WELFARE COMPARISONS FOR BIASED LEARNING

By

Mira Frick, Ryota Iijima, and Yuhta Ishii

February 2021

Revised March 2021

COWLES FOUNDATION DISCUSSION PAPER NO. 2149R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Welfare Comparisons for Biased Learning*

Mira Frick   Ryota Iijima   Yuhta Ishii

First posted version: February 15, 2021

This version: March 11, 2021

Abstract

We study robust welfare comparisons of learning biases, i.e., deviations from correct Bayesian updating. Given a true signal distribution, we deem one bias more harmful than another if it yields lower objective expected payoffs in *all* decision problems. We characterize this ranking in static (one signal) and dynamic (many signals) settings. While the static characterization compares posteriors signal-by-signal, the dynamic characterization employs an "efficiency index" quantifying the speed of belief convergence. Our results yield welfare-founded quantifications of the severity of well-documented biases. Moreover, the static and dynamic rankings can disagree, and "smaller" biases can be worse in dynamic settings.

*Frick: Yale University (mira.frick@yale.edu); Iijima: Yale University (ryota.iijima@yale.edu); Ishii: Pennsylvania State University (yxi5014@psu.edu). We thank Nageeb Ali, Aislinn Bohren, Carlo Cusumano, Drew Fudenberg, Kevin He, Marina Halac, Miho Hong, Johannes Horner, Philippe Jehiel, Vijay Krishna, Bart Lipman, George Mailath, Weicheng Min, Pietro Ortoleva, Larry Samuelson, Rani Spiegler, Philipp Strack, as well as seminar audiences at Emory, Georgetown, Harvard-MIT, ITAM, NYU, Penn, Penn State, UCL, TE session on Misspecified Models, and Yale.

1 Introduction

1.1 Motivation and Overview

A growing literature in behavioral economics studies ways in which individuals' learning departs from correct Bayesian updating, whether due to psychological biases and limitations or due to simplifying assumptions about a complex environment. Experimental work has documented a variety of systematic learning biases, such as overconfidence, under- or overreaction to information, and partisan or confirmation bias (for a recent survey, see Benjamin, 2019). Such learning biases may come to bear on many important economic problems, from career choices to financial investment decisions and voting, by leading to inefficient behavior.

In this paper, we ask how to compare the welfare costs of different learning biases, and how to do so *robustly*, i.e., independently of the particular economic problem a biased agent might face. Given a true signal distribution, we deem one bias more harmful than another if it induces a lower objective expected payoff in *all* decision problems. Our main results characterize this welfare ranking in both a static (one signal observation) and a dynamic (many signal observations) setting. While the static characterization is based on comparing interim beliefs after each signal observation, for the dynamic characterization, we introduce a "learning efficiency index" that quantifies the speed at which beliefs converge. Thus, complementing a large theoretical literature that derives asymptotic beliefs under various learning biases (see Section 1.2), a key ingredient of our analysis is to understand how learning biases affect short- and medium-run beliefs.

Our results provide a welfare-founded approach to quantifying the severity of specific learning biases that are documented and estimated in experiments. For some commonly studied biases, the welfare-founded quantifications we obtain conflict with intuitive measures of severity used in applied work. We also highlight several general implications: In particular, the static and dynamic welfare rankings can disagree; that is, some biases are robustly less harmful than others in the short run, but robustly more harmful in the medium run. Moreover, according to the dynamic ranking, some "large" biases are less harmful than other "vanishingly small" biases; indeed, when agents are uncertain about both payoff-relevant states and the signal structure, some biases can outperform correctly specified learning.

Illustrative example: Asymmetric updating. To illustrate our exercise and some of its implications, consider the following commonly studied learning bias (e.g., Mobius, Niederle, Niehaus, and Rosenblat, 2014). An agent learns about some fixed and unknown state θ (e.g., her ability) that is either low, $\underline{\theta}$, or high, $\overline{\theta}$. Prior probabilities are $p_0(\theta)$. She observes a sequence $(x_1, \cdots, x_T)$ of T signals, drawn conditionally i.i.d. from the binary set $\{\underline{x}, \overline{x}\}$ with probabilities $0 < \mu(\overline{x} \mid \underline{\theta}) < \mu(\overline{x} \mid \overline{\theta}) < 1$. Thus, signal $\overline{x}$ (resp. $\underline{x}$) is "good news" (resp. "bad news") about θ. After each signal observation x_t, the agent updates her belief p_t using a distorted likelihood ratio,

$$\frac{p_t(\overline{\theta})}{p_t(\underline{\theta})} = \frac{p_{t-1}(\overline{\theta})}{p_{t-1}(\underline{\theta})} \left(\frac{\mu(x_t \mid \overline{\theta})}{\mu(x_t \mid \underline{\theta})} \right)^{c(x_t)}, \tag{1}$$

where $c(x) > 0$ for each x. The case $c(\overline{x}) = c(\underline{x}) = 1$ corresponds to standard Bayesian updating. By contrast, (1) can accommodate under- or overinference from some signals and allows for these departures to vary across different signals. For example, $c(\overline{x}) > c(\underline{x})$ captures a form of "ego-biased" updating, where the agent reacts more strongly to good news about her ability than to bad news.

The experimental literature documents bias (1) and has estimated distortion parameters for different subjects.[1] However, based on casual inspection, it might not be obvious how to evaluate the severity of this bias under different distortion functions $c(\cdot)$. This paper provides a welfare-founded approach: Suppose that, upon observing $(x_1, \cdots, x_T)$, the agent faces a decision problem, in which the payoff to each action depends on θ and she maximizes her subjective expected payoff given her posterior p_T. Using our results, we can characterize when the agent's welfare under distortion $c^1(\cdot)$ exceeds that under $c^2(\cdot)$ robustly, i.e., regardless of which decision problem she might face. Here, welfare is defined as the agent's ex-ante expected payoff evaluated according to the true probability measure over signals.

Static ranking: As a benchmark, suppose the agent observes a single signal draw $(T = 1)$. Then our results imply that $c^1(\cdot)$ is less harmful than $c^2(\cdot)$ in every decision problem if and only if $c^1(\cdot)$ distorts each signal less than $c^2(\cdot)$, i.e., for each x,

$$c^2(x) \leq c^1(x) \leq 1 \quad \text{or} \quad 1 \leq c^1(x) \leq c^2(x). \tag{2}$$

[1] The estimated parameters in Mobius, Niederle, Niehaus, and Rosenblat (2014) are $c(\overline{x}) = 0.27$, $c(\underline{x}) = 0.17$. Subsequent work has estimated how distortion parameters vary with gender and other demographic characteristics (see the survey by Benjamin, 2019).

Dynamic ranking: Suppose instead that the agent has access to many signal draws. We show that in every decision problem and for any large enough number of draws T, the ranking depends only on the ratio $c^i(\overline{x})/c^i(\underline{x})$, where welfare is higher the "closer" this ratio is to 1. While it is easy to see that this ratio determines the agent's asymptotic beliefs, our key observation is that this ratio also determines the speed at which beliefs converge. This is crucial in allowing us to compare welfare across the large class of distortions $c^1(\cdot)$ and $c^2(\cdot)$ that induce the same asymptotic beliefs.

Note that the static and dynamic rankings can disagree: For instance, consider any asymmetric distortion $c^1(\underline{x}) \neq c^1(\overline{x})$, no matter how close to 1, and any symmetric distortion $c^2(\overline{x}) = c^2(\underline{x})$, no matter how different from 1. Then, $c^1(\cdot)$ is worse than $c^2(\cdot)$ according to the dynamic ranking, even if $c^1(\cdot)$ distorts each individual signal less than $c^2(\cdot)$ and hence is less harmful based on the static ranking. Moreover, measures that are sometimes used to quantify the severity of bias (1) in applications, such as the difference $c^i(\overline{x}) - c^i(\underline{x})$ (e.g., Coutts, 2019), can be inconsistent with both our welfare-founded rankings.

Overview. Our general model (Section 2) features an arbitrary finite state space and conditionally i.i.d. true signal structure, and we consider any learning bias that can be represented as Bayesian updating under some possibly incorrect perception of signal likelihoods. This encompasses misspecified Bayesian learning, as well as several important cases of non-Bayesian learning, such as the illustrative example above.

Section 3 characterizes our welfare rankings in the binary-state setting (Section 4.1 extends to general finite state spaces). Generalizing the illustrative example, bias 1 is less harmful than bias 2 according to the static ranking if each signal is interpreted more accurately under bias 1, as formalized by a nested likelihood ratio condition (Proposition 1).

Our main focus is on the dynamic welfare ranking. For this purpose, we define a *learning efficiency index*, which for any true and perceived signal structure captures how likely the agent is to encounter signal sequences that make her unable to distinguish between the two states. The key observation, based on arguments from large deviation theory, is that this efficiency index quantifies the speed at which the agent's beliefs converge. Using this observation, Theorems 1 and 2 show that for any biases that give rise to the same asymptotic beliefs, the dynamic welfare ranking is generically complete and is characterized by the learning efficiency index: With correct (resp., incorrect) asymptotic beliefs, higher (resp., lower) learning efficiency is better, as this reduces the medium-run likelihood of suboptimal choices in all decision problems. For biases that do not share the same asymptotic beliefs, the difference in asymptotic beliefs can be used to conduct a dynamic welfare comparison.

Based on these characterizations, we discuss when the static and dynamic rankings disagree. Intuitively, while the static ranking requires that each individual signal is interpreted more accurately, this does not preclude some sequences of signals from being interpreted less accurately. Indeed, if misinferences from different signals go in opposite directions, such opposite errors can "cancel out" and lead to better learning efficiency in dynamic settings. In addition to the illustrative example above, we also apply our welfare rankings to learning under partisan bias and under overconfidence, shedding light on the short- and medium-run inefficiencies these biases induce.

In our baseline model, correctly specified Bayesian learning (trivially) weakly dominates all learning biases according to both the static and dynamic welfare rankings. However, in the dynamic setting, we identify a unique class of biases, namely the special case of (1) with a constant distortion power c (Phillips and Edwards, 1966), that attains the same learning efficiency as the correctly specified case. As we discuss, this implies that, in dynamic settings, any Phillips-Edwards bias, no matter how significant, robustly outperforms any other form of bias, even if the latter is vanishingly small. Moreover, Section 4.2 considers an extension where the agent not only learns about a payoff-relevant state, but is also uncertain about the signal structure. Under such uncertainty, we show that some learning biases strictly outperform correctly specified learning according to the dynamic ranking, by speeding up the rate at which the agent learns the true payoff-relevant state.

1.2 Related Literature

This paper contributes to the theoretical literature on misspecified Bayesian and non-Bayesian learning. Much work studies how various learning biases affect asymptotic beliefs, in both single-agent (e.g., Berk, 1966; Nyarko, 1991; Fudenberg, Romanyuk, and Strack, 2017; Heidhues, Koszegi, and Strack, 2018, 2020; He, 2018; Bushong and Gagnon-Bartsch, 2019) and social learning settings (e.g., Eyster and Rabin, 2010; Bohren, 2016; Gagnon-Bartsch, 2017; Frick, Iijima, and Ishii, 2020a). Several recent papers provide more general criteria to determine convergence to asymptotic beliefs (Bohren and Hauser, 2018; Esponda, Pouzo, and Yamamoto, 2019; Frick, Iijima, and

Ishii, 2020b; Fudenberg, Lanzani, and Strack, 2020).[2] In contrast, our focus is on the welfare implications of learning biases, and a key ingredient of our analysis concerns short-/medium-run beliefs and the speed at which beliefs converge. In this paper, we restrict attention to the case of single-agent learning with exogenous i.i.d. signals, but Section 5 briefly discusses some extensions.

Our exercise is broadly related to a number of papers that examine whether and how specific misspecifications can "survive" based on a variety of selection criteria, including performance in competitive markets (e.g., Sandroni, 2000; Blume and Easley, 2006; Massari, 2020), goodness-of-fit tests (e.g., Cho and Kasa, 2015; Gagnon-Bartsch, Rabin, and Schwartzstein, 2018; Schwartzstein and Sunderam, 2021), voting (e.g., Levy, Razin, and Young, 2019), and subjective welfare (e.g., Montiel Olea, Ortoleva, Pai, and Prat, 2019; Eliaz and Spiegler, 2020). The most closely related papers within this literature are Fudenberg and Lanzani (2021) and He and Libgober (2021), who study selection based on objective welfare. These papers take an evolutionary approach, by characterizing which forms of misspecification are stable against mutations. While they conduct their analysis in a fixed environment and based on long-run outcomes, we compare welfare across all decision problems and consider short-/medium-run beliefs. He and Libgober (2021) show that misspecified agents can be better off than correctly specified agents under strategic externalities. Section 4.2 highlights an alternative mechanism in a single-agent setting, where some forms of misspecified learning outperform correctly specified learning by speeding up belief convergence.[3]

Finally, our focus on comparing welfare in all decision problems is similar in spirit to the literature on comparisons of statistical experiments (e.g., Blackwell, 1951; Moscarini and Smith, 2002; Azrieli, 2014; Mu, Pomatto, Strack, and Tamuz, 2021). This literature conducts robust welfare comparisons across different true signal structures assuming that agents are correctly specified, whereas we fix a true signal structure and compare welfare across different misperceptions of the signal structure. Our

[2] Esponda and Pouzo (2016, 2020) formalize Berk-Nash equilibrium, which captures steady states of general misspecified learning dynamics. Spiegler (2016) formalizes a steady state notion for subjective causal models that are captured by directed acyclic graphs (DAGs); he also asks when one DAG is robustly better than another based on objective steady-state payoffs, and finds that no two DAGs can be ranked in this way (except if one DAG is fully connected, i.e., correctly specified).

[3] Steiner and Stewart (2016); Gossner and Steiner (2018) consider static single-agent settings and show that misspecification can be beneficial when agents cannot perfectly implement their subjectively optimal strategies due to the presence of noise.

static ranking is the counterpart in our setting of Blackwell's (1951) order, while our dynamic ranking is the analog of Moscarini and Smith (2002); indeed, as we discuss in Remark 2, our learning efficiency index can be seen to generalize the efficiency index in Moscarini and Smith (2002). One important difference is that if one information structure dominates another in the sense of Blackwell (1951), this implies dominance in the sense of Moscarini and Smith (2002), because the Blackwell order is preserved under T-fold repetition of signal draws. In contrast, we show that our static ranking over misperceptions can be reversed in the dynamic setting.

2 Setting

A state θ is drawn once and for all from a finite set Θ according to a full-support distribution $p_0 \in \Delta(\Theta)$. An agent does not observe the realized state θ, but learns about θ from signal observations. Specifically, there is a finite set of signals X, and the agent observes a sequence of T draws of signals, $x^T = (x_1, x_2, \cdots, x_T) \in X^T$. Each signal x_t is drawn conditionally i.i.d. according to a ***true signal structure*** $\mu := (\mu_\theta)_{\theta\in\Theta}$, where $\mu_\theta \in \Delta(X)$ denotes the true signal distribution conditional on state θ. Assume that each μ_θ has full support over X and that $\mu_\theta \neq \mu_{\theta'}$ for all $\theta \neq \theta'$.

The agent has the correct prior p_0 over states, but her learning from signals is potentially biased.[4] Specifically, the agent's ***perceived signal structure*** is $\hat{\mu} := (\hat{\mu}_\theta)_{\theta\in\Theta}$, where $\hat{\mu}_\theta \in \mathbb{R}^X_{++}$ captures the agent's perceived likelihood of each signal conditional on state θ, and $\hat{\mu}_\theta \neq \hat{\mu}_{\theta'}$ for all $\theta \neq \theta'$. While we assume that $\hat{\mu}_\theta(x) > 0$ for each $x \in X$, we do not require that $\sum_{x\in X} \hat{\mu}_\theta(x) = 1$. As we discuss in Remark 1, this makes it possible to accommodate both misspecified Bayesian learning as well as certain well-studied classes of non-Bayesian learning. We call the agent ***correctly specified*** if $\hat{\mu} = \mu$.

Upon observing the signal sequence $x^T = (x_1, x_2, \cdots, x_T)$, the agent forms a posterior belief $p_T(\cdot \mid x^T) \in \Delta(\Theta)$ by applying Bayes' rule according to her perceived

[4] It is straightforward to allow agents to update beliefs based on incorrect and/or heterogeneous priors, while still defining ex-ante expected payoffs in (5) based on the true prior p_0. The dynamic welfare characterizations (Theorems 1–2) remain true unchanged. For the static characterization (Proposition 1), we generalize condition (6) to the requirement that after each signal x, the interim belief under $\hat{\mu}^1$ is a convex combination of the beliefs under μ and under $\hat{\mu}^2$.

signal structure. That is, for all $\theta' \in \Theta$,

$$p_T(\theta' \mid x^T) = \frac{p_0(\theta') \prod_{t=1}^{T} \hat{\mu}_{\theta'}(x_t)}{\sum_{\theta'' \in \Theta} p_0(\theta'') \prod_{t=1}^{T} \hat{\mu}_{\theta''}(x_t)}. \tag{3}$$

After forming the posterior belief, the agent faces a ***decision problem***, which is a nonempty finite set A of (payoff) acts.[5] An ***act*** is a vector $a \in \mathbb{R}^\Theta$, where a_θ denotes the agent's payoff from a conditional on state θ. Given any decision problem A and realized signal sequence x^T, the agent chooses an act $a^*(x^T) \in A$ to maximize her subjective expected payoff under her posterior belief $p_T(\cdot \mid x^T)$:

$$a^*(x^T) \in \underset{a \in A}{\operatorname{argmax}} \sum_{\theta \in \Theta} p_T(\theta \mid x^T) a_\theta. \tag{4}$$

For ease of exposition, we assume throughout the main text that $\hat{\mu}$ and A are such that (4) admits a unique solution. All results extend to the case with ties and our proofs in the appendix allow for this possibility.[6]

The agent's ***welfare*** is her ex-ante expected payoff to choosing her $\hat{\mu}$-subjectively optimal act $a^*(x^T)$ at each x^T. Here, taking the perspective of an outside observer, expectations over signal realizations are based on the *true* signal structure μ. That is, letting μ_θ^T denote the true distribution over signal sequences in X^T conditional on state θ, the agent's welfare is given by

$$W_T(\mu, \hat{\mu}, A) = \sum_{\theta \in \Theta} p_0(\theta) \sum_{x^T \in X^T} \mu_\theta^T(x^T) a_\theta^*(x^T). \tag{5}$$

Remark 1. When $\hat{\mu}_\theta \in \Delta(X)$ for each θ, our setting corresponds to misspecified Bayesian learning, capturing, for example, overconfidence (Example 2) or correlation neglect (e.g., Spiegler, 2016; Enke and Zimmermann, 2017). By allowing that $\sum_{x \in X} \hat{\mu}_\theta(x) \neq 1$, the model additionally nests several important classes of non-Bayesian learning, including the illustrative example on asymmetric updating (Section 1.1) and partisan bias (Example 1). At the same time, the assumption that the posterior p_T is formed according to Bayes' rule under $\hat{\mu}$ rules out some other classes of

[5] The finiteness restriction can be relaxed when $T = 1$, but is important for the dynamic case.

[6] Specifically, the static welfare characterization (Proposition 1) remains valid as long as a fixed tie-breaking rule (i.e., strict total order over acts) is used to select among multiple solutions to (4). The dynamic characterizations (Theorems 1–4) remain valid even when tie-breaking rules vary across biases $\hat{\mu}$.

non-Bayesian updating, such as Epstein, Noor, and Sandroni (2010). Finally, while we focus on the simplest possible setting of learning from exogenous i.i.d. signals, Section 5 briefly discusses extensions to active learning or belief-dependent updating (e.g., confirmation bias), as well as to intertemporally correlated signals. ▲

3 Welfare Rankings: Binary States

Given any true signal structure μ, consider two agents $i = 1, 2$ that differ only in their perceived signal structures $\hat{\mu}^i$. We seek to characterize when agent 1's bias is *robustly* less harmful than agent 2's, in the sense that agent 1's welfare exceeds agent 2's welfare at all decision problems. As a benchmark, we first consider a ***static welfare ranking***: This assumes that agents observe only a single signal draw ($T = 1$) and requires that for all A, $W_1(\mu, \hat{\mu}^1, A)$ exceeds $W_1(\mu, \hat{\mu}^2, A)$. Our main focus is on a ***dynamic welfare ranking***: This assumes that agents have access to many signal draws and requires that for all A, $W_T(\mu, \hat{\mu}^1, A)$ exceeds $W_T(\mu, \hat{\mu}^2, A)$ whenever T is large enough.

For ease of exposition, we focus throughout this section on a binary state environment, $\Theta = \{\underline{\theta}, \overline{\theta}\}$. Section 4.1 extends the results to general finite state spaces.

