COLUMBIA UNIVERSITY
Center for Pricing and Revenue Management

WORKING PAPER SERIES: NO. 2013-5

Bayesian Social Learning from Consumer Reviews

Bar Ifrach,
Stanford University

Costis Maglaras,
Columbia University

Marco Scarsini,
Singapore University of Technology and Design

2013

http://www.cprm.columbia.edu

Bayesian Social Learning from Consumer Reviews[*]

Bar Ifrach[†] Costis Maglaras[‡] Marco Scarsini[§]

June 29, 2013

Abstract

Motivated by the proliferation of user generated content online, we study a market with heterogeneous customers who learn in a Bayesian manner the quality of an offered product by observing the reviews of customers who purchased the product earlier in time. The seller, to whom the quality is initially unknown, as well, can adjust her prices dynamically according to the reviews' changing sentiment. We find that this social learning process is successful; agents—the consumers and seller—eventually learn the quality of the product. This holds, under different conditions, both in the case when agents observe an ordered and unordered history of reviews. Analyzing the social learning trajectory, we find that earlier reviews are more influential than later ones. Finally, we study the seller's pricing problem, where we first show that the seller benefits from social learning *ex ante*, i.e., before knowing the quality of her product. Under some conditions, we show that the seller can speed up learning by lowering her price, which is in sharp contrast to Bayesian learning results from private signals as opposed to reviews. Furthermore, we show that the seller's optimal dynamic pricing strategy charges a lower price than the corresponding myopic policy that ignores the effect of pricing on the social learning process.

1 Introduction

Online review sites are playing an increasingly large role in consumers' purchasing decisions. A recent survey by Tripavdvisor, a review site for the hospitality industry, shows that 90% of hoteliers think that reviews are very important for their business and 81% check their reviews at least weekly. Other industries such as online retail, motion pictures, and restaurants have seen

[*]We had very helpful conversations with Ilan Lobel, Dotan Persitz, Yosi Rinott, Kostas Bimpikis, Marco Ottaviani, as well as seminar participants at Informs, MSOM, INSEAD, NetEcon.

[†]Management Science & Engineering, Stanford University (bifrach@stanford.edu)

[‡]Columbia Business School, Columbia University (c.maglaras@gsb.columbia.edu)

[§]Singapore University of Technology and Design (marco_scarsini@sutd.edu.sg)

customers' decisions increasingly influenced by reviews. The proliferation of smartphones is making access to such review sites easier than ever.

Motivated by this trend, this paper studies consumer reviews in the spirit of the literature on social learning. We adopt a stylized model in which a monopolist seller introduces a new product with unknown quality to a market of heterogeneous consumers who have access to product reviews generated by consumers who arrived earlier in time. The key elements in the model are: (a) consumers learn in a Bayesian manner about the product quality from past reviews, (b) the mechanism by which consumers report their reviews resembles that of online review sites, albeit in a simplified way, and (c) consumers have heterogeneous preferences (willingness-to-pay) for the product.

In more detail, the product (or service) features observed attributes, e.g., location, and unobserved attributes that we denote with the term *quality*. To facilitate the Bayesian inference the intrinsic quality of the product is either High (H) or Low (L). The quality experienced by a consumer who purchases the product is a random perturbation around the intrinsic quality, e.g., due to variability in the service delivery process. Both the seller and the consumers are equally uniformed about its realization. As a result the seller's price is not itself an informative signal about quality.

Consumers arrive sequentially over time and make inference about the product quality based on the information available in the market, as described below. Following that, consumers make a once and for all decision of whether to purchase or to forgo the product, depending on the updated quality distribution and on their idiosyncratic preference for the observed attributes that jointly determine their willingness-to-pay. The heterogeneity in preferences is captured by consumers' types that are private information. Prior to each purchase decision, the seller sets a price after making similar inference about the product quality. Consumers and the seller seek to maximize their expected payoff, the latter comprising the discounted revenue contributions from the sequence of consumers.

Buyers report a review stating that they "liked" the product, if their ex-post net utility exceeded that available from the no-purchase option, or that they "disliked" the product, otherwise. The action of consumers who forgo the product are not observed by their predecessors. Reviews are only partially informative due to the heterogeneity in preferences that remains unseen and the fluctuations in the experienced quality of the product. Each agent—consumer or seller—observes the ordered sequence of consumer reviews and their associated prices. Finally, agents are rational and make inferences about the quality from the information available to them in a Bayesian fashion. We formulate this model as a game and analyze its various properties.

