Learning with Misspecified Models*

Dániel Csaba† Bálint Szőke‡

June 2, 2019

Abstract

We consider Bayesian learning about a stable environment when the learner's entertained set of probability distributions (likelihoods) is misspecified. We study how this entertained set affects the agent's welfare through the limit point of learning and the associated best-responding policy. To this end, we introduce a performance measure of likelihoods based on their induced policies' long-run average payoff. Using this measure we define two asymptotic properties of sets of likelihoods that a utility-maximizing agent would find desirable. We show that arbitrary sets of likelihoods coupled with Bayesian learning fail to satisfy both of our properties. However, we characterize a class of decision problems for which one can construct a misspecified set of likelihoods that achieves the highest attainable long-run average payoff irrespective of the data generating process. Our recommendation builds on the payoff-relevant moments—specific to the decision problem at hand—to provide the sufficient statistics of an exponential family of likelihoods. This set of likelihoods coupled with Bayesian learning ensures that the moments targeted by learning coincide with the moments that are relevant for implementing well-performing policies.

Keywords: Bayesian learning; Misspecified models; Statistical Decision Theory

*We thank Timothy Christensen, Toru Kitagawa, Jonathan Payne, and Thomas Sargent for insightful discussions.
†Department of Economics, New York University, E-mail: daniel.csaba@nyu.edu.
‡Department of Economics, New York University, E-mail: balint.szoke@nyu.edu.

1 Introduction

It is self-evident that economists build simple models that are at best crude approximations of the complex environment they wish to describe. Part of the reason for this strategy is that learning the true data generating process seems to be a hopeless endeavor. Nonetheless, it is not obvious that the approach of using carefully chosen misspecified models is necessarily flawed. We do not seek to learn because we are interested in the data generating process *per se*, we learn because knowing certain features of the environment could be necessary for making good decisions. Intuitively, these features should be our main guides when building our models. How can we make sure that these features are appropriately targeted when we confront our models with data? Can we find a carefully chosen misspecified model that gives rise to decisions as good as if we knew the data generating process itself?

To address these questions in a familiar setting, we study a canonical Bayesian decision problem. More precisely, we consider a Bayesian decision maker (DM) learning about the distribution P of an exogenous stochastic process with the aim of uncovering features of the data generating process that are instrumental for the decision problem at hand. Bayesian *learning* entails specifying a parametric family of likelihoods, $\mathcal{M}$, accompanied with a (strictly positive) prior distribution that gets updated according to Bayes' rule as new observations from P arrive. This updating process induces a sequence of posterior distributions over $\mathcal{M}$ that, under well-known regularity conditions, will asymptotically concentrate on the likelihood $Q^{\mathrm{KL}}_{\mathcal{M},P} \in \mathcal{M}$ that minimizes the Kullback-Leibler divergence from P. Bayesian *decision making* entails choosing a sequence of *policies* defined as best response functions with respect to elements of the sequence of posteriors, using the DM's preference as a benchmark to what "best" means. The limit point $a_{\mathcal{M},P}$ of this sequence of policies is the best response with respect to $Q^{\mathrm{KL}}_{\mathcal{M},P}$. This paper investigates how the set of entertained likelihoods, $\mathcal{M}$, can influence the DM's welfare through its impact on the pair of limit points $(Q^{\mathrm{KL}}_{\mathcal{M},P}, a_{\mathcal{M},P})$ when it cannot be guaranteed that $P \in \mathcal{M}$, that is, that $\mathcal{M}$ is correctly specified.

We define two properties of $\mathcal{M}$ that provide uniform performance guarantees with respect to a class $\mathcal{P}$ of potential data generating processes.[1] The key building block of our definitions is a preference-based performance measure of likelihoods: the *long-run average payoff* of likelihood $Q \in \mathcal{P}$ is defined as the expected utility under P induced by a policy that is a best response with respect to Q.[2] We use this measure to define an undominatedness property of $\mathcal{M}$, the so called *misspecification-proofness*, requiring that, irrespective of which $P \in \mathcal{P}$ generates the data, there is no other set of likelihoods that coupled with Bayes learning would asymptotically lead to a likelihood with higher long-run average payoff. This is a demanding property that captures one of the most desirable features of correctly specified sets of likelihoods. We also introduce a weaker performance property of $\mathcal{M}$, the so called *local payoff-optimality*, requiring that, irrespective of which $P \in \mathcal{P}$ generates the data, the likelihood

[1]Importantly, this class can be large. So large that the convergence of posteriors cannot be guaranteed if $\mathcal{M} = \mathcal{P}$ (see e.g. Diaconis and Freedman [1986]). An illustrative example that we use throughout the paper is the class of i.i.d. distributions.

[2]This quantity is meant to capture the idea that learning matters only indirectly, through implementing a given policy.

that the $\mathcal{M}$-implied posteriors asymptotically concentrate on leads to the highest long-run average payoff *within* $\mathcal{M}$. Recognizing that the actual data generating process is unknown, we argue that our P-independent properties are reasonably required from any set of likelihoods that the DM entertains.[3]

We show that an *arbitrary* $\mathcal{M}$ does not satisfy either of our properties. This means that a misspecified set $\mathcal{M}$ can cause the Bayesian learner to implement a policy function that appears suboptimal even *relative to the entertained set*, thereby violating local payoff-optimality. The source of this suboptimality is the misalignment of two implicit loss functions pertinent to Bayesian decision making: one for learning (KL-divergence), and one for decision making (utility function). Intuitively, once a set of likelihoods is specified, its elements tell Bayes' rule what statistical moments of P it should focus on. If that set of moments is different from the payoff-relevant moments that determine the DM's policy function, learning can focus on "wrong" features of the environment. Clearly, under correct specification the misalignment of loss functions is inconsequential, but it can be of first order importance when misspecification is present.

We demonstrate these concepts through a standard consumption-saving problem. After diagnosing the source of suboptimality, we identify a class of decision problems, for which we can construct sets of likelihoods that are both misspecification-proof and locally-payoff-optimal. The key idea is to tailor the misspecified set to preferences. In particular, we recommend constructing an *exponential family* of likelihoods by using the DM's vector of payoff-relevant moments as sufficient statistics. These moments are derived from the underlying decision problem and thus are determined by preferences. By using our exponential family of likelihoods, the decision maker ensures that she learns about the "right" features of the data generating process.

We follow the statistical decision theory proposed by Wald [1950]. The primitives are: (i) a payoff function and (ii) a set of entertained likelihoods with a prior. Despite having a fully-embraced prior to make forward-looking decisions, ideally, the DM would want to maximize her payoff function *under the data generating process*. That said, we deviate from the subjectivist Bayesian view by assuming that the DM treats her likelihoods as instruments for choosing good policies, rather than as indisputable parts of preferences that provide psychological value by themselves. Such an assumption is necessary for any reasonable form of assessment of models that are misspecified. If we were to take the subjectivist view, the DM would always do her best according to the wrong model her decisions are based on. In order to avoid the conclusion of this circular argument, we follow Blume et al. [2018] and take the perspective of an outside observer that allows us to compare potentially misspecified priors based on their asymptotic implications.

