A Model of Two Learning Processes: Online Appendix

Jin Huang [a] and Zheng Gong [b]

[a]NYU Shanghai

[b]The Chinese University of Hong Kong, Shenzhen

Contents

1 Appendix A: Conclusive Good News

In the main paper, we focus on the analysis and results for the case of conclusive bad news. In this appendix, we extend the discussion by examining a market with a biased information source that sends conclusive good news: Poisson signals arrive at a rate of λ if the state is H, and at a rate of 0 if the state is L. We begin by analyzing the benchmark model with an exogenously given product price, and then proceed to examine the case of endogenous pricing.

1.1 Benchmark: Two Learning Processes with Exogenous Price

Suppose the product is sold at a fixed exogenous price p. Let π denote the public belief before a consumer's active learning. During the active learning process, conditional on the absence of news, the evolution of the belief that the state is H is as follows:

$$\pi_m = \frac{\pi e^{-\lambda m}}{\pi e^{-\lambda m} + 1 - \pi}.$$

It is straightforward to confirm that the belief decreases as learning continues: without good news, the consumer becomes increasingly pessimistic about the product's value. Similar to the analysis in the main model for conclusive bad news, given that active learning has started, the consumer stops learning only in two cases: the signal arrives and therefore the consumer buys the product, or the belief is updated to a low enough level that it is more costly to continue learning than to leave the market directly without buying. Thus, the value function is $V_m = \max\{\omega_m, 0\}$

$$\omega_m = \lim_{\xi \to 0^+} \pi_m(1 - e^{-\lambda\xi})(v - p) + (1 - \pi_m(1 - e^{-\lambda\xi}))\max\{\omega_{m+\xi}, 0\} - c\xi.$$

The single-crossing condition is as follows: for an input level m such that $w_m = 0$, we have $w_\tau < 0$ for all $\tau > m$. This implies that the optimal stopping point m^* satisfies:

$$\begin{aligned} 0 &= \lim_{\xi \to 0^+} \pi_m(1 - e^{-\lambda\xi})(v - p) - c\xi, \\ &= \lim_{\xi \to 0^+} \pi_m(\lambda\xi)(v - p) - c\xi. \end{aligned}$$

After rearranging the equation, we can get

$$\overbrace{\underbrace{\lambda\xi\pi_{m^*}}_{\substack{\text{Probability}\\ \text{of signal}}} \times \underbrace{(v-p)}_{\substack{\text{Benefit:}\\ \text{Buy Good Product}}}}^{\text{Benefit of Learning}} = \overbrace{c\xi.}^{\text{Cost of Learning}}$$

Thus, the optimal stopping condition is: $\pi_{m^*} = \frac{c}{\lambda(v-p)}$. The optimal stopping rule is described as follows: when good news arrives, the consumer stops learning and buys the product; if no good news arrives, the consumer continues learning until the stopping point m^* at which the belief reaches $\frac{c}{\lambda(v-p)}$, and leaves the market without buying.

To ensure that active learning occurs, first, the marginal benefit of learning needs to be no lower than the marginal cost of learning at the beginning of the learning process. The following condition needs to be met:

$$\pi \geq \frac{c}{\lambda(v-p)}. \tag{1}$$

In addition, the consumer's expected utility of conducting active learning should be higher than the utility of directly buying (i.e., $\pi v - p$). Let x denote the cumulative learning input required for the first signal to arrive. We have the following:

$$\pi[e^{-\lambda m^*}(-m^*c) + (1-e^{-\lambda m^*})[v-p-cE(x \mid x < m^*, H)]] + (1-\pi)(-m^*c) > \pi v - p,$$

where

$$E(x \mid x < m^*, H) = \frac{1}{\lambda} - \frac{e^{-\lambda m^*}}{1-e^{-\lambda m^*}} m^*.$$

After substituting the expression for π_{m^*} and simplifying the condition, we get:

$$\pi < 1 - \frac{c}{\lambda p + c \ln \frac{(1-\pi)c}{\pi[\lambda(v-p)-c]}}. \tag{2}$$

To simplify the exposition, let l_1^G and l_2^G denote the two thresholds represented on the right-hand side of Equation (1) and Equation (2) (after rearranging the equation and moving π to the left-hand side), respectively. The next lemma summarizes the inactive learning region and active learning region for the public belief π:

LEMMA 1 *In a market with conclusive good news, given product price p, learning cost c and belief π that the state is H after observational learning, a consumer will buy the product directly if $\pi \geq l_2^G$; will conduct active learning if $l_1^G \leq \pi < l_2^G$; will leave the market without learning if $\pi < l_1^G$.*

It is obvious that l_1^G is decreasing in the signal arrival rate λ and increasing in the learning cost c. Though the closed-form expression for l_2^G is too complicated to be presented here, after solving for the implicit differentiation by taking the total derivatives of both sides of Equation (2), we find that l_2^G is increasing in λ and p, and decreasing in v and c if $c < \lambda p(1 - \frac{p}{v})$, which is the same condition for $l_1^G < \frac{p}{v}$ (i.e., consumers will not buy the product at the stopping point of active learning under conclusive good news) and $l_2^B > \frac{p}{v}$ (i.e., consumers will buy the product at the stopping point of active learning under conclusive bad news).

We proceed to study the information transmission between consumers. Suppose the common prior belief π_0 is such that the first consumer conducts active learning. If the state is L, the probability of not buying is 1; if the state is H, the probability of not buying is the probability of no signal arriving before m^*, i.e., $e^{-\lambda m^*}$. When it is time for the second consumer to make a decision, if she observes that the first consumer bought the product, the state H is revealed; if she observes that the first consumer didn't buy the product, the public belief is updated to

$$\Pr(H \mid h = 0) = \frac{\pi_0 e^{-\lambda m^*}}{\pi_0 e^{-\lambda m^*} + 1 - \pi_0}.$$

Substituting m^* into the equation, we can see that it equals $\frac{c}{\lambda(v-p)}$, which is the posterior belief of the first consumer after active learning. As the optimization problem is the same for all consumers, the second consumer has no incentive to acquire additional information on top of the information obtained by the first consumer. Therefore, the learning outcome is that a cascade/herd arises immediately after the first consumer – all later consumers free-ride on the first consumer and do not engage in active learning on their own. The probability of an incorrect herd is the probability of no signal arriving before m^* during the active learning process of the first consumer:

$$\Pr(\text{incorrect herd}) = e^{-\lambda m^*} = \frac{(1-\pi_0)c}{\pi_0(\lambda(v-p)-c)}. \tag{3}$$

It is straightforward to see that the probability of an incorrect herd is increasing in the learning cost c, and is decreasing in the signal arrival rate λ, the common prior belief π_0, and the net utility of a correct purchase $v-p$.

1.2 Two Learning Processes with Dynamic Pricing

In this section, we examine the seller's optimal dynamic pricing strategy in the presence of consumers' two learning processes. We show that the intuitions and results under conclusive good news align closely with those under conclusive bad news.

We begin by analyzing how a consumer responds to different prices. There are three possibilities: opting out of the market directly without active learning, engaging in active learning, or making a direct purchase without active learning. In the second case, the consumer either purchases the product upon receiving good news or, in the absence of news, continues learning until her belief reaches $\pi_{m^*}(p) = \frac{c}{\lambda(v-p)}$, at which point she opts out. The following lemma characterizes the consumer's optimal action when facing price p.

LEMMA 2 *In a market with conclusive good news, for any initial belief π, three cutoffs exist: $p_1^G(\pi)$ that satisfies $1-\pi = \frac{c}{\lambda p_1^G(\pi)+c\ln\frac{(1-\pi)c}{\pi[\lambda(v-p_1^G(\pi))-c]}}$, $p_2^G(\pi) = v - \frac{c}{\lambda\pi}$, and $p_3^G(\pi) = \pi v$.*

1. *When $p_1^G(\pi) < p_2^G(\pi)$, we have: a consumer will opt out without learning if $p > p_2^G(\pi)$; will engage in active learning until her posterior belief reaches $\frac{c}{\lambda(v-p)}$ if $p_1^G(\pi) < p \leq p_2^G(\pi)$; will buy the product directly if $p \leq p_1^G(\pi)$.*

2. *When $p_1^G(\pi) \geq p_2^G(\pi)$, we have: a consumer will opt out without learning if $p > p_3^G(\pi)$; will buy the product directly if $p \leq p_3^G(\pi)$.*

The proof is provided in Section 1.3. Similar to the analysis presented in the main paper, the first cutoff, $p_1^G(\pi)$, is the indifference point between buying directly and active learning. The second cutoff, $p_2^G(\pi)$, is the indifference point between active learning and opt out. The third cutoff, $p_3^G(\pi)$, is the indifference point between direct buying and opt-out. If $p_1^G(\pi) \geq p_2^G(\pi)$, active learning can never be the optimal strategy for any given price. The consumer either purchases directly ($p \leq p_3^G(\pi)$) or opts out ($p > p_3^G(\pi)$), and the public belief will never be updated. In this case, myopic and forward-looking sellers will always set a price at $p = p_3^G(\pi)$. Our following analysis will focus on the case when $p_1^G(\pi) < p_2^G(\pi)$.

We proceed to analyze the seller's optimal dynamic pricing strategy, comparing myopic sellers with forward-looking sellers to examine how sellers may utilize dynamic pricing to leverage information externalities between consumers. The following Lemma characterizes the optimal pricing strategy for a myopic seller.