3.1 Static Ranking

The following result provides a sufficient condition under which agent 1's bias is less harmful than agent 2's according to the static welfare ranking: Condition (6) requires that agent 1's interpretation of each signal x is more accurate than agent 2's interpretation, in the sense that the perceived signal likelihood ratio under $\hat{\mu}^1$ is in between the true likelihood ratio and that under $\hat{\mu}^2$. This guarantees that agent 1's posterior following each signal realization is a convex combination of the true posterior and agent 2's posterior. Hence, agent 1's objective function is more aligned with the true objective function than agent 2's.[7]

Proposition 1. *Suppose that, for each $x \in X$,*

$$\frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)} \geq \frac{\hat{\mu}^1_{\overline{\theta}}(x)}{\hat{\mu}^1_{\underline{\theta}}(x)} \geq \frac{\hat{\mu}^2_{\overline{\theta}}(x)}{\hat{\mu}^2_{\underline{\theta}}(x)} \quad \text{or} \quad \frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)} \leq \frac{\hat{\mu}^1_{\overline{\theta}}(x)}{\hat{\mu}^1_{\underline{\theta}}(x)} \leq \frac{\hat{\mu}^2_{\overline{\theta}}(x)}{\hat{\mu}^2_{\underline{\theta}}(x)}. \tag{6}$$

[7] The logic is similar to results on time inconsistency, where welfare is monotonic in the degree of preference alignment between ex-ante and ex-post preferences (e.g., Gul and Pesendorfer, 2001).

Then $W_1(\mu, \hat{\mu}^1, A) \geq W_1(\mu, \hat{\mu}^2, A)$ for all decision problems A. The converse holds if $\mu, \hat{\mu}^1, \hat{\mu}^2$ satisfy the comonotonic likelikelihood ratio property.

For the necessity direction, we say that $\mu, \hat{\mu}^1, \hat{\mu}^2$ satisfy the ***comonotonic likelihood ratio property*** if there is a linear order $>$ on signals such that for any $x, x' \in X$ with $x > x'$, there exists $\gamma_{xx'} \in \mathbb{R}_+$ such that for all $\nu \in \{\mu, \hat{\mu}^1, \hat{\mu}^2\}$,

$$\frac{\nu_{\overline{\theta}}(x)}{\nu_{\underline{\theta}}(x)} > \gamma_{xx'} > \frac{\nu_{\overline{\theta}}(x')}{\nu_{\underline{\theta}}(x')}. \tag{7}$$

We note that under condition (6), the inequality $W_1(\mu, \hat{\mu}^1, A) \geq W_1(\mu, \hat{\mu}^2, A)$ is strict in any decision problem in which the agents' chosen acts differ at some x.

Asymmetric updating: In the illustrative example from Section 1.1, we have $X = \{\underline{x}, \overline{x}\}$, $\mu_{\overline{\theta}}(\overline{x}) > \mu_{\underline{\theta}}(\overline{x})$, and $\frac{\hat{\mu}^i_{\overline{\theta}}(x)}{\hat{\mu}^i_{\underline{\theta}}(x)} = \left(\frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}\right)^{c^i(x)}$, where $c^i(x) > 0$ for each agent i and signal x. Thus, the sufficient condition (6) for the static welfare ranking is equivalent to the requirement (2) that for each signal x, either $c^2(x) \leq c^1(x) \leq 1$ or $1 \leq c^1(x) \leq c^2(x)$. Moreover, setting $\overline{x} > \underline{x}$ and $\gamma_{\overline{x}\underline{x}} = 1$, true and perceived likelihood ratios satisfy the comonotonic likelihood ratio property (7). Thus, by Proposition 1, condition (2) is both necessary and sufficient for the static welfare ranking.

3.2 Dynamic Ranking: Correct Asymptotic Beliefs

We now characterize the dynamic welfare ranking, where for any decision problem A, agent 1's welfare $W_T(\mu, \hat{\mu}^1, A)$ exceeds agent 2's welfare $W_T(\mu, \hat{\mu}^2, A)$ whenever T is large enough. This ranking is relevant when agents have access to many signal draws, as is natural in most learning settings.

In this section, we present our approach under the assumption that both agents' biases are small enough that they learn the true state as $T \to \infty$; Section 3.4 shows how to extend this approach to the case where agents' beliefs are asymptotically incorrect. When asymptotic beliefs are correct under both biases, they cannot be used to compare welfare. Instead, we show that the dynamic welfare ranking is characterized by a learning efficiency index that captures the *speed* of convergence.

Formally, for any $\nu, \mu \in \mathbb{R}^X_+$, define the ***Kullback-Leibler (KL) divergence*** of ν relative to μ by

$$\mathrm{KL}(\mu, \nu) = \sum_x \mu(x) \log \frac{\mu(x)}{\nu(x)},$$

with the convention that $0 \log 0 = \frac{0}{0} = 0$ and $\log \frac{1}{0} = \infty$. This extends the usual definition of KL divergence between probability measures to arbitrary nonnegative vectors. We assume that both agents' perceived signal structures $\hat{\mu}^i$ satisfy the following consistency condition relative to the true signal structure μ. By standard arguments (as in Berk, 1966), this assumption is necessary and sufficient for an agent's belief to converge almost surely to a point-mass on the true state as $T \to \infty$.

Assumption 1 (Consistency)**.** For any distinct $\theta, \theta' \in \Theta$, we have $\mathrm{KL}(\mu_\theta, \hat{\mu}_\theta) < \mathrm{KL}(\mu_\theta, \hat{\mu}_{\theta'})$.

We introduce the following learning efficiency index:

Definition 1. Given any true and perceived signal structures μ and $\hat{\mu}$, the ***learning efficiency index*** is defined by $w(\mu, \hat{\mu}) := \min_\theta w(\theta, \mu, \hat{\mu})$, where

$$w(\theta, \mu, \hat{\mu}) := \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \quad \text{subject to} \quad \mathrm{KL}(\nu, \hat{\mu}_{\underline{\theta}}) = \mathrm{KL}(\nu, \hat{\mu}_{\overline{\theta}}). \tag{8}$$

We call the constraint $\mathrm{KL}(\nu, \hat{\mu}_{\underline{\theta}}) = \mathrm{KL}(\nu, \hat{\mu}_{\overline{\theta}})$ the ***indistinguishability condition***.[8]

To interpret, suppose that the realized empirical signal distribution $\nu \in \Delta(X)$ satisfies the indistinguishability condition. Then based on observing ν, the agent is unable to tell apart the two states $\underline{\theta}$ and $\overline{\theta}$, because ν comes equally close to her perceived signal distributions in both states. The measure $w(\theta, \mu, \hat{\mu})$ captures how "atypical" such ν are relative to the true signal distribution μ_θ in state θ: The greater $w(\theta, \mu, \hat{\mu})$, the "less likely" it is that the agent will face an empirical distribution ν (when T is large) that does not allow her to distinguish between $\overline{\theta}$ and $\underline{\theta}$. The learning efficiency index $w(\mu, \hat{\mu})$ considers the worst-case over all states of the measure $w(\theta, \mu, \hat{\mu})$.

Note that the learning efficiency index is defined without reference to any decision problem. The following theorem shows that this index characterizes the dynamic welfare ranking. Call a decision problem A ***non-trivial*** if it does not have a dominant act, i.e., there is no $a \in A$ such that $a_\theta \geq b_\theta$ for all $\theta \in \Theta$ and $b \in A$:

Theorem 1. *Fix any true signal structure μ and perceived signal structures $\hat{\mu}^1$ and $\hat{\mu}^2$ satisfying Assumption 1. Suppose $w(\mu, \hat{\mu}^1) > w(\mu, \hat{\mu}^2)$. Then for any non-trivial*

[8] Under Assumption 1, the indistinguishability condition is satisfied for some $\nu \in \Delta(X)$.

decision problem A, there exists T^* such that for all $T \geq T^*$,

$$W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A).$$

Theorem 1 implies that the dynamic welfare ranking is generically complete for all biases satisfying Assumption 1: Given a true signal structure μ, any two such biases $\hat{\mu}^1$ and $\hat{\mu}^2$ can be ranked robustly across all non-trivial decision problems, except when their efficiency indices $w(\mu, \hat{\mu}^1) = w(\mu, \hat{\mu}^2)$ are exactly tied.

We prove Theorem 1 in Appendix A.2. The basic idea is as follows. By Assumption 1, both agents' asymptotic beliefs are correct. Thus, in every decision problem, they choose an ex-post optimal act with probability one as $T \to \infty$. As a result, comparing their welfare amounts to comparing the rate at which their probability of making a suboptimal choice vanishes. The key observation is that, in each state θ and for any decision problem, this probability vanishes at the same rate as the probability that the empirical signal distribution ν approximately satisfies the indistinguishability condition. By applying Sanov's theorem from large deviation theory, we show that the latter probability vanishes exponentially as $T \to \infty$, at a rate given by $w(\theta, \mu, \hat{\mu}^i)$. Finally, the ex-ante welfare comparison is fully determined by the state in which this rate is lowest, i.e., by $w(\mu, \hat{\mu}^i) = \min_\theta w(\theta, \mu, \hat{\mu}^i)$. Note that while the indices $w(\mu, \hat{\mu}^i)$ are independent of the decision problem, the number of draws T^* after which these indices characterize the welfare ranking can depend on A.[9]

To apply Theorem 1 to specific biases, observe that an agent's perceived signal structure $\hat{\mu}$ affects the efficiency index $w(\mu, \hat{\mu})$ only through the indistinguishability condition. Moreover, the set of solutions to the indistinguishability condition, which we refer to as the ***indistinguishability set***, is given by the hyperplane

$$\mathcal{I}(\hat{\mu}) := \{\nu \in \Delta(X) : \mathrm{KL}(\nu, \hat{\mu}_{\underline{\theta}}) = \mathrm{KL}(\nu, \hat{\mu}_{\overline{\theta}})\} = \{\nu \in \Delta(X) : \nu \cdot \hat{\ell} = 0\} \tag{9}$$

whose normal vector $\hat{\ell} := \left(\log \frac{\hat{\mu}_{\overline{\theta}}(x)}{\hat{\mu}_{\underline{\theta}}(x)}\right)_{x \in X}$ is the vector of perceived log-likelihood ratios under $\hat{\mu}$.

Asymmetric updating: Continuing with the illustrative example, we have $\log \frac{\hat{\mu}^i_{\overline{\theta}}(x)}{\hat{\mu}^i_{\underline{\theta}}(x)} = c^i(x) \log \frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}$ for each agent i and signal $x \in \{\overline{x}, \underline{x}\}$. Let $\ell(x) := \log \frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}$.

[9] We expect that, similar to Proposition 5 in Mu, Pomatto, Strack, and Tamuz (2021), an upper bound on T^* for each A can be obtained by considering the cumulant functions of the perceived signal likelihood ratio distributions.

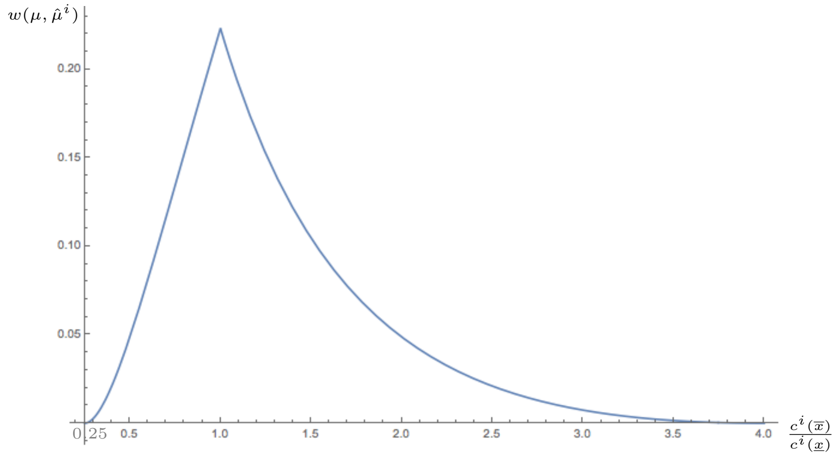

Figure 1: *Efficiency index under asymmetric updating as a function of* $\frac{c^i(\overline{x})}{c^i(\underline{x})}$ *when* $\mu_{\overline{\theta}}(\overline{x}) = 0.8 = \mu_{\underline{\theta}}(\underline{x})$. *Assumption 1 holds iff* $1/4 < \frac{c^i(\overline{x})}{c^i(\underline{x})} < 4$. *When* $\frac{c^i(\overline{x})}{c^i(\underline{x})} \leq 1$, *we have* $w(\mu, \hat{\mu}^i) = w(\underline{\theta}, \mu, \hat{\mu}^i) = \mathrm{KL}(\nu^i, \mu_{\underline{\theta}})$. *When* $\frac{c^i(\overline{x})}{c^i(\underline{x})} \geq 1$, *we have* $w(\mu, \hat{\mu}^i) = w(\overline{\theta}, \mu, \hat{\mu}^i) = \mathrm{KL}(\nu^i, \mu_{\overline{\theta}})$. *Here* ν^i *is given by (10).*

Then, under Assumption 1,[10] agent i's indistinguishability set (9) is a singleton, with unique element $\nu^i \in \Delta(X)$ given by

$$\nu^i(\overline{x}) = \frac{-\ell(\underline{x})}{-\ell(\underline{x}) + \frac{c^i(\overline{x})}{c^i(\underline{x})}\ell(\overline{x})}. \tag{10}$$

Thus, as shown in Figure 1, the efficiency index $w(\mu, \hat{\mu}^i)$ depends on agents' distortion functions $c^i(\cdot)$ only through the ratio $c^i(\overline{x})/c^i(\underline{x})$, and dynamic welfare is higher the closer this ratio is to 1.

Remark 2. Moscarini and Smith (2002) (MS) study robust comparisons of true signal structures for correctly specified agents. MS consider an index over signal structures μ given by

$$w^{\mathrm{MS}}(\mu) := -\min_{\lambda \in [0,1]} \log \sum_x \mu_{\underline{\theta}}(x)^{\lambda} \mu_{\overline{\theta}}(x)^{1-\lambda},$$

and show that if $w^{\mathrm{MS}}(\mu^1) > w^{\mathrm{MS}}(\mu^2)$, then for any non-trivial decision problem, a correctly specified agent's expected payoff under μ^1 is higher than under μ^2 for all sufficiently large T. Using the variational formula (e.g., Dupuis and Ellis, 2011, Lemma 6.2.3(f)), one can show that $w(\mu, \mu) = w^{\mathrm{MS}}(\mu)$, i.e., our efficiency index

[10] Assumption 1 holds if and only if $\frac{-\ell(\underline{x})}{\ell(\overline{x})} \frac{\mu_{\overline{\theta}}(\underline{x})}{\mu_{\overline{\theta}}(\overline{x})} < \frac{c^i(\overline{x})}{c^i(\underline{x})} < \frac{-\ell(\underline{x})}{\ell(\overline{x})} \frac{\mu_{\underline{\theta}}(\underline{x})}{\mu_{\underline{\theta}}(\overline{x})}$. Section 3.4 revisits the example when Assumption 1 is violated.

reduces to MS's index when agents are correctly specified.[11] Moreover, the same argument as in Theorem 1 implies that if $w(\mu^1, \hat{\mu}^1) > w(\mu^2, \hat{\mu}^2)$, then for all non-trivial A, $W_T(\mu^1, \hat{\mu}^1, A) > W_T(\mu^2, \hat{\mu}^2, A)$ for all sufficiently large T. This result nests both Theorem 1 and MS's characterization, and additionally allows for welfare comparisons across agents who differ both in terms of the true signal structures μ^i they face and in their misperceptions $\hat{\mu}^i$. ▲

3.3 Properties of the Welfare Rankings

Inconsistency between static and dynamic rankings. As the asymmetric updating example illustrates, agent 1 can be strictly better off than agent 2 according to the static ranking, but strictly worse off according to the dynamic ranking. Such inconsistencies between the static and dynamic rankings are not specific to asymmetric updating and can arise under many biases, including forms of misspecified Bayesian learning.

To understand the source of these inconsistencies in general, recall that by Proposition 1, $\hat{\mu}^1$ dominates $\hat{\mu}^2$ according to the static ranking if the interpretation of *each* signal under $\hat{\mu}^1$ is closer to the correctly specified case than under $\hat{\mu}^2$, in the sense of the nested likelihood ratio condition (6). Importantly, the fact that each individual signal is interpreted more accurately does not preclude that some *sequences* of signals are interpreted less accurately. The reason is that (6) allows for the possibility that $\frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)} \geq \frac{\hat{\mu}^1_{\overline{\theta}}(x)}{\hat{\mu}^1_{\underline{\theta}}(x)} \geq \frac{\hat{\mu}^2_{\overline{\theta}}(x)}{\hat{\mu}^2_{\underline{\theta}}(x)}$ for some signals x, but $\frac{\mu_{\overline{\theta}}(y)}{\mu_{\underline{\theta}}(y)} \leq \frac{\hat{\mu}^1_{\overline{\theta}}(y)}{\hat{\mu}^1_{\underline{\theta}}(y)} \leq \frac{\hat{\mu}^2_{\overline{\theta}}(y)}{\hat{\mu}^2_{\underline{\theta}}(y)}$ for other signals y, so misinferences from these signals go in opposite directions. When this is the case, then agent 2's inferences from sequences that contain both x and y might be more accurate than agent 1's if agent 2's errors "cancel out" more. Theorem 1 formalizes the sense in which such canceling out of opposite errors affects the dynamic welfare ranking: The learning efficiency indices $w(\mu, \hat{\mu}^i)$ depend on $\hat{\mu}^i$ only through the indistinguishability sets $\mathcal{I}(\hat{\mu}^i)$, but by (9), these sets depend only on the *relative* interpretations of different signals, as captured by the normal vector $\left(\log \frac{\hat{\mu}^i_{\overline{\theta}}(x)}{\hat{\mu}^i_{\underline{\theta}}(x)}\right)_x$.

By contrast, if condition (6) is strengthened to require misinferences from all signals to go in the same direction, then there is no scope for opposite errors to cancel out. In this case, agent 1 is robustly better off than agent 2 after *any* number of

[11] Moreover, $w(\mu, \mu) = w(\underline{\theta}, \mu, \mu) = w(\overline{\theta}, \mu, \mu) = \min_{\nu \in \Delta(X)} \max\{\mathrm{KL}(\nu, \mu_{\underline{\theta}}), \mathrm{KL}(\nu, \mu_{\overline{\theta}})\}$. That is, our index for correctly specified agents coincides with the Chernoff information distance between signal distributions $\mu_{\underline{\theta}}$ and $\mu_{\overline{\theta}}$ (e.g., Cover and Thomas, 2006).

signal draws, which implies both the static and dynamic welfare rankings:

Proposition 2. *Suppose that*

$$\frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)} \geq \frac{\hat{\mu}^1_{\overline{\theta}}(x)}{\hat{\mu}^1_{\underline{\theta}}(x)} \geq \frac{\hat{\mu}^2_{\overline{\theta}}(x)}{\hat{\mu}^2_{\underline{\theta}}(x)} \forall x \in X \quad or \quad \frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)} \leq \frac{\hat{\mu}^1_{\overline{\theta}}(x)}{\hat{\mu}^1_{\underline{\theta}}(x)} \leq \frac{\hat{\mu}^2_{\overline{\theta}}(x)}{\hat{\mu}^2_{\underline{\theta}}(x)} \forall x \in X. \tag{11}$$

Then for all decision problems A and all T, $W_T(\mu, \hat{\mu}^1, A) \geq W_T(\mu, \hat{\mu}^2, A)$.

In the asymmetric updating example, (11) holds if and only if either $c^2(\underline{x}) \geq c^1(\underline{x}) \geq 1 \geq c^1(\overline{x}) \geq c^2(\overline{x})$ or $c^2(\underline{x}) \leq c^1(\underline{x}) \leq 1 \leq c^1(\overline{x}) \leq c^2(\overline{x})$.

Large biases can be less harmful than vanishingly small biases. Since $W_T(\mu, \hat{\mu}, A)$ is the ex-ante expected payoff according to the true signal structure μ, a correctly specified agent's welfare is clearly at least as high as that of any other agent in all decision problems. Thus, by Theorem 1, $w(\mu, \mu) \geq w(\mu, \hat{\mu})$ for all perceived signal structures $\hat{\mu}$. However, generalizing the asymmetric updating example, the following result shows that there is a class of biases $\hat{\mu}$ that achieve the same maximal efficiency index as the correctly specified case: This class coincides exactly with the seminal model of over-/underinference due to Phillips and Edwards (1966), where all likelihood ratios are distorted by a constant power $c > 0$,[12]

$$\mathrm{PE}(\mu) := \left\{ \hat{\mu} \in \left(\mathbb{R}^X_{++}\right)^\Theta : \exists c > 0 \text{ s.t. } \frac{\hat{\mu}_{\overline{\theta}}(x)}{\hat{\mu}_{\underline{\theta}}(x)} = \left(\frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}\right)^c \forall x \in X \right\}.$$

We note that any $\hat{\mu} \in \mathrm{PE}(\mu)$ satisfies Assumption 1.

