Learning with Misspecified Models*

Dániel Csaba† Bálint Szőke‡

July 17, 2023

Abstract

We consider learning with a set of likelihoods when the learner's set is misspecified. We study welfare implications of entertaining a misspecified set by focusing on the limit point of learning and the associated best-responding policy. Building on such policies, we define consistency requirements for sets of likelihoods that a utility-maximizing agent would find desirable. We characterize a class of decision problems for which *exponential families* of likelihoods—with payoff-relevant moments as sufficient statistics—exist that satisfy our consistency requirements therefore guaranteeing the asymptotic implementation of optimal policies irrespective of the data generating process.

JEL classification: C51, D83

Keywords: Learning; Exponential Family; Misspecified models; Statistical Decision Theory

*We thank Timothy Christensen, Toru Kitagawa, Filippo Massari, Francesca Molinari, Jonathan Payne, and Thomas Sargent for insightful discussions and comments.

†QuantCo Inc., E-mail: csaba.daniel@gmail.com.

‡Federal Reserve Board, Division of Monetary Affairs. Email: balint.szoke@frb.gov

1 Introduction

Most economic models aim to capture a simplified and crude approximation of the complex environment they wish to describe. The models are misspecified. Economic agents use these models to interpret observed signals and make decisions given those interpretations. While parsimonious models might have clear benefits for learning in finite samples, when data are abundant it is not obvious if and when using misspecified models for learning has a detrimental impact on the quality of decisions.

To investigate this problem we consider a decision maker (DM) learning about the distribution P of an exogenous stochastic process with the aim of uncovering features of the data generating process (DGP) that are instrumental for her decision problem. While we adopt a Bayesian framework for simplicity our results apply to other likelihood based methods as well. Bayesian *learning* entails specifying a 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 consider parametric families of $\mathcal{M}$ while allowing the DGP to be nonparametric.

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

[1]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.

respect to Q.[2] We use this measure to define an undominatedness property of $\mathcal{M}$, that we call *misspecification-proofness*, requiring that, irrespective of which $P \in \mathcal{P}$ generates the data, there is no other set of likelihoods that 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 models. We also introduce a weaker performance property of $\mathcal{M}$, that we call *local payoff-optimality*, requiring that, irrespective of which $P \in \mathcal{P}$ generates the data, the likelihood 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 generic sets of likelihoods do not satisfy either of our properties under Bayesian learning. This means that a misspecified set $\mathcal{M}$ can cause the 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 implicitly define the statistical moments of P that Bayesian learning will focus on. If those moments are 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

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

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

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 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 introduced 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. As such, while we use 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 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 given by a preference-driven coarsening of the state space.

[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.

1.1 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).[5] 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 problem of misspecification has a rich history in the econometrics and statistics literature. In his seminal paper, Berk (1966) showed that Bayesian learning about a parameter from a series of exchangeable signals asymptotically concentrates on the parameter values for which the KL-divergence of the DGP with respect to the entertained likelihoods is minimal. More 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.[6] 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

[5] See also Esponda and Pouzo (2016).

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

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.[7] 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 expected utility 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).

The observation that misaligned loss functions—one for estimation, and one for evaluation—can lead to undesirable outcomes under misspecification has long been recognized in the forecasting literature. In particular, Granger and Newbold (1973) argued that if misspecification is a serious concern and one believes that a particular loss function (e.g. mean squared error) should be used to evaluate forecasts, then the same loss function should also be used to estimate the model parameters.[8] Our recommendation in section 4 is similar in spirit and can be viewed as a generalization and a tractable operationalization of this idea.

Our recommendation 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 $\exp(-\ell(\theta, x))$ as a quasi-likelihood for Bayesian inference. In our case, the statistics θ can be viewed as our vector of payoff-relevant moments. 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*

[7] In the Bayesian model, optimal inference about P is independent of u. See Hansen and Sargent (2018).
[8] See also the loss-function based proposals of Schorfheide (2001) and Geweke and Amisano (2012).

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.

The rest of the paper is structured as follows. Section 2 introduces our consistency requirements for misspecification-proof learning. In Section 3, we demonstrate how misspecification can lead to suboptimal long-run behavior through an example. Section 4 presents our recommendation that can be used to resolve this suboptimality in some cases. Section 5 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. A typical case would be when $\mathcal{P}$ imposes exchangeability but no further restrictions.

