A Model of Two Learning Processes

Jin Huang[†] and Zheng Gong[‡]

[†]NYU Shanghai

[‡]CUHK-Shenzhen

Abstract

Many studies have shown that consumers, before deciding whether to purchase a new product or service, draw inferences from the choices of other consumers and actively acquire information from other sources. We propose a novel model that integrates two learning processes: observational learning and active learning. Building upon the classic observational learning framework, this model allows each consumer to make dynamic choices of information acquisition. We first analyze consumers' learning behavior in the presence of a given price, and then we endogenize the seller's dynamic pricing strategy in response to the two learning processes of consumers. We show that a forward-looking seller may find it optimal to sacrifice short-term profits by setting a higher price to induce active learning, thereby improving the information transmitted through observational learning and ultimately gaining higher expected future profits. We also investigate consumers' learning behavior, the seller's dynamic pricing strategy, and long-run market learning outcomes when the speed of information acquisition increases with sales.

Keywords: observational learning; information acquisition; dynamic pricing; word of mouth

1 Introduction

When new products and services are introduced to the market, both consumers and sellers may experience uncertainty regarding the product's fit and its value for consumers. Various approaches exist for consumers to gather information about new products before making a purchase decision. Empirical research has consistently shown that consumers often learn through the observation of others' choices or by consulting popularity metrics, such as total/monthly sales and best-seller badges, in a Bayesian-rational manner (see a survey by Sorensen et al. (2017)). Another complementary approach for acquiring information involves investing time and effort in processes such as reading professional reviews and critiques, examining product manuals, or engaging in showrooming (Eliashberg and Shugan, 1997; Reddy et al., 1998; Chen and Xie, 2005; Bar-Isaac and Shelegia, 2023). The literature refers to the process of drawing quality cues from others' actions as "observational learning" (Banerjee, 1992; Bikhchandani et al., 1992), and we will use the term "active learning" to denote the costly information acquisition process. This paper focuses on markets where both observational learning and active learning play significant roles in shaping consumer purchase decisions.[1]

Given these insights into consumer behavior, one might ask: How should a seller dynamically price in such markets? Considering consumers' option to engage in costly information acquisition, price plays a dual role: extracting rent from purchases and serving as an incentive tool to manipulate consumers' active learning behavior. Furthermore, choosing a price not only influences the active learning of current consumers but also has an intertemporal impact on the information acquisition and purchase decisions of all future consumers due to information externalities through observational learning. The main objective of this paper

[1] See Online Appendix B for a motivating example. By exploiting an exogenous shock that alters consumers' search conditions, the example provides suggestive evidence that sales information influences consumers' purchase decisions, and search also plays a role. Consumers tend to rely more on observational learning when the opportunity cost of search is high, and more on their own search when the informativeness of predecessors' decisions is low.

is to provide a novel framework that integrates observational learning and active learning and, based on this framework, to examine a monopolist seller's dynamic pricing strategy in markets with both learning processes.

We first introduce the benchmark model to characterize consumers' two learning processes when they face an exogenously given fixed price. The model builds upon the classic observational learning framework, where each consumer infers private information owned by previous consumers by observing their actions. While the standard assumption in the literature is that each consumer has a private signal about the state of the world, we allow consumers to dynamically decide on information acquisition under a Poisson signal structure in a continuous-time setting. With an exogenous price, the herd behavior described in the classic literature arises immediately, starting with the second consumer. This occurs because, when each consumer can endogenously choose the optimal time to learn, the second consumer lacks the incentive to acquire more information than that inferred from the first consumer through observational learning. This holds true for all future consumers as well. As a result, a herd emerges immediately, with a strictly positive probability of being incorrect.

We then endogenize the price consumers face, focusing particularly on the dynamic pricing strategy of a monopolist seller. The seller shares the same prior belief as consumers about how well the product fits their needs. Like new consumers, the seller observes the purchase decisions of previous consumers and updates its belief accordingly. The seller can dynamically adjust the price, and past prices are observable to every new consumer entering the market. In this game, prices play two roles for the seller. First, for each period, a price extracts rent from the current consumer. Second, a price influences the current consumer's active learning behavior, thereby affecting the beliefs of future consumers through observational learning and the expected future profits of the seller. In any period, inducing active learning allows for charging a higher price compared to inducing a direct purchase, but it comes with the risk of consumers receiving bad news or not receiving any good news. Thus, whether the seller prefers to choose a price that induces active learning is not obvious. To highlight the intertemporal effect of pricing, we compare the optimal dynamic pricing strategy of a forward-looking seller, who is concerned about future profits, with that of a myopic seller, who focuses on current profits. One of our key findings is that a forward-looking seller is more inclined to charge higher prices to induce consumers to engage in active learning. Though this approach sacrifices immediate payoffs for the seller, the increased learning reveals more information about the product to later consumers, ultimately benefiting the seller in the long run.

The tractability of the model allows us to compare the probability of an incorrect herd in different settings (e.g., conclusive bad news and conclusive good news environments, exogenous and endogenous pricing). Our findings suggest that endogenous pricing has a double-edged effect on market learning outcomes: in the long run, the market correctly identifies high-value products but may fail to eliminate low-value products. However, the negative impact of endogenous pricing on market learning outcomes can be mitigated if active learning involves taking the time to learn from various forms of word-of-mouth communication, such as user reviews. Consider, as product sales increase, new consumers can receive informative feedback more quickly from the product users. To formalize this idea, we consider an extension in which the speed of information acquisition increases with sales and find that in this case, the market will eventually achieve complete learning.

The paper unfolds as follows. Section 2 summarizes the related literature. Section 3 describes the framework of the two learning processes and analyzes consumers' learning behavior in the presence of a given price. Section 4 examines endogenous dynamic pricing by the seller and its impacts on market learning outcomes. Section 5 considers the case where the news arrival rate grows with the number of purchases. Finally, Section 6 discusses and concludes the paper. Proofs are deferred to the appendix, and additional model analysis and extensions can be found in the online appendix.

2 Literature Review

Our model builds on the standard observational learning framework introduced by Banerjee (1992) and Bikhchandani et al. (1992). In this framework, Bayesian agents with private information about the state of the world arrive sequentially at the market and can observe the decisions made by their predecessors. The standard observational learning papers assume that private information is costless and exogenous. A few papers have considered the endogenous choice of information acquisition by consumers. For example, Hendricks et al. (2012) examine a scenario where each consumer observes the aggregate purchases of predecessors and decides whether to pay a cost for a private signal (i.e., search). They assume that consumers must search before making a purchase. Mueller-Frank and Pai (2016) study a setting in which agents view the actions taken by predecessors and acquire information about finite samples of predecessors' actions and payoffs. The model by Ali (2018) nests the standard observation learning framework and assumes that each agent chooses whether to acquire a costly signal and how informative the signal is. Our paper differs from the three papers mentioned above in two ways. First, we utilize Poisson learning in continuous time to model endogenous information acquisition. This approach builds upon the strategic experimentation literature (Bolton and Harris, 1999; Keller et al., 2005; Keller and Rady, 2010, 2015) and allows us to tractably characterize belief evolution and the probability of an incorrect herd. Second, and more importantly, we study the seller's dynamic pricing strategy, which can yield rich managerial implications.

Starting with Welch (1992), there is research that explores firms' pricing strategies in the presence of observational learning. The pricing considered in Welch (1992) is static. Caminal and Vives (1996) were the first to examine dynamic pricing by firms in a two-period model. In their model, second-period consumers update their beliefs about two competing products based on the products' market shares in the previous period. Bose et al. (2006) study dynamic pricing by a monopolist in the standard observational learning setting, and Bose

Paper	Observational Learning	Active Learning	Dynamic Pricing
Banerjee (1992)	✓		
Bikhchandani et al. (1992)	✓		
Hendricks et al. (2012)	✓	✓	
Mueller-Frank and Pai (2016)	✓	✓	
Ali (2018)	✓	✓	
Caminal and Vives (1996)	✓		✓
Bose et al. (2006, 2008)	✓		✓
This paper	✓	✓	✓

Table I: Related Literature: Differences in the Topics Covered

et al. (2008) further provide a full characterization of the dynamic pricing policy for the case with binary signals. Differing from them, we allow consumers to actively acquire information about the product alongside observational learning, enriching the strategic interaction between the seller and consumers. To the best of our knowledge, this paper is the first to combine observational learning, active learning, and dynamic pricing. Table I summarizes the differences between this paper and the most closely related research.

Finally, in terms of marketing applications, many papers have examined how firms design or manipulate information to influence consumers' information acquisition. For example, Mayzlin and Shin (2011) study how firms' advertising content choice affects consumers' search decisions. Branco et al. (2016) investigate the optimal amount of information firms provide to maximize the likelihood of purchase. Jerath and Ren (2021) explore firms' strategic choices in providing favorable and unfavorable information. While their marketing contexts and focused research questions prevent a direct comparison with this paper, the biggest difference is that we consider the informational externality among consumers, together with consumers' endogenous information acquisition and firm pricing. This paper offers new insights into markets where observational learning appears to be important, such as markets that provide sales information, bestseller lists, and popularity rankings.

3 Benchmark: Two Learning Processes with Exogenous Price

We first present a benchmark model that integrates observational learning and active learning. To model observational learning, we follow the standard setting where consumers arrive at the market sequentially, and each consumer observes the decisions of all predecessors (Banerjee, 1992; Bikhchandani et al., 1992; Smith and Sørensen, 2000). For active learning, we employ the Poisson learning framework, which is widely used in the literature on experimentation and learning (see a survey by Hörner and Skrzypacz (2017)).

3.1 Model Setting

Consider a newly introduced product in the market, sold at an exogenous price p. The product can possibly offer two values to consumers, corresponding to two states of the world. In the high state (denoted by H), the value of the product is v ($v > 0$). In the low state (denoted by L), the value of the product is, without loss of generality, normalized to zero. An infinite sequence of consumers with homogeneous preferences, indexed as $n = \{1, 2, \dots\}$, enters the market, one in each period. Every consumer is interested in making a once-for-all purchase. The true state is initially unknown to all consumers, and they share a common prior belief where the probability of the state being H is π_0, and the probability of the state being L is $1 - \pi_0$. Opting not to purchase the product results in zero utility, while the expected utility from buying depends on the consumer's belief about the state of the world.

Before making a purchase decision, consumers try to learn about the true state. Their learning journey consists of two phases: initially, each consumer observes the purchase decisions of all predecessors (*observational learning*); subsequently, they decide how long to search for information from an information source without observing the search actions and outcomes of others (*active learning*). We employ the Poisson learning framework in a continuous time setting to model the active learning process, during which consumers have the option at each time point to either continue learning, stop learning and make a purchase, or stop learning and exit the market. The cost of learning follows a linear function of time: ct, where c represents the cost per unit of time.

Information Source The market has an information source that sends signals to consumers according to a *Poisson process* with an exogenous arrival rate of λ. We assume that the information source is biased, either it sends Poisson signals only in state L, in which case a signal conclusively reveals that the true state is L and the signal is called conclusive bad news (*breakdown* news); or it sends Poisson signals only in state H, in which case a signal conclusively reveals that the true state is H and thus is called conclusive good news (*breakthrough* news).[2] For different product categories, market information may be dominated by breakdown news or breakthrough news. For example, when a new high-tech product or medical innovation is introduced to the market, potential consumers may wait to see if any serious adverse events occur before adopting the product. In a different market, such as low-budget independent films, the media only covers films that win awards at film festivals, and consumers typically watch the films only after hearing about the breakthrough news.[3]

As the steps and techniques for solving the model are similar in both the conclusive bad news and conclusive good news settings, we focus on addressing the conclusive bad news scenario in the paper (i.e., the Poisson signal arrival rate is λ if the state is L and zero if the state is H). Detailed analysis and proofs for the case of conclusive good news are provided in Online Appendix A. Any discrepancies in market learning outcomes between the two news environments will be highlighted in the paper. We will now proceed to solve the model using backward induction, starting with active learning

[2] It is technically equivalent to assume that an information source sends two types of signals: a relatively weak signal that occurs in both states and a rare, sufficiently strong signal that occurs in only one state. When a strong signal arrives, it conclusively reveals its corresponding state, causing a jump in consumers' belief. In the absence of a strong signal, the weak signals continuously arrive, which leads to a continuous belief updating based on Bayes' rule.

