Welfare Comparisons for Biased Learning: Non-i.i.d. Environments*

Preliminary and incomplete

Mira Frick Ryota Iijima Yuhta Ishii

December 28, 2022

Abstract

This note extends the characterization of the dynamic welfare ranking over learning biases in Frick, Iijima, and Ishii (2022b) (henceforth, FII) to two simple settings in which signals are not i.i.d. across draws. First, we consider a setting in which signal distributions can depend on the agent's current belief. Second, we consider signals that follow an exogenous Markov process.

1 Environment

Throughout this note, a state θ is drawn once and for all from a binary set $\Theta = \{\overline{\theta}, \underline{\theta}\}$ 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 a sequence of T signal draws $x_1, x_2, \ldots, x_T \in X$, where X is a finite set. In FII, signals at each t are drawn i.i.d. according to a fixed exogenous signal structure and the agent Bayesian updates beliefs based on a fixed incorrect perception of the signal distribution. The following sections move beyond the i.i.d. assumption in two ways.

*Frick: Yale University (mira.frick@yale.edu); Iijima: Yale University (ryota.iijima@yale.edu); Ishii: Pennsylvania State University (yxi5014@psu.edu).

2 Endogenous Signals

2.1 Setting

In this section, we allow for the possibility that the true and/or perceived distributions of each signal draw depend on the agent's current belief. For tractability, we focus on the simplest setting in which the true and perceived signal distributions take two possible forms, depending on whether the agent's current belief is above or below some threshold.

Formally, the agent has a full-support prior belief $p \in \Delta(\Theta)$; for simplicity, we assume that $p = p_0$ is correct, but incorrect priors do not change our results. For each $t = 1, \ldots, T$, the signal sequence $x^t = (x_1, \ldots, x_t) \in X^t$ and the agent's posterior belief $p(\cdot \mid x^t) \in \Delta(\Theta)$ are then determined recursively as follows. Conditional on realized state θ and signal history x^{t-1}, signal x_t is drawn according to distribution $\mu_{+\theta} \in \Delta(X)$ if $p(\overline{\theta} \mid x^{t-1}) \geq \kappa$ and according to distribution $\mu_{-\theta} \in \Delta(X)$ if $p(\overline{\theta} \mid x^{t-1}) < \kappa$, where $\kappa \in (0,1)$ is a fixed threshold and $p(\cdot \mid x^0) := p(\cdot)$. We call $\mu := (\mu_{+\theta}, \mu_{-\theta})_{\theta \in \Theta}$ the ***true signal structure*** and assume that each distribution $\mu_{+\theta}$ and $\mu_{-\theta}$ has full support. The agent's posterior belief $p(\cdot \mid x^t)$ is formed by applying Bayes' rule based on her ***perceived signal structure*** $\hat{\mu} := (\hat{\mu}_{+\theta}, \hat{\mu}_{-\theta})_{\theta \in \Theta}$. That is, for all $\theta' \in \Theta$, $\hat{\mu}_{+\theta'}, \hat{\mu}_{-\theta'} \in \mathbb{R}^X_{++}$ capture the agent's perceived belief-dependent likelihoods of signals in state θ' and her posterior belief is given by

$$p(\theta' \mid x^t) = \begin{cases} \frac{p(\theta'|x^{t-1})\hat{\mu}_{+\theta'}(x_t)}{\sum_{\theta''} p(\theta''|x^{t-1})\hat{\mu}_{+\theta''}(x_t)} & \text{if } p(\overline{\theta} \mid x^{t-1}) \geq \kappa \\ \frac{p(\theta'|x^{t-1})\hat{\mu}_{-\theta'}(x_t)}{\sum_{\theta''} p(\theta''|x^{t-1})\hat{\mu}_{-\theta''}(x_t)} & \text{if } p(\overline{\theta} \mid x^{t-1}) < \kappa. \end{cases}$$

While we assume that $\hat{\mu}_{+\theta}(x), \hat{\mu}_{-\theta}(x) > 0$ for each $x \in X$, we do not require that $\sum_{x \in X} \hat{\mu}_{-\theta}(x) = \sum_{x \in X} \hat{\mu}_{+\theta}(x) = 1$. As in FII, this makes it possible to accommodate both misspecified Bayesian learning as well as certain classes of non-Bayesian learning. We call the agent ***correctly specified*** if $\hat{\mu} = \mu$. The setting reduces to the one studied in FII if $\mu_{+\theta} = \mu_{-\theta}$ and $\hat{\mu}_{+\theta} = \hat{\mu}_{-\theta}$ for each θ.[1]

Upon observing the realized signal sequence $x^T = (x_1, \ldots, x_T)$, the agent faces a ***decision problem***, which is a non-empty finite set A of acts. A (payoff) ***act*** is a vector $a \in \mathbb{R}^\Theta$, where a_θ denotes the agent's payoff from a conditional on state θ.

[1] We focus on binary states while FII allow for general finite state spaces. We leave the analysis of general states as an interesting open question.

To avoid triviality, we assume that 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 $b \in A$. Given any decision problem A, the agent chooses an act $a^*(x^T) \in A$ to maximize her subjective expected payoff under her posterior belief $p(\cdot \mid x^T)$:[2]

$$a^*(x^T) \in \underset{a \in A}{\operatorname{argmax}} \sum_{\theta} p(\theta \mid x^T) a_\theta. \tag{1}$$

The agent's ***welfare*** is her objective ex-ante expected payoff to choosing her subjectively optimal act $a^*(x^T)$ at each x^T, where expectations over signal realizations are based on the *true* signal structure μ. That is, letting $\mathbb{P}_\theta$ denote the true probability measure over signal sequences in X^T conditional on state θ and letting $\mathbb{E}_\theta$ denote the corresponding expectation operator, the agent's welfare is given by

$$W_T(\mu, \hat{\mu}, \kappa, A) = \sum_{\theta \in \Theta} p_0(\theta) \mathbb{E}_\theta[a^*_\theta(x^T)]. \tag{2}$$

Remark 1 (Examples). The above setting can capture simple forms of misspecified active learning, where at each t, the agent chooses an information structure from the binary set μ_+ and μ_- according to some belief-dependent threshold rule (the threshold κ may depend on the decision problem), but the agent misperceives the signal structures to be $\hat{\mu}_+$ and $\hat{\mu}_-$. The setting can also capture various widely studied belief-dependent departures from Bayesian updating, such as confirmation bias (Example 1), where the true signal distributions are exogenous ($\mu_{+\theta} = \mu_{-\theta}$ for all θ) but the agent's reaction to signals depends on her current belief. ▲

Given any true signal structure μ, consider two agents $i = 1, 2$ that differ in their perceived signal structures $\hat{\mu}^i$ (as well as possibly their thresholds κ^i). As in FII, we seek to characterize a robust dynamic welfare ranking, whereby agent 1's bias is less harmful than agent 2's if in *all* decision problems A, agent 1's welfare exceeds agent 2's for any large enough number T of signal draws.