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 measurable policy functions that the DM could implement. 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, 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.[9] To distinguish this notion from 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,[10]

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

In other words, after observing an infinite sequence of signals, Bayes' rule concentrates on a distribution from $\mathcal{Q}^{\text{KL}}_{\mathcal{M},P}$ as the best approximation of P. For ease of notation we consider situations in which the set defined in (4) is a singleton and denote the single KL divergence minimizer as $Q^{\text{KL}}_{\mathcal{M},P}$.

[9]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.

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

Our results can be readily extended to the general case by considering all elements of $\mathcal{Q}^{\text{KL}}_{\mathcal{M},P}$. 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 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[11,12]

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

Our 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.[13] 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 $P \in \mathcal{P}$.

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

[12] 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. Our results can be readily extended to the set-valued case.

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

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.[14] 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 $\mathcal{M}$:[15]

Definition 2 (Locally-payoff-optimal $\mathcal{M}$)**.**
Consider a decision problem characterized by the triplet $(\mathcal{P}, \mathcal{A}, u)$. The family of likelihoods $\mathcal{M} \subseteq \mathcal{P}$ 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 $\mathcal{M}$. Essentially, this amounts to asymptotically vanishing regret relative to the set of policies that can be derived as a best-response to some likelihood in $\mathcal{M}$.

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.

[14] 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.

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

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

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

We can use this set to rephrase local payoff-optimality of $\mathcal{M}$ as $\mathcal{Q}^{\text{KL}}_{\mathcal{M},P} \subseteq \mathcal{Q}^{\mathsf{U}}_{\mathcal{M},P}$ 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 $Q \neq P$.[16] 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 presents an example that demonstrates that under Bayesian learning generic sets of likelihoods fail 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.

3 Illustrative example

The following example illustrates that arbitrary sets of likelihoods coupled with Bayes learning might fail to satisfy both local payoff-optimality and misspecification-proofness. Consider a

[16] In fact, this property is a key feature of our example 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 ($D_{\text{KL}}(P \parallel Q) \neq 0$ and so $Q \neq P$).

consumption-saving problem with a risk-averse agent whose preferences are recursive of the Epstein and Zin (1989) type. Let $\gamma > 0$ be a coefficient of relative risk aversion for atemporal wealth gambles, $\psi > 0$ be a parameter that governs attitude toward substituting goods over time, and β be the agent's discount factor. At the beginning of every period, she observes realization x of the stochastic gross return X on her financial wealth. Suppose that the reference set $\mathcal{P}$ includes distributions asserting that every period X is drawn *i.i.d.* from some unknown distribution $P \in \mathcal{P}$.

The best response function is obtained by solving the Bellman equation for any distribution $Q \in \mathcal{P}$,

$$V(w) = \sup_{c} \ (1-\beta)\frac{c^{1-\psi}}{1-\psi} + \beta \mathbb{E}_Q \left[V(w')^{\frac{1-\gamma}{1-\psi}} \right]^{\frac{1-\psi}{1-\gamma}} \tag{9}$$

$$\text{s.t.} \quad w' = x(w-c), \tag{10}$$

where V is the agent's value function, c denotes her consumption, and w' is her financial wealth *after* realizing next period return x. For a given Q, the above functional equation can be solved for the optimal consumption policy by using standard recursive techniques:[17]

$$b(Q) = w \left[1 - \left(\beta \mathbb{E}_Q \left[X^{1-\gamma} \right]^{\frac{1-\psi}{1-\gamma}} \right)^{\frac{1}{\psi}} \right]. \tag{11}$$

The specific distribution Q affects the best action only through the implied risk-adjusted expected return $m(Q) := \mathbb{E}_Q \left[X^{1-\gamma} \right]$; a particular "perceived moment" of the interest rate process. The fact that we are able to summarize relevant features of the interest rate process with finite moments, m, 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. *If two distributions Q, Q' are such that $m(Q) = m(Q')$, they induce the same policy functions so that $b(Q) = b(Q')$ and $D_{\mathsf{U}}(P \parallel Q) = D_{\mathsf{U}}(P \parallel Q')$.*

As for the entertained set of likelihoods, assume that the agent uses likelihoods that describe X as being *i.i.d.* lognormal parameterized by $\theta = (\mu, \sigma^2)$ denoted by $\mathcal{M}$. 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

[17] Interior solutions require $\beta \mathbb{E}_Q \left[X^{1-\gamma} \right]^{\frac{1-\psi}{1-\gamma}} < 1$, which is a natural restriction on the discounted risk-adjusted return.

best response with respect to P, i.e., $D_{\mathsf{U}}(P \parallel Q_\theta) = 0$, even if all of her likelihoods are wrong. In fact, there are many lognormal distributions with this property.