[3] A number of papers on learning focus on Poisson processes with conclusive news (Keller et al., 2005; Keller and Rady, 2015; Che and Mierendorff, 2019). When there are different sources sending both breakthroughs and breakdowns, consumers never update their beliefs in the absence of news. Che and Mierendorff (2019) discuss a setting where agents can optimally allocate attention across two Poisson information sources that generate opposite conclusive news.

3.2 Active Learning

Assume that after observational learning, a consumer's belief that the state is H is updated to π. This π serves as the prior belief at the beginning of the active learning process. We will now analyze the evolution of belief during active learning.

Evolution of belief In the market with conclusive bad news, the consumer's belief that the state is H jumps from π to 0 upon the arrival of a signal. In the absence of a signal, the belief is continuously updated in the direction towards 1. To see this, let x be the time of the first arrival of the signal. At any t, the updated belief π_t is:

$$\pi_t = \Pr(H \mid x > t) = \frac{\pi}{\pi + (1-\pi)e^{-\lambda t}}. \tag{1}$$

It is clear that in the absence of a signal, belief updating is continuous, with the consumer's belief increasing over time, indicating that "no news" is indeed "good news." Next, we present the optimal stopping rule in the active learning process.

Optimal stopping rule Intuitively, if a consumer has initiated active learning, opting out in the absence of a signal is not optimal. This is because, in the absence of bad news, the consumer's belief will increase over time. The consumer will cease learning in two scenarios: when bad news arrives (leading the consumer to opt out of the market), or when the belief is updated to a sufficiently high level that further learning is no longer worthwhile (leading the consumer to buy immediately). By solving the dynamic optimization problem, our results, summarized in the following proposition, confirm this intuition.

PROPOSITION 1 *In the active learning process, the optimal stopping rule for consumers is: when a signal arrives, stop acquiring information and leave the market without buying; if no signal arrives, continue to acquire information until a stopping time t^* at which the belief reaches a stopping boundary $\pi_{t^*} = 1 - \frac{c}{\lambda p}$, and then buy the product.*

See the appendix for the proof. We find that t^* is the single crossing point when the instantaneous benefit of learning is equal to the instantaneous cost of learning, satisfying the following equation:

$$\overbrace{\underbrace{\lambda\xi(1-\pi_{t^*})}_{\substack{\text{Probability}\\ \text{of signal}}} \quad \times \quad \underbrace{p}_{\substack{\text{Benefit:}\\ \text{Avoid Bad Product}}}}^{\text{Benefit of Learning}} = \overbrace{c\xi.}^{\text{Cost of Learning}} \tag{2}$$

The left-hand side represents the marginal benefit of learning, which arises from the possibility that bad news arrives with conclusive evidence, preventing the consumer from paying p for a zero-value product. The right-hand side represents the marginal cost of learning, the unit cost c. News arrives with probability $\lambda(1-\pi_t)$ in the next instant. The consumer's belief increases as the learning process continues, and thus the expected probability of receiving news in the next instant decreases. As a result, the marginal benefit of learning decreases over time until it equals the marginal cost. Rearranging Equation (2) yields the optimal stopping time t^* such that $\pi_{t^*} = 1 - \frac{c}{\lambda p}$. In Online Appendix A, we follow the same steps to solve for the optimal stopping rule in a market with conclusive good news. Similar to Proposition 1, a consumer terminates the learning process either upon signal arrival or when the belief reaches a stopping boundary.

Thus far, we have derived the optimal stopping rule, assuming that a consumer has initiated active learning. To ensure that active learning occurs, the marginal benefit of learning needs to be higher than the marginal cost of learning at the beginning of the learning process (or, the updated belief at the stopping point must be higher than the prior belief):

$$\pi < 1 - \frac{c}{\lambda p}. \tag{3}$$

Denoting $1 - \frac{c}{\lambda p}$ by l_2^B, we have $\pi < l_2^B$. Additionally, the condition that a consumer's expected utility from active learning exceeds the utility of opting out directly must be met. Let x denote the arrival time of the first signal. We have the following:

$$\pi(v - p - t^*c) + (1 - \pi)[e^{-\lambda t^*}(-p - t^*c) - c(1 - e^{-\lambda t^*})E(x \mid x < t^*, L)] \geq 0, \tag{4}$$

where

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

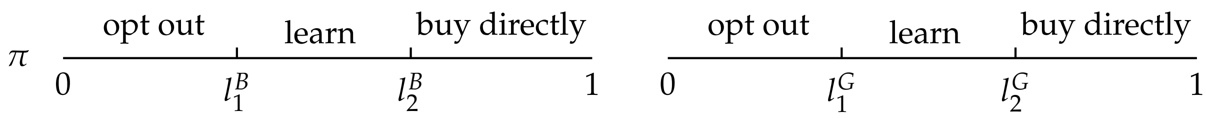
The left-hand side of Inequality (4) represents the expected utility from engaging in active learning. When the true state is H (with probability π), the consumer learns until t^* and buys the product. When the true state is L (with probability $1 - \pi$), the consumer either opts out after receiving bad news before t^* or buys the product if no bad news arrives before t^*. Substituting the expression for π_{t^*} and simplifying the condition, we get:

$$\pi \geq \frac{c}{\lambda(v - p) + c \ln \frac{\pi c}{(1-\pi)(\lambda p - c)}}. \tag{5}$$

The equation uniquely determines the threshold of π. To facilitate exposition, let l_1^B denote the threshold derived from Inequality (5), and thus we must have $\pi \geq l_1^B$. If $l_1^B \geq l_2^B$, the active learning region is empty, meaning that the consumer will either opt out directly or buy the product directly. When $l_1^B < l_2^B$, the following lemma summarizes the inactive learning region and active learning region given belief π:

LEMMA 1 *In a market with conclusive bad 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 $\pi \geq l_2^B$; will engage in active learning if $l_1^B \leq \pi < l_2^B$; will opt out directly if $\pi < l_1^B$.*

Online Appendix A presents the counterpart results when the market is driven by conclusive good news. The figure below illustrates the active and inactive learning regions in the two different news environments.



(a) Conclusive Bad News

(b) Conclusive Good News

Figure I: Active and Inactive Learning Regions in Bad and Good News Environments

3.3 Observational Learning

After studying the active learning of individual consumers, this section focuses on observational learning and examines the long-run market learning outcomes. We first define a few concepts that will appear repeatedly in our analysis: *public belief*, *herd*, and *complete learning*. First, we refer to the updated belief based on past consumers' purchase decisions as the public belief, which is a sufficient statistic summarizing all the information revealed by the purchase history. Second, a herd is defined as action convergence, meaning that all consumers take the same action from a certain point. If a herd forms on "purchase" ("no purchase") in state L (H), we classify it as an incorrect herd. Finally, complete learning means that the public belief converges to the truth. Both *herd* and *complete learning* characterize the long-run learning outcomes.

Let us start with the first consumer. Suppose the market's common prior belief π_0 is such that the first consumer engages in active learning. If the state is H, the probability of purchase is one because a signal will never arrive; if the state is L, the probability of purchase is the probability of no signal arriving before t^*:

$$\Pr(x > t^* \mid L, \pi_0) = e^{-\lambda t^*} = \frac{\pi_0 c}{(1-\pi_0)(\lambda p - c)}. \tag{6}$$

Then, the second consumer observes the first consumer's decision. Let $h = 1$ denote that all consumers in history have purchased, and $h = 0$ denote that at least one consumer in the past chose not to purchase. If the first consumer did not purchase the product, the second consumer easily infers that the first consumer must have received bad news. The second consumer's belief, π_1, is updated to zero after observational learning:

$$\pi_1^{h=0} = 0. \tag{7}$$

Subsequently, all consumers will know that the true state of the world is L, and they will all opt out of the market. If the first consumer purchased the product, the second consumer's belief after observational learning becomes:

$$\pi_1^{h=1} = \Pr(H \mid h = 1, \pi_0) = \frac{\pi_0}{\pi_0 + (1 - \pi_0)e^{-\lambda t^*}}. \tag{8}$$

Substitute Equation (6) into Equation (8), we can get $\pi_1^{h=1} = 1 - \frac{c}{\lambda p}$. Recall that with this belief, the second consumer will buy the product directly without active learning (from Lemma 1). The same goes for the third consumer, the fourth consumer, and so on. There will be no more active learning by individuals, which implies that a herd arises as early as from the second consumer. If the common prior belief π_0 is so low or so high that the first consumer chooses not to engage in active learning, subsequent consumers will do the same, which also implies that a herd arises immediately. Moreover, consumers may herd on the wrong action. In Online Appendix A, we show that the same result is obtained in a market driven by conclusive good news. The following theorem summarizes this finding:

PROPOSITION 2 *For a sequence of consumers with homogeneous preferences, when each consumer is given the option to endogenously optimize information acquisition from an information source, a herd arises immediately after the first consumer. With strictly positive probability, consumers herd on the wrong action (i.e., an incorrect herd).*

The finding of an immediate herd is novel compared to the results of the standard observational learning model. When consumers can make the endogenous choice of information acquisition, the first person follows the optimal stopping rule, searching for information until the marginal benefit of learning equals the marginal cost of learning. The information acquired by the first consumer is transmitted to later consumers through observational learning. As consumer preferences are homogeneous, each consumer faces the same information acquisition problem, and thus later consumers have no incentive to acquire more information than what the first consumer has already obtained.

In terms of learning outcomes, when the true state is H and consumers' prior belief is sufficiently low (lies in the "opt out" region of Figure I), or when the true state is L and consumers' prior belief is sufficiently high (lies in the "buy directly" region of Figure I), all consumers make the incorrect purchase decision because of the lack of learning. Given that the first consumer engages in active learning, an incorrect herd may occur when, in cases where the true state is H (L), no conclusive good (bad) news arrives before the optimal stopping time. The probabilities of an incorrect herd under different true states of the world and information environments are summarized in Table II.

Env. / Truth	Conclusive Bad News	Conclusive Good News
H	If $\pi_0 < l_1^B$, $\Pr = 1$ If $\pi_0 \geq l_1^B$, $\Pr = 0$	If $\pi_0 < l_1^G$, $\Pr = 1$ If $l_1^G \leq \pi_0 < l_2^G$, $\Pr = \frac{(1-\pi_0)c}{\pi_0(\lambda(v-p)-c)}$ If $\pi_0 \geq l_2^G$, $\Pr = 0$
L	If $\pi_0 < l_1^B$, $\Pr = 0$ If $l_1^B \leq \pi_0 < l_2^B$, $\Pr = \frac{\pi_0 c}{(1-\pi_0)(\lambda p-c)}$ If $\pi_0 \geq l_2^B$, $\Pr = 1$	If $\pi_0 < l_2^G$, $\Pr = 0$ If $\pi_0 \geq l_2^G$, $\Pr = 1$

Table II: The Probability of an Incorrect Herd with Exogenous Price

4 Dynamic Pricing with Two Learning Processes

In response to the two consumer learning processes analyzed in the previous section, sellers can strategically employ dynamic pricing to influence the learning incentives and, ultimately, the purchase decisions. In this section, we explore a seller's optimal dynamic pricing strategy to leverage consumer learning for its profit.

We maintain the benchmark model's settings regarding consumers and introduce a monopolistic seller operating infinitely over periods $n = 1, 2, \dots$, discounting future profits with a discount factor of $\beta \in [0, 1)$. The marginal cost of production is normalized to zero. The value of the seller's product to consumers can take on two values, and neither the seller nor consumers know the true value. The seller shares the same prior belief as consumers that the probability of the value being H is π_0. At the beginning of each period, the seller chooses the price p_n, $n = 1, 2, \dots$. The seller aims to maximize its profit: $\Pi(\{p_n\}) = \sum_{n=1}^{\infty} \beta^{n-1} p_n D_n$, where $D_n \in \{0, 1\}$ is a dummy that indicates whether the n-th consumer purchases ($D_n = 1$) or not ($D_n = 0$). In the observation learning phase, consumers can observe the prices set by the seller in previous periods. Similar to the benchmark model, our main analysis focuses on the conclusive bad news case, and a detailed analysis of the conclusive good news case is presented in Online Appendix A.