2.2 Characterization

We restrict attention to the case where agents' biases are small enough that they learn the true state as $T \to \infty$. Thus, asymptotic beliefs are the same as under correctly

[2] Our results do not depend on how an agent breaks ties in case (1) admits multiple solutions.

specified learning and hence cannot be used to compare welfare under the two biases. Instead, we show that, generalizing FII, the dynamic welfare ranking is characterized by a learning efficiency index that captures the speed of convergence. Analogous to Section 5.2 in FII, this approach can also be extended to the case where agents' beliefs are asymptotically incorrect but the same.

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. This assumption implies that an agent's belief converges almost surely to a point-mass on the true state as $T \to \infty$.[3]

Assumption 1. For any distinct $\theta, \theta' \in \Theta$ and $\# \in \{-, +\}$,

$$\mathrm{KL}(\mu_{\#\theta}, \hat{\mu}_{\#\theta}) < \mathrm{KL}(\mu_{\#\theta}, \hat{\mu}_{\#\theta'}).$$

We define the following learning efficiency index, which reduces to the one in FII when $\mu_{+\theta} = \mu_{-\theta}$ and $\hat{\mu}_{+\theta} = \hat{\mu}_{-\theta}$ for each θ:

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

$$w_\theta(\mu, \hat{\mu}) := \min_{\alpha \in [0,1], \nu_+, \nu_- \in \Delta(X)} \alpha \mathrm{KL}(\nu_-, \mu_{-\theta}) + (1-\alpha) \mathrm{KL}(\nu_+, \mu_{+\theta}) \tag{3}$$

subject to

$$\alpha \mathrm{KL}(\nu_-, \hat{\mu}_{-\underline{\theta}}) + (1-\alpha) \mathrm{KL}(\nu_+, \hat{\mu}_{+\underline{\theta}}) = \alpha \mathrm{KL}(\nu_-, \hat{\mu}_{-\bar{\theta}}) + (1-\alpha) \mathrm{KL}(\nu_+, \hat{\mu}_{+\bar{\theta}}), \tag{4}$$

$$\mathrm{KL}(\nu_-, \hat{\mu}_{-\underline{\theta}}) \geq \mathrm{KL}(\nu_-, \hat{\mu}_{-\bar{\theta}}), \tag{5}$$

$$\mathrm{KL}(\nu_+, \hat{\mu}_{+\underline{\theta}}) \leq \mathrm{KL}(\nu_+, \hat{\mu}_{+\bar{\theta}}). \tag{6}$$

We refer to constraint (4) as the ***indistinguishability condition***.

[3] Conditional on each state θ and $\theta' \neq \theta$, a simple calculation shows that $\left(\frac{p(\theta'|x^t)}{p(\theta|x^t)}\right)^q$ is a supermartingale for some small $q > 0$, which converges almost surely (see Frick, Iijima, and Ishii, 2022a). It is easy to see that the limit belief must be the point-mass on θ.

To interpret Definition 1, suppose that $\alpha \in [0,1]$ represents the realized fraction of periods $t \in \{1, \ldots, T\}$ in which the true signal distribution is $\mu_{-\theta}$ and suppose that ν_- represents the empirical signal distribution in these periods, while ν_+ represents the empirical signal distribution in periods where the true signal distribution is $\mu_{+\theta}$. If the realized (α, ν_-, ν_+) satisfy the indistinguishability condition (4), then the agent is unable to distinguish between states $\underline{\theta}$ and $\overline{\theta}$, because the observed data is explained equally well by her perceived signal distributions in both states. The measure $w_\theta(\mu, \hat{\mu})$ captures how "atypical" such realizations of (α, ν_-, ν_+) are under the true signal distributions $(\mu_{+\theta}, \mu_{-\theta})$ in state θ: The greater $w_\theta(\mu, \hat{\mu})$, the more atypical it is for the agent to face such patterns of signals that do not allow her to distinguish between $\underline{\theta}$ and $\overline{\theta}$.

Given (4), the constraints (5)–(6) are redundant if $\alpha \in \{0, 1\}$. However, these constraints discipline ν_-, ν_+ when $\alpha \in (0, 1)$.[4] In particular, when the transition between $\mu_{-\theta}$ and $\mu_{+\theta}$ occurs according to some belief threshold κ, (5) (resp. (6)) ensures that conditional on the posterior satisfying $p(\overline{\theta} \mid x^{t-1}) < \kappa$ (resp. $p(\overline{\theta} \mid x^{t-1}) \geq \kappa$), the signal frequency indicates that state $\overline{\theta}$ (resp. $\underline{\theta}$) is more likely than the other state. This ensures that, consistent with $\alpha \in (0, 1)$, signal distributions do in fact switch between $\mu_{-\theta}$ and $\mu_{+\theta}$ in a stable manner.

Note that the learning efficiency index is defined without reference to any decision problem A (nor an agent's threshold κ). The following theorem shows that this index characterizes the dynamic welfare ranking:

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 decision problem A and $\kappa^1, \kappa^2 \in (0, 1)$, there exists T^* such that $W_T(\mu, \hat{\mu}^1, \kappa^1, A) > W_T(\mu, \hat{\mu}^2, \kappa^2, A)$ for every $T \geq T^*$.*

The key idea behind the proof is to show that for any decision problem and κ^i, $w_\theta(\mu, \hat{\mu}^i)$ represents the exponential rate at which agent i's probability of choosing an ex-post suboptimal act vanishes in state θ. We build on the large deviation result in Dupuis and Ellis (1992), combined with the "variational formula" that provides a dual representation of rate functions based on KL divergence.

[4] Given (4), it is without loss to drop one of (5) or (6).

2.3 Comparison with Exogenous Case

An implication of Theorem 1 is that endogenous signals always lead to (weakly) lower dynamic welfare than the corresponding exogenous signal benchmarks. To see this, for each $\# \in \{-,+\}$, define $w_{\#}(\mu,\hat{\mu}) := \min_\theta w_{\#\theta}(\mu,\hat{\mu})$, where

$$w_{\#\theta}(\mu,\hat{\mu}) := \min_{\nu\in\Delta(X)} \mathrm{KL}(\nu,\mu_{\#\theta}) \text{ s.t. } \mathrm{KL}(\nu,\hat{\mu}_{\#\overline{\theta}}) = \mathrm{KL}(\nu,\hat{\mu}_{\#\underline{\theta}}).$$

That is, $w_{\#}(\mu,\hat{\mu})$ is the learning efficiency index for the exogenous benchmark studied in FII, where the true (resp. perceived) signal distribution in state θ is $\mu_{\#\theta}$ (resp. $\hat{\mu}_{\#\theta}$) at each t.