In principle, the payoff-relevant moment $\mathbb{E}_P[X^{1-\gamma}]$ becomes known asymptotically irrespective of which $P \in \mathcal{P}$ generates the data. This does not mean, however, that this is the moment that Bayes' law focuses on to find the likelihood in $\mathcal{M}$ that minimizes the KL-divergence from P. In fact, asymptotically Bayes learning is converging to θ minimizing $\mathbb{E}_P[-\log q_\theta]$, where q_θ denotes the density of Q_θ with respect to the Lebesgue measure. Given that the entertained likelihoods in $\mathcal{M}$ are lognormals the minimizer matches the first- and second moments of the log-transformed state variable.

One can immediately see a misalignment: while a key component of the agent's policy function is the risk-adjusted expected return, Bayesian learning with lognormal likelihoods aims to match the mean and variance of $\log X$. An immediate consequence of this misalignment is that $\mathcal{M}$ 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 $\log X$ is distributed as a two-component mixture normal distribution.

The left panel of Figure 1 depicts the density of the assumed data generating process (black solid line) with a long left tail capturing rare but disastrous return realizations. The other densities represent lognormal densities that are closest to this data generating process according to D_{KL} and D_{U}. While the blue dash-dotted distribution, used by the decision maker, matches statistical aspects of the black solid line better, the green dashed densities induce higher long-run average payoff, because they are associated with the same risk-adjusted expected return as the black distribution P.

To shed more light on the suboptimality of a lognormal $\mathcal{M}$, the right panel of Figure 1 depicts two sets of level curves corresponding to the projections of D_{KL} and D_{U} on $\mathcal{M}$, respectively. Lognormal distributions on the ellipses have equal KL-divergence relative to P, while the blue straight lines show distributions with equal long-run average payoff. Clearly, the two sets of level curves exhibit strikingly different geometries. This difference emerges from the properties of D_{KL} and D_{U}. While the level curves of D_{U} are influenced by the preference parameter γ, the iso-entropies of D_{KL} depend

