Ambiguity, robust statistical decisions, and Raiffa's critique

Filippo Massari

Department of Decision Sciences, Bocconi University

March 6, 2020

Abstract

Ambiguity-averse decision functionals matched with the multiple-prior learning model are more robust to model misspecification than the standard expected utility with Bayesian learning. However, these criteria fail to deliver robust decisions generically because the multiple-prior learning model inherits the same fragility of Bayesian learning. There are misspecified learning problems in which an ambiguity-averse DM optimally chooses a sequence of ambiguous acts over a sequence of risky acts that would deliver a strictly higher average utility.

Keywords: Ambiguity, misspecified learning, robust statistics, multiple-prior learning.
JEL Classification: D81, D83, C11.

1 Introduction

The maximin criterion is the cornerstone of robust statistical decisions (Wald, 1950) and decisions under ambiguity (MEU Gilboa and Schmeidler, 1989).[1] It provides guidance on how to make robust decisions in a static setting but is silent about how to approach a dynamic problem in which a DM might learn something about the true distribution of outcomes. A natural question is whether existing ambiguity-averse decision functionals, when matched with known learning models, deliver decisions that are robust to model misspecification (Watson et al., 2016; Hansen et al., 2016). An affirmative answer would provide a foundation for the use of ambiguity-averse functionals in settings in which a DM is uncertain about the true data generating process and wants to avoid decisions that would cause him regret and to change his mind were the potential consequences of his behavior explained — a desiderata of *rationality*, according to Gilboa (2009).

[1]See Cerreia-Vioglio et al. (2013) for a formal comparison of the two approaches.

Empirical evidence reinforces the intuition that ambiguity aversion has a robustifying effect on portfolio rules (Garlappi et al., 2007; Branger et al., 2013), however, these findings are mute on the reliability of this effect because they lack a theoretical model to explain them.

In this paper, I focus on a dynamic version of the standard Ellsberg (1961) two urns experiment and show that the MEU/multiple-prior learning decision criterion is more robust to model misspecification than the standards subjective expected utility (SEU, Savage, 1954) with Bayesian (single-prior) learning but yet not robust, generically.

I study the performance of a decision maker (DM) that learns according to the multiple-prior learning models (in possibly misspecified learning environments, Marinacci and Massari, 2019; Massari and Newton, 2019) and makes myopic decisions according to the MEU.[2] The DM faces two challenges. First, although naturally interested in learning, being myopic the DM might be unwilling to sample from the ambiguous process because of his aversion to ambiguity. Second, the learning problem is hard because the ambiguous urn is not kept the same. So, the DM fears being in a misspecified learning environment because of the difficulty of constructing a prior support that is guaranteed to contain the true model. I explore these two issues one by one.

First: *the cost problem can be rendered nil.*

Raiffa (1961)'s critique of Ellsberg's paradox has taught us that, at least in a static setting, a (sophisticated) DM should not fear ambiguity because he can commit his choice to the outcome of a randomization device (a coin).

> *"Suppose you withdraw a ball from the urn with unknown composition but do not look at its color. Now toss a fair (unbiased) coin and call "white" if heads, "black" if tails. The "objective" probability of getting a match is now .5, and therefore it is just as desirable to participate in the second game (ambiguous urn) as in the first (risky urn)".* Raiffa (1961)

The Raiffa argument shows that objective randomization can be used to reshape the choice between two ambiguous acts into a lottery with risky payoffs. Because the randomized act is payoff equivalent to betting on the risky urn and allows for learning, it is reasonable to assume that the DM prefers the a randomized act over bets on the risky urns. I postulate this preference for information as an axiom.

A1 *Preferences for information:* The DM strictly prefers randomizing over ambiguous acts over a risky bet with the same objective expectation.

I consider a DM who makes myopic decisions according to the MEU enriched with axiom **A1** and that learns about the composition of the ambiguous urn according to the multiple-prior

[2]My findings generalize with minimal changes when the MEU is replaced with other ambiguity-averse decision functionals in Definition 2 below — including SEU (Savage, 1954), KMM (Klibanoff et al., 2005); VP (Maccheroni et al., 2006a,b) —, or with decision functionals describing incomplete preferences with inertia (Aumann, 1962; Walley, 1991; Bewley, 2002).

learning model; because of the central role played by randomization, I call this preferences the *Raiffa Strategy* (RS). In every period, the RS prescribes betting on the ambiguous urn on the color that is unambiguously more likely than the other. If there is no color that is unambiguously more likely than the other, then the RS prescribes the randomized act.

Second: *fears about model misspecification are well-justified.*

I find that the RS is more robust than the standard SEU/single-prior learning model to model misspecification, but yet it does not deliver robust statistical decisions. There are generic cases of misspecification in which a DM who follows the RS achieves a strictly lower average utility than betting on the risky urns.

2 Discussion of misspecified learning

The statistical decision problem I study is akin to an online categorization problem. In every period the DM guesses the color of a ball drawn from a new ambiguous urn and gets rewarded only if his prediction is correct. For this type of statistical decision problem, it is known that standard Bayes does not ensure decisions that are robust to model misspecification (Grünwald and Langford, 2007). Standard Bayes, which is equivalent to SEU matched with Bayesian learning, fails to deliver robust decisions under two types of model misspecification.

