

Non-Bayesian Learning in Misspecified Models*

Sebastian Bervoets[†] & Mathieu Faure[‡] & Ludovic Renou[§]

Latest version: October 2, 2025

First version: June 26, 2024

Abstract

Deviations from Bayesian updating are traditionally categorized as biases, errors, or fallacies, thus implying their inherent “sub-optimality.” We offer a more nuanced view. In learning problems with misspecified models, we show that some non-Bayesian updating can outperform Bayesian updating.

KEYWORDS: learning, Bayesian, consistency.

JEL CLASSIFICATION: C73, D82.

*We are grateful to Michel Benaim and Peter Grünwald for their invaluable help and to Aislinn Bohren, Daniel Hauser, Jan Knoepfle, Pietro Ortoleva, Mallesh Pai, Maher Said, Larry Samuelson, Ran Spiegler, Jakub Steiner, Philipp Strack, and seminar audiences for insightful discussions. LR thanks the generous hospitality of AMSE and LUISS. SB and MF acknowledge financial support from the French government under the “France 2030” investment plan managed by the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative of Aix-Marseille University - A*MIDEX.

[†]Aix-Marseille Univ., CNRS, AMSE, sebastian.bervoets(at)univ-amu.fr

[‡]Aix-Marseille Univ., CNRS, AMSE, mathieu.faure(at)univ-amu.fr

[§]ASU, QMUL and CEPR, lrenou.econ(at)gmail.com

1 Introduction

Psychologists, economists, and other social scientists have extensively documented systematic departures from Bayesian updating. These departures include underreaction and overreaction to new information, confirmation bias, base-rate neglect, and numerous other phenomena.¹ While these departures are traditionally characterized as biases, errors, or fallacies – implying their inherent “sub-optimality” – we offer a more nuanced view. Specifically, we demonstrate that in classical learning problems, some non-Bayesian updating can outperform its Bayesian counterpart.

The benefits of non-Bayesian updating have precedents in the literature, particularly in strategic environments where such departures can induce equilibrium behavior favorable to non-Bayesian agents. In sender-receiver games, for instance, non-Bayesian updating by receivers can elicit more information from the senders (De Clippel and Zhang, 2022; Lee et al., 2023).² However, these results rely fundamentally on strategic interactions. In single-agent learning environments, where strategic responses are absent, we identify a distinct mechanism through which non-Bayesian updating may confer a benefit, which we now explain.

We formalize our analysis within the canonical learning framework of Berk (1966). Consider an agent attempting to learn a true data generating process. The agent begins with a set of candidate processes, none of which may exactly match the true process. Then, in each of an infinite number of periods, the agent sequentially observes a new piece of data and updates their beliefs about which process best describes the observations. We model the agent as a *conservative* Bayesian, following the updating rule:

$$(\text{belief at } n + 1) = (1 - \gamma) \times (\text{belief at } n) + \gamma \times (\text{Bayesian update at } n + 1), \quad (\text{C-Bay})$$

where γ may depend on the number of observations, current belief, and current observation. This rule, first introduced by Phillips and Edwards (1966) and Edwards (1968), generalizes Bayesian updating and captures several documented cognitive biases, including under-reaction and over-reaction to new information, confirmation bias, and base-rate neglect. Throughout, we refer to this updating rule as the (C-Bay) rule.

In the first part of the analysis, we assume that the more observations the agent has, the less the agent reacts to new information. A body of evidence supports this assumption; in his meta-analysis, Benjamin (2019) reports it as Stylized Fact 2. Our main result

¹For comprehensive reviews, see Benjamin (2019) and Ortoleva (2022).

²As another instance, in a financial trading problem, Massari (2022) shows that non-Bayesian traders may drive Bayesian traders out of the market. As our analysis will make it clear, this is because non-Bayesian traders may better predict the underlying fundamentals than their Bayesian counterparts.

– Theorem 1 – states that the agent’s *predictive* process – the expectation of candidate processes under the agent’s beliefs – converges almost surely to the *mixture* process the closest to the true data generating process (in the sense of the Kullback-Leibler divergence). In other words, the theorem states that it is *as if* the agent considers all possible mixtures of the candidate processes (thus, “convexifying” his model) and learns as a classical Bayesian. In particular, if the true data generating process is a mixture of the candidate processes, the agent’s predictive process converges to it. This contrasts with the Bayesian agent, whose predictive process converges to the *pure* process the closest to the true data generating process – see [Berk \(1966\)](#). (We stress that when the model is correctly specified, the updating rule (**C-Bay**) converges to the true data-generating process, as Bayesian updating does.)

The intuition is simple. Bayesian agents extract all the information there is from their models.³ However, when their models are wrong, this typically causes them to make wrong inferences. A more conservative approach, as embodied by the updating rule (**C-Bay**), limits the severity of this problem. Moreover, the more observations an agent has, the more conservative they must be to counteract the increasing confidence of their Bayesian counterpart. Ultimately, this can lead to better learning outcomes.

Mathematically, we rewrite the updating rule (**C-Bay**) as the Robbins-Monro algorithm ([Robbins and Monro, 1951](#)):

$$\begin{aligned} (\text{belief at } n + 1) = & (\text{belief at } n) + \gamma_{n+1} \left[\mathbb{E}_{\text{true DGP}}(\text{Bayesian update at } n + 1) - (\text{belief at } n) \right] \\ & + \gamma_{n+1} \left[(\text{Bayesian update at } n + 1) - \mathbb{E}_{\text{true DGP}}(\text{Bayesian update at } n + 1) \right], \end{aligned}$$

where $\mathbb{E}_{\text{true DGP}}$ is the expectation with respect to the true data-generating process. The first term in brackets is deterministic, while the second term is random (a martingale difference). When the agent reacts sufficiently slowly to new information, the random term becomes negligible as the number of observations grows arbitrarily large. The deterministic part then drives the belief dynamics. We show that the cross-entropy of the predictive process with respect to the true data generating process acts as a potential for the deterministic dynamics.⁴ We then show that all critical points of the deterministic flow correspond to either the unconstrained maximizers of the cross-entropy or constrained ones. Finally, we prove that all constrained maximizers repel the dynamics, hence the dynamics converges to the unconstrained maximizers. This last step is the most delicate. The constrained maximizers are neither isolated points nor interior, which precludes us from appealing to classical results in stochastic approximation

³[Zellner \(1988\)](#) proves that Bayesian updating corresponds to extracting all the information from the prior and the statistical model.

⁴The KL divergence of p from q is $\mathbb{E}_p[\log(p/q)]$. Minimizing the KL divergence with respect to q is equivalent to maximizing $\mathbb{E}_p[\log q]$. We call the latter the cross-entropy of q with respect to p .

theory such as [Pemantle \(1990\)](#), [Brandiere \(1998\)](#), or [Brandiere and Duflo \(1996\)](#). We develop new arguments to solve that issue.

In the second part of the analysis, we consider alternative assumptions on the updating weights and show that our main insight, that is, that some non-Bayesian updating can outperform Bayesian updating, continues to hold. For instance, if the updating weights are constant, then the belief dynamics does not converge, but its occupation measure does. As the weight γ becomes arbitrarily small, the occupation measure concentrates on the mixture process the closest to the true data generating process – [Theorem 3](#).

We illustrate our main result with a simple example, inspired by the work of [Spiegler \(2016\)](#) on causal models. Suppose that each observation is a vector of variables (x_1, x_2, x_3) , representing diet, sleep, and health, respectively. The true causal model is such that diet and sleep function are *interdependent* determinants of health, with diet also influencing sleep patterns. However, the agent posits another causal model and assumes that diet and sleep are *independent* causes of health. [Figure 1](#) represents the causal models as two directed acyclic graphs. Thus, the agent considers distributions of the form $p(x^3|x^1, x^2)p(x^2)p(x^1)$, which all differ from the true distribution since x^1 and x^2 are correlated in the true model. Since any correlated distribution is the convex combination of independent distributions, our main result implies that the non-Bayesian agent successfully converges to the correct causal model, while the Bayesian agent fails to do so. This is the paper’s main insight.

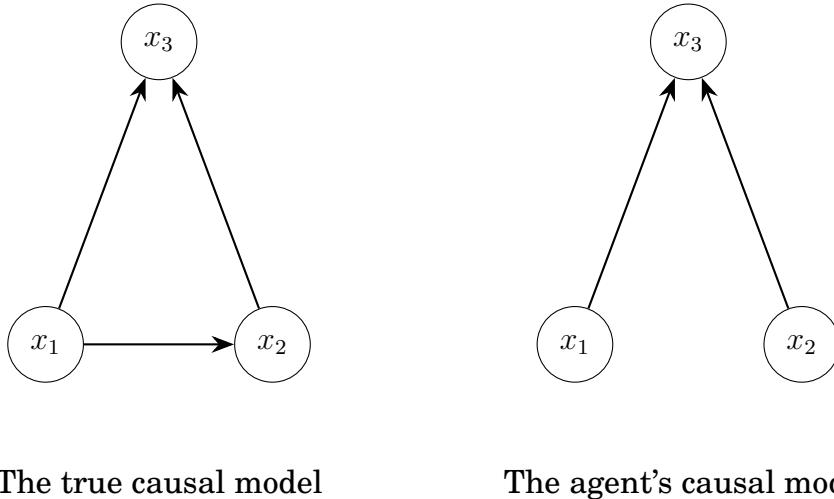


Figure 1: The causal models

Our analysis also suggests an advantage to deliberately adopt a misspecified model. In Bayesian statistics, computing posterior distributions is often challenging. Thus, even when the statistical model is correctly specified, learning the true data-generating process may be computationally intractable – see the large literature on approximate

Bayesian computation. Our analysis suggests an alternative route. Instead of working directly with the distributions of interest, the statistician can restrict attention to families of distributions whose mixtures are rich enough to approximate the distributions of interest, while also yielding posteriors that are easier to compute. Although such a restricted model may be misspecified, applying the updating rule (C-Bay) still guarantees convergence to the true data-generating process, without requiring intractable posterior computations.⁵ For example, restricting attention to Gaussian distributions suffices, since mixtures of Gaussians approximate any distribution arbitrarily well (with respect to the Wasserstein metric).

We close the introduction with a brief discussion of the related literature (Section 5 contains a detailed discussion). The two most closely related papers are Berk (1966) and Epstein et al. (2010). Berk (1966) studies the long-run properties of Bayesian updating when the model is misspecified and shows that the beliefs converge to the process, which minimizes the KL divergence from the true process.⁶ Epstein et al. (2010) study the long-run properties of the updating rule (C-Bay) when the model is correctly specified. They show that if the agent under-reacts to new information, the agent eventually learns the true process. We differ from Berk (1966) by assuming non-Bayesian updating and from Epstein et al. (2010) by assuming a misspecified model. Some recent papers, e.g., Bohren and Hauser (2021, 2024), Heidhues et al. (2021) and Frick et al. (2023), provide a link between Bayesian updating in misspecified problems and non-Bayesian updating. More specifically, it is possible to construct another set of candidate processes such that Bayesian updating in that modified model induces the same beliefs than the updating rule (C-Bay) in the original model. To the best of our understanding, none of these recent results are applicable to our problem – see Section 3 for a detailed discussion.

2 The Problem

At each period $n \in \mathbb{N}$, an agent is observing $x_n \in X$, independently and identically drawn from an *unknown* data-generating process $p^* \in \Delta(X)$. The agent’s model is a set $\mathcal{P} := \{p_\theta \in \Delta(X) : \theta \in \Theta\}$ of data-generating processes, which may not include the true data-generating process p^* – in that case, the model is misspecified. Throughout, we assume that the sets X and Θ are finite and that the support of p^* and each p_θ is X .⁷ The latter assumption guarantees that the agent is never surprised, that is, the agent

⁵As we show later, our main result – Theorem 1 – also holds when the Bayesian posteriors in rule (C-Bay) are replaced with approximations.

⁶See also Freedman (1963), Berk (1970), Diaconis and Freedman (1986), Diaconis and Freedman (1999), Shalizi (2009).

⁷None of our results would change if we were to assume that the support of p^* is included in the support of p_θ for all $\theta \in \Theta$. The only modification would be to substitute X with the support of p^* in the analysis.

cannot observe a realization x , believed impossible under all data-generating processes in \mathcal{P} .

We can think of x as financial returns in investment problems, losses in insurance problems, or demands in pricing problems. For instance, in a canonical portfolio problem, x corresponds to realized asset returns and \mathcal{P} their postulated distributions, e.g., multivariate Gaussian distributions with unknown mean and variance-covariance matrix (the parameter θ). As another instance, in canonical insurance problems, x corresponds to losses or, equivalently, final wealths. More generally, we view x as payoff-relevant consequences, p^* as their true distribution, and \mathcal{P} as the collection of “parameterized” distributions the agent postulates. An alternative is to view x as a (payoff-irrelevant) signal about the (payoff-relevant) state θ , p^* as the signal distribution when the true state is $\theta^* \in \Theta$, and $(p_\theta)_\theta$ as the perceived conditional distributions (an experiment). Both views are common in Economics and Statistics. While our formal results are independent of either perspective, their interpretations differ. Under the first view, successful learning entails the convergence of the *predictive process* to the true distribution p^* , whereas under the alternative view, learning success is characterized by the identification of the true state θ^* . We return to these alternative interpretations later on.

The agent’s prior is $q_0 \in \mathbf{S} := \Delta(\Theta)$. Without loss of generality, we assume that the support of q_0 is Θ , i.e., $q_0 \in \mathbf{S}^* := \text{int}(\mathbf{S})$, the interior of \mathbf{S} . The agent updates his belief from q_n to q_{n+1} upon observing the signal x_{n+1} according to the rule:

$$q_{n+1} = (1 - \gamma_{n+1})q_n + \gamma_{n+1}B(q_n, x_{n+1}), \quad (\text{C-Bay})$$

where $B(q_n, x_{n+1}) \in \Delta(\Theta)$ is the Bayesian posterior of q_n given the new observation x_{n+1} and γ_{n+1} the (possibly random and positive) weight on the Bayesian posterior.⁸ The updating rule (C-Bay) was first discussed in Phillips and Edwards (1966) and Edwards (1968), generalizes the Bayesian rule ($\gamma_{n+1} = 1$) and can accommodate a number of well-documented biases such as underreaction ($\gamma_{n+1} < 1$) and overreaction ($\gamma_{n+1} > 1$) to new information, confirmatory biases, and a few others. We refer the reader to Epstein (2006) and Kovach (2021) for more extensive discussions and axiomatic foundations. Benjamin (2019) documents two additional empirical findings, which the rule can also accommodate. First, the more observations an agent has, the more the agent underreacts. This finding – Stylized Fact 2 in Benjamin (2019) – is consistent with decreasing updating weights $(\gamma_n)_n$. Second, individuals’ beliefs after sequentially observing x_1 to x_n differ from the beliefs they would have formed, had they received the same information simultaneously as a single observation (x_1, \dots, x_n) . This finding – Stylized Fact 8 in

⁸The posterior probability of θ is $B_\theta(q, x) = \frac{q_\theta p_\theta(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)}$, whenever the denominator is positive.

[Benjamin \(2019\)](#) – is also consistent with the updating rule (C-Bay).⁹ Throughout, we refer to the process $(\mathbb{E}_{q_n}[p_\theta])_{n \in \mathbb{N}}$ under the updating rule (C-Bay) as the agent’s *predictive* process.

We conclude this section with a brief review of [Berk \(1966\)](#) and [Epstein et al. \(2010\)](#). [Berk \(1966\)](#) studies the convergence of the *Bayesian* posteriors (i.e., $\gamma_n = 1, \forall n$) and proves the almost-sure convergence to the maximizers of the (negative of the) cross entropy, that is, $\arg \max_{\theta \in \Theta} \sum_x p^*(x) \log p_\theta(x)$. [Epstein et al. \(2010\)](#) study whether the updating rule (C-Bay) induces the agent to eventually learn the true data-generating process, when the model is *correctly* specified. They show that if the agent is always underreacting to new information (i.e., $\gamma_n \leq 1, \forall n$), then the agent eventually learns the true data-generating process, when the weight γ_n is predictable (measurable with respect to the information up to period $n - 1$). With the help of an example, they show that this result is not true when the weight γ_n is only adapted (measurable with respect to the information up to period n).