Proposition 3. *For any true and perceived signal structures μ and $\hat{\mu}$ satisfying Assumption 1, we have $w(\mu, \hat{\mu}) = w(\mu, \mu)$ if and only if $\hat{\mu} \in \mathrm{PE}(\mu)$.*

Proposition 3 again uses the fact that $w(\mu, \hat{\mu})$ depends on the bias $\hat{\mu}$ only through the indistinguishability set $\mathcal{I}(\hat{\mu})$ given by the hyperplane (9). Thus, if $\hat{\mu}$ yields the same indistinguishability set as the correctly specified case, then $w(\mu, \hat{\mu}) = w(\mu, \mu)$. Conversely, we show that any bias with $w(\mu, \hat{\mu}) = w(\mu, \mu)$ must satisfy $\mathcal{I}(\hat{\mu}) = \mathcal{I}(\mu)$. Given this, the result follows from the observation that distorting likelihood ratios by a constant power c is the only operation that does not affect the hyperplane $\mathcal{I}(\mu)$, as it amounts to multiplying its normal vector $\left(\log \frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}\right)_x$ by the constant c.

[12] See Benjamin (2019)(Section 5.2) for a survey of experimental evidence. Some work in behavioral finance (e.g., Daniel, Hirshleifer, and Subrahmanyam, 1998) uses a Phillips-Edwards model with $c > 1$ to capture overconfidence about signal precision.

To interpret Proposition 3, note that $w(\mu, \hat{\mu}) = w(\mu, \mu)$ does not imply that $W_T(\mu, \hat{\mu}, A) = W_T(\mu, \mu, A)$ for all large T.[13] However, by the proof of Theorem 1, $w(\mu, \hat{\mu}) = w(\mu, \mu)$ entails that in any decision problem, the probability of suboptimal choices under $\hat{\mu}$ and μ vanishes at the same exponential rate.

A more important implication of Proposition 3 is that certain forms of "large" biases are robustly better than other "vanishingly small" biases when agents observe many signal draws: Indeed, fix any bias $\hat{\mu} \in \mathrm{PE}(\mu)$ with a distortion factor c that is arbitrarily different from 1. Consider any sequence of biases $\hat{\mu}^n \notin \mathrm{PE}(\mu)$ such that $\hat{\mu}^n \to \mu$ according to an arbitrary metric on signal structures. Then by Proposition 3, $w(\mu, \mu) = w(\mu, \hat{\mu}) > w(\mu, \hat{\mu}^n)$ for all n. Thus, any Phillips-Edwards bias $\hat{\mu}$, no matter how significant, is less harmful according to the dynamic ranking than any vanishingly small non-Phillips-Edwards biases $\hat{\mu}^n$.

3.4 Dynamic Ranking: Incorrect Asymptotic Beliefs

We now consider the dynamic welfare ranking when Assumption 1 is violated, so agents' asymptotic beliefs may be incorrect in some states. The main insight is that, when agents share the same asymptotic beliefs, this ranking is again generically complete and can be characterized using the learning efficiency index from Section 3.2. However, in states in which mislearning occurs, agents with a *lower* efficiency index are better off.

Suppose that at least one agent's perceived signal structure $\hat{\mu} \in \{\hat{\mu}^1, \hat{\mu}^2\}$ violates Assumption 1 in one of the following three ways:

- Case (i): $\mathrm{KL}(\mu_{\overline{\theta}}, \hat{\mu}_{\overline{\theta}}) < \mathrm{KL}(\mu_{\overline{\theta}}, \hat{\mu}_{\underline{\theta}})$ and $\mathrm{KL}(\mu_{\underline{\theta}}, \hat{\mu}_{\overline{\theta}}) < \mathrm{KL}(\mu_{\underline{\theta}}, \hat{\mu}_{\underline{\theta}})$.
- Case (ii): $\mathrm{KL}(\mu_{\underline{\theta}}, \hat{\mu}_{\overline{\theta}}) > \mathrm{KL}(\mu_{\underline{\theta}}, \hat{\mu}_{\underline{\theta}})$ and $\mathrm{KL}(\mu_{\overline{\theta}}, \hat{\mu}_{\overline{\theta}}) > \mathrm{KL}(\mu_{\overline{\theta}}, \hat{\mu}_{\underline{\theta}})$.
- Case (iii): $\mathrm{KL}(\mu_{\underline{\theta}}, \hat{\mu}_{\overline{\theta}}) < \mathrm{KL}(\mu_{\underline{\theta}}, \hat{\mu}_{\underline{\theta}})$ and $\mathrm{KL}(\mu_{\overline{\theta}}, \hat{\mu}_{\overline{\theta}}) > \mathrm{KL}(\mu_{\overline{\theta}}, \hat{\mu}_{\underline{\theta}})$.

Case (i) obtains if and only if the agent's asymptotic belief is almost surely a point-mass on state $\overline{\theta}$ in both states, so mislearning occurs in state $\underline{\theta}$. Similarly, in case (ii),

[13] Indeed, iterated application of Proposition 1 implies that for any $\hat{\mu}^1, \hat{\mu}^2 \in \mathrm{PE}(\mu)$ whose distortion factors satisfy $1 \leq c^1 < c^2$ or $c^2 < c^1 \leq 1$, we have $W_T(\mu, \hat{\mu}^1, A) \geq W_T(\mu, \hat{\mu}^2, A)$ for all A and T, with strict inequality whenever the chosen acts under $\hat{\mu}^1$ and $\hat{\mu}^2$ differ at some signal sequences.

mislearning occurs in state $\overline{\theta}$, while in case (iii), mislearning occurs in both states.[14]

If agents' asymptotic beliefs are different, it is straightforward to determine the dynamic welfare ranking based on these different beliefs: Specifically, if agent 1 satisfies Assumption 1 (correctly learns in both states), but agent 2 does not, then for any non-trivial decision problem A, $W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A)$ for all sufficiently large T. The same is true if agent 1 satisfies case (i) or (ii), but agent 2 satisfies case (iii). Finally, if agent 1 satisfies case (i) but agent 2 satisfies case (ii) (or vice versa), then which agent's welfare is higher depends on the decision problem A and prior p_0.

Thus, the main challenge is how to compare dynamic welfare when agents 1 and 2 share the same asymptotic beliefs, as is the case for many different learning biases when the payoff-relevant state space is coarse (e.g., binary).[15] In this case, the following result shows that the dynamic welfare ranking is again characterized by the state-dependent efficiency indices $w(\theta, \mu, \hat{\mu}^i)$. To ensure that these indices are well-defined, we impose an assumption on both agents that guarantees that their indistinguishability sets are nonempty. This assumption is always satisfied for misspecified Bayesian agents (i.e., if $\hat{\mu}_\theta \in \Delta(X)$ for all θ) and rules out extreme forms of non-Bayesian learning:[16]

Assumption 2 (Non-degeneracy)**.** For each $\theta \in \Theta$, there exists $\nu \in \Delta(X)$ such that $\mathrm{KL}(\nu, \hat{\mu}_\theta) < \mathrm{KL}(\nu, \hat{\mu}_{\theta'})$ for all $\theta' \neq \theta$.

Theorem 2. *Fix $\hat{\mu}^1$ and $\hat{\mu}^2$ satisfying Assumption 2. Suppose one of the following is true:*

1. *Both $\hat{\mu}^1$ and $\hat{\mu}^2$ satisfy case (i) and*

$$\min\left\{w(\overline{\theta}, \mu, \hat{\mu}^2), w(\underline{\theta}, \mu, \hat{\mu}^1)\right\} < \min\left\{w(\overline{\theta}, \mu, \hat{\mu}^1), w(\underline{\theta}, \mu, \hat{\mu}^2)\right\}.$$

2. *Both $\hat{\mu}^1$ and $\hat{\mu}^2$ satisfy case (ii) and*

$$\min\left\{w(\overline{\theta}, \mu, \hat{\mu}^1), w(\underline{\theta}, \mu, \hat{\mu}^2)\right\} < \min\left\{w(\overline{\theta}, \mu, \hat{\mu}^2), w(\underline{\theta}, \mu, \hat{\mu}^1)\right\}.$$

[14] We do not consider the case when Assumption 1 is violated with equality, i.e., $\mathrm{KL}(\mu_\theta, \hat{\mu}_\theta) = \mathrm{KL}(\mu_\theta, \hat{\mu}_{\theta'})$ for some $\theta \neq \theta'$. For any such bias $\hat{\mu}$, beliefs in state θ a.s. cycle indefinitely, which leads to lower dynamic welfare than correct learning but higher dynamic welfare than incorrect learning.

[15] In general finite state spaces, the following approach based on learning efficiency is also relevant in comparing various biases whose asymptotic beliefs differ in some states; see Section 4.1.

[16] Under binary states, Assumption 2 rules out non-Bayesian biases for which all signal realizations lead an agent to update her belief in the same direction.

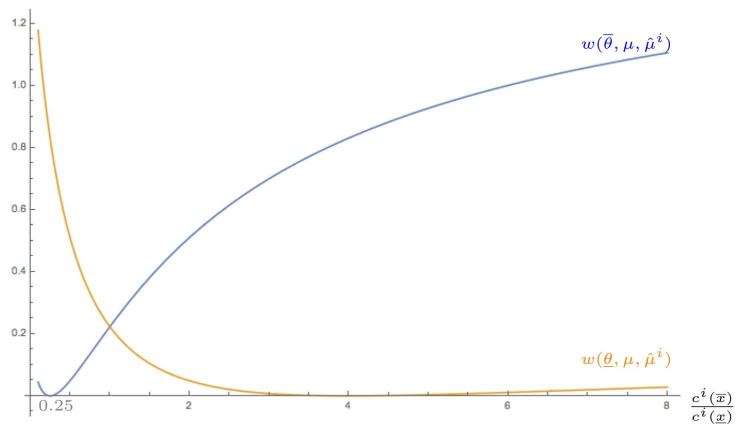

Figure 2: *Efficiency indices when $\mu_{\overline{\theta}}(\overline{x}) = 0.8 = \mu_{\underline{\theta}}(\underline{x})$. Figure 1 discussed the case $1/4 < \frac{c^i(\overline{x})}{c^i(\underline{x})} < 4$ (correct learning). Case (i) corresponds to $\frac{c^i(\overline{x})}{c^i(\underline{x})} > 4$. As depicted, for any $\hat{\mu}, \hat{\mu}'$ in this region, $w(\underline{\theta}, \mu, \hat{\mu}) < w(\overline{\theta}, \mu, \hat{\mu}')$. Thus, by Theorem 2, dynamic welfare is determined by $w(\underline{\theta}, \mu, \hat{\mu}^i)$ and is higher the lower $w(\underline{\theta}, \mu, \hat{\mu}^i)$, i.e., the smaller $\frac{c^i(\overline{x})}{c^i(\underline{x})}$. Analogously, case (ii) corresponds to $\frac{c^i(\overline{x})}{c^i(\underline{x})} < 1/4$, and in this case welfare is higher the lower $w(\overline{\theta}, \mu, \hat{\mu}^i)$, i.e., the greater $\frac{c^i(\overline{x})}{c^i(\underline{x})}$.*

3. *Both $\hat{\mu}^1$ and $\hat{\mu}^2$ satisfy case (iii) and $w(\mu, \hat{\mu}^1) < w(\mu, \hat{\mu}^2)$.*

Then for any non-trivial decision problem A, there exists T^* such that for all $T \geq T^*$, $W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A)$.

To understand the result, consider the first possibility, where both agents' asymptotic beliefs are always a point-mass on $\overline{\theta}$. Then in state $\overline{\theta}$, both agents choose an ex-post optimal act in the long run, so in this case a faster rate of convergence, i.e., a higher efficiency index $w(\overline{\theta}, \mu, \hat{\mu})$, is better for welfare. However, in state $\underline{\theta}$, both agents choose the same ex-post *suboptimal* act in the long run, so in this case *slower* convergence, i.e., a lower $w(\underline{\theta}, \mu, \hat{\mu})$, is better for welfare. Thus, if both (a) $w(\overline{\theta}, \mu, \hat{\mu}^2) < w(\overline{\theta}, \mu, \hat{\mu}^1)$ and (b) $w(\underline{\theta}, \mu, \hat{\mu}^1) < w(\underline{\theta}, \mu, \hat{\mu}^2)$, then agent 1 is better off. Theorem 2 shows that the same conclusion obtains under a weaker condition that only compares the minima of both sides of inequalities (a) and (b).

Asymmetric updating: Returning to the illustrative example, again let $\ell(x) := \log \frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}$. Agents satisfy case (i) iff $\frac{c^i(\overline{x})}{c^i(\underline{x})} > \frac{-\ell(\underline{x})}{\ell(\overline{x})} \frac{\mu_{\underline{\theta}}(\underline{x})}{\mu_{\underline{\theta}}(\overline{x})}$, and satisfy case (ii) iff $\frac{c^i(\overline{x})}{c^i(\underline{x})} < \frac{-\ell(\underline{x})}{\ell(\overline{x})} \frac{\mu_{\overline{\theta}}(\underline{x})}{\mu_{\overline{\theta}}(\overline{x})}$. Case (iii) cannot arise. Assumption 2 holds as all distortion factors $c^i(x)$ are positive. As in Section 3.2, each agent's indistinguishability condition has a unique solution ν^i given by (10). This again implies that the efficiency indices in each state depend on $\hat{\mu}^i$ only through the ratio $c^i(\overline{x})/c^i(\underline{x})$. Figure 2 plots the efficiency indices as a function of $c^i(\overline{x})/c^i(\underline{x})$ and shows that, just as under correct learning, higher dynamic welfare in case (i) and (ii) corresponds to $c^i(\overline{x})/c^i(\underline{x})$ being closer to 1.

3.5 Examples

We present two additional examples that illustrate how our approach can be used to understand the welfare implications of commonly studied learning biases through their effect on short-/medium-run beliefs. For simplicity, we assume binary signals, $X = \{\underline{x}, \overline{x}\}$ with $\mu_{\overline{\theta}}(\overline{x}) > \mu_{\underline{\theta}}(\overline{x})$.

First, we provide an example of a bias for which a natural parametric quantification of its severity is consistent with both the static and dynamic welfare rankings:

Example 1 (Partisan Bias)**.** Consider non-Bayesian updating under partisan bias, where agents $i = 1, 2$ satisfy $\frac{\hat{\mu}^i_{\overline{\theta}}(x)}{\hat{\mu}^i_{\underline{\theta}}(x)} = \eta^i(x)\frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}$, for some distortion functions $\eta^i(\cdot)$ with $\eta^2(x) > \eta^1(x) > 1$ for all x.[17] That is, both agents distort inferences from all signals in the direction of state $\overline{\theta}$ (e.g., the superiority of a right-wing political policy), but agent 2's bias is more severe.

Note that $\frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)} < \frac{\hat{\mu}^1_{\overline{\theta}}(x)}{\hat{\mu}^1_{\underline{\theta}}(x)} < \frac{\hat{\mu}^2_{\overline{\theta}}(x)}{\hat{\mu}^2_{\underline{\theta}}(x)}$ for all x. Hence, Proposition 2 implies that in every decision problem, agent 2 is worse off than agent 1 for all T. This shows that, regardless of whether or not agents' asymptotic beliefs are correct, more severe forms of partisan bias are always robustly welfare-decreasing in both static and dynamic settings, as they lead to every sequence of signals being interpreted less accurately. ▲

For the next bias, more severe forms need not be worse according to the static ranking, but are robustly more harmful according to the dynamic ranking:

Example 2 (Overconfidence)**.** Suppose agents $i = 1, 2$ are misspecified Bayesians with $\hat{\mu}^i \in \Delta(X)$ and $\hat{\mu}^2_\theta(\overline{x}) > \hat{\mu}^1_\theta(\overline{x}) > \mu_\theta(\overline{x})$ for each θ. That is, interpreting signal $\overline{x}$ as "success" and $\underline{x}$ as "failure," both agents are overconfident, in the sense that they overestimate success probabilities in all states, but agent 2's bias is more severe. Assume that agents' perceptions also satisfy $\hat{\mu}^i_{\overline{\theta}}(\overline{x}) > \hat{\mu}^i_{\underline{\theta}}(\overline{x})$, which implies the comonotonic likelihood ratio property (7).

For the static welfare ranking, observe that the fact that $\hat{\mu}^2_\theta(\overline{x}) > \hat{\mu}^1_\theta(\overline{x}) > \mu_\theta(\overline{x})$ for all θ does not imply condition (6). That is, more overconfident agents may interpret

[17] See, e.g., Bohren and Hauser (2018), Thaler (2019), and references therein to applications of the model in political science.

some signals more accurately. Hence, by Proposition 1 and (7), if only a single signal is observed, there may be decision problems in which the more overconfident agent 2 performs strictly better.

For the dynamic ranking, note that for each $i = 1, 2$, there is a unique distribution $\nu^i \in \Delta(X)$ satisfying the indistinguishability condition $\sum_{x=\overline{x},\underline{x}} \nu^i(x) \log \frac{\hat{\mu}^i_{\overline{\theta}}(x)}{\hat{\mu}^i_{\underline{\theta}}(x)} = 0$. A straightforward calculation shows that

$$\nu^*(\overline{x}) < \nu^1(\overline{x}) < \nu^2(\overline{x}), \tag{12}$$

where ν^* is the indistinguishable distribution in the correctly specified case. Intuitively, because an overconfident agent overestimates the probability of high signals in both states, the signal distribution that makes her unable to distinguish between $\underline{\theta}$ and $\overline{\theta}$ must also feature a larger fraction of high signals relative to the correctly specified case, and more so the greater the agent's overconfidence.

This implies that, in any decision problem, the more overconfident agent 2 is worse off if sufficiently many signals are observed. When agent 2 mislearns (i.e., converges to a point-mass on $\underline{\theta}$ in both states) but agent 1 learns the state correctly, the welfare loss is driven by the difference in asymptotic beliefs, which have been the focus of much existing work (e.g., Heidhues, Koszegi, and Strack, 2018). However, by Theorems 1–2, agent 2 is worse off even when both agents' asymptotic beliefs are the same (either both correct or both incorrect): This is because (12) implies that greater overconfidence increases the probability of suboptimal choices in the medium run.[18] ▲

4 Extensions

4.1 General States

Consider a general finite state space Θ. We focus on extending the characterization of the dynamic welfare ranking.[19] Under binary states, Theorems 1–2 showed that the learning efficiency index yields a generically complete ranking over biases that induce the same asymptotic beliefs. With more than two states, a complete ranking

[18] The reasoning is analogous to the asymmetric updating example with $c^i(\overline{x})/c^i(\underline{x}) \leq 1$ for $i = 1, 2$.

[19] For the static ranking, a sufficient condition for $\hat{\mu}^1$ to dominate $\hat{\mu}^2$ is that after each signal, the interim belief under $\hat{\mu}^1$ is a convex combination of the beliefs under μ and $\hat{\mu}^2$, as in Section 3.1.

is no longer possible. For example, consider any decision problem A in which all acts yield the same payoffs in states θ and θ'. Then any bias that only affects inferences between states θ and θ' is payoff-irrelevant in A, but may affect welfare in decision problems in which the payoffs in states θ and θ' differ.

However, we show that, up to controlling for such "redundancies" of equivalent states, the learning efficiency index can be generalized to again yield a generically complete ranking. Formally, given any decision problem A, we consider the partition S_A over Θ whose cells are

$$S_A(\theta) := \{\theta' \in \Theta : \operatorname*{argmax}_{a \in A} a_{\theta'} = \operatorname*{argmax}_{a \in A} a_\theta\} \text{ for each } \theta.$$

That is, S_A divides Θ into equivalence classes of states that share the same ex-post optimal act (but are not necessarily payoff-equivalent for all acts in A). We then extend the learning efficiency index to rank welfare across all decision problems A that induce the same partition S_A. Note that when Θ is binary, all non-trivial decision problems induce the same partition.

Given any non-degenerate partition S over Θ (i.e., with $S(\theta) \neq \Theta$), we define the ***S-learning efficiency index*** by $w(\mu, \hat{\mu}, S) := \min_\theta w(\theta, \mu, \hat{\mu}, S)$, where

$$w(\theta, \mu, \hat{\mu}, S) := \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \quad \text{subject to} \quad \min_{\theta' \in S(\theta)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) = \min_{\theta' \notin S(\theta)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}).$$

Generalizing Definition 1, the indistinguishability condition $\min_{\theta' \in S(\theta)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) = \min_{\theta' \notin S(\theta)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'})$ in state θ now captures the set of empirical distributions ν based on which the agent is unable to distinguish whether the state is in $S(\theta)$ or $\Theta \setminus S(\theta)$. As before, $w(\theta, \mu, \hat{\mu}, S)$ measures how unlikely such empirical distributions ν are under the true signal distribution μ_θ, and the learning efficiency index $w(\mu, \hat{\mu}, S)$ considers the worst-case measure across all states.

The following condition generalizes Assumption 1. Given a partition S of Θ, we require that in each state θ, the agent asymptotically assigns probability one to the correct *cell* $S(\theta)$. However, the condition does not restrict asymptotic beliefs over states in $S(\theta)$, allowing for some forms of mislearning and for comparisons across biases whose asymptotic beliefs need not coincide in each state:

Assumption 3 (S-consistency)**.** For every $\theta \in \Theta$, we have $\min_{\theta' \in S(\theta)} \mathrm{KL}(\mu_\theta, \hat{\mu}_{\theta'}) < \min_{\theta' \notin S(\theta)} \mathrm{KL}(\mu_\theta, \hat{\mu}_{\theta'})$.