First, the learning problem might be misspecified in such a way that the Bayesian posterior converges to a Dirac on a model which is "harmful" for the DM when used to make decisions (Grünwald et al., 2017; Csaba and Szoke, 2018). This problem is symptomatic of tension between the driving force of Bayesian inference — the posterior concentrates on the maximum likelihood model — and the objective of the DM to identify the color with the highest frequency. Standard results in Bayesian statistics guarantee that the Bayesian posterior converges to a Dirac on the model with the lowest average K-L divergence (maximum likelihood) when unique (Berk, 1966). Our DM is, however, interested in making the largest number of correct bets, not on the model with the highest likelihood. When the maximum likelihood model does not attach the highest probability to the color that occurs more often, a Bayesian DM learns a model that is not the one that makes him choose the action that maximizes its average utility according to the true model.

Second, the learning problem might have a misspecified dependence structure. This type of misspecification can generate a mismatch between the sequence of acts chosen by the DM and the colors drawn. There are cases in which the average payoff of the DM is lower than that of the static minimax, even if the frequencies of acts and colors are those we would observe if the DM and Nature were playing the static minimax in every period. This problem is related to the known suboptimal performance of the "fictitious play" dynamics in the repeated game literature (Fudenberg and Levine, 1998). It is symptomatic of the tension between the Bayesian paradigm (on iid models), which treats all sequences with the same frequency as interchangeable, and the objective of the DM, whose average utility might change dramatically between sequences of draws and acts with the same frequencies.

In this paper, I show that both types of misspecification remain problematic even if the DM adopts a more conservative decision criterion, MEU rather than SEU, and a more conservative updating rule, multiple-prior learning rather than standard Bayes. These two types of misspecification errors are qualitatively different. In section 5.2 I show that learning a harmful model equally affects the MEU/multiple-prior learning model and the SEU/single-prior model, but it can be easily avoided universally (i.e., on every path) by introducing a natural symmetry assumption on the support. Conversely, in sections 5.3 and 5.4 I show that while the problems related to a misspecified dependence structure cannot be eliminated in general, the MEU/multiple-prior model provides a better hedge against their occurrence.

3 Multiple-prior learning

Multiple-prior learning describes the attitude of a DM that holds more than one prior distribution over a set of parameters and incorporates new information by independently updating each prior according to Bayes' rule. This learning model has been used to discuss long-run ambiguity in well-specified and misspecified settings, respectively (Marinacci, 2002; Marinacci and Massari, 2019; Massari and Newton, 2019).[3]

In our context, the DM observes a series of draws from the ambiguous urns. Each draw can reveal either a black or a white ball (i.e., for all t, the state space is $X_t = \{w, b\}$, with generic value x_t). The DM believes that draws are iid.[4] His prior support Θ characterizes a family of Bernoulli distribution $\mathcal{M} = \{P_\theta : \theta \in \Theta\}$. All probabilities are defined on the Borel σ-field $(X^\infty, \Sigma^\infty)$ where $X^\infty := \times^\infty X_t$, and Σ^∞ is the usual σ-algebra generated by the cylinders. With an abuse of notation, I use $P(x^t)$ to denote at the same time the probability that a model P attaches to the cylinder with base x^t $(i.e., Cyl(x^t) := \{x_1, ..., x_t, X_{t+1}, X_{t+2}, ...\})$, and the likelihood that model P attaches to the partial sequence $(x_1, ..., x_t)$.

[3]It is a known fact that prior-by-prior updating can lead to dynamically inconsistent choices (Epstein and Schneider, 2003; Epstein and Seo, 2010; Epstein et al., 2011). Concerns about dynamic-consistency have no bite in our setting because we focus exclusively on one-step-ahead decisions.

[4]The iid assumption is made to ease the exposition. By changing definition 3 to accommodate for a Markov structure, the same qualitative results can be derived, allowing models in $\mathcal{M}$ to have arbitrary (but finite) Markov K dependence structure.

P_{θ_0} is the true probability on $(X^\infty, \Sigma^\infty)$; there are no assumptions on P_{θ_0}. My results depend on observable properties of the empirical distribution, through the following notion.

Definition 1. *An event A occurs*

- on almost* every period *(according to the empirical measure) on path x^∞ if the fraction of periods in which the event occurs converges to 1 on x^∞.*
- on a negligible* fraction of periods *(according to the empirical measure) on path x^∞ if the fraction of periods in which the event occurs converges to 0 on x^∞.*
- on a positive* fraction of periods *(according to the empirical measure) on path x^∞ if the fraction of periods in which the event occurs converges to $k \in (0,1)$ on x^∞.*

The prior information about the parameters is summarized by prior distributions $\mu : 2^\Theta \to [0,1]$. The set of prior distributions is $\mathcal{C}$. For any prior distribution $\mu \in \mathcal{C}$ the joint distribution of the parameters and the observations is $P_\mu : 2^\Theta \to [0,1]$. By definition, for all $A \subseteq \Theta$ I have that:

$$P_\mu(A \times x^t) := \int_A P_\theta(x^t) d\mu.$$

I denote by $\mu(.|x^t) : 2^\Theta \to [0,1]$ the usual posterior given observations x^t, while $P_\mu(.|x^t) : \Sigma \to [0,1]$ is the one-step-ahead predictive distribution given observations x^t. By definition, $\forall \mu \in \mathcal{C}$:

$$P_\mu(x_{t+1}|x^t) := \int_\Theta P_\theta(x_{t+1}) d\mu(.|x^t) := \int_\Theta P_\theta(x_{t+1}) \frac{P_\theta(x^t) d\mu}{\int_\Theta P_\theta(x^t) d\mu}.$$

To ensure that all models in $\mathcal{M}$ can be unambiguously learned if correct, we assume that all priors in $\mathcal{C}$ are *regular*. Where a prior is regular if it has full support on a finite Θ (Marinacci, 2002), or if it has full support, and is positive and continuous for learning problems on which Θ has positive Lebesgue measure (Massari and Newton, 2019).