From a modeling viewpoint, the formulation of the problem of learning from reviews is novel and practically relevant. The model is flexible and can be extended to account for more subtle forms of consumer heterogeneity, as well as different mechanisms for writing reviews, e.g., where each buyer decides whether to write a review through some random procedure.

Our first set of results studies the effect of consumer reviews in social learning. To start with, we show that the above-mentioned Bayesian learning process is successful in the sense that eventually consumers almost surely learn the intrinsic quality of the product. Specifically, Proposition 3.4 shows that the conditional beliefs converge to a point mass distribution on the true state of the world. Since no-purchase decisions are not observed, the sequence of beliefs is not a martingale, therefore the convergence result requires some extra argument. In particular, we show that non-purchasing consumers have no informational contribution to the learning process, and, as such, their absence does not affect the decisions of the following consumers.

Second, we study some structural properties of the learning trajectory. We begin by analyzing the effect of the order of reviews on consumers' beliefs, highlighting the importance of early reviews in a Bayesian learning setting. The building block of the analysis is a comparison of two posteriors: one posterior after observing one consumer who liked the product followed by a consumer who disliked it, and the other where the two reviews are reported in reverse order. Proposition 4.1 shows that under general conditions on the distribution of the involved random variables, the posterior after the first sequence (like, dislike) will be greater than the posterior after the second sequence (dislike, like). This result is novel in its own right, has implications for the importance of early reviews on the product's demand, and is of possible interest in other models with learning.

Moreover, we show in Proposition 4.2 that the likelihood that the next review will be positive is decreasing with the belief, or in other words that reviews tend to be negative following high quality expectation produced by positive past reviews, and similarly positive reviews tend to follow negative ones. This result agrees with the empirical findings of Talwar, Jurca, and Faltings (2007) who analyze the influence of past ratings on future reviews. Their explanation is that consumers' reference point in evaluating the product is their quality expectation given by past reviews; in contrast we show that this is the result of a self-selection bias resulting from rational consumer behavior. Finally, Proposition 4.4 shows how increased variability in the perturbations around the intrinsic quality slows down learning, demonstrating that operational policies may have significant effect of consumer learning.

Third, we study the effect of social learning on the seller's pricing decisions. Theorem 5.3 shows that ex-ante the seller benefits from social learning in the sense that her ex-ante expected revenue increases if consumers engage in social learning; however, ex-post, i.e., after the true quality has been revealed, the seller is better off if the quality is high, and worse off, otherwise. Bose, Orosel, Ottaviani, and Vesterlund (2006) have argued that social learning benefits the seller in a setting with signals, and the above result justifies their conclusion in a setting with reviews and heterogenous preferences.

The effect of price on the speed of learning is ambiguous: A lower prices stimulates reviews and consequentially learning, but it also increases consumers surplus, which can weaken the signal conveyed by the review. Assuming consumer types are exponentially distributed, Proposition 5.5 shows that the seller can accelerate learning by lowering her price and stimulating more purchases. This result stands in sharp contrast to Bose et al. (2006) who show that when learning from private signals, increasing the price speeds up learning by screening consumers with the highest signals. In that respect, learning from reviews is different than learning from signals, and the two models call for reversed operational controls to affect learning.

Deriving a complete characterization of the optimal pricing policy is intractable. However, an interesting result that we derive in Proposition 5.6 illustrates its structural properties: the optimal dynamic policy always charges a lower price than the corresponding myopic policy that optimizes instantaneous revenues but disregards the effect of price on the learning trajectory. This result is consistent with Theorem 5.3 and Proposition 5.5 and ultimately suggests that the optimal dynamic price is one that incentivizes learning and, moreover, is willing to tradeoff immediate revenues to accelerate learning and extract higher future expected revenues. Finally, contrasting the pricing results to Proposition 4.4 we note that an alternate lever to accelerate learning is for the seller to operationally focus on reducing variability of experienced quality, to the extent that this is possible.