LEMMA 3 (MYOPIC SELLER) *In a market with conclusive good news, a myopic seller has two possible strategies for each period with a prior belief π. First, setting a low price, $p_{DL}^{G*} = p_1^G(\pi)$, to deter learning and make the consumer buy directly. Second, setting a high price, $p_{IL}^{G*} = \arg\max_{p \in (p_1^G(\pi), p_2^G(\pi))} \frac{\pi - \frac{c}{\lambda(v-p)}}{1 - \frac{c}{\lambda(v-p)}} p$ to induce learning. The seller sets p_{DL}^{G*} if $p_{DL}^{G*} \geq \max_{p \in (p_1^G(\pi), p_2^G(\pi))} \frac{\pi - \frac{c}{\lambda(v-p)}}{1 - \frac{c}{\lambda(v-p)}} p$, and vice versa.*

A proof is provided in Section 1.3. A myopic seller chooses a price that maximizes the current period payoff. For all prices that deter consumer learning (denoted by subscript *DL*), the highest price among them (p_{DL}^{G*}) dominates all others. However, this price may not yield the highest payoff, as it needs to be compared with the optimal price that induces active learning (denoted by subscript *IL*), p_{IL}^{G*}. Additionally, it is straightforward that once the myopic seller starts choosing p_{DL}^{G*}, the seller will consistently set this price in the subsequent periods.

We then analyze the pricing strategy of a forward-looking seller. Similar to the results obtained in the conclusive bad-news market, we find that forward-looking sellers are more likely to induce learning by distorting early prices, thereby transforming a *learning-deterring* scheme into a *learning-inducing* scheme. Proposition 1 summarizes this result.

PROPOSITION 1 (FORWARD-LOOKING SELLER) *In a market with conclusive good news, when it is optimal for a myopic seller to deter learning by charging a low price $p = p_1^G(\pi)$, a forward-looking seller may charge a higher price to induce learning. The learning-inducing parameter region expands with a higher β.*

A proof is provided in Section 1.3. This result echoes Proposition 4 of the main paper, following the same intuition: when consumers learn more about the product, the seller can extract more surplus from them. Consider an extreme case: when a consumer perfectly learns the true state, the seller can extract the full surplus from the consumer ($p = v$ if H state, and $p = 0$ if L). Therefore, a forward-looking seller may sacrifice the current payoff by increasing the price to induce active learning and gain from the more precise information transmitted to subsequent consumers.

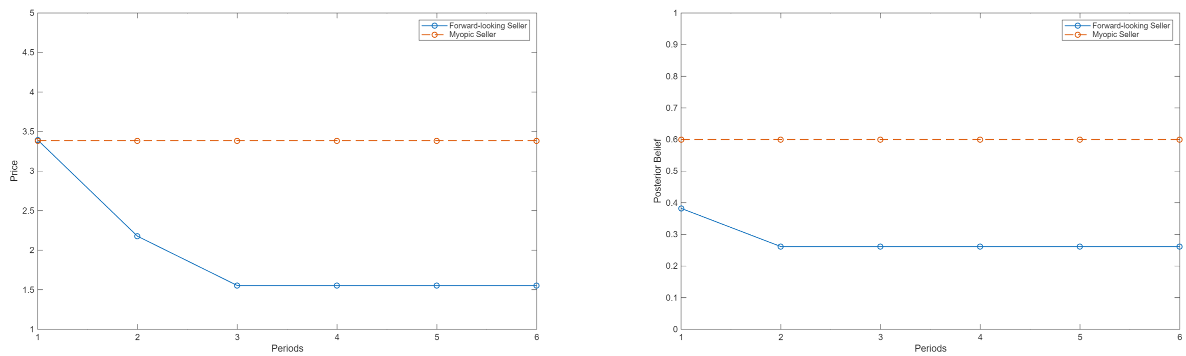

Figure I: Pricing and Posterior Belief Dynamics for Forward-Looking and Myopic Sellers, Conditional on No Good Signal ($\pi_0 = 0.6, c = 1, v = 6, \lambda = 1, \beta = 0.95$)

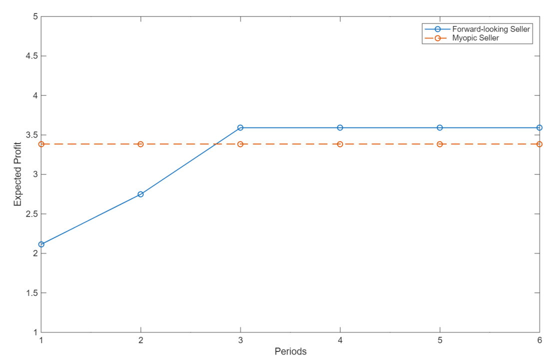

Figure II: Expected Single-Period Profit Dynamics for Forward-Looking and Myopic Sellers ($\pi_0 = 0.6, c = 1, v = 6, \lambda = 1, \beta = 0.95$)

To illustrate the market dynamics under the conclusive good news environment, we present a numerical example. The parameters are set to $\pi_0 = 0.6, c = 1, v = 6, \lambda = 1$, and $\beta = 0.95$, corresponding to a region where the pricing strategies of myopic and forward-looking sellers diverge. Figure I depicts the evolution of pricing and consumers' posterior beliefs at the end of each period, conditional on no good news arriving. Beginning in the first period, the myopic seller always secures a sale by setting a lower, learning-deterring price. The forward-looking seller, in contrast, initially sets a higher price to induce the first consumer to actively search for a breakthrough signal. If the first consumer fails to receive a signal, the belief declines because "no news is bad news," and the forward-looking seller is compelled to reduce the price in the second period to incentivize the new entrant to continue learning. If no signal arrives again, the posterior belief at the end of the second period, which becomes the public belief in the third period, falls to a level where the expected profit from inducing learning is lower than that from deterring learning. The forward-looking seller then switches to the learning-deterring strategy. From this point onward, active learning in the market ceases, and the public belief remains constant. This example illustrates the distinct market outcome under conclusive good news: unlike in the bad news environment, the absence of news gradually lowers belief, which eventually induces the forward-looking seller to adopt the learning-deterring strategy.

It is important to note that Figure I depicts the path where signals do not arrive in the first two periods. Ex ante, however, the expected profit for the forward-looking seller is higher when adopting the learning-inducing strategy at the beginning. The rationale parallels that in the conclusive bad news environment: the forward-looking seller may find it optimal to induce learning early, at the cost of lower expected profit in the current period, in order to obtain potentially higher future returns. If a consumer discovers the good news signal, the public belief jumps to one, enabling the seller to charge the maximum feasible price ($p = v$) to all subsequent consumers indefinitely. As shown in Figure II, the expected long-run profit from adopting a learning-inducing price at the beginning exceeds that from the learning-deterring strategy.

Finally, we examine the probability of an incorrect herd in the conclusive good news market. As the seller will always set prices that lead consumers to either buy directly or engage in active learning, there will never be an incorrect herd when the true state is H. However, when the true state is L, an incorrect herd occurs with certainty. This is because, under the L state, consumers will never receive conclusive good news, leading to a decreasing public belief if the seller induces learning. When the public belief is sufficiently low, it becomes strictly optimal for the seller to deter learning for any $\beta \in (0, 1)$. From that point onward, all subsequent consumers incorrectly purchase the product without engaging in any active learning. We summarize this result in the following corollary; its proof can be found in the next section.

COROLLARY 1 (INCORRECT HERD) *In a market with conclusive good news, an incorrect herd only* *occurs when the true state is L and occurs with probability 1.*

1.3 Proof

Proof of Lemma 2

First, we show that $p_2^G(\pi) = v - \frac{c}{\lambda\pi}$ is the cutoff point that a consumer is indifferent between active learning and opting out. If the consumer opts out without buying, her payoff is 0. If she chooses to engage in active learning, her marginal gain comes from receiving the good news and buying the product. Thus, the consumer's marginal gain is $\lambda\pi(v-p)$. Therefore, the cutoff price $p_2^G(\pi)$ is such that $\lambda\pi(v - p_2^G(\pi)) = c$, which yields $p_2^G(\pi) = v - \frac{c}{\lambda\pi}$.

Second, for $p_1^G(\pi)$, this is a direct result from (2), which characterizes the cutoff point where the consumer is indifferent between buying directly and actively learning. When $p_1^G(\pi) < p_2^G(\pi)$, we have three regions: when $p \leq p_1^G(\pi)$, the consumer buys directly; when $p_1^G(\pi) < p \leq p_2^G(\pi)$, the consumer engages in active learning; when $p > p_2^G(\pi)$, the consumer opts out.

Finally, when $p_1^G(\pi) > p_2^G(\pi)$, there will not be a situation where the consumer engages in active learning because it is either dominated by buying directly (when $p < p_1^G(\pi)$), or dominated by opting out (when $p > p_2^G(\pi)$). There exists the third cutoff point where the consumer is indifferent between buying directly and opting out. As $p = \pi v$ gives the consumer zero expected payoff, it is obviously the cutoff point $p_3^G(\pi)$.

■

Proof of Lemma 3

There are two possible strategies for the seller. First, setting a low price that induces the consumer to buy directly. Since any price $p \leq p_1^G(\pi)$ will deter learning, the seller charges the highest possible price, that is, $p_{DL}^* = p_1^G(\pi)$. Second, the seller can set a high price, $p \in (p_1^G(\pi), p_2^G(\pi))$, to induce active learning by the consumer.

Given price p, a consumer stops learning and leaves the market at the stopping belief $\pi_{m^*} = \frac{c}{\lambda(v-p)} < \pi$. As the belief updating process is a martingale, the probability that the consumer receives the good signal s_g in her learning process is

$$\Pr(s_g = 1) = \frac{\pi - \pi_{m^*}}{1 - \pi_{m^*}}.$$

And if the consumer receives the signal, she purchases at the price p. Hence, the seller makes an expected profit of $\max_{p \in (p_1^G(\pi), p_2^G(\pi))} \frac{\pi - \frac{c}{\lambda(v-p)}}{1 - \frac{c}{\lambda(v-p)}} p$. For a myopic seller, it compares only the static payoffs of the two strategies.

Proof of Proposition 1

We first prove some properties of the three thresholds, $p_1^G(\pi)$, $p_2^G(\pi)$, and $p_3^G(\pi)$, summarized in the following Lemma.