Under Assumption 3, we obtain the following generalization of Theorem 1: The S-learning efficiency index characterizes when agent 1's welfare exceeds agent 2's welfare in all decision problems with partition S. We focus on perceived signal structures that are non-degenerate in the sense of Assumption 2, and on ***regular*** decision problems A, where $a_\theta \neq a'_\theta$ for all θ and distinct $a, a' \in A$.

Theorem 3. *Let S be a non-degenerate partition over Θ. Fix any true signal structure μ and perceived signal structures $\hat{\mu}^1$ and $\hat{\mu}^2$ satisfying Assumptions 2–3. Suppose $w(\mu, \hat{\mu}^1, S) > w(\mu, \hat{\mu}^2, S)$. Then for any non-trivial and regular decision problem A with $S_A = S$, there exists T^* such that for all $T \geq T^*$, $W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A)$.*

Thus, up to restricting to decision problems that feature the same classes of equivalent states, the dynamic welfare ranking is again generically complete.[20] As in Section 3.3, biases of the Phillips-Edwards form achieve the maximal efficiency index for any partition S.

Online Appendix B provides an analogous extension of Theorem 2, where for a given partition S, we assume that in each state θ, both agents' asymptotic beliefs assign probability one to the same, but possibly incorrect, cell $S(\theta')$.

4.2 Uncertainty about Signal Structures

Consider an agent who, instead of dogmatically perceiving a particular signal structure $\hat{\mu}$, entertains multiple possible signal structures and jointly updates about both payoff-relevant states and signal structures. Such uncertainty about the signal structure can be interpreted as capturing some "cautiousness" against misspecification (e.g., Acemoglu, Chernozhukov, and Yildiz, 2016). We extend the dynamic welfare ranking to this setting. Relative to the baseline model without uncertainty about signal structures, a key difference is that now correctly specified agents can be robustly worse off than misspecified agents.

As in Section 2, there is a finite set Θ of payoff-relevant states from which the true state is drawn according to a full-support distribution p_0, and a fixed true signal structure $\mu := (\mu_\theta)_{\theta \in \Theta}$. The agent learns jointly about payoff-relevant states and

[20] Analogous to Remark 2, Theorem 3 extends to the case where both the true signal structure μ^i and perceived signal structures $\hat{\mu}^i$ differ across agents. This result nests the general state characterization for the correctly specified case in Moscarini and Smith (2002), who restrict attention to decision problems A for which the ex-post optimal act in each state is different. Moscarini and Smith's (2002) proof approach relies on correctly specified agents.

signal structures, under a possibly misspecified model that may assign zero probability to the true signal structure: Let $\hat{M}$ denote the set of signal structures the agent deems possible, where $\hat{\mu} := (\hat{\mu}_\theta)_{\theta \in \Theta} \in (\Delta(X))^\Theta$ for each $\hat{\mu} \in \hat{M}$.[21] The agent's prior belief is some full-support $q_0 \in \Delta(\hat{M} \times \Theta)$, where we do not require that $\text{marg}_\Theta q_0 = p_0$ (recall the discussion in footnote 4). The agent is ***correctly specified*** if the true signal structure μ is contained in her subjective model $\hat{M}$.

Upon observing a signal sequence $x^T = (x_1, \cdots, x_T)$, generated i.i.d. according to the true signal structure, the agent Bayesian-updates her belief to

$$q_T(\hat{\mu}, \theta \mid x^T) = \frac{q_0(\hat{\mu}, \theta) \prod_{t=1}^T \hat{\mu}_\theta(x_t)}{\sum_{(\hat{\mu}', \theta') \in \hat{M} \times \Theta} q_0(\hat{\mu}', \theta') \prod_{t=1}^T \hat{\mu}'_{\theta'}(x_t)}. \tag{13}$$

Given any decision problem $A \subseteq \mathbb{R}^\Theta$, the agent chooses a subjectively optimal act according to her posterior belief $\text{marg}_\Theta q_T$ over payoff-relevant states, where as before, we assume away indifferences. Analogous to (5), define the agent's welfare $W_T(\mu, q_0, A)$ as her objective expected payoff according to the true signal structure μ and prior p_0 over Θ.

For simplicity, we focus on binary payoff-relevant states, $\Theta = \{\underline{\theta}, \overline{\theta}\}$. Moreover, we restrict attention to agents who correctly learn the payoff-relevant state θ, as is ensured by imposing Assumption 1 on each $\hat{\mu} \in \hat{M}$. Thus, we isolate the effect of uncertainty about the signal structure on the speed of convergence of beliefs about θ, rather than on asymptotic beliefs about θ:

Assumption 4. For any distinct $\theta, \theta' \in \Theta$ and each $\hat{\mu} \in \hat{M}$, we have $\text{KL}(\mu_\theta, \hat{\mu}_\theta) < \text{KL}(\mu_\theta, \hat{\mu}_{\theta'})$.

We obtain the following generalization of Theorem 1. Define the ***learning efficiency index*** by $w(\mu, \hat{M}) := \min_\theta w(\theta, \mu, \hat{M})$, where

$$w(\theta, \mu, \hat{M}) := \min_{\nu \in \Delta(X)} \text{KL}(\nu, \mu_\theta) \quad \text{subject to} \quad \min_{\hat{\mu} \in \hat{M}} \text{KL}(\nu, \hat{\mu}_{\underline{\theta}}) = \min_{\hat{\mu} \in \hat{M}} \text{KL}(\nu, \hat{\mu}_{\overline{\theta}}).$$

The indistinguishability condition $\min_{\hat{\mu} \in \hat{M}} \text{KL}(\nu, \hat{\mu}_{\underline{\theta}}) = \min_{\hat{\mu} \in \hat{M}} \text{KL}(\nu, \hat{\mu}_{\overline{\theta}})$ again captures the empirical signal distributions ν that do not allow the agent to tell apart states $\underline{\theta}$ and $\overline{\theta}$. Here, in each state θ, the agent uses the signal distribution $\hat{\mu}_\theta$ that

[21] We assume that each perceived signal distribution $\hat{\mu}_\theta \in \Delta(X)$ is a probability measure, but the same analysis extends to the case where $\hat{\mu}_\theta \in \mathbb{R}^X_{++}$, subject to imposing an analog of Assumption 2.

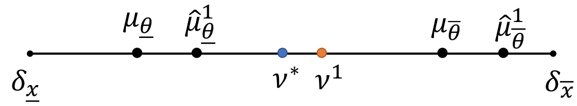

Figure 3: *Indistinguishable distributions ν^* under $\hat{M} = \{\mu\}$ and ν^1 under $\hat{M}^1 = \{\mu, \hat{\mu}^1\}$.*

comes closest to ν among all the distributions she deems possible. Note that the agent's prior belief over signal structures in $\hat{M}$ does not affect the efficiency index. When $\hat{M}$ is a singleton, the index reduces to the one in Definition 1.

Theorem 4. *Fix any μ and $q_0^i \in \Delta(\hat{M}^i \times \Theta)$ $(i = 1, 2)$ with $\hat{M}^i$ satisfying Assumption 4. Suppose $w(\mu, \hat{M}^1) > w(\mu, \hat{M}^2)$. Then for any non-trivial decision problem A, there exists T^* such that for all $T \geq T^*$, $W_T(\mu, q_0^1, A) > W_T(\mu, q_0^2, A)$.*

Clearly, based on objective welfare, an agent with subjective model $\hat{M} = \{\mu\}$, i.e., who is certain of the true signal structure, weakly outperforms any other agent in all decision problems and for all T. Thus, the learning efficiency index $w(\mu, \{\mu\})$ is maximal.

However, a key implication of Theorem 4 is that correctly specified but *uncertain* agents can be robustly worse off than misspecified agents when T is large. Indeed, the following example exhibits $\hat{M}^1, \hat{M}^2$ such that $\mu \in \hat{M}^1 \setminus \hat{M}^2$ but $w(\mu, \hat{M}^2) > w(\mu, \hat{M}^1)$. Thus, in any decision problem, the correctly specified agent 1's welfare is strictly lower than that of the misspecified agent 2 after sufficiently many signal observations. Note that, by Assumption 4, both agents learn the true payoff-relevant state θ asymptotically; however, agent 1 is robustly worse off because her beliefs about θ converge more slowly.[22] Moreover, since the learning efficiency indices $w(\mu, \hat{M}^i)$ do not depend on the priors q_0^i, agent 1 is worse off than agent 2 even if her prior assigns high probability to the true signal structure μ.

Example 3 (Correctly specified agents can be worse off)**.** Suppose $X = \{\overline{x}, \underline{x}\}$. Consider some true signal structure with $\mu_{\underline{\theta}}(\overline{x}) < \mu_{\overline{\theta}}(\overline{x})$. As noted, if $\hat{M} = \{\mu\}$, the learning efficiency index is maximal, and the indistinguishability set in this case consists of a single distribution ν^*. The key observation is that if an agent is correctly

[22]Absent Assumption 4, some correctly specified but uncertain agents may fail to learn the true payoff-relevant state asymptotically, due to (non-generic) identification problems that lead to incomplete learning (e.g., if $\hat{M} = \{\mu, \hat{\mu}\}$ with $\mu_\theta = \hat{\mu}_{\theta'}$ for all $\theta \neq \theta'$). Such agents are dynamically worse off than any agent satisfying Assumption 4.

specified but uncertain about the signal structure, this may reduce her efficiency index by altering her indistinguishability set. For example, suppose agent 1 is uncertain between the true signal structure and some perception $\hat{\mu}^1$ that is overconfident in the sense of Example 2. That is, $\hat{M}^1 = \{\mu, \hat{\mu}^1\}$, where $\hat{\mu}^1_\theta(\overline{x}) > \mu_\theta(\overline{x})$ for all θ. Then, as shown in Figure 3, the unique indistinguishable distribution ν^1 satisfies $\nu^1(\overline{x}) > \nu^*(\overline{x})$.

By contrast, consider any misspecified agent 2 with a unique indistinguishable distribution ν^2 that is closer than ν^1 to ν^*. For example, suppose $\hat{M}^2 = \{\hat{\mu}^2\}$, where $\hat{\mu}^2$ is also overconfident (i.e., $\hat{\mu}^2_\theta(\overline{x}) > \mu_\theta(\overline{x})$ for all θ) but sufficiently less so than $\hat{\mu}^1$. Then $w(\mu, \mu) > w(\mu, \hat{M}^2) > w(\mu, \hat{M}^1)$. ▲

This example highlights a rationale for why larger subjective models $\hat{M}$, i.e., greater "caution," can reduce an agent's welfare: Entertaining an additional possible signal structure may lead to more medium-run mistakes, by increasing the likelihood of observing "confusing" signal sequences that make it impossible to distinguish between different states.[23] At the same time, learning under any model $\hat{M}$ at least outperforms learning under the worst signal structure $\hat{\mu}$ in its support:[24]

Proposition 4. *Fix any μ and $\hat{M}$ satisfying Assumption 4. Then*

$$w(\mu, \hat{M}) \geq \min_{\hat{\mu} \in \hat{M}} w(\mu, \hat{\mu}).$$

5 Concluding Remarks

This paper conducts a robust comparison of objective welfare across a wide range of learning biases. Our core results characterize this welfare comparison in dynamic settings, using a learning efficiency index: Complementing a focus in the literature on asymptotic beliefs, this index determines the speed of belief convergence under each bias, by quantifying the likelihood with which agents encounter signal sequences that do not allow them to distinguish different states. We highlight that learning efficiency

[23] This point is reminiscent of the phenomenon of "overfitting," where statistical predictions can become less accurate if a model uses too many variables relative to the size of observations (e.g., Montiel Olea, Ortoleva, Pai, and Prat, 2019). It also relates to Blume and Easley (2006), who in a competitive market setting with multidimensional states show that agents whose prior is correctly supported on a lower-dimensional set of states outperform agents with a higher-dimensional full-support prior.

[24] We also note that if $\mu \notin \hat{M}$, it is not in general true that $w(\mu, \hat{M}) \leq \max_{\hat{\mu} \in \hat{M}} w(\mu, \hat{\mu})$. That is, in some cases, entertaining uncertainty over multiple misspecified signal structures can robustly outperform learning under each individual signal structure.

can be strictly lower for smaller biases or for biases that are less harmful in static settings, and that correctly specified but uncertain agents can be outperformed by misspecified agents. We apply our results to provide welfare-founded quantifications of the severity of various commonly studied biases.

There are several directions in which to extend our analysis. First, beyond the focus on exogenous i.i.d. signals in this paper, it is natural to consider settings with (i) intertemporally correlated signals or (ii) endogenous signals. For (i), it is straightforward to extend the dynamic characterization based on learning efficiency to exogenous but Markovian signals, by invoking appropriate generalizations of Sanov's theorem. This extension can accommodate additional biases, such as the gambler's and hot-hand fallacies (e.g., Rabin and Vayanos, 2010) and intertemporal correlation neglect (e.g., Ortoleva and Snowberg, 2015). For (ii), our companion note (in preparation) extends the dynamic welfare analysis to simple settings in which the agent's current belief can affect the true and perceived signal distributions, capturing some forms of misspecified active learning or belief-dependent updating (e.g., confirmation bias). One finding is that such endogenous signals always lead to lower learning efficiency than the corresponding exogenous signal structures.

Second, we have focused on learning biases that can be represented as Bayesian updating under some possibly incorrect perception of signal likelihoods. Important features of this setting are that (i) agents' posterior beliefs depend only on the empirical signal distribution (rather than the exact signal sequence), and (ii) for almost all empirical signal distributions, agents' posteriors become confident in a particular state as $T \to \infty$. The same large deviation techniques on which our dynamic characterization relies can also be used to conduct dynamic welfare comparisons for other non-Bayesian learning models that share these features.[25]

Finally, our analysis can serve as a starting point to robustly quantify the welfare implications of some design interventions for biased learning. As an illustration, consider the effect of coarsening signals, i.e., partitioning the signal space X and only revealing the realized cell. For correctly specified agents, coarsening signals always reduces welfare. However, for a misspecified Bayesian agent, some forms of

[25]For example, in Cripps (2018), an agent distorts beliefs based on some homeomorphism $F : \Delta(\Theta) \to \Delta(\Theta)$ before performing Bayesian updating, and then applies the inverse map F^{-1} to the updated belief before making a decision. Assuming that F maps each point-mass belief to itself (to guarantee that asymptotic beliefs are correct), similar arguments as in Theorem 1 imply that this bias has the same learning efficiency index as the correctly specified case.

coarsening can robustly improve welfare in dynamic settings, either by shifting the agent's indistinguishability set closer to the correctly specified case or by affecting the agent's asymptotic beliefs. In the former case, we can use our learning efficiency index to quantify and decompose the effect of coarsening: Let μ, $\hat{\mu}$ (resp. μ', $\hat{\mu}'$) denote the true and perceived signal structures before (resp. after) coarsening. Then

$$w(\mu', \hat{\mu}') - w(\mu, \hat{\mu}) = \underbrace{[w(\mu', \mu') - w(\mu, \mu)]}_{\text{information loss}} + \underbrace{[w(\mu, \mu) - w(\mu, \hat{\mu})] - [w(\mu', \mu') - w(\mu', \hat{\mu}')]}_{\text{difference in welfare costs of misperceptions}}.$$

The first term captures the information loss due to coarsening and is always negative. The second term compares the extent of welfare loss due to misperception before and after coarsening. This term can be positive and overwhelm the first term if the welfare loss before coarsening, $w(\mu, \mu) - w(\mu, \hat{\mu})$, is sufficiently larger than the welfare loss after coarsening, $w(\mu', \mu') - w(\mu', \hat{\mu}')$. Online Appendix C provides an example in the context of overconfidence.

A Proofs

A.1 Proof of Proposition 1

We prove a slight generalization of Proposition 1 that allows decision problems to feature ties: For each decision problem A, we assume some arbitrary strict total order $\succ_A$ over A such that whenever (4) admits multiple solutions for an agent, she chooses the $\succ_A$-optimal act among these solutions. Note that both agents are assumed to use the same tie-breaking rule.

Suppose that (6) holds for all $x \in X$. Fix any decision problem A. For each realized signal x, let p_x denote the posterior belief under μ, and $\hat{p}^i_x$ denote the posterior belief under $\hat{\mu}^i$ $(i = 1, 2)$. Let $a^i_x \in \text{argmax}_{a\in A}\, a \cdot \hat{p}^i_x$ denote the action chosen by agent i (with tie-breaking according to $\succ_A$ in case of indifference).

By (6) and the fact that $\Theta = \{\underline{\theta}, \overline{\theta}\}$ is binary, there is $\beta_x \in [0, 1]$ such that $\hat{p}^1_x = \beta_x p_x + (1 - \beta_x)\hat{p}^2_x$. Thus, for all $a \in A$,

$$a \cdot \hat{p}^1_x = \beta_x a \cdot p_x + (1 - \beta_x)a \cdot \hat{p}^2_x. \tag{14}$$

We claim that

$$a^1_x \cdot p_x \geq a^2_x \cdot p_x. \tag{15}$$

Indeed, if $\beta_x = 0$, then by (14), both agents share the same interim payoffs. Thus, $a^1_x = a^2_x$ (using the assumption that agents follow the same same tie-breaking rule), which implies (15). If $\beta_x > 0$, then (15) follows from (14) and the fact that

$$a^1_x \cdot \hat{p}^1_x \geq a^2_x \cdot \hat{p}^1_x \quad \text{and} \quad a^1_x \cdot \hat{p}^2_x \leq a^2_x \cdot \hat{p}^2_x.$$

Since (15) holds for all x, this ensures

$$W_1(\mu, \hat{\mu}^1, A) = \sum_\theta p_0(\theta) \sum_x \mu_\theta(x) a^1_x \cdot p_x \geq \sum_\theta p_0(\theta) \sum_x \mu_\theta(x) a^2_x \cdot p_x = W_1(\mu, \hat{\mu}^2, A).$$

For the converse direction, assume that $\mu, \hat{\mu}^1, \hat{\mu}^2$ satisfy the comonotonic likelihood ratio property for some linear order $>$ on X. Suppose that (6) is violated at some $x^* \in X$. We will construct a decision problem A such that $W_1(\mu, \hat{\mu}^1, A) < W_1(\mu, \hat{\mu}^2, A)$.

Let $\ell^i(x) := \frac{\hat{\mu}^i_{\overline{\theta}}(x)}{\hat{\mu}^i_{\underline{\theta}}(x)}$ and $\ell(x) := \frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}$ for each x. Since (6) is violated at x^*, we either have (i) $\ell^1(x^*) > \ell^2(x^*)$ and $\ell^1(x^*) > \ell(x^*)$, or (ii) $\ell^1(x^*) < \ell^2(x^*)$ and $\ell^1(x^*) < \ell(x^*)$.

We consider only case (i), as the argument for case (ii) is analogous. Take any $\ell^* \in (\max\{\ell(x^*), \ell^2(x^*)\}, \ell^1(x^*))$. The comonotonic likelihood ratio property ensures that for each $i = 1, 2$, we have $\ell(x), \ell^i(x) > \ell^*$ if $x > x^*$, while we have $\ell(x), \ell^i(x) < \ell^*$ if $x < x^*$.

Consider a decision problem $A = \{\overline{a}, \underline{a}\}$ such that $\overline{a}_{\overline{\theta}} - \underline{a}_{\overline{\theta}} > 0 > \overline{a}_{\underline{\theta}} - \underline{a}_{\underline{\theta}}$ and $(\overline{a}_{\overline{\theta}} - \underline{a}_{\overline{\theta}})\ell^* \frac{p_0(\overline{\theta})}{p_0(\underline{\theta})} + \overline{a}_{\underline{\theta}} - \underline{a}_{\underline{\theta}} = 0$. Let a_x (resp. a_x^i) denote the act that maximizes the conditional expected payoff according to μ (resp. $\hat{\mu}^i$) at signal x. By construction, $a_x = a_x^1 = a_x^2 = \overline{a}$ for $x > x^*$, $a_x = a_x^1 = a_x^2 = \underline{a}$ for $x < x^*$, but $a_x = a_x^2 = \underline{a} \neq \overline{a} = a_x^1$ for $x = x^*$. This implies $W_1(\mu, \mu, A) = W_1(\mu, \hat{\mu}^2, A) > W_1(\mu, \hat{\mu}^1, A)$, as desired. □

A.2 Proof of Theorems 1 and 3

Below, we prove Theorem 3. Theorem 1 follows immediately from Theorem 3: Indeed, Assumption 1 implies Assumptions 2–3; moreover, for binary states, it is without loss of generality to focus on regular decision problems.[26]

Section A.2.1 first analyzes an agent's choices over pairs of acts, while Section A.2.2 considers general decision problems. Section A.2.3 proves Theorem 3. Throughout, we let $\mathbb{P}_\theta$ denote the probability measure over signal sequences induced by repeated i.i.d. draws according to the true signal distribution μ_θ in state θ.