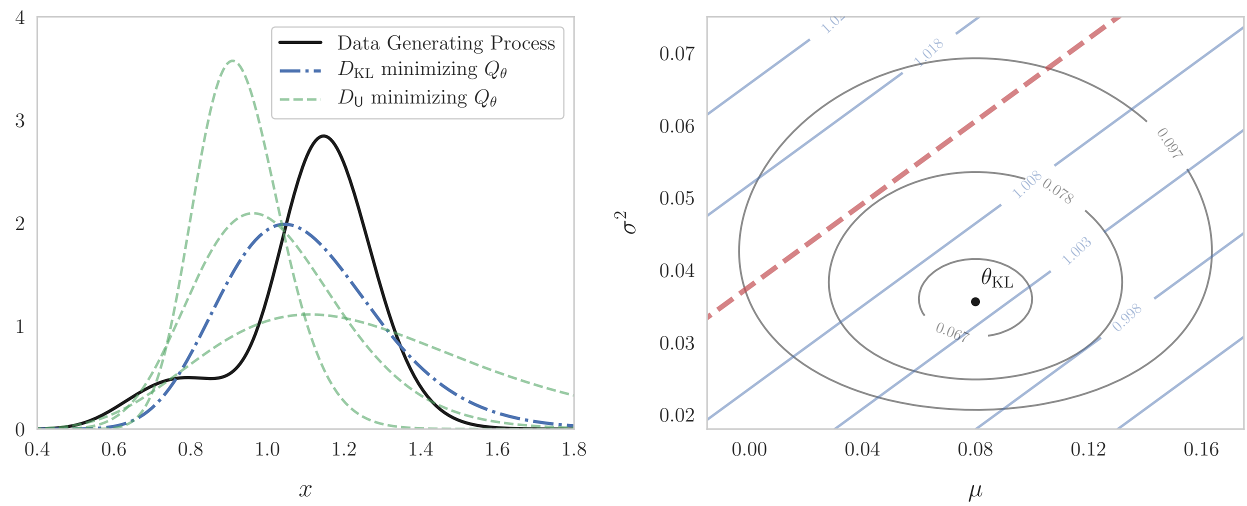

Figure 1: **Left:** Densities of the data generating process along with best approximations within $\mathcal{M}$ according to D_{KL} and D_{U}. Dashed lines show lognormal densities that imply risk-adjusted expected return of $\mathbb{E}_P[X^{1-\gamma}]$. **Right:** Indifference curves over $\mathcal{M}$ according to D_{KL} and D_{U}. The dot $\theta_{\text{KL}} := \left(\mu_{\text{KL}}, \sigma^2_{\text{KL}}\right)$ denotes the KL-divergence minimizing lognormal distribution. The dashed line represents distributions Q with $D_{\mathsf{U}}(P \parallel Q) = 0$.

primarily on the set $\mathcal{M}$.

4 Misspecification-proof learning

The central message of our example is that if the statistical and payoff-relevant aspects of $\mathcal{M}$ are not aligned, learning with misspecified likelihoods can lead to suboptimal policies. The problem with an arbitrary set $\mathcal{M}$ is that KL-divergence might match features of the environment that are irrelevant to the optimal policy function, a^*. We show now that this outcome is not inevitable: 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 payoff-relevant features of the environment. This insight can be generalized and serve as a guideline for specifying $\mathcal{M}$ in situations in which misspecification is an issue.

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, there exist functions* $T_u\colon \mathcal{X} \to \mathbb{R}^d$*, and* $g_u\colon \mathbb{R}^d \to \mathcal{A}$*, such that for any* $Q \in \mathcal{P}$ *the best-responding policy is defined as,*

$$b(Q) = g_u\Big(m(Q; T_u)\Big), \tag{12}$$

with $m(Q; T_u) := \mathbb{E}_Q\left[T_u(X)\right]$ *representing the vector of payoff-relevant moments.*

Decision problems that satisfy Assumption 1 possess two salient features. One concerns the interaction between the decision maker and her environment and it implicitly requires that the DM's actions do not affect the data generating process itself. This property excludes active learning or bandit problems. The other feature concerns the statistical complexity of the decision problem and it requires a degree of simplicity in terms of welfare-relevant characteristics of the DGP. This simple class nonetheless includes a large class of decision problems. For example, every non-active learning problems with a finite policy space satisfies Assumption 1.[18] Moreover, as seen by the example in section 3, there are various decision problems with an infinite policy space that are "simple enough" to satisfy Assumption 1. In that case, Assumption 1 requires limited interaction between implementable policies and the state vector. Any decision problem characterized by a differentiable utility function that can be written in the form $u(a(x), x) = \langle f(a), T_u(x)\rangle$ for some function $f(a) \in \mathbb{R}^d$ also satisfies Assumption 1. This class includes quadratic utility functions, isoelastic utility functions, and many others.

The next proposition states that for learning problems that satisfy Assumption 1 one can construct sets of likelihoods that are both *misspecification-proof* and *locally-payoff-optimal*. In this sense concerns about misspecification should largely depend on the utility function characterizing the decision problem as opposed to the potential complexity of the data generating process.

Proposition 1.

Under Assumption 1, the exponential family of likelihoods defined by the finite sufficient statistics

[18] Enumerating policies as $\mathcal{A} = \{a_1, \dots, a_N\}$ we can define the payoff-relevant statistics as,

$$T_u(x) = \begin{pmatrix} u(a_1(x), x) - u(a_2(x), x) \\ \vdots \\ u(a_{N-1}(x), x) - u(a_N(x), x) \end{pmatrix} \in \mathbb{R}^{N-1}. \tag{13}$$

T_u,[19]

$$\mathcal{M} = \left\{ q_\theta(x) = h_u(x) \exp\left\{ \theta \cdot T_u(x) - A_u(\theta) \right\} \;:\; \theta \in \Theta \subseteq \mathbb{R}^d \right\}, \tag{14}$$

for some $h_u \colon \mathcal{X} \to \mathbb{R}_+$, where $A_u(\theta)$ 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 that we expand on in Appendix A. The KL-divergence minimizing distribution within $\mathcal{M}$ is characterized by,

$$\mathbb{E}_{Q^{\mathrm{KL}}_{\mathcal{M},P}}[T_u(X)] = \mathbb{E}_P[T_u(X)], \tag{15}$$

so that likelihood-based learning coincides with the method of moments in the sense that the KL-divergence minimizing likelihood exactly matches the payoff-relevant moments of the data generating process. In the limit the decision maker implements the policy function $b\left(Q^{\mathrm{KL}}_{\mathcal{M},P}\right)$, which—following the logic in Lemma 1—leads to the same long-run average payoff as $a^* = b\,(P)$.