4 Decisions: Raiffa's Strategy

Let C be the space of consequences on which the DM has a bounded utility function $u : C \to \mathbb{R}$. I consider one-step-ahead acts, i.e., $\Sigma-$measurable maps $f : X \to C$ that associate a consequence to each observation in X. In each period t, the set of acts available to the DM is

$$\mathcal{F} = \begin{cases} f^b_{R,t} : \text{to bet on a black ball from the risky urn;} \\ f^w_{R,t} : \text{to bet on a white ball from the risky urn;} \\ f^b_{A,t} : \text{to bet on a black ball from the ambiguous urn;} \\ f^w_{A,t} : \text{to bet on a white ball from the ambiguous urn;} \\ f^{r^*}_t : \text{to use a fair coin to chose between } f^b_A \text{ and } f^w_A. \end{cases}$$

Each bet pays \$100 if won, \$0 otherwise.

As highlighted by Raiffa, f^{r^*} reshapes the choice between two ambiguous acts into a lottery with risky payoffs. In our learning context, it allows sampling from the ambiguous process facing risky payoffs. I say that a DM follows the *Raiffa Strategy* if he makes myopic decisions according to the MEU enriched with the preference for learning axiom (**A1**), and learns about the composition of the ambiguous urn according to the multiple-prior learning model.

Definition 2. *A DM follows the* RS *if in every period t after history x^{t-1}, he choses act*

$$f^*_t(x^{t-1}) := \text{argmax} \left\{ \min_{\mu \in \mathcal{C}} E_{\mu,t-1} u(f(x_t)), E_{r^*} u(f^{r^*}(x_t)) \right\},$$

where $E_{\mu,t}$ is the expectation according to $P_\mu(x_t|x^{t-1})$ and E_{r^*} is the expectation according to the objective randomization device.

Operationally, the RS prescribes the following:

- the DM bets on the ambiguous urn in every period;
- in those periods in which the set of posteriors of the multiple-prior learning model imply a lower MEU then that of the risky urn, the DM bets on a color at random;
- in those periods (if any) in which the MEU of betting on black (white) in the ambiguous urn is higher than that of the risky urn, the DM bets on black (white).

5 Main result

In this section, I introduce four types of payoff-relevant learning outcomes. Next, I discuss the conditions under which these payoff-relevant learning outcomes occur. This analysis shows that the RS is more robust to model misspecification than the standards SEU/single-prior learning model but yet not robust generically. The RS is not robust because there are misspecified learning problems in which it prescribes a sequence of actions with average payoff below the maximin.

Definition 3. *I say that RS delivers*

- Useful learning *on those paths on which it prescribes randomizing on a negligible** *fraction of periods, and delivers a (weakly) higher average utility than repeated bets on the risky urns.*
- Harmful learning*: on those paths on which it prescribes the same color on almost** *every period and delivers a strictly lower average utility than the risky urns.*
- Harmful inconclusive learning *on those paths on which it prescribes randomizing on a positive** *fraction of periods and delivers a strictly lower average utility than the risky urns.*
- No-learning *on those paths on which it prescribes randomizing on almost** *every period.*[5]

Notably, I show that *harmful learning* and *harmful inconclusive learning* are qualitatively different. *Harmful learning* affects, in the same way, the SEU/Bayesian learning model and the MEU/multiple-prior model but it can be avoided universally (i.e., on every path) by adopting a natural symmetry condition on the parameters in the prior support. Conversely, while *harmful inconclusive learning* is a more severe problem for the SEU/Bayesian learning model than for the MEU/multiple-prior model, it is a necessary evil of both models: there is no parametrization of the learning problem that ensures avoiding *harmful inconclusive learning*.

5.1 Useful learning

If draws are indeed iid, a sufficient condition for useful learning to occur is that the maximum likelihood model exists and attaches the highest probability to the color with the highest true probability of occurring. This assumption guarantees that the DM eventually learns correctly that betting on one color in the ambiguous process delivers a higher average payoff than betting on the risky urn (Berk, 1966). Proposition 1, below, generalizes this intuition to the case in which draws are not iid. It relies on the notion of *strong maximum likelihood*, SML henceforth, introduced by Marinacci and Massari (2019).[6]

Definition 4. *Given a path* $x^\infty \in X^\infty$ *and a family of models* $\mathcal{M} = \{P_\theta : \theta \in \Theta\}$*, I say that* $\hat{\theta} := \hat{\theta}(x^\infty, \Theta) \in \Theta$ *is the* strong maximum likelihood (SML) *model if, for every* $\theta \in \Theta \setminus \hat{\theta}, \lim_{t\to\infty} P_\theta(x^t)/P_{\hat{\theta}}(x^t) = 0$.

Marinacci and Massari (2019); Massari and Newton (2019) show that in a regular learning problem if the SML exists (and is unique) all posteriors concentrate on a Dirac distribution on

[5] It naturally follows from the strong law of large numbers that when no-learning occurs RS delivers the same average utility as the risky urns (Lemma 1).

[6] Definition 4 coincides with that of Marinacci and Massari (2019), provided that the SML is unique.

SML.[7] Proposition 1 makes use of this result to demonstrate that RS delivers useful learning if *i*) the SML model exists and, *ii*), the SML model attaches the highest probability to the color that has the highest empirical frequency. Furthermore, it confirms that useful learning occurs a.s. in well-specified learning problems.