3 Underreaction and Learning

This section presents our main result. Motivated by Stylized Fact 2 in [Benjamin \(2019\)](#), we assume that the more observations the agent has, the more the agent underreacts to new information. Yet, we do not want the underreaction to be so severe that the agent stops reacting after finitely many periods.

Formally, we assume that $(\gamma_n)_n$ is a deterministic sequence of positive real numbers such that

$$(i) : \sum_{n=0}^{+\infty} \gamma_n = +\infty, \text{ and } (ii) : \sum_{n=0}^{+\infty} e^{-c/\gamma_n} < +\infty, \text{ for all } c > 0. \quad (1)$$

Examples of series satisfying conditions (i) and (ii) include $(1/n)_{n \in \mathbb{N}}$, $(1/\sqrt{n})_{n \in \mathbb{N}}$, and $(1/\log^2(n))_{n \in \mathbb{N}}$. More generally, the series $(1/n^\alpha)_n$ with $\alpha \in (0, 1]$ and $(1/\log^\alpha(n))_n$ with $\alpha > 1$ all satisfy the condition.¹⁰ Condition (ii) guarantees that the series $(\gamma_n)_n$ converges to zero, so that the more observations the agent has, the less the agent reacts. Condition (ii) further says $(\gamma_n)_n$ converges to zero faster than the series $(1/\log(n))_n$ does.¹¹ Condition (i) guarantees that the agent does not stop reacting after finitely many periods.

The analysis rests on the observation that the updating rule (C-Bay) can be rewritten

⁹[Cripps \(2018\)](#) characterizes the updating rules, which produce the same update after observing either x_1 to x_n sequentially, or (x_1, \dots, x_n) simultaneously.

¹⁰A more familiar, but stronger, assumption is $\sum_{n=0}^{+\infty} \gamma_n = +\infty$ and $\sum_{n=0}^{+\infty} \gamma_n^2 < +\infty$.

¹¹This is not a demanding condition as the series $(1/\log(n))_n$ goes relatively slowly to zero.

as:

$$q_{n+1} - q_n = \gamma_{n+1} (H(q_n) + U_{n+1}),$$

where $H : \mathbf{S} \rightarrow \mathbb{R}_0^{|\Theta|}$ is defined by

$$H_\theta(q) := q_\theta \left(\underbrace{\sum_x p^*(x) \frac{p_\theta(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)}}_{:=f_\theta(q)} - 1 \right) = \underbrace{\mathbb{E}_{p^*} [B_\theta(q, \mathbf{x})]}_{\text{Expectation of Bayesian posterior of } \theta} - q_\theta, \quad \forall \theta, \quad (2)$$

and $(U_n)_n$ a bounded martingale difference: $U_{n+1} := B(q_n, x_{n+1}) - \mathbb{E}_{p^*} [B(q_n, \mathbf{x}_{n+1})]$.

Before proceeding, we remark that the analysis below continues to apply if the agent updates his belief from q_n to $q_{n+1} := (1 - \gamma_{n+1})q_n + \gamma_{n+1}b(q_n, x_{n+1})$, where $b(q_n, x_{n+1})$ is a *noisy estimate* of $B(q_n, x_{n+1})$ such that $\mathbb{E}[b(q, x)|B(q, x)] = B(q, x)$ for all (q, x) . In words, the updating rule tolerates some errors.

This recursive formulation is a Robbins-Monro algorithm (Robbins and Monro, 1951), a well-known algorithm in the literature on stochastic approximation theory. A popular method for analyzing such algorithms is the ordinary differential equation (ODE) method, which approximates the sequence $(q_n)_n$ with the solution to the associated (deterministic) ODE:

$$\dot{q}(t) = H(q(t)).$$

When applicable (which it is in our case), this method reduces the analysis of the Robbins-Monro algorithm (C-Bay) to the study of the associated ODE. We prove two central results about the ODE. (All proofs are in the Appendix.) The first result – Lemma 1 – states that the function $q \mapsto V(q)$ is a Lyapunov function for the ODE, where $V(q)$ is the (negative of the) cross-entropy of the mixture $\sum_\theta q_\theta p_\theta$ with respect to p^* , that is,

$$V(q) := \sum_x p^*(x) \log \left(\sum_\theta q_\theta p_\theta(x) \right). \quad (3)$$

The second result – Lemma 2 – states that the set $E := \{q \in \mathbf{S} : H(q) = 0\}$ of zeroes of H can be decomposed into a finite union of disjoint, compact and *convex* components C_k , $k = 1, \dots, K$, with the cross-entropy V constant on each of the components.¹² The decomposition is unique, and V attains its maximal value on a unique component, which we denote C_{k^*} :

$$V(C_{k^*}) > V(q), \text{ for all } q \in \mathbf{S} \setminus C_{k^*}.$$

¹²Note that the set of degenerated beliefs $\{\delta_{\{\theta\}} : \theta \in \Theta\}$ belong to E , so that E is non-empty.

To get some intuition for Lemma 2, let $C_{\hat{\Theta}}$ be the maximizers of the cross-entropy V over the face $F_{\hat{\Theta}}$, where $F_{\hat{\Theta}} := \{q \in \mathbf{S} : q_{\theta} > 0 \text{ iff } \theta \in \hat{\Theta}\}$. If non-empty, $C_{\hat{\Theta}}$ is convex and induces a unique $p \in \Delta(X)$, so that the cross-entropy is constant on $C_{\hat{\Theta}}$. Since there are finitely many faces, there are finitely many non-empty sets of maximizers. (At least one is non-empty, since V is continuous and \mathbf{S} non-empty and compact.) Moreover, it is easy to see that each maximizer is a zero of H and, conversely. Intuitively, $H(q) = 0$ are the Kuhn-Tucker conditions for the maximization of the concave function V with respect to $q \in F_{\hat{\Theta}}$. We then construct each C_k as a carefully chosen union of maximizers. Example 1 illustrates the construction. Note that we can have as many components as we have faces.

Example 1 Assume that $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $X = \{x_1, x_2, x_3\}$, and

$$p_{\theta_1} = \frac{1}{5}(1, 2, 2), \quad p_{\theta_2} = \frac{1}{5}(2, 1, 2), \quad p_{\theta_3} = \frac{1}{5}(2, 2, 1), \quad p^* = \frac{1}{3}(1, 1, 1).$$

The set E is:

$$\{\delta_{\{\theta_1\}}\} \cup \{\delta_{\{\theta_2\}}\} \cup \{\delta_{\{\theta_3\}}\} \cup \left\{ \frac{1}{2}(1, 1, 0) \right\} \cup \left\{ \frac{1}{2}(1, 0, 1) \right\} \cup \left\{ \frac{1}{2}(0, 1, 1) \right\} \cup \left\{ \frac{1}{3}(1, 1, 1) \right\}.$$

We have $V(\delta_{\{\theta_1\}}) = V(\delta_{\{\theta_2\}}) = V(\delta_{\{\theta_3\}}) = \frac{1}{3} \log \frac{1}{5} + \frac{2}{3} \log \frac{2}{5}$,

$$V\left(\frac{1}{2}(1, 1, 0)\right) = V\left(\frac{1}{2}(1, 0, 1)\right) = V\left(\frac{1}{2}(0, 1, 1)\right) = \frac{2}{3} \log \frac{3}{10} + \frac{1}{3} \log \frac{4}{10},$$

and $V\left(\frac{1}{3}(1, 1, 1)\right) = \log \frac{1}{3}$. Here, $C_{k^*} = \left\{ \frac{1}{3}(1, 1, 1) \right\}$, the interior point.

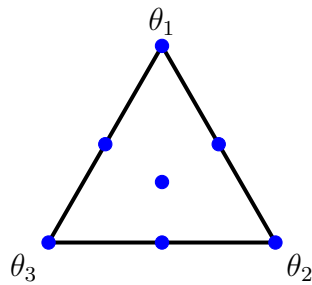


Figure 2: The components in Example 1

To sum up, the cross-entropy V is a Lyapunov function for the flow of the dynamics and maximized in a unique component C_{k^*} of zeroes. This suggests that the component C_{k^*} is globally stable. While our main result – Theorem 1 – states that this is indeed the case, the result is far from immediate. As we shall see, the main technical challenge is that the components C_k are not always *singletons* (isolated points) and, moreover, often exist as *continuuums* on the boundary.

Theorem 1 For all $q_0 \in \text{int}(\mathbf{S})$, the (random) limit set $\mathcal{L}((q_n)_n)$ of the process $(q_n)_n$ is contained in C_{k^*} , with probability one. In words, the beliefs converge to maximizers of the cross-entropy V .

Theorem 1 states that every ω -limit point q induces a unique distribution $\sum_{\theta} q_{\theta} p_{\theta} \in \text{co } \mathcal{P}$, which attains the maximum:

$$\max_{p \in \text{co } \mathcal{P}} \sum_x p^*(x) \log p(x).$$

It then follows immediately that:

Corollary 1 If $p^* \in \text{co } \mathcal{P}$, then the agent’s predictive process converges to the true data-generating process p^* .

Thus, if $p^* \in \text{co } \mathcal{P}$, the learning outcome is the same as the one of a Bayesian learner with the richer set of (non-parametric) models $\text{co } \mathcal{P}$, assuming that the agent’s prior assigns a positive probability to an open neighborhood of the cross-entropy maximizer – see [Berk \(1966\)](#). This observation is reminiscent of [Bohren and Hauser \(2024\)](#), who demonstrate an equivalence between non-Bayesian updating in a correctly specified model and Bayesian updating in a misspecified model. In particular, if the agent’s model is $\{\pi_{\theta}^{q_n, \gamma_n} : \theta \in \Theta\}$, with

$$\pi_{\theta}^{q_n, \gamma_n}(x) := (1 - \gamma_n) \mathbb{E}_{q_n}[p_{\theta}(x)] + \gamma_n p_{\theta}(x)$$

for all x , then the Bayesian posteriors of the modified model $\{\pi_{\theta}^{q_n, \gamma_n} : \theta \in \Theta\}$ coincide with the non-Bayesian posteriors (C-Bay) of the original model $\{p_{\theta} : \theta \in \Theta\}$, when the agent’s prior belief is q_n at period n . We stress that the “*as-if*” model depends on the current belief q_n and weight γ_n . The dependence on current beliefs is a common feature of social learning models, see e.g., [Esponda and Pouzo \(2016\)](#), [Bohren and Hauser \(2021\)](#), and [Frick et al. \(2023\)](#). However, as far as we know, none of the results developed in these (and other related) papers are directly applicable to our problem. Indeed, the main focus of this literature is on the stability of degenerated beliefs and none of the few instances, where mixed beliefs are considered, are directly applicable in our setting. See [Section 5](#) for a more detailed discussion.

Conservative updating thus protects the agent against misspecifications and leads to *better predictions*. Predicting better, even perfectly, does not imply that the agent learns (identifies) the true data-generating process. In fact, the agent *cannot* learn the true data-generating process when his model is misspecified! Yet, when we interpret Θ as (payoff-relevant) states and X as signals, an implication of [Theorem 1](#) is that the agent’s

belief about the states may also come closer to the true state than Bayesian updating would. For an illustration, suppose that there are two states θ^* and θ and two signals ℓ and h . The true conditional distributions (experiment) are $p_{\theta^*}^*(h) = 1 - p_{\theta}^*(h) = 2/3$, while the agent's model is $p_{\theta^*}(h) = 1 - p_{\theta}(h) = 1/10$. Assume that the true state is θ^* . Since the agent wrongly interprets the signal h as evidence of θ and the signal h is the most likely when the state is θ^* , the Bayesian posterior converges almost surely to θ , the wrong state. In contrast, with the updating rule (C-Bay), the agent's belief converges to probability $7/24$ on θ^* , a closer distribution to the truth. (Its cross-entropy with respect to $\delta_{\{\theta^*\}}$ is $\log(7/24)$, while the cross-entropy of the limiting Bayesian posterior is $-\infty$.¹³) It is not always true that the non-Bayesian agent learns better, though. E.g., if we change $1/10$ for $9/10$, the agent learns the true state under Bayesian updating, while the belief converges to $17/24$ on θ^* under updating (C-Bay). In general, all we can say is that the agent's belief comes closer, and sometimes strictly so, to the true state under updating (C-Bay) (in the sense of the Kullback-Leibler divergence) if, and only if, the agent does not learn the true state under Bayesian updating.

The updating rule (C-Bay) may also serve as a “*misspecification test*.” Indeed, if the process $(q_n)_n$ does not converge to a point mass on one of the parameter θ , then the model \mathcal{P} must be misspecified.¹⁴ Thus, a possible test is to reject the hypothesis that the model is correctly specified if $\min_{\theta \in \Theta} \|q_n - \delta_{\{\theta\}}\|_{TV} > \varepsilon$, where ε and n are carefully calibrated to trade-off type I and II errors. We leave this question for future research.

The proof of Theorem 1 uses results from stochastic approximation theory with, however, one major difficulty. The major difficulty is that the convex components of E might neither be isolated points nor interior points. In fact, each component $C_k \neq C_{k^*}$ lies on the boundary of S . Example 2 illustrates that we may have a continuum of zeroes on the boundary.

Example 2 Suppose that $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, $X = \{x_1, x_2, x_3\}$, and

$$p_{\theta_1} = \frac{1}{4}(2, 1, 1), p_{\theta_2} = \frac{1}{4}(1, 2, 1), p_{\theta_3} = \frac{1}{4}(1, 1, 2), p_{\theta_4} = \frac{1}{8}(2, 3, 3), p^* = \frac{1}{8}(6, 1, 1).$$

The set E is

$$\{\delta_{\{\theta_1\}}\} \cup \{\delta_{\{\theta_2\}}\} \cup \{\delta_{\{\theta_3\}}\} \cup \{(0, \lambda/2, \lambda/2, 1 - \lambda) : \lambda \in [0, 1]\},$$

with $C_{k^*} = \{\delta_{\{\theta_1\}}\}$. The component $\{(0, \lambda/2, \lambda/2, 1 - \lambda) : \lambda \in [0, 1]\}$ is a continuum and

¹³We compare distributions in S . Since the parameter identifies distributions in $\Delta(X)$, we are equivalently comparing distributions in $\Delta(\Delta(X))$.

¹⁴More precisely, we need to add the mild requirement that the model is identified, i.e., $p_{\theta} \neq p_{\theta'}$ for all (θ, θ') . When $p^* = p_{\theta} = p_{\theta'}$ with $\theta \neq \theta'$, it is possible for the belief process to converge to an interior point, even though the model is correctly specified.

lies on the boundary of S – it is the line in Figure 3.

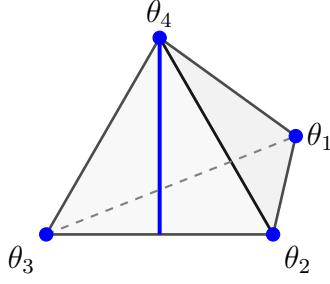


Figure 3: The components in Example 2

As a consequence, none of the classical results in stochastic approximation theory, e.g., Pemantle (1990), Brandiere (1998), Brandiere and Duflo (1996), are immediately applicable. For instance, Pemantle (1990, p. 700-701) states that “if there are uncountably many unstable points, then (q_n) can still converge to one of them even though each point has probability zero of being hit. In this case, all one might conclude is that the law of the limit has no point masses at unstable critical points¹⁵.” Similarly, since the components C_k , $k \neq k^*$, are on the boundary of S , we cannot rely on the Jacobian of H in the neighborhood of each $q \in C_k$ to determine their stability.

We develop novel arguments to tackle that issue. The arguments are quite technical, but the main logic is simple: We show that each component $C_k \neq C_{k^*}$ admits an open neighborhood and a *uniform repelling* direction, so that the dynamics cannot hop indefinitely often from elements of the open neighborhood of C_k to others.

