Robust Maximum Likelihood Updating*

Elchin Suleymanov†

July 11, 2021

Abstract

There is a large body of evidence that decision makers frequently depart from Bayesian updating. This paper introduces a model, *robust maximum likelihood (RML) updating*, where deviations from Bayesian updating are due to multiple priors/ambiguity. The primitive of the analysis is the decision maker's preferences over acts before and after the arrival of new information. The main axioms characterize a representation where the decision maker's probability assessment can be described by a *benchmark prior*, which is reflected in her ex ante ranking of acts, and *a set of plausible priors*, which is revealed from her updated preferences. When new information is received, decision makers revise their benchmark prior within the set of plausible priors via the *maximum likelihood principle* in a way that ensures maximally *dynamically consistent* behavior, and update the new prior using Bayes' rule. RML updating accommodates most commonly observed biases in probabilistic reasoning.

JEL Classification: C11, D81, D91

Keywords: Non-Bayesian updating, multiple priors, ambiguity, maximum likelihood principle, dynamic consistency

*First version: November, 2018. This version: July, 2021. I am indebted to Yusufcan Masatlioglu and Shaowei Ke for their support and guidance. I also want to thank Matias Cattaneo, David Miller, Rocío Titiunik, Victor Aguiar, Christopher Chambers, Federico Echenique, Xiaosheng Mu, Collin Raymond, Philipp Strack, Tomasz Strzalecki, Gerelt Tserenjigmid for helpful comments and suggestions.

†Department of Economics, Purdue University. Email: esuleyma@purdue.edu.

1 Introduction

How do decision makers (DMs) update their beliefs when they receive new information? The answer to this question is critical in economic models and policy analyses where one tries to predict the consequence of releasing new information to market participants. The standard assumption in economics is that beliefs are updated using Bayes' rule. However, there is a large body of experimental and empirical evidence which shows that decision makers frequently deviate from Bayesian updating. For example, many decision makers tend to underweight base rates (*base rate neglect*), ignore informative signals (*conservatism*), interpret contrary evidence as supportive of their original beliefs (*confirmation bias*).[1] This paper introduces a model where deviations from Bayesian updating are due to ambiguity/multiple priors. The model can accommodate previously mentioned and other errors in probabilistic reasoning.

To illustrate how deviations from Bayesian updating can be related to multiple priors, consider the thought experiment due to Ellsberg (1961) where a DM is told that an urn contains 30 red balls and 60 blue or green balls in an unknown proportion. Let f_R, f_B, and f_G stand for bets which yield \$100 if the ball drawn from the urn is red, blue, and green, respectively, and \$0 otherwise. When no further information is given, many decision makers are indifferent between these bets, which is consistent with the prior that assigns equal probability to all colors.[2]

Now suppose the experimenter draws a ball from the urn and conveys to the DM that the ball is not green. How should the DM update her preferences given this information? In particular, should she still be indifferent between f_R and f_B? There are two arguments that can be made. First, following the principle of *dynamic consistency*, one can argue that since both f_R and f_B agree on the payoff assigned to the unrealized event (green), the information that this event is ruled out should not affect the original preference. Hence, indifference should be maintained ex post. On the other hand, the information that the ball is not green may suggest that the number of blue balls in the urn is greater than the number of green balls. Since there are only 30 red balls in the urn and 60 blue or green balls, one can also argue that f_B should be preferred to f_R ex post. This preference

[1] For a review of these findings, see, for example, Camerer (1995), Rabin (1998), Tversky (2004), and Benjamin (2019).

[2] Ellsberg argued that most decision makers would prefer to bet on red rather than blue or green. While Ellsberg style preferences are common, many experimental findings show that a significant number of decision makers are ambiguity neutral (see Binmore, Stewart, and Voorhoeve, 2012; Charness, Karni, and Levin, 2013; Stahl, 2014). In this paper, I consider both ambiguity neutral and ambiguity averse decision makers.

is incompatible with dynamic consistency, which is the key implication of Bayesian updating.

The intuition that decision makers may not perform Bayesian updating when they face ambiguity is confirmed by experiments and observations of practitioners' behavior. For example, in a similar dynamic Ellsberg experiment, Dominiak, Duersch, and Lefort (2012) find that a significant number of decision makers whose behavior can be characterized as ambiguity neutral are not Bayesian. In addition, many statistical tools used in practice (e.g. maximum likelihood estimation, hypothesis testing, etc.) are non-Bayesian even though conceivably many statisticians, and scientists in general, may be ambiguity neutral.

Most existing models tie the DM's ambiguity attitude (rather than ambiguity) to her response to new information, which forces an ambiguity neutral DM to update her beliefs using Bayes' rule. This is not only inconsistent with the intuition and observations described above but also unnatural as ambiguity attitude and belief updating are distinct concepts. The model proposed in this paper allows the DM depart from Bayesian updating when she faces ambiguity even if her attitude towards ambiguity is neutral.

I adopt a dynamic version of the classical Anscombe and Aumann (1963) setup. Let Ω be a finite set of states, and denote by $\Delta(\Omega)$ the set of all probability measures on Ω. An event is a member of $\mathcal{A}$, which is the collection of all subsets of Ω. The set of prizes (e.g. monetary payments) is a convex subset of a metric linear space, and an act is a function that assigns a prize to each state of the world.

The primitive of the analysis is a collection of preferences $\{\succcurlyeq_A\}_{A\in\mathcal{A}}$ where $\succcurlyeq_A$ represents the DM's preference over acts when she learns that event A occurs. The preference when the DM receives no information is $\succcurlyeq_\Omega$, which, for simplicity, is denoted by $\succcurlyeq$. The main axioms in this paper characterize a representation where the DM is endowed with *a benchmark prior* $\pi \in \Delta(\Omega)$, revealed from ex ante preferences $\succcurlyeq$, and *a set of plausible priors* $\mathbb{N}(\pi)$, revealed from updated preferences $\succcurlyeq_A$. For example, in the Ellsberg experiment the benchmark prior may assign equal probability to all colors, while any prior that assigns 1/3 probability to red is plausible. The benchmark prior π is interpreted as the DM's initial best guess where $\pi \in \mathbb{N}(\pi)$.

The axioms yield a novel updating rule, *robust maximum likelihood (RML) updating*, which can be described by two stages. In the first stage, the DM performs maximum likelihood updating within the set of plausible priors. That is, when the DM learns that an event A occurs, she restricts her attention to the subset of plausible priors which maximize the likelihood of this event. This set is denoted by

$$\mathbb{N}_A(\pi) = \arg\max_{\pi' \in \mathbb{N}(\pi)} \pi'(A).$$

Next, the DM chooses a new benchmark prior that induces maximally dynamically consistent behavior among all priors in $\mathbb{N}_A(\pi)$ and updates it using Bayes' rule. Maximal dynamic consistency ensures that the DM stays as "close" to her original benchmark prior as possible. A similar idea is also used in robust control literature where the benchmark prior π is treated as an "approximating model" that is not fully trusted, and models "further away" from π are seen as less appealing (see Hansen and Sargent, 2001; Strzalecki, 2011). When there is no ambiguity (i.e. $\mathbb{N}(\pi)$ is a singleton), RML updating reduces to Bayesian updating.

To illustrate how RML updating can accommodate a strict preference for f_B over f_R after the realization that the ball drawn from the Ellsberg urn is not green, let $\pi, \pi', \pi'' \in \mathbb{N}(\pi)$ where π is the benchmark prior, and π' and π'' represent two plausible priors when there are no green and blue balls in the urn, respectively.

	Red	Blue	Green
π	1/3	1/3	1/3
π'	1/3	2/3	0
π''	1/3	0	2/3

When the DM learns that the ball drawn from the urn is not green, maximum likelihood updating implies $\mathbb{N}_A(\pi) = \{\pi'\}$. The DM endowed with π' as her posterior prefers f_B over f_R.

RML updating provides explanations for most commonly observed biases in probabilistic reasoning. For example, consider confirmation bias, which is the tendency to interpret contrary evidence as supportive of original beliefs (see, for example, Rabin and Schrag, 1999, and references therein). Let $S = \{s_1, s_2\}$ be the set of payoff-relevant states, and denote by $\Sigma = \{\sigma_1, \sigma_2\}$ the set of signals. Suppose the DM assesses s_1 to be more likely than s_2, and σ_i is considered more likely than σ_j when the payoff-relevant state is s_i. The joint state space Ω and the benchmark prior π are illustrated below where $\mu > 1/2$ and $\alpha > 1/2$. A decision maker who displays confirmation bias assigns a higher probability to s_1 than a Bayesian agent with the prior π after observing σ_2.

	σ_1	σ_2
s_1	ω_{11}	ω_{12}
s_2	ω_{21}	ω_{22}

State Space Ω

	σ_1	σ_2
s_1	$\mu\alpha$	$\mu(1-\alpha)$
s_2	$(1-\mu)(1-\alpha)$	$(1-\mu)\alpha$

Benchmark Prior π

Imagine that the decision maker is not fully confident in the link between s_1 and the signals given by her benchmark prior and finds it plausible that the true information structure is given by π' (illustrated below) where $\alpha' < \alpha$. For example, the DM might find it plausible that there is a "bias" in the information source that is potentially unfavorable towards the state she originally finds more likely (i.e. the state s_1). Now suppose the DM observes the realization σ_2. Notice that under the benchmark prior the probability of observing σ_2 is $\mu + \alpha - 2\mu\alpha$. On the other hand, under the alternative plausible prior the probability of observing σ_2 is $\mu + \alpha - \mu\alpha - \mu\alpha'$. Hence, after observing σ_2 the DM performing maximum likelihood updating may change her benchmark prior from π to π'. This will result in a behavior that is consistent with confirmation bias.

	σ_1	σ_2
s_1	$\mu\alpha'$	$\mu(1-\alpha')$
s_2	$(1-\mu)(1-\alpha)$	$(1-\mu)\alpha$

Alternative Plausible Prior π'

An important question is whether one can identify the benchmark prior π and the set of plausible priors $\mathbb{N}(\pi)$ from preferences. The identification of π from ex ante preferences can be done as in Anscombe and Aumann (1963). To illustrate how $\mathbb{N}(\pi)$ can be identified, I first distinguish between *unambiguous* and *ambiguous* events. An event is unambiguous if there is full agreement among all plausible priors on its likelihood.[3] Otherwise, it is ambiguous. A prior is considered plausible if and only if it agrees with the benchmark prior on the likelihood of all unambiguous events.

Since the main axioms in this paper imply subjective expected utility (SEU) preferences, most existing approaches in the literature do not help us identify unambiguous events from preferences.[4]

[3]Several papers have provided a behavioral definition for unambiguous events. Epstein and Zhang (2001), Zhang (2002), and Gul and Pesendorfer (2014) define unambiguous events to be the ones which satisfy some versions of Savage's Sure Thing Principle. Ghirardato, Maccheroni, and Marinacci (2004) argue for a "relation based" approach and provide a definition for an act to be unambiguously preferred to another act.

[4]To be more precise, most existing approaches use only ex ante preferences to reveal ambiguity (see Epstein, 1999; Ghirardato and Marinacci, 2002). Hence, using the standard terminology, the DM whose ex ante preferences satisfy

A novelty in this paper is that unambiguous events are identified by comparing ex ante and ex post preferences. To see how this can be done, notice that the DM may not satisfy dynamic consistency when an ambiguous event is realized, as illustrated by the Ellsberg example. Dynamic consistency requires that if two acts f, g agree outside an event E and f is ex ante preferred to g, then f must still be preferred to g when E is realized. Formally, $f(\omega) = g(\omega)$ for all $\omega \in E^c$ and $f \succsim g$ imply $f \succsim_E g$. Since all plausible priors agree on the likelihood of unambiguous events, preferences are expected to satisfy dynamic consistency when an unambiguous event occurs. More importantly, consider an event $B \supseteq E$ where E is unambiguous. Since E is unambiguous and f, g agree on $B \backslash E$, one also expects that the DM's ex ante preference between f and g should be preserved when B is realized. According to this observation, E is defined to be *perfectly dynamically consistent* if for any two acts f, g that agree on E^c and any event $B \supseteq E$, f is ex ante preferred to g if and only if f is preferred to g when B is realized.

To identify $\mathbb{N}(\pi)$ from preferences, I first define an event E to be unambiguous when both E and E^c are perfectly dynamically consistent. The set $\mathbb{N}(\pi)$ consists of all probability measures on Ω which agree with the benchmark prior π on the likelihood of all unambiguous events. In the Ellsberg example, if the DM is indifferent between f_R and f_B ex ante but strictly prefers f_B to f_R when she is told that the ball drawn from the urn is not green, the definition implies that both $\{R, B\}$ and $\{G\}$ are ambiguous events, and hence there are multiple plausible priors which differ on the likelihood of these events.

The axioms imposed on $\{\succsim_A\}_{A \in \mathcal{A}}$ ensure that both π and $\mathbb{N}(\pi)$ can be identified. In addition to SEU axioms and standard axioms relating ex ante and ex post preferences, two main axioms weakening dynamic consistency are imposed. Consider a *minimal unambiguous event* E, i.e. any nonempty $D \subsetneq E$ is ambiguous. The first main axiom, *robust inference*, requires that the DM's ex ante willingness to bet on E is identical to her willingness to bet on E when $D \subsetneq E$ is ruled out. This reflects the DM's cautious attitude when she updates her prior. Since the DM knows the likelihood of unambiguous events but can only guess the likelihood of ambiguous events, when the DM receives new information, she wants her posterior not to differ too much from her benchmark prior on unambiguous events.

SEU axioms can be characterized as ambiguity neutral. The main difference in this paper is that an ambiguity neutral DM may still have multiple priors which will be reflected in ex post preferences even though it is not reflected in ex ante preferences.