A.2.1 Binary comparisons over acts

Consider an agent with perceived signal structure $\hat{\mu}$. Fix any acts a and a'. Define the set of states where a is preferred to a' ex-post by

$$\Theta_{a \succ a'} := \{\theta \in \Theta : a_\theta > a'_\theta\}.$$

Let $\log \hat{\mu}_\theta := (\log \hat{\mu}_\theta(x))_{x \in X} \in \mathbb{R}^X$. For any $\varepsilon \in \mathbb{R}$, define

$$C^\varepsilon_{aa'} := \left\{ \nu \in \Delta(X) : \max_{\theta \in \Theta_{a \succ a'}} \nu \cdot \log \hat{\mu}_\theta \geq \max_{\theta \in \Theta_{a' \succ a}} \nu \cdot \log \hat{\mu}_\theta + \varepsilon \right\}.$$

Letting ν_t denote the empirical signal distribution following t draws of signals,

[26] To see the latter, suppose A is not regular. Then some distinct acts $a, a' \in A$ have the same payoff in some state. When Θ is binary, this implies $a \geq a'$ or $a' \geq a$. Then removing the dominated act does not change the agent's choices since her posterior after every signal history has full support.

observe that the agent's posterior can be expressed as

$$p_t(\theta) = \frac{e^{t\nu_t \cdot \log \hat{\mu}_\theta + \log p_0(\theta)}}{\sum_{\theta'} e^{t\nu_t \cdot \log \hat{\mu}_{\theta'} + \log p_0(\theta')}}, \quad \forall \theta. \tag{16}$$

Given this, the following lemma shows that for large t, the set of empirical signal distributions under which the agent prefers a to a' can be bounded using $C^\varepsilon_{aa'}$:

Lemma 1. *Fix any $\varepsilon > 0$ and acts a, a' with neither $a \geq a'$ nor $a' \geq a$. There exists $\bar{t}$ such that for all $t \geq \bar{t}$,*

$$\begin{aligned} C^\varepsilon_{aa'} &\subseteq \left\{ \nu \in \Delta(X) : \sum_{\theta \in \Theta} (a_\theta - a'_\theta) e^{t\nu \cdot \log \hat{\mu}_\theta + \log p_0(\theta)} > 0 \right\} \\ &\subseteq \left\{ \nu \in \Delta(X) : \sum_{\theta \in \Theta} (a_\theta - a'_\theta) e^{t\nu \cdot \log \hat{\mu}_\theta + \log p_0(\theta)} \geq 0 \right\} \subseteq C^{-\varepsilon}_{aa'}. \end{aligned}$$

Proof. Since neither $a \geq a'$ nor $a' \geq a$, both $\Theta_{a \succ a'}$ and $\Theta_{a' \succ a}$ are nonempty. Thus, there exists a sufficienctly large $K > 0$ such that

$$\begin{aligned} \frac{1}{K} &\leq \min_{\theta \in \Theta_{a \succ a'}} (a_\theta - a'_\theta), \min_{\theta \in \Theta_{a' \succ a}} (a'_\theta - a_\theta), \\ K &\geq \max_{\theta \in \Theta_{a \succ a'}} (a_\theta - a'_\theta), \max_{\theta \in \Theta_{a' \succ a}} (a'_\theta - a_\theta). \end{aligned}$$

Pick $\bar{t}$ such that

$$\frac{1}{K} e^{\bar{t}\varepsilon} \min_{\theta, \theta' \in \Theta} \frac{p_0(\theta)}{p_0(\theta')} - |\Theta| K > 0.$$

To show the first desired inclusion, note that for any $\nu \in C^\varepsilon_{aa'}$ and $t \geq \bar{t}$,

$$\begin{aligned} &\textstyle\sum_{\theta \in \Theta} \quad (a_\theta - a'_\theta) e^{t\nu \cdot \log \hat{\mu}_\theta + \log p_0(\theta)} \\ &\geq \quad \frac{1}{K} \sum_{\theta \in \Theta_{a \succ a'}} e^{t\nu \cdot \log \hat{\mu}_\theta + \log p_0(\theta)} - K \sum_{\theta \in \Theta_{a' \succ a}} e^{t\nu \cdot \log \hat{\mu}_\theta + \log p_0(\theta)} \\ &\geq \quad \frac{1}{K} \max_{\theta \in \Theta_{a \succ a'}} e^{t\nu \cdot \log \hat{\mu}_\theta + \log p_0(\theta)} - |\Theta| K e^{t(\max_{\theta \in \Theta_{a \succ a'}} \nu \cdot \log \hat{\mu}_\theta - \varepsilon) + \max_\theta \log p_0(\theta)} \\ &\geq \quad e^{t \max_{\theta \in \Theta_{a \succ a'}} \nu \cdot \log \hat{\mu}_\theta + \max_\theta \log p_0(\theta)} \left(\frac{1}{K} \min_{\theta, \theta'} \frac{p_0(\theta)}{p_0(\theta')} - |\Theta| K e^{-t\varepsilon} \right) > 0. \end{aligned}$$

Likewise, to show the final inclusion, note that for any $\nu \notin C_{aa'}^{-\varepsilon}$ and $t \geq \bar{t}$,

$$\begin{aligned}
\textstyle\sum_{\theta\in\Theta} & \quad (a_\theta - a'_\theta)e^{t\nu\cdot\log\hat{\mu}_\theta+\log p_0(\theta)} \\
\leq & \quad K \sum_{\theta\in\Theta_{a\succ a'}} e^{t\nu\cdot\log\hat{\mu}_\theta+p_0(\theta)} - \frac{1}{K}\sum_{\theta\in\Theta_{a'\succ a}} e^{t\nu\cdot\log\hat{\mu}_\theta+\log p_0(\theta)} \\
\leq & \quad K|\Theta|e^{t\max_{\theta\in\Theta_{a\succ a'}}\nu\cdot\log\hat{\mu}_\theta+\max_\theta\log p_0(\theta)} - \frac{1}{K}\max_{\theta\in\Theta_{a'\succ a}} e^{t\nu\cdot\log\hat{\mu}_\theta+\log p_0(\theta)} \\
\leq & \quad K|\Theta|e^{t(\max_{\theta\in\Theta_{a'\succ a}}\nu\cdot\log\hat{\mu}_\theta-\varepsilon)+\max_\theta\log p_0(\theta)} - \frac{1}{K}e^{t\max_{\theta\in\Theta_{a'\succ a}}\nu\cdot\log\hat{\mu}_\theta+\min_\theta\log p_0(\theta)} \\
= & \quad e^{t\max_{\theta\in\Theta_{a'\succ a}}\nu\cdot\log\hat{\mu}_\theta+\max_\theta\log p_0(\theta)}\left(K|\Theta|e^{-t\varepsilon} - \frac{1}{K}\min_{\theta,\theta'}\frac{p_0(\theta)}{p_0(\theta')}\right) < 0.
\end{aligned}$$

□

Lemma 2. *Suppose that neither $a \geq a'$ nor $a' \geq a$ and that Assumption 2 is satisfied. Then for each $\theta^* \in \Theta$, $\min_{\nu\in C_{aa'}^{\varepsilon}} \mathrm{KL}(\nu, \mu_{\theta^*})$ is well-defined and continuous in ε in a neighborhood of $\varepsilon = 0$.*

Proof. Since $a \not\geq a'$ and $a' \not\geq a$, $\Theta_{a\succ a'}, \Theta_{a'\succ a}$ are both nonempty. By Assumption 2, $C_{aa'}^{\varepsilon}$ is nonempty for ε close to 0. Since $C_{aa'}^{\varepsilon}$ is compact and $\mathrm{KL}(\cdot, \mu_{\theta^*})$ is continuous (as μ_{θ^*} has full support), $\min_{\nu\in C_{aa'}^{\varepsilon}} \mathrm{KL}(\nu, \mu_{\theta^*})$ is well-defined for all ε in a neighborhood of 0.

We now verify that $C_{aa'}^{\varepsilon}$ is a continuous correspondence at $\varepsilon = 0$, which ensures that $\min_{\nu\in C_{aa'}^{\varepsilon}} \mathrm{KL}(\nu, \mu_{\theta^*})$ is continuous in ε at $\varepsilon = 0$ by Berge's theorem of the maximum. It is straightforward to prove that $C_{aa'}^{\varepsilon}$ is upper-hemicontinuous in ε.

To see that $C_{aa'}^{\varepsilon}$ is also lower-hemicontinuous at $\varepsilon = 0$, take any $\nu \in C_{aa'}^{0}$ and any sequence $\varepsilon_n \to 0$. If $\max_{\theta\in\Theta_{a\succ a'}} \nu\cdot\log\hat{\mu}_\theta > \max_{\theta\in\Theta_{a'\succ a}} \nu\cdot\log\hat{\mu}_\theta$, then $\nu \in C_{aa'}^{\varepsilon_n}$ for n sufficiently large. Thus, assume $\max_{\theta\in\Theta_{a\succ a'}} \nu\cdot\log\hat{\mu}_\theta = \max_{\theta\in\Theta_{a'\succ a}} \nu\cdot\log\hat{\mu}_\theta$. Pick any $\theta' \in \mathrm{argmax}_{\Theta_{a\succ a'}} \nu\cdot\log\hat{\mu}_\theta$. Then $\nu\cdot\log\hat{\mu}_{\theta'} = \max_{\theta\in\Theta_{a'\succ a}} \nu\cdot\log\hat{\mu}_\theta$. Moreover, by Assumption 2, there exists $\nu' \in \Delta(X)$ such that

$$\nu'\cdot\log\hat{\mu}_{\theta'} > \max_{\theta\in\Theta_{a'\succ a}} \nu'\cdot\log\hat{\mu}_\theta.$$

Thus, for n sufficiently large, there exists $\kappa_n \in [0,1]$ such that $\lim_{n\to\infty}\kappa_n \to 1$ and

$$(\kappa_n\nu + (1-\kappa_n)\nu')\cdot\log\hat{\mu}_{\theta'} = \max_{\theta\in\Theta_{a'\succ a}} (\kappa_n\nu + (1-\kappa_n)\nu')\cdot\log\hat{\mu}_\theta + |\varepsilon_n|.$$

Then $(\kappa_n \nu + (1-\kappa_n)\nu') \in C^{\varepsilon_n}_{aa'}$ and $\lim_{n\to\infty}(\kappa_n \nu + (1-\kappa_n)\nu') = \nu$. This establishes lower hemicontinuity. $\square$

Finally, we use Sanov's theorem to characterize the rate (possibly zero) at which the probability of choosing a over a' vanishes as $t \to \infty$:

Lemma 3. *Suppose that neither $a \geq a'$ nor $a' \geq a$ and that Assumption 2 is satisfied. For any state θ,*

$$
\begin{aligned}
&- \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta \left[\{p_t \cdot (a-a') > 0\}\right] = - \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta \left[\{p_t \cdot (a-a') \geq 0\}\right] \\
&= \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \quad \textit{subject to} \quad \min_{\theta' \in \Theta_{a \succ a'}} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) \leq \min_{\theta' \in \Theta_{a' \succ a}} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}).
\end{aligned}
$$

Proof. For any $\varepsilon > 0$, Lemma 1 and (16) implies that

$$
\begin{aligned}
\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta[\{\nu_t \in C^{\varepsilon}_{aa'}\}] &\leq \liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta[\{p_t \cdot (a-a') > 0\}] \\
&\leq \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta[\{p_t \cdot (a-a') \geq 0\}] \\
&\leq \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta[\{\nu_t \in C^{-\varepsilon}_{aa'}\}].
\end{aligned}
$$

Pick any $\varepsilon > 0$ small enough that $C^{\varepsilon}_{aa'}$ is nonempty, which exists by Lemma 8. Sanov's theorem (e.g., Dembo and Zeitouni, 2010) yields the following inequalities:

$$
\begin{aligned}
\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta[\{\nu_t \in C^{\varepsilon}_{aa'}\}] &\geq - \inf_{\nu \in \mathrm{int}\left(C^{\varepsilon}_{aa'}\right)} \mathrm{KL}(\nu, \mu_\theta) = - \inf_{\nu \in C^{\varepsilon}_{aa'}} \mathrm{KL}(\nu, \mu_\theta), \\
\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta[\{\nu_t \in C^{-\varepsilon}_{aa'}\}] &\leq - \inf_{\nu \in C^{-\varepsilon}_{aa'}} \mathrm{KL}(\nu, \mu_\theta).
\end{aligned}
$$

In the first line, int denotes the interior of a set and the equality follows from the continuity of $\mathrm{KL}(\nu, \mu_\theta)$ in ν (as μ_θ has full support). Thus, taking the limit as $\varepsilon \to 0$ and applying Lemma 8, we obtain:

$$
\begin{aligned}
&- \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta[\{p_t \cdot (a-a') > 0\}] = - \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta[\{p_t \cdot (a-a') \geq 0\}] = \min_{\nu \in C^{0}_{aa'}} \mathrm{KL}(\nu, \mu_\theta) \\
&= \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \quad \text{subject to} \quad \min_{\theta' \in \Theta_{a \succ a'}} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) \leq \min_{\theta' \in \Theta_{a' \succ a}} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}),
\end{aligned}
$$

as claimed. $\square$

A.2.2 General decision problems

Continue to fix a perceived signal structure $\hat{\mu}$. We turn to general decision problems. Building on Lemma 3, the following lemma uses our learning efficiency index to characterize the rate at which the probability of choosing an ex-post suboptimal act vanishes as $t \to \infty$ under Assumption 3.

Lemma 4. *Let A be a regular and non-trivial decision problem and suppose that Assumption 2 and S_A-consistency (Assumption 3) are satisfied. Fix $\theta^* \in \Theta$ and let $a^* \in \operatorname{argmax}_{a \in A} a_{\theta^*}$ denote the ex-post optimal act at θ^*. Then*

$$\begin{aligned} -w(\theta^*, \mu, \hat{\mu}, S_A) &= \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\theta^*}\left[\left\{\max_{a \in A\setminus\{a^*\}} p_t \cdot (a - a^*) > 0\right\}\right] \\ &= \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\theta^*}\left[\left\{\max_{a \in A\setminus\{a^*\}} p_t \cdot (a - a^*) \geq 0\right\}\right]. \end{aligned}$$

Proof. Since a^* is uniquely optimal at θ^* (by regularity of A), we have $a \not\geq a^*$ for all $a \in A \setminus \{a^*\}$. Consider any $a \in A \setminus \{a^*\}$ with $a^* \not\geq a$, which exists as A is non-trivial. Then, by Lemma 3,

$$\begin{aligned} &-\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\theta^*}\left[\{p_t \cdot (a - a^*) > 0\}\right] = -\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\theta^*}\left[\{p_t \cdot (a - a^*) \geq 0\}\right] \\ &= \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_{\theta^*}) \text{ s.t. } \min_{\theta' \in \Theta_{a \succ a^*}} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) \leq \min_{\theta' \in \Theta_{a^* \succ a}} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) \\ &= \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_{\theta^*}) \text{ s.t. } \min_{\theta' \in \Theta_{a \succ a^*}} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) = \min_{\theta' \in \Theta} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}), \end{aligned}$$

where the final equality holds as $\Theta_{a \succ a^*}$ and $\Theta_{a^* \succ a}$ partition Θ by regularity of A. At the same time, for any $a \neq a^*$ with $a^* \geq a$, $\mathbb{P}_{\theta^*}[p_t \cdot (a - a^*) \geq 0] = 0$ for all t, as p_t always has full-support.

Since $\Theta \setminus S_A(\theta^*) = \bigcup_{a \neq a^*} \Theta_{a \succ a^*}$, we therefore have

$$\begin{aligned} &-\max_{a \in A\setminus\{a^*\}} \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\theta^*}\left[\{p_t \cdot (a - a^*) > 0\}\right] = -\max_{a \in A\setminus\{a^*\}} \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\theta^*}\left[\{p_t \cdot (a - a^*) \geq 0\}\right] \\ &= \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_{\theta^*}) \text{ s.t. } \min_{\theta' \notin S_A(\theta^*)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) = \min_{\theta' \in \Theta} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) \\ &= \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_{\theta^*}) \text{ s.t. } \min_{\theta' \notin S_A(\theta^*)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) \leq \min_{\theta' \in S_A(\theta^*)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}). \end{aligned}$$

We claim that the last line is equal to $w(\theta^*, \mu, \hat{\mu}, S_A)$. Consider any solution ν to the minimization problem. We show that $\min_{\theta' \notin S_A(\theta^*)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) = \min_{\theta' \in S_A(\theta^*)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'})$, i.e., the constraint holds with equality. Indeed, note that by S_A-consistency, $\nu \neq \mu_{\theta^*}$. Thus, if $\min_{\theta' \notin S_A(\theta^*)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) < \min_{\theta' \in S_A(\theta^*)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'})$, then by strict convexity of $\mathrm{KL}(\cdot, \mu_{\theta^*})$, we can choose $\varepsilon > 0$ sufficiently small such that $\nu' = (1-\varepsilon)\nu + \varepsilon \mu_{\theta^*}$ still satisfies the constraint but achieves a strictly smaller objective $\mathrm{KL}(\nu', \mu_{\theta^*})$, a contradiction.

Hence, exchanging lim and max, we have

$$-w(\theta^*, \mu, \hat{\mu}, S_A) = \lim_{t\to\infty} \frac{1}{t} \log \max_{a \in A\setminus\{a^*\}} \mathbb{P}_{\theta^*}\left[\{p_t \cdot (a - a^*) > 0\}\right], \tag{17}$$

$$-w(\theta^*, \mu, \hat{\mu}, S_A) = \lim_{t\to\infty} \frac{1}{t} \log \max_{a \in A\setminus\{a^*\}} \mathbb{P}_{\theta^*}\left[\{p_t \cdot (a - a^*) \geq 0\}\right]. \tag{18}$$

Finally, for each t,

$$\begin{aligned}
\max_{a \in A\setminus\{a^*\}} \mathbb{P}_{\theta^*}\left[\{p_t \cdot (a - a^*) > 0\}\right] &\leq \mathbb{P}_{\theta^*}\left[\left\{\max_{a \in A\setminus\{a^*\}} p_t \cdot (a - a^*) > 0\right\}\right] \\
&\leq \mathbb{P}_{\theta^*}\left[\left\{\max_{a \in A\setminus\{a^*\}} p_t \cdot (a - a^*) \geq 0\right\}\right] \\
&\leq (|A| - 1) \max_{a \in A\setminus\{a^*\}} \mathbb{P}_{\theta^*}\left[\{p_t \cdot (a - a^*) \geq 0\}\right].
\end{aligned}$$

Combined with (17) and (18), we obtain the desired conclusion. □

A.2.3 Proof of Theorem 3

Fix any partition S of Θ and $\mu, \hat{\mu}^1, \hat{\mu}^2$ satisfying Assumptions 2–3. Suppose $w(\mu, \hat{\mu}^1, S) > w(\mu, \hat{\mu}^2, S)$ and consider any non-trivial and regular decision problem A with $S_A = S$. Let $p^i_T \in \Delta(\Theta)$ denote the (random) posterior of agent i following T signal draws. Slightly generalizing the main text, we allow for the possibility that (4) features multiple subjectively optimal acts at some posteriors. In this case, agents employ an arbitrary strict total order over A to break ties, where tie-breaking rules can differ across agents.

Since A is regular, for each state θ, there is a unique ex-post optimal act $a^\theta \in \operatorname{argmax}_{a \in A} a_\theta$. Since A and Θ are finite, there exists a constant $K > 0$ such that for all T and $i = 1, 2$,

$$
\begin{aligned}
&-K\sum_{\theta} p_0(\theta)\mathbb{P}_\theta\left[\left\{\max_{a\in A\setminus\{a^\theta\}} p_T^i\cdot(a-a^\theta)\geq 0\right\}\right]\\
&\leq W_T(\mu,\hat{\mu}^i,A)-\sum_{\theta} p_0(\theta)a_\theta^\theta\\
&\leq -\frac{1}{K}\sum_{\theta} p_0(\theta)\mathbb{P}_\theta\left[\left\{\max_{a\in A\setminus\{a^\theta\}} p_T^i\cdot(a-a^\theta)> 0\right\}\right].
\end{aligned}
$$

Note the inequalities hold regardless of how agents break ties in case of indifference.