LEMMA 4 *In the region where $p_1^G(\pi) < p_2^G(\pi)$, $p_2^G(\pi)$ is concavely increasing in π and $p_1^G(\pi)$ is convexly increasing in π. $p_3^G(\pi)$ is linearly increasing in π. If the parameter satisfies $\lambda v > 4c$, when $\pi = \frac{\lambda v \pm \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$, $p_1^G(\pi) = p_2^G(\pi) = p_3^G(\pi)$. When $\pi \in (\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}, \frac{\lambda v + \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v})$, $p_2^G(\pi) > p_3^G(\pi) > p_1^G(\pi)$.*

Proof. $p_3^G(\pi) = \pi v$, which is linearly increasing in π. $p_2^G(\pi) = v - \frac{c}{\lambda \pi}$, which is easy to verify that it is increasing and concave in π. For p_1^G, recall that p_1^G is determined by

$$1 - \pi = \frac{c}{\lambda p_1^G(\pi) + c \ln \frac{(1-\pi)c}{\pi[\lambda(v - p_1^G(\pi)) - c]}} \tag{4}$$

Fixing p, with an increase in π, LHS of (4) decreases and the RHS of (4) increases. As an increase in $p_1^G(\pi)$ decreases the RHS, we must have $\frac{dp_1^G}{d\pi} > 0$. For the concavity, the second-order derivative of (4) is

$$\frac{d^2 p_1^G(\pi)}{d\pi^2} = (1-\pi)(4\pi^2 - 3\pi + 1) \frac{c\lambda^2 (v - p_1^G(\pi))(\lambda(v - p_1^G(\pi)) - c)}{\left[\pi(1-\pi)^2 \lambda^2 (v - p_1^G(\pi))\right]^2}$$

As $p_1^G < p_2^G = v - \frac{c}{\lambda \pi}$, $\lambda(v - p_1^G(\pi)) - c > \frac{c}{\pi} - c > 0$. Therefore, $\frac{d^2 p_1^G(\pi)}{d\pi^2} > 0$, which implies that $p_1^G(\pi)$ is a convexly increasing function in π.

Next, we show that $p_1^G = p_2^G = p_3^G$ when $\pi = \frac{\lambda v \pm \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$. It is easy to verify that these two values equal $p_2^G = v - \frac{c}{\lambda \pi}$ and $p_3^G = \pi v$. Substituting $p_1^G = p_2^G = v - \frac{c}{\lambda \pi}$ into (4). We have

$1-\pi=\frac{c}{\lambda(v-\frac{c}{\lambda\pi})}$. It is easy to verify that $\pi=\frac{\lambda v\pm\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v}$ are two roots.

Lastly, we prove that when $\pi\in(\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v},\frac{\lambda v+\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v})$, $p_2^G(\pi)>p_3^G(\pi)>p_1^G(\pi)$. Substituting $p_1^G=p_2^G=v-\frac{c}{\lambda\pi}$ into (4). We have

$$1-\pi=\frac{c}{\lambda(v-\frac{c}{\lambda\pi})}$$

When $\pi\in(\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v},\frac{\lambda v+\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v})$, $1-\pi>\frac{c}{\lambda(v-\frac{c}{\lambda\pi})}$. From (4), a decrease in $p_1^G(\pi)$ increases the RHS. Thus, $p_1^G(\pi)$ must be lower than $p_2^G(\pi)$. Moreover, $p_2^G(\pi)>p_3^G(\pi)\iff v-\frac{c}{\lambda\pi}>\pi v\iff \lambda v\pi(1-\pi)+c>0$. We get when $\pi\in(\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v},\frac{\lambda v+\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v})$, we also have $p_2^G(\pi)>p_3^G(\pi)$.

If $p_3^G(\pi)<p_1^G(\pi)<p_2^G(\pi)$, at $p_1^G(\pi)$, the consumer is indifferent between direct purchase and active learning, but both options result in a negative expected payoff for the consumer as $p_1^G>p_3^G$. However, at p_2^G, the consumer has a zero expected payoff as it is the indifference point between active learning and opting out. We then obtain that the consumer has a strictly higher utility with a higher price, leading to a contradiction. Hence, we must have $p_1^G(\pi)<p_3^G(\pi)<p_2^G(\pi)$ in the region of $\pi\in(\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v},\frac{\lambda v+\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v})$.

■

Next, we show that when it is optimal for a myopic seller to deter learning, a forward-looking seller may use a higher price to induce learning. From Lemma 4, we know that $p_1^G<p_3^G<p_2^G$ when $\pi\in(\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v},\frac{\lambda v+\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v})$. And if $p_1^G\geq p_2^G$, there exists no price that induces active learning. Thus, if the prior belief is such that $\pi_0\leq\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v}$ or $\pi_0\geq\frac{\lambda v+\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v}$, a myopic seller behaves the same as a forward-looking seller.

When $\pi_0\in(\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v},\frac{\lambda v+\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v})$, charging a price $p\in(p_1^G(\pi_0),p_2^G(\pi_0))$ to induce learning is an option. It is possible that $p_1^G(\pi_0)>\max_{p\in(p_1^G(\pi_0),p_2^G(\pi_0))}\frac{\pi_0-\frac{c}{\lambda(v-p)}}{1-\frac{c}{\lambda(v-p)}}p$, where a myopic seller will charge $p=p_1^G(\pi_0)$ to deter learning. As the public belief does not change, the seller will choose the same price in all subsequent periods, leading to an equilibrium profit of $\Pi_{DL}^*=\frac{p_1^G(\pi_0)}{1-\beta}$.

Now, consider a deviation that is not necessarily the optimal pricing scheme for the forward-looking seller: instead of charging $p_1=p_1^G(\pi_0)$, set the first-period price to be $p_1=p_1^G(\pi_0)+\epsilon$, where $\epsilon\to 0$, which makes the first consumer engage in active learning until the posterior belief reaches $\pi_1=\frac{c}{\lambda(v-p_1)}<\pi_0$. In subsequent periods, if any consumer receives good news, the seller sets the price at $p = v$ forever. Otherwise, the seller charges $p = p_1^G(\pi) + \epsilon$ in all periods with public belief π, until π falls below a threshold $\underline{\pi}$, where $\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v} < \underline{\pi} < \pi_0$ and

$$\frac{\pi_0 - \underline{\pi}}{1 - \underline{\pi}} v + \frac{1 - \pi_0}{1 - \underline{\pi}} p_1^G(\underline{\pi}) \geq p_1^G(\pi_0),$$

and after π falls below $\underline{\pi}$ for the first time, the seller charges $p = p_1^G(\pi)$ forever.

We then demonstrate that this deviation is feasible and profitable for a patient enough seller. From Lemma 4, $p_1^G(\pi_0) < p_3^G(\pi_0) = \pi_0 v$. Note that if the seller always induces consumer learning by a sequence of prices $\{p_1 = p_1^G(\pi_0) + \epsilon, p_2 = p_1^G(\pi_1) + \epsilon, \dots\}$, the limit of the sequence of posterior belief is $\pi_n \to \frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$. By Lemma 4, when $\pi = \frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$, $p_1^G = p_2^G = p_3^G$, which implies that the seller could charge a price $p = \pi v$ and extract all the surplus. Therefore, the ex-ante expected payoff with posterior belief $(\pi_L = \pi, \pi_H = 1)$ when π reaches $\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$ equals πv, which is greater than $p_1^G(\pi_0)$. However, π converges to $\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$ but does not reach the limit in finite periods. Meanwhile, for any $\pi > \frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$, it can be reached in finite time $T(\pi)$. When π changes, the ex-ante expected payoff with posterior belief $(\pi_L = \pi, \pi_H = 1)$ changes continuously. Hence, there must be a threshold $\underline{\pi}$ that can be reached in finite periods $T(\underline{\pi})$, and the threshold satisfies $\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v} < \underline{\pi} < \pi_0$ and $\frac{\pi_0 - \underline{\pi}}{1 - \underline{\pi}} v + \frac{1 - \pi_0}{1 - \underline{\pi}} p_1^G(\underline{\pi}) \geq p_1^G(\pi_0)$.

Finally, it remains to prove that this is a profitable deviation for a patient enough seller. In the event that no consumer receives good news within $T(\underline{\pi})$ periods, the public belief decreases to $\pi < \underline{\pi}$. From the $T(\underline{\pi}) + 1$ period on, the seller charges $p_1^G(\pi)$ forever. The ex-ante expected profit of periods after $T(\underline{\pi})$ is

$$\mathbb{E}(\Pi_\pi) = \frac{\pi_0 - \pi}{1 - \pi} v + \frac{1 - \pi_0}{1 - \pi} p_1^G(\pi) > p_1^G(\pi_0). \tag{5}$$

Denote the profit under this deviation by Π_{IL}. As it is unnecessarily the optimal deviation, $\Pi_{IL}^* \geq \Pi_{IL}$. For Π_{IL}, we must have

$$\Pi_{IL} > \sum_{t=T(\underline{\pi})+1}^{\infty} \beta^{t-1} \mathbb{E}(\Pi_\pi) = \frac{\beta^{T(\underline{\pi})}}{1 - \beta} \mathbb{E}(\Pi_\pi). \tag{6}$$

The RHS of (6) does not include all the expected profits from periods before $T(\underline{\pi})$, and even if we are simply comparing the RHS of (6) with Π_{DL}^*, the RHS can be greater when β is large enough. More specifically, when $\beta^{T(\underline{\pi})} \geq \frac{p_1^G(\pi_0)}{\mathbb{E}(\Pi_\pi)}$, we have

$$\Pi_{IL}^* \geq \Pi_{IL} > \frac{\beta^{T(\underline{\pi})}}{1-\beta}\mathbb{E}(\Pi_\pi) \geq \Pi_{DL}^*.$$

Therefore, when the seller is patient enough, the seller may induce learning with a higher initial price under the same conditions that make a myopic seller choose to deter learning. We also provide Figure III to illustrate the idea of this proof.