The second main axiom, *consistency*, states that every $D \subsetneq E$ is a perfectly dynamically consistent event whenever E is a minimal unambiguous event. As stated earlier, when E is unambiguous both E and E^c must be perfectly dynamically consistent. On the other hand, consistency requires that an ambiguous event $D \subsetneq E$ is perfectly dynamically consistent. Hence, by definition, the realization of D^c must lead to a violation of perfect dynamic consistency. The intuition for this axiom is that when $D \subsetneq E$ is realized, the DM does not learn any new information that can help her make an inference regarding the relative likelihoods of the states within D. To illustrate, consider a DM who is told that an urn contains 25 red (R) balls and 75 blue (B), green (G), or yellow (Y) balls in an unknown proportion. Here, $\{B, G, Y\}$ is a minimal unambiguous event that is known to occur with 0.75 probability, and hence the axiom implies that the DM's preferences are dynamically consistent when, for example, $\{B, G\}$ is realized. This is because the information that $\{B, G\}$ has occurred does not say anything regarding the relative proportion of blue and green balls. Therefore, $f_B \succcurlyeq f_G$ if and only if $f_B \succcurlyeq_{\{B,G\}} f_G$.

This paper lies in the intersection of the literature on non-Bayesian updating and updating under ambiguity. The two most closely related papers are Gilboa and Schmeidler (1993) and Ortoleva (2012). Maximum likelihood updating was introduced by Gilboa and Schmeidler (1993) as a dynamic extension of the maxmin expected utility model. In their model, a DM endowed with a set of priors evaluates acts according to their minimal expected utility, where the minimum is taken over all priors in this set, and the DM performs maximum likelihood updating to revise the set of priors when she receives new information. In Gilboa and Schmeidler (1993), a DM whose behavior is consistent with the subjective expected utility model must follow Bayes' rule. On the other hand, I allow the DM to deviate from Bayesian updating when she faces ambiguity even if she is ambiguity neutral and also show how violations of dynamic consistency can be used to identify the set of priors she considers plausible. The second representation in this paper which has ambiguity averse decision makers is a special case of Gilboa and Schmeidler (1993).

Ortoleva (2012) axiomatizes a novel updating rule, the Hypothesis Testing (HT) model. In his model, the DM follows Bayes' rule for "normal" events but deviates from Bayesian updating when an "unexpected," small probability event occurs. In addition to allowing deviations from Bayesian updating, the HT model also imposes a structure on belief updating when a zero probability event occurs, which is not the case in RML updating. On the other hand, the HT model has two assumptions that are more general than RML updating: (i) the HT model imposes no structure on the set of priors the DM considers, whereas in RML updating every plausible prior must agree with the benchmark prior on unambiguous events, (ii) in the HT model any subjective second-order prior over the set of priors is allowed, while in RML updating it is uniform. In addition, when every state is non-null, as is the case in this paper, the HT model imposes almost no restriction on posteriors, and hence it is significantly more general than RML updating. Due to its generality, the HT model does not have the uniqueness properties of RML updating. In RML updating, both the benchmark prior and the set of plausible priors can be uniquely identified from preferences.

The paper proceeds as follows. Section 2 introduces the updating rule. In Section 3, I take the DM's ex ante and ex post preferences over acts as the primitive and provide a set of behavioral postulates that characterize the updating rule. Section 4 illustrates how the model can explain many well-known biases in probabilistic reasoning. In Section 5, I extend the model to allow for ambiguity averse preferences. Section 6 provides additional discussion on related literature. Section 7 concludes. Appendix includes all the proofs omitted from the main text.

2 Updating Rule

Let Ω be a finite set of states, and denote by $\Delta(\Omega)$ the set of all probability measures on Ω. The collection of all subsets of Ω (i.e. events) is denoted by $\mathcal{A}$. The decision maker's *probability assessment* is characterized by $(\pi, \mathcal{P})$ where $\pi \in \Delta(\Omega)$ is her benchmark prior and $\mathcal{P}$ is a partitioning of Ω that represents the collection of minimal unambiguous events. That is, for any $P \in \mathcal{P}$, the DM assesses that its likelihood is given by $\pi(P)$, and for any nonempty $D \subsetneq P$, the likelihood assigned by π reflects the DM's best guess. Since π is a probability measure, any arbitrary union of the events in $\mathcal{P}$ is unambiguous. A prior is plausible if it agrees with the benchmark prior on unambiguous events. The set of all plausible priors $\mathbb{N}_{\mathcal{P}}(\pi)$ is

$$\mathbb{N}_{\mathcal{P}}(\pi) = \{\pi' \in \Delta(\Omega) |\, \pi'(P) = \pi(P) \text{ for all } P \in \mathcal{P}\}.$$

The set $\mathbb{N}_{\mathcal{P}}(\pi)$ is uniquely defined, given the benchmark prior and the collection of minimal unambiguous events. Throughout this paper, I will maintain the assumption that π has full support.

Suppose the DM learns that event $A \in \mathcal{A}$ is realized. In the model belief updating can be described by two stages. In the first stage, the DM restricts her attention to the subset of plausible priors that maximize the likelihood that A occurs. Let $\mathbb{N}_{\mathcal{P},A}(\pi)$ denote this set. Formally,

$$\mathbb{N}_{\mathcal{P},A}(\pi) = \underset{\pi' \in \mathbb{N}_{\mathcal{P}}(\pi)}{\arg\max}\, \pi'(A). \tag{1}$$

Next, the DM chooses a new benchmark prior from $\mathbb{N}_{\mathcal{P},A}(\pi)$ and updates it using Bayes' rule. Note that when event A is realized, the new collection of minimal unambiguous events becomes $\{A \cap P | \, P \in \mathcal{P}\}$. I will require that the benchmark posterior π_A preserves the relative likelihood of any two states within each event in the new set of minimally unambiguous events:

$$\frac{\pi_A(\omega)}{\pi_A(\omega')} = \frac{\pi(\omega)}{\pi(\omega')} \quad \text{whenever } \omega, \omega' \in A \cap P \text{ for some } P \in \mathcal{P}. \tag{2}$$

For example, if $\pi \in \mathbb{N}_{\mathcal{P},A}(\pi)$, then condition 2 ensures that the DM's posterior is the same as the Bayesian posterior. This is desirable, since the DM has no reason to change her benchmark prior if it maximizes the likelihood of observing the realized event. As shown in Proposition 2, condition 2 is equivalent to requiring that among all Bayesian posteriors of the priors in $\mathbb{N}_{\mathcal{P},A}(\pi)$ the posterior π_A is the "closest" to the Bayesian posterior of π, where closeness is defined in terms of Kullback-Leibler divergence.

Given a probability assessment $(\pi, \mathcal{P})$, the posterior π_A is uniquely defined, and it is potentially distinct from the Bayesian posterior, which is denoted by $\pi(\cdot|A)$. The next proposition illustrates the connection between the posterior π_A and the benchmark prior π.

Proposition 1. *Let $(\pi, \mathcal{P})$ stand for the DM's probability assessment. For any $A \in \mathcal{A}$ and $\omega \in A$, the posterior π_A obtained via equations 1 and 2 satisfies*

$$\pi_A(\omega) = \pi(\omega | A \cap P_\omega) \cdot \pi(P_\omega | \textstyle\bigcup_{P \in \mathcal{P}:\, A \cap P \neq \emptyset} P) \tag{3}$$

where P_ω is the member of $\mathcal{P}$ that contains ω.

Proof. Notice that for any $\pi' \in \mathbb{N}_{\mathcal{P},A}(\pi)$, $\pi'(A \cap P) = \pi(P)$ for all $P \in \mathcal{P}$ with $A \cap P \neq \emptyset$. Since π_A is the Bayesian posterior of some $\pi' \in \mathbb{N}_{\mathcal{P},A}(\pi)$, for any $P, P' \in \mathcal{P}$ that have nonempty intersections with A,

$$\frac{\pi_A(P)}{\pi_A(P')} = \frac{\pi'(P|A)}{\pi'(P'|A)} = \frac{\pi'(A \cap P)}{\pi'(A \cap P')} = \frac{\pi(P)}{\pi(P')}.$$

This together with equation 2 show that the posterior π_A satisfies equation 3. $\square$

Definition 1. *Given a probability assessment $(\pi, \mathcal{P})$, the* ***robust maximum likelihood (RML)*** *updating rule assigns every event $A \in \mathcal{A}$ the posterior π_A given by equation 3.*

The RML updating rule reflects the DM's awareness of potential inaccuracy of her benchmark prior on ambiguous events, which necessitates a revision of the benchmark prior when new information is received, and her willingness to stay as "close" to her benchmark prior as possible. If all events are unambiguous (i.e. $\mathcal{P}$ is the collection of singletons), RML and Bayesian updating coincide. The next proposition provides an alternative representation for the updating rule which formalizes this intuition.

Proposition 2. *Let $(\pi, \mathcal{P})$ be a probability assessment. Denote by $\pi(\cdot|A)$ the Bayesian posterior of π when an event $A \in \mathcal{A}$ occurs. Then, π_A is the RML posterior of π if and only if*

$$\pi_A = \underset{\pi'_A \in B(\mathbb{N}_{\mathcal{P},A}(\pi))}{\arg\min} D_{KL}(\pi(\cdot|A) \,||\, \pi'_A)$$

where

$$D_{KL}(\pi(\cdot|A) \,||\, \pi'_A) = -\sum_{\omega \in A} \pi(\omega|A) \ln\left(\frac{\pi'_A(\omega)}{\pi(\omega|A)}\right)$$

and $B(\mathbb{N}_{\mathcal{P},A}(\pi))$ is the set of Bayesian posteriors of the priors in $\mathbb{N}_{\mathcal{P},A}(\pi)$.

To illustrate the RML updating rule, consider the Ellsberg experiment where the DM is told that an urn contains 30 red (R) balls and 60 blue (B) or green (G) balls in an unknown proportion. Let $(\pi, \mathcal{P})$ stand for the DM's probability assessment and suppose the benchmark prior π assigns equal probability to all colors. According to the information given to the DM, the collection of minimal unambiguous events is $\mathcal{P} = \{\{R\}, \{B, G\}\}$, and the set of plausible priors $\mathbb{N}_{\mathcal{P}}(\pi)$ is

$$\mathbb{N}_{\mathcal{P}}(\pi) = \{\pi' \in \Delta(\{R, B, G\}) |\, \pi'(R) = 1/3\}.$$

Suppose the experimenter draws a ball from the urn and tells the DM that the ball is not green.

In this case, the plausible prior that assigns zero probability to green maximizes the likelihood of the observation. Therefore, the RML posterior $\pi_{\{R,B\}}$ is

$$\pi_{\{R,B\}}(R) = 1/3, \quad \pi_{\{R,B\}}(B) = 2/3, \quad \pi_{\{R,B\}}(G) = 0.$$

When the DM is told that the ball is not red, the first stage of RML updating imposes no restriction on the posterior as all plausible priors agree on the event $\{B, G\}$. Hence, in this case the RML posterior $\pi_{\{B,G\}}$ is the same as the Bayesian posterior:

$$\pi_{\{B,G\}}(R) = 0, \quad \pi_{\{B,G\}}(B) = 1/2, \quad \pi_{\{B,G\}}(G) = 1/2.$$

3 Representation Theorem

Let X stand for the set of prizes which is assumed to be a convex subset of a metric linear space. For example, X can be the set of monetary outcomes the agent may receive ($X \subseteq \mathbb{R}$) or it can be the set of all lotteries over a finite set of outcomes Z (the classical Anscombe and Aumann (1963) setup). An act assigns a prize to each state of the world. The set of all acts is denoted by $\mathcal{F} = X^{\Omega}$. As is standard, constant acts are identified with X. A mixture of two acts is defined statewise: i.e. for any $f, g \in \mathcal{F}$ and $\alpha \in [0, 1]$, the act $\alpha f + (1 - \alpha)g \in \mathcal{F}$ is given by $(\alpha f + (1 - \alpha)g)(\omega) := \alpha f(\omega) + (1 - \alpha)g(\omega)$ for all $\omega \in \Omega$. For any event $A \in \mathcal{A}$ and $f, g \in \mathcal{F}$, $fAg \in \mathcal{F}$ is defined by $(fAg)(\omega) = f(\omega)$ if $\omega \in A$ and $(fAg)(\omega) = g(\omega)$ if $\omega \in A^c$.

I impose axioms on the collection of preferences $\{\succcurlyeq_A\}_{A \in \mathcal{A}}$ where $\succcurlyeq_A$ reflects the DM's preference over acts when she learns that $A \in \mathcal{A}$ is realized. The DM's preference over acts when she receives no information is $\succcurlyeq_\Omega$, which is simply denoted by $\succcurlyeq$. For notational simplicity, it is assumed that the DM is indifferent between all acts when the impossible event occurs, i.e. $f \sim_\emptyset g$ for all $f, g \in \mathcal{F}$.

The first three axioms are standard Weak Order, Archimedean, and Independence.

Axiom 1. *(Weak Order) For any $A \in \mathcal{A}$, $\succcurlyeq_A$ is complete and transitive.*

Axiom 2. *(Archimedean) For any $A \in \mathcal{A}$ and $f, g, h \in \mathcal{F}$ such that $f \succ_A g \succ_A h$, there exist $\alpha, \beta \in (0, 1)$ such that $\alpha f + (1 - \alpha)h \succ_A g$ and $g \succ_A \beta f + (1 - \beta)h$.*

Axiom 3. *(Independence) For any $A \in \mathcal{A}$, if $f \succ_A g$ and $\alpha \in (0,1]$, then $\alpha f + (1-\alpha)h \succ_A \alpha g + (1-\alpha)h$ for all $h \in \mathcal{F}$.*

Axiom 4 states that there exist best and worst alternatives and the DM is not indifferent between them. The existence of best and worst alternatives is not necessary for the representation, but it is assumed for the sake of convenience. The assumption that the DM is not indifferent between all alternatives is necessary for the benchmark prior to be identified from preferences.

Axiom 4. *(Nontriviality) There exist x^* and x_* such that $x^* \succ x_*$ and $x^* \succcurlyeq x \succcurlyeq x_*$ for all $x \in X$.*

The next axiom states that if f assigns a better prize to every state of the world than g does, then f must be preferred to g.