We now present the arguments for the special case where $C_k \neq C_{k^*}$ consists of the single element \hat{q} . (Technical details in this paragraph may be omitted without loss of continuity.) Choose $q^* \in C_{k^*}$ with maximal support, that is, for all $q \in C_{k^*}$, $[q_\theta > 0 \Rightarrow q_\theta^* > 0]$. The *convexity* of each component guarantees that such q^* exists.¹⁶ We partition Θ into three subsets: $\Theta_1 := \{\theta \in \Theta : q_\theta^* = \hat{q}_\theta = 0\}$, $\Theta_2 := \{\theta \in \Theta : \hat{q}_\theta > 0\}$, and $\Theta_3 := \{\theta \in \Theta : \hat{q}_\theta = 0, q_\theta^* > 0\}$. From the construction of the components C_k , Θ_3 is non-empty. (If Θ_3 was empty, then $C_k = C_{k^*}$.) From the strict concavity of V , $\frac{\partial}{\partial \tau} V(\tau q^* + (1 - \tau)\hat{q}) > 0$, $\tau \in (0, 1)$. Therefore,

$$0 < \sum_{\theta \in \Theta} (q_\theta^* - \hat{q}_\theta) f_\theta(\hat{q}) = \sum_{\theta \in \Theta} q_\theta^* f_\theta(\hat{q}) - 1 = \sum_{\theta \in \Theta_2 \cup \Theta_3} q_\theta^* f_\theta(\hat{q}) - 1 = \sum_{\theta \in \Theta_2} q_\theta^* + \sum_{\theta \in \Theta_3} q_\theta^* f_\theta(\hat{q}) - 1,$$

where the first and third equalities follows from $\hat{q}_\theta (f_\theta(\hat{q}) - 1) = 0$ for all θ , since \hat{q} is a zero of H . (See Equation 2 for the definition of $f_\theta(\hat{q})$.) Since $\sum_{\theta \in \Theta_2 \cup \Theta_3} q_\theta^* = 1$, there exists

¹⁵Here, the term “critical points” is relative to the flow. In other terms, a critical point of the flow associated to the EDO $\dot{q} = H(q)$ is simply a zero of H .

¹⁶This is where having convex, and not merely connected, components helps.

$\hat{\theta} \in \Theta_3$ such that $f_{\hat{\theta}}(\hat{q}) > 1$, that is, $\dot{q}_{\hat{\theta}} > 0$ in the neighborhood of \hat{q} – the dynamics cannot converge to \hat{q} . The logic extends to arbitrary convex components.

The observation that the components are not necessarily isolated points is a direct consequence of the fact that distributions in $\text{co } \mathcal{P}$ do not always have a unique decomposition in terms of its extreme points \mathcal{P} . For future reference, let Λ be the set of all possible convex combinations of the p_{θ} 's equal to p^* , that is,

$$\Lambda := \left\{ \lambda = (\lambda_{\theta})_{\theta \in \Theta} \in \mathbf{S} : p^* = \sum_{\theta \in \Theta} \lambda_{\theta} p_{\theta} \right\}.$$

We now discuss the implications of assuming different notions of independences of the family \mathcal{P} . We say that the family $\mathcal{P} = \{p_{\theta} : \theta \in \Theta\}$ is *full* if the vectors $(p_{\theta} - p_{\theta'})_{(\theta, \theta') \in \Theta \times \Theta}$ span the tangent space associated to $\Delta(X)$, i.e., $\mathbb{R}_0^{|X|} := \{v \in \mathbb{R}^{|X|} : \sum_{x \in X} v_x = 0\}$. In words, if the family is full, then differences between its members can point in any direction allowed by the simplex structure. This property is common in statistics, as it allows for identification.

Proposition 1 *If the family \mathcal{P} is full, then $E \cap \mathbf{S}^* \subseteq \Lambda$. If, in addition, $p^* \notin \text{co } \mathcal{P}$, then $E \subseteq \partial(\mathbf{S})$.*

Thus, when the family is full, interior zeroes, if they exist, “identify” the true data-generating process p^* ; the agent would learn to predict correctly. We stress that the inclusion $E \cap \mathbf{S}^* \subseteq \Lambda$ may be strict, because Λ may intersect the boundary of \mathbf{S} , as the following example demonstrates.

Example 3 *Suppose that $\Theta = \{\theta_1, \theta_2, \theta_3\}$, $X = \{x_1, x_2\}$, and*

$$p_{\theta_1} = \frac{1}{4}(1, 3); p_{\theta_2} = \frac{1}{3}(1, 2); p_{\theta_3} = \frac{1}{4}(3, 1), p^* = \frac{1}{2}(1, 1).$$

Then, $E = \{\delta_{\{\theta_1\}}\} \cup \{\delta_{\{\theta_2\}}\} \cup \{\delta_{\{\theta_3\}}\} \cup \Lambda$, where

$$\Lambda = \{\lambda \in \mathbf{S} : \lambda_1 p_{\theta_1} + \lambda_2 p_{\theta_2} + \lambda_3 p_{\theta_3} = p^*\} = \left\{ \left(\lambda_1, \frac{1}{5}(3 - 6\lambda_1), \frac{1}{5}(2 + \lambda_1) \right) : \lambda_1 \in \left[0, \frac{1}{2} \right] \right\}.$$

We say that the family \mathcal{P} satisfies convex independence if there does not exist $\hat{\theta} \in \Theta$ such that $p_{\hat{\theta}} \in \text{co } \{p_{\theta} : \theta \in \Theta \setminus \{\hat{\theta}\}\}$. The next result states that all components are singletons when the family satisfies convex independence and $|\Theta| \leq |X|$.

Proposition 2 *If $|\Theta| \leq |X|$ and the family \mathcal{P} satisfies convex independence, then every C_k is a singleton.*

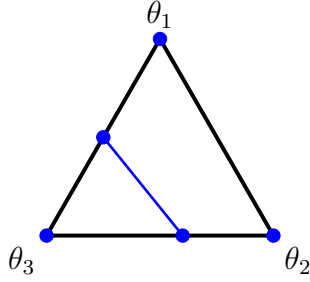


Figure 4: The four components in Example 3

Finally, we say that the family \mathcal{P} is *tight* if it is full and *minimal*, in the sense that any proper sub-family is not full. Note that a family is tight if and only if $|\Theta| = |X|$ and it satisfies convex independence. We have the following theorem when the family is tight.

Theorem 2 *Assume that the family \mathcal{P} is tight. Then, there exists a unique λ^* such that $(q_n)_n$ converges to λ^* almost surely. If $\Lambda \neq \emptyset$, then $\Lambda = \{\lambda^*\}$.*

4 Discussion

4.1 Updating Weights

Section 3 assumes that the updating weights $(\gamma_n)_n$ are deterministic and vanishing not too slowly. In line with some empirical evidences (Benjamin, 2019), these assumptions guarantees that the agent does not stop learning, but does so at a progressively slower rate. This section discusses alternative assumptions and their consequences.

We start with an observation. The updating weights can be random: Theorem 1 remains true if each random weight γ_n is measurable with respect to \mathcal{F}_{n-1} , $\mathbb{E}[\sum_n \gamma_n] = +\infty$, and $\mathbb{E}[\sum_n e^{-c/\gamma_n}] < +\infty$, for all $c > 0$.¹⁷ See Remark 4.3 of Benaïm (1999). We now turn our attention to two more substantive alternatives.

CONSTANT WEIGHTS. Assume that the updating weights are constant, i.e., $\gamma_n = \gamma \in (0, 1)$ for all n . With a constant weight, the agent continues to underreact to information, but does so at a constant rate. The analysis in Epstein et al. (2010) covers this case and shows that the agent eventually learns the data-generating process, that is, if p_{θ^*} is the true data-generating process, the agent's beliefs converge almost surely to $\delta_{\{\theta^*\}}$. The argument is standard. Under the true data-generating process p_{θ^*} , the random process $(\log q_n(\theta^*))_n$ is a sub-martingale, bounded from above, and thus converges almost surely by the martingale convergence theorem. To see that $(\log q_n(\theta^*))_n$ is a sub-martingale,

¹⁷We denote $(\mathcal{F}_n)_n$ the filtration adapted to the sequence $(q_n, x_n)_n$.

note that

$$\begin{aligned} \mathbb{E}_{p_{\theta^*}}[\log q_{n+1}(\theta^*)] - \log q_n(\theta^*) &= \mathbb{E}_{p_{\theta^*}} \left[\log \left((1 - \gamma) + \gamma \frac{p_{\theta^*}(x)}{\sum_{\theta'} q_n(\theta') p_{\theta'}(x)} \right) \right] \\ &\geq \gamma \sum_x p_{\theta^*}(x) \log \left(\frac{p_{\theta^*}(x)}{\sum_{\theta'} q_n(\theta') p_{\theta'}(x)} \right) \geq 0, \end{aligned}$$

where the first inequality follows from the concavity of \log and the second from the positivity of the relative entropy. It follows that $(q_n(\theta^*))_n$ converges almost surely to one.

However, when the model is misspecified (i.e., $p^* \notin \mathcal{P}$), the argument does not apply. Indeed, when the model is misspecified, the term

$$\gamma \sum_x p^*(x) \log \left(\frac{p_{\theta^*}(x)}{\sum_{\theta'} q_n(\theta') p_{\theta'}(x)} \right),$$

is not always positive. For some q , the predictive distribution $\sum_{\theta} q_{\theta} p_{\theta}$ is strictly closer to p^* than p_{θ^*} is, in which case the term above is strictly negative.

In fact, the belief process is not even guaranteed to converge. For an example, suppose that $X = \{0, 1\}$, $\Theta = \{\theta, \theta'\}$, $p^*(1) = 2/3$, $p_{\theta}(1) = 3/4 = p_{\theta'}(0)$. We plot the belief processes in Figure 5. The blue curve represents the predictive process induced by the updating rule (C-Bay) with a constant weight, while the black curve represents the true data-generating process. Clearly, the process does not converge, but oscillates around the true data-generating process. For completeness, we also plot the predictive process induced by the updating rule (C-Bay) with decreasing weights (resp., Bayesian updating) as the green (resp., purple) curve.

We therefore follow a different approach than in the previous section. We start by noting that the updating rule (C-Bay) induces a Markov chain on \mathbf{S} , where the probability to transition from q to q' is $\sum_{x \in X_{q,q'}} p^*(x)$, $X_{q,q'} := \{x \in X : q' = (1 - \gamma)q + \gamma B(q, x)\}$. From any q , the Markov chain transitions to finitely many q' . We then study the long-run properties of the Markov chain, which hold for almost all possible realizations of the observations. More precisely, we study the convergence properties of the sequences of occupation measures $(\Pi_n^\gamma)_n$, where for any Borel set $A \subset \mathbf{S}$,

$$\Pi_n^\gamma(A) := \frac{1}{n} \sum_{m=1}^n \delta_{q_m}(A).$$

The notation stresses that the chain and, therefore, its occupation measure is parameterized by γ .

We first prove that the sequences of occupation measures converge to invariant distributions $\pi \in \Delta(\Delta(\Theta))$ of the Markov chains. This result is, however, of limited value as

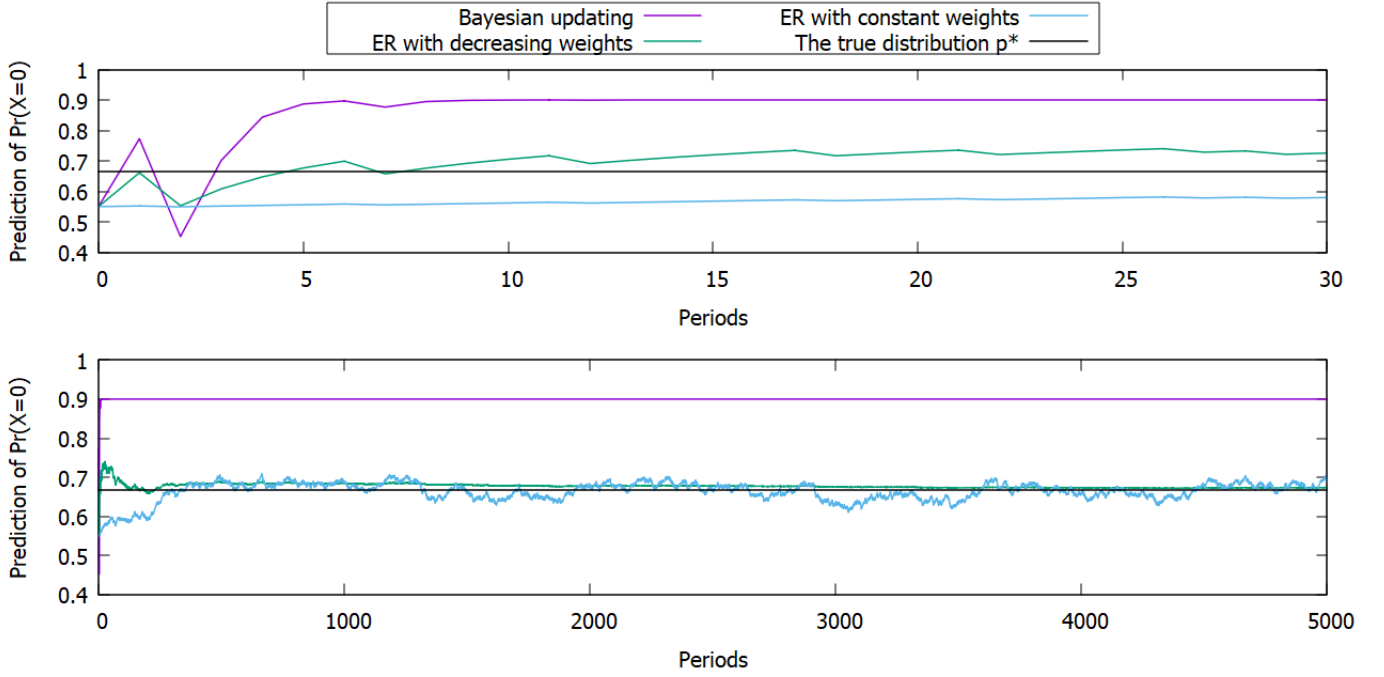


Figure 5: Constant weight and non-convergence

there are multiple invariant measures. For instance, for each $\theta \in \Theta$, $\delta_{\{\delta_{\{\theta\}}\}}$ is an invariant measure. We, however, prove that as γ becomes arbitrarily small, the occupation measure converges to an invariant measure, which concentrates on C_{k^*} . (Recall that C_{k^*} is the component of zeroes, which maximize the cross-entropy.) Thus, when the reaction to new information is small enough, the process is qualitatively similar to the process with decreasing weights studied in Section 3. We now turn to a formal definition and statement.

Definition 1 *The measure $\pi^* \in \Delta(\mathcal{S})$ is a limiting measure for the updating rule (C-Bay) if there exists a sequence $(\gamma_\ell, \pi_\ell)_{\ell \in \mathbb{N}}$ such that:*

1. $\gamma_\ell \downarrow 0$,
2. π_ℓ is a weak* limit point of $(\Pi_n^{\gamma_\ell})_n$, for all ℓ , and
3. $\lim_\ell \pi_\ell = \pi^*$ for the weak* topology.