Proposition 1. *Let $\mathcal{M} = \{P_\theta : \theta \in \Theta\}$ be a family of Bernoulli distributions parametrized by Θ. For all $\Theta \subset (0,1)$ and every compact and regular set $\mathcal{C}$ of prior distributions, the RS delivers*

a) useful learning on all and only those paths $x^\infty \in X^\infty$ on which a unique $\hat{\theta}(x^\infty)$ exists and

$$P_{\hat{\theta}(x^\infty)}(b) > (<).5 \Leftrightarrow \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=1}^{t} I_{\{x_\tau=b\}} = \bar{t}_b(x^\infty) > (<).5;$$

b) useful learning P_{θ_0}-a.s. if the learning problem is well-specified (i.e., $\theta_0 \in \Theta$).

Proof.

- By Marinacci and Massari (2019) Theorem 1 and Massari and Newton (2019), there exists a unique $\hat{\theta}(x^\infty) \Rightarrow$ all posteriors converge to a Dirac on $\hat{\theta}(x^\infty)$. Because $P_{\hat{\theta}(x^\infty)}(b) \neq .5$ the same color is selected by the RS for all large t. WLOG, let x^∞ be such that

$$i)\ P_{\hat{\theta}(x^\infty)}(b) > .5 \ \text{ and } \ ii)\ \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=1}^{t} I_{\{x_\tau=b\}} = \bar{t}_b(x^\infty) > .5;$$

  i) $\Rightarrow$ the RS prescribes betting on black on all large t;
  ii) $\Rightarrow \bar{u}(f^*) = \bar{t}_b(x^\infty)u(\$100) + (1-\bar{t}_b(x^\infty))u(\$0) > .5u(\$100) + .5u(\$0) =^{P\text{- a.s.}} \bar{u}(f^r)$.
  A similar argument can be used to show the opposite implication via
  $P_{\hat{\theta}(x^\infty)}(b) > .5$ and $\lim_{t\to\infty} \frac{1}{t}\sum_{\tau=1}^{t} I_{\{x_\tau=b\}} = \bar{t}_b(x^\infty) < .5 \Rightarrow \bar{u}(f^*) <^{P\text{-a.s.}} \bar{u}(f^r)$.

- Let P_{θ_0} be the true probability of black, by Marinacci and Massari (2019) (Theorem 2), (and Massari and Newton (2019)) the probability mass of each prior in $\mathcal{C}$ concentrate on θ_0 P_{θ_0}-a.s.. So, if $P_{\theta_0} \neq .5$ condition *a*) above is satisfied almost surely. If $P_{\theta_0} = .5$ any adopted betting strategy, including the RS, delivers the same average payoff almost surely by (Williams, 1991). Thus, its average payoff coincides (almost surely) with that of randomizing in every period, which coincides (almost surely by Lemma 1) with that of betting on the risky urns in every period.

□

Although Proposition 1 is a positive result, its conditions depend on properties of the true data generating process that cannot be known a priori. Its first line tells us that if all posteriors concentrate to a Dirac on a model that does not attach equal probability to both outcomes, the

[7] More precisely, Massari and Newton (2019) show that ambiguity fades away on all paths, even if $\lim_t \hat{\theta}_t$ does not exists.

RS delivers useful learning provided that that model attaches higher probability to the color with the highest empirical frequency — a requirement that cannot be a priori guaranteed without knowledge of the truth. Its second line depends on the truth in an even more stringent way. It tells us that a sufficient condition for useful learning to occur is that the learning problem is correctly specified because the DM eventually learns the true model.[8]

Example 1 illustrates Proposition 1.

- **Beliefs**: the DM believes draws are iid from two possible models: $\Theta = \{\theta_1; \theta_2\}$, where, for all t, $P_{\theta_1}(b_t) = .2$ and $P_{\theta_2}(b_t) = .8$. He has two full support priors on Θ: $\mathcal{C} := \{\mu^1, \mu^2\}$.
- *Useful learning* occurs, with strict inequality, on all paths in which the average number of black (white) exceeds that of white (black).

Proof: On all x^∞ such that $\bar{t}_b(x^\infty) = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=1}^{t} I_{\{x_\tau = b\}} > .5$
the posteriors calculated from both prior converge to $\theta = .8$.
$\Rightarrow$ the RS prescribes black, in almost every period: for all large t, $f_t^* = f^b$

$$\Rightarrow \bar{u}(f^*(x^\infty)) = \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=1}^{t} u(f_t^*) = \bar{t}_b(x^\infty)u(\$100) + (1 - \bar{t}_b(x^\infty))u(0)$$

$$> .5u(\$100) + .5u(0) =^{P\text{a.s.}} \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=1}^{t} u(f_t^r) = \bar{u}(f^r).$$

5.2 Harmful learning

Harmful learning occurs when the recommended color according to the most conservative prior is not the color that has the highest frequency almost surely according to the truth. Proposition 2 shows that a sufficient condition for harmful learning is the presence of an asymmetry between the most accurate conservative and anti-conservative model on the probability of a color. This asymmetry is potentially problematic because it guarantees the existence of a true model such that the model in $\mathcal{M}$ with the highest likelihood does not attach the highest probability to the color with the highest (true probability) empirical frequency.