Axiom 5. *(Monotonicity) For any $A \in \mathcal{A}$, if $f(\omega) \succcurlyeq_A g(\omega)$ for all $\omega \in \Omega$, then $f \succcurlyeq_A g$. If, in addition, f is a constant act and $f(\omega) \succ_A g(\omega)$ for some $\omega \in A$, then $f \succ_A g$.*

Axiom 5 also requires that if the prize associated with a constant act is replaced in some state with a prize that is strictly worse, the DM considers the new act as strictly inferior. In addition to guaranteeing that the utility function derived from preferences is state independent, Axiom 5 also ensures that every state is assigned positive probability. Notice that by itself this axiom is weaker than strict monotonicity, which requires that if $f(\omega) \succcurlyeq_A g(\omega)$ for all $\omega \in \Omega$ and $f(\omega) \succ_A g(\omega)$ for some $\omega \in A$, then $f \succ_A g$. For example, if the DM evaluates acts according to their worst prize on A, then Axiom 5 is still satisfied even though strict monotonicity is violated. In the presence of previous axioms, Axiom 5 and strict monotonicity are equivalent.

The next axiom states that the ranking of two constant acts does not change when new information is received. This is because the prize associated with a constant act is the same regardless of the realized state and the utility of a prize is not affected by new information.

Axiom 6. *(Constant Act Preference Invariance) For any $A \in \mathcal{A} \backslash \emptyset$ and $x, y \in X$, $x \succcurlyeq y \Leftrightarrow x \succcurlyeq_A y$.*

Axiom 7 requires that when A is realized the DM must be indifferent between acts that agree on A. This axiom is known as *consequentialism*. In the literature, deviations from consequentialism are usually allowed to accommodate non-expected utility preferences (e.g. Machina, 1989) or to model a decision maker with an imperfect understanding of the state space (e.g. Minardi and

Savochkin, 2017). Since the main goal of this paper is to explore non-Bayesian updating when the DM has expected utility preferences and perfect understanding of the state space, consequentialism is retained in the representation.

Axiom 7. *(Consequentialism) For any $A \in \mathcal{A}$, if $f(\omega) = g(\omega)$ for all $\omega \in A$, then $f \sim_A g$.*

If, in addition to Axioms 1–7, one also assumes dynamic consistency, then belief updating must be Bayesian (e.g. see Ghirardato, 2002).[5] Dynamic consistency requires that if two acts agree outside an event, then the ranking of these acts should not change when this event occurs. In other words, this says that ex ante optimal plans must be optimal ex post. It can formally be stated as follows.

Dynamic Consistency: For any non-null $A \in \mathcal{A}$ and $f, g \in \mathcal{F}$, $fAg \succsim g \Leftrightarrow f \succsim_A g$.[6]

Axioms 1–5 guarantee that the benchmark prior can be uniquely revealed from ex ante preferences as in Anscombe and Aumann (1963). If the DM considers the benchmark prior as the only plausible prior (i.e. no ambiguity), then dynamic consistency is natural. In contrast, if the DM considers multiple priors plausible, it seems natural to revise the benchmark prior when new information arrives. Since the new prior may be distinct from the original benchmark prior, preferences may violate dynamic consistency. However, dynamic consistency should still be satisfied when an unambiguous event is realized. This is because the realization of such an event is not useful in distinguishing between plausible priors as all priors agree on the likelihood of these events, and hence there is no reason for the DM to deviate from her benchmark prior. Therefore, every unambiguous event must be dynamically consistent defined as below.

Definition 2. *$A \in \mathcal{A}$ is* ***dynamically consistent*** *if for any $f, g \in \mathcal{F}$, $fAg \succsim g \Leftrightarrow f \succsim_A g$.*[7]

[5]The connection between dynamic consistency and Bayesian updating is very general as shown in Epstein and Le Breton (1993). They show that if conditional preferences are derived in a way to ensure dynamic consistency and both ex ante and conditional preferences are "based on beliefs" (i.e. an event A is considered to be more likely than B if the DM prefers to bet on A rather than B), then standard axioms (Savage (1954) axioms except the Sure Thing Principle) guarantee that there exists a unique prior that represents beliefs and conditional beliefs are obtained using Bayes' rule.

[6]An event is *null* if the DM assigns it zero probability. Behaviorally, A is null if the DM is indifferent between any two acts that agree outside it: A is null if $f \sim g$ for any $f, g \in \mathcal{F}$ such that $f(w) = g(w)$ on A^c. Axiom 5 ensures that the only null event is $\emptyset$.

[7]Notice that dynamic consistency is a feature of preferences, not events. However, I use this terminology for the sake of brevity. Also note that $\emptyset$ is a dynamically consistent event by this definition, since we assumed that $f \sim_\emptyset g$ for all $f, g \in \mathcal{F}$. This only plays a role in simplifying the notation.

Suppose the analyst observes that the DM's preferences are dynamically consistent upon realization of an event. Can the analyst conclude that the DM considers this event as unambiguous? I provide an example which shows that this conclusion may not be accurate and then define a stronger version of dynamic consistency that captures unambiguous events.

Example 1. *Consider a DM who is told that an urn contains 50 red or blue balls and 50 green or yellow balls in unknown proportions. Let* $\Omega = \{R, B, G, Y\}$ *where* R*,* B*,* G*, and* Y *stand for states when the ball drawn from an urn is red, blue, green, and yellow, respectively. The set of plausible priors is*

$$\{\pi' \in \Delta(\{R, B, G, Y\}) |\, \pi'(R) + \pi'(B) = \pi'(G) + \pi'(Y) = 1/2\}.$$

When no further information is given, the DM may choose her benchmark prior as the one that assigns equal probability to all colors. Now suppose that a ball is drawn from the urn and the DM is told that the ball is either blue or green. This information does not favor either blue or green relative to the original information. Hence, it makes sense to assume that the benchmark posterior also assigns equal probability to blue and green. But then the event $\{B, G\}$ *is dynamically consistent. On the other hand, given the set of plausible priors, it is not possible to tell the exact probability that* $\{B, G\}$ *occurs.*

In Example 1, even though $\{B, G\}$ is a dynamically consistent event, it is still possible that $\{B, G\}$ is ambiguous. Consider two bets $f_B = (0, 100, 0, 0)$ and $f_G = (0, 0, 100, 0)$. If the DM's benchmark prior and posterior are as in the example, it must be that $f_B \sim f_G$ and $f_B \sim_{\{B,G\}} f_G$. Now suppose before learning that the ball drawn from the urn is either blue or green, the DM first learns that the ball is not yellow. The information that the ball is not yellow may suggest that the number of green balls in the urn is greater than the number of yellow balls. Since this information does not say anything regarding the relative proportion of red and blue balls, there is no reason for the DM to deviate from her original evaluation of the relative likelihood of R and B. But then the DM strictly prefers f_G over f_B when she learns that the ball is not yellow. Hence, $f_G \succ_{\{R,B,G\}} f_B$ even though $f_B \sim_{\{B,G\}} f_G$ and f_B and f_G agree on $\{R\}$. This would not be expected if $\{B, G\}$ was unambiguous.

This example motivates a new definition that captures unambiguous events via a stronger version of dynamic consistency

Definition 3. *$A \in \mathcal{A}$ is* ***perfectly dynamically consistent*** *if for any event $B \supseteq A$ and $f, g \in \mathcal{F}$,*

$$fAg \succcurlyeq_B g \Leftrightarrow f \succcurlyeq_A g.$$

For a Bayesian decision maker, every event should be perfectly dynamically consistent. Indeed, perfect dynamic consistency is implicitly assumed in the previous characterizations of Bayesian updating as $\{\succcurlyeq_A\}_{A\in\mathcal{A}}$ satisfies dynamic consistency only if every non-null event is perfectly dynamically consistent.[8]

Example 1 illustrates that when $\{\succcurlyeq_A\}_{A\in\mathcal{A}}$ does not satisfy dynamic consistency, there may be events that are dynamically consistent but not perfectly dynamically consistent. Since unambiguous events are expected to be perfectly dynamically consistent and it is possible to find a violation of perfect dynamic consistency for ambiguous events as in Example 1, an event is defined to be unambiguous if the event as well as its complement are perfectly dynamically consistent. The reason for requiring the complement to be perfectly dynamically consistent comes from the observation that the complement of an unambiguous event must be unambiguous.

Definition 4. *E is an* ***unambiguous event*** *if both E and E^c are perfectly dynamically consistent. The collection of all unambiguous events is denoted by $\mathcal{E}$. An event that does not belong to $\mathcal{E}$ is an* ***ambiguous event****.*

The next axiom ensures that the collection of unambiguous events form an *algebra*. A collection of events $\mathcal{E}$ is an algebra over Ω if (i) $\Omega \in \mathcal{E}$, (ii) $E \in \mathcal{E}$ implies $E^c \in \mathcal{E}$, and (iii) $E, E' \in \mathcal{E}$ implies $E \cap E' \in \mathcal{E}$.

Axiom 8. *(Algebra of Unambiguous Events) If $E, E' \in \mathcal{E}$, then $E \cap E' \in \mathcal{E}$.*

Intuitively, this axiom requires the following. Suppose events E, E^c, E', and E'^c are perfectly dynamically consistent so that both E and E' are unambiguous. Let f and g be two acts which agree outside $E \cap E'$. By the definition of perfect dynamic consistency, ex ante preference between f and g must be preserved when the DM learns either $B \supseteq E$ or $B' \supseteq E'$. But then it makes sense

[8] To see this, suppose $\{\succcurlyeq_A\}_{A\in\mathcal{A}}$ satisfies dynamic consistency. Let $A \in \mathcal{A}$, $B \supseteq A$ and $f, g \in \mathcal{F}$ be given. Dynamic consistency implies that $fAg \succcurlyeq g \Leftrightarrow f \succcurlyeq_A g$. On the other hand, since $B \supseteq A$, $fAg = (fAg)Bg$. Hence, dynamic consistency also implies that $fAg \succcurlyeq g \Leftrightarrow (fAg)Bg \succcurlyeq g \Leftrightarrow fAg \succcurlyeq_B g$. Therefore, A is perfectly dynamically consistent.

to assume that ex ante preference between f and g must still be preserved when the DM learns B and B' simultaneously. Therefore, perfectly dynamically consistent events are expected to be closed under intersection.

Since $\mathcal{E}$ is an algebra, there exists a unique partitioning of the state space that generates $\mathcal{E}$. A partition $\mathcal{P}$ of Ω generates the algebra $\mathcal{E}$ if $E \in \mathcal{E} \Leftrightarrow$ there exist $P_1, \ldots, P_k \in \mathcal{P}$ such that $P_1 \cup \cdots \cup P_k = E$. Let $\mathcal{P}_{\mathcal{E}}$ denote the partition that generates $\mathcal{E}$. The members of $\mathcal{P}_{\mathcal{E}}$ are minimal unambiguous events, i.e. any nonempty $D \subsetneq P$ where $P \in \mathcal{P}_{\mathcal{E}}$ is ambiguous. The next two axioms rely on $\mathcal{P}_{\mathcal{E}}$.

The following definitions will be useful for the statement of the next axiom.

Definition 5. *1. For any event A,* ***a bet on*** *A is an act f_A that yields the best prize on A and the worst prize outside A, i.e. $f_A = x^* A x_*$.*

2. For any $f \in \mathcal{F}$ and $A \in \mathcal{A}$, ***a certainty equivalent of*** *f* ***given*** *A is a sure outcome $c_A(f) \in X$ such that $f \sim_A c_A(f)$.*[9]

Since the DM knows the likelihood of unambiguous events but can only guess the likelihood of ambiguous events, she may want her posterior not to differ too much from her benchmark prior on unambiguous events. The next axiom, *robust inference*, reflects this cautious attitude when the DM updates her benchmark prior. Consider a minimal unambiguous event $P \in \mathcal{P}_{\mathcal{E}}$ and let f_P denote a bet on P. Suppose A is realized, and hence $A \cap P$ is a new minimal unambiguous event. Robust inference requires that the DM's willingness to bet on P is not affected when $D \subsetneq A \cap P$ is ruled out. That is, $c_A(f_P) \sim c_{A \setminus D}(f_P)$. In other words, since D is a proper subset of a minimal unambiguous event, when it is ruled out, the DM's considers the plausibility that it was a null event in the first place.

Axiom 9. *(Robust Inference) For any $A \in \mathcal{A}$ and $D \subsetneq A \cap P$ where $P \in \mathcal{P}_{\mathcal{E}}$,*

$$c_A(f_P) \sim c_{A \setminus D}(f_P).$$

In general, it is desirable if the DM's preferences are dynamically consistent unless there is a justifiable reason for deviation. The next axiom, *consistency*, requires that every $D \subsetneq P$, where P is

[9] The axioms stated so far and the assumption that X is convex guarantee that every act has a certainty equivalent.

a minimal unambiguous event, is perfectly dynamically consistent. Since an event E is unambiguous when both E and E^c are perfectly dynamically consistent and consistency requires an ambiguous event $D \subsetneq P$ to be perfectly dynamically consistent, the implication of the axiom is that D^c is not perfectly dynamically consistent.

Axiom 10. *(Consistency) Every $D \subsetneq P$ where $P \in \mathcal{P}_{\mathcal{E}}$ is perfectly dynamically consistent.*

Intuitively, when $D \subsetneq P$ is realized, the DM does not learn any information that can help her make an inference regarding the relative likelihoods of the states within D, and hence there is no reason for the DM to deviate from her original evaluation. To illustrate, suppose the DM is told that an urn contains 25 red (R) balls and 75 blue (B), green (G), or yellow (Y) balls in an unknown proportion. Since $\{B, G, Y\}$ is a minimal unambiguous event that is known to occur with 0.75 probability, consistency requires that the DM's preferences are dynamically consistent when, for example, $\{B, G\}$ is realized. This is because the information that $\{B, G\}$ is realized does not say anything regarding the relative proportion of blue and green balls, and hence there is no justification for deviation from the benchmark prior. Therefore, if the DM is ex ante indifferent between betting on blue and betting on green, she should remain indifferent when she learns that the ball drawn from the urn is either blue or green.

The next theorem provides a characterization result for the RML updating model.

Theorem 1. *The collection of preferences $\{\succsim_A\}_{A \in \mathcal{A}}$ satisfies Axioms 1-10 if and only if there exist a non-constant, affine utility function $u : X \to \mathbb{R}$ with $u(X) = [u(x_*), u(x^*)]$ and a probability assessment $(\pi, \mathcal{P})$, where π has full support on Ω, such that for any $A \in \mathcal{A}$,*

$$f \succsim_A g \Leftrightarrow \sum_{\omega \in \Omega} \pi_A(\omega) u(f(\omega)) \geq \sum_{\omega \in \Omega} \pi_A(\omega) u(g(\omega)) \tag{4}$$

and π_A is the RML posterior of π. Moreover, u is unique up to a positive affine transformation, π_A is unique for all $A \in \mathcal{A} \setminus \emptyset$, and $\mathcal{P}$ is uniquely revealed as $\mathcal{P}_{\mathcal{E}}$ unless $\mathcal{P} = \{\Omega, \emptyset\}$.