In words, the measure π^* is a limiting measure if we can find a sequence of updating weights $(\gamma_\ell)_\ell$ and a corresponding sequence $(\pi_\ell)_\ell$ of limit points of $(\Pi_n^{\gamma_\ell})_n$, which converges to π^* for the weak* topology. We then prove the following:

Theorem 3 *Let π^* be a limiting measure for the updating rule (C-Bay). The support of π^* is included in C_{k^*} .*

Theorem 3 states that the limiting measure concentrates on C_{k^*} as the updating weight becomes smaller, the natural analogue of Theorem 1. To get some intuition, recall that

a limit point π_ℓ of $(\Pi_n^{\gamma_\ell})_n$ is an invariant measure of the Markov chain, parameterized by γ_ℓ . Since invariant measures capture the long-run stationary properties of the belief process, beliefs should not grow at the limit. More precisely, let

$$r_\theta^\gamma(q) := \sum_x p^*(x) \log \left(\underbrace{(1 - \gamma) + \gamma \frac{p_\theta(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)}}_{\approx \frac{q_{n+1}(\theta|x) - q_n(\theta)}{q_n(\theta)}} \right).$$

be the expected growth rate of the belief in θ , when the current belief is q . We show that

$$\int_{\mathbf{S}} \sum_{\theta} \lambda_{\theta} r_{\theta}^{\gamma_\ell}(q) \pi_\ell(dq) \leq 0, \quad (4)$$

for all positive weight $\lambda \in \mathbb{R}_+^{|\Theta|}$, for all limit point π_ℓ . Now, choose $\lambda^* \in C_{k^*}$. For any $q \notin C_{k^*}$, we have that

$$\sum_x p^*(x) \left(\frac{\sum_{\theta'} \lambda_{\theta'}^* p_{\theta'}(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)} - 1 \right) \geq \sum_x p^*(x) \log \left(\frac{\sum_{\theta'} \lambda_{\theta'}^* p_{\theta'}(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)} \right) > 0,$$

where the first inequality follows from the concavity of the log and the second from the definition of C_{k^*} . The left-hand side is the derivative of $\gamma \rightarrow \sum_{\theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q)$ evaluated at $\gamma = 0$. Hence, for all $q \notin C_{k^*}$, there exists $\gamma(q) > 0$ such that for all $\gamma \leq \gamma(q)$, $\sum_{\theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q) > 0$. Now, if we had a uniform $\gamma_0 > 0$ such that $\sum_{\theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q) > 0$ for all $q \notin C_{k^*}$, for all $\gamma \leq \gamma_0$, and Equation (4) was holding in equality, this would prove that no invariant distribution can put strictly positive probability outside the set C_{k^*} for all γ small enough. For instance, when C_{k^*} is the singleton $\{\lambda^*\}$, this would say that λ^* is an invariant distribution of the Markov chain for all γ small enough. This is not true in general. What we can show, however, is that the same logic applies to appropriately chosen neighborhoods of C_{k^*} . Choosing a sequence of such neighborhoods converging to C_{k^*} then completes the proof. In simple terms, for small enough γ_ℓ , the invariant measure π_ℓ puts most of its mass on C_{k^*} , but not all of it.

OBSERVATION-DEPENDENT WEIGHTS. Biases such as the self-confirmation bias require the weights of the updating rule (**C-Bay**) to depend on the realized signal x_n at period n and, possibly, the belief at period $n - 1$ (see, e.g., [Rabin and Schrag \(1999\)](#)). Now, if each $\gamma_n(x, q)$ is an arbitrary function of (x, q) , there is little hope in characterizing the belief process. As [Epstein et al. \(2010\)](#) have already shown, there are instances, where the process converges to a Dirac on a wrong state, even when the model is correctly specified. Similarly, as the blue curve in [Figure 6](#) shows, the process may oscillate perpetually.¹⁸

¹⁸To draw [Figure 6](#), we have assumed that $X = \{0, 1\}$, $\Theta = \{\theta, \theta'\}$, $p^*(1) = 2/3$, $p_\theta(1) = 9/10 = p_{\theta'}(0)$, and $\gamma_n(x, q) = 1$ if $(B(x, q) - q)(1/2 - q) > 0$, and $1/n$, otherwise, with $q \in [0, 1]$ the belief about θ . In words, the

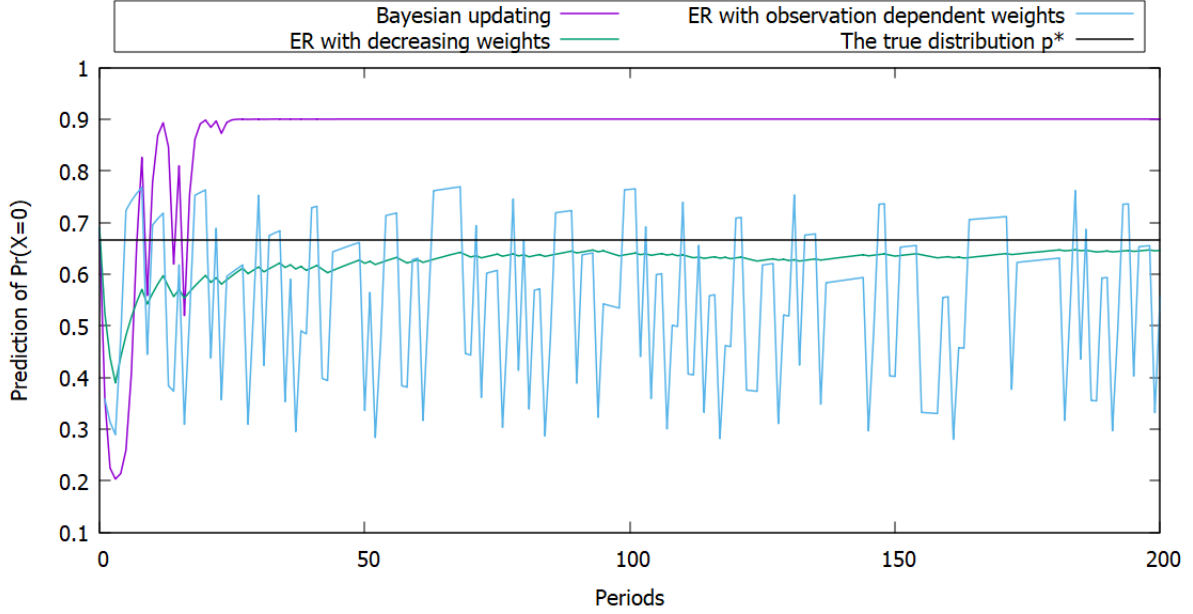


Figure 6: Observation-dependent weights and oscillation

There is, however, a special case we can deal with. Assume that the weights can be written as $\gamma_n(x) = \gamma_n \times \alpha(x)$, where $\alpha(x) \in (0, 1]$ for all $x \in X$. With this formulation, the weights are independent of the current belief. In addition, the relative weights $\gamma_n(x)/\gamma_n(y)$ are independent of the number n of observations. If the agent reacts more to the observation x than y , he continues to do so throughout. With this formulation, we can rewrite the Robbins-Monro algorithm as:

$$q_{n+1} - q_n = \gamma_{n+1} \left(\widehat{H}(q_n) + \widehat{U}_{n+1} \right),$$

where \widehat{H} is the function from \mathbf{S} to $\mathbb{R}^{|\Theta|}$, whose θ -coordinate is

$$\widehat{H}_\theta(q) := q_\theta \left(\sum_x p^*(x) \alpha(x) \left(\frac{p_\theta(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)} - 1 \right) \right),$$

and $(\widehat{U}_n)_n$ the bounded martingale difference

$$\widehat{U}_{n+1} := \alpha(x_{n+1}) B(q_n, x_{n+1}) - \mathbb{E}_{p^*} [\alpha(x_{n+1}) B(q_n, x_{n+1})].$$

The analysis is then identical to the one in Section 3. In particular, if we denote $p_\alpha^*(x) := \frac{p^*(x)\alpha(x)}{\sum_{x'} p^*(x')\alpha(x')}$ the *as-if* data generating process, it is immediate to see that the predictive process converges almost surely to:

$$\arg \max_{p \in \text{co } \mathcal{P}} \sum_{x \in X} p_\alpha^*(x) \log p(x).$$

agent updates as a Bayesian when the observation contradicts his belief and is conservative, otherwise.

In the special case where $p_\alpha^* \in \text{co } \mathcal{P}$, the predictive process converges to p_α^* , which can be closer to p^* than any of the p_θ 's are.

Example 4 *There are two states A and B , two possible observations a and b , and $p_A(a) > p_B(a)$, that is, observing a (resp., b) is more likely when the state is A (resp., B). Assume that the agent initially believes that A is more likely, i.e., $q_0(A) > 1/2$ and let $\alpha(a) > \alpha(b)$. This simple setting captures a form of prior dependence, where the agent reacts more to information that confirms his initial belief. It is an extreme form of prior dependence, though, as the agent continues to react more to observation a , even when his current belief q_n indicates that B is the most likely state. The as-if data-generating process is:*

$$p_\alpha^*(a) = \frac{p^*(a)\alpha(a)}{p^*(a)\alpha(a) + p^*(b)\alpha(b)} > p^*(a),$$

i.e., it is as-if the agent learns in an environment where the data-generating process is distorted towards a , the observation most indicative of A . For a concrete example, let $p^(a) = 2/3$, $p_A(a) = p_B(b) = 3/4$, $\alpha(a) = 3/4$ and $\alpha(b) = 1/4$, so that $p_\alpha^*(a) = 6/7$. The agent's predictive process converges to p_A , the same limiting predictive process a Bayesian agent would converge to. If, however, the prior dependence is weaker, say $\alpha(a) = 2/3$ and $\alpha(b) = 1/2$, then the predictive process converges to $(8/11, 3/11)$, which is strictly closer to the true data-generating process than p_A .*

So far, we have assumed that the weights are independent of the current belief q_n . Within the context of Example 4, suppose instead that we write the updating weights as $\gamma_{n+1} \times \alpha(x_{n+1}, q_n)$, with $\alpha(a, q) = \alpha^+ = \alpha(b, 1 - q)$ when $q > 1/2$, $\alpha(a, 1 - q) = \alpha^- = \alpha(b, q)$ when $q < 1/2$, and $1 \geq \alpha^+ > \alpha^- \geq 0$. When $q = 1/2$, let $\alpha(a, 1/2) = \alpha(b, 1/2) = \alpha$. Thus, if the agent observes a (resp. b) when he currently believes A (resp. B) to be the most likely state, he revises his belief towards the Bayesian posterior more than when he observes b (resp., a). This is a weaker form of “prior” dependence: The agent reacts more to information at period $n + 1$ that confirms the belief q_n they had at the end of period n . This updating process continues to be a Robbins-Monro algorithm with the function \widehat{H} :

$$\widehat{H}(q) := \begin{cases} p^*(a) \left[\frac{p_A(a)q}{p_A(a)q + p_B(a)(1-q)} \right] \alpha^+ + p^*(b) \left[\frac{p_A(b)q}{p_A(b)q + p_B(b)(1-q)} \right] \alpha^-, & q > 1/2, \\ p^*(a) \left[\frac{p_A(a)}{p_A(a) + p_B(a)} \right] \alpha + p^*(b) \left[\frac{p_A(b)}{p_A(b) + p_B(b)} \right] \alpha, & q = 1/2, \\ p^*(a) \left[\frac{p_A(a)q}{p_A(a)q + p_B(a)(1-q)} \right] \alpha^- + p^*(b) \left[\frac{p_A(b)q}{p_A(b)q + p_B(b)(1-q)} \right] \alpha^+, & q < 1/2. \end{cases}$$

For a concrete example, let $p^*(a) = 2/3$, $p_A(a) = p_B(b) = 3/4$, $\alpha^+ = 3/4$, $\alpha^- = 1/4$, and

$\alpha = 1/2$. The function \widehat{H} is then

$$\widehat{H}(q) := \begin{cases} \frac{2}{3} \left[\frac{3q}{1+2q} - q \right] \frac{3}{4} + \frac{1}{3} \left[\frac{q}{3-2q} - q \right] \frac{1}{4} & q > 1/2, \\ \frac{1}{24} & q = 1/2, \\ \frac{2}{3} \left[\frac{3q}{1+2q} - q \right] \frac{1}{4} + \frac{1}{3} \left[\frac{q}{3-2q} - q \right] \frac{3}{4} & q < 1/2. \end{cases}$$

We plot the function \widehat{H} in Figure 7. It is immediate to verify that \widehat{H} has three zeroes: $q = 0$, $q = 3/10$, and $q = 1$, the latter two being locally stable.

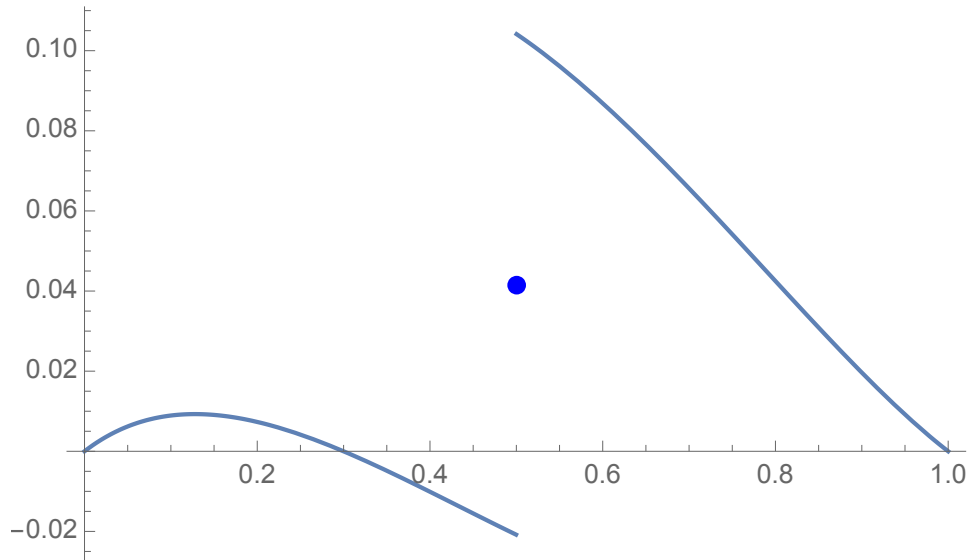


Figure 7: The function \widehat{H}

The local stability of $q = 3/10$ and $q = 1$ is not surprising. Since $\alpha(x, \cdot)$ is constant in q for all $q < 1/2$, the cross-entropy

$$\sum_{x \in \{a,b\}} \frac{p^*(x)\alpha(x,0)}{\sum_{x' \in \{a,b\}} p^*(x')\alpha(x',0)} \log(qp_A(x) + (1-q)p_B(x))$$

is maximized with respect to $q \in [0, 1/2)$ in a stable zero of \widehat{H} , restricted to that sub-domain. It is immediate to verify that $q = 3/10$ is indeed the maximizer. A symmetric argument applies for all $q > 1/2$.

To sum up, our main insight, that some non-Bayesian updating can outperform Bayesian updating in misspecified environments, is not limited to the setting in Section 3.

4.2 Generalized Bayes Rule

A related updating rule is the generalized Bayes rule, where the revised belief $q_{n+1}(\theta)$ is proportional to

$$q_n(\theta)p_\theta(x_{n+1})^{\gamma_{n+1}},$$

when the agent observes x_{n+1} , $\gamma_{n+1} \in [0, 1]$. The generalized Bayes rule is a special case of the [Grether \(1980\)](#) rule, captures some form of underreaction to new information and corresponds to Bayes rule when $\gamma_{n+1} = 1$. Recent contributions in statistics, e.g., [Grünwald and Van Ommen \(2017\)](#) and [Grünwald and Mehta \(2020\)](#), advocate the use of the generalized Bayes rule, particularly as a method to correct for inconsistencies in Bayesian inference problems.