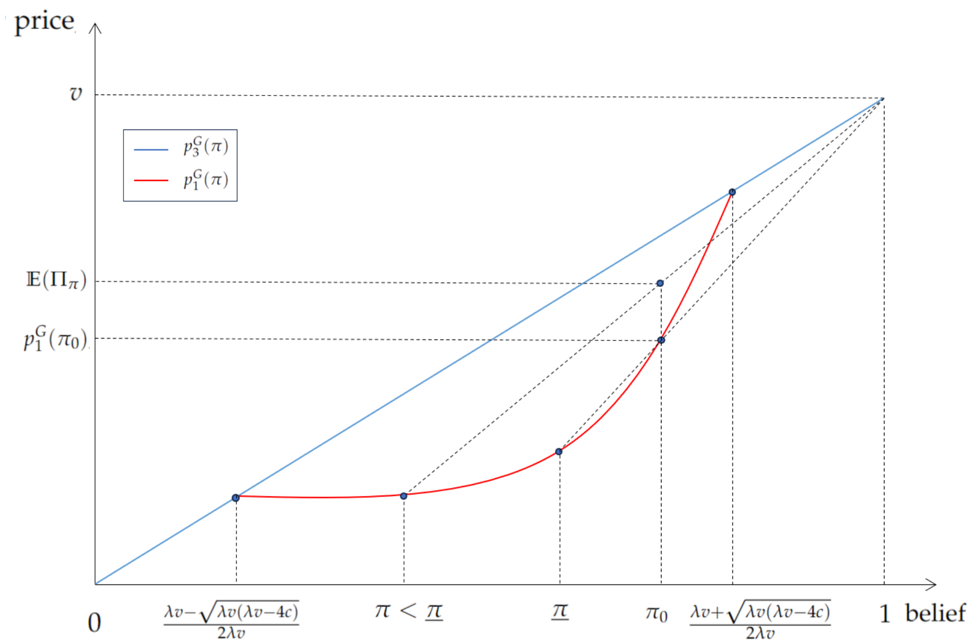

Figure III: Illustration of the proof of Proposition 1

■

Proof of Corollary 1

In the conclusive good news market, we only need to show that if the seller keeps charging learning-inducing prices and good news never arrives (resulting in all consumers engaging in active learning not purchasing), the seller will deter learning for any $\beta \in (0,1)$ when the public belief is sufficiently low. More specifically, from Lemma 4, the seller can induce learning when $\pi \in (\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v}, \frac{\lambda v+\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v})$. We aim to show the existence of a lower threshold $\underline{\pi} > \frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v}$, such that for any $\pi \in (\frac{\lambda v-\sqrt{\lambda v(\lambda v-4c)}}{2\lambda v}, \underline{\pi}]$, although $p_2^G(\pi) > p_1^G(\pi)$, it is still optimal for the seller to charge $p_1^G(\pi)$ to deter learning.

At belief π, the learning inducing profit $\Pi_{IL}(\pi)$ is

$$\Pi^*_{IL}(\pi) = \max_{p \in (p_1^G(\pi), p_2^G(\pi))} \frac{\pi - \frac{c}{\lambda(v-p)}}{1 - \frac{c}{\lambda(v-p)}} p,$$

where $\frac{\pi - \frac{c}{\lambda(v-p)}}{1 - \frac{c}{\lambda(v-p)}}$ is the probability that the consumer will make a purchase at price p, which is strictly decreasing in p. Also note that since $p > p_1^G(\pi) > p_1^G(\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}) = p_2^G(\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}) = v - \frac{2cv}{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}$, we have

$$\frac{\pi - \frac{c}{\lambda(v-p)}}{1 - \frac{c}{\lambda(v-p)}} < \frac{\pi - \frac{c}{\lambda(v - p_2^G(\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}))}}{1 - \frac{c}{\lambda(v - p_2^G(\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}))}} = \frac{\pi - \frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}}{1 - \frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}}.$$

Substituting it into $\Pi^*_{IL}(\pi)$, we have

$$\Pi^*_{IL}(\pi) = \max_{p \in (p_1^G(\pi), p_2^G(\pi))} \frac{\pi - \frac{c}{\lambda(v-p)}}{1 - \frac{c}{\lambda(v-p)}} p < \frac{\pi - \frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}}{1 - \frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}} (v - \frac{c}{\lambda \pi}) \equiv \bar{\Pi}_{IL}(\pi),$$

where we analytically find an upper limit of the profit under learning-inducing strategy, $\bar{\Pi}_{IL}(\pi)$.

On the other hand, if the seller chooses to deter learning, the price charged is $p = p_1^G(\pi) > p_1^G(\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}) = p_2^G(\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}) = v - \frac{2cv}{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}$. And the consumer purchases with probability 1. Therefore, we can obtain a lower bound of the profit under the learning-deterring strategy, $\underline{\Pi}_{DL}(\pi) = v - \frac{2cv}{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}$.

Under any $\beta \in (0,1)$, if the seller deters learning, the profit is lower-bounded by $\frac{\underline{\Pi}_{DL}(\pi)}{1-\beta}$. If the seller induces learning, the profit is upper-bounded by $\bar{\Pi}_{IL}(\pi) + \frac{\beta}{1-\beta}\pi v$. However, when $\pi \to \frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$, $\bar{\Pi}_{IL}(\pi) \to 0$ and $\pi v \to \underline{\Pi}_{DL}(\pi)$ (by Lemma 4), and $\underline{\Pi}_{DL}(\pi) > 0$ and does not change with π. Therefore, there must exist a cutoff $\underline{\pi}$, such that $\frac{\underline{\Pi}_{DL}(\pi)}{1-\beta} = \bar{\Pi}_{IL}(\underline{\pi}) + \frac{\beta}{1-\beta}\underline{\pi} v$.

From the above analysis, we can see that the learning-inducing scheme is not sustainable indefinitely. When the public belief falls below $\underline{\pi}$, it becomes optimal for the seller to consistently set a learning-deterring price. In summary, two scenarios may unfold: either early consumers receive conclusive good news, leading the public belief to jump to 1, or in the absence of news, the public belief drops below $\underline{\pi}$, prompting the seller to deter learning. While early consumers might choose not to make a purchase, in the long run, later consumers will always buy the product. Therefore, when the true state is H, the probability of an incorrect herd is 0. When the true state is L, the probability of an incorrect herd is 1. ■

2 Appendix B: Inconclusive Discrete Signal

In our main paper, we model the consumer's active learning process as a stopping problem over incremental search inputs, with information arriving according to a Poisson process that yields conclusive signals. This formulation, which is common in the strategic experimentation literature, provides tractability in characterizing dynamic belief updating. An alternative way to formalize active learning is to endogenize the private signal in the standard framework of the observational learning literature. In this setting, each consumer decides whether to pay a cost to obtain one informative but inconclusive signal. This appendix develops the analysis under this alternative information structure.

Our analysis proceeds in two steps. In Sections 1–3, for analytical clarity, we first consider a benchmark case in which the consumer pays a fixed cost to generate a signal of exogenous precision. In Section 4, we extend the model by allowing the consumer to choose the level of costly effort to invest, which endogenously determines the precision of the signal. Although this extension is analytically intractable, we solve it numerically. Throughout the analysis, we show that the central result of our paper, that a forward-looking seller has stronger incentives to induce learning than a myopic seller, remains robust under the inconclusive discrete signal setting.

2.1 Model Setup

We assume that each consumer can pay a fixed cost $c > 0$ to acquire a single private signal $s \in h, l$, where h indicates good news and l indicates bad news. The signal is symmetric, informative yet noisy.[1] The quality or precision of the signal is defined by

$$\rho = \Pr(s = h|H) = \Pr(s = l|L) \in (\frac{1}{2}, 1).$$

In each period n, both consumer n and the seller observe all previous transaction outcomes, which determine the prior belief/the public belief $\pi = \pi_{n-1}$ via Bayes' rule. After observing the seller's price $p = p_n$, the consumer chooses one of three actions:

[1] For analytical simplification, we restrict attention to the symmetric case; the analysis of asymmetric signals where $\Pr(s = h|H) \neq \Pr(s = l|L)$ is analogous.

1. Buy directly: Purchase the product without acquiring a signal, yielding an expected payoff of $V_B(\pi, p) = \pi v - p$.

2. Exit: Leave the market without learning or purchasing, which gives a payoff of $V_L = 0$.

3. Learn: Pay the cost c, observe the signal s, and update the belief accordingly. Afterward, the consumer decides whether to buy or exit.

If the consumer acquires a signal, the posterior belief is updated via Bayes' rule. Observing a signal $s = h$ leads to the posterior belief:

$$\pi_h(\pi) = Pr(H|s = h, \pi) = \frac{\rho\pi}{\rho\pi + (1-\rho)(1-\pi)}.$$

Observing a signal $s = l$ leads to the posterior belief:

$$\pi_l(\pi) = \frac{(1-\rho)\pi}{(1-\rho)\pi + \rho(1-\pi)}.$$

For any prior belief $\pi \in (0, 1)$ and signal quality $\rho \in (\frac{1}{2}, 1)$, we have $\pi_l(\pi) < \pi < \pi_h(\pi)$.

2.2 The Consumer's Decision Problem

A rational consumer chooses to pay the cost c only if the acquired information can potentially alter her purchase decision. This means that the consumer will purchase upon receiving good news ($s = h$) and will exit upon receiving bad news ($s = l$). Such scenario is possible only if the price p lies in the interval $\pi_l(\pi)v < p < \pi_h(\pi)v$. Given a price in this region, the expected payoff from acquiring the signal equals the expected value of the posterior outcomes minus the cost:

$$V_S(\pi, p) = \Pr(h|\pi)(\pi_h(\pi)v - p) - c = \rho\pi v - p\Pr(h|\pi) - c.$$

The consumer's optimal action depends on the price and her prior belief, as characterized in the following lemma.

LEMMA 1 (THE LEARNING REGION) *Given price p, the consumer's decision rule is determined by her prior belief π. There exists a learning region of intermediate beliefs, (π_L^*, π_H^*), in which acquiring a signal is* *optimal. For beliefs $\pi \geq \pi_H^*$, the consumer buys directly without learning. For beliefs $\pi \leq \pi_L^*$, the consumer exits the market without learning. The thresholds π_H^* and π_L^* are defined by the indifference conditions: : $V_S(\pi_H^*, p) = V_B(\pi_H^*, p)$ and $V_S(\pi_L^*, p) = V_L = 0$.*

2.3 The Seller's Dynamic Pricing Strategy

We now endogenize the price and compare the optimal strategies of a myopic seller and a forward-looking seller.