Our proposed properties of $\mathcal{M}$ are meant to capture these asymptotic implications without considering the transition. Therefore, assuming that the standard regularity conditions are satisfied, they are restricted to the *support* of the prior ignoring the specific weighting. This permits to incorporate potential non-data information through flexible prior weightings over $\mathcal{M}$. As such, while we use

[3]Similar to the notion of *admissibility*, the strength of our properties is to help abandoning strategies that are undesirable.

Bayesian terminology throughout the paper, our findings are applicable to any converging learning rule such that the entertained hypotheses are representable with probability distributions over the observables and such that the likelihood of each hypothesis is assessed in light of the data.[4]

Our long-run average payoff based performance measure is related to the notion of *asymptotic frequentist risk* which is an object of interest in Müller [2013]. He devises a procedure that improves the frequentist risk of the Bayes estimator of parameter θ by substituting the original posterior with an artificial normal posterior centered at θ with "sandwich" covariance matrix. While this procedure modifies the set of likelihoods only through higher-order properties and keeps the targeted parameter of interest fixed, our proposed $\mathcal{M}$ changes the targeted parameters themselves.

Our results illustrate that the set of likelihoods that the decision maker entertains should be tailored to the DM's objective as opposed to the environment. Even if the data generating process is highly complex in the statistical sense, there are decision problems for which a simple misspecified set can implement the optimal policy if it is targeted at the appropriate features of the environment. Our results can thus be viewed as a recipe for model building.

The rest of the paper is structured as follows. Section 2 introduces key concepts and notation. In section 3, we demonstrate how misspecification can lead to suboptimal long-run behavior through a sequence of examples. Section 4 presents our recommendation that can be used to resolve this suboptimality in certain cases. Related literature is summarized in section 5. Section 6 concludes.

2 General framework and notation

The environment is described by a probability space $(\Omega, \mathcal{F}, P)$ such that a strictly stationary and ergodic observable state vector X takes values in the measurable space $(\mathcal{X}, \Xi)$ with distribution P. Alternative descriptions of the environment can be obtained by replacing P with some other strictly stationary and ergodic distribution. Loosely, we use $\mathcal{P}$ to denote the set of distributions over $\mathcal{X}$ that the DM deems to be plausible for the specific problem at hand.

The decision maker chooses a *policy function*, $a \colon \mathcal{X} \to \mathcal{C}$, that assigns a particular action from some choice set $\mathcal{C}$ to every realization x of the state vector X. Let $\mathcal{A}$ denote the collection of all possible policy functions. Payoffs are described by the period *utility function*, $u \colon \mathcal{C} \times \mathcal{X} \to \mathbb{R}$, and possibly depend on the state. To describe the decision maker's objective, we introduce a functional, $\mathsf{U} \colon \mathcal{A} \times \mathcal{P} \to \mathbb{R}$, defined as,

$$\mathsf{U}(a, Q) := \int_{\mathcal{X}} u(a(x), x) \mathrm{d}Q(x). \tag{1}$$

U defines the *expected payoff* induced by a policy function $a \in \mathcal{A}$ under a distribution $Q \in \mathcal{P}$. Ideally,

[4] This includes both Bayesian learning and anticipated utility learning (see Kreps [1998]) accompanied with some frequentist procedures (e.g. MLE). On the other hand, we do not consider *active* learning. That is, we assume that the implemented policy function does not affect the information that the DM observes or the probability distribution that she wants to learn about.

the decision maker would want to implement a policy function a^* that maximizes the *expected payoff under P*

$$a^* \in \arg\max_{a\in\mathcal{A}} \ \mathsf{U}(a,P). \tag{2}$$

Since we assume that the environment is strictly stationary and ergodic, $\mathsf{U}(\cdot,P)$ is equal to the average payoff that the decision maker realizes in the long-run.[5] To distinguish this notion from the expected payoff under arbitrary $Q\in\mathcal{P}$, we will henceforth call $\mathsf{U}(\cdot,P)$ the decision maker's *long-run average payoff.*

However, P is unknown, so the decision maker has to solve an alternative problem having objective $\mathsf{U}(\cdot,Q)$ in which P is replaced by some approximating distribution Q. To this end, she entertains a set of *likelihoods* $\mathcal{M}$, i.e., a family of strictly stationary and ergodic probability distributions $Q_\theta\in\mathcal{P}$, each indexed by a finite parameter vector θ,

$$\mathcal{M} := \{Q_\theta \ : \ \theta\in\Theta\}, \qquad \text{where} \quad \Theta\subseteq\mathbb{R}^p. \tag{3}$$

We are interested in situations in which there is no guarantee that $\mathcal{M}$ is *correctly specified*, i.e., that $P\in\mathcal{M}$. We call the set $\mathcal{M}$ *misspecified* if $P\notin\mathcal{M}$.

Initialized with some prior distribution over $\mathcal{M}$, Bayes' rule induces a sequence of posteriors that summarize the decision maker's best guesses for P at every point in time after the available data is taken into account. It is well known that under certain regularity conditions [Shalizi, 2009], the sequence of posteriors will eventually concentrate on the likelihoods in $\mathcal{M}$ that minimize the Kullback-Leibler (KL) divergence from the data generating process,[6]

$$\mathcal{Q}^{\mathrm{KL}}_{\mathcal{M},P} := \arg\min_{\theta\in\Theta} \ D_{\mathrm{KL}}\left(P \parallel Q_\theta\right). \tag{4}$$

In other words, after observing an infinite sequence of signals, Bayes' rule will suggest using a distribution from $\mathcal{Q}^{\mathrm{KL}}_{\mathcal{M},P}$ as the best approximation of P. While in principle (4) defines a set, we will only consider $\mathcal{M}$'s that imply a singleton $\mathcal{Q}^{\mathrm{KL}}_{\mathcal{M},P}$. To emphasize this, we use the notation $Q^{\mathrm{KL}}_{\mathcal{M},P}$. The KL-divergence minimizing likelihood has interesting information theoretic interpretations, but it is not obvious what its properties are in terms of long-run average payoff: the quantity that the decision maker ultimately cares about.