The moral of Proposition 1 is that even if the environment described by the underlying DGP 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.

In this sense our result can also be viewed as a recommendation for constructing the relevant state space in a given decision problem. As in our example, considering the whole set $\mathcal{P}$ as the state space would necessitate the use of non-parametric models. However, this turns out to be unnecessary if one can partition $\mathcal{P}$ into parsimonious equivalence classes by means of the inverse-image of the

[19]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.

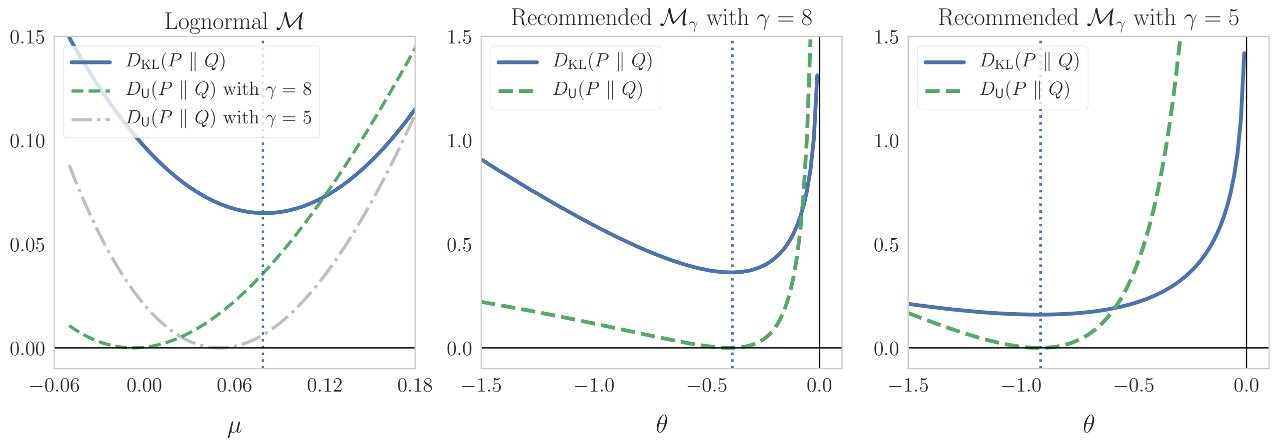

Figure 2: Similarity measures D_{KL} and D_{U} over $\mathcal{M}$. Vertical dotted lines represent the KL-divergence minimizing parameters. Left panel shows the case when $\mathcal{M}$ includes lognormal distributions with fixed $\sigma^2 = \sigma^2_{\mathrm{KL}}$ so that the set is indexed only by μ. Middle and right panels depict cases when the recommended set $\mathcal{M}_\gamma$ in (16) is used with different degrees of risk aversion.

best response function.[20] In this scenario, states can be defined through moments of payoff-relevant statistics much like in the case of method of moments approaches.

Example revisited

To see how to apply the introduced concepts, we revisit the consumption-saving problem of section 3. Recall that we identified the risk-adjusted expected return as the only payoff-relevant moment of X, therefore, Assumption 1 is satisfied and misspecification-proof learning is feasible. The sufficient statistic can be written as

$$T_u(x) := x^{1-\gamma} \qquad \Rightarrow \qquad m(Q; T_u) = \mathbb{E}_Q\left[T_u(X)\right].$$

Moreover, using standard integration by substitution logic,[21] it is straightforward to show that

$$A_u(\theta) := -\log\left(\theta(1-\gamma)\right) \qquad \text{and} \qquad h_u(x) := x^{-\gamma}.$$

[20] Using the introduced notation the state space is defined as $\left\{b^{-1}(a) \subset \mathcal{P} \colon a \in \mathcal{A}\right\}$.