Observe that restricting to $\alpha = 1$ (resp. $\alpha = 0$) in the minimization problem (3) yields the value $w_{-\theta}(\mu,\hat{\mu})$ (resp. $w_{+\theta}(\mu,\hat{\mu})$). Thus,

$$w_\theta(\mu,\hat{\mu}) \le \min\{w_{-\theta}(\mu,\hat{\mu}), w_{+\theta}(\mu,\hat{\mu})\} \text{ for each } \theta. \tag{7}$$

This immediately implies the following:

Corollary 1. *For any true and perceived signal structures μ and $\hat{\mu}$, we have*

$$w(\mu,\hat{\mu}) \le \min\{w_{-}(\mu,\hat{\mu}), w_{+}(\mu,\hat{\mu})\}. \tag{8}$$

That is, the efficiency index $w(\mu,\hat{\mu})$ (and hence dynamic welfare, by Theorem 1) under endogenous learning is weakly lower than *both* of the exogenous benchmarks $w_{-}(\mu,\hat{\mu})$ and $w_{+}(\mu,\hat{\mu})$. Intuitively, relative to the exogenous benchmarks, endogenous signals give rise to more possibilities of "confusing" signal realizations that do not allow the agent to distinguish different states, and hence lead to a higher probability of medium-run mistakes. Specifically, index $w(\mu,\hat{\mu})$ allows for the possibility that the agent cannot distinguish $\underline{\theta}$ and $\overline{\theta}$ based on the *combined* realization of (α,ν_-,ν_+) (as captured by (4)), even if the two empirical distributions ν_- and ν_+ *individually* distinguish between states (i.e., conditions (5)–(6) hold strictly). This cannot arise under exogenous signals. Example 2 below highlights a concrete instance where endogenous signals lead to strictly lower learning efficiency than both exogenous benchmarks.

2.4 Examples

We illustrate the dynamic welfare ranking for two classes of learning biases. For simplicity, we consider binary signals, $X = \{\overline{x}, \underline{x}\}$, and assume that the true signal distributions are exogenous and symmetric, i.e., for some $\gamma \in (1/2, 1)$ and each $\# \in \{-, +\}$, we have $\mu_{\#\overline{\theta}}(\overline{x}) = \mu_{\#\underline{\theta}}(\underline{x}) = \gamma$.

We first consider confirmation bias, an important class of belief-dependent updating:

Example 1 (Confirmation bias)**.** As in Charness and Dave (2017) and Benjamin (2019), suppose that for some $c > d > 0$, perceived signal distributions satisfy:[5]

$$\frac{\hat{\mu}_{+\overline{\theta}}(\underline{x})}{\hat{\mu}_{+\underline{\theta}}(\underline{x})} = \left(\frac{\mu_{+\overline{\theta}}(\underline{x})}{\mu_{+\underline{\theta}}(\underline{x})}\right)^{d}, \quad \frac{\hat{\mu}_{+\overline{\theta}}(\overline{x})}{\hat{\mu}_{+\underline{\theta}}(\overline{x})} = \left(\frac{\mu_{+\overline{\theta}}(\overline{x})}{\mu_{+\underline{\theta}}(\overline{x})}\right)^{c},$$

$$\frac{\hat{\mu}_{-\overline{\theta}}(\underline{x})}{\hat{\mu}_{-\underline{\theta}}(\underline{x})} = \left(\frac{\mu_{-\overline{\theta}}(\underline{x})}{\mu_{-\underline{\theta}}(\underline{x})}\right)^{c}, \quad \frac{\hat{\mu}_{-\overline{\theta}}(\overline{x})}{\hat{\mu}_{-\underline{\theta}}(\overline{x})} = \left(\frac{\mu_{-\overline{\theta}}(\overline{x})}{\mu_{-\underline{\theta}}(\overline{x})}\right)^{d}.$$

That is, perceived signal likelihood ratios are distorted according to a distortion factor c for "confirming" signals and a distortion factor d for "disconfirming" signals, where confirming signals are overweighted relative to disconfirming signals. Specifically, when the agent's belief on $\overline{\theta}$ is high (i.e., $\hat{\mu}_+$ governs perceived signals), signal $\overline{x}$ which is indicative of $\overline{\theta}$ is overweighted relative to signal $\underline{x}$. The opposite pattern holds when the belief on $\overline{\theta}$ is low (i.e., $\hat{\mu}_-$ governs perceptions).

In this example, one can show (see Appendix A.3 for details) that

$$w_{\overline{\theta}}(\mu, \hat{\mu}) = w_{-\overline{\theta}}(\mu, \hat{\mu}) < w_{+\overline{\theta}}(\mu, \hat{\mu}), \quad w_{\underline{\theta}}(\mu, \hat{\mu}) = w_{+\underline{\theta}}(\mu, \hat{\mu}) < w_{-\underline{\theta}}(\mu, \hat{\mu}).$$

Thus, (7) holds with equality.

Note that in state $\overline{\theta}$, the agent's belief eventually remains in the $+$ region forever. However, the efficiency index $w_{\overline{\theta}}(\mu, \hat{\mu})$ is equal to the index $w_{-\overline{\theta}}(\mu, \hat{\mu})$ under the exogenous perception benchmark corresponding to the $-$ region. Intuitively, this is because in the latter region, the agent overweights the "incorrect" signal $\underline{x}$ relative to the "correct" signal $\overline{x}$, so this bias is more harmful for welfare than the opposite perception. Analogously, in state $\underline{\theta}$, the efficiency index $w_{\underline{\theta}}(\mu, \hat{\mu})$ is equal to $w_{+\underline{\theta}}(\mu, \hat{\mu})$. ▲

[5] Rabin and Schrag (1999) proposed a different formulation of confirmation bias, which can also be nested by our general model.

The next example illustrates a belief-dependent learning bias for which the inequalities (7)–(8) are strict:

Example 2 (Belief-dependent over-/underinference)**.** Consider a belief-dependent variation of Phillips and Edwards's (1966) model of over-/underinference. There exist $c > d > 0$ such that for all $x \in X$,

$$\frac{\hat{\mu}_{-\overline{\theta}}(x)}{\hat{\mu}_{-\underline{\theta}}(x)} = \left(\frac{\mu_{-\overline{\theta}}(x)}{\mu_{-\underline{\theta}}(x)}\right)^d,$$

$$\frac{\hat{\mu}_{+\overline{\theta}}(x)}{\hat{\mu}_{+\underline{\theta}}(x)} = \left(\frac{\mu_{+\overline{\theta}}(x)}{\mu_{+\underline{\theta}}(x)}\right)^c.$$

That is, perceived signal likelihood ratios of all signals are distorted, where the distortion factor is higher when the current belief on $\overline{\theta}$ is high (in the $+$ region) than when it is low (in the $-$ region).