Sketch of the Proof

While showing the necessity of Axioms 1–7 is standard, the necessity of Axioms 8–10 is not trivial. The key step in the proof is showing that if $\{\succsim_A\}_{A \in \mathcal{A}}$ can be represented by equation 4, then the collection of minimal unambiguous events $\mathcal{P}_{\mathcal{E}}$ that is derived from preferences is exactly $\mathcal{P}$ unless $\mathcal{P} = \{\Omega, \emptyset\}$. This is achieved by showing that an event belongs to the algebra generated by $\mathcal{P}$ if and only if both this event and its complement satisfy perfect dynamic consistency. Once this is shown, Axioms 8–10 directly follow from the representation.

The proof of the claim that when E belongs to the algebra generated by $\mathcal{P}$ both E and E^c are perfectly dynamically consistent can be done using standard arguments. To prove the opposite, suppose E does not belong to the algebra generated by $\mathcal{P}$. As an illustration, suppose $E \subsetneq P$ for some $P \in \mathcal{P}$. Since Axiom 10 is satisfied only if E is perfectly dynamically consistent, it needs to be shown that E^c is not perfectly dynamically consistent. Consider bets on P and $P \setminus E$, i.e. $f_P = x^* P x_*$ and $f_{P \setminus E} = x_* E (x^* P x_*)$. Using the representation in equation 4, it is possible to find $\bar{z}, z \in X$ such that

$$f_P = \begin{pmatrix} x^* & \text{if } \omega \in E \\ x^* & \text{if } \omega \in P \setminus E \\ x_* & \text{if } \omega \in \Omega \setminus P \end{pmatrix} \sim \bar{z} \quad \text{and} \quad f_{P \setminus E} = \begin{pmatrix} x_* & \text{if } \omega \in E \\ x^* & \text{if } \omega \in P \setminus E \\ x_* & \text{if } \omega \in \Omega \setminus P \end{pmatrix} \sim \begin{pmatrix} x_* & \text{if } \omega \in E \\ z & \text{if } \omega \in P \setminus E \\ z & \text{if } \omega \in \Omega \setminus P \end{pmatrix}$$

where $\bar{z} \succ z$. Now suppose E^c is realized. From the representation, $\bar{z} \succ_{E^c} z \sim_{E^c} x_* E z$. Since π_{E^c} the RML posterior of π, $\bar{z} \sim_{E^c} f_P \sim_{E^c} f_{P \setminus E}$. But then, $f_{P \setminus E} = x_* E (x^* P x_*) \succ_{E^c} x_* E z$, in violation of dynamic consistency. Therefore, E^c is not perfectly dynamically consistent. This shows that E is an ambiguous event, i.e. $E \notin \mathcal{E}$. The case when $E \cap P \neq \emptyset$ and $E \cap P' \neq \emptyset$ for at least two distinct $P, P' \in \mathcal{P}$ is similar.

To prove sufficiency, first observe that Axioms 1–5 yield an SEU representation for each $A \in \mathcal{A}$ as in Anscombe and Aumann (1963). Moreover, Axiom 6 guarantees that the same utility function can be used for all $\succcurlyeq_A$, and Axioms 5 and 7 guarantee that π_A has full support on A and $\pi_A(A^c) = 0$. Therefore, it only needs to be shown that each π_A is the RML posterior of π. Let $\mathcal{E}$ be given by Definition 4, and $\mathcal{P}_{\mathcal{E}}$ is the partition that generates $\mathcal{E}$. The DM's probability assessment is $(\pi, \mathcal{P}_{\mathcal{E}})$. Since preferences are dynamically consistent on $E \in \mathcal{E}$, standard arguments show that updating is Bayesian when E is realized, consistent with RML updating.

Consider an event $A \notin \mathcal{E}$. The next step is to construct an unambiguous event $B \in \mathcal{E}$ such that $B \supseteq A$ and B is the smallest such event with respect to set inclusion. To construct B, let

$\mathcal{P}_\mathcal{E} = \{P_1, \ldots, P_n\}$ and consider $J \subseteq \{1, \ldots, n\}$ such that $P_j \cap A \neq \emptyset$ for all $j \in J$. The event B is given by $B = \cup_{j \in J} P_j$. Since $B \in \mathcal{E}$, according to the previous paragraph, the DM performs Bayesian updating when B is realized. Hence, for any $j, j' \in J$,

$$\frac{\pi_B(P_j)}{\pi_B(P_{j'})} = \frac{\pi(P_j)}{\pi(P_{j'})}.$$

On the other hand, Axiom 9 implies that $\pi_A(P_j) = \pi_{A \cup P_j}(P_j)$ for all $j \in J$. Therefore, iterative application of Axiom 9 yields $\pi_A(P_j) = \pi_B(P_j)$ for all $j \in J$, which implies

$$\frac{\pi_A(P_j)}{\pi_A(P_{j'})} = \frac{\pi(P_j)}{\pi(P_{j'})}.$$

Consider an event $A \cap P_j$ where $j \in J$. By Axiom 10, $A \cap P_j$ is perfectly dynamically consistent, and hence standard arguments guarantee that $\pi_{A \cap P_j}$ is the Bayesian posterior of π. Moreover, perfect dynamic consistency also ensures that $\pi_{A \cap P_j}(\omega) = \pi_A(\omega | P_j)$ for all $\omega \in A \cap P_j$. Therefore, for any $\omega, \omega' \in A \cap P_j$,

$$\frac{\pi_A(\omega)}{\pi_A(\omega')} = \frac{\pi_{A \cap P_j}(\omega)}{\pi_{A \cap P_j}(\omega')} = \frac{\pi(\omega)}{\pi(\omega')}.$$

This together with the conclusion of the previous paragraph and Proposition 1 show that π_A is the RML posterior of π, concluding the proof of sufficiency.

Lastly, the uniqueness of u up to a positive affine transformation and the uniqueness of π_A for each $A \in \mathcal{A}$ are standard results. The uniqueness of $\mathcal{P}$ is implied by the proof of necessity where the equivalence of $\mathcal{P}$ and $\mathcal{P}_\mathcal{E}$ is shown.

4 Applications

In this section, I show how the RML updating rule can help explain commonly observed biases in probabilistic reasoning. While all the examples in this section only use the first (maximum likelihood) stage of RML updating, in more realistic examples with a larger state space the first stage of RML updating by itself will not produce a unique posterior in general, and hence the second stage of RML updating is needed to make meaningful predictions.

Let $\Omega \equiv S \times \Sigma$ where $S = \{s_1, s_2\}$ is the set of payoff-relevant states and $\Sigma = \{\sigma_1, \sigma_2\}$ is the set of possible signals. The DM's benchmark prior π is represented by two parameters (μ, α) where $\mu > 1/2$ is the probability that the payoff relevant state is s_1 and $\alpha > 1/2$ denotes the probability that the DM receives signal σ_i when the payoff-relevant state is s_i. $\mu > 1/2$ reflects the DM's initial evaluation that s_1 is more likely.

	σ_1	σ_2
s_1	ω_{11}	ω_{12}
s_2	ω_{21}	ω_{22}

State Space Ω

	σ_1	σ_2
s_1	$\mu\alpha$	$\mu(1-\alpha)$
s_2	$(1-\mu)(1-\alpha)$	$(1-\mu)\alpha$

Benchmark Prior π

4.1 Confirmation Bias

I now revisit the confirmation bias phenomenon illustrated in the introduction. The DM who displays confirmation bias interprets contrary evidence as supportive of her original beliefs (Rabin and Schrag, 1999). That is, when the DM observes σ_2, she may find s_1 to be at least as likely as before.[10] Formally,

$$\pi_{\text{conf. bias}}(s_1|\sigma_2) \geq \mu > \frac{\mu - \mu\alpha}{\mu + \alpha - 2\mu\alpha} = \pi(s_1|\sigma_2).$$

Confirmation bias is frequently reported in experiments (e.g. see Lord, Ross, and Lepper, 1979; Darley and Gross, 1983).

To see how RML updating can accommodate confirmation bias, suppose the DM finds it plausible that there is a "bias" in the information source that is potentially unfavorable towards the state she originally finds more likely. Even though such a DM unambiguously knows the probability that the payoff-relevant state is s_1, the event that σ_i occurs when s_1 is the payoff-relevant state is ambiguous. Therefore, when the DM observes σ_2, she revises her benchmark prior to account for the possibility that σ_2 might be more likely than σ_1 when the payoff-relevant state is s_1. Notice that the DM might find the existence of a bias plausible even though it is not her benchmark belief. Once the bias is seen as plausible, the DM is endowed with multiple priors and uses signal realizations to distinguish between plausible priors.

[10] The milder version of confirmation bias states that after observing σ_2 the agent finds s_1 more likely than a Bayesian agent with the prior π. RML updating can accommodate both the milder version of confirmation bias and the more extreme version as illustrated in this section.

To formalize the intuition, let $\mathcal{P}$ stand for the set of minimal unambiguous events in this example, which is given by

$$\mathcal{P} = \{\{\omega_{11}, \omega_{12}\}, \{\omega_{21}\}, \{\omega_{22}\}\}.$$

In RML updating, the DM uses the maximum likelihood method to make an inference regarding the direction of the bias. Given the benchmark prior and the set of minimal unambiguous events, the RML posterior π_{σ_2} is

$$\pi_{\sigma_2}(s_1) = \frac{\mu}{\mu + \alpha - \mu\alpha} \quad \text{and} \quad \pi_{\sigma_2}(s_2) = \frac{\alpha - \mu\alpha}{\mu + \alpha - \mu\alpha}$$

Hence, the DM performing RML updating believes that s_1 is strictly more likely than before when she observes σ_2, consistent with confirmation bias.

4.2 Other Behavioral Biases

I consider three other commonly observed deviations from Bayesian updating: base rate neglect, conservatism, and overconfidence.

Base Rate Neglect: In a series of experiments, Kahneman and Tversky (1973) and Bar-Hillel (1980) show that decision makers tend to ignore the base rate μ in their predictions. In the well-known "cab problem," DMs are told that there are two cab companies, Blue and Green, one of which has been involved in a hit-and-run accident. The proportion of Blue cabs in the city is 85%, and the cab involved in the accident was identified as Green by a witness who is accurate 80% of the time. When DMs are asked to predict the probability that the car involved in the accident is Green, the median and modal response is 0.8, much higher than the Bayesian posterior (≈ 0.41).

To see how RML updating can explain this phenomenon, imagine that the DM has full confidence in the likelihood information α but does not have full confidence in the base rate μ. Even when the DM does not have full confidence in the base rate, the event that consists of states in which she gets "correct" signals is still unambiguous and known to occur with $\alpha = 0.8$ probability. Similarly, the event that corresponds to states in which she gets "wrong" signals is unambiguously assigned $1 - \alpha = 0.2$ probability. Given this set of minimal unambiguous events, the RML posterior is exactly equal to the median response in the cab problem (see the figure below).

Conservatism: DMs display conservatism bias when they overweight the base rate and underweight the likelihood information (see Edwards, 1968, for the classical experimental findings). RML updating results in conservatism bias when decision makers have full confidence in the base rate information but not in the likelihood information. This is exactly the mirror image of the base rate neglect phenomenon.

Overconfidence: Decision makers who treat their private information as more precise than it actually is are described as overconfident (see Odean, 1998, for a review of psychology literature on overconfidence and its implications for asset markets). Suppose s_1 = good market, s_2 = bad market, σ_1 = good jobs report, and σ_2 = bad jobs report. Overconfident investors tend to over-invest when they observe good jobs report and under-invest when they observe bad jobs report. RML updating results in overconfidence when the DM has full confidence in the likelihood information but is not completely sure whether the "correct" signal is more likely when the state is s_1 or s_2.

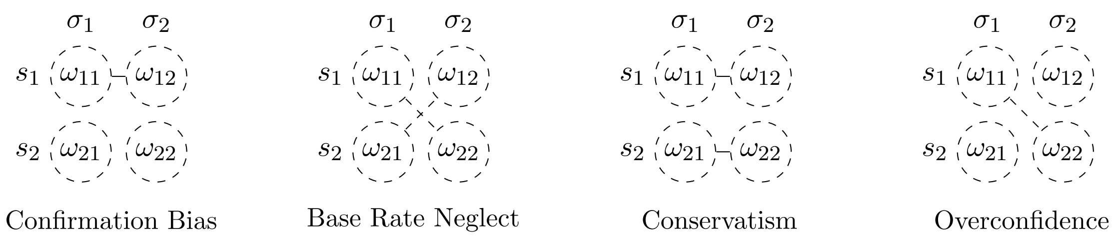

Figure 1: Behavioral biases. This figure illustrates partitions used to explain each behavioral bias. States connected by a line belong to the same partition element.

5 Ambiguity Averse Preferences

In Section 3, ambiguity is reflected in the DM's belief updating even though the DM's preferences display neutral attitude towards ambiguity. In this section, I consider an ambiguity averse DM whose preferences are consistent with the maxmin expected utility model of Gilboa and Schmeidler (1989).

An ambiguity averse DM is expected to satisfy all the axioms that characterize the subjective expected utility model in Section 3 except independence and consistency. An ambiguity averse DM may not satisfy independence due to strict preference for randomization, which may arise as randomization potentially limits exposure to ambiguity. In the Ellsberg example, the DM may be indifferent between betting on blue (f_B) and betting on green (f_G) but may strictly prefer the 50-50 randomization of these bets, in violation of independence (see below). This is because the 50-50 randomization of f_B and f_G gives the DM the same monetary outcome regardless of whether the ball drawn from the urn is blue or green, and hence it can be seen as a perfect hedge against ambiguity.