We conclude the paper with an important result motivated by a practical consideration. Namely, while it is possible for consumers to extract and study the chronological sequence of reviews from review aggregator sites, this is cumbersome and consumers may be more likely to react to aggregate review information, i.e., the cumulative numbers of positive and negative reviews, as opposed to detailed review sequence information. There is empirical support for that observation in various papers, e.g., Luca (2011). This partial information affects in a crucial way the learning dynamics, but, as Proposition 6.2 establishes, asymptotic learning continues to occur in this restricted information setting. This result is of interest in its own right and its derivation and bounding approach, which leverage Proposition 4.1, may be of interest in related settings. A model of social learning from signals where agents only observed unordered samples from the action history is studied in Smith and Sørensen (2008).

Our model retains some aspects of the literature on social learning that started with Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992) who show that social learning may fail—even asymptotically—resulting in herding to a bad outcome. The informational structure in our model differs from theirs considerably. In our case, consumers have no private information that is associated with the state of the world. Namely, if one lets all consumers exchange any information they may have *before* any of them makes a purchasing decision, they would be equally uninformed about the state of the world.

Many papers have extended these principal models. Smith and Sørensen (2000) show that asymptotic learning holds if agents' signals have unbounded strength. Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) consider agents who are embedded on a general social network and find conditions on the social network and signal structure for asymptotic learning. Herrera and Hörner (2013) consider a model where, like in this paper, agents make a binary choice—"buy" and "not buy"—but only the "buy" decision can be observed by predecessors. They show that asymptotic learning occurs when signals have unbounded strength and agents observe the time passed since the product was launched. Asymptotic learning is at the core of the model in Acemoglu, Bimpikis, and Ozdaglar (2013) where agents can collect information by forming costly communication links or by delaying their irreversible action. Goeree, Palfrey, and Rogers (2006) consider a model where agents make choices sequentially and their payoff depends not only on the state of the world and their action, but also on an idiosyncratic privately observed shock. Like in our model, this heterogeneity of the consumers makes herding phenomena impossible. The above-mentioned papers do not consider pricing decisions.

Many papers have considered social learning with word-of-mouth communication. Banerjee and Fudenberg (2004) consider such communication with repeated interactions, however they do not model word-of-mouth explicitly. Bergemann and Välimäki (1997) consider a duopoly and heterogeneous consumers on a line, who report their experienced utility and analyze the resulting price path. Bose et al. (2006) were the first to consider pricing problems in the presence of social learning. In their model agents have private signals, and this marks the main difference with respect to the model we study here. Candogan, Bimpikis, and Ozdaglar (2012) study pricing strategies for a seller when consumers are part of a social network and face network externalities, but do not consider social learning. Ifrach, Maglaras, and Scarsini (2012) have the closest informational structure to ours, however their consumers are not Bayesian and their main focus is on the seller's pricing decision.

The paper is organized as follows. Section 2 introduces the Bayesian learning model from reviews, and Sections 3 and 4 establish the asymptotic learning result and study the structural properties of the learning trajectory, respectively. Section 5 examines the seller's pricing problem, and finally Section 6 extends the asymptotic learning result to the case where the review sequence is unobservable.

Notation

Given any sequence $\{X_t\}$ of i.i.d. random variables, the distribution function of X_1 is denoted by F_X, its survival function by $\bar{F}_X$, its density by f_X, that is,

$$F_X(t) = \int_{-\infty}^{t} f_X(s)\ \mathrm{d}s = 1 - \bar{F}_X(t).$$

The symbol $\mathbb{1}_{\{A\}}$ denotes the indicator of the event A.

2 The model

A monopolist introduces a product or service of unknown quality to a market of heterogeneous consumers who will learn about this quality through a social learning mechanism and will make their respective purchase decisions accordingly. Specifically, the monopolist introduces a product of intrinsic quality Q that for simplicity is assumed to take one of two possible values L or H, where $H > L$. The intrinsic quality of the product is determined through a random draw at time $t = 0$, and takes value H with probability π_0 and value L with probability $1 - \pi_0$. The realization of Q is assumed to be unknown to the potential consumers and to the seller (cf. Remark 2.2).