As shown by FII, for each $\# \in \{-, +\}$ and $\theta \in \Theta$,

$$w_{\#}(\mu, \hat{\mu}) = w_{\#\theta}(\mu, \hat{\mu}) = w_{\#\theta}(\mu, \mu) = w_{\#}(\mu, \mu). \tag{9}$$

Thus, if the distortion factor is belief-independent, this form of bias does not affect learning efficiency relative to the correctly specified case. However, belief-dependent distortion factors can lead to strictly lower learning efficiency relative to the correctly specified case. Indeed, as we show in Appendix A.4,

$$w_{\underline{\theta}}(\mu, \hat{\mu}) < \min\{w_{+\underline{\theta}}(\mu, \hat{\mu}), w_{-\underline{\theta}}(\mu, \hat{\mu})\},$$

which, combined with (9), implies

$$w(\mu, \hat{\mu}) < \min\{w_{+}(\mu, \hat{\mu}), w_{-}(\mu, \hat{\mu})\}. \quad \blacktriangle$$

While the above examples assume that the true signal structure is exogenous, strict inequalities (7)-(8) similar to Example 2 can also be shown to arise under active learning. This highlights a way in which active learning can exacerbate the efficiency loss due to learning biases. This point complements Heidhues, Koszegi, and Strack (2018) (HKS) who show in the context of overconfidence that active learning can lower long-run welfare by leading asymptotic beliefs to be farther away from the truth than under exogenous learning. Our analysis focuses on the inefficiency due to a slower speed of learning rather than incorrect asymptotic beliefs. Moreover, while HKS show that welfare under active learning is lower than *some* exogenous signal benchmark, we show that efficiency can be lower than *all* the corresponding exogenous benchmarks (i.e., both + and −).

3 Markovian Signals

3.1 Setting

This section considers a generalization of FII in which signals are correlated over time. This setting includes several important learning biases that feature misperceptions of the intertemporal correlation of signals (e.g., the gambler's/hot-hand fallacies à la Rabin and Vayanos (2010) and forms of intertemporal correlation neglect).

Specifically, we assume that signals follow an exogenous Markov process. Each signal x_t ($t = 1, \ldots, T$) is drawn from the full-support distribution $\mu_{x\theta} \in \Delta(X)$ in state θ when $x_{t-1} = x$, where we fix some arbitrary initial signal $x_0 \in X$. Call $\mu = (\mu_{x\theta})_{x\in X,\theta\in\Theta}$ the ***true signal structure***. Call $\hat{\mu} = (\hat{\mu}_{x\theta})_{x\in X,\theta\in\Theta}$ the ***perceived signal structure***, where $\hat{\mu}_{x\theta} \in \mathbb{R}^X_{++}$ denotes the perceived signal likelihoods following signal x in state θ. Let $m_\theta \in \Delta(X)$ denote the stationary distribution of signals under the true distributions $(\mu_{x\theta})_{x\in X}$, which is unique by the full-support assumption.

The agent's posterior belief on state θ after signal sequence $x^T = (x_1, \ldots, x_T)$ is given by

$$p(\theta|x^T) = \frac{p_0(\theta)\prod_{t=1}^T \hat{\mu}_{x_{t-1}\theta}(x_t)}{\sum_{\theta'\in\Theta} p_0(\theta')\prod_{t=1}^T \hat{\mu}_{x_{t-1}\theta'}(x_t)}.$$

Welfare $W_T(\mu, \hat{\mu}, A)$ at decision problem A is defined as before by (1)-(2).

3.2 Characterization

As in the previous setting, we restrict attention to the case where agents' biases are small enough that they learn the true state as $T \to \infty$. This is ensured by the following condition:

Assumption 2. For any distinct $\theta, \theta' \in \Theta$, we have $\sum_x m_\theta(x)\mathrm{KL}(\mu_{x\theta}, \hat{\mu}_{x\theta}) < \sum_x m_\theta(x)\mathrm{KL}(\mu_{x\theta}, \hat{\mu}_{x\theta'})$.

This requires that the agent's perceived signal distributions in the true state θ achieve a smaller weighted KL-divergence relative to the true signal distributions than do her perceived signal distributions in any other state θ', where the weights are given by the true stationary distribution m_θ.[6]

For each $\nu \in \Delta(X \times X)$, we let $\overline{\nu} \in \Delta(X)$ denote the marginal distribution on the first coordinate, and $\nu_x \in \Delta(X)$ denote the conditional distribution given that the first coordinate is x (which is specified arbitrarily in case $\overline{\nu}(x) = 0$). We call ν ***shift-invariant*** if it has the same marginal distributions for the two coordinates.

We define the following learning efficiency index, which reduces to the one in FII when $\mu_{x\theta} = \mu_{x'\theta}$ and $\hat{\mu}_{x\theta} = \hat{\mu}_{x'\theta}$ for each θ and x, x' (i.e., the true and perceived signal structures are i.i.d.):

Definition 2. For any true and perceived signal structures μ and $\hat{\mu}$, the ***learning efficiency index*** (under Markovian signals) is defined by $w(\mu, \hat{\mu}) := \min_\theta w_\theta(\mu, \hat{\mu})$, where

$$w_\theta(\mu, \hat{\mu}) := \inf_{\substack{\nu \in \Delta(X \times X) \text{ s.t.} \\ \nu \text{ is shift-invariant}}} \sum_x \overline{\nu}(x) \mathrm{KL}(\nu_x, \mu_{x\theta}) \tag{10}$$

subject to

$$\sum_x \overline{\nu}(x) \mathrm{KL}(\nu_x, \hat{\mu}_{x\overline{\theta}}) = \sum_x \overline{\nu}(x) \mathrm{KL}(\nu_x, \hat{\mu}_{x\underline{\theta}}). \tag{11}$$

We refer to constraint (4) as the ***indistinguishability condition***.

The following result shows that this index characterizes the dynamic welfare ranking under Markovian signals:

Theorem 2. *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 decision problem A, there exists T^* such that $W_T(\mu, \hat{\mu}^1, A) > W_T(\mu, \hat{\mu}^2, A)$ for every $T \geq T^*$.*

The proof of Theorem 2 again shows that the learning efficiency index characterizes the asymptotic rate at which mistake probabilities vanish. For this, we build on large-deviation results for irreducible Markov chains (e.g., Section 3.1 in Dembo and Zeitouni (2010)).

[6]Such weighted KL-divergences also appear in the formulation of Berk-Nash equilibrium in Markovian environments (Esponda and Pouzo, 2021).

A Proofs

A.1 Proof of Theorem 1

Let $L_t^i := \log \frac{p^i(\bar{\theta}|x^t)}{p^i(\underline{\theta}|x^t)}$ denote the posterior log-likelihood ratio of agent i in period t. In the following lemmas, we analyze large deviations of L_t^i. For simplicity, we drop superscript i.

For each $\theta \in \Theta$ and $\# \in \{-, +\}$, construct the Legendre-Fenchel transform of the cumulant functions of perceived signal log-likelihood ratios:

$$I_{\theta,\#}(\beta) := \sup_{\alpha \in \mathbb{R}} \alpha\beta - \log \sum_x \mu_{\#\theta}(x) \left(\frac{\hat{\mu}_{\#\bar{\theta}}(x)}{\hat{\mu}_{\#\underline{\theta}}(x)} \right)^{\alpha}$$

for each $\beta \in \mathbb{R}$. By applying the variational formula (Lemma 6.2.3 (f) in Dupuis and Ellis (2011)), we can write

$$I_{\theta,\#}(\beta) = \inf_{\nu \in \Delta(X)} \mathrm{KL}(\nu, \mu_{\#\theta}) \text{ s.t. } \mathrm{KL}(\nu, \hat{\mu}_{\#\underline{\theta}}) = \mathrm{KL}(\nu, \hat{\mu}_{\#\bar{\theta}}) + \beta \tag{12}$$

for each $\beta \in \mathbb{R}$.