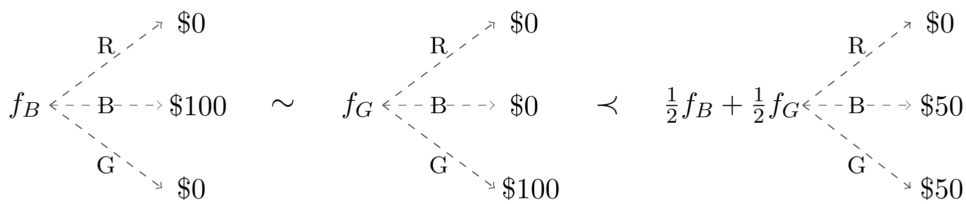

Figure 2: Preference for randomization due to ambiguity aversion.

To illustrate why an ambiguity averse DM may not satisfy consistency, recall the example where the DM is told that an urn contains 25 red (R) balls and 75 blue (B), green (G), or yellow (Y) balls in an unknown proportion. Here, $\{B, G, Y\}$ is a minimal unambiguous events, and hence consistency implies that $\{B, G\}$ is dynamically consistent. Let f_1 be the act that yields \$100 if the ball drawn from the urn is blue, and \$0 otherwise. Let f_2 be the act that yields \$25 if the ball drawn from the urn is either blue or green, and \$0 otherwise. When the DM learns that $\{B, G\}$ is realized, she may have a strict preference for f_2 over f_1 as f_2 perfectly hedges against ambiguity but f_1 does not. Should this DM have a strict preference for f_2 over f_1 ex ante as required by consistency? This is not obvious because ex ante f_2 is not a perfect hedge against ambiguity and it has much lower expected value than f_1 for many plausible priors.

In addition to the axioms in Section 3 except for independence and consistency, I impose two new axioms on preferences that characterize RML updating for ambiguity averse DMs. Let $\mathcal{E}$ stand for the collection of unambiguous events as in Definition 4, and $\mathcal{P}_{\mathcal{E}}$ is the collection of minimal unambiguous events. Acts that are constant on minimal unambiguous events are unambiguous acts.

Definition 6. *$f \in \mathcal{F}$ is an **unambiguous act** if $f(\omega) = f(\omega')$ whenever $\omega, \omega' \in P$ for some $P \in \mathcal{P}_{\mathcal{E}}$. The set of all unambiguous acts is denoted by $\mathcal{F}^{ua} \subseteq \mathcal{F}$.*

Axiom 11 imposes independence on the set of all unambiguous acts. Since unambiguous acts have no exposure to ambiguity, strict preference for randomization between unambiguous acts cannot be justified by ambiguity aversion.

Axiom 11. *(Weak Independence) For any $A \in \mathcal{A}$, $f, g, h \in \mathcal{F}^{ua}$ and $\alpha \in (0,1]$, $f \succ_A g$ implies $\alpha f + (1-\alpha)h \succ_A \alpha g + (1-\alpha)h$.*

Weak independence is consistent with both ambiguity averse and ambiguity loving attitude. The next axiom imposes that the DM is ambiguity averse.

Axiom 12. *(Ambiguity Aversion) For any $A \in \mathcal{A}$, $D \subsetneq A \cap P$ where $P \in \mathcal{P}_{\mathcal{E}}$, and $f \in \mathcal{F}$, $x^* D f \sim_A f$.*

Consider a minimal unambiguous event P and suppose A is realized. After this realization, $A \cap P$ is a minimal unambiguous event, and hence $D \subsetneq A \cap P$ is ambiguous. Axiom 12 requires that the DM is indifferent between an act f and an act which agrees with f outside D and yields the best prize on D. This is an extreme attitude that completely disregards that in the second act the DM receives the best prize when D occurs. This is due to two assumptions: (i) in the model, every $D \subsetneq A \cap P$ is treated as maximally ambiguous, (ii) given the set of plausible priors, the DM evaluates acts according to their worst case utility as in Gilboa and Schmeidler (1989).[11]

The next theorem provides a characterization for RML updating for ambiguity averse DMs.

Theorem 2. *The collection of preferences $\{\succsim_A\}_{A\in\mathcal{A}}$ satisfies Axioms 1, 2, 4-9, 11, and 12 if and only if there exist a non-constant, affine utility function $u : X \to \mathbb{R}$ with $u(X) = [u(x_*), u(x^*)]$, and a probability assessment $(\pi, \mathcal{P})$, where π has full support on $\mathcal{P}$, such that for any $A \in \mathcal{A}$,*

$$f \succsim_A g \Leftrightarrow \min_{\pi_A \in B(\mathbb{N}_{\mathcal{P},A}(\pi))} \sum_{\omega\in\Omega} \pi_A(\omega) u(f(\omega)) \geq \min_{\pi_A \in B(\mathbb{N}_{\mathcal{P},A}(\pi))} \sum_{\omega\in\Omega} \pi_A(\omega) u(g(\omega)) \tag{5}$$

where $B(\mathbb{N}_{\mathcal{P},A}(\pi))$ is the set of Bayesian posteriors of the priors in $\mathbb{N}_{\mathcal{P},A}(\pi)$. Moreover, u is unique up to a positive affine transformation, $\mathcal{P}$ is uniquely revealed as $\mathcal{P}_{\mathcal{E}}$, and the set $\mathbb{N}_{\mathcal{P},A}(\pi)$ is unique for all $A \in \mathcal{A} \setminus \emptyset$.

When the DM's preferences are consistent with the maxmin expected utility model, the benchmark prior can no longer be identified from preferences. In fact, the only role of the benchmark

[11] A natural extension of the model analyzed here may impose a less extreme version of ambiguity aversion. For example, the DM may perform maxmin within a *subset* of plausible priors rather than the full set. Analysis of such extensions is left for future work.

prior in Theorem 2 is to define the likelihood of unambiguous events. Since the DM uses the worst case scenario to evaluate acts, even if she has a guess for the likelihood of ambiguous events, this will not be reflected in her preferences. This distinguishes Theorem 2 from Theorem 1 where the benchmark prior can be revealed from ex ante preferences. Because of the limited role the benchmark prior plays in Theorem 2, with maxmin expected utility preferences RML updating coincides with the maximum likelihood updating rule of Gilboa and Schmeidler (1993).

Recall that an event E is defined to be unambiguous if and only if both E and E^c are perfectly dynamically consistent, which holds for both ambiguity neutral and ambiguity averse DMs. Example 1 shows that an event may fail to be perfectly dynamically consistent even when it is dynamically consistent. However, if $\{\succsim_A\}_{A \in \mathcal{A}}$ satisfies the axioms in Theorem 2, then every dynamically consistent event is perfectly dynamically consistent. Therefore, if the DM is ambiguity averse, it is possible to identify unambiguous events via dynamic consistency.

Proposition 3. *If the collection of preferences $\{\succsim_A\}_{A \in \mathcal{A}}$ can be represented by equation 5, then every dynamically consistent event is perfectly dynamically consistent, and hence an event E is unambiguous if and only if both E and E^c are dynamically consistent.*

6 Related Literature

This paper lies in the intersection of the literature on non-Bayesian updating and updating under ambiguity. As discussed in the introduction, the two most closely related papers are Gilboa and Schmeidler (1993) and Ortoleva (2012). In this section I provide a discussion on other related work.

Non-Bayesian Updating

Epstein (2006) and Epstein, Noor, and Sandroni (2008) axiomatize a non-Bayesian updating model where decision makers may be tempted to update their beliefs using a prior different from their original prior. For example, they might be tempted to overreact to new information. In RML updating, decision makers also revise their original prior when they receive new information, but this is not due to temptation but rather due to ambiguity and willingness to make an inference. In addition, the primitive of the analysis is different in these papers.

Zhao (2017) proposes a model that allows DMs to update their beliefs when they receive new information of the form "event A is more likely than event B." In his model, the posterior minimizes Kullback-Leibler (KL) divergence from the prior subject to the constraint that the posterior assigns a higher probability to A than B.[12] In RML updating, the idea is similar as the DM chooses her posterior by minimizing KL divergence from the Bayesian posterior of the benchmark prior subject to the constraint that the new prior assigns the maximal likelihood to the observed event among all plausible priors. However, the main focus in Zhao (2017) is different from the one in this paper, and these papers use different primitives of analysis.

A few recent decision theoretic papers axiomatize models that can accommodate belief updating biases. Zhao (2018) axiomatizes a non-Bayesian updating model, called similarity-based updating, that builds on the representativeness heuristic of Kahneman and Tversky (1972). Kovach (2021b) axiomatizes an updating rule, called conservative updating, where the DM's posterior is a mixture of her prior and Bayesian posterior as in Epstein, Noor, and Sandroni (2010). As opposed to RML updating, conservative updating may violate both consequentialism and dynamic consistency. While similarity-based updating, conservative updating, and RML updating can accommodate some of the same behavioral biases, the underlying behavioral motivations for these models are completely different.

Many behavioral models in the literature explain non-Bayesian updating by assuming some type of bounded rationality. This includes assuming imperfect memory (Mullainathan, 2002a; Gennaioli and Shleifer, 2010; Wilson, 2014), coarse thinking (Mullainathan, 2002b; Mullainathan, Schwartzstein, and Shleifer, 2008), the use of representativeness heuristic (Kahneman and Tversky, 1972), or incorrect modeling (Barberis, Shleifer, and Vishny, 1998; Rabin and Schrag, 1999) by decision makers. All of these models are non-axiomatic and focus on particular applications.

A natural setup where the DM might have a benchmark prior and a set of plausible priors is when the DM receives forecasts or recommendations from different experts. Levy and Razin (2021) study the problem of a DM who receives forecasts from multiple Bayesian forecasters and uses the maximum likelihood method to form an explanation for these forecasts. The decision maker then forms her posterior by applying Bayes' rule to the most likely explanation. They show that this

[12] More recently, Dominiak, Kovach, and Tserenjigmid (2021) consider more general forms of information structures and more general distance measures.

updating can lead to reliance on extreme forecasts and ignoring moderate forecasts. Ke, Wu, and Zhao (2021) take a decision theoretic approach and study a problem where recommendations may not necessarily come from a Bayesian agent. They show that there exists no updating rule that satisfies certain desirable axioms. The main difference in this paper is that I impose axioms on the DM's preferences which allows me to derive the set of plausible priors endogenously from the DM's preferences.

Updating Under Ambiguity

There is a growing literature on updating under ambiguity, as reviewed in Machina and Siniscalchi (2014) and Gilboa and Marinacci (2016). As an alternative to maximum likelihood updating, one natural way to update under ambiguity is updating each prior according to Bayes' rule. This method is known as full Bayesian updating and was axiomatized by Pires (2002) using the maxmin expected utility model. Epstein and Schneider (2007) propose another model where the DM applies Bayes' rule only to a subset of priors that are considered sufficiently likely. A few recent papers axiomatize updating rules that generalize full Bayesian and maximum likelihood updating (see, for example, Cheng, 2021; Hill, 2021; Kovach, 2021a). The main difference between these papers and RML updating is that when the DM has subjective expected utility preferences, these updating rules coincide with Bayesian updating.

As discussed in Section 3, updating rules under ambiguity that satisfy consequentialism may violate dynamic consistency. Dynamic consistency is usually considered a desirable property, and the fact that it may not be satisfied under ambiguity has been a subject of criticism in the literature (Al-Najjar and Weinstein, 2009). Epstein and Schneider (2003) retain dynamic consistency by restricting the set of events on which the DM can update her beliefs. Hanany and Klibanoff (2007) characterize dynamically consistent maxmin expected utility preferences without any restriction on the set of conditioning events and show that updated preferences must depend on the initial menu the DM is offered. Siniscalchi (2011) allows deviations from dynamic consistency but assumes that the DM can anticipate her future deviations. Gul and Pesendorfer (2018) impose a weaker version of dynamic consistency which can be interpreted as "not all news can be bad news" and show that neither maximum likelihood updating nor full Bayesian updating satisfies this property. RML updating also treats dynamic consistency as a desirable property by ensuring that the DM is maximally dynamically consistent. That is, any deviation from dynamic consistency is due to the DM's willingness to use new information to make an inference on the set of plausible priors.

A few recent papers use ambiguity to explain deviations from Bayesian updating. Baliga, Hanany, and Klibanoff (2013) show that belief polarization can arise when decision makers are ambiguity averse. Fryer, Harms, and Jackson (2018) explain confirmation bias by assuming that when decision makers receive ambiguous signals they interpret it as favorable to their original beliefs. In a social learning experiment, Filippis, Guarino, Jehiel, and Kitagawa (2016) find that decision makers frequently deviate from Bayesian updating when they receive private information that contradicts their original beliefs. Their explanation for this phenomenon assumes multiple priors. All of these papers are non-axiomatic and address specific deviations from Bayesian updating.

7 Conclusion

Many real life economic problems involve ambiguity. In this paper, it is argued that departing from Bayesian updating is natural when one faces ambiguity. I axiomatize a non-Bayesian updating model, robust maximum likelihood (RML) updating, where the DM's probability assessment can be represented by a benchmark prior, which reflects the DM's initial best guess, and a set of priors the DM considers plausible. The DM responds to new information by revising the benchmark prior via the maximum likelihood principle in a way that ensures maximally dynamically consistent behavior, and updates the new prior using Bayes' rule. I show that RML updating can accommodate many commonly observed deviations from Bayesian updating.

I take the DM's preferences over acts before and after the arrival of new information as the primitive of the analysis. In addition to standard axioms, the two main axioms imposed on preferences are robust inference and consistency. Robust inference requires that when a proper subset of a minimal unambiguous event is ruled out, the DM's willingness to bet on this minimal unambiguous event is not affected. This reflects the DM's cautious attitude when she updates her benchmark prior. Consistency states that every proper subset of a minimal unambiguous event is perfectly dynamically consistent. This reflects the intuition that when such an event is realized, the DM does not learn any new information that can help her make an inference regarding the relative likelihoods of the states within this event. I show that if the DM satisfies these axioms, both the benchmark prior and the set of plausible priors are uniquely identified from preferences. Lastly, I provide a characterization of RML updating with ambiguity averse preferences.

A Proof of Theorem 1

A.1 Necessity

The necessity of Axioms 1–5 is standard. The necessity of constant act preference invariance follows from the observation that the same utility function is used for all $\succsim_A$ in the representation. Consequentialism is necessary as the RML posterior satisfies $\pi_A(\omega) > 0$ only if $\omega \in A$.