2.3.1 The Myopic Seller

A myopic seller with belief π chooses a price p to maximize expected profit from the current consumer. The price induces one of two possible responses: a direct purchase or a learning decision. To deter learning and guarantee an immediate purchase, the seller sets a learning-deterring price $\tilde{p}(\pi)$ that makes the consumer indifferent between buying directly and acquiring a signal ($V_B = V_S$). This price is given by

$$\tilde{p}(\pi) = \frac{(1-\rho)\pi v + c}{\Pr(l|\pi)},$$

where $\Pr(l|\pi)$ denotes the probability of observing bad news given prior belief π. At this price, the seller's one-period profit is $\Pi^*_{Deter}(\pi) = \tilde{p}(\pi)$.

To induce learning, the seller sets a learning-inducing price $\bar{p}(\pi)$ that makes the consumer indifferent between acquiring the signal and exiting ($V_S = V_L$). This price is given by

$$\bar{p}(\pi) = \frac{\rho\pi v - c}{\Pr(h|\pi)}.$$

At this price, the consumer chooses to acquire the signal, and a purchase occurs only if the signal is h. The seller's expected one-period profit is $\Pi^*_{Induce}(\pi) = \Pr(h|\pi)\bar{p}(\pi) = \rho\pi v - c$.

The learning region exists if and only if $\tilde{p}(\pi) < \bar{p}(\pi)$, which is equivalent to $c < \pi(1-\pi)v(2\rho - 1)$. If this condition fails, for example when the search cost c is too high or the belief π is too close to zero or one, no price can induce learning. In such cases, the myopic seller's optimal policy is to set the highest price at which the consumer is indifferent between buying and exiting. The following proposition summarizes the myopic seller's dynamic pricing strategy:

PROPOSITION 1 (MYOPIC SELLER) *For a myopic seller with belief π:*

- *If $c \geq \pi(1-\pi)v(2\rho-1)$, no learning region exists. The seller sets $p = \pi v - c$ to extract all surplus from a direct purchase.*
- *If $c < \pi(1-\pi)v(2\rho-1)$, a learning region exists:*
  - *If $\frac{(1-\rho)\pi v + c}{\rho(1-\pi)+(1-\rho)\pi} \geq \rho\pi v - c$, the seller sets the learning-deterring price $\tilde{p} = \frac{(1-\rho)\pi v + c}{\rho(1-\pi)+(1-\rho)\pi}$, and the consumer buys directly.*
  - *Otherwise, the seller sets the learning-inducing price $\bar{p} = \frac{\rho\pi v - c}{\rho\pi+(1-\rho)(1-\pi)}$. The consumer acquires a signal and purchases only if she receives a good signal ($s = h$).*

2.3.2 The Forward-Looking Seller

We now turn to the problem of a forward-looking seller who maximizes the discounted sum of future profits, with discount factor $\beta \in (0,1)$. Let $W(\pi)$ denote the seller's value function. Define $\mathcal{C}$ as the space of bounded and continuous functions on the belief space. The Bellman operator T maps a function $W \in \mathcal{C}$ into a new function TW as follows:

$$(TW)(\pi) = \max \left\{ \underbrace{\Pi^*_{Deter}(\pi) + \beta W(\pi)}_{\text{Value of Deterring}}, \quad \underbrace{\Pi^*_{Induce}(\pi) + \beta E(\pi; W)}_{\text{Value of Inducing}} \right\},$$

where $E(\pi; W) = \Pr(h|\pi)W(\pi_h(\pi)) + \Pr(l|\pi)W(\pi_l(\pi))$ denotes the expected continuation value if the consumer acquires a signal.

It is straightforward to verify that the Bellman operator T satisfies Blackwell's conditions for a contraction mapping, namely monotonicity and discounting. By the Banach Fixed-Point Theorem, T therefore admits a unique fixed point $W(\pi) \in \mathcal{C}$, which is the seller's value function. Moreover, this fixed point can be obtained by iterating T starting from any bounded initial function W_0. This guarantees that the value function exists, is unique, and is independent of the choice of initial condition in the iterative procedure.

The main difference between the myopic and the forward-looking seller lies in how they value the distribution of future public beliefs. A key step in the analysis is the following lemma.

LEMMA 2 *The forward-looking seller's value function $W(\pi)$ is a convex function of the public belief π.*

Proof. The value function $W(\pi)$ has a support on $[0,1]$, and must be bounded by $[0, \frac{v}{1-\beta}]$. Let $\mathcal{C}$ denote the space of bounded continuous functions on $[0,1]$, equipped with the sup-norm $\|f\| = \sup_{\pi\in[0,1]} |f(\pi)|$. This space is a Banach space and hence a complete metric space. The Bellman operator $\mathcal{T}$ is defined as

$$(\mathcal{T}W)(\pi) = \max\left\{\Pi^*_{\text{Deter}}(\pi) + \beta W(\pi),\ \Pi^*_{\text{Induce}}(\pi) + \beta\mathbb{E}(\pi; W)\right\},$$

where $\mathbb{E}(\pi; W) = \Pr(h|\pi)W(\pi_h(\pi)) + \Pr(l|\pi)W(\pi_l(\pi))$. We first verify that $\mathcal{T}$ maps $\mathcal{C}$ into itself. For any $W \in \mathcal{C}$, since $\Pi^*_{\text{Deter}}(\pi)$, $\Pi^*_{\text{Induce}}(\pi)$, $\Pr(h|\pi)$, $\Pr(l|\pi)$, $\pi_h(\pi)$, and $\pi_l(\pi)$ are all continuous in π, and W is continuous and bounded, both branches of the maximand are continuous and bounded in π. The pointwise maximum of two continuous functions is continuous, so $\mathcal{T}W \in \mathcal{C}$.

We next verify that $\mathcal{T}$ satisfies Blackwell's sufficient conditions for a contraction mapping (Blackwell, 1965):

- *Monotonicity:* If $W_1(\pi) \le W_2(\pi)$ for all $\pi \in [0,1]$, then both $\Pi^*_{\text{Deter}}(\pi) + \beta W_1(\pi) \le \Pi^*_{\text{Deter}}(\pi) + \beta W_2(\pi)$ and $\Pi^*_{\text{Induce}}(\pi) + \beta\mathbb{E}(\pi; W_1) \le \Pi^*_{\text{Induce}}(\pi) + \beta\mathbb{E}(\pi; W_2)$, where the second inequality follows from $\mathbb{E}(\pi; \cdot)$ being a positive linear operator. Taking the pointwise maximum, $(\mathcal{T}W_1)(\pi) \le (\mathcal{T}W_2)(\pi)$.

- *Discounting:* For any constant $a \ge 0$, $\mathcal{T}(W + a)(\pi) \le (\mathcal{T}W)(\pi) + \beta a$, since adding a to W increases each branch of the maximand by at most βa.

By Blackwell's theorem, $\mathcal{T}$ is a contraction mapping on $(\mathcal{C}, \|\cdot\|)$ with modulus β. By the Banach Fixed-Point Theorem, $\mathcal{T}$ has a unique fixed point $W^* \in \mathcal{C}$, and the iterates $W_{k+1} = \mathcal{T}W_k$ converge uniformly to W^* from any initial function $W_0 \in \mathcal{C}$.

We now show that W^* is convex. For any convex value function $W_k(\pi) : [0,1] \to \mathbb{R}$, $W_{k+1}(\pi) = \max\{A(\pi), B(\pi)\}$, where $A(\pi) = \Pi^*_{Deter}(\pi) + \beta W_k(\pi)$, and $B(\pi) = \Pi^*_{Induce}(\pi) + \beta E(\pi; W_k)$. As the pointwise maximum of convex functions is convex, it suffices to show that both $A(\pi)$ and $B(\pi)$ are convex in π.

First, $A(\pi)$ is the sum of two functions. By the inductive assumption, $W_k(\pi)$ is convex. For

$\Pi^*_{Deter}(\pi)$, which is given by $\tilde{p}(\pi) = \frac{(1-\rho)\pi v + c}{\pi(1-\rho)+(1-\pi)\rho}$, the second-order derivative is:

$$\frac{d^2\tilde{p}(\pi)}{d\pi^2} = \frac{-2(1-2\rho)\left[(1-\rho)v\rho - c(1-2\rho)\right]}{((1-2\rho)\pi+\rho)^3}.$$

Because $\rho \in (\frac{1}{2}, 1)$ and $\pi \in [0,1]$, it is easy to verify that $\frac{d^2\tilde{p}(\pi)}{d\pi^2} \geq 0$ as both the denominator and the numerator are positive. Hence, $\Pi^*_{\text{Deter}}(\pi)$ is convex, and therefore $A(\pi)$ is convex.

Next, $B(\pi)$ is also the sum of two functions. The term $\Pi^*_{Induce}(\pi) = \rho\pi v - c$ is linear and thus weakly convex. For the continuation term $\beta E(\pi; W_k)$, take any $\pi_a, \pi_b \in (0,1)$ and $\alpha \in (0,1)$, and let $\pi_c = \alpha\pi_a + (1-\alpha)\pi_b$. We want to show that

$$E(\pi_c; W_k) \leq \alpha E(\pi_a; W_k) + (1-\alpha)E(\pi_b; W_k).$$

Since $\Pr(s|\pi)$ is linear in π, we have

$$\Pr(s|\pi_c) = \alpha \Pr(s|\pi_a) + (1-\alpha)\Pr(s|\pi_b), \quad s \in \{h, l\}.$$

Moreover, the posterior $\pi_h(\pi_c)$ can be expressed as a convex combination of other posteriors:

$$\pi_h(\pi_c) = \frac{\alpha \Pr(h|\pi_a)}{\Pr(h|\pi_c)}\pi_h(\pi_a) + \frac{(1-\alpha)\Pr(h|\pi_b)}{\Pr(h|\pi_c)}\pi_h(\pi_b).$$

Define $\gamma_h = \frac{\alpha \Pr(h|\pi_a)}{\Pr(h|\pi_c)} \in [0,1]$. Then $\pi_h(\pi_c) = \gamma_h\pi_h(\pi_a) + (1-\gamma_h)\pi_h(\pi_b)$. A similar representation holds for $\pi_l(\pi_c)$ with γ_l.