To begin our analysis, we first characterize several properties of the equilibrium prices. Apparently, in order to incentivize consumers to engage in active learning or make direct purchases instead of leaving the market without buying, a price ceiling exists. This price ceiling represents the price at which consumers expect zero payoff. For any consumer with a prior belief π, let $\bar{p}(\pi)$ denote the price ceiling, we have

LEMMA 2 *$\bar{p}(\pi)$ is continuous and increasing in the prior belief π with $\bar{p}(0) = 0$ and $\bar{p}(1) = v$. When $p = \bar{p}(\pi)$, the consumer's expected surplus is zero.*

A detailed proof is provided in the appendix. A simple explanation is that, as can be clearly seen from Figure I, when consumers hold higher prior beliefs, they are more likely to make direct purchases or engage in active learning rather than opting out. Therefore, compared to a consumer with a lower prior belief, the seller can extract more expected surplus from a consumer with a higher prior belief by charging a higher price until the two consumers face the same tradeoff. Next, we characterize the optimal actions of consumers under different prices in the following lemma.

LEMMA 3 *For a consumer with a prior belief $\pi \in (0,1)$, let $\tilde{p}(\pi) = \frac{c}{\lambda(1-\pi)}$. If $\tilde{p}(\pi) \geq \bar{p}(\pi)$, the consumer will not engage in active learning: she buys directly when $p \leq \bar{p}(\pi)$ and opts out directly when $p > \bar{p}(\pi)$. If $\tilde{p}(\pi) < \bar{p}(\pi)$, we have*

1. *When $p \leq \tilde{p}(\pi)$, the consumer makes a purchase directly. The posterior belief remains π.*

2. *When $\tilde{p}(\pi) < p \leq \bar{p}(\pi)$, the consumer engages in active learning. If no conclusive bad news arrives, the consumer's belief gradually increases, and when the posterior belief reaches $1 - \frac{c}{\lambda p}$, the consumer makes a purchase.*

3. *When $p > \bar{p}(\pi)$, the consumer opts out directly. The posterior belief remains π.*

This lemma is, in fact, an alternative presentation of Lemma 1, with a focus on varying prices instead of prior beliefs, and its proof is immediate from Lemma 1. The condition under which the consumer prefers active learning to buying directly is shown in Equation (3). After rearranging the equation, we can find $\tilde{p}(\pi) = \frac{c}{\lambda(1-\pi)}$ is the cutoff point.[4] Intuitively, the benefits of active learning as opposed to buying directly come from the possibility of discovering that the product is not worth buying. At a lower price, the gain from avoiding a wrong purchase is smaller, and therefore the consumer has a stronger incentive to purchase directly without active learning. When $p \leq \tilde{p}(\pi)$, the price is low enough to not incentivize learning, and the posterior belief remains unchanged. We hereafter refer to such a price as a *learning-deterring* price. When $\tilde{p}(\pi) < p \leq \bar{p}(\pi)$, the consumer engages in active learning, and we hereafter refer to price in this region as a *learning-inducing* price. In the learning region, a higher price leads to a longer learning time. Figure II depicts the three possible pricing regions when $\tilde{p}(\pi) < \bar{p}(\pi)$ and consumers' actions in these regions.

[4]Note that $\bar{p}$ represents two different thresholds, corresponding to two scenarios: one is when $\tilde{p} < \bar{p}$, $\bar{p}$ corresponds to l_1^B in Figure I, which is the point where consumers are indifferent between engaging in active learning and opting out. The other scenario is when $\tilde{p} \geq \bar{p}$, and in this case, there are only two regions: buying directly and opting out. $\bar{p}$ is the threshold between them and thus equal to πv. For ease of exposition, we use $\bar{p}$ to denote the price ceiling in both cases.

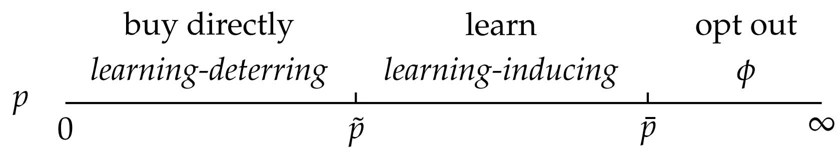


Figure II: Pricing Regions

Now, suppose that the prices are set in a manner that encourages early consumers to engage in learning. In this scenario, late consumers' prior beliefs will increase in the absence of bad news, which creates an opportunity for the seller to charge higher prices as indicated by Lemma 2. Therefore, it is interesting to explore how the seller can strategically employ dynamic pricing to leverage the intertemporal externalities among consumers. Moving forward, we will formally study the seller's optimal pricing strategy, with a particular focus on comparing the optimal pricing between myopic and forward-looking sellers.

4.1 A Myopic Seller's Pricing Problem

First, we consider the problem of a myopic seller with $\beta = 0$ who is essentially solving the static pricing problem in each period independently. Without loss of generality, we analyze the pricing problem of the seller in the first period. Proposition 1 shows that consumer 1 will stop active learning and buy the product when the posterior belief reaches $\pi_1 = 1 - \frac{c}{\lambda p_1}$, where p_1 is the first-period price. As the belief updating process is mean-preserving, the probability that the consumer will eventually buy after learning is $\frac{\pi_0}{\pi_1} = \frac{\pi_0}{1-\frac{c}{\lambda p_1}}$. Thus, the seller's optimization problem at $t = 1$ is

$$\max_{p_1} \Pi_1(p_1) = \begin{cases} \frac{\pi_0}{1-\frac{c}{\lambda p_1}} p_1 & \text{if } p_1 > \tilde{p}(\pi_0), \\ p_1 & \text{if } p_1 \leq \tilde{p}(\pi_0), \end{cases} \tag{9}$$

subject to

$$p_1 \in \left[\min\left\{\bar{p}(\pi_0), \tilde{p}(\pi_0)\right\}, \bar{p}(\pi_0)\right], \tag{10}$$

where $\tilde{p}(\pi_0) = \frac{c}{\lambda(1-\pi_0)}$ is the highest price at which the consumer prefers direct purchase over active learning, as defined in Lemma 3. It is never optimal for a seller to charge a price higher than $\bar{p}(\pi_0)$, which yields zero profit, or to charge a price below $\min\{\bar{p}(\pi_0), \tilde{p}(\pi_0)\}$, as raising the price to the threshold will still result in a direct purchase by the consumer.

The first-order condition of (9) is

$$\frac{\partial \Pi_1(p_1)}{\partial p_1} = \frac{\pi_0(1 - \frac{2c}{\lambda p_1})}{(1 - \frac{c}{\lambda p_1})^2}, \tag{11}$$

which is negative when $p_1 < \frac{2c}{\lambda}$ and positive when $p_1 > \frac{2c}{\lambda}$. This implies that the profit function is U-shaped on p_1, with the turning point at $p_1 = \frac{2c}{\lambda}$ where the posterior belief reaches $\frac{1}{2}$.

In the subsequent periods, the seller faces the same optimization problem as in the first period, except that consumers' prior beliefs might be different due to belief updating after observational learning. The following proposition summarizes the seller's optimal pricing policy, and the proof can be found in the appendix.

PROPOSITION 3 (MYOPIC SELLER) *For a myopic seller with $\beta = 0$, the optimal pricing policy is as follows:*

1. *When $\tilde{p}(\pi_0) \geq \bar{p}(\pi_0)$, the seller sets the price at $\bar{p}(\pi_0)$ for each period. All consumers buy directly without active learning.*

2. *When $\tilde{p}(\pi_0) < \min\left\{\bar{p}(\pi_0), \frac{\pi_0}{1-\frac{c}{\lambda \bar{p}(\pi_0)}}\bar{p}(\pi_0)\right\}$, the seller sets $p_n = \bar{p}(\pi_{n-1})$ for consumer n with a prior belief π_{n-1} $(n = 1, 2, \dots)$. If no bad news ever arrives, each consumer n engages in active learning until their posterior belief reaches $\pi_n = 1 - \frac{c}{\lambda p_n}$. When the first bad news arrives, the public belief becomes zero, and all consumers exit the market without making a purchase.*

3. *When $\frac{\pi_0}{1-\frac{c}{\lambda \bar{p}(\pi_0)}}\bar{p}(\pi_0) \leq \tilde{p}(\pi_0) < \bar{p}(\pi_0)$, the seller sets the price at $\tilde{p}(\pi_0)$ for each period. All consumers buy directly without active learning.*

In cases 1 and 3 of Proposition 3, the seller optimally sets a price in the first period to induce the first consumer to make a direct purchase without active learning. As a result, the public belief remains constant, and the seller will continue to set the same price in all future periods. When the seller's optimal strategy in the first period is to encourage active learning, we find that the learning-inducing pricing policy remains optimal in subsequent periods, as stated in case 2 of the proposition. In this scenario, in the absence of bad news, the public belief monotonically increases over time.

4.2 A Forward-Looking Seller's Pricing Problem

How does the seller's optimal dynamic pricing strategy change when the seller takes future profits into account? To answer this question, we analyze the case in which the seller is forward-looking with $\beta \in (0, 1)$ and compare the results with those from the myopic seller case.

Again, if consumer n with prior belief π_{n-1} engages in active learning at price p_n, the consumer will stop learning when her posterior belief reaches $\pi_n = 1 - \frac{c}{\lambda p_n}$. As the belief updating process is mean-preserving, we can calculate the probability of purchase after active learning for each consumer $n = 1, 2, \ldots$ as follows:

$$\text{For consumer 1,} \quad \Pr = \frac{\pi_0}{\pi_1^{h=1}},$$

$$\text{For consumer 2,} \quad \Pr = \frac{\pi_0}{\pi_1^{h=1}} \times \frac{\pi_1^{h=1}}{\pi_2^{h=1}} = \frac{\pi_0}{\pi_2^{h=1}},$$

$$\ldots$$

$$\text{For consumer } n, \quad \Pr = \frac{\pi_0}{\pi_1^{h=1}} \times \frac{\pi_1^{h=1}}{\pi_2^{h=1}} \times \cdots \times \frac{\pi_{n-1}^{h=1}}{\pi_n^{h=1}} = \frac{\pi_0}{\pi_n^{h=1}},$$

where the superscript $h = 1$ denotes that all previous consumers have purchased in history.

Therefore, the expected profit Π_n from period n is

$$\Pi_n(p_n|\pi_{n-1}^{h=1}) = \begin{cases} \frac{\pi_0}{1-\frac{c}{\lambda p_n}} p_n & \text{if } p_n > \tilde{p}(\pi_{n-1}^{h=1}), \\ p_n & \text{if } p_n \leq \tilde{p}(\pi_{n-1}^{h=1}), \end{cases} \tag{12}$$

And the dynamic pricing problem that the seller faces at the beginning of the game is:

$$\max_{\{p_n\}} \Pi(\{p_n\}) = \sum_{n=1}^{\infty} \beta^{n-1} \Pi_n(p_n|\pi_{n-1}^{h=1}), \tag{13}$$

subject to

$$\pi_n^{h=1} = \begin{cases} 1 - \frac{c}{\lambda p_n} & \text{if } p_n > \tilde{p}(\pi_{n-1}^{h=1}), \\ \pi_{n-1}^{h=1} & \text{if } p_n \leq \tilde{p}(\pi_{n-1}^{h=1}), \end{cases} \tag{14}$$

$$p_n \in \left[\min\left\{\bar{p}(\pi_{n-1}^{h=1}), \tilde{p}(\pi_{n-1}^{h=1})\right\}, \bar{p}(\pi_{n-1}^{h=1})\right], \tag{15}$$

where (14) describes the belief evolution process under the condition that bad news has not yet arrived. If $p_n > \tilde{p}(\pi_{n-1}^{h=1})$, consumer n engages in active learning and thus $\pi_n^{h=1} > \pi_{n-1}^{h=1}$. If $p_n \leq \tilde{p}(\pi_{n-1}^{h=1})$, consumer n finds it optimal to either make a direct purchase or opt out, rather than engaging in active learning, and therefore $\pi_n^{h=1} = \pi_{n-1}^{h=1}$.

Equation (12) shows that the profit obtained in each period depends solely on the price charged during that period. For each period, if the upper bound of the price range is higher, the seller can potentially earn a higher profit. Hence, a forward-looking seller may choose early-period prices strategically to manipulate public beliefs and thereby influence price ranges in subsequent periods. After analyzing all cases, we summarize our findings in the following proposition.

PROPOSITION 4 (FORWARD-LOOKING SELLER) *For a forward-looking seller with $\beta \in (0, 1)$, we have*

1. *When $\frac{\pi_0}{1-\frac{c}{\lambda \bar{p}(\pi_0)}} \bar{p}(\pi_0) \leq \tilde{p}(\pi_0) < \bar{p}(\pi_0)$, unlike the myopic seller, the forward-looking seller*

may find it optimal to set price $p_1 \in (\tilde{p}(\pi_0), \bar{p}(\pi_0)]$ to incentivize the first consumer to choose active learning, and the optimal prices in all subsequent periods are also learning-inducing. The higher the β, the more likely the seller is to do so.