Let $(\pi, \mathcal{P}, u)$ be a representation of $\{\succsim_A\}_{A\in\mathcal{A}}$ given by Theorem 1. To prove the necessity of Axioms 8–10, it needs to be shown that the set of unambiguous events, as defined in Definition 4 and denoted by $\mathcal{E}$, is equal to $\sigma(\mathcal{P})$, the algebra generated by the partition $\mathcal{P}$, as long as $\mathcal{P}$ is not degenerate (i.e. $\mathcal{P} \neq \{\Omega, \emptyset\}$). This is proved in Claim 3 after two preliminary observations. The next claim shows that if $A \subseteq B$ and π_A and π_B are the RML posteriors of π, then π_A is the RML posterior of π_B where the partition of B is given by $\{B \cap P |\, P \in \mathcal{P}\}$.

Claim 1. *Let $(\pi, \mathcal{P})$ represent the probability assessment and suppose for any non-null $A \in \mathcal{A}$, π_A is the RML posterior of π. Then for any non-null $A \subseteq B$, π_A is the RML posterior of π_B where the partition of B is given by $\{B \cap P |\, P \in \mathcal{P}\}$.*

Proof. Notice that since $A \cap (B \cap P) = A \cap P$ for any $P \in \mathcal{P}$, it suffices to show that for any $P, P' \in \mathcal{P}$ with $A \cap P \neq \emptyset$ and $A \cap P' \neq \emptyset$,

$$\frac{\pi_A(P)}{\pi_A(P')} = \frac{\pi_B(P)}{\pi_B(P')}$$

and for any $\omega, \omega' \in A \cap P$ where $P \in \mathcal{P}$,

$$\frac{\pi_A(\omega)}{\pi_A(\omega')} = \frac{\pi_B(\omega)}{\pi_B(\omega')}.$$

The second equation follows from the observation that since π_A and π_B are the RML posteriors of π and $A \subseteq B$, the above ratio is equal to $\frac{\pi(\omega)}{\pi(\omega')}$. Similarly, the first equation follows from the observation that $B \cap P \neq \emptyset$ whenever $A \cap P \neq \emptyset$, and hence the above ratio is equal to $\frac{\pi(P)}{\pi(P')}$. $\square$

Consider an act that gives the best prize on $P \setminus E$, where P is a minimal unambiguous event and $E \subsetneq P$, and the worst prize outside $P \setminus E$, i.e. $f_{P\setminus E} = x_* E(x^* P x_*)$. The next claim shows that for any realized event $A \supseteq P$, we can find a sure outcome z_A such that the DM is indifferent between $f_{P\setminus E}$ and an act that gives the worst prize on E and z_A outside E.

Claim 2. *Let $E \subsetneq P$ for some $P \in \mathcal{P}$ and $f_{P\setminus E} = x_* E(x^* P x_*)$. Then for any $A \supseteq P$, there exists $z_A \in X$ such that $f_{P\setminus E} \sim_A x_* E z_A$.*

Proof. Let z_A be defined by

$$z_A = \frac{\pi_A(P \setminus E)}{1 - \pi_A(E)} x^* + \frac{1 - \pi_A(P)}{1 - \pi_A(E)} x_*.$$

By assumption, $A \supsetneq E$ and $\pi_A(\omega) > 0$ for each $\omega \in A$. Hence, $\pi_A(E) < 1$. Since X is convex, $z_A \in X$. Moreover, since u is affine,

$$u(z_A) = \frac{\pi_A(P \setminus E)}{1 - \pi_A(E)} u(x^*) + \frac{1 - \pi_A(P)}{1 - \pi_A(E)} u(x_*).$$

Let U_A be a representation of $\succcurlyeq_A$ given by Theorem 1. By the representation,

$$\begin{aligned}
U_A(f_{P\setminus E}) &= \pi_A(P \setminus E)u(x^*) + (1 - \pi_A(P \setminus E))u(x_*) \\
&= \pi_A(P \setminus E)u(x^*) + (1 - \pi_A(P))u(x_*) + \pi_A(E)u(x_*) \\
&= (1 - \pi_A(E)) \left[\frac{\pi_A(P \setminus E)}{1 - \pi_A(E)} u(x^*) + \frac{1 - \pi_A(P)}{1 - \pi_A(E)} u(x_*) \right] + \pi_A(E)u(x_*) \\
&= (1 - \pi_A(E))u(z_A) + \pi_A(E)u(x_*).
\end{aligned}$$

Therefore, $f_{P\setminus E} \sim_A x_* E z_A$. □

The next claim is the key step in concluding the necessity of axioms by showing that unambiguous events are fully revealed from preferences unless all proper nontrivial events are ambiguous. The case when all proper nontrivial events are ambiguous is treated in the same way in the model as the case when no event is ambiguous. Hence, unambiguous events cannot be revealed if $\mathcal{P} = \{\Omega, \emptyset\}$.

Claim 3. *Suppose $\mathcal{P} \neq \{\Omega, \emptyset\}$. Then $E \in \mathcal{E}$ if and only if $E \in \sigma(\mathcal{P})$.*

Proof. Suppose $E \notin \sigma(\mathcal{P})$. It needs to be shown that $E \notin \mathcal{E}$. There are two cases to consider.

Case 1: $E \subsetneq P$ for some $P \in \mathcal{P}$. Let $x^* \succ x_*$ be given and define $f_{P \setminus E} = x_* E(x^* P x_*)$. Let z be such that $f_{P \setminus E} \sim x_* E z$ as in the previous claim, and let $\bar{z}$ be such that $f_P = x^* P x_* \sim \bar{z}$.

$$f_P = \begin{pmatrix} x^* & \text{if } \omega \in E \\ x^* & \text{if } \omega \in P \setminus E \\ x_* & \text{if } \omega \in \Omega \setminus P \end{pmatrix} \sim \bar{z} \quad \text{and} \quad f_{P \setminus E} = \begin{pmatrix} x_* & \text{if } \omega \in E \\ x^* & \text{if } \omega \in P \setminus E \\ x_* & \text{if } \omega \in \Omega \setminus P \end{pmatrix} \sim \begin{pmatrix} x_* & \text{if } \omega \in E \\ z & \text{if } \omega \in P \setminus E \\ z & \text{if } \omega \in \Omega \setminus P \end{pmatrix}$$

Since each state in Ω is non-null, $\pi(P) > \pi(P \setminus E)/(1 - \pi(E))$. Hence, $u(x^*) > u(x_*)$ implies

$$u(\bar{z}) = \pi(P)u(x^*) + (1 - \pi(P))u(x_*) > \frac{\pi(P \setminus E)}{1 - \pi(E)} u(x^*) + \frac{1 - \pi(P)}{1 - \pi(E)} u(x_*) = u(z).$$

Therefore, by the representation, $\bar{z} \succ z$ and $\bar{z} \succ_{E^c} z$. Moreover, since π_{E^c} is the RML posterior of π, $\pi_{E^c}(P) = \pi(P)$. This implies $f_P \sim_{E^c} \bar{z}$ because

$$U_{E^c}(f_P) = \pi_{E^c}(P)u(x^*) + (1 - \pi_{E^c}(P))u(x_*) = u(\bar{z}).$$

In addition, since $\pi_{E^c}(E) = 0$, $f_{P \setminus E} \sim_{E^c} f_P \sim_{E^c} \bar{z}$ and $z \sim_{E^c} x_* E z$. Therefore, $\bar{z} \succ_{E^c} z$ implies $f_{P \setminus E} \succ_{E^c} x_* E z$. On the other hand, $f_{P \setminus E} \sim x_* E z$ and $f_{P \setminus E}$ and $x_* E z$ agree on E, violating dynamic consistency. Since E^c is not dynamically consistent, neither E nor E^c belongs to $\mathcal{E}$.

Case 2: There are $P, P' \in \mathcal{P}$ such that $P \cap E \neq \emptyset$, $P' \cap E \neq \emptyset$ and either $P \cap E^c \neq \emptyset$ or $P' \cap E^c \neq \emptyset$. Without loss of generality, assume that $P \cap E^c \neq \emptyset$. Let $A = E \cup P$. As before, let $x^* \succ x_*$ be given. Consider a bet on $P \cap E$ given by $f_{P \cap E} = x_* P \setminus E(x^* P x_*)$. Let z_A be as in the previous claim such that $x_* P \setminus E(x^* P x_*) \sim_A x_* P \setminus E z_A$, and let $\bar{z}_A \in X$ be such that $f_P = x^* P x_* \sim_A \bar{z}_A$. Note that $\bar{z}_A \succ z_A$ as in the previous case.

$$f_P = \begin{pmatrix} x^* & \text{if } \omega \in P \setminus E \\ x^* & \text{if } \omega \in P \cap E \\ x_* & \text{if } \omega \in \Omega \setminus P \end{pmatrix} \sim_A \bar{z}_A \quad \text{and} \quad f_{P \cap E} = \begin{pmatrix} x_* & \text{if } \omega \in P \setminus E \\ x^* & \text{if } \omega \in P \cap E \\ x_* & \text{if } \omega \in \Omega \setminus P \end{pmatrix} \sim_A \begin{pmatrix} x_* & \text{if } \omega \in P \setminus E \\ z_A & \text{if } \omega \in P \cap E \\ z_A & \text{if } \omega \in \Omega \setminus P \end{pmatrix}$$

By Claim 1, π_E is the RML posterior of π_A with the partition $\{A \cap P | \, P \in \mathcal{P}\}$. Now note that if $P'' \in \mathcal{P}$ has a nonempty intersection with A, it must also have a nonempty intersection with E: if $P'' \cap A \neq \emptyset$, then either $P'' = P$ in which case $P'' \cap E \neq \emptyset$ follows by assumption or $P'' \subseteq P^c$

in which case $P'' \cap E \neq \emptyset$ follows from $A = E \cup P$. Hence, it must be that $\pi_E(P) = \pi_A(P)$. Therefore, the representation implies $f_{P \cap E} = x_* P \setminus E(x^* P x_*) \sim_E \bar{z}_A \succ_E z_A$. On the other hand, $x_* P \setminus E z_A \sim_E z_A$. Hence, $x_* P \setminus E(x^* P x_*) \sim_A x_* P \setminus E z_A$ but $x_* P \setminus E(x^* P x_*) \succ_E x_* P \setminus E z_A$, violating perfect dynamic consistency. This proves that $E \notin \mathcal{E}$.

Now suppose $E \in \sigma(\mathcal{P})$. It needs to be shown that $E \in \mathcal{E}$. Let $A \supseteq E$ be given. By Claim 1, π_E is the RML posterior of π_A with the partition $\{A \cap P | \, P \in \mathcal{P}\}$. To prove that $E \in \mathcal{E}$, it suffices to show that for any $\omega \in E$,

$$\pi_E(\omega) = \frac{\pi_A(\omega)}{\pi_A(E)}.$$

Since $E \in \sigma(\mathcal{P})$, $\cup_{P \in \mathcal{P}:\, E \cap P \neq \emptyset} P = E$. By Proposition 1 and Claim 1,

$$\pi_E(\omega) = \pi_A(\omega | P_\omega) \cdot \pi_A(P_\omega | \cup_{P \in \mathcal{P}:\, E \cap P \neq \emptyset} P) = \pi_A(\omega | P_\omega) \cdot \pi_A(P_\omega | E) = \frac{\pi_A(\omega)}{\pi_A(E)}$$

as desired. This concludes the proof of the claim. □

Since $E \in \mathcal{E}$ if and only if $E \in \sigma(\mathcal{P})$, the necessity of Axiom 8 is obvious. To see the necessity of Axiom 9, let $A \in \mathcal{A}$ and $D \subsetneq A \cap P$ for some $P \in \mathcal{P}$. Since $\pi_{A \setminus D}$ is the RML posterior of π_A with the partition $\{A \cap P | P \in \mathcal{P}\}$, we have $\pi_{A \setminus D}(P) = \pi_A(P)$. Hence, Axiom 9 follows. Lastly, to see the necessity of Axiom 10, let $D \subsetneq P$ for some $P \in \mathcal{P}$, $A \supseteq D$ and $\omega \in D$ be given. Since π_D is the RML posterior of π_A,

$$\pi_D(\omega) = \pi_A(\omega | D) \cdot \pi_A(P | P) = \pi_A(\omega | D).$$

Hence, D satisfies perfect dynamic consistency.

A.2 Sufficiency

The first claim is a standard result. For a proof, see Fishburn (1970) or Kreps (1988).

Claim 4. *Suppose Axioms 1–5 are satisfied. Then, for any $A \in \mathcal{A}$, there exist a subjective probability measure $\pi_A \in \Delta(\Omega)$ and a non-constant, affine utility function $u_A : X \to \mathbb{R}$ such that for any $f, g \in \mathcal{F}$,*

$$f \succsim_A g \Leftrightarrow \sum_{\omega \in \Omega} \pi_A(\omega) u_A(f(\omega)) \geq \sum_{\omega \in \Omega} \pi_A(\omega) u_A(g(\omega)).$$

By constant act preference invariance, $\succcurlyeq_A$ and $\succcurlyeq$ agree on all constant acts for all $A \in \mathcal{A} \setminus \emptyset$. Using the standard uniqueness result, for any $A \in \mathcal{A} \setminus \emptyset$, u_A is a positive affine transformation of u_Ω, which is denoted by u. Hence, it is without loss to let $u_A = u$ for all $A \in \mathcal{A} \setminus \emptyset$. Moreover, Axioms 4 and 6 imply that $\succcurlyeq_A$ is nontrivial, and hence π_A is unique for each $A \in \mathcal{A} \setminus \emptyset$ as in Anscombe and Aumann (1963). By consequentialism, for any $f \in \mathcal{F}$, $fAx^* \sim_A fAx_*$. Since by the representation $u(x^*) > u(x_*)$, we must have $\pi_A(A^c) = 0$. Moreover, by monotonicity, for any $\omega \in A$, $x^* \succ_A x_*\omega x^*$. Hence, the representation implies that $\pi_A(\omega) > 0$ for all $\omega \in A$. This establishes the following claim.

