Learning From Reviews: Endogenous Histories and Information Bounds

Chae-Yeon Xon
Yonsei University

September 2026

Abstract

Acemoglu, Makhdoumi, Malekian, and Ozdaglar (2022) characterize the learning rate from a terminal review summary by a Kullback–Leibler divergence that evaluates the true state at its limiting belief and the alternative state at the opposite limiting belief. A terminal summary is a function of the review history generated under the same summary-based policy. Data processing yields a finite-horizon information bound, and any constant almost-sure learning exponent is bounded by the information in that history. The history's asymptotic rate evaluates both states at the same limiting belief. Uniform primitives satisfying the paper's assumptions make the stated formulas equal 0.2176 and 0.2662, while the corresponding information bounds are 0.0288 and 0.1181. The example has positive selection, so it also contradicts the strict speed comparison in Proposition 2(2). Terminal-event likelihoods admit an exact finite-horizon relative-entropy representation. A model-specific large-deviation contraction of this representation supplies the corresponding asymptotic rate problem.

JEL Codes: C11, C44, D83, L15

Keywords: Bayesian learning, endogenous selection, information bounds, online reviews, summary statistics

Yonsei University. eunixon@yonsei.ac.kr. Early stage draft; please do not cite or circulate without author's written permission. All errors are my own.

1. Introduction

Acemoglu et al. (2022) study learning when purchase decisions select the customers who produce reviews. With a full review history, each customer observes the belief attached to every preceding action. With a summary statistic, the customer observes terminal review frequencies and integrates over the histories that generated them. Theorem 4 assigns an exact exponential rate to this second learning problem. Under quality zero, its divergence compares the review distribution at $(Q, q) = (0, 0)$ with the distribution at $(Q, q) = (1, 1)$. The reverse comparison gives the rate under quality one.

The likelihood of a terminal summary sums over full action histories. Each history in this sum carries the sequence of posteriors generated by its intermediate summaries. Under the counterfactual quality, a history that reaches a quality-zero-typical terminal summary can follow posteriors near zero. The proof of Theorem 4 applies quality-one typical convergence, $q \to 1$, to the counterfactual likelihoods of histories inside that terminal-summary fiber.

This paper derives an information bound that keeps the two counterfactual laws on the same realized history. Let H_n denote the category sequence generated when customers observe the summary statistic, and let S_n be its terminal summary. The chain rule gives the information in H_n as a sum of conditional divergences evaluated at the posterior generated by each realized prefix. Under quality zero, complete learning makes its average converge to $D(\pi_0(\cdot, 0)\|\pi_1(\cdot, 0))$. Since S_n is a function of H_n, data processing bounds the terminal-summary rate by this same-belief divergence. The quality-one bound is analogous.

A binary-review example makes both inequalities strict relative to the rates in Theorem 4. Take independent uniform ex ante and ex post valuations, price 1/5, and two review thresholds. The resulting top-review probabilities are

$$\pi_0(q) = \frac{32/25 + (6/5)q - (1/2)q^2}{8}, \qquad \pi_1(q) = \pi_0(q) + \frac{4/5 + q}{8}. \tag{1}$$

The partition has positive selection and strict separation. Theorem 4's quality-zero divergence is 0.2176, while the information bound is 0.0288. The quality-one figures are 0.2662 and 0.1181. These inequalities contradict both rate formulas. They also contradict Proposition 2(2), which gives summary-statistic learning a strictly larger rate than full-history learning under positive selection. The full-history rate equals the same-belief information bound.

The correction changes the form of the rate calculation. For each finite horizon, the probability of reaching a terminal-summary set has a variational representation as the minimum relative entropy of a controlled history law that reaches that set. The transition cost at every date uses the posterior attached to the same realized summary path. Any large-deviation limit is therefore a path-space contraction. A constant frozen-belief divergence arises when the minimizing counterfactual histories share that belief limit; the assumptions of Theorem 4 allow lower-cost histories with the true-state belief limit.

Theorem 3's complete-learning conclusion remains in place and supplies the convergence used in the information bound. Theorem 2's full-history rate remains in place and coincides with the bound. The findings concern the summary-statistic speed formula, the positive-selection comparison based on it, and subsequent rate calculations that use the same cross-boundary substitution. Empirical work has already linked review provision to learning speed and used the framework of Acemoglu et al. (2022) to organize the role of selection in review markets; see Fradkin and Holtz (2023).

