlag, Robust Monopoly Pricing: The Case of Regret, Cowles Foundation Discussion Paper No. 1527RR (2008).

S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, L. Montrucchio, Uncertainty Averse Preferences, Working Paper No. 77, *Collegio Carlo Alberto* (2008).

G. Chamberlain, Econometrics and Decision Theory, *Journal of Econometrics* **95** (2000), 255-283.

A.P. Dempster, A generalization of Bayesian inference, *Journal of the Royal Statistical Society B* **30** (1968) 205-247.

J. Eichberger, S. Grant, Dynamically Consistent Preferences with Quadratic Beliefs, *Journal of Risk and Uncertainty* **12** (1997), 189-207.

J. Eichberger, S. Grant, D. Kelsey, Updating Choquet Beliefs, *Journal of Mathematical Economics* **43** (2007), 888-899.

L.G. Epstein, A Definition of Uncertainty Aversion, *Review of Economic Studies* **66** (1999), 579-608.

L.G. Epstein, M. Le Breton, Dynamically Consistent Beliefs Must be Bayesian, *Journal of Economic Theory* **61** (1993), 1-22.

L.G. Epstein, M. Schneider, Recursive Multiple-Priors, *Journal of Economic Theory* **113** (2003), 1-31.

H. Ergin, F. Gul, A Theory of Subjective Compound Lotteries, *Journal of Economic Theory* **144** (2009), 899–929.

R. Fagin, J.Y. Halpern, A new approach to updating belief, in *Proceedings of the 6th Conference on Uncertainty and AI, 1990.* I.B.M. Research Report RJ 7222 (1989).

P. Ghirardato, Revisiting Savage in a Conditional World, *Economic Theory* **20** (2002), 83-92.

P. Ghirardato, M. Marinacci, Ambiguity Made Precise: A Comparative Foundation, *Journal of Economic Theory* **102** (2002), 251-289.

I. Gilboa, D. Schmeidler, Maxmin Expected Utility with Non-Unique Prior, *Journal of Mathematical Economics* **18** (1989), 141-153.

I. Gilboa, D. Schmeidler, Updating Ambiguous Beliefs, *Journal of Economic Theory* **59** (1993), 33-49.

S. Grant, A. Kajii, B. Polak, Preference for Information and Dynamic Consistency, *Theory and Decision* **48** (2000), 263-286.

J. Green, "Making Book Against Oneself," The Independence Axiom, and Nonlinear Utility Theory, *The Quarterly Journal of Economics* **102** (1987), 785-796.

H.J. Greenberg, W.P. Pierskalla, Quasi-Conjugate Functions and Surrogate Duality, *Cahiers Centre Etudes Recherche Operationnelle* **15(4)** (1973), 437-448.

E. Hanany, P. Klibanoff, Updating Preferences with Multiple Priors, *Theoretical Economics* **2** (2007), 261-298.

E. Hanany, P. Klibanoff, E. Marom, Dynamically Consistent Updating of Multiple Prior Beliefs - an Algorithmic Approach, mimeo., *Tel Aviv University* (2008).

L.P. Hansen, Beliefs, Doubts and Learning: The Valuation of Macroeconomic Risk, *American Economic Review* **97(2)** (2007), 1-30.

L.P. Hansen and T. Sargent, Robust Control and Model Uncertainty, *American Economic Review* **91(2)** (2001), 60-66.

T. Hayashi, Regret aversion and opportunity dependence, *Journal of Economic Theory* **139** (2008), 242-268.

T. Hayashi, Context dependence and consistency in dynamic choice under uncertainty: the case of anticipated regret, *Theory and Decision*, forthcoming, DOI 10.1007/s11238-009-9175-x.

J.Y. Jaffray, Bayesian Updating and Belief Functions, *IEEE Transactions on Systems, Man and Cybernetics*, **22** (1992), 1144-1152.

J.Y. Jaffray, Dynamic Decision Making with Belief Functions, in R.R. Yager et al. (Eds), *Advances in the Dempster-Shafer Theory of Evidence* (1994), 331-352, Wiley, New York.