We now argue that its long-run properties are markedly different from the rule (C-Bay). To ease the discussion, assume that there exists a unique p_{θ^*} , which maximizes the cross-entropy V , that is, $V(p_{\theta^*}) > V(p_\theta)$ for all $\theta \neq \theta^*$. When the observations up to period $n + 1$ are (x_1, \dots, x_{n+1}) , we have that

$$\begin{aligned} \log \left(\frac{q_{n+1}(\theta^*)}{q_{n+1}(\theta)} \right) &= \log \left(\frac{q_0(\theta^*)}{q_0(\theta)} \right) + \sum_{i=1}^{n+1} \gamma_i \log \left(\frac{p_{\theta^*}(x_i)}{p_\theta(x_i)} \right), \\ &\geq \log \left(\frac{q_0(\theta^*)}{q_0(\theta)} \right) + \\ &\quad n \left[\left(\max_{i=1, \dots, n+1} \gamma_i \right) \frac{1}{n+1} \sum_{i=1}^{n+1} \log p_{\theta^*}(x_i) - \left(\min_{i=1, \dots, n+1} \gamma_i \right) \frac{1}{n+1} \sum_{i=1}^{n+1} \log p_\theta(x_i) \right]. \end{aligned}$$

The strong law of large numbers implies that $\lim_{n \rightarrow +\infty} \frac{1}{n+1} \sum_{i=1}^{n+1} \log p_{\hat{\theta}}(x_i) = V(p_{\hat{\theta}})$ almost surely for all $\hat{\theta}$, hence the term into bracket converges to a strictly positive number since $V(p_{\theta^*}) > V(p_\theta)$. Therefore, $\left(\log \left(\frac{q_{n+1}(\theta^*)}{q_{n+1}(\theta)} \right) \right)_n$ converges to $+\infty$ and, consequently, $(q_{n+1}(\theta^*))_n$ converges to one almost surely. The generalized Bayes rule thus converges to the same degenerated belief as Bayes rule, in sharp contrast with the updating rule (C-Bay).

Thus, despite the similarities of the two rules, their asymptotic properties are fundamentally different. Intuitively, this is because the two rules underreact to new information on two different scales (linear vs. logarithmic), with the updating rule (C-Bay) underreacting more and, consequently, protecting more against model's misspecification.

4.3 Overreaction

We have assumed that the agent underreacts to news, i.e., $\gamma_n < 1$. If the agent overreacts to news, the convergence to a limit point or cycle is not guaranteed in general. Intuitively, overreaction may lead to big swings in beliefs, which prevents them from stabilizing. For an illustration, see Figure 8. Clearly, the belief process (the green curve) is quite chaotic, with the beliefs varying wildly.¹⁹ There is therefore little hope for general statements. We conjecture, however, that if $(\gamma_n)_n$ converges to 1, then we recover the convergence to the closest distribution in \mathcal{P} to p^* .

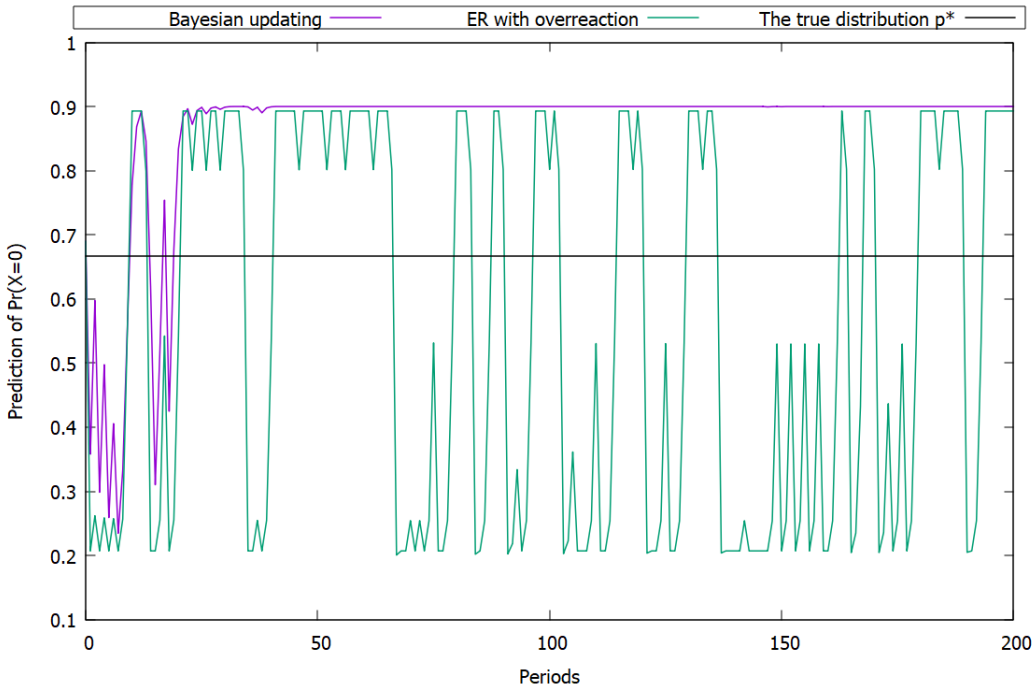


Figure 8: Overreaction and Oscillation

5 Literature

Two closely related papers are [Heidhues et al. \(2021\)](#) and [Frick et al. \(2023\)](#). [Heidhues et al. \(2021\)](#) consider the following model. In each period n , the agent makes a decision a_n and observes a signal q_n , which depends on the decision a_n and an (unobservable) external factor x_n . The authors assume that $x_n = \theta + \epsilon_n$, where θ is a fixed parameter and ϵ_n a realization of a normal distribution with mean zero and precision ρ_ϵ . (The realizations are independent.) When the agent makes the decision a_n and the external factor is x_n , the signal is $q_n = Q(a_n, x_n)$, with Q strictly monotonic in x for all a . The agent, however, perceives the relationship between decisions, external factors and signals to be \widehat{Q} , also

¹⁹To draw Figure 8, we have assumed that $X = \{0, 1\}$, $\Theta = \{\theta, \theta'\}$, $p^*(1) = 2/3$, $p_\theta(1) = 9/10 = p_{\theta'}(0)$, and $\gamma_n = 2 + 1/n$. In instances where $(1 - \gamma_n)q_n + \gamma_n B(x_{n+1}, q_n) > 1$ (resp., < 0), we let $q_{n+1} = 0.99$ (resp., 0.01).

strictly monotonic in x for all a . The agent is Bayesian and initially believes that the unknown mean θ is normally distributed with mean zero and precision ρ_0 . Thus, upon observing q_n after having chosen a_n , the agent thinks that the realized external factor is \hat{x}_n , the unique solution to $\hat{Q}(a_n, x) = q_n$, which may differ from the truly realized x_n . At period n , the agent's expectation $\hat{\theta}_n$ of the mean therefore satisfies:

$$\hat{\theta}_n = \hat{\theta}_{n-1} + \gamma_n[\hat{x}_n - \hat{\theta}_{n-1}], \text{ where } \gamma_n = \frac{\rho_\epsilon}{n\rho_\epsilon + \rho_0}.$$

[Heidhues et al. \(2021\)](#) rewrite the evolution of $\hat{\theta}_n$ as a Robbins-Monro algorithm and prove almost sure convergence. (This discussion sidesteps an important issue, namely that, \hat{x}_n depends on the decision made a_n .) While [Heidhues et al. \(2021\)](#) frame their problem as an active learning problem, they also explain how we can reframe it as a learning problem with non-Bayesian updating. For instance, the model accommodates underreaction for specifications of Q and \hat{Q} such that $x_n > \hat{x}_n > 0$ and $x_n < \hat{x}_n < 0$. The model can also accommodate base-rate neglect and the confirmatory bias. If we interpret any deviation from x_n as a departure from Bayesian updating, as the authors do, then Bayesian updating corresponds to $x_n = \hat{x}_n$, that is, for all decisions made, the agent's model is correct (and, thus, learns θ). Unlike our framework, their framework is not flexible enough to cleanly decouple the updating rule from the agent's misspecification. We stress that non-Bayesian updating in misspecified models was not their focus. We differ in two other important ways. First, we do not restrict attention to a particular family of distributions, such as the normal distributions. Second, we not only prove that the belief process converges almost surely, but also fully characterize the limit set. [Heidhues et al. \(2021\)](#) only prove convergence. In common, we use the ODE method to characterize the limit set of our stochastic approximation algorithm. See [Bervoets et al. \(2020\)](#), [Esponda et al. \(2021\)](#), [Gagnon-Bartsch and Bushong \(2022\)](#) and [Danenberg and Fudenberg \(2024\)](#) for some other recent applications of stochastic approximation in Economics and [Marcet and Sargent \(1989\)](#) for an earlier application.

[Frick et al. \(2023\)](#) propose a general model of Bayesian learning, where the true and perceived data generating processes $p_q^* \in \Delta(X)$ and $p_{\theta,q} \in \Delta(X)$, $\theta \in \Theta$, may depend on the current belief $q \in \mathbf{S}$. The model is flexible enough to accommodate some departures from Bayesian updating as well as several models of active and social learning (in the latter case the dependence on the current beliefs is through the actions chosen). In particular, as we have already argued, the updating rule (**C-Bay**) induces the same posteriors than the Bayesian updating of the “as-if” data generating processes $p_{\theta,q}^\gamma \in \Delta(X)$, where $p_{\theta,q}^\gamma(x) := (1 - \gamma)(\sum_{\theta'} q_{\theta'} p_{\theta'}(x)) + \gamma p_\theta(x)$ for all (x, θ) , when the current belief is q . The results of [Frick et al. \(2023\)](#) are not applicable to our problem, however. First, their model cannot capture weights that depend on the number n of observations, as we

do. Second, and most importantly, their results provide conditions for the (in)stability of degenerated beliefs δ_θ , $\theta \in \Theta$. None of their results addresses the stability properties of interior beliefs. For other related papers on learning in misspecified models, see [Esponda and Pouzo \(2016\)](#), [Fudenberg et al. \(2017\)](#), [Heidhues et al. \(2018\)](#), [Frick et al. \(2020\)](#), [Fudenberg et al. \(2021\)](#), [Bohren and Hauser \(2021\)](#), and [Lanzani \(2022\)](#).

Relatedly, [Frick et al. \(2024\)](#) compare misspecified models in the spirit of [Blackwell \(1951\)](#). More precisely, consider two statistical models $\mathcal{P}^1 := (p_\theta^1)_\theta$ and $\mathcal{P}^2 := (p_\theta^2)_\theta$ and assume that the true model is $\mathcal{P}^* := (p_\theta^*)_\theta$. The idea is to compare the expected payoff the agent would obtain if his decisions are based on either model. Formally, let \mathcal{F} be a non-empty set of payoff acts $f : \Theta \rightarrow \mathbb{R}$ and $f_{\mathcal{F}}^i(x_1, \dots, x_n)$ a solution to $\max_{f \in \mathcal{F}} \mathbb{E}[f | (x_1, \dots, x_n), \mathcal{P}^i]$, that is, $f_{\mathcal{F}}^i(x_1, \dots, x_n)$ maximizes the agent's expected payoff given the observations (x_1, \dots, x_n) and the model \mathcal{P}^i . (The authors assume Bayesian updating, but as already explained, some choices of \mathcal{P}^i are equivalent to non-Bayesian updating.) The model \mathcal{P}^1 outperforms the model \mathcal{P}^2 if there exists n^* such that for all $n \geq n^*$, $\mathbb{E}_{\mathcal{P}^*}[f_{\mathcal{F}}^1(x_1, \dots, x_n)] \geq \mathbb{E}_{\mathcal{P}^*}[f_{\mathcal{F}}^2(x_1, \dots, x_n)]$ for all decision problems \mathcal{F} . [Frick et al. \(2024\)](#) provide tight conditions for one model to outperform the other. There are no difficulties to adapt their definitions to our framework. It is, however, unclear how to adapt their results, since the updating rule [\(C-Bay\)](#) necessitates a model, which changes with the current belief and the number of observations. Intuitively, though, when $p^* \in \text{co } \mathcal{P}$, the updating rule [\(C-Bay\)](#) converges to the true data-generating process, while Bayesian updating does not. For large enough n , the updating rule [\(C-Bay\)](#) therefore outperforms Bayesian updating, in the sense of [Frick et al. \(2024\)](#). Even when $p^* \notin \text{co } \mathcal{P}$, the updating rule [\(C-Bay\)](#) comes closer to the true data-generating process, hence there are decision problems (e.g., scoring problems), where it again outperforms Bayesian updating.

Finally, we have established that the updating rule [\(C-Bay\)](#) induces (random) sequences of beliefs converging to solutions of

$$\max_{q \in \Delta(\Theta)} \sum_{x \in X} p^*(x) \ln \left(\sum_{\theta} q_{\theta} p_{\theta}(x) \right). \quad (\mathcal{M})$$

This maximization problem (\mathcal{M}) has a well-known antecedent in Information Theory and Finance, where it is known as the [Kelly \(1956\)](#) problem.²⁰ In Kelly model, there is an investor with initial wealth W_0 , $|\Theta|$ assets, and $|X|$ states, with p^* the distribution over states. States are i.i.d drawn. Asset θ returns $p_{\theta}(x)$, when the state is x . At each period, the investor chooses a portfolio q , with q_{θ} the fraction of the current wealth allocated to asset θ . Kelly assumes that the investor's objective is to maximize the long-term growth rate of wealth and shows that there exists an optimal stationary strategy

²⁰We thank Jakub Steiner and Larry Samuelson for drawing our attention to this literature.

that consists in choosing the portfolio q^* at each period, where q^* is a solution to (\mathcal{M}) . For a detailed exposition, see [Cover and Thomas \(2006\)](#).

More closely related to our work, [Cover \(1984\)](#) and [Cover and Gluss \(1986\)](#) develop algorithms for solving (\mathcal{M}) . Neither, however, coincides with the updating rule (C-Bay). [Cover \(1984\)](#) assumes that the distribution p^* is known and proposes a gradient-ascent procedure. The procedure depends on p^* and, thus, differs from the procedure (C-Bay). Like us, [Cover and Gluss \(1986\)](#) assume that p^* is unknown. Their approach, however, differs substantially from ours. Roughly, their algorithm computes sequences of solutions to (\mathcal{M}) , substituting p^* for the corresponding empirical distributions.

6 Proofs

6.1 Proof of Theorem 1

We equip the space S with the metric $d(q, q') := \max_{\theta \in \Theta} |q_\theta - q'_\theta|$.

The proof uses methods and results from stochastic approximation theory. We refer the reader to [Benaïm \(1999\)](#) for a short exposition and to [Borkar \(2008\)](#) and [Kushner and Yin \(2003\)](#) for book-length treatments.

The roadmap is as follows. We first rewrite the updating rule (C-Bay) as a Robbins-Monro algorithm, with associated ODE $\dot{q}(t) = H(q(t))$. We then analyze the associated ODE. In particular, we show that the cross-entropy map V is a Lyapunov function for the associated flow and that the set E can be decomposed into finitely many closed convex (hence, connected) components (the decomposition is unique). By Proposition 6.4 in [Benaïm \(1999\)](#), the (random) limit set of $(q_n)_n$ is included in a closed connected subset of E . Finally, we show that the probability to converge to the component that maximizes the cross-entropy is one.

We rewrite the updating rule (C-Bay) as:

$$\begin{aligned} q_{n+1} &= q_n + \gamma_{n+1} [B(q_n, x_{n+1}) - q_n] \\ &= q_n + \gamma_{n+1} [H(q_n) + U_{n+1}], \end{aligned}$$

where $(U_n)_n$ is the bounded martingale difference with $U_{n+1} := B(q_n, x_{n+1}) - \mathbb{E}_{p^*}[B(q_n, \mathbf{x}_{n+1})]$ when the realized signal is x_{n+1} , and $H(q_n) := \mathbb{E}_{p^*}[B(q_n, \mathbf{x}_{n+1})] - q_n$.

The ODE associated with the random process $(q_n)_n$ is

$$\dot{q}(t) = H(q(t)). \tag{5}$$

For further reference, note that the component-wise system is given by

$$\dot{q}_\theta(t) = q_\theta(t)(-1 + f_\theta(q(t))), \text{ with } f_\theta(q) := \sum_x p^*(x) \frac{p_\theta(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)}, \forall \theta \in \Theta. \quad (6)$$

We first show that the cross-entropy map V grows along the trajectory of the dynamics (5).

Lemma 1 *V is a strict Lyapunov function for the flow ϕ_t associated to (5), meaning that, along any non-stationary solution curve $(q(t))_{t \geq 0}$ we have $\frac{d}{dt}V(q(t)) > 0$.*

Proof of Lemma 1 Note that

$$\frac{d}{dt}V(q(t)) = \langle \nabla V(q(t)), \dot{q}(t) \rangle = G(q(t))$$

for all $t \geq 0$, where $G(q) := \sum_\theta \frac{\partial V}{\partial q_\theta}(q) q_\theta (-1 + f_\theta(q))$.