The analysis begins by mapping the published rate argument onto the two likelihoods defined along a realized history. Information bounds then show what a terminal summary can retain. A primitive example connects those bounds to Proposition 2, and a finite-horizon variational representation makes the underlying path problem explicit. A recursive evaluation illustrates the same mechanism numerically before the final section records the scope of the correction.

2. The published rate argument

Let $Q \in \{0,1\}$ denote quality, let $q \in [0,1]$ denote a customer's posterior probability of $Q = 1$, and let $\pi_Q(i,q)$ be the probability of category i conditional on an action in the reported set T. A summary system reports the empirical frequencies of a partition $T_1, \dots, T_m$ after τ reported actions. Write q_τ for the posterior computed from this terminal summary. The paper normalizes the common prior to 1/2.

Under Assumptions 1 and 2 and strict separation, Theorem 3 of Acemoglu et al. (2022) gives complete learning. Theorem 4 then states

$$\lim_{\tau\to\infty} \frac{1}{\tau} \log q_\tau = -D_{\mathrm{KL}}(\pi_0(\cdot,0) \,\|\, \pi_1(\cdot,1)) \qquad \text{under } Q = 0, \tag{2}$$

$$\lim_{\tau\to\infty} \frac{1}{\tau} \log(1 - q_\tau) = -D_{\mathrm{KL}}(\pi_1(\cdot,1) \,\|\, \pi_0(\cdot,0)) \qquad \text{under } Q = 1. \tag{3}$$

The proof appears on pages 2892–2894. For a realized vector of terminal category counts, it lets $\mathcal{H}_\tau$ collect every full history consistent with those counts. Equation (A-27) bounds the quality-one likelihood of the histories in $\mathcal{H}_\tau$ with transition probabilities evaluated over $q \geq 1 - \varepsilon$. The convergence $q \to 1$ used to introduce that range holds along quality-one-typical histories. A fiber $\mathcal{H}_\tau$ chosen by quality-zero-typical counts also contains quality-one histories following the posterior sequence generated by those same counts. Lemma A-1 assigns one posterior $q(k,\tau)$ to each summary state and evaluates both quality likelihoods at that posterior.

The distinction can be written directly at the history level. For a prefix h_{t-1}, let $s_{t-1} =$

$g_{t-1}(h_{t-1})$ be its displayed summary. The two conditional laws used in Bayes' rule are

$$\mathbb{P}_Q(Y_t = i \mid H_{t-1} = h_{t-1}) = \pi_Q(i, q(s_{t-1}))\,, \qquad Q \in \{0,1\}. \tag{4}$$

The posterior in (4) belongs to the observed prefix. Switching the quality changes π_Q while keeping that prefix and its posterior fixed.

3. An information bound for terminal summaries

For each horizon n, let $P_Q^{H_n}$ be the law of the category history $H_n = (Y_1, \dots, Y_n)$ generated by the summary system under quality Q, and let $P_Q^{S_n}$ be the law of $S_n = g_n(H_n)$. The history H_n here records the actions underlying the summary-based policy. It serves as the information envelope for S_n.

Define

$$d_0 = D_{\mathrm{KL}}(\pi_0(\cdot, 0) \,\|\, \pi_1(\cdot, 0))\,, \qquad d_1 = D_{\mathrm{KL}}(\pi_1(\cdot, 1) \,\|\, \pi_0(\cdot, 1))\,. \tag{5}$$

These are also the full-history exponents in Theorem 2 for the corresponding reported categories.

THEOREM 1 (Finite-horizon information and rate bounds). *Suppose the category set is finite, every $\pi_Q(i, q)$ is continuous and positive on $q \in [0,1]$, and complete learning holds. Let $\eta \in (0,1)$ be the prior probability of $Q = 1$. At every horizon,*

$$\mathbb{E}_0\left[\log \frac{1-q_n}{q_n}\right] = \log \frac{1-\eta}{\eta} + D_{\mathrm{KL}}(P_0^{S_n} \| P_1^{S_n}), \tag{6}$$

$$\mathbb{E}_1\left[\log \frac{q_n}{1-q_n}\right] = \log \frac{\eta}{1-\eta} + D_{\mathrm{KL}}(P_1^{S_n} \| P_0^{S_n}), \tag{7}$$

and

$$D_{\mathrm{KL}}(P_0^{S_n} \| P_1^{S_n}) \le \sum_{t=1}^{n} \mathbb{E}_0[D_{\mathrm{KL}}(\pi_0(\cdot, q_{t-1}) \,\|\, \pi_1(\cdot, q_{t-1}))]\,, \tag{8}$$

$$D_{\mathrm{KL}}(P_1^{S_n} \| P_0^{S_n}) \le \sum_{t=1}^{n} \mathbb{E}_1[D_{\mathrm{KL}}(\pi_1(\cdot, q_{t-1}) \,\|\, \pi_0(\cdot, q_{t-1}))]\,. \tag{9}$$