For each $\theta \in \Theta$, we define a function I_θ as follows. First,

$$I_\theta(0) := \inf_{\alpha \in [0,1], \beta_- \geq 0, \beta_+ \leq 0} \alpha I_{\theta,-}(\beta_-) + (1-\alpha) I_{\theta,+}(\beta_+) \text{ s.t. } \alpha\beta_- + (1-\alpha)\beta_+ = 0. \tag{13}$$

Then for $\beta < 0$,

$$I_\theta(\beta) := \inf_{\alpha \in (0,1]} \alpha I_{\theta,-}(\beta/\alpha) + (1-\alpha) I_\theta(0)$$

and for $\beta > 0$,

$$I_\theta(\beta) := \inf_{\alpha \in (0,1]} \alpha I_{\theta,+}(\beta/\alpha) + (1-\alpha) I_\theta(0).$$

Assumption 1 ensures that each $I_{\theta,\#}(\cdot)$ is finite-valued and continuous in a neighborhood of 0. This guarantees that each $I_\theta(\cdot)$ is also finite-valued and continuous in a neighborhood of 0. To see this, first note that by Assumption 1, there exists some $\varepsilon > 0$ such that $I_{\theta,-}(-\varepsilon), I_{\theta,+}(\varepsilon) < +\infty$. Then for each $\beta \in (-\varepsilon, 0)$,

$$I_\theta(\beta) \leq |\beta/\varepsilon| I_{\theta,-}(-\varepsilon) + (1 - |\beta/\varepsilon|) I_\theta(0).$$

Furthermore, because $I_{\theta,-}$ is convex and differentiable at 0,

$$I_\theta(\beta) \geq \inf_{\alpha\in[0,1]} \beta I'_{\theta,-}(0) + \alpha I_{\theta,-}(0) + (1-\alpha)I_\theta(0) = \beta I'_{\theta,-}(0) + I_\theta(0).$$

Thus, we have the following bounds:

$$\beta I'_{\theta,-}(0) + I_\theta(0) \leq I_\theta(\beta) \leq |\beta/\varepsilon| I_{\theta,-}(-\varepsilon) + (1 - |\beta/\varepsilon|)I_\theta(0).$$

Similarly, for $\beta \in (0,\varepsilon)$, we have:

$$\beta I'_{\theta,+}(0) + I_\theta(0) \leq I_\theta(\beta) \leq |\beta/\varepsilon| I_{\theta,+}(\varepsilon) + (1 - |\beta/\varepsilon|)I_\theta(0).$$

Clearly, this implies that I_θ is finite-valued in a neighborhood of 0 and moreover, taking the limit as $\beta \to 0$, we have $I_\theta(\beta) \to I_\theta(0)$.

The following lemma is based on a large-deviation result of Dupuis and Ellis (1992):

Lemma 1. *For any* $\gamma \in \mathbb{R}$*,* $\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\underline{\theta}}[L_t \geq \gamma] = -I_{\underline{\theta}}(0)$ *and* $\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] = -I_{\overline{\theta}}(0)$.

Proof. We consider the case of state $\overline{\theta}$ (the case of state $\underline{\theta}$ is analogous). Consider the modified process $\tilde{L}_t := L_t - \log \frac{\kappa^i}{1-\kappa^i}$. Assumption 1 implies that for each $\# \in \{-,+\}$, there exist $x, x' \in X$ such that $\log \frac{\hat{\mu}_{\#\overline{\theta}}(x)}{\hat{\mu}_{\#\underline{\theta}}(x)} < 0 < \log \frac{\hat{\mu}_{\#\overline{\theta}}(x')}{\hat{\mu}_{\#\underline{\theta}}(x')}$. Thus, the process $(\tilde{L}_t)$, conditional on each θ, satisfies Hypothesis H in Dupuis and Ellis (1992).

Fix any $\gamma \in \mathbb{R}$. Notice that $\mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] = \mathbb{P}_{\overline{\theta}}\left[\tilde{L}_t \leq \gamma - \log \frac{\kappa^i}{1-\kappa^i}\right]$. Thus, for any $\varepsilon > 0$,

$$\mathbb{P}_{\overline{\theta}}[\tilde{L}_t/t \leq -\varepsilon] \leq \mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] \leq \mathbb{P}_{\overline{\theta}}[\tilde{L}_t/t \leq \varepsilon]$$

for all large enough t. Since $I_{\overline{\theta}}$ is finite-valued and continuous in a neighborhood of 0, Theorem 2.1 in Dupuis and Ellis (1992) implies that for all small enough $\varepsilon > 0$,

$$-t \inf_{\beta\leq-\varepsilon} I_{\overline{\theta}}(\beta) + o(t) \leq \log \mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] \leq -t \inf_{\beta\leq\varepsilon} I_{\overline{\theta}}(\beta) + o(t)$$

as $t \to \infty$. Thus, $\log \mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] = -t \inf_{\beta\leq 0} I_{\overline{\theta}}(\beta) + o(t)$ by continuity of $I_{\overline{\theta}}$ at 0.

By standard arguments, $I_{\overline{\theta},-}$ is convex and minimized at $\sum_x \mu_{\overline{\theta},-}(x) \log \frac{\hat{\mu}_{\overline{\theta},-}(x)}{\hat{\mu}_{\underline{\theta},-}(x)}$ (e.g., Lemma 2.2.5 in Dembo and Zeitouni, 2010)), which is strictly positive by Assumption 1. Thus, for each $\beta \leq 0$, $I_{\overline{\theta},-}(\beta) \geq I_{\overline{\theta},-}(0)$, and thus $I_{\overline{\theta}}(\beta) \geq I_{\overline{\theta}}(0)$.

Therefore $\inf_{\beta \leq 0} I_{\overline{\theta}}(\beta) = I_{\overline{\theta}}(0)$, which establishes the claim. □

The next lemma represents each $I_\theta(0)$ in terms of KL-divergence:

Lemma 2. *For each $\theta \in \Theta$, $I_\theta(0) = w_\theta(\mu, \hat{\mu})$.*

Proof. By (12), (13) can be written as

$$I_\theta(0) = \inf_{\alpha \in [0,1], \nu_-, \nu_+ \in \Delta(X)} \alpha \mathrm{KL}(\nu_-, \mu_{-\theta}) + (1-\alpha)\mathrm{KL}(\nu_+, \mu_{+\theta})$$

subject to

$$\alpha \mathrm{KL}(\nu_-, \hat{\mu}_{-\underline{\theta}}) + (1-\alpha)\alpha \mathrm{KL}(\nu_+, \hat{\mu}_{+\underline{\theta}}) = \alpha \mathrm{KL}(\nu_-, \hat{\mu}_{-\overline{\theta}}) + (1-\alpha)\alpha \mathrm{KL}(\nu_+, \hat{\mu}_{+\overline{\theta}})$$

$$\mathrm{KL}(\nu_-, \hat{\mu}_{-\underline{\theta}}) \geq \mathrm{KL}(\nu_-, \hat{\mu}_{-\overline{\theta}})$$

$$\mathrm{KL}(\nu_+, \hat{\mu}_{+\underline{\theta}}) \leq \mathrm{KL}(\nu_+, \hat{\mu}_{+\overline{\theta}})$$

by replacing each variable $\beta_\#$ with $\mathrm{KL}(\nu_\#, \hat{\mu}_{\#\underline{\theta}}) - \mathrm{KL}(\nu_\#, \hat{\mu}_{\#\overline{\theta}})$. Moreover, inf can be replaced with min since $I_\theta(0)$ is finite-valued and the domain is closed. □