In order to investigate the payoff-relevant properties we define a performance measure of different

[5]For a Ξ-measurable policy a, such that the function $\bar{u}\colon x\mapsto u(a(x),x)$ is P-integrable, the ergodic theorem implies

$$\mathsf{U}(a,P) \overset{\text{a.s.}}{=} \lim_{t\to\infty} \ \frac{1}{t}\sum_{k=0}^{t-1}\bar{u}\left(\mathbb{T}^k x\right) = \lim_{\beta\nearrow 1} \ (1-\beta)\sum_{k=0}^{\infty}\beta^k\bar{u}\left(\mathbb{T}^k x\right)$$

where $\mathbb{T}\colon\mathcal{X}\to\mathcal{X}$ is the ergodic, measure-preserving shift operator on $(\Omega,\mathcal{F},P)$. The last equality uses the Abel summation formula to illustrate that $\mathsf{U}(\cdot,P)$ can be also viewed as the zero-discounting-limit of the utility of someone who knows P.

[6]The KL divergence is $D_{\mathrm{KL}}(P\parallel Q) = \int\log\frac{p(x)}{q(x)}\mathrm{d}P(x) = \int\log p(x)\mathrm{d}P(x) - \int\log q(x)\mathrm{d}P(x)$.

likelihoods in terms of the long-run average payoff of the policy function they induce. Correspondingly, a crucial component of this object is the *best response function* $b \colon \mathcal{P} \to \mathcal{A}$, defined as[7,8]

$$b(Q) \in \arg\max_{a \in \mathcal{A}} \; \mathsf{U}(a, Q). \tag{5}$$

Our proposed performance measure of likelihood Q combines b with the expected payoff function under P:

$$\mathsf{U}\left(b(Q), P\right). \tag{6}$$

This gives the realized long-run average payoff induced by an arbitrary likelihood Q. By using P to evaluate the performance of Q, we effectively take the perspective of an outside observer and intend to capture the idea that learning influences the decision maker's welfare only indirectly through its induced policy functions.[9] While we find it instructive, this notion suffers from the fact that it hinges on the unknown data generating process. Nevertheless, it can be used to construct a P-independent property of $\mathcal{M}$.

Definition 1 (Misspecification-proof $\mathcal{M}$)**.**
Consider a decision problem characterized by the triplet $(\mathcal{P}, \mathcal{A}, u)$. The family of likelihoods $\mathcal{M} \subseteq \mathcal{P}$ is misspecification-proof with respect to $(\mathcal{P}, \mathcal{A}, u)$, if there exists no $\mathcal{M}' \subseteq \mathcal{P}$, such that

$$\mathsf{U}\left(b\left(Q^{KL}_{\mathcal{M}',P}\right), P\right) \geq \mathsf{U}\left(b\left(Q^{KL}_{\mathcal{M},P}\right), P\right) \qquad \forall P \in \mathcal{P},$$

with strict inequality for some $P \in \mathcal{P}$.

Definition 1 describes an asymptotic performance property of a family of likelihoods $\mathcal{M}$ relative to a specific decision problem. Being misspecification-proof guarantees that there is no other (potentially misspecified) set of likelihoods that provides *uniformly* better asymptotic performance over the entire set $\mathcal{P}$ of potential data generating processes.[10] Intuitively, a decision maker with an objective (2), fearing misspecification, would find this property desirable.

However, misspecification-proofness is an admittedly demanding property, as it is defined globally in relation to all families of likelihoods. A related, but significantly weaker, performance property can be defined in a local manner, that is, conditionally on the set $\mathcal{M}$:[11]

Definition 2 (Locally-payoff-optimal $\mathcal{M}$)**.**
Consider a decision problem characterized by the triplet $(\mathcal{P}, \mathcal{A}, u)$. The family of likelihoods $\mathcal{M} \subseteq \mathcal{P}$

[7] Considering only learning rules that converge to a single likelihood asymptotically, we can determine the best-responding functions without having to consider mixture distributions.

[8] While potentially there could be a set of best-responding policy functions, for ease of exposition, we assume that the best-responding policy function is unique.

[9] The standard Bayesian approach would use $\mathsf{U}\left(b(Q), Q\right)$ to evaluate the implications of Q.

[10] The concept is akin to undominatedness in game theory. In our context, the decision maker choosing among sets of likelihoods $\mathcal{M}$ plays against Nature whose action space is the set $\mathcal{P}$ of potential data generating processes.

[11] This property is related to the loss-function-based consistency notion typically used in the statistical learning theory literature (see Vapnik [1995]).

is locally-payoff-optimal with respect to $(\mathcal{P}, \mathcal{A}, u)$*, if for all* $P \in \mathcal{P}$*, there exists no* $Q' \in \mathcal{M}$*, such that,*

$$\mathsf{U}\left(b\left(Q'\right), P\right) > \mathsf{U}\left(b\left(Q^{KL}_{\mathcal{M},P}\right), P\right).$$

Local payoff-optimality requires that the likelihood $Q^{\mathrm{KL}}_{\mathcal{M},P}$, that the Bayesian posterior asymptotically concentrates on, generates the highest long-run average payoff *within the set* $\mathcal{M}$. In this sense, if there is no other likelihood within the entertained set $\mathcal{M}$ with a higher long-run average payoff, Bayesian learning is successful in maximizing the unknown objective (2) at least over the entertained set $\mathcal{M}$.

While there are many trivial sets of likelihoods that are locally payoff-optimal—such as any singleton—not satisfying this property is symptomatic of a deeper issue which underlies the violation of misspecification-proofness. This issue is the incompatibility of two loss functions relevant to the decision problem at hand: one governing the decision maker's learning, and the other governing her decisions.

To shed light on the source of this problem, we construct a preference-based similarity measure and contrast it with the analogous KL-divergence D_{KL}, that captures statistical similarity. In particular, by normalizing $\mathsf{U}(b(Q), P)$ we define $D_{\mathsf{U}} \colon \mathcal{P} \times \mathcal{P} \to \mathbb{R}_+$ as

$$D_{\mathsf{U}}(P \parallel Q) := \mathsf{U}(b(P), P) - \mathsf{U}(b(Q), P). \tag{7}$$

For completeness, we denote the likelihoods that yield the highest long-run average payoff in $\mathcal{M}$ by

$$\mathcal{Q}^{\mathsf{U}}_{\mathcal{M},P} := \arg\min_{\theta \in \Theta} \; D_{\mathsf{U}}\left(P \parallel Q_\theta\right). \tag{8}$$

We can use this set to rephrase local payoff-optimality of $\mathcal{M}$ as $\mathcal{Q}^{\mathrm{KL}}_{\mathcal{M},P} \subseteq \mathcal{Q}^{\mathsf{U}}_{\mathcal{M},P}$ for all $P \in \mathcal{P}$.

For fixed P, both D_{KL} and D_{U} attain their global minima (zero) at $Q = P$. Nevertheless, as opposed to D_{KL}, D_{U} is *not* a divergence, because it can also take zero values at $Q \neq P$.[12] This suggests that in principle the two measures D_{KL} and D_{U} can induce quite different level curves over the space of likelihoods. In the following sections we further explore the implications of this difference.