Since

$$\frac{\partial V}{\partial q_\theta}(q) = \sum_{x \in X} p^*(x) \frac{p_\theta(x)}{\sum_{\theta' \in \Theta} p_{\theta'}(x) q_{\theta'}} = f_\theta(q),$$

we have

$$G(q) = \sum_{\theta \in \Theta} q_\theta f_\theta(q) (-1 + f_\theta(q)) = \sum_{\theta \in \Theta} q_\theta (-1 + f_\theta(q))^2 + \sum_{\theta \in \Theta} q_\theta f_\theta(q) - \sum_{\theta \in \Theta} q_\theta.$$

Notice that $\sum_{\theta \in \Theta} q_\theta f_\theta(q) = 1$, and since $\sum_{\theta \in \Theta} q_\theta = 1$, we have $G(q) \geq 0$, with equality if and only if $f_\theta(q) = 1$ for all θ such that $q_\theta > 0$. Hence, $G(q) \geq 0$ with equality if and only if $q \in E$, which concludes the proof. \blacksquare

Lemma 1 implies that the only invariant sets for the flow are connected components of E , the set of zeroes of H . By Proposition 6.4 in [Benaïm \(1999\)](#), it also implies that $\mathcal{L}(q_n)$, the (random) limit set of $(q_n)_n$, is included in a closed connected subset of E .²¹ We now characterize E .

Lemma 2 (i) *The set E is a finite union of disjoint non-empty compact and convex components: there exists $K < +\infty$ such that $E = \bigcup_{k=1}^K C_k$, where each C_k is a non-empty compact convex subset of \mathbb{S} and $C_k \cap C_{k'} = \emptyset$ for all (k, k') with $k \neq k'$.*
(ii) *The cross-entropy V takes the constant value $V(C_k)$ on each component C_k , and there exists a unique k^* such that $V(C_{k^*}) > V(q)$ for all $q \notin C_{k^*}$.*

²¹The decomposition of a set into its connected components is always unique.

Proof of Lemma 2 For any non-empty subset $\widehat{\Theta}$ of Θ , we write $F_{\widehat{\Theta}}$ for the subset of \mathbf{S} :

$$F_{\widehat{\Theta}} := \left\{ q \in \mathbf{S} : q_{\theta} = 0 \text{ iff } \theta \notin \widehat{\Theta} \right\}.$$

In particular, $F_{\Theta} = \mathbf{S}^*$ and $F_{\{\theta\}} = \{\delta_{\theta}\}$.

Recall that $q \in E$ if and only if

$$q_{\theta}(f_{\theta}(q) - 1) = 0,$$

for all θ . Consider the constrained maximization problem $\max_{q \in F_{\widehat{\Theta}}} V(q)$, that is,

$$\max_{q \in \mathbb{R}^{|\Theta|}} V(q),$$

subject to $\sum_{\theta \in \widehat{\Theta}} q_{\theta} = 1$, $q_{\theta} > 0$ for all $\theta \in \widehat{\Theta}$ and $q_{\theta} = 0$ for all $\theta \notin \widehat{\Theta}$. If a solution exists, for all $\theta \in \widehat{\Theta}$, the first-order condition with respect q_{θ} satisfies

$$f_{\theta}(q) := \sum_x p^*(x) \frac{p_{\theta}(x)}{\sum_{\theta'} q_{\theta'} p_{\theta'}(x)} = \lambda,$$

where λ is the Kuhn-Tucker multiplier associated with $\sum_{\theta \in \widehat{\Theta}} q_{\theta} = 1$. (The complementary slackness condition implies that the multipliers associated with $q_{\theta} > 0$ for all $\theta \in \widehat{\Theta}$ are all zero.) Multiplying each equation by q_{θ} and summing over $\theta \in \widehat{\Theta}$, we get

$$\sum_{\theta \in \Theta} q_{\theta} f_{\theta}(q) = \lambda \sum_{\theta \in \Theta} q_{\theta} = \lambda.$$

Since the left hand-side is equal to 1, it follows that

$$f_{\theta}(q) = 1 \quad \text{for all } \theta \in \widehat{\Theta}.$$

Therefore, at a maximum, we have

$$q_{\theta}(f_{\theta}(q) - 1) = 0,$$

for all $\theta \in \Theta$, i.e., it is a zero of H . Since V is concave, the converse is also true by the Kuhn-Tucker-Karush theorem. It follows that E is the union of the sets $S_{\widehat{\Theta}} := \arg \max_{q \in F_{\widehat{\Theta}}} V(q)$ over all possible non-empty subsets $\widehat{\Theta}$ of Θ .

Since V is concave, $S_{\widehat{\Theta}}$ is convex for all $\widehat{\Theta}$. We write $V(S_{\widehat{\Theta}})$ for the value taken by V on $S_{\widehat{\Theta}}$. We now construct the decomposition from the sets $S_{\widehat{\Theta}}$.

Let $T := \left\{ \widehat{\Theta} \subseteq \Theta : S_{\widehat{\Theta}} \neq \emptyset \right\}$. Given $(\widehat{\Theta}, \widehat{\Theta}') \in T \times T$ say that $\widehat{\Theta} \sim \widehat{\Theta}'$ if there exists a family

$\{\widehat{\Theta}_k\}_{k=0,\dots,K}$ such that $\widehat{\Theta} = \widehat{\Theta}_0$, $\widehat{\Theta}' = \widehat{\Theta}_K$, and $\text{dist}(S_{\widehat{\Theta}_k}, S_{\widehat{\Theta}_{k+1}}) = 0$ for $k = 0, \dots, K-1$.²² The binary relation \sim is an equivalence relation. Let $(C_k)_{k=1,\dots,K}$ be its equivalence classes (note that $K \leq 2^{|\Theta|} - 1$) and $C_k := \bigcup_{\widehat{\Theta} \in C_k} S_{\widehat{\Theta}}$. To conclude the proof, we need to show that C_k is closed and convex, and that the global maximum of V is attained on only one of these sets.

We first show that C_k is closed: let $(q^i)_i$ be a sequence in C_k and suppose that $\lim_{i \rightarrow +\infty} q^i = q$, with $q \in F_{\widehat{\Theta}}$. By continuity, $V(q) = V(C_k)$. Moreover, for i large enough, $q_\theta^i > 0$, for all $\theta \in \widehat{\Theta}$. Assume by contradiction that $q \notin S_{\widehat{\Theta}}$, then there exists $\tilde{q} \in F_{\widehat{\Theta}}$ such that $V(\tilde{q}) > V(q)$. There exists a sequence \tilde{q}^i such that \tilde{q}^i is in the same face as q^i , and $\lim_{i \rightarrow +\infty} \tilde{q}^i = \tilde{q}$. Therefore

$$\lim_{i \rightarrow +\infty} V(q^i) = V(q) < V(\tilde{q}) = \lim_{i \rightarrow +\infty} V(\tilde{q}^i),$$

implying that, for large enough i , $V(\tilde{q}^i) > V(q^i)$. This contradicts the fact that q^i maximizes V on the face it belongs to. Therefore, $q \in S_{\widehat{\Theta}}$ for some $\widehat{\Theta} \in C_k$. Consequently, the collection $\{C_k\}$ is composed of closed, connected, pairwise disjoint sets, and

$$E = \bigcup_{k=1}^K C_k. \quad (7)$$

We now argue that C_k is convex, for $k = 1, \dots, K$. Define $W : \Delta(X) \rightarrow \mathbb{R}$ as follows:

$$p \in \Delta(X) \mapsto W(p) := \sum_x p^*(x) \log p(x).$$

Since p^* has full support on X , W is strictly concave on $\Delta(X)$.²³ Let $L : \mathbf{S} \rightarrow \Delta(X)$ be the linear map $q \mapsto (\sum_\theta q_\theta p_\theta(x))_{x \in X}$. By linearity, $L(S_{\widehat{\Theta}})$ is a convex subset of $\Delta(X)$. By strict concavity of W , W is constant on $L(S_{\widehat{\Theta}})$ if, and only if, $L(S_{\widehat{\Theta}})$ is a singleton. Since V is constant on $S_{\widehat{\Theta}}$, and $V = W \circ L$, this implies that $L(S_{\widehat{\Theta}})$ is a singleton. Call $p_{\widehat{\Theta}} := \{L(S_{\widehat{\Theta}})\} \in \Delta(X)$. By continuity of L , $p_{\widehat{\Theta}} = p_{\widehat{\Theta}'}$ for $(\widehat{\Theta}, \widehat{\Theta}') \in C_k \times C_k$, so that L is constant on C_k (notice that this is stronger than stating that V is constant along C_k). Let $(q, q') \in C_k \times C_k$, with $q \in S_{\widehat{\Theta}}$, $q' \in S_{\widehat{\Theta}'}$, and $\lambda \in (0, 1)$. Then, if $q'' := \lambda q + (1 - \lambda)q'$, we have $L(q'') = L(\lambda q + (1 - \lambda)q') = \lambda L(q) + (1 - \lambda)L(q') = L(C_k)$. Since q and q' are zeroes of H , we have

$$1 = f_\theta(q) = \sum_x p^*(x) \frac{p_\theta(x)}{L(C_k)(x)}, \quad \forall \theta \in \widehat{\Theta}, \quad \text{and} \quad 1 = f_\theta(q) = \sum_x p^*(x) \frac{p_\theta(x)}{L(C_k)(x)}, \quad \forall \theta \in \widehat{\Theta}'.$$

²²The distance is the Hausdorff distance between subsets of \mathbf{S} , with \mathbf{S} equipped with the total variation metric.

²³More generally, if the support of p^* is not X , the argument remains valid if we restrict W to the support of p^* .

Now, let θ be such that $q''_\theta \neq 0$. then either $q_\theta \neq 0$ or $q'_\theta \neq 0$. In other terms, $\theta \in \widehat{\Theta} \cup \widehat{\Theta}'$. Since $\lambda \in (0, 1)$, $q''_\theta > 0$ if and only if $\theta \in \widehat{\Theta} \cup \widehat{\Theta}'$. Moreover,

$$f_\theta(q'') = \sum_x p^*(x) \frac{p_\theta(x)}{L(q'')(x)} = \sum_x p^*(x) \frac{p_\theta(x)}{L(C_k)(x)} = 1,$$

for all $\theta \in \widehat{\Theta} \cup \widehat{\Theta}'$. Hence, q'' is a zero of H in $S_{\widehat{\Theta} \cup \widehat{\Theta}'}$ and, consequently, $\text{co}(C_k) \subseteq E$. It follows that $\cup_{k'} C_{k'} = E = \cup_{k'} \text{co}(C_{k'})$. Now, if $C_k \neq \text{co}(C_k)$, then $\text{co}(C_k) \subseteq \cup_{k' \neq k} C_{k'}$, a contradiction with the fact that the components are pairwise disjoint. Therefore, $C_k = \text{co}(C_k)$, that is, C_k is convex.

Finally, since $\arg \max_{q \in S} V(q)$ is a convex set, there exists a unique index $k^* \in \{1, \dots, K\}$ such that $C_{k^*} = \arg \max_{q \in S} V(q)$. This concludes the proof. \blacksquare

The final step consists in proving that the convergence to any C_k with $k \neq k^*$ has probability zero. The following observations play a crucial role. Since C_k is convex (and not merely connected), the set C_k admits a maximal element, that is, there exists $\widehat{\Theta}(k) \in C_k$ such that $\widehat{\Theta} \subset \widehat{\Theta}(k)$, for all $\widehat{\Theta} \in C_k \setminus \{\widehat{\Theta}(k)\}$. We call $\widehat{\Theta}(k)$ the *support* of C_k . By construction, for $k \neq k^*$, $\widehat{\Theta}(k) \neq \Theta$. We also have:

Lemma 3 *For all $k \neq k^*$, the set $\widehat{\Theta}(k^*) \setminus \widehat{\Theta}(k)$ is non-empty.*

Proof of Lemma 3 If it were not the case, we would have $\widehat{\Theta}(k^*) \subseteq \widehat{\Theta}(k)$. Therefore, $\max_{q \in \Delta(\widehat{\Theta}(k^*))} V \leq \max_{q \in \Delta(\widehat{\Theta}(k))} V$ and, thus, $V(C_{k^*}) \leq V(C_k)$, a contradiction. \blacksquare

We are now ready to state and prove the final step in the proof.

Lemma 4 *For all $k \neq k^*$, $\mathbb{P}(\mathcal{L}(q_n) \subseteq C_k) = 0$.*

Proof of Lemma 4 Recall that $C_k = \bigcup_{\widehat{\Theta} \in \mathcal{C}_k} S_{\widehat{\Theta}}$. Pick $q^* \in F_{\widehat{\Theta}(k^*)}$, and let

$$\Theta_3 := \bigcup_{q \in C_k} \{\theta \in \Theta : q_\theta^* > 0, q_\theta = 0\}.$$

For $\hat{q} \in C_k$, define $\Theta_1(\hat{q}) := \{\theta \in \Theta : q_\theta^* = \hat{q}_\theta = 0\}$, and $\Theta_2(\hat{q}) := \{\theta \notin \Theta_3 : \hat{q}_\theta > 0\}$.

By Lemma 3, there exists $\theta^* \in \Theta_3$ such that $\hat{q}_{\theta^*} = 0$ for all $\hat{q} \in C_k$. We stress that the sets $\Theta_1(\hat{q})$ and $\Theta_2(\hat{q})$ depend on the face \hat{q} belongs to.

By strict concavity of V on $\text{co}\{q^*, \hat{q}\}$, we have $\frac{\partial}{\partial t} V(tq^* + (1-t)\hat{q}) > 0$. Hence,

$$0 < \sum_{\theta \in \Theta} (q_\theta^* - \hat{q}_\theta) f_\theta(\hat{q}) = \sum_{\theta \in \Theta} q_\theta^* f_\theta(\hat{q}) - 1 = \sum_{\theta \in \Theta_2(\hat{q}) \cup \Theta_3} q_\theta^* f_\theta(\hat{q}) - 1 = \sum_{\theta \in \Theta_2(\hat{q})} q_\theta^* + \sum_{\theta \in \Theta_3} q_\theta^* f_\theta(\hat{q}) - 1.$$

Thus, $\sum_{\theta \in \Theta_3} q_\theta^* f_\theta(\hat{q}) > 1 - \sum_{\theta \in \Theta_2(\hat{q})} q_\theta^*$. Since $\sum_{\theta \in \Theta_2(\hat{q}) \cup \Theta_3} q_\theta^* = 1$,

$$\sum_{\theta \in \Theta_3} \mu_\theta f_\theta(\hat{q}) > 1, \quad \forall \hat{q} \in C_k,$$

where $\mu_\theta := \frac{q_\theta^*}{\sum_{\theta \in \Theta_3} q_\theta^*}$, for all $\theta \in \Theta_3$.