[21] For simplicity, we assume $\gamma > 1$, so that $\theta < 0$ is required for q_θ to be a probability density function.

Therefore, our recommended set of (misspecified) likelihoods is,

$$\mathcal{M}_\gamma := \Big\{ q_\theta(x) = x^{-\gamma} \exp\left\{\theta x^{1-\gamma} + \log(\theta(1-\gamma))\right\} \quad : \quad \theta < 0,\ x > 0 \Big\}, \tag{16}$$

where the γ-index emphasizes the prior support's dependence on the decision maker's risk aversion. Figure 2 illustrates how $\mathcal{M}_\gamma$ renders the implied information geometry aligned with the utility geometry. The left panel shows again that lognormal likelihoods are inconsistent with the consumption-saving example in Section 3. Because the best response function depends on a single moment, a one-parameter family should be sufficient for learning. To this end, we take a subset of $\mathcal{M}$ by fixing $\sigma^2 = \sigma^2_{\text{KL}}$ so that the entertained set of lognormal distributions can be indexed by μ. The left panel of Figure 2 demonstrates how the D_{KL}- and D_{U}-minimizing lognormal distributions differ from each other.[22] Since the entertained set does not depend on preferences, D_{KL} remains unchanged as γ varies.

The middle and right panels of Figure 2 depict cases when the agent entertains our recommended set $\mathcal{M}_\gamma$ parametrized by θ for different levels of risk aversion. Evidently, KL-divergence now depends on γ reflecting the fact that we chose $\mathcal{M}_\gamma$ so that D_{KL} is focused on the risk-adjusted expected return. As a result, the D_{KL}- and D_U-minimizing distributions coincide irrespective of γ (or P). Nevertheless, the set $\mathcal{M}_\gamma$ is still misspecified—as can be seen from D_{KL} not taking the value zero over $\mathcal{M}_\gamma$.

Discussion

We have considered the asymptotic welfare implications of misspecification in the context of Bayesian learning. We have argued that in a wide class of decision problems the DM can avoid the harm from misspecification by tailoring her entertained likelihoods to her preferences. In the introduction we started out by considering the DM treating her likelihoods as instruments to implement well-performing policies, where the performance is judged against an unknown DGP. In light of Proposition 1 a successful recipe for learning might take the following form. Instead of defining beliefs over the state space of all potential *i.i.d.* distributions over gross returns, the Bayesian DM

[22] The left panel of Figure 2 can be viewed as a horizontal slice of value $\sigma^2 = \sigma^2_{\text{KL}}$ in the right panel of Figure 1.

should define her beliefs over the state space of equivalent distributions where equivalence is defined relative to the implied best-responding policies. In case the DM's preferences are described by her relative risk aversion parameter γ, these equivalent distributions are defined by having the identical moment $\mathbb{E}\left[X^{1-\gamma}\right]$. When the best-responding policies depend on finitely many such moments, the exponential family provides an immediate way to define a set of likelihoods that is misspecification-proof. Importantly, we do not require the Bayesian DM to believe that the true DGP lies in the defined exponential family, only that she is able to define her beliefs over the set of distributions with identical welfare-relevant moments. In this way, Bayesian learning can combine prior information, while still entertain likelihoods that are only instrumental in implementing well-performing policies.

5 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.[23] That said, it seems sensible to impose consistency among beliefs and preferences even if learning is correctly specified.

An important advantage of our approach is that adopting the proposed misspecification-proof model, and thereby ensuring no asymptotic bias in the implemented policy,[24] does not necessarily lead to increased finite sample variance. In fact, in the example of Section 3, a natural alternative to our recommended exponential family would be to use a nonparametric kernel estimator. Nevertheless, this kernel density estimator would exhibit the error decay rate of $O\left(n^{-4/5}\right)$, while

[23] This definition of irrationality is motivated by Gilboa and Schmeidler (2001) and Gilboa (2009).

[24] Note that bias in the space of implemented policies and bias in the space of data generating processes differ.

our recommended exponential family—built from payoff-relevant statistics—converges at the faster rate of $O\left(n^{-1}\right)$. Other parametric families would have the same $O\left(n^{-1}\right)$ error rate, but as demonstrated by our example they run the risk of targeting welfare-irrelevant statistics and hence inducing suboptimal policies.