Thus, to show that there exists T^* such that $W_T(\mu,\hat{\mu}^1,A) > W_T(\mu,\hat{\mu}^2,A)$ for all $T \geq T^*$, it suffices to show that for all sufficiently large T,

$$
\frac{\sum_\theta p_0(\theta)\mathbb{P}_\theta\left[\left\{\max_{a\in A\setminus\{a^\theta\}} p_T^2\cdot(a-a^\theta)>0\right\}\right]}{\sum_\theta p_0(\theta)\mathbb{P}_\theta\left[\left\{\max_{a\in A\setminus\{a^\theta\}} p_T^1\cdot(a-a^\theta)\geq 0\right\}\right]} > K^2. \tag{19}
$$

By Lemma 4,

$$
\begin{aligned}
\frac{\sum_\theta p_0(\theta)\mathbb{P}_\theta\left[\left\{\max_{a\in A\setminus\{a^\theta\}} p_T^2\cdot(a-a^\theta)>0\right\}\right]}{\sum_\theta p_0(\theta)\mathbb{P}_\theta\left[\left\{\max_{a\in A\setminus\{a^\theta\}} p_T^1\cdot(a-a^\theta)\geq 0\right\}\right]} &= \frac{\sum_\theta p_0(\theta)e^{-Tw(\theta,\mu,\hat{\mu}^2,S_A)+o(T)}}{\sum_\theta p_0(\theta)e^{-Tw(\theta,\mu,\hat{\mu}^1,S_A)+o(T)}}\\
&\geq \frac{p_0(\theta)e^{-Tw(\theta,\mu,\hat{\mu}^2,S_A)+o(T)}}{e^{-T\min_\theta w(\theta,\mu,\hat{\mu}^1,S_A)+o(T)}},\ \forall\theta.
\end{aligned}
$$

Since $\min_\theta w(\theta,\mu,\hat{\mu}^1,S_A) > \min_\theta w(\theta,\mu,\hat{\mu}^2,S_A)$, this implies (19) for all T sufficiently large. □

A.3 Proof of Proposition 2

Consider any T. Condition (11) implies that either

$$
\prod_{t=1}^T \frac{\mu_{\overline{\theta}}(x_t)}{\mu_{\underline{\theta}}(x_t)} \geq \prod_{t=1}^T \frac{\hat{\mu}^1_{\overline{\theta}}(x_t)}{\hat{\mu}^1_{\underline{\theta}}(x_t)} \geq \prod_{t=1}^T \frac{\hat{\mu}^2_{\overline{\theta}}(x_t)}{\hat{\mu}^2_{\underline{\theta}}(x_t)} \quad \forall x^T\in X^T,
$$

or

$$
\prod_{t=1}^T \frac{\mu_{\overline{\theta}}(x_t)}{\mu_{\underline{\theta}}(x_t)} \leq \prod_{t=1}^T \frac{\hat{\mu}^1_{\overline{\theta}}(x_t)}{\hat{\mu}^1_{\underline{\theta}}(x_t)} \leq \prod_{t=1}^T \frac{\hat{\mu}^2_{\overline{\theta}}(x_t)}{\hat{\mu}^2_{\underline{\theta}}(x_t)} \quad \forall x^T\in X^T.
$$

Thus, the same argument as in Proposition 1 yields that $W_T(\mu,\hat{\mu}^1,A) \geq W_T(\mu,\hat{\mu}^2,A)$ for any decision problem A (where we can again allow for ties as long as both agents use the same tie-breaking rule). $\square$

A.4 Proof of Proposition 3

We first establish the following lemma, which will be used to prove the "only if" direction of Proposition 3:

Lemma 5. *Fix any μ. There exists a unique $\nu^* \in \mathcal{I}(\mu)$ such that $w(\mu,\mu) = \mathrm{KL}(\nu^*, \mu_{\underline{\theta}}) = \mathrm{KL}(\nu^*, \mu_{\overline{\theta}})$. Moreover, $\nu^*(x) > 0$ for all $x \in X$ and there exist $\lambda_{\underline{\theta}}, \lambda_{\overline{\theta}} > 0$ and $C_{\underline{\theta}}, C_{\overline{\theta}} \in \mathbb{R}$ such that*

$$\nabla_\nu \mathrm{KL}(\nu^*, \mu_{\underline{\theta}}) = \left(1 + C_{\underline{\theta}} + \lambda_{\underline{\theta}} \log \frac{\mu_{\underline{\theta}}(x)}{\mu_{\overline{\theta}}(x)}\right)_{x \in X},$$
$$\nabla_\nu \mathrm{KL}(\nu^*, \mu_{\overline{\theta}}) = \left(1 + C_{\overline{\theta}} + \lambda_{\overline{\theta}} \log \frac{\mu_{\overline{\theta}}(x)}{\mu_{\underline{\theta}}(x)}\right)_{x \in X}.$$

Proof. By definition, $\nu \in \mathcal{I}(\mu)$ if and only if $\mathrm{KL}(\nu, \mu_{\underline{\theta}}) = \mathrm{KL}(\nu, \mu_{\overline{\theta}})$. Thus, $w(\mu,\mu) = \mathrm{KL}(\nu^*, \mu_{\underline{\theta}}) = \mathrm{KL}(\nu^*, \mu_{\overline{\theta}})$ if and only if $\nu^* \in \mathrm{argmin}_{\nu \in \mathcal{I}(\mu)} \mathrm{KL}(\nu, \mu_{\underline{\theta}}) = \mathrm{argmin}_{\nu \in \mathcal{I}(\mu)} \mathrm{KL}(\nu, \mu_{\overline{\theta}})$. Moreover, $\mathrm{argmin}_{\nu \in \mathcal{I}(\mu)} \mathrm{KL}(\nu, \mu_{\underline{\theta}}) = \{\nu^*\}$ is a singleton, since $\mathcal{I}(\mu)$ is nonempty, compact and convex and $\mathrm{KL}(\cdot, \mu_{\underline{\theta}})$ is continuous and convex.

To see that $\nu^*(x) > 0$ for all $x \in X$, suppose instead that $\nu^*(\overline{x}) = 0$ for some $\overline{x} \in X$. Consider any other $\hat{\nu} \in \mathcal{I}(\mu)$ with $\hat{\nu}(\overline{x}) > 0$.[27] For each $\varepsilon \in [0,1]$, consider $\nu_\varepsilon := (1-\varepsilon)\nu^* + \varepsilon\hat{\nu} \in \mathcal{I}(\mu)$. Then for each θ,

$$\lim_{\varepsilon \searrow 0} \frac{\mathrm{KL}(\nu_\varepsilon, \mu_\theta) - \mathrm{KL}(\nu^*, \mu_\theta)}{\varepsilon} = -\infty,$$

contradicting the fact that $\nu^* \in \mathrm{argmin}_{\nu \in \mathcal{I}(\mu)} \mathrm{KL}(\nu, \mu_\theta)$.

For the final part, consider state $\underline{\theta}$; the argument for state $\overline{\theta}$ is analogous. Since $\nu^*(x) > 0$ for all $x \in X$, $\mathrm{KL}(\nu, \mu_{\underline{\theta}})$ is differentiable in ν at ν^*. As we saw in the proof

[27] Such a $\hat{\nu}$ exists: If $\mu_{\overline{\theta}}(\overline{x}) = \mu_{\underline{\theta}}(\overline{x})$, set $\hat{\nu}(\overline{x}) = 1$. If $\mu_{\overline{\theta}}(\overline{x}) > \mu_{\underline{\theta}}(\overline{x})$, set $\hat{\nu}(\overline{x}) + \hat{\nu}(\underline{x}) = 1$ and $\frac{\hat{\nu}(\overline{x})}{\hat{\nu}(\underline{x})} = \frac{\log \frac{\mu_{\underline{\theta}}(\underline{x})}{\mu_{\overline{\theta}}(\underline{x})}}{\log \frac{\mu_{\overline{\theta}}(\overline{x})}{\mu_{\underline{\theta}}(\overline{x})}}$ for some $\underline{x}$ such that $\mu_{\underline{\theta}}(\underline{x}) > \mu_{\overline{\theta}}(\underline{x})$. The case $\mu_{\overline{\theta}}(\overline{x}) < \mu_{\underline{\theta}}(\overline{x})$ is analogous.

of Lemma 4, $\nu^* \in \Delta(X)$ also solves the following relaxed problem:

$$\min_{\nu \in \Delta(X)} \text{KL}(\nu, \mu_{\underline{\theta}}) \text{ s.t. } \text{KL}(\nu, \mu_{\overline{\theta}}) \leq \text{KL}(\nu, \mu_{\underline{\theta}})$$
$$= \min_{\nu \in \Delta(X)} \text{KL}(\nu, \mu_{\underline{\theta}}) \text{ s.t. } \sum_{x \in X} \nu(x) \log \frac{\mu_{\underline{\theta}}(x)}{\mu_{\overline{\theta}}(x)} \leq 0.$$

Thus, we have the following first-order conditions at ν^*: There exist Lagrange multipliers $\lambda_{\underline{\theta}} \geq 0$ and $C_\theta \in \mathbb{R}$ such that for all x,

$$\log \frac{\nu^*(x)}{\mu_{\underline{\theta}}(x)} = \lambda_{\underline{\theta}} \log \frac{\mu_{\underline{\theta}}(x)}{\mu_{\overline{\theta}}(x)} + C_{\underline{\theta}}. \tag{20}$$

If $\lambda_{\underline{\theta}} = 0$, then (20) implies $\nu^* = \mu_{\underline{\theta}}$, which violates the constraint $\sum_{x \in X} \nu(x) \log \frac{\mu_{\underline{\theta}}(x)}{\mu_{\overline{\theta}}(x)} \leq 0$, since $\mu_{\underline{\theta}} \neq \mu_{\overline{\theta}}$. Thus, $\lambda_{\underline{\theta}} > 0$. Since $\nabla_\nu \text{KL}(\nu^*, \mu') = \left(1 + \log \frac{\nu^*(x)}{\mu'(x)}\right)_{x \in X}$ for any $\mu' \in \text{int}\Delta(X)$, (20) becomes

$$\nabla_\nu \text{KL}(\nu^*, \mu_{\underline{\theta}}) = \left(1 + C_{\underline{\theta}} + \lambda_{\underline{\theta}} \log \frac{\mu_{\underline{\theta}}(x)}{\mu_{\overline{\theta}}(x)}\right)_{x \in X},$$

as claimed. □

Proof of Proposition 3: For the "if" direction, suppose $\hat{\mu} \in \text{PE}(\mu)$. Then, as noted in the main text, $\mathcal{I}(\hat{\mu}) = \mathcal{I}(\mu)$. Thus, $w(\mu, \mu) = w(\mu, \hat{\mu})$.

For the "only if" direction, suppose that $\hat{\mu} \notin \text{PE}(\mu)$. By Lemma 5, there exists a unique $\nu^* \in \mathcal{I}(\mu)$ such that $\nu^*(x) > 0$ for all $x \in X$ and

$$w(\mu, \mu) = \text{KL}(\nu^*, \mu_{\underline{\theta}}) = \text{KL}(\nu^*, \mu_{\overline{\theta}}).$$

First, suppose that $\nu^* \notin \mathcal{I}(\hat{\mu})$. Then either $\text{KL}(\nu^*, \hat{\mu}_{\underline{\theta}}) < \text{KL}(\nu^*, \hat{\mu}_{\overline{\theta}})$ or $\text{KL}(\nu^*, \hat{\mu}_{\overline{\theta}}) < \text{KL}(\nu^*, \hat{\mu}_{\underline{\theta}})$. We only consider the former case, as the latter case is analogous. As in the proof of Lemma 4, $w(\overline{\theta}, \mu, \hat{\mu})$ can equivalently be written as the value of the following relaxed problem:

$$w(\overline{\theta}, \mu, \hat{\mu}) = \min_{\nu' \in \Delta(X)} \text{KL}(\nu', \mu_{\overline{\theta}}) \text{ s.t. } \text{KL}(\nu', \hat{\mu}_{\underline{\theta}}) \leq \text{KL}(\nu', \hat{\mu}_{\overline{\theta}}). \tag{21}$$

Since $\mathrm{KL}(\cdot, \hat{\mu}_{\underline{\theta}})$ and $\mathrm{KL}(\cdot, \hat{\mu}_{\overline{\theta}})$ are continuous, for $\varepsilon \in (0,1)$ sufficiently small,

$$\mathrm{KL}((1-\varepsilon)\nu^* + \varepsilon\mu_{\overline{\theta}}, \hat{\mu}_{\underline{\theta}}) < \mathrm{KL}((1-\varepsilon)\nu^* + \varepsilon\mu_{\overline{\theta}}, \hat{\mu}_{\overline{\theta}}).$$

Thus, $(1-\varepsilon)\nu^* + \varepsilon\mu_{\overline{\theta}}$ satisfies the constraint in (21). But, by convexity of $\mathrm{KL}(\cdot, \mu_{\overline{\theta}})$ and since $w(\mu, \mu) > 0$ (because $\mu_{\overline{\theta}} \neq \mu_{\underline{\theta}}$)

$$\mathrm{KL}((1-\varepsilon)\nu^* + \varepsilon\mu_{\overline{\theta}}, \mu_{\overline{\theta}}) \leq (1-\varepsilon)\mathrm{KL}(\nu^*, \mu_{\overline{\theta}}) + \varepsilon\mathrm{KL}(\mu_{\overline{\theta}}, \mu_{\overline{\theta}}) < w(\mu, \mu).$$

Thus, $w(\mu, \hat{\mu}) \leq w(\overline{\theta}, \mu, \hat{\mu}) < w(\mu, \mu)$.

Next, suppose that $\nu^* \in \mathcal{I}(\hat{\mu})$. Since $\hat{\mu} \notin \mathrm{PE}(\mu)$ and

$$\mathcal{I}(\mu) = \left\{\nu \in \Delta(X) : \nu \cdot \log \frac{\mu_{\overline{\theta}}}{\mu_{\underline{\theta}}} = 0\right\}, \quad \mathcal{I}(\hat{\mu}) = \left\{\nu \in \Delta(X) : \nu \cdot \log \frac{\hat{\mu}_{\overline{\theta}}}{\hat{\mu}_{\underline{\theta}}} = 0\right\},$$

there exists $\nu \in \mathcal{I}(\hat{\mu}) \setminus \mathcal{I}(\mu)$. Then either $\nu \cdot \log \frac{\mu_{\underline{\theta}}}{\mu_{\overline{\theta}}} < 0$ or $\nu \cdot \log \frac{\mu_{\overline{\theta}}}{\mu_{\underline{\theta}}} < 0$. We only consider the former case, as the latter case is analogous. For each $\varepsilon \in (0,1)$, define

$$\nu_\varepsilon = (1-\varepsilon)\nu^* + \varepsilon\nu.$$

Since both $\nu^*, \nu \in \mathcal{I}(\hat{\mu})$, we have $\nu_\varepsilon \in \mathcal{I}(\hat{\mu})$. Moreover, by Lemma 5, there exists $\lambda_{\underline{\theta}} > 0$ such that

$$\lim_{\varepsilon \to 0} \frac{\mathrm{KL}(\nu_\varepsilon, \mu_{\underline{\theta}}) - \mathrm{KL}(\nu^*, \mu_{\underline{\theta}})}{\varepsilon \|\nu - \nu^*\|_2} = (\nu - \nu^*) \cdot \nabla_\nu \mathrm{KL}(\nu^*, \mu_{\underline{\theta}}) = \lambda_{\underline{\theta}} \nu \cdot \log \frac{\mu_{\underline{\theta}}}{\mu_{\overline{\theta}}} < 0.$$

Thus, for $\varepsilon > 0$ sufficiently small, we have $\nu_\varepsilon \in \mathcal{I}(\hat{\mu})$ and $\mathrm{KL}(\nu_\varepsilon, \mu_{\underline{\theta}}) < \mathrm{KL}(\nu^*, \mu_{\underline{\theta}}) = w(\mu, \mu)$. Hence, $w(\mu, \hat{\mu}) \leq w(\underline{\theta}, \mu, \hat{\mu}) < w(\mu, \mu)$. □

A.5 Proof of Theorem 2

Fix $\hat{\mu}^1$ and $\hat{\mu}^2$ satisfying Assumption 2. Take any non-trivial decision problem A. It is without loss to assume that A contains no dominated acts, since agents would never choose such acts given that their posteriors p_i^T have full support at each signal history. As in the proof of Theorem 3, we allow for the possibility that (4) features multiple subjectively optimal acts at some posteriors, in which case agents employ (possibly different) tie-breaking rules.

Let $\overline{a}$ (resp. $\underline{a}$) denote the unique act in A that is ex-post optimal at $\overline{\theta}$ (resp. $\underline{\theta}$). Since A does not contain dominated acts, we have

$$[\overline{a}_{\underline{\theta}} < a_{\underline{\theta}}, \forall a \in A \setminus \{\overline{a}\}] \quad \text{and} \quad [\underline{a}_{\overline{\theta}} < a_{\overline{\theta}}, \forall a \in A \setminus \{\underline{a}\}]. \tag{22}$$

First part: By (22), there exists a constant $K > 0$ such that for all T and $i = 1, 2$,

$$\begin{aligned}
& -Kp_0(\overline{\theta})\mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a-\overline{a}) \geq 0\right\}\right] + \frac{1}{K}p_0(\underline{\theta})\mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a-\overline{a}) > 0\right\}\right] \\
\leq \;& W_T(\mu, \hat{\mu}^i, A) - p_0(\overline{\theta})\overline{a}_{\overline{\theta}} - p_0(\underline{\theta})\overline{a}_{\underline{\theta}} \\
\leq \;& -\frac{1}{K}p_0(\overline{\theta})\mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a-\overline{a}) > 0\right\}\right] + Kp_0(\underline{\theta})\mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a-\overline{a}) \geq 0\right\}\right]
\end{aligned}$$

Thus, to show that there exists T^* such that $W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A)$ for all $T \geq T^*$, it suffices to show that for all sufficiently large T,

$$\frac{p_0(\overline{\theta})\mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^1 \cdot (a-\overline{a}) \geq 0\right\}\right] + p_0(\underline{\theta})\mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^2 \cdot (a-\overline{a}) \geq 0\right\}\right]}{p_0(\overline{\theta})\mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^2 \cdot (a-\overline{a}) > 0\right\}\right] + p_0(\underline{\theta})\mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^1 \cdot (a-\overline{a}) > 0\right\}\right]} < \frac{1}{K^2}. \tag{23}$$

Since both agents satisfy case (i), the same argument as in Lemma 4 implies that for all θ and $i = 1, 2$,

$$\begin{aligned}
-w(\theta, \mu, \hat{\mu}^i) &= \lim_{T\to\infty} \frac{1}{T} \log \mathbb{P}_\theta\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a-\overline{a}) > 0\right\}\right] \\
&= \lim_{T\to\infty} \frac{1}{T} \log \mathbb{P}_\theta\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a-\overline{a}) \geq 0\right\}\right].
\end{aligned}$$

Hence, the fact that $\min\{w(\underline{\theta}, \mu, \hat{\mu}^1), w(\overline{\theta}, \mu, \hat{\mu}^2)\} < \min\{w(\underline{\theta}, \mu, \hat{\mu}^2), w(\overline{\theta}, \mu, \hat{\mu}^1)\}$ implies (23) for all sufficiently large T.

Second part: The proof is analogous to the first part.

Third part: By (22), there exists a constant $K > 0$ such that for all T and

$i = 1, 2,$

$$\begin{aligned}
&\frac{1}{K} p_0(\overline{\theta})\mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\underline{a}\}} p_T^i \cdot (a - \underline{a}) > 0\right\}\right] + \frac{1}{K} p_0(\underline{\theta})\mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a - \overline{a}) > 0\right\}\right] \\
\leq\ & W_T^i(\mu, \hat{\mu}^i, A) - p_0(\overline{\theta})\underline{a}_{\overline{\theta}} - p_0(\underline{\theta})\overline{a}_{\underline{\theta}} \\
\leq\ & K p_0(\overline{\theta})\mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\underline{a}\}} p_T^i \cdot (a - \underline{a}) \geq 0\right\}\right] + K p_0(\underline{\theta})\mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a - \overline{a}) \geq 0\right\}\right].
\end{aligned}$$

Thus, to show that there exists T^* such that $W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A)$ for all $T \geq T^*$, it suffices to show that for all sufficiently large T,

$$\frac{p_0(\overline{\theta})\mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\underline{a}\}} p_T^2 \cdot (a - \underline{a}) \geq 0\right\}\right] + p_0(\underline{\theta})\mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^2 \cdot (a - \overline{a}) \geq 0\right\}\right]}{p_0(\overline{\theta})\mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\underline{a}\}} p_T^1 \cdot (a - \underline{a}) > 0\right\}\right] + p_0(\underline{\theta})\mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^1 \cdot (a - \overline{a}) > 0\right\}\right]} < \frac{1}{K^2}. \tag{24}$$

Since both agents satisfy case (iii), the same argument as in Lemma 4 implies that for each $i = 1, 2$,

$$\begin{aligned}
-w(\overline{\theta}, \mu, \hat{\mu}^i) &= \lim_{T\to\infty} \frac{1}{T} \log \mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\underline{a}\}} p_T^i \cdot (a - \underline{a}) > 0\right\}\right] \\
&= \lim_{T\to\infty} \frac{1}{T} \log \mathbb{P}_{\overline{\theta}}\left[\left\{\max_{a\in A\setminus\{\underline{a}\}} p_T^i \cdot (a - \underline{a}) \geq 0\right\}\right] \\
-w(\underline{\theta}, \mu, \hat{\mu}^i) &= \lim_{T\to\infty} \frac{1}{T} \log \mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a - \overline{a}) > 0\right\}\right] \\
&= \lim_{T\to\infty} \frac{1}{T} \log \mathbb{P}_{\underline{\theta}}\left[\left\{\max_{a\in A\setminus\{\overline{a}\}} p_T^i \cdot (a - \overline{a}) \geq 0\right\}\right].
\end{aligned}$$

Hence, the fact that $\min_\theta w(\mu, \hat{\mu}^1) < \min_\theta w(\mu, \hat{\mu}^2)$ implies (24) for all sufficiently large T. □

A.6 Proof of Theorem 4

The proof follows the same arguments as in Theorem 3 and is thus omitted. In particular, analogous to Lemma 4, one can show that, in any non-trivial decision problem, $w(\theta, \mu, \hat{M}^i)$ corresponds to the exponential rate at which the probability that agent i chooses an ex-post suboptimal act in state θ vanishes as $T \to \infty$.