By compactness of C_k , there exists $\varepsilon > 0$ such that $\sum_{\theta \in \Theta_3} \mu_\theta f_\theta(\hat{q}) > 1 + 3\varepsilon$, for all $\hat{q} \in C_k$. By continuity of f_θ , there exists an open neighborhood U of C_k (in \mathbf{S}) such that

$$\inf_{\hat{q} \in U} \sum_{\theta \in \Theta_3} \mu_\theta f_\theta(\hat{q}) > 1 + 2\varepsilon.$$

Note that, if $q_\theta > 0$, then $0 \leq \frac{B_\theta(q, x)}{q_\theta} = \frac{p_\theta(x)}{\sum_{\theta' \in \Theta} q_{\theta'} p_{\theta'}(x)} \leq \frac{p_\theta(x)}{\min_{\theta' \in \Theta} p_{\theta'}(x)} < +\infty$ since $p_{\theta'}(x) > 0$ for all (θ', x) .²⁴ Let $C(\theta) := \max_{x \in X} \frac{p_\theta(x)}{\min_{\theta' \in \Theta} p_{\theta'}(x)}$. If $q_{n, \theta} > 0$, it follows that

$$1 - \gamma_{n+1} + \gamma_{n+1} C(\theta) \geq \frac{q_{n+1, \theta}}{q_{n, \theta}} = 1 - \gamma_{n+1} + \frac{\gamma_{n+1}}{q_{n, \theta}} B_\theta(q_n, x_{n+1}) \geq 1 - \gamma_{n+1}. \quad (8)$$

Since

$$\frac{q_{n+1, \theta}}{q_{n, \theta}} = 1 + \gamma_{n+1} \left(-1 + f_\theta(q_n) + \frac{U_{n+1, \theta}}{q_{n, \theta}} \right). \quad (9)$$

we can combining (8) and (9) to obtain $C(\theta) - 1 \geq -1 + f_\theta(q_n) + \frac{U_{n+1, \theta}}{q_{n, \theta}} \geq -1$, for all n .

From the Stochastic Taylor Theorem (see e.g. [Aliprantis and Border \(1999\)](#), Theorem 17.17 p. 569), for any $n \in \mathbb{N}^*$, there exists a random variable ζ_n , such that $|\zeta_n| \leq \gamma_{n+1} \left(-1 + f_\theta(q_n) + \frac{U_{n+1, \theta}}{q_{n, \theta}} \right)$, and

$$\log \left(\frac{q_{n+1, \theta}}{q_{n, \theta}} \right) = \gamma_{n+1} \left(-1 + f_\theta(q_n) + \frac{U_{n+1, \theta}}{q_{n, \theta}} \right) - \frac{1}{(1 + \zeta_n)^2} \gamma_{n+1}^2 \left(-1 + f_\theta(q_n) + \frac{U_{n+1, \theta}}{q_{n, \theta}} \right)^2.$$

Defining $z_n := \sum_{\theta \in \Theta_3} \mu_\theta \log q_{n, \theta}$, it follows that

$$\begin{aligned} z_{n+1} - z_n &= \sum_{\theta \in \Theta_3} \mu_\theta \log \left(\frac{q_{n+1, \theta}}{q_{n, \theta}} \right) \\ &= \sum_{\theta \in \Theta_3} \mu_\theta \log \left(1 + \gamma_{n+1} \left(-1 + f_\theta(q_n) + \frac{U_{n+1, \theta}}{q_{n, \theta}} \right) \right) \\ &= \gamma_{n+1} \sum_{\theta \in \Theta_3} \mu_\theta \left(-1 + f_\theta(q_n) + \frac{U_{n+1, \theta}}{q_{n, \theta}} \right) - K_{n+1} \gamma_{n+1}^2, \end{aligned}$$

where $K_{n+1} := \frac{1}{(1 + \zeta_n)^2} \left(-1 + f_\theta(q_n) + \frac{U_{n+1, \theta}}{q_{n, \theta}} \right)^2$. Note that for n large enough $|K_{n+1}| \leq K :=$

²⁴Again, if the support of p^* was a strict subset the support of each $p_{\theta'}$, the analysis would hold for all x in the support of p^* .

$2 \max\{1, (C(\theta) - 1)^2\}$. $|\zeta_n| \leq 1/2$. Taking the conditional expectation, we get

$$\mathbb{E}(z_{n+1} | \mathcal{F}_n) - z_n = \gamma_{n+1} \sum_{\theta \in \Theta_3} \mu_\theta (-1 + f_\theta(q_n)) - \gamma_{n+1}^2 \mathbb{E}(K_{n+1} | \mathcal{F}_n) \geq \gamma_{n+1} \varepsilon - K \gamma_{n+1}^2,$$

on the event $\{q_n \in U\}$, for large enough n . Assume by contradiction that $\mathbb{P}(\mathcal{L}(q_n) \subseteq U) > 0$. Since

$$\{\mathcal{L}(q_n) \subseteq U\} \subseteq \bigcup_{N \in \mathbb{N}} \{q_n \in U, \forall n \geq N\},$$

there exists $\bar{N} \in \mathbb{N}$ such that

$$\mathbb{P}(q_n \in U, \forall n \geq \bar{N}) > 0.$$

However, on the event $\{q_n \in U, \forall n \geq \bar{N}\}$, we have $\sum_{\theta \in \Theta_3} \mu_\theta (-1 + f_\theta(q_n)) > \varepsilon$ for all $n \geq \bar{N}$, implying that, for any $n > \bar{N}$,

$$\mathbb{E}(z_{n+1}) \geq \mathbb{E}(z_{\bar{N}}) + \varepsilon \sum_{m=\bar{N}}^n \gamma_{m+1} - K \sum_{m=\bar{N}}^n \gamma_{m+1}^2.$$

Thus, on the event $\{q_n \in U, \forall n \geq \bar{N}\}$, we have $\lim_{n \rightarrow +\infty} \mathbb{E}(z_n) = +\infty$, which cannot happen since $z_n \leq 0$. Therefore $\mathbb{P}(q_n \in U, \forall n \geq \bar{N}) = 0$, the desired contradiction. \blacksquare

6.2 Proof of Proposition 1

Let $\tilde{q} \in \mathbf{S}^*$ be a zero of H , that is, for all $\theta \in \Theta$,

$$1 = f_\theta(\tilde{q}) = \sum_{x \in X} \frac{p^*(x)}{\sum_{\theta'} p_{\theta'}(x) \tilde{q}_{\theta'}} p_\theta(x).$$

Define $u \in \mathbb{R}^{|X|}$ as $u(x) := \frac{p^*(x)}{\sum_{\theta'} p_{\theta'}(x) \tilde{q}_{\theta'}}$. We have

$$\langle u, p_\theta - p_{\theta'} \rangle = 0, \quad \forall \theta, \theta'.$$

Since the family $\{p_\theta\}_{\theta \in \Theta}$ is full, this means that u is orthogonal to all vectors in the tangent space $\{v \in \mathbb{R}^{|X|} : \sum_{x \in X} v_x = 0\}$, that is, u is proportional to the unit vector $(1, \dots, 1)$. Therefore, there exists a real number μ such that $p^*(x) = \mu \sum_{\theta'} p_{\theta'}(x) \tilde{q}_{\theta'}$, for all $x \in X$. Summing over X , we obtain that $\mu = 1$. Hence $p^* = \sum_{\theta \in \Theta} p_\theta \tilde{q}_\theta$, and $\tilde{q} \in \Lambda$, which completes the proof.

6.3 Proof of Proposition 2

Let $\widehat{\Theta}$ be a non-empty subset of Θ . The argmax of the map $p \in \text{co}\{p_\theta\}_{\theta \in \widehat{\Theta}} \mapsto W(p)$ is a singleton, $\{\bar{p}\}$, by convexity of $\text{co}\{p_\theta\}_{\theta \in \widehat{\Theta}}$ and strict concavity of W . Since the family is free and $|\widehat{\Theta}| \leq |\Theta| \leq |X|$, the map $\mu : q \mapsto \sum_{\theta \in \Theta} q_\theta p_\theta$ is injective from $F_{\widehat{\Theta}}$ to $\Delta(X)$. Therefore, there exists a unique $\bar{q} \in F_{\widehat{\Theta}}$ such that $\mu(\bar{q}) = \bar{p}$.

6.4 Proof of Theorem 3

Recall that Theorem 3 characterizes the long-run properties of the updating process when the updating weights are constant, i.e., $\gamma_n = \gamma > 0$ for all n . The proof borrows and adapts ideas from the recent work of [Benaim and Schreiber \(2019\)](#).

We consider the family of random processes $\{(q_n^\gamma)_n\}$, where $q_0^\gamma = q_0 \in \mathbf{S}^*$, and $(q_n^\gamma)_n$ is recursively given by

$$q_{n+1}^\gamma = (1 - \gamma)q_n^\gamma + \gamma B(q_n^\gamma, x_{n+1}). \quad (10)$$

It will be convenient to rewrite the recursive formula (10) component-wise as follows: for $\theta \in \Theta$,

$$\frac{q_{n+1, \theta}^\gamma}{q_{n, \theta}^\gamma} = F_\theta^\gamma(q_n^\gamma, x_{n+1}), \text{ where } F_\theta^\gamma(q, x) := (1 - \gamma) + \gamma \frac{p_\theta(x)}{\sum_{\theta'} p_{\theta'}(x) q_{\theta'}}. \quad (11)$$

Let P_γ be the transition kernel associated to the Markov chain $(q^\gamma(n))_n$. Given $h \in \mathcal{C}(\mathbf{S})$, the set of continuous maps from \mathbf{S} to \mathbb{R} , define $P_\gamma h$ as

$$q \in \mathbf{S} \mapsto P_\gamma h(q) := \mathbb{E}[h(q_1^\gamma) \mid q_0^\gamma = q] = \int_{\mathbf{S}} h(\lambda) P_\gamma(q, d\lambda).$$

Given $\gamma > 0$, define the *occupation measure up to time n* as

$$\Pi_n^\gamma := \frac{1}{n} \sum_{m=0}^{n-1} \delta_{q_m^\gamma}, \text{ where } \delta_q \text{ denotes the Dirac measure at } q. \quad (12)$$

We are interested in the asymptotic behavior of $(\Pi_n^\gamma)_n$, when γ gets small.

Following [Benaim and Schreiber \(2019\)](#), define

$$r_\theta^\gamma(q) := \sum_{x \in X} p^*(x) \log F_\theta^\gamma(q, x).$$

Intuitively, this quantity represents the expected growth of the belief in θ , when current belief is q and observations are distributed according to p^* .

Lemma 5 For all $\lambda^* \in C_{k^*}$ and $q \in \mathbf{S}$, the derivative of the mapping $\gamma \mapsto \sum_{\theta \in \Theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q)$ evaluated in zero is non-negative, and equal to zero if, and only if, $q \in C_{k^*}$.

Consequently, for any $q \in \mathbf{S} \setminus C_{k^*}$, there exists $\gamma_0(q) > 0$ such that, for all $0 < \gamma < \gamma_0(q)$ and all $\lambda^* \in C_{k^*}$, we have $\sum_{\theta \in \Theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q) > 0$.

Proof of Lemma 5 Pick $q \in \mathbf{S}$ and $\lambda^* \in C_{k^*}$. The derivative of $r_{\theta}^{\gamma}(q)$ with respect to γ , evaluated in $\gamma = 0$ is equal to

$$\sum_{x \in X} p^*(x) \left(\frac{p_{\theta}(x)}{\sum_{\theta' \in \Theta} q_{\theta'} p_{\theta'}(x)} - 1 \right)$$

Hence, the derivative of $\gamma \mapsto \sum_{\theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q)$ evaluated in $\gamma = 0$ is equal to

$$\sum_{x \in X} p^*(x) \left(\frac{\sum_{\theta \in \Theta} \lambda_{\theta}^* p_{\theta}(x)}{\sum_{\theta \in \Theta} q_{\theta} p_{\theta}(x)} - 1 \right) \geq \sum_{x \in X} p^*(x) \log \left(\frac{\sum_{\theta \in \Theta} \lambda_{\theta}^* p_{\theta}(x)}{\sum_{\theta \in \Theta} q_{\theta} p_{\theta}(x)} \right) \geq 0,$$

where the inequality follows from the concavity of the log. Moreover, by definition of C_{k^*} , the right-hand side holds in equality if, and only if, $q \in C_{k^*}$. In addition, the quantity $\sum_{\theta \in \Theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q)$ is equal to zero if $\gamma = 0$. Therefore, if $q \notin C_{k^*}$, there exists $\gamma(q) > 0$ such that, for any $0 < \gamma < \gamma(q)$,

$$\sum_{\theta \in \Theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q) > 0.$$

This concludes the proof. ■

Remark 1 Note that there does not exist a uniform threshold $\gamma_0 > 0$ on $\mathbf{S} \setminus C_{k^*}$: in words, for any $\gamma > 0$, there might exist $q \in \mathbf{S} \setminus C_{k^*}$ such that $\sum_{\theta \in \Theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q) < 0$. However, a direct consequence of this lemma is that, for any open set V such that $C_{k^*} \subset V$, there exists $\gamma_0 > 0$ such that, for any $\gamma < \gamma_0$

$$\inf_{q \in \mathbf{S} \setminus V} \sum_{\theta \in \Theta} \lambda_{\theta}^* r_{\theta}^{\gamma}(q) > 0.$$

An invariant probability measure for $(q_n^{\gamma})_n$ is a probability measure π on \mathbf{S} such that

$$\int_{\mathbf{S}} h(q) \pi(dq) = \int_{\mathbf{S}} P_{\gamma} h(q) \pi(dq), \text{ for all } h \in \mathcal{C}(\mathbf{S}). \quad (13)$$

The set of such invariant probability distributions is denoted $\text{Inv}(P_{\gamma})$. It is a compact and convex subset of $\mathcal{P}(\mathbf{S})$. Its extreme points are the so-called *ergodic measures* of P_{γ} .

Given $\pi \in \mathcal{P}(\mathbf{S})$ and $\theta \in \Theta$, define the quantity $r_{\theta}^{\gamma}(\pi) := \int_{\mathbf{S}} r_{\theta}^{\gamma}(q) \pi(dq)$. It is the expected growth of the belief in θ if the current belief is distributed according to π . Proposition 1

in [Benaïm and Schreiber \(2019\)](#) states the following: if π is an ergodic measure (for P_γ) then

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{m=0}^{n-1} \log F_\theta^\gamma(q_m^\gamma, x_{m+1}) = r_\theta^\gamma(\pi) \text{ for } \pi - \text{almost every initial condition } q_0.$$

Moreover $r_\theta^\gamma(\pi) = 0$ for any $\theta \in \text{Supp}(\pi)$, where

$$\text{Supp}(\pi) := \{\theta \in \Theta : \pi(\{q \in \mathbf{S} : q_\theta > 0\}) = 1\}.$$

The intuition behind this is that, under stationary regime, the supported states have a null expected growth.

Lemma 6 *Let π^γ be a weak* limit point of the sequence $(\Pi_n^\gamma)_n$. Then $\pi^\gamma \in \text{Inv}(P_\gamma)$, and*

$$\sum_{\theta} \lambda_\theta^* r_\theta^\gamma(\pi^\gamma) \leq 0, \text{ for all } \lambda^* \in C_{k^*}.$$

Proof of Lemma 6 Let π^γ be a weak* limit point of the sequence $(\Pi_n^\gamma)_n$. By definition, there exists an increasing sequence n_k such that, for any $h \in \mathcal{C}(\mathbf{S})$,

$$\int_{\mathbf{S}} h(q) \pi^\gamma(dq) = \lim_k \int_{\mathbf{S}} h(q) \Pi_{n_k}^\gamma(dq) = \lim_k \frac{1}{n_k} \sum_{m=0}^{n_k-1} h(q_m^\gamma)$$

and

$$\int_{\mathbf{S}} P_\gamma h(q) \pi^\gamma(dq) = \lim_k \frac{1}{n_k} \sum_{m=0}^{n_k-1} P_\gamma h(q_m^\gamma)$$

By Lemma 3(i) in [Benaïm and Schreiber \(2019\)](#), for any $g \in \mathcal{C}(\mathbf{S} \times X)$, denoting $\bar{g}(q) := \sum_x p^*(x)g(q, x)$, we have

$$\lim_{n \rightarrow +\infty} \left| \frac{\sum_{m=0}^{n-1} g(q_m^\gamma, x_{m+1})}{n} - \frac{\sum_{m=0}^{n-1} \bar{g}(q_m^\gamma)}{n} \right| = 0, \text{ almost surely.} \quad (14)$$

Let $g(q, x) := h(q \circ F_\gamma(q, x))$.²⁵ Note that $g(q_m^\gamma, x_{m+1}) = h(q_m^\gamma \circ F_\gamma(q_m^\gamma, x_{m+1})) = h(q_{m+1}^\gamma)$. Since $P_\gamma h(q) = \sum_{x \in X} p^*(x)h(q \circ F_\gamma(q, x)) = \bar{g}(q)$, and using (14), we get

$$\lim_{n \rightarrow +\infty} \left| \frac{\sum_{m=0}^{n-1} h(q_{m+1}^\gamma)}{n} - \frac{\sum_{m=0}^{n-1} P_\gamma h(q_m^\gamma)}{n} \right| = 0.$$

²⁵We let $q \circ q'$ denote the Hadamard product: $(q \circ q')_\theta := q_\theta q'_\theta$

Hence

$$\int_{\mathbf{S}} P_\gamma h(q) \pi^\gamma(dq) = \lim_k \frac{1}{n_k} \sum_{m=0}^{n_k-1} P_\gamma h(q_m^\gamma) = \lim_k \frac{1}{n_k} \sum_{m=0}^{n_k-1} h(q_m^\gamma) = \int_{\mathbf{S}} h(q) \pi^\gamma(dq),$$

and (13) follows.