2. *In all other cases, the optimal pricing policy for a forward-looking seller is the same as for a myopic seller.*

A detailed proof of Proposition 4 is provided in the appendix. We find that for a forward-looking seller, similar to a myopic seller, the optimal pricing policy will either be consistently *learning-deterring*, under which all consumers make a purchase directly in equilibrium, or consistently *learning-inducing*, under which all consumers engage in active learning in equilibrium. Due to the analytical complexity, we do not provide a closed-form expression for the parameter regions in which each strategy is chosen. However, we can demonstrate that under certain parameter conditions, the forward-looking seller finds it optimal to induce learning while the myopic seller finds it optimal to deter learning. The intuition for the difference in pricing strategies between the two types of sellers is as follows.

In this model, the seller's profit is constrained by the possibility of consumers choosing active learning over direct purchases. This constraint is reflected in two potential pricing strategies: the seller has to lower the price to encourage direct purchases or, when encouraging direct purchases is too costly, increase the price to a level where consumers slightly prefer active learning over opting out. In the second case, the price should not be set too high, as it needs to compensate for the learning cost. In short, profit per period is influenced by consumer learning incentives. Due to informational externalities across consumers over time, a forward-looking seller can potentially manipulate early learning dynamics to influence future consumers' learning incentives and, thus, future profits. While inducing learning by early consumers hurts early profits, it creates a more informative public belief (either 0 or a higher belief), which decreases the marginal gain and the duration of active learning for later consumers. A lower marginal gain from active learning implies that the seller can charge a higher price while still being able to induce direct purchases, and a shorter learning time implies lower compensation for the total learning cost. In both scenarios, the seller earns a higher expected profit from a more informative public belief. We find similar intuitions and results in the context of conclusive good news, as analyzed in Online Appendix A.

In terms of market learning outcomes, shifting from a learning-deterring strategy to a learning-inducing strategy reduces the probability of an incorrect herd. Therefore, under certain parameter conditions, a forward-looking seller can improve the learning results and thus consumer welfare compared to a myopic seller. The following corollary summarizes the probability of an incorrect herd with respect to different pricing strategies.

COROLLARY 1 (LEARNING OUTCOMES) *When the seller employs the learning-deterring pricing strategy, an incorrect herd occurs only when the true state is L, with probability 1. When the seller employs the learning-inducing pricing strategy, it reduces the likelihood of an incorrect herd, yet complete learning cannot be achieved. The public belief converges to* $\bar{\pi} = \frac{\lambda v + \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$ *in the long run. An incorrect herd occurs only when the true state is L, with probability* $\frac{\pi_0}{1-\pi_0}\frac{1-\bar{\pi}}{\bar{\pi}}$.

The proof is available in the appendix. When the true state is H, there can be no incorrect herd as the seller will never set prices too high to drive away consumers. When the equilibrium pricing strategy is learning-deterring, all consumers buy the product directly, and thus an incorrect herd definitely occurs if the true state is L. When the equilibrium pricing strategy is learning-inducing, a herd does not occur immediately after $t = 1$ as in the benchmark model. However, complete learning is also not achieved. The reason is that, given all past consumers have made purchases, the public belief gradually increases. Nevertheless, consumers will eventually learn for an arbitrarily short time, and the public belief is bounded above. An upper bound on the public belief exists because, with a sufficiently high prior belief, the seller cannot set a price to induce consumers to choose active learning over direct purchase (Inequality (3) cannot hold). Therefore, the public belief is bounded, leading to a positive probability of an incorrect herd in the long run.

In the conclusive good news setting, the learning outcomes remain the same when the true state is H because the seller can ensure that all consumers purchase the product. When the true state is L and the optimal pricing strategy is learning-deterring, the result that an incorrect herd must emerge still applies. The difference compared to the conclusive bad news setting occurs when the true state is L and the pricing strategy is learning-inducing. In the case of conclusive good news, when a consumer engages in active learning, the consumer will only make a purchase if good news arrives; otherwise, the consumer exits the market, resulting in a lower public belief. In the absence of good news, the public belief decreases, and the cost of inducing learning becomes higher for the seller. The seller will eventually adopt the learning-deterring pricing strategy if the public belief becomes low enough. As a result, when the true state is L, under the seller's optimal pricing policy, early consumers may engage in active learning but make no purchases in the absence of news. From a certain point onwards, all consumers purchase the product directly without active learning due to the lowered price. A detailed analysis is provided in Online Appendix A. We summarize the probabilities of an incorrect herd in the two news environments in Table III.

Truth \ Env.	Conclusive Bad News	Conclusive Good News
H	$\Pr = 0$	$\Pr = 0$
L	Define $\bar{\pi} = \frac{\lambda v + \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$ With active learning, $\Pr = \frac{\pi_0}{1-\pi_0}\frac{1-\bar{\pi}}{\bar{\pi}}$ Without active learning, $\Pr = 1$	$\Pr = 1$

Table III: The Probability of an Incorrect Herd with Endogenous Pricing

A Numerical Example We provide a numerical example to illustrate how a forward-looking seller may behave differently from a myopic seller in equilibrium. In Figure III, the left panel illustrates the dynamics of single-period profits for both forward-looking and myopic sellers, and the right panel displays the dynamics of the posterior belief of each period. With the specified parameters, the optimal strategy for the myopic seller (represented by the dashed

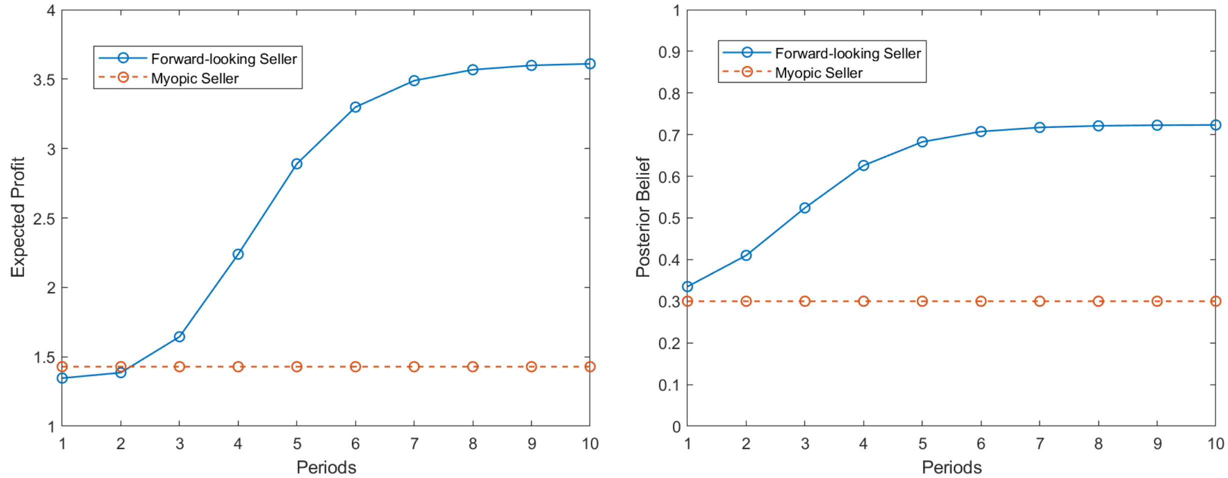


Figure III: Comparison of forward-looking and myopic sellers ($\pi_0 = 0.3, c = 1, v = 5, \lambda = 1$)

orange curve) is to deter learning. As a result, the expected profit remains constant across all periods, and the posterior belief remains unchanged. In contrast, the forward-looking seller (represented by the solid blue curve) finds it optimal to sacrifice early profits at $t = 1, 2$ in exchange for greater gains in later periods, as shown in the left panel. This strategy leads to an increase in the posterior belief and a reduced likelihood of an incorrect herd compared to the myopic seller's case. In line with Corollary 1, the public belief converges to $\frac{5+\sqrt{5}}{10} \approx 0.724$.

5 Increasing News Arrival Rate Over Time

In the previous sections, the information source for Poisson learning is considered static and unaffected by past consumers' behavior, resulting in a fixed speed of information acquisition for all consumers. However, when active learning involves reading detailed customer reviews or consulting past users about the product, it is reasonable to expect that the likelihood of users sharing their experiences will increase as more and more consumers purchase the product, allowing future consumers to learn about the product more rapidly. In this section, we incorporate a dynamic information source that is affected by sales. To capture the dynamics, we specify the signal arrival rate as $f(\lambda, N)$, where λ is a fixed rate, and N

denotes the total number of consumers who have bought the product in the past. In cases of no previous purchases, $f(\lambda, 0) = \lambda$. We assume $f_N > 0$, indicating that news about the product arrives more quickly as sales increase.

5.1 Exogenous Price

We first analyze in detail consumers' learning dynamics in the presence of a given price. To focus on non-trivial cases, let us assume that the parameters fall within the region where the first consumer engages in active learning. If the first consumer chooses not to buy the product, the true state L is perfectly revealed, leading to subsequent consumers exiting the market without making a purchase. If the first consumer buys the product, the public belief is updated to

$$\pi_1^{h=1} = 1 - \frac{c}{\lambda p}.$$

And the signal arrival rate increases to $f(\lambda, 1)$. As $f(\lambda, 1) > f(\lambda, 0) = \lambda$, it follows that $1 - \frac{c}{f(\lambda,1)p} > 1 - \frac{c}{\lambda p}$. Recall the condition on the belief that guarantees active learning, as stated in Lemma 1. We can infer that the second consumer will also engage in active learning due to the increased signal arrival rate. Following this, if the second consumer does not buy the product, the true state L is revealed. If the second consumer buys the product, the public belief is updated to

$$\pi_2^{h=1} = \Pr(H \mid h = 1, \pi_1^{h=1}) = \frac{\pi_1^{h=1}}{\pi_1^{h=1} + (1 - \pi_1^{h=1}) \Pr(X > t^* \mid L, \pi_1^{h=1})}.$$

After simplifying the equation, we obtain that $\pi_2^{h=1} = 1 - \frac{c}{f(\lambda,1)p}$, which is also the posterior belief of the second consumer. Afterward, the third consumer will conduct active learning because the marginal benefit of learning becomes $f(\lambda, 2)(1 - (1 - \frac{c}{f(\lambda,1)p}))p$ according to Equation (2) and it is higher than the marginal cost c. And so on and so forth, the learning dynamics continue in the same pattern.

In terms of market learning outcomes, an incorrect herd occurs only when the true state is H and the prior belief is sufficiently low that no active learning takes place from the first consumer onwards, which is essentially a cold-start problem for marketers. When the true state is L, an incorrect herd cannot form because, after some consumers purchase the product, the signal arrival rate increases, shifting the public belief into the active learning region (both l_2^B and l_2^G in Figure I increase as the signal arrival rate increases). From a certain point onwards, consumers all engage in active learning until someone receives bad news. Therefore, the public belief will almost surely converge to the truth (i.e., complete learning).

With conclusive good news, the above-mentioned results also apply. Furthermore, an incorrect herd occurs if the true state is H and the first consumer engages in active learning but does not receive the good news. In such a scenario, the first consumer opts out, and an immediate herd on "no purchase" occurs due to the unchanged news arrival rate.

5.2 Endogenous Dynamic Pricing

Next, we follow the same procedures as in Section 4 to analyze the seller's optimal pricing strategy when pricing is endogenous, considering both myopic and forward-looking seller scenarios. In line with previous notations, for a given history of N purchases, $\bar{p}(\pi, N)$ denotes the highest price that keeps the current consumer with a prior belief π from opting out, and $\tilde{p}(\pi, N) = \frac{c}{f(\lambda,N)(1-\pi)}$ denotes the price that makes the consumer indifferent between direct purchase and active learning. The following proposition summarizes our findings.