Consumers arrive sequentially and are indexed by their arrival time $t \in \{1, 2, \dots\}$. They are heterogeneous with respect to their preference for the product. Consumer t's preference is represented by his type Θ_t. Types are i.i.d. random variables with a strictly increasing continuous distribution function F_Θ. The type Θ_t is known to consumer t, but not to the other consumers. A consumer t who purchases the product will experience a quality level $Q_t = Q + \varepsilon_t$, where ε_t is a random fluctuation around the nominal and initially unknown quality level Q. This fluctuation could be the result of variations in the product itself, or even variations in the way individuals experience or perceive quality. The random variables ε_t are i.i.d. with a continuous, zero mean distribution function F_ε, independent of the types Θ_t.

As an example one could consider a dinning experience in a new restaurant. Some of the characteristics of the restaurant are observable and as such reflected in the consumer type Θ_t, but the quality of the cuisine or of the overall experience is still unknown. In the restaurant example, fluctuations in the experienced quality could be the result of variability in the quality of the prepared menu items, of the table service rendered, the ambience, etc.

Each consumer t makes a once-and-for-all purchase decision denoted by $B_t \in \{0, 1\}$: he either buys the product ($B_t = 1$) or does not buy it ($B_t = 0$). If a consumer buys the product, his payoff is given by the following simple additive form

$$V_t := \Theta_t + Q_t - p_t, \tag{2.1}$$

where p_t is the price of the product at time t. If he chooses to forgo the product his payoff is given by 0, without loss of generality. That is, the payoff of consumer t is given by $B_t V_t$. Whatever the purchase decision is, consumers do not revisit it in later periods.

If $B_t = 1$, once consumer t has bought the product and experienced its quality, he publicly posts a review R_t, where

$$R_t = \begin{cases} \text{👍} & \text{if } B_t = 1, \text{ and } V_t \geq 0, \\ \text{👎} & \text{if } B_t = 1, \text{ and } V_t < 0, \\ \text{✗} & \text{if } B_t = 0. \end{cases}$$

The review is either positive (thumbs up), if consumer t's ex-post net-utility is non-negative (i.e., exceeds that of the no-purchase option), and negative (thumbs down), otherwise.[1]. Although consumers who do not buy the product do not review it, it is useful to suppose that they report a blank review ✗. We will show that, in contrast to models with private signals, in our model ✗'s are not informative.

Define the time indices of consumers who choose to purchase the product

$$\tau_1 = \min(t \mid B_t = 1) \quad \text{and} \quad \tau_k = \min(t \mid t > \tau_{k-1}, B_t = 1)$$

and let the corresponding (price, review) histories be, for $\tau_k \leq t < \tau_{k+1}$,

$$h_t = (p_{\tau_1}, R_{\tau_1}, \ldots, p_{\tau_k}, R_{\tau_k}), \tag{2.2}$$

$$h_t^+ = (h_t, p_{t+1}) = (p_{\tau_1}, R_{\tau_1}, \ldots, p_{\tau_k}, R_{\tau_k}, p_{t+1}). \tag{2.3}$$

At each period t the seller observes h_{t-1}, while consumer t observes history h_{t-1}^+. Let $\mathscr{H}_t$ the set of all histories h_t and $\mathscr{H}_t^+$ the set of all histories h_t^+. Note that the realization of Θ_t and ε_t is never revealed to consumers different from t. Here consumers *generate* signals about the quality of the product when they review it, whereas in the literature on social learning with signals, they *reveal* privately held information when making a purchase decision.

The form of the utility function, review decision, information structure, and all the distributions of the relevant random variables are assumed to be common knowledge.

Remark 2.1 (Reviews)**.** In practice, some online review systems follow similar binary reviews (e.g., youtube.com, ebay.com), but most follow a finer review scale with 5-star scale being the most common (e.g., Amazon.com, Yelp.com). However, it is has been well documented that reviews typically follow a bimodal distribution even with finer scales, indicating that consumers tend to submit binary reviews regardless of the review scale[2] (e.g., see Hu, Pavlou, and Zhang, 2006).

Remark 2.2 (Seller's information)**.** The seller does not hold any private information about the quality of the product or value of the disturbances around it, but rather is as informed as consumers

[1] All our results extend to the case when each buying consumer writes a review with probability $\eta \in (0, 1]$, as long as this is independent of the reviews of past consumers and the experienced utility V_t.