We now prove the second claim. Using identity (14) with $g(q, x) = \sum_{\theta \in \Theta} \lambda_\theta^* \log F_\theta^\gamma(q, x)$ and using the fact that

$$\sum_{m=0}^{n-1} \log F_\theta^\gamma(q_m^\gamma, x_{m+1}) = \sum_{m=0}^{n-1} \log \frac{q_{m+1, \theta}^\gamma}{q_{m, \theta}^\gamma} = \log \frac{q_{n, \theta}^\gamma}{q_{0, \theta}^\gamma},$$

we obtain

$$\lim_{n \rightarrow +\infty} \frac{\sum_{\theta \in \Theta} \lambda_\theta^* (\log q_{n, \theta}^\gamma - \log q_{0, \theta}^\gamma) - \sum_{m=0}^{n-1} \sum_{\theta \in \Theta} \lambda_\theta^* \sum_{x \in X} p^*(x) \log (F_\theta^\gamma(q_m^\gamma, x))}{n} = 0.$$

Note that $\limsup_n \frac{\sum_{\theta \in \Theta} \lambda_\theta^* \log q_{n, \theta}^\gamma}{n} \leq 0$ and $\lim_{n \rightarrow +\infty} \frac{\sum_{\theta \in \Theta} \lambda_\theta^* \log q_{0, \theta}^\gamma}{n} = 0$. Moreover

$$\lim_{k \rightarrow +\infty} \frac{1}{n_k} \sum_{m=0}^{n_k-1} \sum_{x \in X} p^*(x) \log F_\theta^\gamma(q_m^\gamma, x) = \int_{\mathbf{S}} \sum_{x \in X} p^*(x) \log F_\theta^\gamma(q, x) \pi(dq) = r_\theta^\gamma(\pi^\gamma).$$

As a result, we obtain that $\sum_{\theta \in \Theta} \lambda_\theta^* r_\theta^\gamma(\pi^\gamma) \leq 0$. ■

Remark 2 *It is worth noting that, even though $\pi \in \mathcal{P}(\mathbf{S})$ is such that $\pi(\mathbf{S} \setminus C_{k^*}) > 0$, it does not imply that there exists $\gamma_0 > 0$ such that, for all $\gamma < \gamma_0$,*

$$\sum_{\theta \in \Theta} \lambda_\theta^* r_\theta^\gamma(\pi) > 0.$$

Indeed, regardless how small γ is, $\text{Supp}(\pi)$ might contain points q such that $\sum_{\theta \in \Theta} \lambda_\theta^ r_\theta^\gamma(q) < 0$.*

Importantly, Lemmas 5 and 6 do not allow us to conclude that there exists $\gamma_0 > 0$ such that, for any $0 < \gamma < \gamma_0$, we have $\pi^\gamma(C_{k^}) = 1$. As a matter of fact, this statement cannot hold, because an element of $\text{Inv}(P_\gamma)$ cannot be supported by C_{k^*} . To see this, suppose for the sake of simplicity that $C_{k^*} = \{\lambda^*\}$. Then the distribution $\delta_{\{\lambda^*\}}$ is not invariant for P_γ .*

Definition 2 *The measure $\pi^* \in \Delta(\mathbf{S})$ is a limiting measure for the updating rule (C-Bay) if there exists a sequence $(\gamma_\ell, \pi_\ell)_{\ell \in \mathbb{N}}$ such that:*

1. $\gamma_\ell \downarrow 0$,
2. π_ℓ is a weak* limit point of $(\Pi_n^{\gamma_\ell})_n$, for all ℓ ,
3. $\lim_\ell \pi_\ell = \pi^*$ for the weak* topology.

In words, the measure π^* is a limiting measure if we can find a sequence of updating weights $(\gamma_\ell)_\ell$ and a corresponding sequence $(\pi_\ell)_\ell$ of limit points of $(\Pi_n^{\gamma_\ell})_n$, which converges to π^* for the weak* topology. We then prove the following:

Theorem 3. Let π^* be a limiting measure for the updating rule (C-Bay). The support of π^* is included in C_{k^*} .

Proof of Theorem 3 Let π^* be a limiting measure and $(\gamma_\ell, \pi_\ell)_{\ell \in \mathbb{N}}$ the corresponding sequence of weights and invariant measures mentioned in Definition 2.

Pick any $\lambda^* \in C_{k^*}$. Given $\gamma > 0$, $q \in \mathbf{S}$, define

$$R^*(\gamma, q) := \sum_{\theta \in \Theta} \lambda_\theta^* r_\theta^\gamma(q).$$

Note that $R^*(0, q) = 0$ and, by first point of Lemma 5,

$$R^*(\gamma, q) = R^*(0, q) + \gamma \frac{\partial R^*}{\partial \gamma}(0, q) + o(\gamma), \text{ for all } q \in \mathbf{S}.$$

Let O^* be an open subset of \mathbf{S} such that

$$\bigcup_{k \neq k^*} C_k \subseteq O, \quad C_{k^*} \subseteq O^*,$$

where $O := \mathbf{S} \setminus \text{Cl}(O^*)$. (Recall that the components $(C_k)_{k=1, \dots, k^*}$ are compact, connected and disjoint.)

Fix $\varepsilon > 0$. Choose O^* small enough so that $\left| \sup_{q \in \text{Cl}(O^*)} \frac{\partial R^*}{\partial \gamma}(0, q) \right| \leq \frac{\varepsilon}{2}$. (Recall that the derivative of R^* is zero for all $q \in C_{k^*}$ and the derivative is continuous. Hence, such a choice is always possible.) This implies that $|R^*(\gamma, q)| \leq \varepsilon \gamma$ for all $q \in \text{Cl}(O^*)$ and small enough γ .

Now define $c := \frac{1}{2} \inf_{q \in O} \frac{\partial R^*}{\partial \gamma}(0, q) > 0$. We then have $R^*(\gamma, q) > c\gamma$ for all $q \in O$ and small enough γ .

Finally, for large enough ℓ ,

$$\begin{aligned} R^*(\gamma_\ell, \pi_\ell) &= \int_{q \in \text{Cl}(O^*)} R^*(\gamma_\ell, q) \pi_\ell(dq) + \int_{q \in O} R^*(\gamma_\ell, q) \pi_\ell(dq) \\ &> -\varepsilon \gamma_\ell \pi_\ell(\text{Cl}(O^*)) + c \gamma_\ell \pi_\ell(O) \end{aligned}$$

By Lemma 6, $R^*(\gamma_\ell, \pi_\ell) \leq 0$, meaning that $\pi_\ell(O) < \frac{\varepsilon}{c} \pi_\ell(\text{Cl}(O^*)) \leq \frac{\varepsilon}{c}$ for large enough ℓ . Since O is open,

$$\pi^*(O) \leq \liminf_{\ell \rightarrow +\infty} \pi_\ell(O) < \frac{\varepsilon}{c}.$$

Since this inequality holds for any $\varepsilon > 0$, we have $\pi^*(O) = 0$. In particular, this proves

that $\pi^*(N^\delta(C_{k^*})) = 1$, where $N^\delta(C_{k^*})$ is the δ -neighborhood of the set C_{k^*} . This concludes the proof. ■

Remark 3 For any $\delta > 0$, we have $\liminf_{\ell \rightarrow +\infty} \pi_\ell(N^\delta(C_{k^*})) = 1$.

References

- ALIPRANTIS, C. D. AND K. C. BORDER (1999): *Infinite dimensional analysis: a hitchhiker's guide*, Springer, 2nd ed.
- BENAÏM, M. (1999): "Dynamics of stochastic approximation algorithms," *Le Seminaire de Probabilites. Lectures Notes in Mathematics*, 1709, 1–68.
- BENAÏM, M. AND S. J. SCHREIBER (2019): "Persistence and extinction for stochastic ecological models with internal and external variables," *Journal of Mathematical Biology*, 79, 393–431.
- BENJAMIN, D. J. (2019): "Errors in probabilistic reasoning and judgment biases," *Handbook of Behavioral Economics: Applications and Foundations 1, 2*, 69–186.
- BERK, R. H. (1966): "Limiting behavior of posterior distributions when the model is incorrect," *The Annals of Mathematical Statistics*, 37, 51–58.
- (1970): "Consistency a posteriori," *The Annals of Mathematical Statistics*, 41, 894–906.
- BERVOETS, S., M. BRAVO, AND M. FAURE (2020): "Learning with minimal information in continuous games," *Theoretical Economics*, 15, 1471–1508.
- BLACKWELL, D. (1951): "Comparison of experiments," in *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, vol. 1, 26.
- BOHREN, J. A. AND D. N. HAUSER (2021): "Learning with heterogeneous misspecified models: Characterization and robustness," *Econometrica*, 89, 3025–3077.
- (2024): "The behavioral foundations of model misspecification," Tech. rep.
- BORKAR, V. S. (2008): *Stochastic approximation: a dynamical systems viewpoint*, vol. 9, Springer.
- BRANDIERE, O. (1998): "Some pathological traps for stochastic approximation," *SIAM Journal on Control and Optimization*, 36, 1293–1314.

- BRANDIERE, O. AND M. DUFLO (1996): “Les algorithmes stochastiques contournent-ils les pièges?” in *Annales de l’IHP Probabilités et Statistiques*, vol. 32, 395–427.
- COVER, T. (1984): “An algorithm for maximizing expected log investment return,” *IEEE Transactions on Information Theory*, 30, 369–373.
- COVER, T. M. AND D. H. GLUSS (1986): “Empirical Bayes stock market portfolios,” *Advances in applied mathematics*, 7, 170–181.
- COVER, T. M. AND J. A. THOMAS (2006): *Elements of information theory*, Wiley-Interscience, 2 ed.
- CRIPPS, M. W. (2018): “Divisible updating,” UCL.
- DANENBERG, T. AND D. FUDENBERG (2024): “Endogenous attention and the spread of false news,” *arXiv preprint arXiv:2406.11024*.
- DE CLIPPEL, G. AND X. ZHANG (2022): “Non-bayesian persuasion,” *Journal of Political Economy*, 130, 2594–2642.
- DIACONIS, P. AND D. FREEDMAN (1986): “On the consistency of Bayes estimates,” *The Annals of Statistics*, 1–26.
- (1999): “Iterated random functions,” *SIAM review*, 41, 45–76.
- EDWARDS, W. (1968): “Conservatism in human information processing,” *Formal representation of human judgment*.
- EPSTEIN, L. G. (2006): “An axiomatic model of non-Bayesian updating,” *The Review of Economic Studies*, 73, 413–436.
- EPSTEIN, L. G., J. NOOR, AND A. SANDRONI (2010): “Non-bayesian learning,” *The BE Journal of Theoretical Economics*, 10.
- ESPONDA, I. AND D. POUZO (2016): “Berk–Nash equilibrium: A framework for modeling agents with misspecified models,” *Econometrica*, 84, 1093–1130.
- ESPONDA, I., D. POUZO, AND Y. YAMAMOTO (2021): “Asymptotic behavior of Bayesian learners with misspecified models,” *Journal of Economic Theory*, 195, 105260.
- FREEDMAN, D. A. (1963): “On the asymptotic behavior of Bayes’ estimates in the discrete case,” *The Annals of Mathematical Statistics*, 34, 1386–1403.
- FRICK, M., R. IJIMA, AND Y. ISHII (2020): “Misinterpreting others and the fragility of social learning,” *Econometrica*, 88, 2281–2328.

- (2023): “Belief convergence under misspecified learning: A martingale approach,” *The Review of Economic Studies*, 90, 781–814.
- (2024): “Welfare comparisons for biased learning,” *American Economic Review*, 114, 1612–1649.
- FUDENBERG, D., G. LANZANI, AND P. STRACK (2021): “Limit points of endogenous misspecified learning,” *Econometrica*, 89, 1065–1098.
- FUDENBERG, D., G. ROMANYUK, AND P. STRACK (2017): “Active learning with a misspecified prior,” *Theoretical Economics*, 12, 1155–1189.
- GAGNON-BARTSCH, T. AND B. BUSHONG (2022): “Learning with misattribution of reference dependence,” *Journal of Economic Theory*, 203, 105473.
- GREYER, D. M. (1980): “Bayes rule as a descriptive model: The representativeness heuristic,” *The Quarterly Journal of Economics*, 95, 537–557.
- GRÜNWARD, P. AND T. VAN OMMEN (2017): “Inconsistency of Bayesian inference for misspecified linear models, and a proposal for repairing it,” *Bayesian Analysis*, 12, 1069–1103.
- GRÜNWARD, P. D. AND N. A. MEHTA (2020): “Fast rates for general unbounded loss functions: from ERM to generalized Bayes,” *Journal of Machine Learning Research*, 21, 1–80.
- HEIDHUES, P., B. KÓSZEGI, AND P. STRACK (2018): “Unrealistic expectations and misguided learning,” *Econometrica*, 86, 1159–1214.
- (2021): “Convergence in models of misspecified learning,” *Theoretical Economics*, 16, 73–99.
- KELLY, J. L. (1956): “A new interpretation of information rate,” *the bell system technical journal*, 35, 917–926.
- KOVACH, M. (2021): “Conservative updating,” *arXiv preprint arXiv:2102.00152*.
- KUSHNER, H. J. AND G. G. YIN (2003): *Stochastic Approximation and Recursive Algorithms and Applications*, Springer.
- LANZANI, G. (2022): “Dynamic concern for misspecification,” *Available at SSRN 4454504*.
- LEE, Y.-J., W. LIM, AND C. ZHAO (2023): “Cheap talk with prior-biased inferences,” *Games and Economic Behavior*, 138, 254–280.

- MARCET, A. AND T. J. SARGENT (1989): “Convergence of least squares learning mechanisms in self-referential linear stochastic models,” *Journal of Economic Theory*, 48, 337–368.
- MASSARI, F. (2022): “Under-reaction: Irrational behavior or robust response to model misspecification?” University of Bologna.
- ORTOLEVA, P. (2022): “Alternatives to Bayesian updating,” *Annual Review of Economics*, 16.
- PEMANTLE, R. (1990): “Nonconvergence to unstable points in urn models and stochastic approximations,” *The Annals of Probability*, 18, 698–712.
- PHILLIPS, L. D. AND W. EDWARDS (1966): “Conservatism in a simple probability inference task.” *Journal of Experimental Psychology*, 72, 346.
- RABIN, M. AND J. L. SCHRAG (1999): “First impressions matter: A model of confirmatory bias,” *The Quarterly Journal of Economics*, 114, 37–82.
- ROBBINS, H. AND S. MONRO (1951): “A stochastic approximation method,” *The Annals of Mathematical Statistics*, 400–407.
- SHALIZI, C. R. (2009): “Dynamics of Bayesian updating with dependent data and misspecified models,” *Electronic Journal of Statistics*, 3, 1039–1074.
- SPIEGLER, R. (2016): “Bayesian networks and boundedly rational expectations,” *The Quarterly Journal of Economics*, 131, 1243–1290.
- ZELLNER, A. (1988): “Optimal information processing and Bayes’s theorem,” *The American Statistician*, 42, 278–280.