To complete the proof of Theorem 1, take any decision problem A. It is without loss to assume that A contains no dominated acts (as such acts are never chosen). Thus, given binary states, there is a unique optimal act $a^\theta \in \operatorname{argmax}_{a \in A} a_\theta$ in each state θ. Then

$$\begin{aligned}
& \limsup_{T \to \infty} \frac{1}{T} \log \left(\sum_\theta p_0(\theta) a^\theta_\theta - W_T(\mu, \hat{\mu}^i, \kappa^i, A) \right) \\
= \;& \limsup_{T \to \infty} \frac{1}{T} \log \left(\sum_\theta p_0(\theta) \sum_{a \neq a^\theta} (a^\theta_\theta - a_\theta) \mathbb{P}_\theta[a^{*i}(x^T) = a] \right) \\
= \;& \limsup_{T \to \infty} \frac{1}{T} \log \left(\sum_\theta p_0(\theta) \mathbb{P}_\theta[a^{*i}(x^T) \neq a^\theta] \right) \\
= \;& \limsup_{T \to \infty} \max_\theta \frac{1}{T} \log \left(\mathbb{P}_\theta[a^{*i}(x^T) \neq a^\theta] \right),
\end{aligned}$$

and the analogous claim holds for lim inf. Lemmas 1-2 show that

$$\lim_{T \to \infty} \frac{1}{T} \log \left(\mathbb{P}_\theta[a^{*i}(x^T) \neq a^\theta] \right) = -w_\theta(\mu, \hat{\mu}^i)$$

for each θ. Therefore,

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T} \log \left(\sum_{\theta} p_0(\theta) a_\theta^\theta - W_T(\mu, \hat{\mu}^1, \kappa^1, A) \right) \\
&= -w(\mu, \hat{\mu}^1) \\
&< -w(\mu, \hat{\mu}^2) = \lim_{T\to\infty} \frac{1}{T} \log \left(\sum_{\theta} p_0(\theta) a_\theta^\theta - W_T(\mu, \hat{\mu}^2, \kappa^2, A) \right).
\end{aligned}$$

This implies the desired result. $\square$

A.2 On interior solutions to (3) under binary signals

Suppose $X = \{\overline{x}, \underline{x}\}$. We provide a necessary condition for (3) to have an interior solution $\alpha \in (0, 1)$, which will be used in the analysis of Examples 1–2. For each $\# \in \{-, +\}$, let $\phi_\# := \nu_\#(\overline{x})$. Then the efficiency index at each state θ takes the form

$$\min_{\alpha,\phi_-,\phi_+\in[0,1]} \alpha \left(\phi_- \log \frac{\phi_-}{\mu_{-\theta}(\overline{x})} + (1-\phi_-) \log \frac{1-\phi_-}{\mu_{-\theta}(\underline{x})} \right) + (1-\alpha) \left(\phi_+ \log \frac{\phi_+}{\mu_{+\theta}(\overline{x})} + (1-\phi_+) \log \frac{1-\phi_+}{\mu_{+\theta}(\underline{x})} \right)$$

$$\text{s.t. } \alpha \left(\phi_- \log \frac{\hat{\mu}_{-\overline{\theta}}(\overline{x})}{\hat{\mu}_{-\underline{\theta}}(\overline{x})} + (1-\phi_-) \log \frac{\hat{\mu}_{-\overline{\theta}}(\underline{x})}{\hat{\mu}_{-\underline{\theta}}(\underline{x})} \right) + (1-\alpha) \left(\phi_+ \log \frac{\hat{\mu}_{+\overline{\theta}}(\overline{x})}{\hat{\mu}_{+\underline{\theta}}(\overline{x})} + (1-\phi_+) \log \frac{\hat{\mu}_{+\overline{\theta}}(\underline{x})}{\hat{\mu}_{+\underline{\theta}}(\underline{x})} \right) = 0,$$

$$\phi_- \log \frac{\hat{\mu}_{-\overline{\theta}}(\overline{x})}{\hat{\mu}_{-\underline{\theta}}(\overline{x})} + (1-\phi_-) \log \frac{\hat{\mu}_{-\overline{\theta}}(\underline{x})}{\hat{\mu}_{-\underline{\theta}}(\underline{x})} \geq 0,$$

$$\phi_+ \log \frac{\hat{\mu}_{+\overline{\theta}}(\overline{x})}{\hat{\mu}_{+\underline{\theta}}(\overline{x})} + (1-\phi_+) \log \frac{\hat{\mu}_{+\overline{\theta}}(\underline{x})}{\hat{\mu}_{+\underline{\theta}}(\underline{x})} \leq 0.$$

Suppose we have a solution at which $\alpha \in (0, 1)$ and the latter two inequality constraints are slack. Then the first-order conditions with respect to ϕ_-, ϕ_+ are

$$\alpha \left(\log \frac{\phi_-}{\mu_{-\theta}(\overline{x})} - \log \frac{1-\phi_-}{\mu_{-\theta}(\underline{x})} \right) + \alpha\lambda \left(\log \frac{\hat{\mu}_{-\overline{\theta}}(\overline{x})}{\hat{\mu}_{-\underline{\theta}}(\overline{x})} - \log \frac{\hat{\mu}_{-\overline{\theta}}(\underline{x})}{\hat{\mu}_{-\underline{\theta}}(\underline{x})} \right) = 0,$$

$$(1-\alpha) \left(\log \frac{\phi_+}{\mu_{+\theta}(\overline{x})} - \log \frac{1-\phi_+}{\mu_{+\theta}(\underline{x})} \right) + (1-\alpha)\lambda \left(\log \frac{\hat{\mu}_{+\overline{\theta}}(\overline{x})}{\hat{\mu}_{+\underline{\theta}}(\overline{x})} - \log \frac{\hat{\mu}_{+\overline{\theta}}(\underline{x})}{\hat{\mu}_{+\underline{\theta}}(\underline{x})} \right) = 0,$$

where λ denotes the Lagrange multiplier on the first constraint. If $\lambda = 0$ then the first-order conditions imply $\phi_\# = \mu_{\#\theta}(\overline{x})$ for each $\#$, which cannot satisfy the first constraint (by Assumption 1). Thus, combining the first-order conditions yields

$$\frac{\log \frac{\phi_- \mu_{-\theta}(\underline{x})}{(1-\phi_-)\mu_{-\theta}(\overline{x})}}{\log \frac{\phi_+ \mu_{+\theta}(\underline{x})}{(1-\phi_+)\mu_{+\theta}(\overline{x})}} = \frac{\log \frac{\hat{\mu}_{-\overline{\theta}}(\overline{x})\hat{\mu}_{-\underline{\theta}}(\underline{x})}{\hat{\mu}_{-\underline{\theta}}(\overline{x})\hat{\mu}_{-\overline{\theta}}(\underline{x})}}{\log \frac{\hat{\mu}_{+\overline{\theta}}(\overline{x})\hat{\mu}_{+\underline{\theta}}(\underline{x})}{\hat{\mu}_{+\underline{\theta}}(\overline{x})\hat{\mu}_{+\overline{\theta}}(\underline{x})}}. \tag{14}$$