The two right sides divided by n converge to d_0 and d_1. If $C_{0,n}$ is any sequence of terminal-summary events with $P_0^{S_n}(C_{0,n}) \to 1$, and $C_{1,n}$ is defined analogously, then

$$\limsup_n -\frac{1}{n} \log P_1^{S_n}(C_{0,n}) \le d_0, \qquad \limsup_n -\frac{1}{n} \log P_0^{S_n}(C_{1,n}) \le d_1. \tag{10}$$

Consequently, if the posterior log losses converge almost surely to constants

$$R_0 = -\lim_n \frac{1}{n}\log q_n \quad (Q=0), \qquad R_1 = -\lim_n \frac{1}{n}\log(1-q_n) \quad (Q=1), \tag{11}$$

then

$$R_0 \le d_0, \qquad R_1 \le d_1. \tag{12}$$

PROOF. Bayes' rule gives (6)–(7). The chain rule for relative entropy and (4) give

$$D_{\mathrm{KL}}(P_0^{H_n} \| P_1^{H_n}) = \sum_{t=1}^{n} \mathbb{E}_0[D_{\mathrm{KL}}(\pi_0(\cdot, q_{t-1}) \,\|\, \pi_1(\cdot, q_{t-1}))], \tag{13}$$

with the analogous identity after reversing the states. Data processing for $S_n = g_n(H_n)$ proves (8)–(9). Under $Q = 0$, complete learning gives $q_t \to 0$ almost surely. Positivity and continuity over the compact belief interval bound every log likelihood increment. Dominated convergence followed by Cesàro averaging yields

$$\lim_{n\to\infty} \frac{1}{n} D_{\mathrm{KL}}(P_0^{H_n} \| P_1^{H_n}) = d_0. \tag{14}$$

The state-reversed argument yields the limit d_1.

Set $a_n = P_0^{S_n}(C_{0,n})$ and $b_n = P_1^{S_n}(C_{0,n})$. Applying data processing to the indicator of $C_{0,n}$ gives

$$D_{\mathrm{KL}}(P_0^{S_n} \| P_1^{S_n}) \ge a_n \log\frac{a_n}{b_n} + (1-a_n)\log\frac{1-a_n}{1-b_n} \ge -a_n \log b_n - h(a_n), \tag{15}$$

where h is binary entropy. Since $a_n \to 1$ and $h(a_n) \le \log 2$, equations (8), (14), and (15) prove the first bound in (10); reversing the states proves the second.

For the last claim, set $X_n = n^{-1}\log(dP_0^{S_n}/dP_1^{S_n})(S_n)$. A history likelihood ratio lies between e^{-Mn} and e^{Mn} for some finite M. A summary likelihood ratio is a conditional average of history likelihood ratios in its fiber, so $|X_n| \le M$. Under $Q = 0$, an almost-sure constant posterior exponent equals the almost-sure limit of X_n because the prior log odds contribute $O(1/n)$. Bounded convergence, (8), and (14) give $R_0 \le d_0$. The quality-one argument gives $R_1 \le d_1$. □

The event bound also displays the counterfactual histories suppressed in (A-27). Let $L_n(H_n) = \log\{dP_0^{H_n}/dP_1^{H_n}\}(H_n)$. Bounded martingale differences and complete learning give $L_n/n \to d_0$ under P_0. For any quality-zero-typical terminal-summary events $C_{0,n}$, let

$$B_n(\varepsilon) = \left\{|L_n/n - d_0| \le \varepsilon,\ S_n \in C_{0,n}\right\}, \qquad A_n = g_n(B_n) \subseteq C_{0,n}. \tag{16}$$

Then $P_0(B_n) \to 1$ and

$$P_1^{S_n}(A_n) \geq P_1^{H_n}(B_n) = \mathbb{E}_0[e^{-L_n}\mathbf{1}_{B_n}] \geq e^{-n(d_0+\varepsilon)}P_0^{H_n}(B_n). \tag{17}$$

Thus a subset of the quality-zero-typical terminal summaries is generated under quality one along the same low-belief histories at exponent at most d_0. This is the history class omitted when the quality-one typical-belief restriction is applied throughout the terminal-summary fiber.