[2] An extreme example of bimodal reviews is used in *Rotten Tomatoes*, a popular review site for movies, where reviews on any scale are converted to either a "ripe" tomato or a "rotten" one, corresponding to thumbs up and thumbs down, respectively.

are on the realization of the product's quality, i.e., the seller does not know *ex-ante* whether the product is of high quality. Under this assumption, the price is not itself a signal of quality that consumers could use to learn Q. The analysis of the paper would readily extend to a setting where the seller may in fact know Q but due to external factors, such as competition or other marketing considerations, the price of the product is again non indicative of the realization of Q.

Strategies. Consumer t's pure strategy is a measurable functions $B_t : \mathscr{T} \times \mathscr{H}_{t-1}^{+} \to \{0, 1\}$, where $\mathscr{T}$ is the set of consumer types. Call the set of these strategies $\mathscr{B}_t$. Given his type and conditional on h_{t-1}^{+}, consumer t chooses either to buy or forgo the product to maximize his expected payoff

$$B_t \mathsf{E}[V_t | \Theta_t, h_t^{+}].$$

Note that V_t is independent of the actions of the other consumers, including the ones taken by consumers $1, \ldots, t-1$; past actions affect player t's inference, not his payoff. We call $B = (B_1, B_2, \ldots)$ the profile of all consumers' strategies.

At every time t, based on the history h_{t-1}, the seller chooses the price for that period, denoted by p_t. Hence, the seller's strategy set Φ is the set of all measurable functions $\phi : \mathscr{H} \to \mathbb{R}_+$, where $\mathscr{H} := \cup_{t \geq 0} \mathscr{H}_t$ is the set of all possible histories. The seller seeks to maximize her expected discounted profit given by

$$\mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t p_t B_t\right], \tag{2.4}$$

where $\beta \in (0, 1)$ is the discount factor. It is assumed that the seller's production costs are zero. We comment in Section 3 about the case with positive marginal costs.

Equilibrium. We formalize the model as a Bayesian game. There is a probability space $(\Omega, \mathscr{F}, \mathsf{P})$ on which all the random quantities are defined. Moreover, the probability measure P is common across players. The set of players is $\mathscr{L} := \mathbf{S} \cup \mathbb{N}_+$, where $\mathbf{S}$ is the seller and the remaining players are the consumers. Consumer t's type is given by Θ_t, while the seller's type is degenerate. The set of payoff-relevant states of the world is $\{H, L\}$. Each player forms beliefs on the state of the world given the history available to him/her and the strategies played by other players. In a Bayesian Nash equilibrium we require that these beliefs are computed using Bayes rule, whenever possible.

Definition 2.3. A strategy profile (ϕ^*, B^*) is a perfect Bayesian equilibrium (PBE) if

1.

$$\mathsf{E}\left[\sum_{t=k}^{\infty}\beta^t\phi^*(h_t)B_t^*\right] \geq \mathsf{E}\left[\sum_{t=k}^{\infty}\beta^t\phi(h_t)B_t^*\right] \tag{2.5}$$

for $k = 0, 1, \ldots$ and for all $\phi \in \Phi$;

2.

$$B_t^*(\Theta_t, h_{t-1}^+)\mathsf{E}[\Theta_t + Q_t - \phi^*(h_t)|\Theta_t, h_{t-1}^+] \geq B_t(\Theta_t, h_{t-1}^+)\mathsf{E}[\Theta_t + Q_t - \phi^*(h_t)|\Theta_t, h_{t-1}^+] \tag{2.6}$$

for all $t \geq 1$ $h_{t-1}^+ \in \mathscr{H}_{t-1}^+$ and $B_t \in \mathscr{B}_t$;

3. Beliefs are computed via Bayes' rule, whenever possible.

A peculiar feature of this game is that consumers act once and their actions do not affect the payoffs of consumers who arrive prior to them. As a result, their optimal action is always their best response, which always exists. Consequently, the equilibrium computation can be collapsed into the seller's optimization problem, which can be formulated as a dynamic program with the belief as a continuous state variable (see Section 5). An equilibrium of the game exists whenever this dynamic program attains a solution, which is not easily verified in a continuous state space specification. However, an ϵ-equilibrium (see Osborne and Rubinstein, 1994, Exercise 108.1), where the seller's strategy is arbitrarily close to optimal and the buyers' strategies are exactly optimal, always exists (see Bertsekas and Shreve, 2007, Proposition 9.19). Moreover, it can be found by successive iterations of the Bellman operator (see Bertsekas and Shreve, 2007, Proposition 9.14).