Proposition 2. *Let $\mathcal{M} = \{P_\theta : \theta \in \Theta\}$ be a family of Bernoulli distributions parametrized by Θ, for every compact and regular set $\mathcal{C}$ of prior distributions if*

$$P_*(b) = \sup_{\theta\in\Theta}\{P_\theta(b) \leq .5\} \neq 1 - \inf_{\theta\in\Theta}\{P_\theta(b) \geq .5\} = 1 - P^*(b);$$

then, $\exists P_{\theta_0}$: RS delivers harmful-learning P_{θ_0}-a.s..

[8] The second condition of Proposition 1 holds in general: it does not require models in $\mathcal{M}$ to be iid.

Proof. Let recall that the average K-L divergence from P_θ to the true probability P_{θ_0} is

$$\bar{d}(P_{\theta_0}||P_\theta) := \lim_{t\to\infty} \frac{1}{t} E_{P_{\theta_0}}\left[\ln \frac{P_{\theta_0}(x^t)}{P_\theta(x^t)}\right].$$

WLOG, let $P_*(b) > 1-P^*(b)$ so that $\bar{d}(.5||P_*) < \bar{d}(.5||P^*)$. Continuity of the K-L divergence guarantees that $\exists\bar{\epsilon} > 0 : \forall\epsilon \in (0,\bar{\epsilon}), P_{\theta_0} := .5+\epsilon \Rightarrow \bar{d}(P_{\theta_0}||P_*) < \bar{d}(P_{\theta_0}||P^*)$. By Marinacci and Massari (2019) (Theorem 2), $\bar{d}(P_{\theta_0}||P_*) < \bar{d}(P_{\theta_0}||P^*) \Rightarrow$ for every compact set of regular priors $\mathcal{C}$, ambiguity fades away and all posteriors concentrate to a Dirac on P_*. Because $P_*(b) < .5$, RS recommends betting on white in all large t. Since the true probability of white is $.5-\epsilon < .5$ the claim follows by the strong law of large numbers:

$$\bar{u}(f^*) =^{P_{.5-\epsilon}\text{-a.s.}} (.5-\epsilon)u(\$100) + (.5+\epsilon)u(0) < (.5)u(\$100) + (.5)u(0) =^{P\text{-a.s.}} \bar{u}(f^r).$$

□

Proposition 2 does not depend on $|\mathcal{C}|$ because its condition is sufficient for the posteriors calculated from all priors to concentrate on the same (harmful) model. Therefore it applies to both the MEU/multiple-prior decision criterion and the SEU/Bayes' criterion. If the model with the lowest K-L divergence does not attach the highest probability to the color with the highest true probability to occur, both criteria recommend betting on the same incorrect color for all large t.

Example 2 illustrates a case in which the DM unambiguously learns a "harmful model" that makes him choose to bet on the color with the lowest frequency.

- **Beliefs**: The DM believes draws are iid from two possible models: $\Theta = \{\theta_1, \theta_2\}$, where, for all $t, P_{\theta_1}(b_t) = .9, P_{\theta_2}(b_t) = .4$. He has two full-support priors on Θ: $\mathcal{C} := \{\mu^1, \mu^2\}$.
- *Harmful learnig* occurs if draws are iid with probability $P_{\theta_0}(b) = .65$.

Proof By standard arguments,

$$P_{\theta_0}(b) = .65 \Rightarrow \lim_{t\to\infty} \frac{P_{\theta_1}(x^t)}{P_{\theta_2}(x^t)} =^{P\text{-a.s.}} \lim_{t\to\infty}\left(\frac{.9^{.65}.1^{.35}}{.4^{.65}.6^{.35}}\right)^t = \lim_{t\to\infty}(.90...)^t = 0,$$

Thus, P_{θ_2} is the SML model and ambiguity fades away with both posteriors attaching probability 1 to model P_{θ_2} (by Marinacci and Massari, 2019, Theorem 1).

The DM learns that model P_{θ_2} is unambiguously the maximum likelihood model in Θ. Accordingly, the DM chooses to bet on white on all large t. Ironically,
betting on white the DM obtains an average utility of $.35u(\$100) + .65u(\$0)$ P_{θ_0}-a.s.;
less than what he would get randomizing in every period: $.5u(\$100) + .5u(\$0)$ P_{θ_0}-a.s.;
less than what he would have gotten had he learned P_{θ_1}: $.65u(\$100) + .35u(\$0)$ P_{θ_0}-a.s..

Next, we present a sufficient condition on the primitives of the learning problem which prevents harmful learning on any path.

Proposition 3. *Let $\mathcal{M} = \{P_\theta : \theta \in \Theta\}$ be a family of Bernoulli distributions parametrized by Θ, for every compact and regular set $\mathcal{C}$ of prior distributions, if*

$$\begin{cases} P_*(b) = \sup_{\theta\in\Theta}\{P_\theta(b) \leq .5\} = 1 - \inf_{\theta\in\Theta}\{P_\theta(b) \geq .5\} = 1 - P^*(b), \\ \textit{and } \ \forall\hat{\theta} \in \Theta, \exists\mu^i \in \mathcal{C} : \mu^i(\hat{\theta}) = \max_{\theta\in\Theta} \mu^i(\theta) \end{cases},$$

then, there are no paths $x^\infty \in X^\infty$ on which the RS delivers harmful learning.

Proof. The symmetry assumption on the models in the prior support ensures that the maximum likelihood model attaches the highest probability to the color with the highest true probability; the assumption on the set of priors ensures that a color cannot be recommended when it does not correspond to the recommendation of the maximum likelihood model. □

Proposition 3's condition has two parts. First, it requires symmetry between the most accurate conservative and anti-conservative models on the probability of a color. This condition guarantees that if there is an SML model this model must be useful. Second, it requires a symmetry condition on the set of priors to prevent those cases in which ambiguity persists and an asymmetry on the set of priors dominates the symmetry in the prior support.