4. A primitive counterexample

Set $K = 1$ and take

$$\begin{gathered} \theta \sim U[-1,1], \quad \zeta \sim U[-2,2], \quad \theta \perp \zeta, \\ p = \frac{1}{5}, \quad \lambda_{-1} = \frac{2}{5}, \quad \lambda_1 = \frac{4}{5}. \end{gathered} \tag{18}$$

Use the partition $T = A$, $T_2 = \{+1\}$, and $T_1 = A \setminus \{+1\}$. Thus the displayed summary is the fraction of top reviews among all customer actions. Footnote 11 permits the no-purchase action to enter a reported group, and Example 2 uses the same all-action binary partition $T = \{-1, 0, 1\} \cup \{N\}$, $T_1 = \{-1, 0\} \cup \{N\}$, and $T_2 = \{1\}$. Since $T = A$, review time and calendar time coincide.

A customer purchases when $\theta \geq p - q$. Conditional on purchase, she leaves the top review when $\zeta \geq \lambda_1 + p - \theta - Q = 1 - \theta - Q$. These cutoffs remain in the support of ζ over the purchase region. Integration gives

$$\pi_0(q) = \frac{1}{8}\int_{1/5-q}^{1}(1+\theta)\,d\theta = \frac{32/25 + (6/5)q - (1/2)q^2}{8}, \tag{19}$$

$$\pi_1(q) = \frac{1}{8}\int_{1/5-q}^{1}(2+\theta)\,d\theta = \pi_0(q) + \frac{4/5+q}{8}. \tag{20}$$

Both probabilities increase on $[0, 1]$, and $\pi_1(q) > \pi_0(q)$ throughout this interval. The partition therefore satisfies the positive-selection definition. Its endpoint probabilities are

$$\pi_0(0) = \frac{4}{25}, \qquad \pi_1(0) = \frac{13}{50}, \qquad \pi_0(1) = \frac{99}{400}, \qquad \pi_1(1) = \frac{189}{400}. \tag{21}$$

Strict separation follows from $13/50 - 99/400 = 1/80$. Assumption 1's two support inequalities become $14/5 > 4/5$ and $-1/5 < 2/5$. Assumption 2 becomes $4/5 > 0$. The continuous uniform distributions, ordered cutoffs, independence, and interior category probabilities supply the remaining premises.

PROPOSITION 1 (Failure of the cross-boundary rates). *For the primitives in* (18)*, both rates* *in Theorem 4 of Acemoglu et al. (2022) exceed the bounds in Theorem 1. Proposition 2(2) also fails.*

PROOF. For Bernoulli probabilities, write $K(p,q) = D_{\mathrm{KL}}(\mathrm{Ber}(p) \| \mathrm{Ber}(q))$. Its derivative in the second argument is

$$\frac{\partial K(p,q)}{\partial q} = \frac{q-p}{q(1-q)}. \tag{22}$$

Under $Q = 0$, the relevant order is $4/25 < 13/50 < 189/400$. Hence

$$K(4/25, 189/400) > K(4/25, 13/50) = d_0. \tag{23}$$

Under $Q = 1$, the order is $4/25 < 99/400 < 189/400$, so

$$K(189/400, 4/25) > K(189/400, 99/400) = d_1. \tag{24}$$

The left sides are the two rates in Theorem 4, while Theorem 1 bounds the actual summary exponents by the right sides. This proves the first claim.

The example has positive selection. Theorem 2 gives full-history exponents d_0 and d_1 for the binary partition. The summary exponents are bounded above by the same two values, which rules out the strict summary-over-full-history ordering in Proposition 2(2). □

Table 1 gives the magnitudes. The gap arises in both quality states and exceeds the information bound by 0.1888 and 0.1481.

TABLE 1. Published rates and same-history information bounds

True quality	Theorem 4	Information bound
$Q = 0$	0.217554	0.028790
$Q = 1$	0.266232	0.118136

5. Finite-horizon path costs

The terminal-summary likelihood can be written as a control problem over the histories that reach the observed summary. Let C_n be a set of terminal summaries and let $\mathcal{R}(C_n)$ contain probability laws R on category histories satisfying $R(S_n \in C_n) = 1$. The conditional category law under R is denoted $u_t(\cdot \mid H_{t-1})$.