P. Klibanoff, M. Marinacci, S. Mukerji, A Smooth Model of Decision Making under Ambiguity, *Econometrica* **73** (2005), 1849-1892.

P. Klibanoff, M. Marinacci, S. Mukerji, Recursive Smooth Ambiguity Preferences, *Journal of Economic Theory* **144** (2009), 930-976.

M.J. Machina, Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty, *Journal of Economic Literature* **27** (1989), 1622-1668.

E.F. McClennen, *Rationality and dynamic choice: foundational explorations* (1990), Cambridge University Press, Cambridge.

M.J. Machina, D. Schmeidler, A More Robust Definition of Subjective Probability, *Econometrica* **60** (1992), 745-780.

F. Maccheroni, M. Marinacci, A. Rustichini, Ambiguity Aversion, Robustness, and the Variational Representation of Preferences, *Econometrica* **74** (2006), 1447-1498.

F. Maccheroni, M. Marinacci, A. Rustichini, Dynamic Variational Preferences, *Journal of Economic Theory* **128** (2006), 4-44.

R. Nau, Uncertainty Aversion with Second-Order Utilities and Probabilities, *Management Science* **52(1)** (2006), 136-145.

W.S. Neilson, A Simplified Axiomatic Approach to Ambiguity Aversion, mimeo., *University of Tennessee* (2009).

K.G. Nishimura, H. Ozaki, A Simple Axiomatization of Iterated Choquet Objectives, mimeo., *University of Tokyo* (2003).

T. O'Donoghue, M. Rabin, Doing It Now or Later, *American Economic Review* **89(1)** (1999), 103-124.

E. Ozdenoren, J. Peck, Ambiguity Aversion, Games Against Nature, and Dynamic Consistency, *Games and Economic Behavior* **62** (2008), 106–115.

C.P. Pires, A Rule for Updating Ambiguous Beliefs, *Theory and Decision* **53** (2002), 137-152.

R.T. Rockafellar, *Convex Analysis* (1970), Princeton University Press, Princeton, New Jersey.

R. Sarin, P. Wakker, Dynamic Choice and Nonexpected Utility, *Journal of Risk and Uncertainty* **17** (1998), 87-119.

L.J. Savage, The Theory of Statistical Decision, *Journal of the American Statistics Association* 46 (1951), 55-67.

L.J. Savage, *The Foundations of Statistics* (1954), Wiley, New York.

D. Schmeidler, Subjective Probability and Expected Utility Without Additivity, *Econometrica* **57** (1989), 571-587.

U. Segal, Dynamic Consistency and Reference Points, *Journal of Economic Theory* **72** (1997), 208-219.

K. Seo, Ambiguity and Second-Order Belief, *Econometrica* **77(5)** (2009), 1575-1605.

T. Seidenfeld, A contrast between two decision rules for use with (convex) sets of probabilities: $\gamma-$maximin versus E-admissibility, *Synthese* **140** 1-2 (2004), 69-88.

G. Shafer, A Mathematical Theory of Evidence, *Princeton University Press* (1976), New Jersey.

M. Siniscalchi, Dynamic Choice under Ambiguity, mimeo., *Northwestern University* (2009).

J. Stoye, Statistical Decisions under Ambiguity, mimeo., *New York University* (2008).

J. Stoye, Axioms for Minimax Regret Choice Correspondences, mimeo., *New York University* (2008).

R. Strotz, Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies* **23** (3) (1955-6), 165-180.

T. Strzalecki, Axiomatic Foundations of Multiplier preferences, mimeo., *Harvard University* (2008).

P. Wakker, Are Counterfactual Decisions Relevant for Dynamically Consistent Updating under Nonexpected Utility?, mimeo., http://people.few.eur.nl /wakker/pdf/stratupd.pdf. (1997).

P. Walley, *Statistical Reasoning with Imprecise Probabilities* (1991), Chapman and Hall, London.

T. Wang, Conditional Preferences and Updating, *Journal of Economic Theory* **108** (2003), 286-321.

L.A. Wasserman, J. Kadane, Bayes' theorem for Choquet capacities, *Annals of Statistics* **18** (1990), 1328–1339.