PROPOSITION 5 (PRICING WITH AN INCREASING NEWS ARRIVAL RATE)
In any period when the consumer holds a prior belief π and N previous consumers have purchased the product, we have

1. *A myopic seller follows the same optimal pricing policy as outlined in Proposition 3, except that the news arrival rate changes to $f(\lambda, N)$.*

2. *When* $\frac{\pi}{1-\frac{c}{f(\lambda,N)\bar{p}(\pi,N)}}\bar{p}(\pi,N) \leq \tilde{p}(\pi,N) < \bar{p}(\pi,N)$, *a myopic seller charges* $p = \tilde{p}(\pi,N)$ *to deter learning, whereas a forward-looking seller with* $\beta \in (0,1)$ *may find it optimal to set a higher price* $p \in (\tilde{p}(\pi,N), \bar{p}(\pi,N)]$ *to induce learning.*

3. *In all other cases, the optimal pricing policy for a forward-looking seller is the same as for a myopic seller.*

The proof is provided in the appendix. For a myopic seller, the pricing decisions are made independently for each period. Therefore, the increase in the news arrival rate does not change the optimization problem qualitatively. The seller follows the same pricing policy, adjusting it based on the corresponding news arrival rate. One point to note is that the equilibrium price cannot always be learning-deterring. This is because the learning-deterring price $\tilde{p}(\pi)$ decreases as the news arrival rate increases. Thus, the equilibrium price must induce active learning from some point, whether the seller is myopic or forward-looking.

Moreover, we discover a result consistent with Proposition 4: a forward-looking seller may choose to induce learning when a myopic seller chooses to deter learning. In the absence of bad news, all consumers eventually make purchases. This implies that the news arrival rate in period n is independent of the previous pricing strategy, whether learning-deterring or learning-inducing. As early pricing decisions do not impact future news arrival rates, the previous logic still holds, i.e., the forward-looking seller may choose to induce learning in early periods, which entails sacrificing early profits but ultimately benefits from a more informative public belief in future periods.

In terms of market learning outcomes, when the seller can dynamically set the product's price, complete learning always occurs. An incorrect herd arises when consumers cease learning and/or when belief updating is bounded. Based on the analysis in Section 4, in both H and L states and across both conclusive good news and conclusive bad news environments, as long as no conclusive bad news arrives, consumers will always make purchases in the long run under endogenous pricing. Therefore, as sales increase, the news arrival rate will rise to a level at which no price can incentivize consumers to choose direct purchase over active learning, ultimately resulting in complete learning.[5] The following table and corollary summarize the learning outcomes for all cases discussed in this section.

Pricing	Env. / Truth	Conclusive Bad News	Conclusive Good News
Exogenous	H	If $\pi_0 < l_1^B$, $\Pr = 1$ If $\pi_0 \geq l_1^B$, $\Pr = 0$	If $\pi_0 < l_1^G$, $\Pr = 1$ If $l_1^G \leq \pi_0 < l_2^G$, $\Pr = \frac{(1-\pi_0)c}{\pi_0(\lambda(v-p)-c)}$ If $\pi_0 \geq l_2^G$, $\Pr = 0$
	L	$\Pr = 0$	$\Pr = 0$
Endogenous	H	$\Pr = 0$	$\Pr = 0$
	L	$\Pr = 0$	$\Pr = 0$

Table IV: The Probability of an Incorrect Herd with an Increasing News Arrival Rate

COROLLARY 2 (LEARNING OUTCOMES) *In a market with an increasing and unbounded news arrival rate, at a given price, a low-value product will inevitably be identified and eliminated in the long run; however, there remains a strictly positive probability of incorrectly eliminating a high-value product from the market. When endogenous pricing is employed, the market achieves complete learning, ensuring that all products are correctly identified in the long run.*

6 Discussion and Conclusion

Learning plays a crucial role in various markets, particularly in the context of new products and services. Consumers often acquire information by observing others' choices or by actively seeking detailed product information. This paper introduces a model that integrates consumers' two learning processes and explores firms' pricing strategies in this context. We first examine the market outcomes when the product price is exogenous and the amount of information in the market remains constant. Depending on the prior belief upon the product

[5] As the news arrival rate approaches infinity, consumers are instantaneously informed about the product's true value. This situation mirrors real-world scenarios where aggregate reviews, such as five-star ratings, can accurately reflect the product's value when the volume of reviews is sufficiently high.

launch, one of the following scenarios occurs: either all consumers make a purchase directly, all consumers opt out directly, or the first consumer engages in active learning while all subsequent consumers simply follow the first consumer's choice. In any case, a herd begins no later than the second consumer, and the public belief is no longer updated thereafter.

When endogenizing the product price, we consider a situation where the seller is uncertain about the new product's fit to consumers' needs. Over time, the seller updates its belief in response to consumers' purchasing decisions and adjusts the price dynamically. In equilibrium, all consumers opting out due to a low prior belief wouldn't occur, as the seller can simply lower the price to ensure the product gets sold. However, whether or not the seller prefers to choose a price that induces active learning is not straightforward. Interestingly, encouraging early consumers to engage in active learning has a positive intertemporal effect on future expected profits. Therefore, a forward-looking seller may find it optimal to sacrifice early profits to induce learning and enhance future gains.

Overall, endogenous pricing has a double-edged effect on long-run market learning outcomes. Given that conclusive news has not arrived, the public belief will either gradually increase (in the case of conclusive bad news) or gradually decrease (in the case of conclusive good news). As the public belief increases, regardless of whether equilibrium prices encourage or discourage learning, all consumers will purchase the product. As the public belief decreases, the seller will, at a certain point, begin to deter learning and offer lower prices to induce all future consumers to make direct purchases. Therefore, in the absence of conclusive news, consumers herd on ″purchase″ in the long run. This implies that if the true state is H (L), endogenous pricing enhances (worsens) the market's learning outcomes.

The negative impact of endogenous pricing on long-run market learning outcomes can be mitigated when the availability of information in the market increases with sales. As discussed earlier, consumers herd on ′′purchase″ in the long run under endogenous pricing. The accumulated purchases result in an increasing news arrival rate. Ultimately, it becomes infeasible for the seller to use price to induce consumers to choose direct purchase over active learning, and therefore, complete learning is achieved.

Our theoretical insights can yield practical implications for addressing real-world business challenges, such as the *cold-start problem* encountered when launching a new product. A straightforward solution to the cold-start problem is to lower the initial price with the aim of expediting sales in order to create a herd among later consumers, or to enhance word-of-mouth, potentially building a positive reputation. However, this simple solution overlooks the active role of consumers in seeking information. Based on our analysis, considering consumers' active learning option, sellers can achieve higher profits through a more strategic pricing approach. Specifically, sellers should initially set a higher price to induce more learning among early consumers, rather than charging a low price that encourages direct purchases. Although inducing learning comes with the risk of consumers discovering bad news in the process, it is expected to be more effective in the long run.

Finally, there are some limitations of this paper that should be acknowledged. To make the model tractable and closely linked to the literature, we adopt the standard framework for observational learning in which consumers enter the market one by one in sequence. One might be concerned that this framework has limited applicability, as markets characterized by consumers making decisions sequentially are not prevalent. To partially address this concern, Online Appendix C presents an extension where N consumers enter the market in each period. The results of this extension are consistent with those of the main model, suggesting that the key insights of the main model are applicable to the mass product market. However, we acknowledge that while the new assumption of N consumers per period relaxes the limitations of sequential decision making, for observational learning to work, consumers need to be informed about the size of the target market and the total sales realized in each period. If consumers are unaware of the market size, or if obtaining sales information is difficult, it becomes challenging for consumers to draw quality cues from observing others. In such cases, active learning might be the sole channel for consumers to acquire information, simplifying this model to a standard static game with pricing and costly information acquisition.

References

S Nageeb Ali. Herding with costly information. *Journal of Economic Theory*, 175:713–729, 2018.

Abhijit V Banerjee. A simple model of herd behavior. *Quarterly Journal of Economics*, pages 797–817, 1992.

Heski Bar-Isaac and Sandro Shelegia. Search, showrooming, and retailer variety. *Marketing Science*, 42(2):251–270, 2023.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, pages 992–1026, 1992.

Patrick Bolton and Christopher Harris. Strategic experimentation. *Econometrica*, 67(2):349–374, 1999.

Subir Bose, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. Dynamic monopoly pricing and herding. *The RAND Journal of Economics*, 37(4):910–928, 2006.

Subir Bose, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. Monopoly pricing in the binary herding model. *Economic Theory*, 37(2):203–241, 2008.

Fernando Branco, Monic Sun, and J Miguel Villas-Boas. Too much information? information provision and search costs. *Marketing Science*, 35(4):605–618, 2016.

Ramon Caminal and Xavier Vives. Why market shares matter: An information-based theory. *The RAND Journal of Economics*, pages 221–239, 1996.

Yeon-Koo Che and Konrad Mierendorff. Optimal dynamic allocation of attention. *American Economic Review*, 109(8):2993–3029, 2019.

Yubo Chen and Jinhong Xie. Third-party product review and firm marketing strategy. *Marketing Science*, 24(2):218–240, 2005.

Jehoshua Eliashberg and Steven M Shugan. Film critics: Influencers or predictors? *Journal of marketing*, 61(2):68–78, 1997.

Kenneth Hendricks, Alan Sorensen, and Thomas Wiseman. Observational learning and demand for search goods. *American Economic Journal: Microeconomics*, 4(1):1–31, 2012.

Johannes Hörner and Andrzej Skrzypacz. Learning, experimentation and information design. *Advances in Economics and Econometrics*, 1:63–98, 2017.

Kinshuk Jerath and Qitian Ren. Consumer rational (in) attention to favorable and unfavorable product information, and firm information design. *Journal of Marketing Research*, 58 (2):343–362, 2021.

Godfrey Keller and Sven Rady. Strategic experimentation with poisson bandits. *Theoretical Economics*, 5(2):275–311, 2010.

Godfrey Keller and Sven Rady. Breakdowns. *Theoretical Economics*, 10(1):175–202, 2015.

Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

Dina Mayzlin and Jiwoong Shin. Uninformative advertising as an invitation to search. *Marketing Science*, 30(4):666–685, 2011.

Manuel Mueller-Frank and Mallesh M Pai. Social learning with costly search. *American Economic Journal: Microeconomics*, 8(1):83–109, 2016.

Srinivas K Reddy, Vanitha Swaminathan, and Carol M Motley. Exploring the determinants of broadway show success. *Journal of Marketing Research*, 35(3):370–383, 1998.

Lones Smith and Peter Sørensen. Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398, 2000.

Alan T Sorensen et al. Bestseller lists and the economics of product discovery. *Annual Review of Economics*, 9(1):87–101, 2017.

Ivo Welch. Sequential sales, learning, and cascades. *The Journal of finance*, 47(2):695–732, 1992.

Appendix: Proofs

[PROOF OF PROPOSITION 1] To derive the optimal stopping rule, we first characterize the consumer's value function. Define the belief $\pi_t = \Pr(H \mid X > t)$ as the state variable. At every point in time, the consumer chooses to continue learning, buying the product, or leaving the market without buying. Let w_t denote the instantaneous value of continuing learning at time t. The value function is expressed as follows:

$$V_t = \max\{w_t, \pi_t v - p, 0\}. \tag{16}$$

Note that w_t defines the expected continuation payoff with the optimal state-contingent action taken in the next instant based on information at time t. By waiting for a ξ length of time, the expected probability of bad news arrival is $(1 - \pi_t)(1 - e^{-\lambda\xi})$. If bad news arrives, the consumer will leave the market without buying. Otherwise, the consumer faces the same problem in the next instant. We have,

$$w_t = \lim_{\xi \to 0^+} (1 - \pi_t)(1 - e^{-\lambda\xi})0 + (1 - (1 - \pi_t)(1 - e^{-\lambda\xi}))V_{t+\xi} - c\xi, \tag{17}$$

where

$$V_{t+\xi} = \max\{w_{t+\xi}, \pi_{t+\xi} v - p, 0\},$$

and

$$\pi_{t+\xi} = \frac{\pi_t}{\pi_t + (1 - \pi_t)e^{-\lambda\xi}}.$$

The problem has a nice single crossing property, which allows us to pin down the optimal stopping time.

LEMMA 4 *The value function satisfies a single crossing condition: for a time t such that $w_t = \pi_t v - p$, we have $w_\tau < \pi_\tau v - p$ for all $\tau > t$.*

This single crossing property implies that, from time zero to the single crossing point, the expected benefit of learning is consistently higher than the expected benefit of buying. Beyond this point, buying strictly dominates learning. In order to prove the single crossing property, we introduce four additional lemmas.

LEMMA 5 *w_t is continuous in t*

Proof: It is immediately obtained from Equation (17).

∎

LEMMA 6 $\lim_{t \to \infty} \pi_t v - p > \lim_{t \to \infty} w_t$.