In particular, using the two performance properties introduced in Definition 1 and 2, section 3 demonstrates through a sequence of examples, that under Bayesian learning an *arbitrary* set of likelihoods fails to satisfy both *local payoff-optimality* and *misspecification-proofness*. While *a priori* our properties might appear so strong that only correctly specified families of likelihoods could satisfy them, in section 4 we characterize a class of decision problems for which it is possible (and straightforward) to construct a set $\mathcal{M}$ with $P \notin \mathcal{M}$ that are both misspecification-proof and locally-payoff-optimal.

[12] In fact, this property is a key feature of some of our examples in section 3. Moreover, we will provide conditions under which certain likelihoods Q and the data generating process P can share all "payoff-relevant" moments ($D_{\mathsf{U}}(P \parallel Q) = 0$), but otherwise differ in terms of their "statistical" moments ($D_{\mathrm{KL}}(P \parallel Q) \neq 0$ and so $Q \neq P$).

3 Examples

Our first example illustrates that arbitrary sets of likelihoods coupled with Bayes learning might fail to satisfy local payoff-optimality. That is, unless the entertained set $\mathcal{M}$ of likelihoods is carefully tailored to our objective, Bayes' rule itself cannot guarantee the highest long-run average payoff that is attainable with $\mathcal{M}$. Consider an income fluctuation problem with a risk-averse agent who can borrow or lend at a constant rate of interest, r, subject to the period budget constraint (and the corresponding transversality condition),

$$c_t + w_t = (1+r)w_{t-1} + x_t,$$

where c_t denotes consumption, x_t is the agent's uncertain labor income, and w_t is her financial wealth at the end of period t. For simplicity, we impose $\beta(1+r) = 1$, where β is the agent's discount factor. In this example, the agent wants to learn about the unknown distribution of the exogenous random variable x_t. Suppose that the reference set $\mathcal{P}$ includes distributions asserting that every period labor income is drawn *i.i.d.* from some unknown distribution $P \in \mathcal{P}$.

Crucially, the agent's preferences determine which features of labor income process x_t are relevant to the savings problem. To illustrate this point, we derive best response functions for two well-known utility specifications,[13]

(i) quadratic preferences with $u(c) = -\frac{1}{2}(c-\bar{c})^2$ and $\bar{c} > 0$ that imply:

$$b_{\mathrm{q}}(Q) = rw + \mathbb{E}_Q[X]; \tag{9}$$

(ii) constant absolute risk-averse (CARA) preferences with $u(c) = -\eta^{-1}\exp(-\eta c)$ and $\eta > 0$ that imply:

$$b_{\mathrm{c}}(Q) = rw + \mathbb{E}_Q[X] - \frac{1}{r\eta}\log \mathbb{E}_Q\left[\exp\left(-\frac{r\eta}{1+r}X\right)\right]. \tag{10}$$

Evidently, in both cases, the specific distribution Q affects the agent's best action through the implied expected labor income. In fact, as the well-known *certainty equivalence* property of linear-quadratic control problems suggests, with quadratic preferences the first moment $m_{\mathrm{q}}(Q) := \mathbb{E}_Q[X]$ is the only feature of the income process that influences the agent's decisions and her log-run average payoffs.

On the other hand, with CARA preferences precautionary motives are also present, so the agent's optimal action depends on higher moments as well. Nonetheless, the effect of higher moments are conveniently summarized by the last term of (10), so similar to the quadratic case, we can define a

[13] See the seminal papers by Hall [1978] and Caballero [1990] for the quadratic and CARA specifications, respectively.

vector

$$m_{\mathrm{c}}(Q) := \left(\mathbb{E}_Q\left[X\right], \quad \mathbb{E}_Q\left[\exp\left(-\frac{r\eta}{1+r}X\right)\right]\right)'$$

that captures all payoff-relevant features of X. The fact that in both cases we are able to summarize relevant features of X with a vector of moments permits an intuitive characterization of our preference-based measure, D_{U}, expressing the reduction in long-run average payoff owing to using Q instead of P.

Lemma 1. *Let $i \in \{q, c\}$. If two distributions Q, Q' are such that $m_i(Q) = m_i(Q')$, they induce the same policy functions so that $b_i(Q) = b_i(Q')$ and $D_{\mathsf{U}}\left(P \parallel Q\right) = D_{\mathsf{U}}\left(P \parallel Q'\right)$.*

As for the entertained set of likelihoods, assume first that the agent uses likelihoods that describe X as being *i.i.d.* lognormal parameterized by $\theta = \left(\mu, \sigma^2\right)$ so that

$$\mathcal{M} = \left\{ \ \log\mathcal{N}\left(\mu, \sigma^2\right) \ : \ (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+ \ \right\}. \tag{11}$$

During the learning process, Q is a mixture model arising from mixing the distributions in $\mathcal{M}$ using posterior probabilities. However, since the agent is assumed to settle on a single likelihood asymptotically, for our purposes, it is sufficient to focus only on the individual parametric distributions of $\mathcal{M}$. In particular, Lemma 1 implies that if there is a $Q_\theta \in \mathcal{M}$ such that $m(Q_\theta) = m(P)$, then the agent is able to implement the action that is the best response with respect to P, thereby attaining the maximum possible long-run average payoff, $D_{\mathsf{U}}(P \parallel Q_\theta) = 0$, even if all of her likelihoods are wrong. Indeed, because the moments in m enter the policy functions in an additive manner, there are in principle many likelihoods in $\mathcal{M}$ that can satisfy this property.

When the payoff function is quadratic, the D_{U}-minimizing likelihoods are characterized by,

$$\mathcal{Q}^{\mathsf{U}_{\mathrm{q}}}_{\mathcal{M},P} = \left\{ \ \left(\mu, \sigma^2\right) \ : \ \exp\left(\mu + \frac{\sigma^2}{2}\right) = \mathbb{E}_P[X] \ \right\}. \tag{12}$$

As for CARA preferences, notice that the extra moment that the agent cares about is the Laplace transform of the perceived distribution of labor income evaluated at $\eta r/(1+r)$; we use $\mathcal{L}(Q)$ to denote this object. Although there is no closed form solution for the Laplace transform of the lognormal distribution, good approximations exist that we can use to compute $\mathcal{L}(Q)$ for the likelihoods in $\mathcal{M}$. For the sake of transparency, we use $\mathcal{L}(\mu, \sigma^2)$ to denote these values. That said, the D_{U}-minimizing likelihoods with CARA utility are

$$\mathcal{Q}^{\mathsf{U}_{\mathrm{c}}}_{\mathcal{M},P} = \left\{ \ \left(\mu, \sigma^2\right) \ : \ \exp\left(\mu + \frac{\sigma^2}{2}\right) - \frac{1}{\eta r}\log\mathcal{L}\left(\mu, \sigma^2\right) = \mathbb{E}_P[X] - \frac{1}{\eta r}\log\mathcal{L}(P) \ \right\}. \tag{13}$$