Our results do not hinge on the decision maker being fully Bayesian. As we have seen, initialized with the recommended misspecification-proof set of likelihoods, Bayesian updating, frequentist MLE, and moment-based learning all lead to implementing the same policy. In a similar vein, it is not necessary that the decision maker chooses her policies in a Bayesian manner. For instance, provided that ambiguity vanishes (as in Marinacci (2002)) a max-min decision rule collapses to expected utility maximization under the limit point of learning.

References

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

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

Blume, L. E., Cogley, T., Easley, D. A., Sargent, T. J., and Tsyrennikov, V. (2018). A case for incomplete markets. *Journal of Economic Theory*, 178:191 – 221.

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

Bray, M. and Kreps, D. (1987). Rational learning and rational expectations. In Feiwel, G. R., editor, *Arrow and the Ascent of Modern Economic Theory*, pages 597–625. Palgrave Macmillan, London.

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

Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *The Annals of Statistics*, pages 1–26.

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

Epstein, L. G. and Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969.

Esponda, I. and Pouzo, D. (2016). Berk-Nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica*, 84(3):1093–1130.

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

Geweke, J. and Amisano, G. (2012). Prediction with misspecified models. *American Economic Review*, 102(3):482–86.

Gilboa, I. (2009). *Theory of Decision under Uncertainty*. Econometric Society Monographs. Cambridge University Press.

Gilboa, I. and Schmeidler, D. (2001). *A Theory of Case-Based Decisions*. Cambridge University Press.

Granger, C. W. J. and Newbold, P. (1973). Some comments on the evaluation of economic forecasts. *Applied Economics*, 5(1):35–47.

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

Hansen, L. P. and Sargent, T. J. (2018). *Risk, Uncertainty, and Value*. Princeton, New Jersey: Princeton University Press.

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

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

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

Marinacci, M. (2002). Probabilistic sophistication and multiple priors. *Econometrica*, 70(2):755–764.

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

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

Schorfheide, F. (2001). Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, 15(6):645–670.

Shalizi, C. R. (2009). Dynamics of Bayesian updating with dependent data and misspecified models. *Electronic Journal of Statistics*, 3:1039–1074.

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

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer Science & Business Media.

White, H. (1996). *Estimation, Inference and Specification Analysis*. Number 9780521574464 in Cambridge Books. Cambridge University Press.

A Exponential families and misspecification-proofness

Proposition 1 follows from well-known properties of exponential families. For an excellent exposition see Shao (2003). For completeness we spell out the arguments below.

Consider the exponential family defined in Proposition 1 given in its natural parametrization. Under the regularity conditions mentioned before (Shalizi, 2009) Bayesian learning concentrates on the likelihood minimizing the KL-divergence relative to the DGP.

$$\begin{aligned} D_{\mathrm{KL}}\left(P \parallel Q_\theta\right) &= \int \log \frac{p(x)}{q_\theta(x)} \mathrm{d}P(x) \\ &= -\int \log h_u(x) + \theta \cdot T_u(x) - A_u(\theta) \mathrm{d}P(x) \end{aligned} \tag{17}$$

The log partition or cumulant function $A_u(\theta)$ ensure that the likelihoods q_θ are properly normalized. It is defined as,

$$A_u(\theta) \coloneqq \log \int h_u(x) \exp\{\theta \cdot T_u(x)\} \mathrm{d}x. \tag{18}$$

Taking the derivative of the cumulant we obtain,

$$\begin{aligned} \nabla A_u(\theta) &= \frac{\int h_u(x) \exp\{\theta \cdot T_u(x)\} T_u(x) \mathrm{d}x}{\int h_u(x) \exp\{\theta \cdot T_u(x)\} \mathrm{d}x} \\ &= \int h_u(x) \exp\{\theta \cdot T_u(x) - A_u(\theta)\} T_u(x) \mathrm{d}x = \mathbb{E}_{Q_\theta}\left[T_u(X)\right]. \end{aligned} \tag{19}$$

Hence, the first-order condition characterizing the KL-divergence minimizing likelihood implies,

$$\mathbb{E}_P\left[T_u(X)\right] - \mathbb{E}_{Q_\theta}\left[T_u(X)\right] = 0. \tag{20}$$

That is, Bayesian learning concentrates on the likelihood which matches the payoff-relevant moments under the true DGP.