Proof: It is easy to show that $\lim_{t \to \infty} \pi_t = 1$ from Equation (1). And $\forall\, t$, we have

$$w_t \leq \lim_{\xi \to 0} (v - p) - c\xi.$$

Hence,

$$\lim_{t \to \infty} (\pi_t v - p - w_t) = v - p - \lim_{t \to \infty} w_t > 0.$$

∎

LEMMA 7 *When $w_t > \pi_t v - p$ for some $t > 0$, then there exists some $\epsilon > 0$, such that w_t is continuously differentiable on the interval $(t - \epsilon, t + \epsilon)$, and $\dot{w}_t = (1 - \pi_t)\lambda w_t + c$.*

Proof: As $w_t > \pi_t v - p$, we have

$$w_t = \lim_{\xi \to 0} (1 - \lambda\xi(1 - \pi_t))(w_t + \xi\dot{w}_t) - c\xi.$$

Therefore,

$$\lim_{\xi \to 0} \lambda\xi(1 - \pi_t) w_t = \lim_{\xi \to 0} \xi\dot{w}_t(1 - \lambda\xi(1 - \pi_t)) - c\xi.$$

We get,

$$\dot{w}_t = \lambda(1 - \pi_t) w_t + c.$$

■

LEMMA 8 *When $w_t < \pi_t v - p$ for some $t > 0$, then there exists some $\epsilon > 0$, such that w_t is continuously differentiable on the interval $(t-\epsilon, t+\epsilon)$, and $\dot{w}_t = \dot{\pi}_t v$*

Proof: As $w_t < \pi_t v - p$, we have

$$w_t = \lim_{\xi \to 0}(1 - \lambda\xi(1-\pi_t))(\pi_t v + \dot{\pi}_t \xi v - p) - c\xi.$$

Therefore,

$$\dot{w}_t = \lim_{\xi \to 0} \lambda\xi\dot{\pi}_t(\pi_t v + \dot{\pi}_t\xi v - p) + (1 - \lambda\xi(1-\pi_t))\dot{\pi}_t v + (1 - \lambda\xi(1-\pi_t))\dot{\dot{\pi}}_t \xi v.$$

After simplifying it, we obtain

$$\dot{w}_t = \dot{\pi}_t v.$$

■

Finally, we prove Lemma 4 by contradiction. Recall that the lemma is: for a t such that $w_t = \pi_t v - p$, we have $w_\tau < \pi_\tau v - p$ for all $\tau > t$.

Proof: Suppose $w_t = \pi_t v - p$ and $\exists\, l' > t$ such that $w_{l'} \geq \pi_{l'} v - p$. Because of Lemma 5, we can define a $\bar{l} = sup\{l > l' : w_l \geq \pi_l v - p\}$. By Lemma 6, it is straightforward that $\bar{l} < \infty$. We also define $\underline{l} = max\{t, inf\{l < l' : w_l \geq \pi_l v - p\}\}$. Then $\underline{l} \leq \bar{l}$, and $w_l \geq \pi_l v - p$ for all $l \in [\underline{l}, \bar{l}]$. Moreover, we know $\exists\, \epsilon$ such that for $l \in (\underline{l} - \epsilon, \underline{l})$ and $l \in (\bar{l}, \bar{l} + \epsilon)$, we have $w_l < \pi_l v - p$. By applying Lemma 7, we get two limits $\Lambda_{\underline{l}}$ and $\Lambda_{\bar{l}}$.

$$\dot{\pi}_l = \lambda\pi_l(1-\pi_l),$$

$$\Lambda_{\underline{l}} = \lim_{l \to \underline{l}^-}(\dot{w}_l - \frac{d}{dl}(\pi_l v - p)) = \lambda(1-\pi_{\underline{l}})(\pi_{\underline{l}} v - p) + c - \lambda\pi_{\underline{l}}(1-\pi_{\underline{l}})v,$$

$$\Lambda_{\bar{l}} = \lim_{l \to \bar{l}^+}(\dot{w}_l - \frac{d}{dl}(\pi_l v - p)) = \lambda(1-\pi_{\bar{l}})(\pi_{\bar{l}} v - p) + c - \lambda\pi_{\bar{l}}(1-\pi_{\bar{l}})v.$$

We should have $\Lambda_{\underline{l}} \geq 0$ and $\Lambda_{\bar{l}} \leq 0$. Hence, the two equations need to be satisfied:

$$c \geq \lambda(1 - \pi_{\underline{l}})p,$$

$$c \leq \lambda(1 - \pi_{\bar{l}})p.$$

Because π_l is monotonically increasing in l, $\pi_{\bar{l}} \geq \pi_{\underline{l}}$. The two equations cannot be satisfied at the same time unless $\underline{l} = \bar{l} = l^*$. To rule out l^*, we know that $\forall l \in (t, t+\epsilon)$, $w_l < \pi_l v - p$, and moreover, $\dot{w}_l = (\pi_l v \dot{-} p)$ according to Lemma 8. Therefore, it is impossible that $l^* > t$ exists where $w_{l^*} = \pi_{l^*} v - p$.

∎

We will proceed to prove Proposition 1. The single crossing condition in Lemma 4 implies that the optimal stopping time t^* is when:

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

Rearrange the equation, we can get

$$\overbrace{\underbrace{\lambda\xi(1-\pi_t)}_{\substack{\text{Probability}\\ \text{of signal}}} \times \underbrace{p}_{\substack{\text{Benefit:}\\ \text{Avoid Bad Product}}}}^{\text{Benefit of Learning}} = \overbrace{c\xi.}^{\text{Cost of Learning}}$$

Finally, we have the optimal stopping condition $\pi_{t^*} = 1 - \frac{c}{\lambda p}$.

∎

[PROOF OF LEMMA 2] The proof is analogous to proving that l_1^B in the benchmark model (which appears in Lemma 1 and Figure I) is increasing in p. Although we do not provide a closed-form expression for l_1^B due to its complexity, we can use implicit differentiation by taking the total derivatives of both sides of Equation (5) to prove that l_1^B is increasing in p. In this proof, we will show another method that does not require finding the expression for the price ceiling.

At any moment, when a consumer holds belief π and faces price p, the consumer has three options. First, the consumer can simply opt out and receive zero payoff. Second, the consumer can buy the product directly (denoted by the subscript BD) and receive an expected payoff of $W_{BD}(\pi, p) = \pi v - p$. Third, the consumer can engage in active learning until the next instant (denoted by the subscript AL). From Proposition 1, we know that once the consumer starts active learning, she will keep learning until bad news arrives or until her belief reaches $\pi_{t^*}(p) = 1 - \frac{c}{\lambda p}$.

Let $C(\pi, \pi_{t^*}(p))$ denote the expected cost of the whole active learning process (from the prior belief π to the arrival of bad news or to the posterior belief π_{t^*}). As the belief updating process is mean-preserving, the probability that the consumer will eventually make a purchase is $\frac{\pi}{\pi_{t^*}(p)}$, and therefore, the expected return from learning is $\pi v - \frac{\pi}{\pi_{t^*}(p)} p$. The expected payoff from active learning is then $W_{AL}(\pi, p) = \pi v - \frac{\pi p}{\pi_{t^*}(p)} - C(\pi, \pi_{t^*}(p))$, and the overall expected payoff to the consumer is $W(\pi, p) = \max\{0, W_{BD}(\pi, p), W_{AL}(\pi, p)\}$. As both $W_{BD}(\pi, p)$ and $W_{AL}(\pi, p)$ are continuous in π and p, $W(\pi, p)$ is also continuous in π and p everywhere. Next, we will show that $W(\pi, p)$ is decreasing in p.

Consider two prices $p > p'$. For any π, it is obvious that $W_{BD}(\pi, p) < W_{BD}(\pi, p')$. And for $W_{AL}(\pi, p)$, we have:

$$W_{AL}(\pi, p) < \pi v - \frac{\pi p'}{\pi_{t^*}(p)} - C(\pi, \pi_{t^*}(p)) \leq \pi v - \frac{\pi p'}{\pi_{t^*}(p')} - C(\pi, \pi_{t^*}(p')) \equiv W_{AL}(\pi, p').$$

The second inequality holds because $\pi_{t^*}(p')$ is the optimal stopping belief of active learning under p'. Hence, we know that $W(\pi, p) = \max\{0, W_{BD}(\pi, p), W_{AL}(\pi, p)\}$ is strictly decreasing in p. As $W(\pi, p)$ is both continuous and decreasing in p, if $W(\pi, p) > 0$, the seller can always charge a higher price until $W(\pi, \bar{p}(\pi)) = 0$, where $\bar{p}(\pi)$ is the price ceiling that extracts all the surplus from the consumer.

Finally, we also need to prove that $\bar{p}(\pi)$ is increasing in π. Consider any two different prior beliefs, $\pi > \pi'$, and the corresponding price ceilings $\bar{p}(\pi)$ and $\bar{p}(\pi')$. If with a prior belief π' and price $\bar{p}(\pi')$, the consumer buys directly or engages in active learning until $\pi_{t^*}(\bar{p}(\pi'))$. And if $\pi_{t^*}(\bar{p}(\pi')) < \pi$, then when with a prior belief π and price $\bar{p}(\pi')$, the consumer can at least buy directly and gains an expected payoff that is higher than $W(\pi', \bar{p}(\pi'))$. In the case where $\pi_{t^*}(\bar{p}(\pi')) \geq \pi$, then when with a prior belief π and price $\bar{p}(\pi')$, the consumer can at least engages in active learning utill $\pi_{t^*}(\bar{p}(\pi'))$. By doing so, the consumer obtains the same expected gain but saves the learning cost from π' to π. Therefore, we have $W(\pi', \bar{p}(\pi')) < W(\pi, \bar{p}(\pi'))$. From the above analysis, we know that $W(\pi, \bar{p}(\pi)) = W(\pi', \bar{p}(\pi')) = 0$. Assume that $\bar{p}(\pi) < \bar{p}(\pi')$, we get the following:

$$0 = W(\pi, \bar{p}(\pi)) > W(\pi, \bar{p}(\pi')) > W(\pi', \bar{p}(\pi')) = 0,$$

and this is a contradiction. Therefore, $\bar{p}(\pi) > \bar{p}(\pi')$ must hold, i.e., $\bar{p}(\pi)$ is increasing in π.

■

[PROOF OF PROPOSITION 3] When $\pi_0 \geq \frac{1}{2}$, $\tilde{p}(\pi_0) = \frac{c}{\lambda(1-\pi_0)} \geq \frac{2c}{\lambda}$, and therefore over the entire range of p_1 in Condition (10), $\frac{\partial \Pi_1(p_1)}{\partial p_1} > 0$. If $\tilde{p}(\pi_0) < \bar{p}(\pi_0)$, it is optimal for the seller to set $p_1 = \bar{p}(\pi_0) > \tilde{p}(\pi_0)$. The first consumer will engage in active learning at the chosen price. From Lemma 3, if $\tilde{p}(\pi_0) \geq \bar{p}(\pi_0)$, it is also optimal for the seller to set $p_1 = \bar{p}(\pi_0)$. The first consumer will buy directly.

When $\pi_0 < \frac{1}{2}$ and $\tilde{p}(\pi_0) < \bar{p}(\pi_0)$, the seller will compare the profit of selling at $\bar{p}(\pi_0)$ (i.e., $\Pi_1(\bar{p}(\pi_0)) = \frac{\pi_0}{1-\frac{c}{\lambda \bar{p}(\pi_0)}} \bar{p}(\pi_0)$) with the profit of letting the consumer buy directly without learning (i.e., $\Pi_1(\tilde{p}(\pi_0)) = \frac{c}{\lambda(1-\pi_0)}$) and choose the more profitable price. Whether the first consumer actively learns or directly buys depends on the chosen price. If $\tilde{p}(\pi_0) \geq \bar{p}(\pi_0)$, the seller chooses $p_1 = \bar{p}(\pi_0)$, and the first consumer will buy directly.