In principle, moments on the right hand side of (12) and (13) become known asymptotically irrespective of which $P \in \mathcal{P}$ generates the data. This does not mean, however, that these are the moments that Bayes' law focuses on to find the likelihood in $\mathcal{M}$ that minimizes the KL-divergence from P. Using the definition of KL-divergence, this amounts to minimizing $\mathbb{E}_P[-\log q_\theta]$, where q_θ denotes the density of Q_θ with respect to Lebesque measure. Given that the entertained likelihoods in $\mathcal{M}$ are lognormals we obtain,

$$\mathcal{Q}^{\mathrm{KL}}_{\mathcal{M},P} = \left\{ \ \left(\mu, \sigma^2\right) \ : \ \mu = \mathbb{E}_P\left[\ln X\right], \quad \sigma^2 = \mathbb{E}_P\left[\,(\ln X)^2\,\right] - \mathbb{E}_P\left[\ln X\right]^2 \ \right\}. \tag{14}$$

Comparing (12) and (13) with (14), one can immediately see a misalignment: while a key component of the agent's policy function is the mean of X, Bayesian learning with lognormal likelihoods aims to match the mean of $\ln X$. In other words, with both utility functions, the vector m of relevant moments is incompatible with the set $\mathcal{M}$ that tells KL-divergence which moments to match. An immediate consequence of this misalignment is that $\mathcal{M}$ in (11) is neither locally-payoff-optimal, nor misspecifiction-proof.

Although the specific P is irrelevant for this conclusion, further insight can be gained by looking at a particular example with $P \notin \mathcal{M}$. Suppose that P is such that $\ln X$ is distributed as a two-component mixture normal distribution,

$$\ln X \overset{\text{iid}}{\sim} \lambda \cdot \mathcal{N}(\mu_1, \sigma_1^2) + (1-\lambda) \cdot \mathcal{N}(\mu_2, \sigma_2^2). \tag{15}$$

Hence the agent's long-run behavior is determined by the likelihood $Q^{\mathrm{KL}}_{\mathcal{M},P} \in \mathcal{M}$ with parameters:

$$\mu_{\mathrm{KL}} = \lambda\mu_1 + (1-\lambda)\mu_2, \qquad\qquad \sigma^2_{\mathrm{KL}} = \lambda(\sigma_1^2 + \mu_1^2) + (1-\lambda)(\sigma_2^2 + \mu_2^2) - \mu^2_{\mathrm{KL}}.$$

Notice that for a general quintuple $(\lambda, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$, $m_{\mathrm{i}}\left(Q^{\mathsf{U}_{\mathrm{i}}}_{\mathcal{M},P}\right) \neq m_{\mathrm{i}}(P)$ for $i \in \{\mathrm{q}, \mathrm{c}\}$, so that the asymptotically implemented policy function is suboptimal relative to those induced by likelihoods in $\mathcal{Q}^{\mathsf{U}_{\mathrm{q}}}_{\mathcal{M},P}$ or $\mathcal{Q}^{\mathsf{U}_{\mathrm{c}}}_{\mathcal{M},P}$, that is, $\mathcal{M}$ is not locally-payoff-optimal. The extent to which this suboptimality matters depends on the utility function, here captured by the risk aversion parameter, η, and the agent's environment, here captured by the interest rate r.

For simplicity, we focus on CARA preferences, but the case of quadratic utility is very similar. The left panel of Figure 1 depicts densities of the data generating process along with two misspecified likelihoods that are closest to it according to D_{KL} and D_{U}. While the blue distribution, used by the Bayesian learner for making decisions, matches statistical aspects of the data generating process better, the green dashed likelihood induces higher long-run average payoff given that it lies in $\mathcal{Q}^{\mathsf{U}_{\mathrm{c}}}_{\mathcal{M},P}$. In fact, the value generated by the green likelihood equals that implied by the black distribution P.

To shed more light on the reason why $\mathcal{M}$ in (11) is neither locally-payoff-optimal, nor misspecification-proof, the right panel of Figure 1 depicts level curves corresponding to the projections of D_{KL} and D_{U} on $\mathcal{M}$. Distributions on the ellipses have equal KL-divergence relative to P. The farther an ellipsis is from $\theta_{\mathrm{KL}} \coloneqq \left(\mu_{\mathrm{KL}}, \sigma^2_{\mathrm{KL}}\right)$ the higher is the corresponding KL-divergence. On the other hand, the level curves of D_{U} exhibit a strikingly different geometry. The red line represents $\mathcal{Q}^{\mathsf{U}_{\mathrm{c}}}_{\mathcal{M},P}$, i.e., likelihoods in $\mathcal{M}$ that perform identically to P in terms of induced long-run average payoff. Similarly, the blue lines correspond to likelihoods with equal long-run average payoff. The farther we are from the red line,

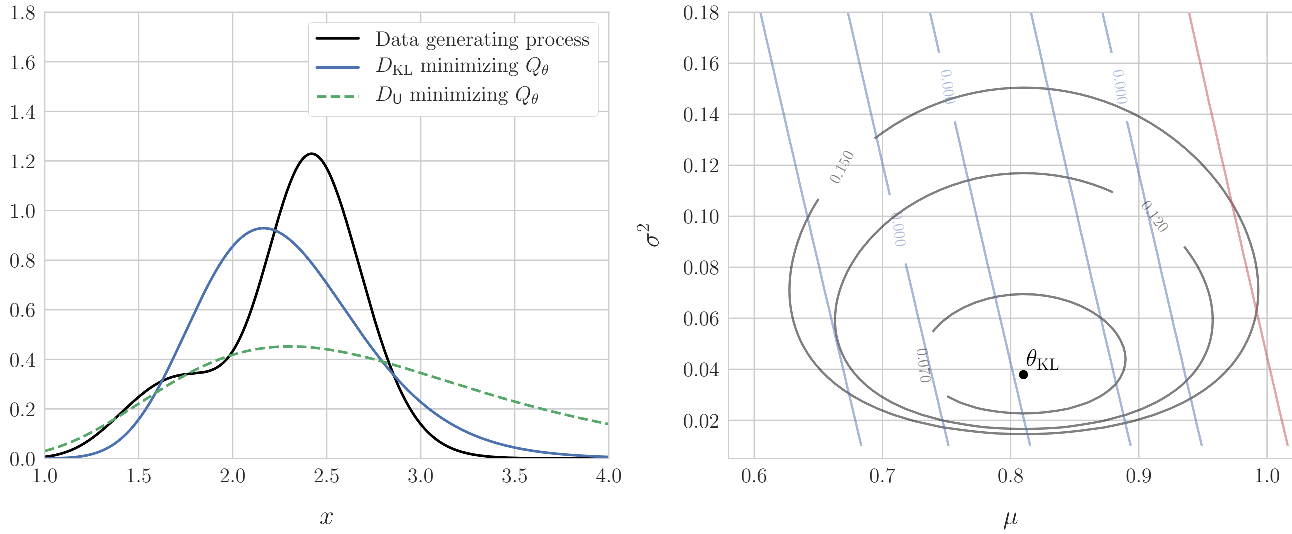

Figure 1: Left: Densities of the data generating process and best approximations within $\mathcal{M}$ according to D_{KL} and D_{U}. Right: Indifference curves with CARA preferences and the information geometry over $\mathcal{M}$. Note: The parameters are $(\lambda, \mu_1, \sigma_1, \mu_2, \sigma_2) = (0.3, 0.6, 0.2, 0.9, 0.1)$, $\beta = 0.98$, and $\eta = 2.5$. This implies that under P the mean of X is 2.29 and the standard deviation is 0.41.

the larger the loss from not using P.