A.7 Proof of Proposition 4

Observe that

$$
\begin{aligned}
w(\underline{\theta}, \mu, \hat{M}) &= \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_{\underline{\theta}}) \quad \text{subject to} \quad \min_{\hat{\mu} \in \hat{M}} \mathrm{KL}(\nu, \hat{\mu}_{\underline{\theta}}) \geq \min_{\hat{\mu} \in \hat{M}} \mathrm{KL}(\nu, \hat{\mu}_{\overline{\theta}}) \\
&\geq \min_{\hat{\mu} \in \hat{M}} \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_{\underline{\theta}}) \quad \text{subject to} \quad \min_{\hat{\mu}' \in \hat{M}} \mathrm{KL}(\nu, \hat{\mu}'_{\underline{\theta}}) \geq \mathrm{KL}(\nu, \hat{\mu}_{\overline{\theta}}) \\
&\geq \min_{\hat{\mu} \in \hat{M}} \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_{\underline{\theta}}) \quad \text{subject to} \quad \mathrm{KL}(\nu, \hat{\mu}_{\underline{\theta}}) \geq \mathrm{KL}(\nu, \hat{\mu}_{\overline{\theta}}) \\
&= \min_{\hat{\mu} \in \hat{M}} w(\underline{\theta}, \mu, \hat{\mu}).
\end{aligned}
$$

Here, the first and the last equalities follow from the definition of the efficiency index and Assumption 4 by analogous arguments as in the proof of Lemma 4. A similar argument shows $w(\overline{\theta}, \mu, \hat{M}) \geq \min_{\hat{\mu} \in \hat{M}} w(\overline{\theta}, \mu, \hat{\mu})$. Thus, $w(\mu, \hat{M}) \geq \min_{\hat{\mu} \in \hat{M}} w(\mu, \hat{\mu})$. □

References

Acemoglu, D., V. Chernozhukov, and M. Yildiz (2016): "Fragility of asymptotic agreement under Bayesian learning," *Theoretical Economics*, 11(1), 187–225.

Azrieli, Y. (2014): "Comment on "the law of large demand for information"," *Econometrica*, 82(1), 415–423.

Benjamin, D. J. (2019): "Errors in probabilistic reasoning and judgment biases," *Handbook of Behavioral Economics-Foundations and Applications 2*, p. 69.

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

Blackwell, D. (1951): "Comparison of experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93–102. University of California Press.

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

Bohren, J. A. (2016): "Informational herding with model misspecification," *Journal of Economic Theory*, 163, 222–247.

Bohren, J. A., and D. Hauser (2018): "Social Learning with Model Misspecification: A Framework and a Robustness Result," working paper.

Bushong, B., and T. Gagnon-Bartsch (2019): "Learning with Misattribution of Reference Dependence," working paper.

Cho, I.-K., and K. Kasa (2015): "Learning and model validation," *The Review of Economic Studies*, 82(1), 45–82.

Coutts, A. (2019): "Good news and bad news are still news: Experimental evidence on belief updating," *Experimental Economics*, 22(2), 369–395.

Cover, T. M., and J. A. Thomas (2006): *Elements of Information Theory*. Wiley-Interscience.

Cripps, M. W. (2018): "Divisible updating," working paper.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998): "Investor psychology and security market under-and overreactions," *Journal of Finance*, 53(6), 1839–1885.

Dembo, A., and O. Zeitouni (2010): *Large Deviations Techniques and Applications*. Springer.

Dupuis, P., and R. S. Ellis (2011): *A weak convergence approach to the theory of large deviations*, vol. 902. John Wiley & Sons.

Eliaz, K., and R. Spiegler (2020): "A Model of Competing Narratives," *American Economic Review*, 110, 3786–3816.

Enke, B., and F. Zimmermann (2017): "Correlation Neglect in Belief Formation," mimeo.

Epstein, L. G., J. Noor, and A. Sandroni (2010): "Non-bayesian learning," *The BE Journal of Theoretical Economics*, 10(1).

Esponda, I., and D. Pouzo (2016): "Berk–Nash Equilibrium: A Framework for Modeling Agents With Misspecified Models," *Econometrica*, 84(3), 1093–1130.

——— (2020): "Equilibrium in misspecified Markov decision processes," *Theoretical Economics*, forthcoming.

Esponda, I., D. Pouzo, and Y. Yamamoto (2019): "Asymptotic Behavior of Bayesian Learners with Misspecified Models," working paper.

Eyster, E., and M. Rabin (2010): "Naive herding in rich-information settings," *American economic journal: microeconomics*, 2(4), 221–43.

Frick, M., R. Iijima, and Y. Ishii (2020a): "Misinterpreting Others and the Fragility of Social Learning," *Econometrica*, 88(6), 2281–2328.

——— (2020b): "Stability and Robustness in Misspecified Learning Models," Cowles Foundation Discussion Paper.

Fudenberg, D., and G. Lanzani (2021): "Which Misperceptions Persist?," working paper.

Fudenberg, D., G. Lanzani, and P. Strack (2020): "Limits Points of Endogenous Misspecified Learning," working paper.

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

Gagnon-Bartsch, T. (2017): "Taste Projection in Models of Social Learning," working paper.

Gagnon-Bartsch, T., M. Rabin, and J. Schwartzstein (2018): "Channeled Attention and Stable Errors," working paper.

Gossner, O., and J. Steiner (2018): "On the cost of misperception: General results and behavioral applications," *Journal of Economic Theory*, 177, 816–847.

Gul, F., and W. Pesendorfer (2001): "Temptation and self-control," *Econometrica*, 69(6), 1403–1435.

He, K. (2018): "Mislearning from Censored Data: The Gambler's Fallacy in Optimal-Stopping Problems," working paper.

He, K., and J. Libgober (2021): "Evolutionarily Stable (Mis) specifications: Theory and Applications," working paper.

Heidhues, P., B. Koszegi, and P. Strack (2018): "Unrealistic expectations and misguided learning," *Econometrica*, 86(4), 1159–1214.

——— (2020): "Convergence in models of misspecified learning," *Theoretical Economics*.

Levy, G., R. Razin, and A. Young (2019): "Misspecified Politics and the Recurrence of Populism," Discussion paper.

Massari, F. (2020): "Under-reaction: Irrational behavior or robust response to model misspecification?," working paper.

Mobius, M. M., M. Niederle, P. Niehaus, and T. Rosenblat (2014): "Managing Self-Confidence: Theory and Experimental Evidence," working paper.

Montiel Olea, J. L., P. Ortoleva, M. Pai, and A. Prat (2019): "Competing Models," working paper.

Moscarini, G., and L. Smith (2002): "The law of large demand for information," *Econometrica*, 70(6), 2351–2366.

Mu, X., L. Pomatto, P. Strack, and O. Tamuz (2021): "From Blackwell Dominance in Large Samples to Rényi Divergences and Back Again," *Econometrica*, 89(1), 475–506.

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

Ortoleva, P., and E. Snowberg (2015): "Overconfidence in political behavior," *American Economic Review*, 105(2), 504–35.

Phillips, L. D., and W. Edwards (1966): "Conservatism in a simple probability inference task.," *Journal of Experimental Psychology*, 72(3), 346.

Rabin, M., and D. Vayanos (2010): "The gambler's and hot-hand fallacies: Theory and applications," *The Review of Economic Studies*, 77(2), 730–778.

Sandroni, A. (2000): "Do markets favor agents able to make accurate predictions?," *Econometrica*, 68(6), 1303–1341.

Schwartzstein, J., and A. Sunderam (2021): "Using models to persuade," *American Economic Review*, 111(1), 276–323.

Spiegler, R. (2016): "Bayesian networks and boundedly rational expectations," *The Quarterly Journal of Economics*, 131(3), 1243–1290.

Steiner, J., and C. Stewart (2016): "Perceiving prospects properly," *American Economic Review*, 106(7), 1601–31.

Thaler, M. (2019): "The "fake news" effect: An experiment on motivated reasoning and trust in news," working paper.

Online Appendix to "Welfare Comparisons for Biased Learning"

Mira Frick, Ryota Iijima, and Yuhta Ishii

B Extension of Theorem 2 under General States

We extend Theorem 2 to general finite state spaces Θ. As in Section 3.4, we consider perceived signal structures $\hat{\mu}$ satisfying Assumption 2 (Non-degeneracy). We allow for arbitrary correct or incorrect asymptotic beliefs: We only require that in each state θ, asymptotic beliefs under $\hat{\mu}$ are a point-mass on some state θ', ruling out cycling.

Assumption 5 (No cycling)**.** For every $\theta \in \Theta$, there exists $\theta' \in \Theta$ with $\mathrm{KL}(\mu_\theta, \hat{\mu}_{\theta'}) < \mathrm{KL}(\mu_\theta, \hat{\mu}_{\theta''})$ for all $\theta'' \neq \theta$.

Given any partition S over Θ, Assumption 5 implies that in each state θ, asymptotic beliefs under $\hat{\mu}$ assign probability one to some (possibly incorrect) cell $S_{\hat{\mu}}(\theta) \in S$ given by

$$S_{\hat{\mu}}(\theta) = S(\theta'), \text{ where } \{\theta'\} = \operatorname*{argmin}_{\theta'' \in \Theta} \mathrm{KL}(\mu_\theta, \hat{\mu}_{\theta''}).$$

Theorem 5 below will compare dynamic welfare across any biases $\hat{\mu}^1$ and $\hat{\mu}^2$ under which asymptotic beliefs concentrate on the same cell $S_{\hat{\mu}^1}(\theta) = S_{\hat{\mu}^2}(\theta) =: \hat{S}(\theta)$ in each state θ; however, asymptotic beliefs within cells may differ across $\hat{\mu}^1$ and $\hat{\mu}^2$.

As in Section 4.1, we obtain a generically complete welfare ranking over such biases up to restricting to subclasses of decision problems. Given any regular decision problem A, let $\alpha(\theta) = \operatorname{argmax}_{a \in A} a_\theta$ denote the unique ex-post optimal act in state θ. Let S_A denote the partition over Θ whose cells are $S_A(\theta) = \{\theta' \in \Theta : \alpha(\theta) = \alpha(\theta')\}$. If asymptotic beliefs in state θ concentrate on cell $\hat{S}_A(\theta) \in S_A$, let $\hat{a}(\theta) := \alpha(\theta')$ for some $\theta' \in \hat{S}_A(\theta)$ denote the induced (asymptotically) subjectively optimal act. Then A induces a partition of each $\Theta \setminus \hat{S}_A(\theta)$ into

$$S_A^+(\theta) := \{\theta'' \in \Theta : \alpha_\theta(\theta'') > \hat{a}_\theta(\theta)\}, \quad S_A^-(\theta) := \{\theta'' \in \Theta : \alpha_\theta(\theta'') < \hat{a}_\theta(\theta)\},$$

where $S_A^+(\theta)$ (resp. $S_A^-(\theta)$) consists of all states θ'' whose ex-post optimal act $\alpha(\theta'')$ yields higher (resp. lower) payoffs in state θ than act $\hat{a}(\theta)$.

We define a learning efficiency index that allows one to robustly rank dynamic welfare in all decision problems A that induce the same collections $S^+ := (S_A^+(\theta))_{\theta \in \Theta}$ and $S^- := (S_A^-(\theta))_{\theta \in \Theta}$. Given any such collections S^+ and S^-, define

$$w^+(\mu, \hat{\mu}, S^+) := \min_{\theta \in \Theta} w(\theta, \mu, \hat{\mu}, S^+(\theta)), \quad w^-(\mu, \hat{\mu}, S^-) := \min_{\theta \in \Theta} w(\theta, \mu, \hat{\mu}, S^-(\theta)),$$

where for any $\Theta' \subseteq \Theta$, the index

$$w(\theta, \mu, \hat{\mu}, \Theta') := \min_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \quad \text{subject to} \quad \min_{\theta' \in \Theta'} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) = \min_{\theta' \notin \Theta'} \mathrm{KL}(\nu, \hat{\mu}_{\theta'})$$

measures how unlikely it is in state θ to encounter empirical signal distributions ν that make it impossible to distinguish states in Θ' vs. $\Theta \setminus \Theta'$. We set $w(\theta, \mu, \hat{\mu}, \emptyset) = \infty$.

Based on the learning efficiency indices $w^+(\mu, \hat{\mu}, S^+)$ and $w^-(\mu, \hat{\mu}, S^-)$, we obtain the following characterization of the dynamic welfare ranking. We strengthen non-triviality to the requirement that decision problems are ***essential***, in the sense that $|A| > 1$ and $A \subseteq \alpha(\Theta)$, i.e., each act in A is ex-post optimal in some state.

Theorem 5. *Fix any μ, $\hat{\mu}^1$, $\hat{\mu}^2$ satisfying Assumptions 2 and 5. Consider any partition S of Θ such that $S_{\hat{\mu}^1}(\theta) = S_{\hat{\mu}^2}(\theta) =: \hat{S}(\theta)$ for all θ, and any partition $\{S^+(\theta), S^-(\theta)\}$ of $\Theta \setminus \hat{S}(\theta)$ for each θ. Suppose that*

$$\min\left\{w^+(\mu, \hat{\mu}^2, S^+), w^-(\mu, \hat{\mu}^1, S^-)\right\} > \min\left\{w^+(\mu, \hat{\mu}^1, S^+), w^-(\mu, \hat{\mu}^2, S^-)\right\}. \qquad (25)$$

Then for any regular and essential decision problem A with $S_A(\theta) = S(\theta)$, $S_A^+(\theta) = S^+(\theta)$ and $S_A^-(\theta) = S^-(\theta)$ for all θ, there exists T^* such that for all $T \geq T^*$, $W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A)$.

Remark 3. Theorem 5 generalizes both Theorems 2 and 3 (up to strengthening non-triviality to essentiality). When $\Theta = \{\underline{\theta}, \overline{\theta}\}$, it is easy to verify that all three cases in Theorem 2 yield the ranking (25).

Theorem 3 assumes S-consistency (Assumption 3). In this case, $S_{\hat{\mu}^i}(\theta) = S(\theta)$ for all θ and $i = 1, 2$, so $S_A^+(\theta) = \emptyset$ and $S_A^-(\theta) = \Theta \setminus S(\theta)$ for all A with $S_A = S$. Thus, $w^+(\mu, \hat{\mu}^i, S^+) = \infty$ and $w^-(\mu, \hat{\mu}^i, S^-) = w(\mu, \hat{\mu}^i, S)$. Hence, (25) reduces to $w(\mu, \hat{\mu}^1, S) > w(\mu, \hat{\mu}^2, S)$. ▲

B.1 Proof of Theorem 5

B.1.1 Preliminary Lemmas

Fix any regular and essential decision problem A with induced partition $S_A =: S$. Fix any $\hat{\mu}$ satisfying Assumptions 2 and 5. Write $\hat{S}(\theta) := S_{\hat{\mu}}(\theta)$, $S^+(\theta) := S_A^+(\theta)$, $S^-(\theta) := S_A^-(\theta)$, and $w^+(\theta, \mu, \hat{\mu}) := w(\theta, \mu, \hat{\mu}, S^+(\theta))$, $w^-(\theta, \mu, \hat{\mu}) := w(\theta, \mu, \hat{\mu}, S^-(\theta))$. Let $A^+(\theta) := \{a \in A : a_\theta > \hat{a}_\theta(\theta)\}$, $A^-(\theta) := \{a \in A : a_\theta < \hat{a}_\theta(\theta)\}$.

As in Appendix A.2, for any distinct $a, a' \in A$, define $\Theta_{a \succ a'} := \{\theta \in \Theta : a_\theta > a'_\theta\}$. By regularity of A, $\Theta_{a \succ a'} \cup \Theta_{a' \succ a} = \Theta$. For any $\varepsilon \in \mathbb{R}$, define

$$C_{a,a'}^{\varepsilon,\hat{\mu}} := \left\{\nu \in \Delta(X) : \max_{\theta \in \Theta_{a \succ a'}} \nu \cdot \log \hat{\mu}_\theta > \max_{\theta \in \Theta_{a' \succ a}} \nu \cdot \log \hat{\mu}_\theta + \varepsilon\right\}, \quad C_a^{\varepsilon,\hat{\mu}} := \bigcap_{a' \in A \setminus \{a\}} C_{a,a'}^{\varepsilon,\hat{\mu}}.$$

For each $a \in A$, we observe that

$$C_a^{0,\hat{\mu}} = \left\{ \nu \in \Delta(X) : \max_{\theta \in \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta > \max_{\theta \notin \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta \right\}. \tag{26}$$

Moreover, the closure of $C_a^{0,\hat{\mu}}$ is given by

$$\overline{C}_a^{0,\hat{\mu}} = \left\{ \nu \in \Delta(X) : \max_{\theta \in \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta \geq \max_{\theta \notin \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta \right\}, \tag{27}$$

because if $\nu \cdot \log \hat{\mu}_{\theta'} = \max_{\theta \notin \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta$ for some $\theta' \in \alpha^{-1}(a)$, then $((1-\varepsilon)\nu + \varepsilon\nu') \cdot \log \hat{\mu}_{\theta'} > \max_{\theta \notin \alpha^{-1}(a)}((1-\varepsilon)\nu + \varepsilon\nu') \cdot \log \hat{\mu}_\theta$ for any $\varepsilon \in (0,1)$ and any ν' such that $\nu' \cdot \log \hat{\mu}_{\theta'} > \nu' \cdot \log \hat{\mu}_\theta$ for all $\theta \neq \theta'$ (which exists by Assumption 2).

Lemma 6. *For each $\theta \in \Theta$ and $a \in A$, we have*

$$w^+(\theta, \mu, \hat{\mu}) = \inf_{a \in A^+(\theta), \nu \in C_a^{0,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta), \qquad w^-(\theta, \mu, \hat{\mu}) = \inf_{a \in A^-(\theta), \nu \in C_a^{0,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta).$$

Proof. We prove the result for w^+. The proof for w^- is analogous. Observe that

$$\begin{aligned}
\inf_{a \in A^+(\theta), \nu \in C_a^{0,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta) &= \inf_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \text{ s.t. } \nu \in \bigcup_{a \in A^+(\theta)} \overline{C}_a^{0,\hat{\mu}} \\
&= \inf_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \text{ s.t. } \max_{\theta' \in S^+(\theta)} \nu \cdot \log \hat{\mu}_{\theta'} \geq \max_{\theta' \notin S^+(\theta)} \nu \cdot \log \hat{\mu}_{\theta'} \\
&= \inf_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \text{ s.t. } \max_{\theta' \in S^+(\theta)} \nu \cdot \log \hat{\mu}_{\theta'} = \max_{\theta' \notin S^+(\theta)} \nu \cdot \log \hat{\mu}_{\theta'}.
\end{aligned}$$

Indeed, the first equality follows by continuity of KL, and the second equality follows from (27) and the fact that $S^+(\theta) = \alpha^{-1}(A^+(\theta))$. To see the last equality, consider the case in which the inequality constraint in the second line is feasible (if infeasible, then the equality constraint in the third line is also infeasible) and suppose for a contradiction that under any ν that achieves the infimum value the constraint holds with a strict inequality. Note that $\nu \neq \mu_\theta$ for any such ν, as $\mathrm{argmax}_{\theta' \in \Theta} \mu_\theta \cdot \log \hat{\mu}_{\theta'} \in S_{\hat{\mu}}(\theta)$ and $S_{\hat{\mu}}(\theta) \cap S^+(\theta) = \emptyset$. Now consider $\hat{\nu} = \lambda\nu + (1-\lambda)\mu_\theta$, which satisfies $\max_{\theta' \in S^+(\theta)} \hat{\nu} \cdot \log \hat{\mu}_{\theta'} > \max_{\theta' \notin S^+(\theta)} \hat{\nu} \cdot \log \hat{\mu}_{\theta'}$ if $\lambda < 1$ is sufficiently close to 1. By strict convexity of Kullback-Leibler divergence, we then have $\mathrm{KL}(\hat{\nu}, \mu_\theta) < \lambda \mathrm{KL}(\nu, \mu_\theta) < \mathrm{KL}(\nu, \mu_\theta)$, contradicting the choice of ν.