LEMMA 1 (Finite-horizon entropy representation). *For every terminal-summary event with* *positive probability under quality one,*

$$-\log P_1^{S_n}(C_n) = \inf_{R\in\mathcal{R}(C_n)} D_{\mathrm{KL}}(R\|P_1^{H_n}) \tag{25}$$

$$= \inf_{R\in\mathcal{R}(C_n)} \mathbb{E}_R \sum_{t=1}^{n} D_{\mathrm{KL}}(u_t(\cdot \mid H_{t-1}) \,\|\, \pi_1(\cdot, q(S_{t-1}))) \,. \tag{26}$$

PROOF. The first variational identity follows from conditioning $P_1^{H_n}$ on the event $S_n \in C_n$. For any R concentrated on that event, the chain rule decomposes $D_{\mathrm{KL}}(R\|P_1^{H_n})$ into the conditional divergences in (26). □

Dividing Lemma 1 by n gives the exact finite-horizon cost

$$I_{1,n}(C_n) := -\frac{1}{n}\log P_1^{S_n}(C_n) = \frac{1}{n}\inf_{R\in\mathcal{R}(C_n)} \mathbb{E}_R \sum_{t=1}^{n} D_{\mathrm{KL}}(u_t(\cdot \mid H_{t-1}) \,\|\, \pi_1(\cdot, q(S_{t-1}))) \,. \tag{27}$$

Equation (27) keeps the alternative transition law at the posterior generated by the controlled summary path. For every sequence $C_{0,n}$ reached with probability approaching one under quality zero, Theorem 1 gives $\limsup_n I_{1,n}(C_{0,n}) \le d_0$. A closed-form posterior exponent calls for a model-specific link between posterior neighborhoods and the limit of these finite-horizon control problems. The quality-one expression follows by reversing the states.

6. Finite recursion for the example

The binary partition permits a finite recursion. Let $p_Q(k,n)$ be the probability under quality Q of observing k top reviews after n actions, and set

$$q(k,n) = \frac{p_1(k,n)}{p_0(k,n) + p_1(k,n)}. \tag{28}$$

With $p_Q(0,0) = 1$, the recursion is

$$\begin{aligned} p_Q(k,n+1) &= p_Q(k,n)\{1 - \pi_Q(q(k,n))\} \\ &\quad + p_Q(k-1,n)\pi_Q(q(k-1,n)), \end{aligned} \tag{29}$$

where missing boundary terms equal zero. The same $q(k,n)$ enters the two quality transitions at each state.

Table 2 reports the evaluation of (29). The posterior columns are medians under the indicated true quality. The relative-entropy columns compare the two laws of the terminal count, while the history columns accumulate the conditional divergences in (13). Every horizon satisfies the finite-sample data-processing inequality. At $n = 4000$, the median

TABLE 2. Finite recursion under the uniform primitives

n	med. $R_{0,n}$	$D(S_n)_0/n$	$D(H_n)_0/n$	med. $R_{1,n}$	$D(S_n)_1/n$	$D(H_n)_1/n$
500	0.029414	0.029384	0.032424	0.112550	0.109461	0.112289
1000	0.029087	0.028808	0.030672	0.115111	0.113285	0.115107
2000	0.028939	0.028727	0.029738	0.116382	0.115576	0.116610
4000	0.028866	0.028743	0.029265	0.117013	0.116826	0.117372

posterior exponents and the two information averages lie within 0.0014 of the bounds $d_0 = 0.028790$ and $d_1 = 0.118136$. This finite-horizon proximity illustrates the low-cost same-history channel. Proposition 1 rests on the analytic information bound and endpoint inequalities.

7. Scope

Theorem 4's two formulas are contradicted by Proposition 1, as is the positive-selection speed ordering in Proposition 2(2). Proposition 2(1) uses the published summary rate and remains for separate derivation. Extensions that import the cross-boundary rate, including Theorem B-1 of the online appendix, require the same-history check.

Theorem 3 continues to deliver complete learning under strict separation. Theorem 2 continues to characterize learning from the full history. These results supply the policy-belief convergence and same-history information rate used in Theorem 1.

For the primitives in Section 4, a quality-zero-typical path costs 0.0288 per observation under quality one, while the cross-boundary formula assigns 0.2176. The corresponding quality-one figures are 0.1181 and 0.2662. The data-processing bounds determine both comparisons directly.

References

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar. 2022. "Learning From Reviews: The Selection Effect and the Speed of Learning." *Econometrica* 90 (6): 2857–2899. 10.3982/ECTA15847.

Fradkin, Andrey and David Holtz. 2023. "Do Incentives to Review Help the Market? Evidence from a Field Experiment on Airbnb." *Marketing Science* 42 (5): 853–865. 10.1287/mksc.2023.1439.