By the convexity of W_k, we obtain

$$W_k(\pi_s(\pi_c)) \leq \gamma_s W_k(\pi_s(\pi_a)) + (1-\gamma_s)W_k(\pi_s(\pi_b)), \quad s \in \{h, l\}.$$

Substituting all these expressions into $E(\pi_c; W_k)$, we have

$$
\begin{aligned}
E(\pi_c; W_k) &= \Pr(h|\pi_c)W_k(\pi_h(\pi_c)) + \Pr(l|\pi_c)W_k(\pi_l(\pi_c)) \\
&\le Pr(h|\pi_c)[\gamma_h W_k(\pi_h(\pi_a)) + (1-\gamma_h)W_k(\pi_h(\pi_b))] \\
&\quad + Pr(l|\pi_c)[\gamma_l W_k(\pi_l(\pi_a)) + (1-\gamma_l)W_k(\pi_l(\pi_b))] \\
&= \alpha\big[\Pr(h|\pi_a)W_k(\pi_h(\pi_a)) + \Pr(l|\pi_a)W_k(\pi_l(\pi_a))\big] \\
&\quad + (1-\alpha)\big[\Pr(h|\pi_b)W_k(\pi_h(\pi_b)) + \Pr(l|\pi_b)W_k(\pi_l(\pi_b))\big] \\
&= \alpha E(\pi_a; W_k) + (1-\alpha)E(\pi_b; W_k).
\end{aligned}
$$

Thus, $E(\pi; W_k)$ is convex if W_k is convex. Consequently, $B(\pi) = \Pi^*_{Induce}(\pi) + \beta E(\pi; W_k)$ is convex. As $W_{k+1}(\pi) = \max\{A(\pi), B(\pi)\}$ is the pointwise maximum of two convex functions, it is also convex. By induction, starting from $W_0 = 0$, $W_k(\pi)$ is convex for all k. As the value function $W(\pi)$ is the limit of the sequence $\{W_k\}$, $W(\pi)$ is convex.

■

Now that we have established the convexity of the value function $W(\pi)$, a direct implication follows: inducing learning generates a positive intertemporal ”information premium” for the seller. Formally, let $\Delta(\pi) = E(\pi; W) - W(\pi)$. By convexity, $\Delta(\pi) > 0$. This result connects back to the main model: a forward-looking seller has an additional incentive to induce learning, even if it comes at the expense of current-period profits. By creating a more dispersed posterior belief distribution via consumer learning, the seller benefits from the convexity of $W(\pi)$. The comparison between the two seller types can now be summarized in the following proposition:

PROPOSITION 2 (FORWARD-LOOKING SELLER) *A forward-looking seller is more inclined to set a learning-inducing price than a myopic seller. There exists a region of parameters where a myopic seller finds it optimal to deter learning, while a forward-looking seller finds it optimal to induce learning. The size of this region increases with the seller's discount factor β.*

Proof. A myopic seller induces learning if and only if $\Pi^*_{Induce} > \Pi^*_{Deter}$. A forward-looking seller chooses to induce learning if

$$\Pi^*_{Induce} + \beta E(\pi; W) > \Pi^*_{Deter} + \beta W(\pi).$$

Rearranging it, we have

$$\Pi^*_{Induce} + \beta\Delta(\pi) > \Pi^*_{Deter}.$$

where $\Delta(\pi) = E(\pi; W) - W(\pi) > 0$ represents the information premium. As this premium is multiplied by the discount factor β, the parameter region in which learning is induced expands as β increases. ■

Proposition 2 shows that even if inducing learning is not profitable in a static sense, it can be dynamically optimal for the forward-looking seller whenever the short-term loss is outweighed by the long-term information premium. The intuition, consistent with the main model, is that improving the precision of public belief increases the seller's future profits. Consider the extreme case in which the signal is perfectly informative: if a consumer can learn the true state with certainty (signal precision $\rho \to 1$), then after one consumer's search, the public belief jumps to either $\pi = 1$ or $\pi = 0$. Once uncertainty is eliminated, the seller can charge the state-contingent monopoly price and extract all surplus thereafter. The expected profit per period in this case is $\pi_0 v$, since with probability π_0 the product is high value and yields profit v, and with probability $1 - \pi_0$ the low-value product yields zero profit. This payoff strictly exceeds what the seller can obtain under the prior belief π_0, where it either deters learning but must leave information rent to consumers, or induces learning but must compensate them for the learning cost.

2.4 Endogenous Signal Precision

The previous analysis assumed, for tractability, that consumers decide whether to acquire a private signal of exogenous quality ρ. In contrast, a key feature of our main paper's active learning process is that consumers can choose how much effort to devote to learning. A natural analogue in the discrete signal setting is to allow the consumer to endogenously select an effort level $e \geq 0$, which determines the precision of the signal, $\rho(e)$.

Although endogenizing effort brings the discrete signal model closer to the spirit of our main model, it makes the analysis analytically intractable. The seller's optimization problem becomes a nested maximization over price p. For any given price, the seller must anticipate the consumer's optimal effort choice $e^*(p)$, which itself is the solution of an inner optimization problem. This effort level $e^*(p)$ then determines the signal quality $\rho(e^*(p))$, the probability of purchase, and ultimately the seller's profit. The resulting profit function is generally not well-behaved and does not admit a closed-form solution.

To bridge this gap and demonstrate the robustness of our mechanism, we solve this endogenous model numerically. We specify standard, well-behaved functional forms for the consumer's effort cost and the signal precision function:

- A linear cost of effort function: $c(e) = k_c \cdot e$, where k_c is the consumer's marginal cost of effort.
- A signal precision function: $\rho(e) = 1 - \frac{1}{2}\exp(-\alpha e)$, where α measures the effectiveness of learning. This functional form is chosen for its standard and economically intuitive properties: (i) with zero effort ($e = 0$), the signal is uninformative ($\rho(0) = \frac{1}{2}$); (ii) as effort grows large ($e \to \infty$), the signal becomes perfectly informative ($\rho(e) \to 1$), which guarantees that any discrete signal with finite effort remains imperfect; and (iii) the improvement in signal quality exhibits diminishing marginal returns to effort.

With these specifications, we solve the infinite-horizon model using value function iteration for both a myopic seller and a forward-looking seller. In parallel with Figure 3 in the main paper, we compute the optimal strategy for each seller type across two parameters: the consumer's learning cost, k_c, and the product's high-state value, v. The remaining parameters of the simulation are fixed at representative values: the prior belief is $\pi = 0.45$, the learning effectiveness is $\alpha = 3.0$, and the seller's discount factor is $\beta = 0.9$.

The results of this numerical simulation are presented in Figure I. The figure shows three distinct strategy regions: one where both seller types deter learning (white), one where both induce learning (dark grey), and an intermediate region where the forward-looking seller finds it optimal to induce learning while the myopic seller deters it (light grey). These results confirm that the main insight of our paper remains valid under the inconclusive discrete signal framework with endogenous precision.

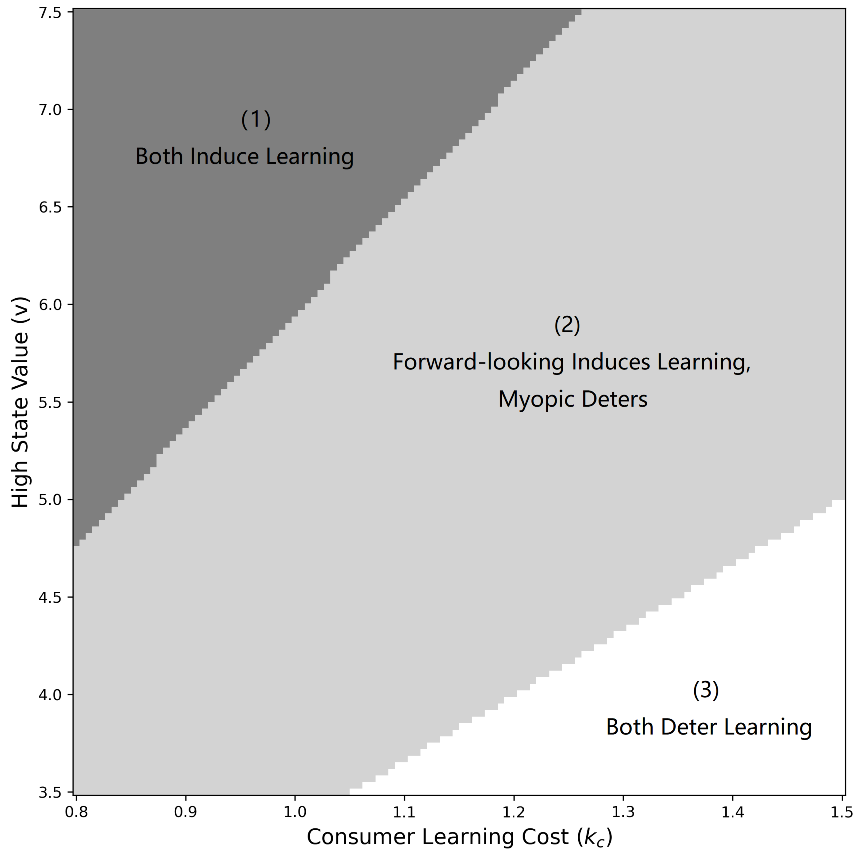

Figure I: Forward-Looking vs. Myopic Seller Pricing under Inconclusive Discrete Signal

2.5 Conclusion

In this online appendix, we considered an alternative information structure to connect more directly with the observational learning literature. Specifically, we modeled consumer learning as the acquisition of a costly private signal. We first analyzed the case in which consumers choose whether to acquire a signal of exogenous quality, and then extended the model to allow consumers to endogenously determine signal precision through costly effort. Across both settings, we find that our paper's central result continues to hold: a forward-looking seller values the information premium generated by early consumer learning and thus has stronger incentives to induce learning than a myopic seller. The consistency of this result across the Poisson framework, the inconclusive discrete signal model, and the extension of discrete signals with endogenous precision demonstrates that the mechanism we identify is not dependent on a particular modeling device but constitutes a broad finding in markets with costly information acquisition.