The difference between the two geometries emerges from the properties of D_{KL} and D_{U}. While the level curves of D_{U} are influenced by the preference parameters, (η, β), the iso-entropies of D_{KL} depend primarily on the learner's set of models $\mathcal{M}$. As we saw, lognormal distributions imply that Bayes learning focuses on the moments $\mathbb{E}_P\left[\ln X\right]$ and $\mathbb{E}_P\left[(\ln X)^2\right]$, thereby directing the agent's attention to irrelevant features of the data generating process.

An alternative $\mathcal{M}$

Our second example illustrates that for some decision problems, finding a misspecification-proof set $\mathcal{M}$ is feasible even if $P \notin \mathcal{M}$. That is, while misspecification-proofness is a demanding property, it is not true that only $\mathcal{P}$ can satisfy it. To this end, we replace the agent's set in (11) with an admittedly more common "default prior". Suppose that $\mathcal{M}$ contains likelihoods that describe X as if it was drawn *i.i.d.* from a normal distribution parameterized by $\theta = (\mu, \sigma^2)$,

$$\mathcal{M} = \left\{ \mathcal{N}\left(\mu, \sigma^2\right) \ : \ (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+ \right\}. \tag{16}$$

For simplicity, let P be the same as above. The key difference from the previous analysis is the way KL-divergence determines $Q^{\text{KL}}_{\mathcal{M},P}$. With Gaussian likelihoods, D_{KL} targets the mean and variance of X:

$$\mathcal{Q}^{\text{KL}}_{\mathcal{M},P} = \left\{ \left(\mu, \sigma^2\right) \ : \ \mu = \mathbb{E}_P\left[X\right], \quad \sigma^2 = \mathbb{E}_P\left[X^2\right] - \mathbb{E}_P\left[X\right]^2 \right\}. \tag{17}$$

These targets are consistent with quadratic preferences, in the sense that $m_{\mathrm{q}}(Q^{\mathrm{KL}}_{\mathcal{M},P}) = m_{\mathrm{q}}(P)$. That is, with quadratic utility the Gaussian set $\mathcal{M}$ is misspecification-proof and locally-payoff-optimal, because KL-divergence focuses on a moment that is relevant to the agent's objective. In fact, by learning about both the mean and the variance of X, the agent is wasting resources: σ^2 is irrelevant to her long-run payoffs, so it appears inefficient to include models with different variances in $\mathcal{M}$.

The situation is different with CARA utility. In this case, the best-responding policy is a function of $\mathcal{L}(P)$ that depends on all higher order moments of P not just the mean and variance, so it cannot be guaranteed that $m_{\mathrm{q}}(Q^{\mathrm{KL}}_{\mathcal{M},P}) = m_{\mathrm{q}}(P)$. In other words, while with quadratic preferences Gaussian likelihoods constitute a misspecification-proof set, the same does not hold if preferences are CARA.

4 Misspecification-proof learning

The central message of our examples is that if the statistical and payoff-relevant aspects of $\mathcal{M}$ are not aligned, learning with misspecified likelihoods can lead to suboptimal policies. As demonstrated by the negative examples in section 3, the problem with the DM's set is that KL divergence matches features of the environment that are irrelevant to the optimal policy function, a^*. In this sense, Bayesian learning is focusing on wrong features of the data generating process.

Inspired by the positive example in section 3, we show now that under certain conditions, constructing sets of likelihoods that are both *misspecification-proof* and *locally-payoff-optimal* is feasible. As we saw before, Bayes rule designates KL-divergence as an implicit loss function for learning. While taking this loss function as given, we can choose $\mathcal{M}$ so that D_{KL} focuses on pertinent features of the environment. To this end, the selection of entertained likelihoods should be based on the payoff-relevant moments of the data generating process. This insight can be generalized and serve as a guideline for specifying $\mathcal{M}$ in situations in which misspecification is an issue.

We give conditions on $\mathcal{M}$ such that the KL divergence minimizing model induces the optimal policy function even if the entertained set of likelihoods is misspecified.

Assumption 1 (Moment-dependent policy function)**.**
For a given utility function, u, suppose that the optimal policy function (relative to $\mathcal{P}$) can be expressed as a function of finitely many moments of the data generating process. That is, for any $Q \in \mathcal{P}$, the best-response policy function, $\hat{a} = b\,(Q)$, can be defined implicitly as

$$G_u\left(\hat{a}, \mathbb{E}_Q\left[T_u(X)\right]\right) = \mathbf{0}, \tag{18}$$

where $T_u \colon \mathcal{X} \to \mathbb{R}^d$ defines moments of Q, and $G_u \colon \mathcal{A} \times \mathbb{R}^d \to \mathbb{R}^k$ specifies the dependence of the optimal policy function on these moments. Note that both of these functions can depend on the DM's preferences, u.

Assumption 1 implies that the data generating process does not enter the optimal policy function other than through the moments, $m(P) = \mathbb{E}_P\left[T_u(X)\right]$. As a result, T_u defines the sufficient statistic that is needed to implement the optimal policy function.