If consumer 1 engages in active learning and makes a purchase, her posterior belief conditional on no bad news arriving, $\pi_1^{h=1} = 1 - \frac{c}{\lambda \bar{p}(\pi_0)}$ must exceed $\frac{1}{2}$. This is because $\frac{\partial \Pi_1(p_1)}{\partial p_1} < 0$ when p_1 is such that posterior belief $\pi_1 \leq \frac{1}{2}$. Otherwise, it is optimal for the seller to decrease p_1 until $\tilde{p}(\pi_0)$ to prevent learning (i.e., learning-deterring). Subsequently, for any consumer $n \geq 2$, it is always optimal for the seller to set a price at $p_n = \bar{p}(\pi_{n-1}^{h=1})$, and $\bar{p}(\pi_{n-1}^{h=1}) > \tilde{p}(\pi_{n-1}^{h=1})$, which implies that all the subsequent consumers induce active learning. To see the latter, note that

$$\pi_n^{h=1} = 1 - \frac{c}{\lambda \bar{p}(\pi_{n-1}^{h=1})}. \tag{18}$$

Rearranging (18), we have

$$\pi_{n-1}^{h=1} = \frac{c}{\lambda(1 - \pi_n^{h=1})} = \tilde{p}(\pi_n^{h=1}) > \tilde{p}(\pi_{n-1}^{h=1}). \tag{19}$$

Apparently, the last inequality holds when $n = 1$ as $\pi_1^{h=1} > \pi_0$. And by induction, it holds for all $n > 1$. Therefore, we find that $\bar{p}(\pi_n^{h=1}) > \tilde{p}(\pi_n^{h=1})$, which implies that there will be active learning for all subsequent consumers conditional on no bad news arriving.

■

[PROOF OF PROPOSITION 4] We break down this proposition into three parts and prove each of them. First, only in the region where $\frac{\pi_0}{1-\frac{c}{\lambda \bar{p}(\pi_0)}} \tilde{p}(\pi_0) \leq \bar{p}(\pi_0)$, the forward-looking seller may behave differently from the myopic seller by charging a learning-inducing price. Second, the greater β is, the more likely the forward-looking seller is to induce learning. Third, in the learning-inducing scheme, the seller consistently induces learning in all subsequent periods. Below, we prove the three parts separately.

Part 1: The forward-looking seller may induce active learning. First, if the parameters fall into the region where a myopic seller chooses $p_1 = \bar{p}(\pi_0)$ to induce the first consumer to actively learn (case 1 of Proposition 3), the posterior belief in the absence of bad news must exceed $\frac{1}{2}$. And for all the subsequent consumers, we have $\pi_n^{h=1} > \frac{1}{2}, \forall n > 1$. Therefore, $\tilde{p}(\pi_{n-1}^{h=1}) = \frac{c}{\lambda(1-\pi_{n-1}^{h=1})} > \frac{2c}{\lambda}, \forall n > 1$. Hence, for all $n > 1$ periods, $\frac{\partial \Pi_n(p_n|\pi_{n-1}^{h=1})}{\partial p_n} = \frac{\pi_0(1-\frac{2c}{\lambda p_n})}{(1-\frac{c}{\lambda p_n})^2} > 0$, and thus the optimal price plan that maximizes the static profit of each period is: $p_n = \bar{p}(\pi_{n-1}^{h=1})$. When the seller is forward-looking, as $\bar{p}(\pi)$ is increasing in π, charging $p_1 = \bar{p}(\pi_0)$ at $t = 1$ not only yields the highest static profit but also generates the highest possible posterior belief $\pi_1^{h=1}$, which benefits all future periods. Following the same logic, the forward-looking seller will consistently find it optimal to set prices $p_n = \bar{p}(\pi_{n-1})$ in all periods. Therefore, in this case, the forward-looking seller employs the same pricing strategy as that of the myopic seller.

Second, if the parameters fall into the region where $\tilde{p}(\pi_0) \geq \bar{p}(\pi_0)$ (case 2 of Proposition 3), for any price p_1, consumer 1 either buys directly or opts out. Hence, the public belief will not evolve under any price p_1, and therefore early prices have no intertemporal effect on later profits. In this case, the forward-looking seller's optimal pricing strategy remains the same as that of the myopic seller. In all subsequent periods, the seller must charge the highest price that induces consumers to buy directly, $p_n = \bar{p}(\pi_0)$, and none of the consumers engage in active learning.

Third, when $\frac{\pi_0}{1-\frac{c}{\lambda \bar{p}(\pi_0)}}\bar{p}(\pi_0) \leq \tilde{p}(\pi_0) < \bar{p}(\pi_0)$ (case 3 of Proposition 3), the forward-looking seller may behave differently from the myopic seller. It might be optimal for the forward-looking seller to charge $p_1 > \tilde{p}(\pi_0)$ at $t = 1$ to induce active learning by consumer 1. By doing so, the seller sacrifices the static payoff at $t = 1$, but gains by increasing consumer 2's prior belief to $\pi_1^{h=1}(p_1) = \frac{\pi_0}{1-\frac{c}{\lambda p_1}}$, allowing for a higher price at $t = 2$ and beyond. It is worth noting that, in the absence of bad news arrival, after the first time when the public belief surpasses $\frac{1}{2}$, the static profit of a single period becomes monotonically increasing in price. Therefore, in all subsequent periods, the seller will optimally set prices at

$p_n = \bar{p}(\pi_{n-1})$, leading to active learning by all subsequent consumers.

Part 2: Greater β, more likely to induce learning. In this part, we demonstrate that the incentive to induce active learning strengthens with a higher β. We begin by proving the following lemma.

LEMMA 9 *For any increasing sequence $\{a_n\}$, if there exists $x > 0$, such that $\sum_{n=1}^{\infty} x^n a_n \geq 0$ and finite. Then for any $y > x$, $\sum_{n=1}^{\infty} y^n a_n \geq \sum_{n=1}^{\infty} x^n a_n$.*

Proof. When $a_1 \geq 0$, as $\{a_n\}$ is an increasing sequence, all elements are positive. As $y > x$, it is trivial that $\sum_{n=1}^{\infty} y^n a_n \geq \sum_{n=1}^{\infty} x^n a_n$. When $a_1 < 0$, there must exist a turning point $\hat{n}$, such that $a_n < 0$ if and only if $n \leq \hat{n}$.

We prove this by contradiction. Let $X = \sum_{n=1}^{\infty} x^n a_n$ and $Y = \sum_{n=1}^{\infty} y^n a_n$. Assume that there exists an $y > x$ such that $Y < X$. As $y > x$

$$\frac{Y - X}{y - x} = a_1 + \frac{y^2 - x^2}{y - x} a_2 + \frac{y^3 - x^3}{y - x} a_3 + \cdots < 0. \tag{20}$$

As $X \geq 0$ and $y > x > 0$, $\frac{\frac{Y-X}{y-x} - \frac{X}{x}}{y} < 0$. Thus, we have

$$\frac{\frac{Y-X}{y-x} - \frac{X}{x}}{y} = a_2 + \frac{y^2 - x^2}{y - x} a_3 + \frac{y^3 - x^3}{y - x} a_4 + \cdots < 0. \tag{21}$$

If $\hat{n} = 1$, all terms in (21) are positive. We find a contradiction. If $\hat{n} > 1$, as $a_1 < 0$, $X \geq 0$, and $y > x > 0$, we have

$$\frac{\frac{\frac{Y-X}{y-x} - \frac{X}{x}}{y} - \frac{X}{x^2} + \frac{a_1}{x}}{y} = a_3 + \frac{y^2 - x^2}{y - x} a_4 + \frac{y^3 - x^3}{y - x} a_5 + \cdots < 0. \tag{22}$$

If $\hat{n} = 2$, all terms in (22) are positive. We find a contradiction. If $\hat{n} > 2$, by induction, we have

$$\frac{1}{y^n}\frac{Y-X}{y-x}-\sum_{i=1}^{\hat{n}}\frac{X}{y^i x^{\hat{n}+1-i}}+\sum_{i=2}^{\hat{n}}\left(\sum_{j=1}^{\hat{n}-i+1}\frac{a_{i-1}}{y^j x^{\hat{n}-i+2-j}}\right)=a_{\hat{n}+1}+\frac{y^2-x^2}{y-x}a_{\hat{n}+2}+\frac{y^3-x^3}{y-x}a_{\hat{n}+3}+\cdots<0. \tag{23}$$

As all terms on the right-hand side of (23) are positive, the sum cannot be negative, leading to a contradiction.

■

When $\frac{\pi_0}{1-\frac{c}{\lambda\bar{p}(\pi_0)}}\bar{p}(\pi_0) \leq \tilde{p}(\pi_0) < \bar{p}(\pi_0)$, a myopic seller would charge $p_n = \tilde{p}(\pi_0)$ in all periods, and the expected payoff remains the same across all periods, $\Pi_n(\tilde{p}(\pi_0)|\pi_0) = \tilde{p}(\pi_0)$. For any $\beta_1 < \beta_2$, if it is optimal under β_1 to induce active learning, denote the optimal prices in the absence of bad news by $\{p_{n,\beta_1^*}\}$. The corresponding posterior beliefs are denoted by $\{\pi_{n,\beta_1^*}^{h=1}\}$. The equilibrium profit of a seller with β_1 is:

$$\Pi^*(\{p_{n,\beta_1^*}\}|\beta_1)=\sum_{n=1}^{\infty}\beta_1^{n-1}\frac{\pi_0}{1-\frac{c}{\lambda p_{n,\beta_1^*}}}p_{n,\beta_1^*}. \tag{24}$$

It must exceed the profit under the myopic seller's pricing. Let $\Delta_n = \frac{\pi_0}{1-\frac{c}{\lambda p_{n,\beta_1^*}}}p_{n,\beta_1^*} - \tilde{p}(\pi_0)$, we have

$$\Pi^*(\{p_{n,\beta_1^*}\}|\beta_1)-\Pi(p_n=\tilde{p}(\pi_0)|\beta_1)=\sum_{n=1}^{\infty}\beta_1^{n-1}\Delta_n\geq 0 \tag{25}$$

As the expected one-period profit is increasing, Δ_n is increasing in n, and there exists a threshold $\hat{n}$, such that $\Delta_n < 0$ when $n < \hat{n}$, and $\Delta_n \geq 0$ when $n \geq \hat{n}$. Now, we consider a seller with β_2, who values the future more than the seller with β_1. For the β_2 seller, by Lemma 9, if it uses the same pricing strategy, $\{p_{n,\beta_1^*}\}$, as the β_1 seller, it also gains a higher profit than statically optimal prices $p_n = \tilde{p}(\pi_0)$. That is,

$$\Pi(\{p_{n,\beta_1^*}\}|\beta_2)-\Pi(p_n=\tilde{p}(\pi_0)|\beta_1)=\sum_{n=1}^{\infty}\beta_2^{n-1}\Delta_n>\sum_{n=1}^{\infty}\beta_1^{n-1}\Delta_n\geq 0. \tag{26}$$

Moreover, note that $\{p_{n,\beta_1^*}\}$ may not be the optimal prices for the β_2 seller. The equilibrium profit for the β_2 seller must be weakly greater than the profit under $\{p_{n,\beta_1^*}\}$. Thus, we have

$$\Pi^*(\{p_{n,\beta_2^*}\}|\beta_2) \geq \Pi(\{p_{n,\beta_1^*}\}|\beta_2) > \Pi(\{p_{n,\beta_1^*}\}|\beta_1) \geq \Pi(p_n = \tilde{p}(\pi_0)|\beta_1). \tag{27}$$

Therefore, a consumer with a higher β has a stronger incentive to induce active learning.

Part 3: Learning-inducing in all subsequent periods. We continue to show that if the seller induces learning in the first period, it is optimal for the seller to choose learning-inducing prices in all subsequent periods. First, if the seller can set a learning-inducing price at $t = 1$, from Lemma 3, we must have $\bar{p}(\pi_0) > \tilde{p}(\pi_0) = \frac{c}{\lambda(1-\pi_0)}$. $\bar{p}(\pi_0)$ is determined by (5). Recall that (5) states

$$\pi \geq \frac{c}{\lambda(v-p) + c\ln\frac{\pi c}{(1-\pi)(\lambda p - c)}}. \tag{28}$$

It is easy to verify that the RHS is strictly increasing in p. Therefore, $\bar{p}$ is the highest price to make (5) binding. If $\bar{p}(\pi_0) > \tilde{p}(\pi_0)$, substituting p by $\tilde{p}(\pi_0)$ and π by π_0 in (5), the inequality must have slack. That is,

$$\pi_0 > \frac{c}{\lambda v - \frac{c}{1-\pi_0}}. \tag{29}$$

It simplifies to $\lambda v \pi_0(1-\pi_0) - c > 0$. When $v \leq \frac{4c}{\lambda}$, there is no solution for (29). It implies that for any prior belief π_0, $\bar{p}(\pi_0) \leq \tilde{p}(\pi_0)$, and thus, there cannot be an initial learning-inducing price. When $v > \frac{4c}{\lambda}$, the solution is $\pi_0 \in (\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}, \frac{\lambda v + \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v})$. Hence, we conclude that the prior belief π_0 must lie within this region.