Claim 5. *Suppose Axioms 1–7 are satisfied. Then, there exist a non-constant, affine utility function $u : X \to \mathbb{R}$ with $u(X) = [u(x_*), u(x^*)]$ and a family of probability measures $\{\pi_A\}_{A\in\mathcal{A}}$ such that for any $f, g \in \mathcal{F}$ and $A \in \mathcal{A}$,*

$$f \succcurlyeq_A g \Leftrightarrow \sum_{\omega\in\Omega} \pi_A(\omega)u(f(\omega)) \geq \sum_{\omega\in\Omega} \pi_A(\omega)u(g(\omega)).$$

Moreover, π_A has full support on A and is unique for all $A \in \mathcal{A}\setminus\emptyset$, and u is unique up to a positive affine transformation.

Let $\mathcal{E}$ be the collection of events which are perfectly dynamically consistent and whose complements are also perfectly dynamically consistent as in Definition 4. By definition, $\mathcal{E}$ is closed under complements. By Axiom 8, $\mathcal{E}$ is closed under intersections. Moreover, $\Omega \in \mathcal{E}$. Hence, $\mathcal{E}$ is an algebra over Ω. Let $\mathcal{P}_\mathcal{E}$ be the partitioning of the state space that generates $\mathcal{E}$, and let $(\pi, \mathcal{P}_\mathcal{E})$ denote the probability assessment. To establish the representation, it suffices to show that each π_A is the RML posterior of π.

Claim 6. *For any non-null $A \in \mathcal{A}$, π_A is the RML posterior of π.*

Proof. First, consider $A \in \mathcal{E}$. By definition, $f \succcurlyeq_A g \Leftrightarrow fAg \succcurlyeq g$. $fAg \succcurlyeq g$ is equivalent to

$$\sum_{\omega\in A} \pi(\omega)u(f(\omega)) \geq \sum_{\omega\in A} \pi(\omega)u(g(\omega)) \quad \Leftrightarrow \quad \sum_{\omega\in A} \frac{\pi(\omega)}{\pi(A)}u(f(\omega)) \geq \sum_{\omega\in A} \frac{\pi(\omega)}{\pi(A)}u(g(\omega)).$$

By the uniqueness of π_A in the representation, for all $\omega \in A$,

$$\pi_A(\omega) = \frac{\pi(\omega)}{\pi(A)} = \pi(\omega|A).$$

That is, π_A is the Bayesian posterior of π, which also corresponds to the RML posterior since A is unambiguous.

Now consider $A \notin \mathcal{E}$. Let $\mathcal{P}_{\mathcal{E}} = \{P_1, \ldots, P_n\}$ and choose an index set $J \subseteq \{1, \ldots, n\}$ such that $P_j \cap A \neq \emptyset \Leftrightarrow j \in J$. Let $B = \cup_{j \in J} P_j$. Since $B \in \mathcal{E}$, by the first part of the claim, π_B is the Bayesian posterior of π. Moreover, by Axiom 9, for any $j \in J$, $c_A(f_{P_j}) \sim c_{A \cup P_j}(f_{P_j})$ where $f_{P_j} = x^* P_j x_*$. Given the representation in Claim 5, this is possible only if $\pi_A(P_j) = \pi_{A \cup P_j}(P_j)$. Hence, iterative application Axiom 9 implies that $\pi_A(P_j) = \pi_B(P_j)$ for all $j \in J$. Therefore, for any $j, j' \in J$,

$$\frac{\pi_A(P_j)}{\pi_A(P_{j'})} = \frac{\pi_B(P_j)}{\pi_B(P_{j'})} = \frac{\pi(P_j)}{\pi(P_{j'})}.$$

Let $A \cap P_j$ for some $j \in J$ be given. By Axiom 10, $A \cap P_j$ is a perfectly dynamically consistent event. Hence, using the same reasoning as in the first part of the claim, we get

$$\pi(\omega|A \cap P_j) = \pi_{A \cap P_j}(\omega) = \pi_A(\omega|P_j).$$

But then for any $\omega, \omega' \in A \cap P_j$,

$$\frac{\pi(\omega)}{\pi(\omega')} = \frac{\pi_{A \cap P_j}(\omega)}{\pi_{A \cap P_j}(\omega')} = \frac{\pi_A(\omega)}{\pi_A(\omega')}.$$

This together with the conclusion of the previous paragraph and Proposition 1 imply that π_A is the RML posterior of π. This concludes the proof of sufficiency. $\square$

Lastly, to show that $\mathcal{P}$ is uniquely revealed as $\mathcal{P}_{\mathcal{E}}$, assume that $\mathcal{P}$ is not degenerate ($\mathcal{P} \neq \{\Omega, \emptyset\}$). By Claim 3, $\mathcal{P}$ and $\mathcal{P}_{\mathcal{E}}$ are two partitions of the state space that generate the same algebra $\mathcal{E}$. But then $\mathcal{P} = \mathcal{P}_{\mathcal{E}}$. $\square$

B Proof of Proposition 3

Let $(\pi, \mathcal{P}, u)$ be a representation of $\{\succcurlyeq_A\}_{A\in\mathcal{A}}$ given by Theorem 2. Observe that for any A,

$$\pi_A \in B(\mathbb{N}_{\mathcal{P},A}(\pi)) \Leftrightarrow \pi_A(A^c) = 0 \text{ and } \pi_A(P) = \frac{\pi(P)}{\sum_{P'\in\mathcal{P}:P'\cap A\neq\emptyset}\pi(P')} \text{ for any } P \in \mathcal{P} \text{ with } A\cap P \neq \emptyset.$$

That is, all plausible posteriors agree on minimal unambiguous events. Moreover, the utility function defined by

$$U_A(f) = \sum_{P\in\mathcal{P}} \pi_A(P) \min_{\omega\in A\cap P} u(f(\omega)),$$

where $\pi_A \in B(\mathbb{N}_{\mathcal{P},A}(\pi))$, represents $\succcurlyeq_A$.

Let E be a dynamically consistent event. Suppose E is not perfectly dynamically consistent so that there exist $A \supseteq E$ and $f, g \in \mathcal{F}$ such that $fEg \succcurlyeq_A g$ and $g \succ_E f$. Let h and h' be given as below.

$$h = \begin{pmatrix} f(\omega) & \text{if } \omega \in E \\ g(\omega) & \text{if } \omega \in A \setminus E \\ x^* & \text{if } \omega \in A^c \end{pmatrix} \quad \text{and} \quad h' = \begin{pmatrix} g(\omega) & \text{if } \omega \in E \\ g(\omega) & \text{if } \omega \in A \setminus E \\ x^* & \text{if } \omega \in A^c \end{pmatrix}$$

Then, the representation implies $h \succcurlyeq_A h'$ and $h' \succ_E h$. Also note that $h(\omega) = h'(\omega)$ for all $\omega \in E^c$. Next, it is shown that $h \succcurlyeq_A h'$ implies $h \succcurlyeq h'$. Let $\mathcal{P} = \{P_1, \dots, P_n\}$ and $J \subseteq \{1, \dots, n\}$ be the index set such that $j \in J \Leftrightarrow P_j \cap E \neq \emptyset$. Then, for any $i \notin J$, $P_i \cap E = \emptyset$, and hence $\min_{\omega\in A\cap P_i} u(h) = \min_{\omega\in A\cap P_i} u(h')$. Therefore, by the representation, $h \succcurlyeq_A h'$ implies

$$\sum_{j\in J} \pi_A(P_j) \min_{\omega\in A\cap P_j} u(fEg(\omega)) \geq \sum_{j\in J} \pi_A(P_j) \min_{\omega\in A\cap P_j} u(g(\omega)).$$

Notice that since $h(\omega) = h'(\omega) = x^*$ for all $\omega \in A^c$, for any $j \in J$, $\min_{\omega\in A\cap P_j} u(h(\omega)) = \min_{\omega\in P_j} u(h(\omega))$ and $\min_{\omega\in A\cap P_j} u(h'(\omega)) = \min_{\omega\in P_j} u(h'(\omega))$. In addition, for $i \notin J$, $\min_{\omega\in P_i} u(h) = \min_{\omega\in P_i} u(h')$ since h and h' agree on E^c. Lastly, for any $j \in J$, $\pi(P_j) = c \cdot \pi_A(P_j)$ where

$c = \sum_{P' \in \mathcal{P}: P' \cap A \neq \emptyset} \pi(P')$. Hence,

$$\begin{aligned} \sum_{j \in J} \pi_A(P_j) \min_{\omega \in A \cap P_j} u(fEg(\omega)) &\geq \sum_{j \in J} \pi_A(P_j) \min_{\omega \in A \cap P_j} u(g(\omega)) \\ \Rightarrow \quad \sum_{j \in J} c \cdot \pi_A(P_j) \min_{\omega \in A \cap P_j} u(h(\omega)) &\geq \sum_{j \in J} c \cdot \pi_A(P_j) \min_{\omega \in A \cap P_j} u(h'(\omega)) \\ \Rightarrow \quad \sum_{j \in J} \pi(P_j) \min_{\omega \in P_j} u(h(\omega)) &\geq \sum_{j \in J} \pi(P_j) \min_{\omega \in P_j} u(h'(\omega)) \\ \Rightarrow \quad \sum_{P_i \in \mathcal{P}} \pi(P_i) \min_{\omega \in P_i} u(h(\omega)) &\geq \sum_{P_i \in \mathcal{P}} \pi(P_i) \min_{\omega \in P_i} u(h'(\omega)), \end{aligned}$$

which implies $h \succcurlyeq h'$. But then since $h(\omega) = h'(\omega)$ for all $\omega \in E^c$ and $h' \succ_E h$, this contradicts the original hypothesis that E is dynamically consistent. Hence, E must be perfectly dynamically consistent. □

C Proof of Theorem 2

C.1 Necessity

The necessity of Axioms 1, 2, 4, 5, 6, and 7 is standard. To prove the necessity of Axioms 8, 9, 11, and 12, it is shown that $\mathcal{E} = \sigma(\mathcal{P})$, where $\sigma(\mathcal{P})$ is the algebra generated by $\mathcal{P}$, as in the proof of Theorem 1.

Claim 7. *Let $(\pi, \mathcal{P}, u)$ be a representation of $\{\succcurlyeq_A\}_{A \in \mathcal{A}}$ given by equation 5. Then $E \in \mathcal{E}$ if and only if $E \in \sigma(\mathcal{P})$.*

Proof. First, I show that if $E \notin \sigma(\mathcal{P})$, then $E \notin \mathcal{E}$. Notice that since $E \notin \sigma(\mathcal{P})$, there exists $P \in \mathcal{P}$ such that $P \cap E \neq \emptyset$ and $P \setminus E \neq \emptyset$. Let $f = x^* P \cap E x_*$. Then the representation implies that $f \sim x_*$ but $f \succ_E x_*$ even though $f(\omega) = x_*$ for $\omega \in E^c$. $f \succ_E x_*$ holds because of the assumption that $\pi(P) > 0$, which implies $\pi_E(P) > 0$ for all $\pi_E \in B(\mathbb{N}_{\mathcal{P},E}(\pi))$. Hence, E is not a dynamically consistent event. By definition, $E \notin \mathcal{E}$.

Now suppose $E \in \sigma(\mathcal{P})$. To show that $E \in \mathcal{E}$, let $A \supseteq E$ be given. It needs to be shown that $fEg \succcurlyeq_A g$ if and only if $f \succcurlyeq_E g$. Let $\mathcal{P} = \{P_1, \ldots, P_n\}$. Since $E \in \sigma(\mathcal{P})$, there exists an index set $J \subseteq \{1, \ldots, n\}$ such that $E = \cup_{j \in J} P_j$. Notice that if $i \notin J$, then $\min_{\omega \in A \cap P_i} u(fEg(\omega)) =$

$\min_{\omega \in A \cap P_i} u(g(\omega))$. Hence, $fEg \succsim_A g$ if and only if

$$\sum_{j \in J} \pi_A(P_j) \min_{\omega \in P_j} u(fEg(\omega)) \geq \sum_{j \in J} \pi_A(P_j) \min_{\omega \in P_j} u(g(\omega))$$

where π_A is an arbitrary member of $B(\mathbb{N}_{\mathcal{P},A}(\pi))$. On the other hand, it is easy to see that for any $\pi_E \in B(\mathbb{N}_{\mathcal{P},E}(\pi))$, $\pi_E(P_j) > 0$ if and only if $j \in J$, and $\pi_E(P_j) = c \cdot \pi_A(P_j)$ where $c = \frac{1}{\sum_{j \in J} \pi_A(P_j)}$. Hence, the above inequality holds if and only if

$$\sum_{j \in J} \pi_E(P_j) \min_{\omega \in P_j} u(fEg(\omega)) \geq \sum_{j \in J} \pi_E(P_j) \min_{\omega \in P_j} u(g(\omega))$$

which is true if and only if $f \succsim_E g$. $\square$

Since $\mathcal{E} = \sigma(\mathcal{P})$, the necessity of Axiom 8 is obvious. The necessity of Axiom 9 is the same as in Theorem 1. To see the necessity of Axiom 11, let $\Gamma : \Omega \to S$ be a surjective mapping that satisfies $\Gamma(\omega) = \Gamma(\omega')$ for all $\omega, \omega' \in P$ and $\Gamma(\omega) \neq \Gamma(\omega')$ whenever $\omega \in P$ and $\omega' \in P'$ for distinct P and P'. Now for each $A \in \mathcal{A}$, define a probability measure on S by $\pi_A \circ \Gamma^{-1}$ where π_A is an arbitrary member of $B(\mathbb{N}_{\mathcal{P},A}(\pi))$. Let $\hat{\mathcal{F}}$ be the set of all acts X^S. Then, $\hat{\mathcal{F}}$ is isomorphic $\mathcal{F}^{ua}$. Since $(\pi_A \circ \Gamma^{-1}, u)$ is an SEU representation of $\succsim_A$ restricted to $\hat{\mathcal{F}}$, Axiom 11 follows. Lastly, Axiom 12 is necessary because within each partition element only the minimal payoff matters.

C.2 Sufficiency

Axiom 8 implies that $\mathcal{E}$ is an algebra and $\mathcal{P}_{\mathcal{E}}$ is a partitioning of the state space. As in Claim 5, it is easy to see that the axioms imply the following claim.