Finally, rearranging yields

$$\begin{aligned}
&\inf_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \text{ s.t. } \max_{\theta' \in S^+(\theta)} \nu \cdot \log \hat{\mu}_{\theta'} = \max_{\theta' \notin S^+(\theta)} \nu \cdot \log \hat{\mu}_{\theta'} \\
&= \inf_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_\theta) \text{ s.t. } \min_{\theta' \in S^+(\theta)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) = \min_{\theta' \notin S^+(\theta)} \mathrm{KL}(\nu, \hat{\mu}_{\theta'}) \\
&= w^+(\theta, \mu, \hat{\mu}),
\end{aligned}$$

as required. ☐

Lemma 7. *For each $\varepsilon > 0$ and $a \in A$, there exists t^* such that for all $t \geq t^*$,*

$$C_a^{\varepsilon,\hat{\mu}} \subseteq \left\{ \nu \in \Delta(X) : \sum_{\theta \in \Theta} a_\theta p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} > \max_{a' \in A \setminus \{a\}} \sum_{\theta \in \Theta} a'_\theta p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} \right\}$$
$$\subseteq \left\{ \nu \in \Delta(X) : \sum_{\theta \in \Theta} a_\theta p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} \geq \max_{a' \in A \setminus \{a\}} \sum_{\theta \in \Theta} a'_\theta p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} \right\} \subseteq C_a^{-\varepsilon,\hat{\mu}}.$$

Proof. Since A is essential, $\alpha^{-1}(a) \neq \emptyset$. Since A is regular, there exists $K > 0$ sufficiently large such that

$$\frac{1}{K} < \min_{\theta \in \Theta} \min_{a'' \neq a'} |a'_\theta - a''_\theta| \leq \max_{\theta \in \Theta} \max_{a'' \neq a'} |a'_\theta - a''_\theta| < K.$$

Choose t^* sufficiently large such that

$$e^{t^* \varepsilon} \frac{1}{K} \min_{\theta \in \Theta} p_0(\theta) > |\Theta| K \max_{\theta \in \Theta} p_0(\theta).$$

To show the first set inclusion, consider any $\nu \in C_a^{\varepsilon,\hat{\mu}}$ and $t \geq t^*$. For any $a' \neq a$, note that $\nu \in C_{a,a'}^{\varepsilon,\hat{\mu}}$ ensures $\nu \cdot \log \hat{\mu}_{\theta'} \geq \max_{\theta \in \Theta_{a' \succ a}} \nu \cdot \log \hat{\mu}_\theta + \varepsilon$, where $\theta' \in \operatorname{argmax}_{\theta \in \Theta_{a \succ a'}} \nu \cdot \log \hat{\mu}_\theta$. Then

$$\begin{aligned}
\sum_{\theta \in \Theta} (a_\theta - a'_\theta) p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} &= \sum_{\theta \in \Theta_{a \succ a'}} (a_\theta - a'_\theta) p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} + \sum_{\theta \in \Theta_{a' \succ a}} (a_\theta - a'_\theta) p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} \\
&\geq \frac{1}{K} p_0(\theta') e^{t\nu \cdot \log \hat{\mu}_{\theta'}} - |\Theta_{a' \succ a}| K \max_{\theta \in \Theta_{a' \succ a}} p_0(\theta) e^{t(\nu \cdot \log \hat{\mu}_{\theta'} - \varepsilon)} \\
&= e^{t\nu \cdot \log \hat{\mu}_{\theta'}} \left(\frac{1}{K} p_0(\theta') - |\Theta_{a' \succ a}| K \max_{\theta \in \Theta_{a' \succ a}} p_0(\theta) e^{-t\varepsilon} \right) > 0.
\end{aligned}$$

To show the last set inclusion, consider any $\nu \notin C_a^{-\varepsilon,\hat{\mu}}$ and $t \geq t^*$. Since $\nu \notin C_a^{-\varepsilon,\hat{\mu}}$, $\nu \notin C_{a,a'}^{-\varepsilon,\hat{\mu}}$ for some $a' \neq a$. Thus $\nu \cdot \log \hat{\mu}_{\theta'} - \varepsilon \geq \max_{\theta \in \Theta_{a \succ a'}} \nu \cdot \log \hat{\mu}_\theta$, where $\theta' \in \operatorname{argmax}_{\theta \in \Theta_{a' \succ a}} \nu \cdot \log \hat{\mu}_\theta$. Then observe that

$$\begin{aligned}
\sum_{\theta \in \Theta} (a_\theta - a'_\theta) p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} &= \sum_{\theta \in \Theta_{a \succ a'}} (a_\theta - a'_\theta) p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} + \sum_{\theta \in \Theta_{a' \succ a}} (a_\theta - a'_\theta) p_0(\theta) e^{t\nu \cdot \log \hat{\mu}_\theta} \\
&\leq K |\Theta_{a \succ a'}| \max_{\theta \in \Theta_{a \succ a'}} p_0(\theta) e^{t(\nu \cdot \log \hat{\mu}_{\theta'} - \varepsilon)} - \frac{1}{K} p_0(\theta') e^{t\nu \cdot \log \hat{\mu}_{\theta'}} \\
&= e^{t\nu \cdot \log \hat{\mu}_{\theta'}} e^{-\varepsilon t} \left(K |\Theta_{a \succ a'}| \max_{\theta \in \Theta_{a \succ a'}} p_0(\theta) - \frac{1}{K} p_0(\theta') e^{\varepsilon t} \right) < 0.
\end{aligned}$$

☐

Lemma 8. *For each $\theta^* \in \Theta$ and $a \in A$, $\inf_{\nu \in C_a^{\varepsilon,\hat{\mu}}} \mathrm{KL}(\nu, \mu_{\theta^*})$ is finite and continuous in ε in a neighborhood of* 0.

Proof. Since A is regular and essential, $\alpha^{-1}(a)$ and $\Theta \setminus \alpha^{-1}(a)$ are both non-empty. Thus, by Assumption 2, $C_a^{\varepsilon,\hat{\mu}}$ is nonempty for ε close to 0. Since $\overline{C}_a^{\varepsilon,\hat{\mu}}$ is compact and $\mathrm{KL}(\cdot, \mu_{\theta^*})$ is continuous (as μ_{θ^*} has full support), $\inf_{\nu \in C_a^{\varepsilon,\hat{\mu}}} \mathrm{KL}(\nu, \mu_{\theta^*}) = \inf_{\nu \in \overline{C}_a^{\varepsilon,\hat{\mu}}} \mathrm{KL}(\nu, \mu_{\theta^*})$ is finite for all ε in a neighborhood of 0.

We now verify that $\overline{C}_a^{\varepsilon,\hat{\mu}}$ is a continuous correspondence at $\varepsilon = 0$, which ensures that $\inf_{\nu \in C_a^{\varepsilon,\hat{\mu}}} \mathrm{KL}(\nu, \mu_{\theta^*})$ is continuous in ε at $\varepsilon = 0$ by Berge's theorem of the maximum. It is straightforward to prove that $\overline{C}_a^{\varepsilon,\hat{\mu}}$ is upper-hemicontinuous in ε.

To see that $\overline{C}_a^{\varepsilon,\hat{\mu}}$ is also lower-hemicontinuous at $\varepsilon = 0$, take any $\nu \in \overline{C}_a^{0,\hat{\mu}}$ and any sequence $\varepsilon_n \to 0$. Note that (27) holds at $\varepsilon = 0$. If $\max_{\theta \in \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta > \max_{\theta \notin \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta$, then $\nu \in C_a^{\varepsilon_n,\hat{\mu}}$ for n sufficiently large. Thus, assume $\max_{\theta \in \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta = \max_{\theta \notin \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta$. Pick any $\theta' \in \operatorname{argmax}_{\theta \in \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta$. Then $\nu \cdot \log \hat{\mu}_{\theta'} = \max_{\theta \notin \alpha^{-1}(a)} \nu \cdot \log \hat{\mu}_\theta$. Moreover, by Assumption 2, there exists $\nu' \in \Delta(X)$ such that $\nu' \cdot \log \hat{\mu}_{\theta'} > \max_{\theta \notin \alpha^{-1}(a)} \nu' \cdot \log \hat{\mu}_\theta$. Thus, for n sufficiently large, there exists $\kappa_n \in [0, 1]$ such that $\lim_{n\to\infty} \kappa_n = 1$ and

$$(\kappa_n \nu + (1 - \kappa_n)\nu') \cdot \log \hat{\mu}_{\theta'} = \max_{\theta \notin \alpha^{-1}(a)} (\kappa_n \nu + (1 - \kappa_n)\nu') \cdot \log \hat{\mu}_\theta + |\varepsilon_n|.$$

Then $(\kappa_n \nu + (1 - \kappa_n)\nu') \in \overline{C}_a^{\varepsilon_n,\hat{\mu}}$ and $\lim_{n\to\infty}(\kappa_n \nu + (1 - \kappa_n)\nu') = \nu$. This establishes lower hemicontinuity. □

Finally, consider any $\hat{\mu}^1$ and $\hat{\mu}^2$ satisfying Assumptions 2 and 5 and such that $S_{\hat{\mu}^i}(\theta) = \hat{S}(\theta)$ for all θ and $i = 1, 2$.

Lemma 9. *For any ε in some open neighborhood of* 0*,*

$$\lim_{t\to\infty} \frac{1}{t} \log \left(\sum_{\theta \in \Theta} p_0(\theta) \left(\mathbb{P}_\theta \left[\left\{ \nu_t \in \bigcup_{a \in A^+(\theta)} C_a^{\varepsilon,\hat{\mu}^1} \right\} \right] + \mathbb{P}_\theta \left[\left\{ \nu_t \in \bigcup_{a \in A^-(\theta)} C_a^{\varepsilon,\hat{\mu}^2} \right\} \right] \right) \right)$$
$$= - \min_{\theta \in \Theta} \min \left\{ \inf_{a \in A^+(\theta), \nu \in C_a^{\varepsilon,\hat{\mu}^1}} \mathrm{KL}(\nu, \mu_\theta), \inf_{a \in A^-(\theta), \nu \in C_a^{\varepsilon,\hat{\mu}^2}} \mathrm{KL}(\nu, \mu_\theta) \right\}.$$

Proof. By Lemma 8, $C_a^{\varepsilon,\hat{\mu}^1}$ and $C_a^{\varepsilon,\hat{\mu}^2}$ are non-empty for any ε in some open neighborhood of 0. Fix any such ε. By Sanov's theorem, for each $\hat{\mu} \in \{\hat{\mu}^1, \hat{\mu}^2\}$, a, and θ,

$$\inf_{\nu \in C_a^{\varepsilon,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta) \leq - \lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta \left[\{\nu_t \in C_a^{\varepsilon,\hat{\mu}}\} \right] \leq \inf_{\nu \in \mathrm{int} C_a^{\varepsilon,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta) = \inf_{\nu \in C_a^{\varepsilon,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta)$$

where the equality follows from the continuity of $\mathrm{KL}(\nu, \mu_\theta)$ in ν (as μ_θ has full support). Thus, $\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\theta \left[\{\nu_t \in C_a^{\varepsilon,\hat{\mu}}\} \right] = - \inf_{\nu \in C_a^{\varepsilon,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta)$.

The above implies the desired claim based on standard arguments from large deviation theory (e.g., Lemma 1.2.15 in Dembo and Zeitouni, 2010). $\square$

B.1.2 Proof of Theorem 5

Take $\hat{\mu}^1, \hat{\mu}^2$, S, S^+, and S^- as in Theorem 5. Consider any regular and essential A with $S_A = S$, $S_A^+ = S^+$, $S_A^- = S^-$. By regularity of A, there exist $\underline{M} < \overline{M}$ such that

$$0 < \underline{M} < \min_{\theta\in\Theta} \min_{a\in A^+(\theta)\cup A^-(\theta)} |a_\theta - \hat{a}_\theta(\theta)| \leq \max_{\theta\in\Theta} \max_{a\in A^+(\theta)\cup A^-(\theta)} |a_\theta - \hat{a}_\theta(\theta)| < \overline{M}.$$

For each $\hat{\mu} \in \{\hat{\mu}^1, \hat{\mu}^2\}$, Lemma 6 implies that

$$w^+(\theta, \mu, \hat{\mu}) = \inf_{a\in A^+(\theta),\nu\in C_a^{0,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta), \qquad w^-(\theta, \mu, \hat{\mu}) = \inf_{a\in A^-(\theta),\nu\in C_a^{0,\hat{\mu}}} \mathrm{KL}(\nu, \mu_\theta).$$

Since $\min\{w^+(\mu, \hat{\mu}^2, S^+), w^-(\mu, \hat{\mu}^1, S^-)\} > \min\{w^+(\mu, \hat{\mu}^1, S^+), w^-(\mu, \hat{\mu}^2, S^-)\}$, Lemma 8 yields some $\varepsilon > 0$ such that

$$\begin{aligned}&\min_{\theta\in\Theta} \min\left\{\inf_{a\in A^+(\theta),\nu\in C_a^{\varepsilon,\hat{\mu}^1}} \mathrm{KL}(\nu, \mu_\theta), \inf_{a\in A^-(\theta),\nu\in C_a^{\varepsilon,\hat{\mu}^2}} \mathrm{KL}(\nu, \mu_\theta)\right\}\\ &> \min_{\theta\in\Theta} \min\left\{\inf_{a\in A^+(\theta),\nu\in C_a^{-\varepsilon,\hat{\mu}^2}} \mathrm{KL}(\nu, \mu_\theta), \inf_{a\in A^-(\theta),\nu\in C_a^{-\varepsilon,\hat{\mu}^1}} \mathrm{KL}(\nu, \mu_\theta)\right\}.\end{aligned}$$

By Lemma 9, this implies that there exists some $\bar{T}$ such that for all $T \geq \bar{T}$,

$$\begin{aligned}&\sum_{\theta\in\Theta} p_0(\theta) \left(\mathbb{P}_\theta\left[\left\{\nu_T \in \bigcup_{a\in A^+(\theta)} C_a^{\varepsilon,\hat{\mu}^1}\right\}\right] + \mathbb{P}_\theta\left[\left\{\nu_T \in \bigcup_{a\in A^-(\theta)} C_a^{\varepsilon,\hat{\mu}^2}\right\}\right]\right) \underline{M}\\ &> \sum_{\theta\in\Theta} p_0(\theta) \left(\mathbb{P}_\theta\left[\left\{\nu_T \in \bigcup_{a\in A^+(\theta)} C_a^{-\varepsilon,\hat{\mu}^2}\right\}\right] + \mathbb{P}_\theta\left[\left\{\nu_T \in \bigcup_{a\in A^-(\theta)} C_a^{-\varepsilon,\hat{\mu}^1}\right\}\right]\right) \overline{M}.\end{aligned}$$

Rearranging the above inequality yields that, for all $T \geq \bar{T}$,

$$\begin{aligned}&\sum_{\theta\in\Theta} p_0(\theta) \left(\mathbb{P}_\theta\left[\left\{\nu_T \in \bigcup_{a\in A^+(\theta)} C_a^{\varepsilon,\hat{\mu}^1}\right\}\right] \underline{M} - \mathbb{P}_\theta\left[\left\{\nu_T \in \bigcup_{a\in A^-(\theta)} C_a^{-\varepsilon,\hat{\mu}^1}\right\}\right] \overline{M}\right)\\ &> \sum_{\theta\in\Theta} p_0(\theta) \left(\mathbb{P}_\theta\left[\left\{\nu_T \in \bigcup_{a\in A^+(\theta)} C_a^{-\varepsilon,\hat{\mu}^2}\right\}\right] \overline{M} - \mathbb{P}_\theta\left[\left\{\nu_T \in \bigcup_{a\in A^-(\theta)} C_a^{\varepsilon,\hat{\mu}^2}\right\}\right] \underline{M}\right).\end{aligned} \tag{28}$$

Let $W_T(\mu, \hat{\mu}^i, A, \theta)$ denote objective expected payoffs conditional on state θ. Then, applying Lemma 7, we have that, for all sufficiently large T,

$$\begin{aligned} W_T(\mu, \hat{\mu}^2, A, \theta) - \hat{a}_\theta(\theta) &\leq \mathbb{P}_\theta \left[\bigcup_{a \in A^+(\theta)} \left\{ \max_{a' \in A \setminus \{a\}} \sum_{\theta \in \Theta} (a'_\theta - a_\theta) p_0(\theta) e^{T \nu_T \cdot \log \hat{\mu}^2_\theta} \leq 0 \right\} \right] \overline{M} \\ &\quad - \mathbb{P}_\theta \left[\bigcup_{a \in A^-(\theta)} \left\{ \max_{a' \in A \setminus \{a\}} \sum_{\theta \in \Theta} (a'_\theta - a_\theta) p_0(\theta) e^{T \nu_T \cdot \log \hat{\mu}^2_\theta} < 0 \right\} \right] \underline{M} \\ &\leq \mathbb{P}_\theta \left[\left\{ \nu_T \in \bigcup_{a \in A^+(\theta)} C_a^{-\varepsilon, \hat{\mu}^2} \right\} \right] \overline{M} - \mathbb{P}_\theta \left[\left\{ \nu_T \in \bigcup_{a \in A^-(\theta)} C_a^{\varepsilon, \hat{\mu}^2} \right\} \right] \underline{M}. \end{aligned}$$

Similarly, for all sufficiently large T,

$$\begin{aligned} W_T(\mu, \hat{\mu}^1, A, \theta) - \hat{a}_\theta(\theta) &\geq \mathbb{P}_\theta \left[\bigcup_{a \in A^+(\theta)} \left\{ \max_{a' \in A \setminus \{a\}} \sum_{\theta \in \Theta} (a'_\theta - a_\theta) p_0(\theta) e^{T \nu_T \cdot \log \hat{\mu}^1_\theta} < 0 \right\} \right] \underline{M} \\ &\quad - \mathbb{P}_\theta \left[\bigcup_{a \in A^-(\theta)} \left\{ \max_{a' \in A \setminus \{a\}} \sum_{\theta \in \Theta} (a'_\theta - a_\theta) p_0(\theta) e^{T \nu_T \cdot \log \hat{\mu}^1_\theta} \leq 0 \right\} \right] \overline{M} \\ &\geq \mathbb{P}_\theta \left[\bigcup_{a \in A^+(\theta)} C_a^{\varepsilon, \hat{\mu}^1} \right] \underline{M} - \mathbb{P}_\theta \left[\bigcup_{a \in A^-(\theta)} C_a^{-\varepsilon, \hat{\mu}^1} \right] \overline{M}. \end{aligned}$$

Thus, there exists $T^* \geq \bar{T}$ such that for all $T \geq T^*$,

$$\begin{aligned} &\sum_{\theta \in \Theta} p_0(\theta) \left(\mathbb{P}_\theta \left[\left\{ \nu_T \in \bigcup_{a \in A^+(\theta)} C_a^{\varepsilon, \hat{\mu}^1} \right\} \right] \underline{M} - \mathbb{P}_\theta \left[\left\{ \nu_T \in \bigcup_{a \in A^-(\theta)} C_a^{-\varepsilon, \hat{\mu}^1} \right\} \right] \overline{M} \right) \\ &\leq W_T(\mu, \hat{\mu}^1, A) - \sum_{\theta \in \Theta} p_0(\theta) \hat{a}_\theta(\theta); \\ &\sum_{\theta \in \Theta} p_0(\theta) \left(\mathbb{P}_\theta \left[\left\{ \nu_T \in \bigcup_{a \in A^+(\theta)} C_a^{-\varepsilon, \hat{\mu}^2} \right\} \right] \overline{M} - \mathbb{P}_\theta \left[\left\{ \nu_T \in \bigcup_{a \in A^-(\theta)} C_a^{\varepsilon, \hat{\mu}^2} \right\} \right] \underline{M} \right) \\ &\geq W_T(\mu, \hat{\mu}^2, A) - \sum_{\theta \in \Theta} p_0(\theta) \hat{a}_\theta(\theta). \end{aligned}$$

These bounds together with inequality (28) imply that for all $T \geq T^*$, $W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A)$. □

$\mu, \hat{\mu}$	x_0	x_1	x_2
$\overline{\theta}$	$0.3, 0.1$	$0.4, 0.5$	$0.3, 0.4$
$\underline{\theta}$	$0.4, 0.3$	$0.5, 0.6$	$0.1, 0.1$

$\mu', \hat{\mu}'$	$\{x_0, x_1\}$	$\{x_2\}$
$\overline{\theta}$	$0.7, 0.6$	$0.3, 0.4$
$\underline{\theta}$	$0.9, 0.9$	$0.1, 0.1$

Figure 4: *Left: True and perceived signal distributions $\mu, \hat{\mu}$ before coarsening. Right: True and perceived signal distributions $\mu', \hat{\mu}'$ after coarsening signals by bunching together x_0 and x_1.*

C Coarsening Signals: An Example

We provide an example where coarsening signals (see Section 5) can robustly improve welfare by strictly increasing a biased agent's learning efficiency. Consider the state space $\Theta = \{\underline{\theta}, \overline{\theta}\}$ and signal space $X = \{x_0, x_1, x_2\} \subseteq \mathbb{R}$ with $x_0 < x_1 < x_2$.

Suppose the true and perceived signal distributions $\mu, \hat{\mu}$ before coarsening are as shown in the left-hand panel of Figure 4. Note that $\hat{\mu}_\theta \geq_{\text{FOSD}} \mu_\theta$ for all θ, i.e., the agent is overconfident. The agent's asymptotic beliefs are correct (Assumption 1 holds), and her learning efficiency index can be computed as $w(\mu, \hat{\mu}) = w(\overline{\theta}, \mu, \hat{\mu}) \approx 0.00009$.

Now, consider the effect of coarsening signals by bunching together x_0 and x_1. This results in the true and perceived signal structures $\mu', \hat{\mu}'$ shown in the right-hand panel of Figure 4. Assumption 1 continues to hold, but the efficiency index increases to $w(\mu', \hat{\mu}') = w(\overline{\theta}, \mu', \hat{\mu}') \approx 0.013$. Thus, by the generalization of Theorem 1 in Remark 2, coarsening robustly improves dynamic welfare. Intuitively, before coarsening, the agent's misinferences from signals x_0 and x_1 go in opposite directions; bunching these two signals together shifts the agent's indistinguishability set closer to the correctly specified case.