If the seller induces learning at $t = 1$, conditional on the first consumer's purchase (with no bad news arriving), the prior belief is $\pi_1^{h=1} > \pi_0$. There are two cases: $\pi_1^{h=1} \geq \frac{1}{2}$, or $\pi_1^{h=1} < \frac{1}{2}$.

1. If $\pi_1^{h=1} > \frac{1}{2}$, following (11) in the analysis of Proposition 3, it is even optimal for the myopic seller to set a learning-inducing price $p_2 = \bar{p}(\pi_1^{h=1})$. As the forward-looking seller also gains from a more informative posterior belief, it sets the highest possible learning-inducing price, $p_2 = \bar{p}(\pi_1^{h=1})$.

2. If $\pi_1^{h=1} \leq \frac{1}{2}$, assume that it is optimal for the seller to deter learning from $t = 2$ onward. That is, $p_2 = \tilde{p}(\pi_1^{h=1})$. However, this is dominated by charging a uniform price at $p = \tilde{p}(\pi_0)$ from $t = 1$ onward. The probability that no bad news arrives is $\frac{\pi_0}{\pi_1^{h=1}}$; otherwise, the first consumer receives bad news, and all the subsequent consumers do not buy. Hence, the expected profit from $t = 2$ onward is

$$\frac{\pi_0}{\pi_1^{h=1}} \frac{c}{\lambda(1 - \pi_1^{h=1})}.$$

   However, this is lower than the profit gained if the seller charges $p = \tilde{p}(\pi_0)$ from $t = 1$ onward, which equals $\frac{c}{\lambda(1-\pi_0)}$. To see it, note that

$$\frac{\pi_0}{\pi_1^{h=1}} \frac{c}{\lambda(1 - \pi_1^{h=1})} \geq \frac{c}{\lambda(1 - \pi_0)} \iff \pi_0(1 - \pi_0) \geq \pi_1^{h=1}(1 - \pi_1^{h=1}).$$

   It cannot be true as $\pi_0 < \pi_1^{h=1} \leq \frac{1}{2}$.

Inductively, if $\pi_1^{h=1} < \frac{1}{2}$, following the same argument, we know that whether $\pi_2^{h=1} \geq \frac{1}{2}$ or $\pi_2^{h=1} < \frac{1}{2}$, the seller still induces learning at $t = 3$, and this strategy continues in all future periods. The above analysis also suggests that if it is optimal for the seller to induce learning, the public belief sequence will eventually exceed $\frac{1}{2}$. Otherwise, it will be dominated by deterring learning from $t = 1$.

Finally, we need to show that when the seller charges learning-inducing prices, the public belief will never exceed the region where $\bar{p}(\pi) > \tilde{p}(\pi)$. We prove it by showing that the belief updating process under $p = \bar{p}(\pi)$ is a contraction. With a slight abuse of notation, for any prior belief $\pi \in (\frac{\lambda v - \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}, \frac{\lambda v + \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v})$ (the region where $\bar{p}(\pi) > \tilde{p}(\pi)$), let the posterior belief conditional on no bad news arriving under $p = \bar{p}(\pi)$ be π'. Under the learning-inducing regime, we have

$$\pi' = 1 - \frac{c}{\bar{p}(\pi)}, \tag{30}$$

where $\bar{p}(\pi)$ is determined by the binding inequality (5). Rearranging (30), we have $\lambda\bar{p}(\pi) = \frac{c}{1-\lambda'}$. Substituting it into the binding inequality (5), we get

$$\lambda v - \frac{c}{1-\pi'} + c\ln\frac{\pi}{1-\pi} + c\ln\frac{1-\pi'}{\pi'} = \frac{c}{\pi}. \tag{31}$$

Equation (31) determines a mapping from π to π'. We denote it as $\pi' = f(\pi)$. Take full derivative with respect to π and π', we have

$$f'(\pi) = \frac{d\pi'}{d\pi} = \frac{\pi'(1-\pi')^2}{\pi^2(1-\pi)}. \tag{32}$$

It is easy to see that $f'(\pi) < 1$ when $1 > \pi' > \pi > \frac{1}{2}$. By the mean value Theorem, $f(\pi)$ is a contraction over $(\frac{1}{2}, 1)$. By the Banach fixed-point theorem, $f(\pi)$ has unique fixed point $\bar{\pi}$, such that starting from any π_0, the sequence $\pi_n = f(\pi_{n-1})$ converges to $\bar{\pi}$ as $n \to \infty$. To find the convergence limit $\bar{\pi}$, note that at the fixed point, $\bar{\pi} = f(\bar{\pi})$. Thus, we have

$$\bar{\pi} = \frac{c}{\lambda v - \lambda\bar{p}(\bar{\pi}) + c\ln\frac{\bar{\pi}c}{(1-\bar{\pi})(\lambda\bar{p}(\bar{\pi})-c)}}. \tag{33}$$

It has a unique root, $\bar{\pi} = \frac{\lambda v + \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$ over $(\frac{1}{2}, 1)$, which happens to be the upper-bound of the region where $\bar{p}(\pi) > \tilde{p}(\pi)$. Therefore, we show that as n increases, $\pi_n^{h=1}$ will always fall into the region where the seller finds it optimal to induce active learning, and $\lim_{n\to\infty} \pi_n^{h=1} = \bar{\pi} \equiv \frac{\lambda v + \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$.

■

[PROOF OF COROLLARY 1] The proof of Proposition 4 already derives the limit of convergence of public belief, $\bar{\pi} = \frac{\lambda v + \sqrt{\lambda v(\lambda v - 4c)}}{2\lambda v}$. We proceed to derive the probability of an incorrect herd.

First, when the seller deters learning from $t = 1$, there is no active learning, and the public belief remains to be π_0. As all consumers purchase under the learning-deterring scheme, an incorrect herd occurs with probability 1 when the true state is L, and it occurs with probability 0 when the true state is H.

Second, when the seller uses the learning-inducing strategy, from Proposition 4, all consumers engage in active learning and the public belief converges to $\bar{\pi}$. When the true state is H, conclusive bad news never arrives, and the seller will always set prices that make consumers purchase after learning, eliminating the possibility of an incorrect herd. An incorrect herd can only occur when the true state is L. The probability of no bad signal arriving during the learning process of all consumers, which leads to the public belief converging to $\bar{\pi}$, by Bayes rule, is

$$\Pr(h=1) = \frac{\pi_0}{\bar{\pi}} = \Pr(H) + \Pr(L)\Pr(h=1|L) = \pi_0 + (1-\pi_0)\Pr(h=1|L). \tag{34}$$

Thus, we obtain $\Pr(h = 1|L) = \frac{\pi_0}{1-\pi_0}\frac{1-\bar{\pi}}{\bar{\pi}}$, which is the probability of an incorrect herd conditional on state L.

■

[PROOF OF PROPOSITION 5] For a myopic seller with $\beta = 0$, in any period with a prior belief π and news arrival rate $f(\lambda, N)$, the seller only sets a price that optimizes current profits. Thus, the optimal pricing strategy aligns with the results of the main model with a fixed news arrival rate, simply replacing λ with $f(\lambda, N)$.

When $\tilde{p}(\pi, N) \geq \bar{p}(\pi, N)$, the seller cannot induce learning by any price. The consumer either buys directly or opts out. Therefore, a forward-looking seller also chooses to deter learning. When $\tilde{p}(\pi, N) < \frac{\pi}{1-\frac{c}{f(\lambda,N)\bar{p}(\pi,N)}}\bar{p}(\pi, N)$, both the myopic seller and the forward-looking seller charge $p = \bar{p}(\pi, N)$, which is the highest price that induces learning. Thus, the only non-trivial case lies in the region where $\frac{\pi}{1-\frac{c}{f(\lambda,N)\bar{p}(\pi,N)}}\bar{p}(\pi, N) \leq \tilde{p}(\pi, N) < \bar{p}(\pi, N)$.

For a forward-looking seller with $\beta \in (0, 1)$, the seller optimizes the sum of the expected future payoffs. At period n, if all consumers have made a purchase in the history ($h = 1$), the public belief is updated to $\pi_{n-1}^{h=1}$, and the news arrival rate becomes $f(\lambda, n-1)$. If not all previous consumers have purchased ($h = 0$), the consumer infers that the product is of type L and consequently rejects any positive price. The optimization problem for the seller is

$$\max_{\{p_n\}} \Pi(\{p_n\}) = \sum_{n=1}^{\infty} \beta^{n-1} \Pi_n(p_n | \pi_{n-1}^{h=1}), \tag{35}$$

subject to

$$\Pi_n(p_n|\pi_{n-1}^{h=1}) = \begin{cases} \frac{\pi_0}{1-\frac{c}{f(\lambda,n-1)p_n}} p_n & \text{if } p_n > \tilde{p}(\pi_{n-1}^{h=1}, n-1), \\ p_n & \text{if } p_n \leq \tilde{p}(\pi_{n-1}^{h=1}, n-1), \end{cases} \tag{36}$$

$$\pi_n^{h=1} = \begin{cases} 1 - \frac{c}{\lambda p_n} & \text{if } p_n > \tilde{p}(\pi_{n-1}^{h=1}, n-1), \\ \pi_{n-1}^{h=1} & \text{if } p_n \leq \tilde{p}(\pi_{n-1}^{h=1}, n-1), \end{cases} \tag{37}$$

$$p_n \in \left[\min\left\{\bar{p}(\pi_{n-1}^{h=1}, n-1), \tilde{p}(\pi_{n-1}^{h=1}, n-1)\right\}, \bar{p}(\pi_{n-1}^{h=1}, n-1)\right], \tag{38}$$

Note that this is analogous to the main model, except for replacing λ with $f(\lambda, n-1)$. Conditional on no bad signal arriving in the history ($h = 1$), the evolution of $f(\lambda, n)$ is deterministic. Therefore, switching from learning-deterring to learning-inducing at time t will not affect $f(\lambda, n)$ in any period $n > t$ but will only increase the public belief $\pi_n^{h=1}$ in all subsequent periods. Therefore, it is sufficient for us to show that $\Pi_n(p_n|\pi_{n-1}^{h=1})$ is increasing in $\pi_{n-1}^{h=1}$ for any $f(\lambda, n-1)$. Assume that in any period n, the public belief (when $h = 1$) is $\pi_{n-1}^{h=1}$ under the optimal strategy of the myopic seller. If the forward-looking seller induces learning in a previous period, the public belief is $\tilde{\pi}_{n-1}^{h=1} \geq \pi_{n-1}^{h=1}$. There are two cases:

1. If for the myopic seller, it is optimal to set a learning-deterring price $p_n = \tilde{p}(\pi_{n-1}^{h=1}, n-1)$. Then, a forward-looking seller can at least set the price $p_n = \tilde{p}(\tilde{\pi}_{n-1}^{h=1}, n-1)$. As $\tilde{p}(\pi, n)$ is strictly increasing in π, the profit will be strictly higher.

2. If for the myopic seller, it is optimal to set a learning-inducing price $p_n = \bar{p}(\pi_{n-1}^{h=1}, n-1)$. Then, for the forward-looking seller, if $\bar{p}(\pi_{n-1}^{h=1}, n-1) < \tilde{p}(\tilde{\pi}_{n-1}^{h=1}, n-1)$, the seller can set the price at $\tilde{p}(\tilde{\pi}_{n-1}^{h=1}, n-1)$, ensuring sales to all consumers at a higher price. Thus, the profit is strictly higher. If $\bar{p}(\pi_{n-1}^{h=1}, n-1) \geq \tilde{p}(\tilde{\pi}_{n-1}^{h=1}, n-1)$, the seller can at least set the same price of $\bar{p}(\pi_{n-1}^{h=1}, n-1)$ and guarantee the same profit. This is feasible as $\bar{p}(\pi, n)$ is increasing in π.

Summarizing the above two cases, we show that $\Pi_n(p_n|\pi_{n-1}^{h=1})$ is increasing in $\pi_{n-1}^{h=1}$ for any $f(\lambda, n-1)$, implying that the seller benefits from generating a more dispersed public belief in later periods. Therefore, the seller may find it optimal to sacrifice early profits by charging a higher price to induce learning.

■