Proposition 1.
If Assumption 1 is satisfied, the exponential family of likelihoods defined by the finite sufficient statistics T_u[14]

$$\mathcal{M} = \left\{ q_\theta(x) = h(x) \exp\left\{\theta \cdot T_u(x) - A(\theta)\right\} : \theta \in \Theta \subseteq \mathbb{R}^d \right\}, \tag{19}$$

for some $h \colon \mathcal{X} \to \mathbb{R}_+$*, where* $A(\theta)$ *is the cumulant function, is both misspecification-proof and locally-payoff-optimal.*

Proposition 1 follows from the tractable relationship between exponential families and their finite sufficient statistics.[15] The KL-divergence minimizing distribution within $\mathcal{M}$ is characterized by,

$$\mathbb{E}_{Q^{\mathrm{KL}}_{\mathcal{M},P}}[T_u(X)] = \mathbb{E}_P[T_u(X)]. \tag{20}$$

As a result, the Bayesian learner in the limit implements the policy function $\hat{a}_{\mathrm{KL}} = b\left(Q^{\mathrm{KL}}_{\mathcal{M},P}\right)$ that satisfies

$$G_u\left(\hat{a}_{\mathrm{KL}}, \mathbb{E}_{Q^{\mathrm{KL}}_{\mathcal{M},P}}[T_u(X)]\right) = \mathbf{0}. \tag{21}$$

Given equation (20) the KL-divergence minimizing likelihood exactly matches the payoff-relevant moments of the data generating process, and by Assumption 1 the optimal policy function is implemented,

$$G_u\left(\hat{a}_{\mathrm{KL}}, \mathbb{E}_P[T_u(X)]\right) = \mathbf{0}. \tag{22}$$

The moral of Proposition 1 is as follows. Even if the environment described by the underlying data generating process is complex in the statistical sense—i.e. infinite dimensional—the relevant complexity of the learner's problem is defined through her objective and reflects the properties of u. If the purpose of learning is to aide decisions, the learner should select her likelihoods $Q \in \mathcal{M}$ based on their ability to capture features of the environment that matter for good decisions. In this sense, we treat the DM's set $\mathcal{M}$ as an instrument for making decisions rather than an indisputable part of preferences. Given that the environment's statistical complexity can seriously limit the ability to learn, we advocate calibrating $\mathcal{M}$ not to the intricacies of the "true" environment but to features pertinent to making good decisions.

Example revisited

To see how to apply the above concepts, we revisit the income fluctuation problem of section 3. Recall that we identified a vector of payoff-relevant moments, $m_i(Q)$, for both the quadratic ($i = \mathrm{q}$)

[14] For ease of notation we define the exponential family through density functions. q_θ is the density of Q_θ with respect to the Lebesgue measure. We use the canonical parametrization.

[15] See any standard reference on mathematical statistics, e.g. Shao [2003].

and CARA ($i = \mathrm{c}$) utility specifications. In terms of the introduced notation, these vectors can be rewritten as,

$$T_{\mathrm{q}}(x) \doteq x \qquad \Rightarrow \qquad m_{\mathrm{q}}(Q) = \mathbb{E}_Q\left[T_{\mathrm{q}}(X)\right];$$
$$T_{\mathrm{c}}(x) \doteq \left(x,\ \exp\left(-\frac{r\eta}{1+r}x\right)\right)' \qquad \Rightarrow \qquad m_{\mathrm{c}}(Q) = \mathbb{E}_Q\left[T_{\mathrm{c}}(X)\right].$$

Evidently, in both cases, Assumption 1 is satisfied, so misspecification-proof Bayes learning is feasible. With quadratic utility our recommended set of (misspecified) likelihoods is

$$\mathcal{M}_{\mathrm{q}} = \Big\{\exp\left(\theta \cdot x - A_{\mathrm{q}}(\theta)\right) \ : \ \theta \in \Theta \subseteq \mathbb{R}\Big\},$$

which is consistent with our previous finding that Gaussian distributions with unknown mean and known variance can lead to the implementation of a^* in the long-run.

Regarding CARA utility, we specify the following two-parameter family

$$\mathcal{M}_{\mathrm{c}} = \left\{\exp\left(\theta_1 \cdot x + \theta_2 \cdot \exp\left(-\frac{\eta r}{1+r} \cdot x\right) - A_{\mathrm{c}}(\theta_1, \theta_2)\right) \ : \ (\theta_1, \theta_2) \in \Theta \subseteq \mathbb{R}^2\right\}$$

that renders the implied information geometry aligned with the utility geometry.

5 Related literature

Two key features of our analysis are: (i) we take an outside observer's perspective by considering standard Bayesian decision making but assessing the set of likelihoods entertained by the DM according to their long-run implications under some "objective reality", and (ii) we allow the set of likelihoods (prior support) to be misspecified.

While Bayesian learning occupies a prominent place in the economics literature, most papers focus on the case of correct specification. In their classic paper, Bray and Kreps [1987] argue that this benchmark is too "sterile" and call for models that "have in place some level of inconsistency with reality". Important examples of such models are Nyarko [1991] and Fudenberg et al. [2017].[16] Similar to us, these papers analyze Bayesian learning under the assumption that the decision maker's prior is misspecified. Nevertheless, instead of assessing the usefulness of these priors as we do, these papers focus on the non-trivial dynamics of beliefs that may arise when learning is active.

The econometrics and statistics literature paid relatively more attention to the idea of misspecification. In his seminal paper, Berk [1966] showed that when a Bayesian is learning about a parameter from a series of exchangeable signals, asymptotically, her posterior concentrates on the parameter values for which the KL-divergence of the DGP with respect to the entertained likelihoods is minimal. More

[16] See also Esponda and Pouzo [2016].

recently, Shalizi [2009] arrives at the same conclusion in a much more general setting. Following the frequentist tradition, White [1996] provides a thorough analysis of maximum-likelihood techniques when the model is misspecified.[17] In this case, the KL-divergence minimizing parameter, θ_{KL}, is typically called the *pseudo-true parameter*. By studying large sample properties of Bayesian inference about θ_{KL}, Müller [2013] shows that one can reduce the Bayes estimator's expected loss (under the DGP) by replacing the original posterior with an artificial normal posterior centered at θ_{KL} with the sandwich covariance matrix. Although similar, our approach is different in the sense that our decision maker is not interested in P or θ_{KL} *per se*. Statistical closeness is important for her only to the extent that it helps to make better decisions.

As for point (i), an example is Blume et al. [2018] who use an objective welfare criterion—similar in spirit to ours—to rank alternative market structures in the presence of belief heterogeneity (without learning). In addition, measuring the implications of learning relative to the DGP is of a similar flavor to the question of survival in financial markets analyzed by Blume and Easley [2006]. We illustrate that a learner's value function can carry invaluable information about the relative usefulness of different likelihoods when the prior is misspecified. This echoes the literature on max-min expected utility that breaks a key feature of Bayesian decision making: the separation of inference and control.[18] This literature—in contrast to the Bayesian decision rule we use—alters the manner in which optimal policies are chosen: instead of trying to maximize (5) under a single distribution, a max-min decision maker seeks policies that work well (not necessarily optimally) under a whole set of reasonable distributions. Attempts to marry such behavior with learning can be found in Hansen and Sargent [2007], Klibanoff et al. [2009], and Epstein and Schneider [2007].