Remark: *how pervasive is the harmful learning problem?*
The harmful learning problem is not specific to categorization problems. The symmetry condition to avoid harmful learning of Proposition 3 applies to categorization problems only, but do not have a natural counterpart for the second and third prediction tasks below. Harmful learning might occur in any misspecified learning problem in which

- the utility function is not continuous (De Rooij et al., 2014, categorization problems);
- the utility function does not depend exclusively on the moments of the learned distribution (e.g., the CARA utility, Csaba and Szoke, 2018);
- the model family is not convex (e.g., a set of nested linear (or ridge) regression models, Grünwald et al., 2017).

These results exhort caution regarding misspecification in practical decision problems and in constructing theoretical models. They highlight that when the learning problem is misspecified the use of the K-L divergence and Bayes' rule as a measure of distance between probabilistic models and learning is neither "natural", nor robust.

5.3 Harmful inconclusive learning

With inconclusive learning, I denote the situation in which the RS prescribes randomizing a positive* fraction of periods. When harmful inconclusive learning occurs, the DM has strictly lower average utility following the RS than betting on the risky urns because when he does not bet on a color at random, he is later proven incorrect by the data. For example, the RS might prescribe betting on black after every period in which there are K_b more black draws than white draws, while the true data generating process for the ambiguous urns is not iid and attaches 0 probability to black when there are K_b more black draws than white draws.

Proposition 4 shows that harmful inconclusive learning is a "necessary evil" of the MEU/multiple-prior decision criterion. There are no parameter specifications which guarantee that the RS average utility is at least as high as the maximin on any sequence.

Proposition 4. *Let $\mathcal{M} = \{P_\theta : \theta \in \Theta\}$ be a family of Bernoulli distributions parametrized by Θ. For all $\Theta \subset (0,1)$ and for every compact and regular set $\mathcal{C}$ of prior distributions on Θ,there exists a true probability P_{θ_0} such that harmful inconclusive learning occurs P_{θ_0}-a.s.*

Proof. It is easy to verify that the RS is inconclusive P_{θ_0}-a.s. when

$$P_{\theta_0}(b_t|x^{t-1}) = \begin{cases} P_-(b) = 0 & \text{if } f_t^*(x^{t-1}) = b \\ P_+(b) = 1 & \text{if } f_t^*(x^{t-1}) \neq b; \end{cases} .$$

The RS delivers harmful inconclusive learning P_{θ_0}-a.s. because it recommends the incorrect act every time it prescribes b and the correct act at most half of the time it does not prescribe b (half of the time only if it prescribes randomizing), thus its average utility is lower than that of betting on the risky urn.[9] □

Example 3 illustrates Proposition 4.

- **Beliefs**: the DM believes draws are iid from two possible models: $\Theta = \{\theta_1; \theta_2\}$, where, for all $t, P_{\theta_1}(b_t) = .2$ and $P_{\theta_2}(b_t) = .8$. He has two full-support priors on Θ: $\mathcal{C} := \{\mu^1, \mu^2\}$ with $\mu^1(\theta_1) = .1 = 1 - \mu^2(\theta_1)$.
- *harmful inconclusive learning* occurs on path $x^\infty := \{b, b, w, b, w, b, ...\}$.

Proof: On $x^\infty = \{b, b, w, b, w, b, ...\}$, for $i = 1, 2$,
$\forall t$ odd $\geq 3, P_{\mu_i}(b_t|x^{t-1}) = \frac{\mu_i 2^3\left(\sqrt{.2.8}\right)^{t-3} + (1-\mu_i).8^3\left(\sqrt{.2.8}\right)^{t-3}}{\mu_i.2^2\left(\sqrt{.2.8}\right)^{t-3} + (1-\mu_i).8^2\left(\sqrt{.2.8}\right)^{t-3}}$

[9] Hajek (1982) result on the occupation time that a mean reverting process spends close to its mean reverting point can be used to show that the RS delivers harmful inconclusive learning P_{θ_0}-a.s.for

$$P_\theta : P_\theta(b_t|\sigma^{t-1}) = \begin{cases} P_-(b) < .5 & \text{if } f_t^*(x^{t-1}) = b \\ P_+(b) > .5 & \text{if } f_t^*(x^{t-1}) \neq b \end{cases} .$$

$$\forall t \text{ even} \quad , P_{\mu_i}(b_t|x^{t-1}) = \frac{\mu_i 2^2\left(\sqrt{.2.8}\right)^{t-2} + (1-\mu_i).8^2\left(\sqrt{.2.8}\right)^{t-2}}{\mu_i.2\left(\sqrt{.2.8}\right)^{t-2} + (1-\mu_i).8\left(\sqrt{.2.8}\right)^{t-2}}$$

$$\Rightarrow \begin{cases} \min\{P_{\mu_1}(b_t|x^{t-1}), P_{\mu_2}(b_t|x^{t-1})\} > .5 & \forall t \text{ odd} \geq 3, \\ \begin{cases} \min\{P_{\mu_1}(b_t|x^{t-1}), P_{\mu_2}(b_t|x^{t-1})\} < .5 \\ \max\{P_{\mu_1}(b_t|x^{t-1}), P_{\mu_2}(b_t|x^{t-1})\} > .5 \end{cases} & \forall t \text{ even.} \end{cases}$$

$$\Rightarrow \text{the RS prescribes} f^*_t = \begin{cases} f^b & \forall t \text{ odd} \geq 3, \\ f^{r^*} & \forall t \text{ even.} \end{cases}$$