3 Asymptotic learning

This section shows that *asymptotic learning* occurs, that is, both the consumers and the seller eventually learn the quality of the product in any PBE of this game.

Given history h, define

$$\pi(h) := \mathsf{P}(Q = H|h). \tag{3.1}$$

So, for instance, $\pi(h_{t-1}^+)$ is the belief of consumer t that the quality Q is high based on the history h_{t-1}^+ and prior to making his purchase decision. Note that $\pi(h_t^+) = \pi(h_t)$, since the seller's posted price does not reveal information not already contained in h_t. We frequently use the shorthand notation

$$\pi_t := \pi(h_t). \tag{3.2}$$

The belief determines the buyers' purchase decision. Consumer t will buy the product if and only if the expected net utility from buying is greater than zero:

$$\begin{aligned} \mathsf{E}[V_t|h_{t-1}^{+}, \Theta_t] &= \Theta_t + \mathsf{E}[Q_t|h_{t-1}^{+}] - p_t \\ &= \Theta_t + \pi_{t-1}H + (1-\pi_{t-1})L - p_t \\ &\geq 0, \end{aligned}$$

or, alternatively, if and only if $\Theta_t \geq \theta(\pi_{t-1}, p_t)$, where

$$\theta(\pi, p) := p - (\pi H + (1-\pi)L) = p - \mathsf{E}_{\pi}[Q]. \tag{3.3}$$

Note that ε_t does not affect the purchase decision since it has zero mean and is independent of the history and Θ_t.

Given the purchase criterion, in each period t the seller faces an expected demand function $\bar{F}_{\Theta}(\theta(\pi_{t-1}, p_t))$. For each belief, define the set of prices for which the probability of purchase is strictly positive,

$$\mathscr{P}(\pi) := \{p | \bar{F}_{\Theta}(\theta(\pi, p)) > 0\}. \tag{3.4}$$

If the seller chooses a price outside $\mathscr{P}(\pi)$ when the belief is π, the probability of purchase is zero and the belief remains unchanged.

The following assumption will hold throughout the paper.

Assumption 3.1. (a) $\text{supp}(\varepsilon_t) = \mathbb{R}$.

(b) $\bar{F}_{\Theta}(-L) > 0$.

Assumption 3.1(a) assures that there is always a positive probability that a consumer will derive positive or negative net utility from buying the product, irrespective of whether the intrinsic quality Q is high or low and of the value of the consumer type[3]. Assumption 3.1(b) is necessary in order for social learning to occur. It assures that $0 \in \mathscr{P}(\pi)$ for all beliefs, i.e., the seller can choose a nonnegative price under which some consumers will always purchase the product, regardless of the true state of the world. The assumption trivially holds when the distribution of Θ_t is unbounded, and it is reminiscent of Assumption 3 in Goeree et al. (2006). Assumption 3.1(b) is fundamental in the next proposition that establishes that the stream of purchasing consumers and reviews will never cease in equilibrium.

[3]This assumption allows us to focus on the more interesting case where a single review is never fully informative, which is the case when the support of ϵ_t is small.

Proposition 3.2. *The prices set by the seller in equilibrium are such that the probability of purchase is positive in every period, i.e.,* $p_t \in \mathscr{P}(\pi_{t-1})$ *for all* $t \geq 1$*, with probability* 1.

The object of interest in our analysis is the belief π_t, which is a random variable in $[0, 1]$. Lemma A.1 demonstrates some basic properties of the belief. First, starting from any history h_t with corresponding price $p_{t+1} \in \mathscr{P}(\pi_t)$ the belief increases after a 👍 and decreases after a 👎,

$$\pi((h_t, p_{t+1}, 👎)) \leq \pi(h_t) \leq \pi((h_t, p_{t+1}, 👍)),$$

and this inequality is weak only when $\pi(h_t) \in \{0, 1\}$. This is expected, since the ex-post net utility of a buyer is higher when the intrinsic quality is high, which itself increases the probability of a positive review. The opposite result holds in case of 👎.