Our recommendation in section 4 also resembles the idea of *Gibbs posteriors* advocated by Jiang and Tanner [2008] and Bissiri et al. [2016]. Instead of trying to model the DGP directly, this approach starts with some statistics of interest, θ, accompanied with a corresponding loss function, $\ell(\theta, x)$, such that θ minimizes the expected $\ell(\theta, x)$ under the DGP. It then proposes to use $\exp\left(-\ell(\theta, x)\right)$ as a likelihood for Bayesian inference. In our case, the statistics θ can be viewed as our vector of payoff-relevant moments, $m(Q)$. In this sense, the main difference relative to our analysis is that we do not take these moments as given but derive them from primitives (preferences and market structure) of an economic decision problem. Similar comments apply to the so called *focused information criterion* developed by Claeskens and Hjort [2003]. It is a model selection tool that evaluates candidate models based on their ability to efficiently estimate a particular parameter of interest, instead of comparing their overall fit.

[17] Ideas similar to the justification that we give in section 4 can be spotted in various chapters of White [1996].

[18] In the Bayesian model, optimal inference about P is independent of the utility function u. See Hansen and Sargent [2018].

6 Concluding remarks

This paper shows that in a setting in which misspecification is a major concern, Bayesian learning with arbitrary likelihoods can lead to outcomes that appear irrational from an objective point of view. Importantly, we do not mean this as a critique of Bayesian decision making. Instead, our result is meant to shed light on the advantages of viewing the decision maker's likelihoods as instruments rather than part of her preferences. In a truly unknown environment, entertaining a set that is inconsistent with the agent's payoff function is "irrational" in the sense that the decision maker would feel regret and change her mind if she were told the potential consequences of her behavior.[19] Assuming that beliefs are of the same nature as attitudes toward risk or the rate of time preference makes this irrationality an unchallangeable axiom of behavior that we find unreasonable. In our view, there are such things as "good" and "bad" beliefs, just like there are "good" and "bad" models. Having said that, it seems sensible to impose consistency among the agent's beliefs and preferences even if learning is correctly specified.

We see at least two avenues for future research. First, we assumed that the decision maker fully embraces her likelihoods and uses them to derive policy functions as if they were subjective beliefs. In other words, our decision maker does not explicitly acknowledge her misspecification concerns. In contrast, a growing literature pioneered by Hansen and Sargent [2008] endows agents inside economic models with the same kind of misspecification concerns as we, econometricians, face. It would be interesting to see how our results change in the presence of ambiguity aversion, that is, when policy functions are chosen in a max-min fashion. Hansen and Marinacci [2016] discuss related results and challenges. Second, for the sake of clarity, we focused on single agent problems, but it seems feasible to extend some of our insights to more general market settings as well. We leave for future research the analysis of general equilibrium with multiple agents who are learning about a common exogenous stochastic process.

[19]This definition of irrationality is motivated by Gilboa and Schmeidler [2001] and Gilboa [2009].

References

Robert H. Berk. Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist.*, 37(1):51–58, 02 1966.

P. G. Bissiri, C. C. Holmes, and S. G. Walker. A general framework for updating belief distributions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(5):1103–1130, 2016.

Lawrence E. Blume and David Easley. If you're so smart, why aren't you rich? belief selection in complete and incomplete markets. *Econometrica*, 74(4):929–966, 2006.

Lawrence E. Blume, Timothy Cogley, David A. Easley, Thomas J. Sargent, and Viktor Tsyrennikov. A case for incomplete markets. *Journal of Economic Theory*, 178:191 – 221, 2018. ISSN 0022-0531.

Margaret Bray and David Kreps. Rational learning and rational expectations. In G. R. Feiwel, editor, *Arrow and the Ascent of Modern Economic Theory*, pages 597–625. Palgrave Macmillan, London, 01 1987. ISBN 978-1-349-07241-5.

Ricardo J. Caballero. Consumption puzzles and precautionary savings. *Journal of Monetary Economics*, 25(1):113–136, January 1990.

Gerda Claeskens and Nils Lid Hjort. The focused information criterion. *Journal of the American Statistical Association*, 98(464):900–945, 2003.

Persi Diaconis and David Freedman. On the consistency of bayes estimates. *The Annals of Statistics*, pages 1–26, 1986.

Larry G. Epstein and Martin Schneider. Learning under ambiguity. *The Review of Economic Studies*, 74(4):1275–1303, 2007.

Ignacio Esponda and Demian Pouzo. Berk–nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica*, 84(3):1093–1130, 2016.

Drew Fudenberg, Gleb Romanyuk, and Philipp Strack. Active learning with a misspecified prior. *Theoretical Economics*, 12(3):1155–1189, 2017.

Itzhak Gilboa. *Theory of Decision under Uncertainty.* Econometric Society Monographs. Cambridge University Press, 2009. doi: 10.1017/CBO9780511840203.

Itzhak Gilboa and David Schmeidler. *A theory of case-based decisions.* Cambridge University Press, 2001.

Robert E. Hall. Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, 86(6):971–987, 1978.

Lars Peter Hansen and Massimo Marinacci. Ambiguity aversion and model misspecification: An economic perspective. *Statist. Sci.*, 31(4):511–515, 11 2016.

Lars Peter Hansen and Thomas J. Sargent. Recursive robust estimation and control without commitment. *Journal of Economic Theory*, 136(1):1 – 27, 2007.

Lars Peter Hansen and Thomas J. Sargent. *Robustness*. Princeton University Press, 2008.

Lars Peter Hansen and Thomas J. Sargent. *Risk, Uncertainty, and Value*. Princeton, New Jersey: Princeton University Press., 2018.

Wenxin Jiang and Martin A. Tanner. Gibbs posterior for variable selection in high-dimensional classification and data mining. *Ann. Statist.*, 36(5):2207–2231, 10 2008.

Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji. Recursive smooth ambiguity preferences. *Journal of Economic Theory*, 144(3):930 – 976, 2009.

David M. Kreps. Anticipated utility and dynamic choice. In E. Kalai D.P. Jacobs and M. Kamien, editors, *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures*, pages 242–274. Cambridge University Press, Cambridge, England, 1998.

Ulrich K. Müller. Risk of bayesian inference in misspecified models, and the sandwich covariance matrix. *Econometrica*, 81(5):1805–1849, 2013.

Yaw Nyarko. Learning in mis-specified models and the possibility of cycles. *Journal of Economic Theory*, 55(2):416 – 427, 1991.

Cosma Rohilla Shalizi. Dynamics of Bayesian updating with dependent data and misspecified models. *Electronic Journal of Statistics*, 3:1039–1074, 2009.

Jun Shao. *Mathematical Statistics*. Springer-Verlag New York Inc, 2nd edition, 2003.

Vladimir Vapnik. *The nature of statistical learning theory*. Springer science & business media, 1995.

Abraham Wald. *Statistical Decision Functions*. Wiley: New York, 1950.

Halbert White. *Estimation, Inference and Specification Analysis*. Number 9780521574464 in Cambridge Books. Cambridge University Press, December 1996.