Therefore,

$$\begin{aligned} \bar{u}(f^*(x^\infty)) &= O\left(\frac{1}{t}\right) + \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=2}^{t} u(f^*_t) \\ &= \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=3}^{t} u(f^b)\Bigg|_{\tau \ odd} + \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=2}^{t} u(f^{r^*})\Bigg|_{\tau \ even} \\ \dagger \quad &=^{P^*\text{a.s.}} .5u(\$0) + .5(.5u(\$0) + .5u(\$100)) \\ &< .5u(\$0) + .5u(\$100) \\ &=^{P\text{a.s.}} \bar{u}(f^r) \end{aligned}$$

†) The $.5u(\$0)$ term reflects the fact that the DM choses f^b incorrectly in every $\tau \geq 3$ odd. The $.5(.5u(\$0) + .5u(\$100))$ term follows from the strong law of large numbers because in every even period the DM randomizes and wins with probability .5.

Remark: *comparing the RS with fictitious play dynamics*

A way to understand the suboptimal performance of the RS under "harmful inconclusive learning" is to draw a parallel between RS and known results about the possibility of underperformance of the *fictitious play* dynamics, FP henceforth, in repeated games (Brown, 1951; Fudenberg and Levine, 1998). When Θ is the whole simplex and $\mathcal{C}$ contains only one Beta prior, the choices prescribed by the RS coincide with the best response dynamic of FP (for some initial weights). In every period, the RS (FP) prescribes choosing the act that is the best response to the weighted frequency of past realizations (action of the opponent). As per FP, the sequences in which the RS underperforms against randomizing in every period — the minimax Nash equilibrium — are those in which there is a positive fraction of periods in which the action recommended by the best response function changes. This eventuality can occur only if the data generating process is not iid as erroneously assumed by the DM (player).

When $\mathcal{C}$ is not a singleton, the main difference between the RS and FP is that the RS is

more cautious than FP since it prescribes randomized actions more often than the FP. While FP prescribes randomizing only if the posterior calculated from a unique prior does not support bets on a color, RS prescribes randomizing in all those cases in which there are at least two posteriors that do not support bets on the same color. This difference makes the RS MEU/multiple prior setting more robust to *harmful inconclusive learning* than the standard SEU/Bayesian model.
Example 4: illustrates these observations.

- **Beliefs**: The DM believes draws are iid and that all compositions of the urn are possible: $\Theta = (0, 1)$. The DM's prior information of the composition of the urn is captured by the family of Beta priors with parameters in two strictly positive, finite intervals $[a, b], [c, d]$: $\mathcal{C} = \{Beta(\alpha, \beta),\ \alpha \in [a, b], \beta \in [c, d]$ on Θ.

  Taking advantage of the conjugacy of the Beta prior, standard calculation shows that

  $$\text{the RS prescribes as follows:} f_t^* = \begin{cases} f^b & \text{if } \sum_{\tau=1}^{t-1} I_{\{\underline{b}(x^\tau)>.5\}} > 0 \\ f^w & \text{if } \sum_{\tau=1}^{t-1} I_{\{\underline{w}(x^\tau)>.5\}} > 0 \\ f^{r^*} & \text{otherwise} \end{cases} ;$$

  where $\underline{b}(x^t) := \min_{\alpha,\beta} \frac{\alpha + t_b(x^t)}{t+\alpha+\beta}$ and $\underline{w}(x^t) := \min_{\alpha,\beta} \frac{\beta + t_w(x^t)}{t+\alpha+\beta}$ is the expected probability of black and the expected probability of white according to the most conservative posterior obtained from the priors in $\mathcal{C}$, respectively, and $t_b(x^t) = t - t_w(x^t)$ is the number of black balls drawn on x^t. If $\mathcal{C}$ is a singleton $\underline{b}(x^\tau) = 1 - \underline{w}(x^\tau)$, the RS coincides with FP, and randomization cannot be chosen in almost every period. Otherwise, there are numbers K_b and K_w such that RS prescribes randomizing whenever the difference between the number of black draws and white draws belongs to the interval (K_w, K_b).

5.4 No-learning

In this section, I show that the possibility of no-learning is a special feature of the MEU/multiple prior criterion. It is an outcome that cannot occur in the absence of ambiguity, i.e., when the set $\mathcal{C}$ of prior distributions is a singleton. So, the MEU/multiple prior model is more robust than the standard SEU/Bayesian model on some sequences because an ambiguity-averse DM who follows the MEU/multiple prior criterion is indeed more prudent than a SEU/Bayesian DM.

Proposition 5. *Let $\mathcal{M} = \{P_\theta : \theta \in \Theta\}$ be a family of iid Bernoulli distributions parametrized by Θ, there exists x^∞ such that RS delivers no-learning only if $\mathcal{C}$ is not singleton.*