Claim 8. *Suppose Axioms 1, 2, 4-8, and 11 are satisfied. Then there exist a non-constant, affine utility function $u : X \to \mathbb{R}$ with $u(X) = [u(x_*), u(x^*)]$ and a family of subjective probability measures $\{\pi_A\}_{A \in \mathcal{A}}$ on $\sigma(\mathcal{P}_{\mathcal{E}})$ such that for any $f, g \in \mathcal{F}^{ua}$ and $A \in \mathcal{A}$,*

$$f \succsim_A g \Leftrightarrow \sum_{P \in \mathcal{P}} \pi_A(P) u(f(P)) \geq \sum_{P \in \mathcal{P}} \pi_A(P) u(g(P)).$$

Moreover, $\pi_A(P) > 0$ for any $P \in \mathcal{P}_{\mathcal{E}}$ with $A \cap P \neq \emptyset$ and $\pi_A(P) = 0$ whenever $A \cap P = \emptyset$.

The second part of the claim is implied by Axioms 5 and 7. Extend π_A to $\sigma(\mathcal{P}_{\mathcal{E}} \cup A)$ (i.e. the algebra generated by sets of the form $A \setminus P$, $A \cap P$, and $P \setminus A$) by letting $\pi_A(A \cap P) = \pi_A(P)$ whenever $A \cap P \neq \emptyset$.

The next claim shows that for any $f \in \mathcal{F}$ and $A \in \mathcal{A}$, there exists an unambiguous act $f^{ua} \in \mathcal{F}^{ua}$ such that the DM is indifferent between f and f^{ua} given A.

Claim 9. *Suppose Axioms 1, 2, 4-8, 11, and 12 are satisfied. Then, for any $A \in \mathcal{A}$ and $f \in \mathcal{F}$, there exists $f^{ua} \in \mathcal{F}^{ua}$ such that $f \sim_A f^{ua}$.*

Proof. Let $P \in \mathcal{P}_{\mathcal{E}}$ and $\omega^* = \arg\min_{\omega \in A \cap P} u(f(\omega))$. By Axioms 7 and 12, $x^* P \setminus \{\omega^*\} f \sim_A f(\omega^*) P f$. On the other hand, by Axiom 5, $x^* P \setminus \{\omega^*\} f \succcurlyeq_A f \succcurlyeq_A f(\omega^*) P f$. Hence, $f \sim_A f(\omega^*) P f$. Now let f^{ua} denote an act that assigns the worst prize of f in $A \cap P$ to P for all $P \in \mathcal{P}_{\mathcal{E}}$ with $A \cap P \neq \emptyset$. Let f^{ua} be constant on P' with $A \cap P' = \emptyset$. This act belongs to $\mathcal{F}^{ua}$, and $f \sim_A f^{ua}$ by iterative application of the previous argument and Axiom 7. $\square$

For any $f \in \mathcal{F}$, let

$$U_A(f) = \sum_{P \in \mathcal{P}} \pi_A(A \cap P) \min_{\omega \in A \cap P} u(f(\omega)).$$

Notice that for f^{ua} defined as in Claim 9, $U_A(f) = U_A(f^{ua})$. We already know that U_A represents $\succcurlyeq_A$ on $\mathcal{F}^{ua}$. Hence, $f \succcurlyeq_A g$ if and only if $f^{ua} \succcurlyeq_A g^{ua}$ if and only if $U_A(f^{ua}) \geq U_A(g^{ua})$ if and only if $U_A(f) \geq U_A(g)$. Hence, U_A represents $\succcurlyeq_A$ on all $\mathcal{F}$.

The only thing left to prove is that $\pi_A \in B(\mathbb{N}_{\mathcal{P},A}(\pi))$. This is implied by Axiom 9. The proof is identical to the first part of Claim 6. Lastly, the uniqueness result for u is standard. The uniqueness of $\mathcal{P}$ is a consequence of Claim 7, and $\mathbb{N}_{\mathcal{P},A}(\pi)$ is uniquely defined given π on $\mathcal{P}$. $\square$

D Proof of Proposition 2

Notice that $\pi'_A \in B(\mathbb{N}_{\mathcal{P},A}(\pi))$ if and only if for all $P \in \mathcal{P}$ with $A \cap P \neq \emptyset$,

$$\sum_{\omega \in A \cap P} \pi'_A(\omega) = \frac{\pi(P)}{\sum_{P' \in \mathcal{P}:\, A \cap P' \neq \emptyset} \pi(P')}. \tag{D.1}$$

The objective is to minimize Kullback-Leibler divergence $D_{\mathrm{KL}}(\pi(\cdot|A) \,||\, \pi'_A)$ subject to these constraints for each $P \in \mathcal{P}$ with $A \cap P \neq \emptyset$. The Lagrangian for the minimization problem is

$$\mathcal{L}\big(\{\pi'_A(\omega)\}_{\omega \in A}, \{\lambda_P\}_{P \in \mathcal{P}: A \cap P \neq \emptyset}\big) = -\sum_{\omega \in A} \pi(\omega|A) \ln\Big(\frac{\pi'_A(\omega)}{\pi(\omega|A)}\Big) + \sum_{P \in \mathcal{P}: A \cap P \neq \emptyset} \lambda_P \Big(\sum_{\omega \in A \cap P} \pi'_A(\omega) - \frac{\pi(P)}{\sum_{P' \in \mathcal{P}:\, A \cap P' \neq \emptyset} \pi(P')}\Big).$$

The first order conditions imply that for any $P \in \mathcal{P}$ with $A \cap P \neq \emptyset$ and any $\omega, \omega' \in A \cap P$,

$$\frac{\pi(\omega|A)}{\pi'_A(\omega)} = \lambda_P = \frac{\pi(\omega'|A)}{\pi'_A(\omega')}, \quad \text{and hence} \quad \frac{\pi'_A(\omega)}{\pi'_A(\omega')} = \frac{\pi(\omega|A)}{\pi(\omega'|A)} = \frac{\pi(\omega)}{\pi(\omega')}. \tag{D.2}$$

Since the objective function is strictly convex, equations D.1 and D.2 characterize the solution to the minimization problem. □

References

Al-Najjar, N. I. and J. Weinstein (2009): "The ambiguity aversion literature: a critical assessment," *Economics & Philosophy*, 25(3), 249–284.

Anscombe, F. J. and R. J. Aumann (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34(1), 199–205.

Baliga, S., E. Hanany, and P. Klibanoff (2013): "Polarization and Ambiguity," *American Economic Review*, 103(7), 3071–83.

Bar-Hillel, M. (1980): "The Base-Rate Fallacy in Probability Judgments," *Acta Psychologica*, 44(3), 211–233.

Barberis, N., A. Shleifer, and R. Vishny (1998): "A Model of Investor Sentiment," *Journal of Financial Economics*, 49(3), 307–343.

Benjamin, D. J. (2019): "Errors in Probabilistic Reasoning and Judgment Biases," in *Handbook of Behavioral Economics-Foundations and Applications 2*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, Elsevier, 69–186.

Binmore, K., L. Stewart, and A. Voorhoeve (2012): "How Much Ambiguity Aversion?" *Journal of Risk and Uncertainty*, 45(3), 215–238.

Camerer, C. (1995): "Individual Decision Making," in *Handbook of Experimental Economics*, ed. by J. H. Kagel and A. E. Roth, Princeton University Press, 587–703.

Charness, G., E. Karni, and D. Levin (2013): "Ambiguity Attitudes and Social Interactions: An Experimental Investigation," *Journal of Risk and Uncertainty*, 46(1), 1–25.

Cheng, X. (2021): "Relative Maximum Likelihood Updating of Ambiguous Beliefs," *arXiv preprint arXiv:1911.02678*.

Darley, J. M. and P. H. Gross (1983): "A Hypothesis-Confirming Bias in Labeling Effects," *Journal of Personality and Social Psychology*, 44(1), 20–33.

Dominiak, A., P. Duersch, and J.-P. Lefort (2012): "A Dynamic Ellsberg Urn Experiment," *Games and Economic Behavior*, 75(2), 625–638.

Dominiak, A., M. Kovach, and G. Tserenjigmid (2021): "Minimum Distance Belief Updating with General Information," *Working Paper*.

Edwards, W. (1968): "Conservatism in Human Information Processing," in *Formal Representation of Human Judgment*, ed. by B. Kleinmuntz, Wiley, New York, 17–52.

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

Epstein, L. G. (1999): "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66(3), 579–608.

——— (2006): "An Axiomatic Model of Non-Bayesian Updating," *Review of Economic Studies*, 73(2), 413–436.

Epstein, L. G. and M. Le Breton (1993): "Dynamically Consistent Beliefs Must Be Bayesian," *Journal of Economic Theory*, 61(1), 1–22.

Epstein, L. G., J. Noor, and A. Sandroni (2008): "Non-Bayesian Updating: A Theoretical Framework," *Theoretical Economics*, 3(2), 193–229.

——— (2010): "Non-bayesian learning," *The BE Journal of Theoretical Economics*, 10(1), 1–20.

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

——— (2007): "Learning Under Ambiguity," *Review of Economic Studies*, 74(4), 1275–1303.

Epstein, L. G. and J. Zhang (2001): "Subjective Probabilities on Subjectively Unambiguous Events," *Econometrica*, 69(2), 265–306.

Filippis, R. D., A. Guarino, P. Jehiel, and T. Kitagawa (2016): "Updating Ambiguous Beliefs in a Social Learning Experiment," *Working Paper, Centre for Microdata Methods and Practice.*

Fishburn, P. C. (1970): *Utility Theory for Decision Making*, Wiley, New York.

Fryer, R. G., P. Harms, and M. O. Jackson (2018): "Updating Beliefs When Evidence is Open to Interpretation: Implications for Bias and Polarization," *Working Paper.*

Gennaioli, N. and A. Shleifer (2010): "What Comes to Mind," *Quarterly Journal of Economics*, 125(4), 1399–1433.

Ghirardato, P. (2002): "Revisiting Savage in a Conditional World," *Economic Theory*, 20(1), 83–92.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004): "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory*, 118(2), 133–173.

Ghirardato, P. and M. Marinacci (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102(2), 251–289.

Gilboa, I. and M. Marinacci (2016): "Ambiguity and the Bayesian Paradigm," in *Readings in Formal Epistemology: Sourcebook*, ed. by H. Arló-Costa, V. F. Hendricks, and J. van Benthem, Springer International Publishing, Cham, 385–439.

Gilboa, I. and D. Schmeidler (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18(2), 141–153.

——— (1993): "Updating Ambiguous Beliefs," *Journal of Economic Theory*, 59(1), 33–49.

Gul, F. and W. Pesendorfer (2014): "Expected Uncertain Utility Theory," *Econometrica*, 82(1), 1–39.

——— (2018): "Evaluating Ambiguous Random Variables and Updating by Proxy," *Working Paper, Princeton University.*

Hanany, E. and P. Klibanoff (2007): "Updating Preferences with Multiple Priors," *Theoretical Economics*, 2(3), 261–298.

Hansen, L. and T. J. Sargent (2001): "Robust Control and Model Uncertainty," *American Economic Review*, 91(2), 60–66.

Hill, B. (2021): "Updating Confidence in Beliefs," *Journal of Economic Theory*, 105209.

Kahneman, D. and A. Tversky (1972): "Subjective Probability: A Judgment of Representativeness," *Cognitive Psychology*, 9, 430–454.

——— (1973): "On the Psychology of Prediction," *Psychological Review*, 80(4), 237.

Ke, S., B. Wu, and C. Zhao (2021): "Learning from a Black Box," *Available at SSRN 3848205.*

Kovach, M. (2021a): "Ambiguity and Partial Bayesian Updating," *arXiv preprint arXiv:2102.11429.*

——— (2021b): "Conservative Updating," *arXiv preprint arXiv:2102.00152.*

Kreps, D. M. (1988): *Notes on the Theory of Choice*, Westview Press: Underground Classics in Economics.

Levy, G. and R. Razin (2021): "A Maximum Likelihood Approach to Combining Forecasts," *Theoretical Economics*, 16(1), 49–71.

Lord, C. G., L. Ross, and M. R. Lepper (1979): "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of Personality and Social Psychology*, 37(11), 2098–2109.

Machina, M. J. (1989): "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *Journal of Economic Literature*, 27(4), 1622–1668.

Machina, M. J. and M. Siniscalchi (2014): "Ambiguity and Ambiguity Aversion," in *Handbook of the Economics of Risk and Uncertainty*, ed. by M. Machina and K. Viscusi, North-Holland, 729 – 807.

Minardi, S. and A. Savochkin (2017): "Subjective Contingencies and Limited Bayesian Updating," *HEC Paris Research Paper No. ECO-SCD-2017-1203*.

Mullainathan, S. (2002a): "A Memory-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 117(3), 735–774.

——— (2002b): "Thinking Through Categories," *Unpublished Paper*.

Mullainathan, S., J. Schwartzstein, and A. Shleifer (2008): "Coarse Thinking and Persuasion," *Quarterly Journal of Economics*, 123(2), 577–619.

Odean, T. (1998): "Volume, Volatility, Price, and Profit When All Traders Are Above Average," *Journal of Finance*, 53(6), 1887–1934.

Ortoleva, P. (2012): "Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News," *American Economic Review*, 102(6), 2410–36.

Pires, C. P. (2002): "A Rule for Updating Ambiguous Beliefs," *Theory and Decision*, 53(2), 137–152.

Rabin, M. (1998): "Psychology and Economics," *Journal of Economic Literature*, 36(1), 11–46.

Rabin, M. and J. L. Schrag (1999): "First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics*, 114(1), 37–82.

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

Siniscalchi, M. (2011): "Dynamic Choice Under Ambiguity," *Theoretical Economics*, 6(3), 379–421.

Stahl, D. O. (2014): "Heterogeneity of Ambiguity Preferences," *Review of Economics and Statistics*, 96(4), 609–617.

Strzalecki, T. (2011): "Axiomatic Foundations of Multiplier Preferences," *Econometrica*, 79(1), 47–73.

Tversky, A. (2004): *Preference, Belief, and Similarity: Selected Writings*, MIT Press, 193–402.

Wilson, A. (2014): "Bounded Memory and Biases in Information Processing," *Econometrica*, 82(6), 2257–2294.

Zhang, J. (2002): "Subjective Ambiguity, Expected Utility and Choquet Expected Utility," *Economic Theory*, 20(1), 159–181.

Zhao, C. (2017): "Pseudo-Bayesian Updating," *Working Paper, The University of Hong Kong.*

——— (2018): "Representativeness and Similarity," *Working Paper, The University of Hong Kong.*