A.3 Details for Example 1

Let $\phi_\# := \nu_\#(\overline{x})$ for each $\# \in \{-, +\}$. First, we consider the two boundary cases. For the choice of $\alpha = 1$, the first constraint $\phi_- \log \frac{\hat{\mu}_{-\overline{\theta}}(\overline{x})}{\hat{\mu}_{-\underline{\theta}}(\overline{x})} + (1-\phi_-) \log \frac{\hat{\mu}_{-\overline{\theta}}(\underline{x})}{\hat{\mu}_{-\underline{\theta}}(\underline{x})} = 0$ pins down

$$\phi_- = \frac{\log \left(\frac{\gamma}{1-\gamma}\right)^c}{\log \left(\frac{\gamma}{1-\gamma}\right)^{c+d}} > 1/2.$$

For the choice of $\alpha = 0$, the first constraint $\phi_+ \log \frac{\hat{\mu}_{+\overline{\theta}}(\overline{x})}{\hat{\mu}_{+\underline{\theta}}(\overline{x})} + (1-\phi_+) \log \frac{\hat{\mu}_{+\overline{\theta}}(\underline{x})}{\hat{\mu}_{+\underline{\theta}}(\underline{x})} = 0$ pins down

$$\phi_+ = \frac{\log \left(\frac{\gamma}{1-\gamma}\right)^d}{\log \left(\frac{\gamma}{1-\gamma}\right)^{c+d}} < 1/2.$$

Since $\phi_- + \phi_+ = 1$, we have $\phi_- \log \phi_- + (1-\phi_-) \log(1-\phi_-) = \phi_+ \log \phi_+ + (1-\phi_+) \log(1-\phi_+)$. Thus, by letting $\mu_\theta(\cdot) := \mu_{-\theta}(\cdot) = \mu_{+\theta}(\cdot)$,

$$\begin{aligned} & \phi_- \log \frac{\phi_-}{\mu_\theta(\overline{x})} + (1-\phi_-) \log \frac{1-\phi_-}{\mu_\theta(\underline{x})} < \phi_+ \log \frac{\phi_+}{\mu_\theta(\overline{x})} + (1-\phi_+) \log \frac{1-\phi_+}{\mu_\theta(\underline{x})} \\ \iff \quad & (\phi_- - \phi_+)(\log \mu_\theta(\underline{x}) - \log \mu_\theta(\overline{x})) < 0, \end{aligned}$$

which is true iff $\theta = \overline{\theta}$. That is, $\alpha = 1$ achieves lower objective value than $\alpha = 0$ under $\overline{\theta}$, while $\alpha = 0$ archives lower objective value than $\alpha = 1$ under $\underline{\theta}$. That is,

$$w_-(\overline{\theta}, \mu, \hat{\mu}) < w_+(\overline{\theta}, \mu, \hat{\mu}), \quad w_+(\underline{\theta}, \mu, \hat{\mu}) < w_-(\underline{\theta}, \mu, \hat{\mu}).$$

Now consider the choice of $\alpha \in (0, 1)$. We consider the case in which the two inequality constraints are non-binding, as otherwise the same objective value can be achieved by setting $\alpha = 0$ or 1. Then the first-order condition (14) implies $\phi_- = \phi_+$. However, the first inequality constraint, $\phi_- \log \frac{\hat{\mu}_{-\overline{\theta}}(\overline{x})}{\hat{\mu}_{-\underline{\theta}}(\overline{x})} + (1-\phi_-) \log \frac{\hat{\mu}_{-\overline{\theta}}(\underline{x})}{\hat{\mu}_{-\underline{\theta}}(\underline{x})} \geq 0$, implies $\phi_- > \frac{1}{2}$, while the second one, $\phi_+ \log \frac{\hat{\mu}_{+\overline{\theta}}(\overline{x})}{\hat{\mu}_{+\underline{\theta}}(\overline{x})} + (1-\phi_+) \log \frac{\hat{\mu}_{+\overline{\theta}}(\underline{x})}{\hat{\mu}_{+\underline{\theta}}(\underline{x})} \leq 0$, implies $\phi_+ < \frac{1}{2}$, a contradiction. Thus, the choice of $\alpha \in (0,1)$ cannot be optimal. Therefore,

$$w(\overline{\theta},\mu,\hat{\mu}) = w_-(\overline{\theta},\mu,\hat{\mu}), \quad w(\underline{\theta},\mu,\hat{\mu}) = w_+(\underline{\theta},\mu,\hat{\mu}).$$

A.4 Details for Example 2

Consider the true state $\underline{\theta}$. Set $\alpha = 0.5$ in the minimization problem (3). For any small $\varepsilon \geq 0$, choosing

$$\nu_+(\overline{x}) = \frac{1}{2} - \frac{\varepsilon}{c}, \quad \nu_-(\overline{x}) = \frac{1}{2} + \frac{\varepsilon}{d}$$

satisfies the constraints (4)-(6). We note that choosing $\varepsilon = 0$, or $\nu_+(\overline{x}) = \nu_-(\overline{x}) = 1/2$, yields the same objective value that can be obtained under the boundary cases $\alpha \in \{0,1\}$. Letting $\mu_\theta(\cdot) := \mu_{-\theta}(\cdot) = \mu_{+\theta}(\cdot)$, observe that each $\mathrm{KL}(\nu_\#, \mu_{\underline{\theta}})$, $\# \in \{-,+\}$, has the same positive slope in $\nu_\#(\overline{x})$ at $\nu_\#(\overline{x}) = 1/2$ since $\mu_{\underline{\theta}}(\overline{x}) = 1-\gamma < 1/2$. Therefore, setting $\varepsilon > 0$ small enough that $\nu_+(\overline{x}) + \nu_-(\overline{x}) < 1$ strictly lowers the objective.

A.5 Proof of Theorem 2

Let $\nu_t \in \Delta(X \times X)$ denote the empirical frequency of signal pairs in period t, i.e., $\nu_t(x,x') = \frac{1}{t}\sum_{\tau=1}^t \mathbf{1}_{\{x_{\tau-1}=x, x_\tau = x'\}}$ for each $(x,x') \in X \times X$. Below we provide preliminary lemmas analyzing large deviations of $L_t^i := \log \frac{p^i(\overline{\theta}|x^t)}{p^i(\underline{\theta}|x^t)}$, i.e., the posterior log-likelihood ratio of agent i in period t. For simplicity, we drop superscript i.