Proof. Let $\mathcal{C}$ be a singleton. First note that on every path in which the limiting average does not exist the RS does not prescribe randomizing in almost every period. Second, on every path in which the limiting average exists, at most two models in Θ can have positive weight in a positive fraction of periods. Because of the iid assumption, the likelihoods of each model depends only on past frequencies, and given a frequency there can only be two models equidistant in K-L divergence. Lastly, with two models, randomizing in most periods is impossible. The RS with a unique prior and two models prescribes randomizing if and only if $\frac{\mu_{\theta^1} P_{\theta^1}(x^t)}{\mu_{\theta^2} P_{\theta^2}(x^t)} = 1$. The RS prescribing randomizing a positive fraction of periods implies $\exists t, t+1$:

$$\frac{\mu_{\theta^1} P_{\theta^1}(x^t)}{\mu_{\theta^2} P_{\theta^2}(x^t)} = 1 = \frac{\mu_{\theta^1} P_{\theta^1}(x^{t+1})}{\mu_{\theta^2} P_{\theta^2}(x^{t+1})} := \frac{P_{\theta^1}(x_{t+1}) \mu_{\theta^1} P_{\theta^1}(x^t)}{P_{\theta^2}(x_{t+1}) \mu_{\theta^2} P_{\theta^2}(x^t)} = \frac{P_{\theta^1}(x_{t+1})}{P_{\theta^2}(x_{t+1})},$$

which is impossible because $P_{\theta^1}(x_{t+1}) \neq P_{\theta^2}(x_{t+1})$ □

Proposition 5 shows that randomizing in every period is never an outcome of the SEU/Bayes decision criterion, while examples (e.g., example 4) show that it is a possible outcome of the MEU/multiple prior criterion. Broadly speaking, a decision maker with a rich set of priors randomizes more often than if he had a unique prior because he needs more evidence in favor of each color to be persuaded to bet on that color than if he had a unique prior.

6 Conclusion

I show that the MEU/multiple-prior criterion is more robust than the SEU/Bayes' criterion to model misspecification but it does not guarantee robust decisions generically. Adding a more conservative decision criterion and a more conservative learning rule increases the robustness of the decision criterion but is not enough to completely hedge against dynamic ambiguity. I identify two types of misspecification problems that may occur: harmful learning and harmful inconclusive learning. These two problems are qualitatively different. The first one can be avoided by using natural symmetry assumptions; the second one is a "necessary evil" of both the SEU/single prior and the MEU/multiple-prior decision criteria, but it affects the former model more severely than the latter.

A Appendix

Lemma 1. *If RS delivers no-learning then $\bar{u}(f^*(x^\infty)) = \bar{u}(f^r(x^\infty))$ $P^* \times P$-a.s..; where P and P^* are the probability of the risky urn and of the coin, respectively.*

Proof. If RS delivers no-learning the DM randomizes in a unitary fraction of periods and the result follows by the strong law of large numbers. □

References

Aumann, R. J. (1962). Utility theory without the completeness axiom. *Econometrica: Journal of the Econometric Society*, pages 445–462.

Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *The Annals of Mathematical Statistics*, 37(1):51–58.

Bewley, T. F. (2002). Knightian decision theory. part i. *Decisions in economics and finance*, 25(2):79–110.

Branger, N., Larsen, L. S., and Munk, C. (2013). Robust portfolio choice with ambiguity and learning about return predictability. *Journal of Banking & Finance*, 37(5):1397–1411.

Brown, G. W. (1951). Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Montrucchio, L. (2013). Ambiguity and robust statistics. *Journal of Economic Theory*, 148(3):974–1049.

Csaba, D. and Szoke, B. (2018). Learning with misspecified models.

De Rooij, S., Van Erven, T., Grünwald, P. D., and Koolen, W. M. (2014). Follow the leader if you can, hedge if you must. *The Journal of Machine Learning Research*, 15(1):1281–1316.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, pages 643–669.

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

Epstein, L. G. and Seo, K. (2010). Symmetry of evidence without evidence of symmetry. *Theoretical Economics*, 5(3):313–368.

Epstein, L. G., Seo, K., et al. (2011). Symmetry or dynamic consistency. *BEJ Theor. Econ.(Advances)*.

Fudenberg, D. and Levine, D. (1998). Learning in games. *European economic review*, 42(3):631–639.

Garlappi, L., Uppal, R., and Wang, T. (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies*, 20(1):41–81.

Gilboa, I. (2009). *Theory of decision under uncertainty*, volume 1. Cambridge university press Cambridge.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics*, 18(2):141–153.

Grünwald, P. and Langford, J. (2007). Suboptimal behavior of bayes and mdl in classification under misspecification. *Machine Learning*, 66(2-3):119–149.

Grünwald, P., Van Ommen, T., et al. (2017). Inconsistency of bayesian inference for misspecified linear models, and a proposal for repairing it. *Bayesian Analysis*, 12(4):1069–1103.

Hajek, B. (1982). Hitting-time and occupation-time bounds implied by drift analysis with applications. *Advances in Applied probability*, 14:502–525.

Hansen, L. P., Marinacci, M., et al. (2016). Ambiguity aversion and model misspecification: An economic perspective. *Statistical Science*, 31(4):511–515.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2006a). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2006b). Dynamic variational preferences. *Journal of Economic Theory*, 128(1):4–44.

Marinacci, M. (2002). Learning from ambiguous urns. *Statistical Papers*, 43(1):143–151.

Marinacci, M. and Massari, F. (2019). Learning from ambiguous and misspecified models. *Journal of Mathematical Economics*, 84:144–149.

Massari, F. and Newton, J. (2019). When does ambiguity fade away? *Available at SSRN 3473625*.

Raiffa, H. (1961). Risk, ambiguity, and the savage axioms: comment. *The Quarterly Journal of Economics*, 75(4):690–694.

Savage, L. J. (1954). *The Foundations of Statistics*. Courier Corporation.

Wald, A. (1950). Statistical decision functions.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. Chapman & Hall.

Watson, J., Holmes, C., et al. (2016). Approximate models and robust decisions. *Statistical Science*, 31(4):465–489.

Williams, D. (1991). *Probability with martingales*. Cambridge University Press.