3 Appendix C: Brownian Learning

In this appendix, we explore an alternative formulation of gradual learning, following Branco et al. (2012). We consider a continuous-time Brownian-motion search model, where a consumer's expected valuation of a product evolves according to a Brownian motion as the consumer continues searching. This setup allows the unknown state to be continuous rather than binary and captures smooth, gradual (i.e., inconclusive) learning and belief updating. In this way, the Brownian formulation incorporates distinct features of information acquisition that complement our Poisson framework and provides a robustness check of our main findings. We show that our central result, that a forward-looking seller is more inclined than a myopic seller to set a learning-inducing price, remains valid in this setting, underscoring the generality of the economic trade-offs highlighted in the main paper.

3.1 Model Setup

We consider a product with an unknown true value to consumers. An infinite sequence of consumers enters the market sequentially, indexed by $n = 1, 2, \ldots$. The public belief about the product value at the beginning of period n is denoted by v_{n-1}, which summarizes all information revealed through prices and purchase decisions in the first $n-1$ periods. Formally, v_{n-1} equals the expected value of the product given the public history. Consumer n's prior expected valuation of the product equals the public belief v_{n-1}. The consumer may then invest a continuous measure of learning input m to refine her valuation at a flow cost $c > 0$ per unit of input. Following the continuous search model of Branco et al. (2012), we model the evolution of the consumer's expected valuation, denoted by u, during active learning as a standard Brownian motion with zero drift and variance σ^2 per unit of input.[1] The stochastic differential equation for the expected valuation is:

$$du = \sigma dw,$$

where w is a standardized Brownian motion and σ is the standard deviation of the Brownian motion. σ^2 represents the "informativeness" or volatility of the search process; a higher σ^2 implies that new

[1] Branco et al. (2012) interpret the learning process as investigating product attributes, where the learning input m corresponds to the number of attributes examined. They assume that there are infinitely many attributes, each providing an infinitesimal piece of information, and in the limit the updating process of the expected valuation u converges to a Brownian motion. See Branco et al. (2012) for the detailed microfoundation.

information causes larger revisions to the consumer's valuation.

3.2 Two Learning Processes

We analyze the two learning processes beginning with the active learning stage. A consumer who chooses to learn must decide when to stop learning and make a purchase decision. Consider now that the seller sets a price p_n in period n. The optimal stopping rule is characterized by two thresholds on the consumer's net valuation $x = u - p_n$: an upper "purchase" threshold $\overline{U}$ and a lower "exit" threshold $\underline{U}$.

PROPOSITION 1 (OPTIMAL STOPPING RULE AND PURCHASE PROBABILITY) *A consumer with initial net valuation $x_0 = v_{n-1} - p_n$ follows an optimal stopping rule characterized by symmetric thresholds:*

$$\overline{U} = -\underline{U} = \frac{\sigma^2}{4c}.$$

The consumer engages in active learning if $x_0 \in (\underline{U}, \overline{U})$, purchases immediately if $x_0 \geq \overline{U}$, and exits immediately if $x_0 \leq \underline{U}$. If the consumer learns, the probability of eventually purchasing the product is:

$$P_n(v_{n-1}, p_n) = \frac{1}{2}\left(1 + \frac{v_{n-1} - p_n}{\overline{U}}\right).$$

Proof. The stopping rule and purchase probability follow directly from the standard results in Brownian-motion search models (see Branco et al. 2012). We therefore omit the detailed derivation. ■

Next, we turn to the observational learning stage. The following lemma shows that, consistent with our main paper, the next consumer's initial belief (equivalently, the public belief) is updated to the posterior belief of the previous consumer at her stopping point.

LEMMA 1 (BELIEF UPDATING) *If consumer $n+1$ expects consumer n to engage in active learning ($v_{n-1} - p_n \in (\underline{U}, \overline{U})$), then the public belief updates as follows:*

- *If consumer n purchases, then $v_n = p_n + \overline{U}$.*
- *If consumer n exits, then $v_n = p_n - \overline{U}$.*

The belief updating process is a martingale, i.e., $E(v_n | v_{n-1}, p_n) = v_{n-1}$.

Proof. Conditional on the current belief v_{n-1} and price p_n, the expected public belief in the next period is:

$$E[v_n | v_{n-1}, p_n] = P_n(v_{n-1}, p_n) \cdot v_n^{buy} + (1 - P_n(v_{n-1}, p_n)) \cdot v_n^{exit}.$$

Substituting the expressions for the probabilities and updated beliefs yields:

$$\begin{aligned} E[v_n] &= \frac{1}{2}\left(1 + \frac{v_{n-1} - p_n}{\overline{U}}\right)(p_n + \overline{U}) + \frac{1}{2}\left(1 - \frac{v_{n-1} - p_n}{\overline{U}}\right)(p_n - \overline{U}) \\ &= \frac{1}{2\overline{U}}[2v_{n-1}\overline{U}] \\ &= v_{n-1}. \end{aligned}$$

■

3.3 The Seller's Dynamic Pricing Problem

Building on the analysis of consumers' learning and purchase decisions, we now endogenize the price. For consistency with the main model, we assume the seller's marginal production cost is zero. We compare the dynamic pricing strategies of a myopic seller and a forward-looking seller.

PROPOSITION 2 (MYOPIC SELLER) *Given public belief v, a myopic seller chooses between a learning-inducing and a learning-deterring price. The optimal price $p^M(v)$ is*

$$p^M(v) = \begin{cases} 0, & v < -\overline{U} \quad \textit{(no-sale region)}, \\ \dfrac{v + \overline{U}}{2}, & -\overline{U} \le v < 3\overline{U} \quad \textit{(learning-inducing region)}, \\ v - \overline{U}, & v \ge 3\overline{U} \quad \textit{(learning-deterring region)}. \end{cases}$$

Proof. The myopic seller's problem is to choose a price that maximizes single-period profit. This static optimization problem is identical to the one analyzed in Branco et al. (2012). Under zero marginal cost, their results directly yield the pricing rule stated above. ■

Next, we analyze the dynamic pricing strategy of a forward-looking seller with discount factor $\beta > 0$. The key is to show the convexity of the value function. To establish it, we work with a finite-horizon formulation of the seller's problem. This approach is both sufficient and natural for the following reason. In the discrete-signal model of Online Appendix B, the value function is bounded on the compact belief space $[0, 1]$, which allows a direct application of the Banach Fixed-Point Theorem in the space of bounded continuous functions. In the Brownian learning environment, however, the public belief v lies on the real line, and the seller's value function grows without bound as $v \to \infty$ (since the seller can charge an arbitrarily high learning-deterring price when the product's expected value is sufficiently high). This makes the standard contraction mapping argument less straightforward. Instead, we consider a finite-horizon version of the seller's problem with T periods remaining and establish convexity of the value function for every finite T by induction on the Bellman operator. Because the economic comparison between the forward-looking and myopic seller's incentive to induce learning applies at each finite horizon, this approach establishes our main result without requiring infinite-horizon convergence arguments.

LEMMA 2 (CONVEXITY OF VALUE FUNCTION) *For any finite horizon $T \geq 0$, the seller's T-horizon value function $\Pi_T(v)$ is a convex function of the public belief v.*

Proof. Consider a finite-horizon version of the seller's problem with T periods remaining and terminal payoff $\Pi_0(v) = 0$. The value function satisfies the recursion $\Pi_{t+1}(v) = (\mathcal{T}\Pi_t)(v)$, where the Bellman operator $\mathcal{T}$ is defined as

$$(\mathcal{T}\Pi_t)(v) = \max_p \{p \cdot P(v,p) + \beta [P(v,p)\, \Pi_t(p+U) + (1 - P(v,p))\, \Pi_t(p-U)]\},$$

with $P(v,p) = \frac{1}{2}\left(1 + \frac{v-p}{U}\right) \in [0,1]$ and $U = \frac{\sigma^2}{4c}$.

We prove by induction that $\Pi_t(v)$ is convex for all $t \geq 0$. The base case $\Pi_0(v) = 0$ is trivially convex. Suppose Π_t is convex. For any fixed price p, define the maximand

$$J(p, v, \Pi_t) = p \cdot P(v,p) + \beta [P(v,p)\, \Pi_t(p+U) + (1 - P(v,p))\, \Pi_t(p-U)].$$

Since $P(v,p) = \frac{1}{2}\left(1 + \frac{v-p}{U}\right)$ is linear in v, and $\Pi_t(p+U)$ and $\Pi_t(p-U)$ do not depend on v, the function $J(p, v, \Pi_t)$ is linear in v for each fixed p, and hence convex. Therefore,

$$\Pi_{t+1}(v) = \max_p J(p, v, \Pi_t),$$

as the pointwise maximum of a family of convex (in fact, linear) functions indexed by p, is convex. By induction, $\Pi_t(v)$ is convex for every $t \geq 0$. ■

Finally, the following proposition shows that, in line with the main model, a forward-looking seller has stronger incentives to induce consumer learning, even at the expense of current-period profit.

Proposition 3 (Forward-Looking Seller) *With $\beta \in (0, 1)$, a forward-looking seller has stronger incentives to induce consumer learning than a myopic seller. In particular, there exist belief regions in which the forward-looking seller sets a learning-inducing price while the myopic seller would deter learning.*

Proof. The proof compares the total discounted value of the learning-inducing and learning-deterring strategies for a seller with a finite horizon of T periods remaining.