Following a no-buy, the posterior will not change, since the sequence of reviews remains unchanged. However, this would have been the case even if ✗ reviews had been observable. A no-buy decision of consumer $t + 1$ merely reveals that his type is lower than $\theta(\pi(h_t), p_{t+1})$. This carries no information about the quality of the product. This observation is in sharp contrast with the literature on social learning from signals, where any action can be informative by revealing the agent's private signal. The next proposition plays a significant role in subsequent analysis.

Proposition 3.3. *We have*

$$\mathsf{E}[\pi_{t+1} | h_t] = \pi_t. \tag{3.5}$$

It is worth noting that $\{h_t\}_{t \geq 1}$ is not a filtration, since ✗ reviews are not observable, therefore (3.5) is not the usual martingale property of posterior distributions. Nevertheless, as argued above, ✗ reviews carry no information on the quality of the product, and so conditioning on the full history, which is a filtration, does not alter the belief.

We next present the main result of this section.

Proposition 3.4. *If* $\pi_0 \in (0, 1)$*, then* $\pi_t \to \mathbb{1}_{\{Q=H\}}$ *with probability* 1.

The result builds on two observations. First, as long as consumers purchase the product, the drift of the belief process is positive when the quality is high and negative drift when it is low. Second, the seller will always choose a price such that the probability of purchase is strictly positive, as asserted in Proposition 3.2. As the number of reviews grows large, the posterior will converge and correctly identify the intrinsic quality Q of the product.

Proposition 3.4 can be modified to account for positive marginal production cost $c > 0$. The result would continue to hold under a stronger Assumption 3.1(b) stating that $\bar{F}_\Theta(-L - c) > 0$. Under the stricter condition we would need $c \in \mathscr{P}(0)$ and Proposition 3.2 would still hold—the probability of purchase will be positive even under the lowest possible belief.

4 Structural properties of the learning trajectory

This section explores structural properties of the dynamics of the learning process, and, in part, provides a theoretical foundation for several empirical findings on the effect of consumer reviews. Specifically, in the spirit of comparative statics analysis, we compare outcomes of the social learning process under a single change in model parameters or review histories.

To start with, we explore the effect of the order in which reviews are submitted on agents' beliefs. Identical reviews may carry different information, because the reviewers observed different histories. Thus, reviews are not exchangeable random variables and potentially carry different weights on the posterior distributions of a future consumer. Do earlier or later reviews carry more weight in forming a posterior belief? We begin by comparing the belief resulting from two histories: one where a positive review is followed by a negative one, $h_{\text{👍},\text{👎}} := (p, \text{👍}, p', \text{👎})$ and the reverse sequence $h_{\text{👎},\text{👍}} := (p', \text{👎}, p, \text{👍})$. The following proposition shows that, under some weak assumption on the distribution of ε_t, the earlier review is more influential[4].

Proposition 4.1. *If f_ε is log-concave, then for any histories $h' \in \mathscr{H}_{t'}$ and $h'' \in \mathscr{H}_{t''}$ with $t', t'' \in \mathbb{N} \cup \{0\}$ we have*

$$\pi(h', h_{\text{👍},\text{👎}}, h'') \geq \pi(h', h_{\text{👎},\text{👍}}, h'').$$

This result holds for any distribution of types. Proposition 4.1 should be understood in the context of consumers' self-selection given their unobservable types (for example see Li and Hitt, 2008). As reviews and prices vary over time, the corresponding cutoff for purchase, $\theta(\pi_t, p_{t+1})$ varies as well. Holding Q fixed, a consumer who purchases when the cutoff $\theta(\pi_t, p_{t+1})$ is low is more likely to be disappointed than a consumer who purchases when the cutoff is high, since the latter is more likely to be a high type. As a result, 👎 is a weaker negative signal in $h_{\text{👍},\text{👎}}$ than in $h_{\text{👎},\text{👍}}$. Similarly, a purchasing consumer with a low cutoff is less likely to be satisfied, because on average his type is low, resulting is a stronger positive effect of the 👍 in $h_{\text{👍},\text{👎}}$ than in $h_{\text{👎},\text{👍}}$. In conclusion, earlier reviews have a higher effect on the posterior belief.