Let

$$I_{\overline{\theta}}(\alpha) := \inf_{\substack{\nu \in \Delta(X\times X) \text{ s.t.} \\ \nu \text{ is shift-invariant}}} \sum_x \overline{\nu}(x)\mathrm{KL}(\nu_x, \mu_{x\overline{\theta}}) \tag{15}$$

$$\text{s.t.} \sum_x \overline{\nu}(x)\mathrm{KL}(\nu_x, \hat{\mu}_{x\underline{\theta}}) \leq \sum_x \overline{\nu}(x)\mathrm{KL}(\nu_x, \hat{\mu}_{x\overline{\theta}}) + \alpha$$

and

$$I_{\underline{\theta}}(\alpha) := \inf_{\substack{\nu \in \Delta(X\times X) \text{ s.t.} \\ \nu \text{ is shift-invariant}}} \sum_x \overline{\nu}(x)\mathrm{KL}(\nu_x, \mu_{x\underline{\theta}}) \tag{16}$$

$$\text{s.t.} \sum_x \overline{\nu}(x)\mathrm{KL}(\nu_x, \hat{\mu}_{x\underline{\theta}}) \geq \sum_x \overline{\nu}(x)\mathrm{KL}(\nu_x, \hat{\mu}_{x\overline{\theta}}) + \alpha$$

for each $\alpha \in \mathbb{R}$. Let $\alpha_\theta := \sum_x m_\theta(x) \sum_{x'} \mu_{x\theta}(x') \log \frac{\hat{\mu}_{x\overline{\theta}}(x')}{\hat{\mu}_{x\underline{\theta}}(x')}$ for each θ. By Assumption 2, $\alpha_{\underline{\theta}} < 0 < \alpha_{\overline{\theta}}$.

Lemma 3. *Take any $\alpha \in (\alpha_{\underline{\theta}}, \alpha_{\overline{\theta}})$.*

1. *In (15) and (16), the minimum is achieved and the constraints bind.*

2. *$I_{\underline{\theta}}$ and $I_{\overline{\theta}}$ are continuous and positive at α.*

Proof. The proof follows analogous arguments as the proof of Lemma 3 in FII, using Assumption 2 and the fact that the objective function in (15) and (16) is convex over shift-invariant ν (Section 3.1.3 in Dembo and Zeitouni (2010)). $\square$

Lemma 4. *For any $\gamma \in \mathbb{R}$, $\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\underline{\theta}}[L_t \geq \gamma] = -I_{\underline{\theta}}(0)$ and $\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] = -I_{\overline{\theta}}(0)$.*

Proof. We consider the case of $\overline{\theta}$ (the case of $\underline{\theta}$ is analogous). Fix any $\gamma \in \mathbb{R}$. Observe that

$$\begin{aligned}\mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] &= \mathbb{P}_{\overline{\theta}}[\log \frac{p_0(\overline{\theta})}{p_0(\underline{\theta})} + t \sum_{x,x'} \nu_t(x,x') \log \frac{\hat{\mu}_{x\overline{\theta}}(x')}{\hat{\mu}_{x\underline{\theta}}(x')} \leq \gamma] \\ &= \mathbb{P}_{\overline{\theta}}[\sum_x \overline{\nu}_t(x) \mathrm{KL}(\nu_{tx}, \hat{\mu}_{x\underline{\theta}}) \leq \sum_x \overline{\nu}_t(x) \mathrm{KL}(\nu_{tx}, \hat{\mu}_{x\overline{\theta}}) + \frac{\gamma - \log \frac{p_0(\overline{\theta})}{p_0(\underline{\theta})}}{t}].\end{aligned}$$

Thus, for any $\varepsilon > 0$,

$$\begin{aligned}&\mathbb{P}_{\overline{\theta}}[\sum_x \overline{\nu}_t(x) \mathrm{KL}(\nu_{tx}, \hat{\mu}_{x\underline{\theta}}) \leq \sum_x \overline{\nu}_t(x) \mathrm{KL}(\nu_{tx}, \hat{\mu}_{x\overline{\theta}}) - \varepsilon] \\ \leq\ & \mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] \leq \mathbb{P}_{\overline{\theta}}[\sum_x \overline{\nu}_t(x) \mathrm{KL}(\nu_{tx}, \hat{\mu}_{x\underline{\theta}}) \leq \sum_x \overline{\nu}_t(x) \mathrm{KL}(\nu_{tx}, \hat{\mu}_{x\overline{\theta}}) + \varepsilon]\end{aligned}$$

for all large t. Thus, for any $\varepsilon > 0$ such that $I_{\overline{\theta}}(\cdot)$ is continuous on $(-\varepsilon, \varepsilon)$, Sanov's theorem for pair empirical measures under Markov chains (Theorem 3.1.13 in Dembo and Zeitouni (2010)) implies that

$$-tI_{\overline{\theta}}(-\varepsilon) + o(t) \leq \log \mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] \leq -tI_{\overline{\theta}}(\varepsilon) + o(t)$$

as $t \to \infty$. Since this holds for all small $\varepsilon > 0$ and $I_{\overline{\theta}}$ is continuous around 0 (second part of Lemma 3), we have

$$\log \mathbb{P}_{\overline{\theta}}[L_t \leq \gamma] = -tI_{\overline{\theta}}(0) + o(t),$$

as $t \to \infty$, which establishes the desired claim. $\square$

Given the above lemmas, the remaining proof follows the same arguments as that of Theorem 1. $\square$

References

BENJAMIN, D. J. (2019): "Errors in probabilistic reasoning and judgment biases," *Handbook of Behavioral Economics-Foundations and Applications 2*, pp. 69–186.

CHARNESS, G., AND C. DAVE (2017): "Confirmation bias with motivated beliefs," *Games and Economic Behavior*, 104, 1–23.

DEMBO, A., AND O. ZEITOUNI (2010): *Large Deviations: Techniques and Applications*. Springer.

DUPUIS, P., AND R. S. ELLIS (1992): "Large deviations for Markov processes with discontinuous statistics, II: random walks," *Probability Theory and Related Fields*, 91(2), 153–194.

——— (2011): *A Weak Convergence Approach to the Theory of Large Deviations*, vol. 902. John Wiley & Sons.

ESPONDA, I., AND D. POUZO (2021): "Equilibrium in misspecified Markov decision processes," *Theoretical Economics*, 16(2), 717–757.

FRICK, M., R. IIJIMA, AND Y. ISHII (2022a): "Belief Convergence under Misspecified Learning: A Martingale Approach," *Review of Economic Studies*.

——— (2022b): "Welfare Comparisons for Biased Learning," working paper.

HEIDHUES, P., B. KOSZEGI, AND P. STRACK (2018): "Unrealistic expectations and misguided learning," *Econometrica*, 86(4), 1159–1214.

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

RABIN, M., AND J. L. SCHRAG (1999): "First impressions matter: A model of confirmatory bias," *Quarterly Journal of Economics*, 114(1), 37–82.

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.