Case 1: Myopic seller induces learning ($v < 3U$). In this region, the myopic seller finds it optimal to induce learning based solely on single-period profits. A forward-looking seller additionally values the continuation payoff. By Lemma 2, $\Pi_t(v)$ is convex for every t. Since learning generates a non-degenerate lottery over future beliefs (the public belief moves to either $p + U$ or $p - U$, which is a mean-preserving spread relative to the current belief v), Jensen's inequality implies

$$P(v, p)\,\Pi_t(p + U) + (1 - P(v, p))\,\Pi_t(p - U) \geq \Pi_t(v),$$

for any t, whenever $P(v, p) \in (0, 1)$. This information premium from the belief lottery reinforces the forward-looking seller's incentive to induce learning in this region.

Case 2: Myopic seller deters learning ($v \geq 3U$). Consider the threshold belief $v^* = 3U$, where the myopic seller is indifferent between inducing and deterring learning. At this cutoff, the maximized single-period profits under the two strategies are equal: $\pi_I(v^*) = \pi_D(v^*) = 2U$.

For a forward-looking seller with T periods remaining, if the seller deters learning, the belief does not change, and the value is

$$V_D^T(v^*) = \pi_D(v^*) + \beta \Pi_{T-1}(v^*).$$

If the seller induces learning at price p_I, the value is

$$V_I^T(v^*) = \max_{p_I} \left\{ p_I P(v^*, p_I) + \beta \left[P(v^*, p_I)\, \Pi_{T-1}(p_I + U) + (1 - P(v^*, p_I))\, \Pi_{T-1}(p_I - U) \right] \right\}.$$

By Lemma 2, Π_{T-1} is convex for any $T \geq 1$. Since learning generates a non-degenerate distribution over beliefs while deterring learning keeps the belief fixed at v^*, and since the belief updating process is a martingale (Lemma 1), Jensen's inequality gives

$$P(v^*, p_I)\, \Pi_{T-1}(p_I + U) + (1 - P(v^*, p_I))\, \Pi_{T-1}(p_I - U) > \Pi_{T-1}(v^*),$$

whenever $P(v^*, p_I) \in (0, 1)$. Evaluating at the myopic seller's indifference point where $\pi_I(v^*) = \pi_D(v^*) = 2U$:

$$V_I^T(v^*) \geq 2U + \beta \left[P(v^*, p_I)\, \Pi_{T-1}(p_I + U) + (1 - P(v^*, p_I))\, \Pi_{T-1}(p_I - U) \right] > 2U + \beta \Pi_{T-1}(v^*) = V_D^T(v^*).$$

Therefore, at v^*, the forward-looking seller strictly prefers to induce learning for any horizon $T \geq 1$.

Since $V_I^T(v) - V_D^T(v)$ is continuous in v and strictly positive at $v^* = 3U$, there exists an interval $(v^*, v^* + \varepsilon)$ with $\varepsilon > 0$ such that the forward-looking seller prefers to induce learning even though the myopic seller would deter it. In this region, the information premium from the belief lottery outweighs the static loss from inducing learning. ■

4 Appendix D: Observational Learning from Sales

Our main model follows the classic observational learning framework, assuming that consumers enter the market one by one in sequence, and each consumer observes the public history of previous consumers' purchase decisions. While this framework facilitates a simple and tractable illustration of informational externalities and herd behavior among consumers, one might be concerned that it is not realistic to assume that individual consumers' decisions, particularly non-purchase decisions, are observable, and that markets where consumers make decisions strictly in sequence, with each consumer entering after the previous one has completed their decision, are relatively uncommon.

In this appendix, we extend the model setup to make it applicable to a mass market, where N consumers enter the market in each period.[1] Each consumer has unit demand and can observe the prices and total sales from previous periods. Each consumer's Poisson learning process is independent. Conclusive bad news arrives according to a Poisson process, with arrival rate λ in state L and zero in state H. All other assumptions are the same as in the main model. The following proposition presents the main results under endogenous dynamic pricing.

PROPOSITION 1 *In a market with many consumers per period, a myopic seller's pricing policy remains the same as in Proposition 3 of the main model. When it is optimal for a myopic seller to deter learning by charging a low price, a forward-looking seller may choose to charge a higher price to induce active learning. The higher is β and N, the more likely the seller will do so.*

A detailed proof is provided below. For each consumer, as other consumers' decisions do not affect their utility, their optimal learning behavior remains the same as in the main model. Even though the N consumers may have different learning outcomes after engaging in active learning, they are ex-ante homogeneous, and thus the expected static profit obtained from each of the N consumers is the same and equals the static profit per period derived in the main model. Therefore, the myopic seller's pricing policy remains unchanged.

Furthermore, when at least one out of the N consumers receives conclusive bad news and chooses

[1] We assume that the cohort size is constant across periods. Allowing the market size to vary over time would introduce an additional tradeoff for a forward-looking seller. A larger current cohort increases the static profit loss when the seller sets a price to induce learning rather than immediate purchases, but it also increases the amount of information generated in the market, which improves future belief precision and raises long-run profit. As a result, the impact of market size on dynamic pricing remains ambiguous.

to opt out, the public belief decreases to 0. As a result, with a larger N, there is a higher likelihood that consumers will receive bad news if the true state is L, making the public belief more informative. Thus, for a forward-looking seller, the incentive to induce learning is stronger when N is larger, as the seller can extract more rent with increased informativeness.

[PROOF OF PROPOSITION 1] For consumers' optimal learning behavior and the myopic seller's pricing strategy, the proof follows the preceding discussion directly. Our main focus in this proof is on the forward-looking seller. The only difference between Proposition 1 and Proposition 3 in the main model lies in the comparative statics of N. The proof for other statements is analogous to the one presented in the main paper, and we will not repeat it here.

If a consumer in period n with a prior belief π_{n-1} engages in active learning when facing price p_n, the consumer stops learning when posterior belief reaches $1-\frac{c}{\lambda p_n}$. However, this no longer serves as the prior belief of consumers in period $n+1$ because other $N-1$ consumers in period n also have engaged in active learning. We denote the public belief updated after period n by $\pi_n^{h=1}(N)$, conditional on all past consumers having made a purchase. We have

$$\pi_n^{h=1}(N) = \frac{\pi_{n-1}^{h=1}(N)}{\pi_{n-1}^{h=1}(N) + (1-\pi_{n-1}^{h=1}(N))\left(\frac{\frac{\pi_{n-1}^{h=1}(N)}{1-\frac{c}{\lambda p_n}} - \pi_{n-1}^{h=1}(N)}{1-\pi_{n-1}^{h=1}(N)}\right)^N}.$$

Since we assume that the price p_n is learning inducing, we have $\pi_{n-1}^{h=1}(N) < 1-\frac{c}{\lambda p_n}$. Therefore, $\pi_n^{h=1}(N) > \pi_{n-1}^{h=1}(N)$, and is increasing in N. It is easy to verify that given p_n and N, $\pi_n^{h=1}(N)$ increases with $\pi_{n-1}^{h=1}(N)$. Thus, given any sequence of prices $p_1, p_2, p_3, \ldots$, we have $\pi_n^{h=1}(N) \geq \pi_n^{h=1}(N'), \forall n, \forall N > N'$.

Finally, we show that if the seller finds it optimal to induce learning by N' consumers per period, the seller will induce learning by N consumers, $\forall N > N'$. Similar to Equation (12) in the main model, the expected profit Π_n from period n is

$$\Pi_n(p_n|\pi_{n-1}^{h=1}(N)) = \begin{cases} N\frac{\pi_0}{1-\frac{c}{\lambda p_n}}p_n & \text{if } \bar{p}(\pi_{n-1}^{h=1}(N)) \geq p_n > \tilde{p}(\pi_{n-1}^{h=1}(N)), \\ Np_n & \text{if } p_n \leq \tilde{p}(\pi_{n-1}^{h=1}(N)), \\ 0 & if p_n > \bar{p}(\pi_{n-1}^{h=1}(N)). \end{cases}$$

And the total profit, given the price plan $\{p_n\}$, is

$$\Pi(\{p_n\}) = \sum_{n=1}^{\infty} \beta^{n-1}\Pi_n(p_n|\pi_{n-1}^{h=1}(N)).$$

From the expression, we find that beliefs $\pi_n^{h=1}(N)$ do not enter the profit function directly but rather affect the two thresholds, $\tilde{p}$ and $\bar{p}$. Since $\bar{p}(\pi)$ and $\tilde{p}(\pi) = \frac{c}{\lambda(1-\pi)}$ are increasing in π, $\forall N > N'$, when $p_n \in (\tilde{p}(\pi_{n-1}^{h=1}(N')), \tilde{p}(\pi_{n-1}^{h=1}(N)))$ or $p_n \in (\bar{p}(\pi_{n-1}^{h=1}(N')), \bar{p}(\pi_{n-1}^{h=1}(N))]$, we have $\frac{1}{N}\Pi_n(p_n|\pi_{n-1}^{h=1}(N)) > \frac{1}{N'}\Pi_n(p_n|\pi_{n-1}^{h=1}(N'))$. Otherwise, $\frac{1}{N}\Pi_n(p_n|\pi_{n-1}^{h=1}(N)) = \frac{1}{N'}\Pi_n(p_n|\pi_{n-1}^{h=1}(N'))$. In other words, for any given period and any given price, a seller with a higher number of consumers per period has a higher expected profit per consumer.

If a forward-looking seller chooses to induce learning in a market with N' consumers by setting prices $\{p_n^*(N')\}$, it implies that the profit per consumer gained is greater than the profit gained when prices are learning-deterring. Then, even if the seller in a market with N consumers also charges $\{p_n(N)\} = \{p_n^*(N')\}$, it should yield a weakly higher profit compared to the profit obtained with learning-deterring prices. Moreover, $\{p_n^*(N')\}$ is unnecessarily the optimal price plan that induces learning for the market with N consumers. Hence, if a seller with N' consumers per period would induce learning, a seller with $\forall N > N'$ consumers per period would also induce learning.

■

References

David Blackwell. Discounted dynamic programming. *The Annals of Mathematical Statistics*, 36(1): 226–235, 1965.

Fernando Branco, Monic Sun, and J Miguel Villas-Boas. Optimal search for product information. *Management Science*, 58(11):2037–2056, 2012.