Self-selection drives another result summarized in the next proposition: namely, that the likelihood that the next review will be positive is decreasing with the belief, or with the expected quality.

Proposition 4.2. *For any $\pi \in [0,1]$, $p \in \mathscr{P}(\pi)$ and $q \in \{L, H\}$ we have that*

$$\mathsf{P}\big(R_{t+1} = \text{👍} \,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\big)$$

is decreasing in π and $\mathsf{P}\big(R_{t+1} = \text{👎} \,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\big)$ is increasing in π.

[4] As will become clear in the proof, Proposition 4.1 could be stated under more general assumptions. We present it this way for the sake of simplicity.

Talwar et al. (2007) empirically analyze the influence of past ratings on future reviews. Their results confirm Proposition 4.2: reviews tend to be negative following high quality expectation produced by positive past reviews, and similarly positive reviews tend to follow negative ones. Talwar et al. (2007) offer the explanation that consumers' reference point in evaluating the product is their quality expectation given by past reviews.

Our model provides a different explanation. At times when the belief is high, many consumers choose to purchase, including many with lower types who are more likely to have a negative experience and as such write negative reviews. Similarly, only high types buy the product when the belief is low, but they are more likely to write positive reviews. In particular, in our model consumers do not form a reference point that is affected by past reviews, but instead always compare their net utility to the value of the no purchase option that is zero. This negative serial correlation in reviews stems from consumers' self-selection in their purchase decision and the fact that the types of "other" consumers are unobservable.

Finally, our next result studies the effect of the variability of quality disturbances, $\{\varepsilon_t\}_{t\geq 1}$ on the speed of the social learning process. Recall that given h_{t-1}, the belief π_t is a random variable that can take three values following a 👍, 👎 and ✗. Proposition 3.4 shows that as t grows large, this random variable converges either to 0 or to 1. Learning is fast if the random variable converges quickly to either of these points, i.e., if the dispersion of π_t is large. With that in mind, we formalize the notion of learning speed using the convex order of random variables, whose definition is provided below,

Definition 4.3. Random variables X and Y are ordered by the *convex order* $X \leq_{\text{cx}} Y$, if

$$\mathsf{E}[\psi(X)] \leq \mathsf{E}[\psi(Y)]$$

for all convex functions ψ for which the expectations exist.

Let a and b denote different model specifications, i.e., different primitive assumptions such as the noise distribution, different price, etc. We say that learning is faster under specification a than specification b, if $\pi_t^b \leq_{\text{cx}} \pi_t^a$ starting from any history h_{t-1}. We analyze the speed of learning under two simplifying assumptions. First, that consumers are homogeneous with respect to their types, i.e., that the distribution of Θ_t is degenerate, and that the price charged by the seller is not a subsidy price, i.e., is not smaller than L.

Proposition 4.4. *Suppose Θ_t is degenerate and consider two sequences $\{\varepsilon_t^a\}_{t\geq 1}$ and $\{\varepsilon_t^b\}_{t\geq 1}$ of i.i.d. random variables with symmetric, zero mean distributions that single cross with the sequence of signs of $F_{\varepsilon^a} - F_{\varepsilon^b}$ being $-\,+$. Fix any history h_{t-1} and let π_t^a and π_t^b denote the beliefs under specifications a, b. Then, $\pi_t^b \leq_{\text{cx}} \pi_t^a$.*

Notice that the assumptions on the distributions of ε_t^a and ε_t^b imply $\varepsilon_t^a \leq_{\text{cx}} \varepsilon_t^b$. The result shows that learning is faster in settings where the experienced qualities are closer to the intrinsic quality Q, since, intuitively, less variability in quality disturbances results in more accurate reviews that speed up social learning. In the extreme case when quality disturbances are degenerate, learning is resolved with certainty after a single 👎 for arbitrary type distributions, since a negative review will never be submitted under $Q = H$.

One implication of this result is that a seller can speed up the social learning process by focusing on reducing the variability of her service or production process. This will tend to result in consumers experiencing a similar quality product, which accelerates the learning process. The opposite holds true if the process variability is increased.

5 Pricing

5.1 Is social learning beneficial to the seller?

Consider first the seller's problem of maximizing the expected revenue from a single consumer with belief π. Concretely, she is interested in maximizing the revenue function

$$W(\pi, p) := pD(\pi, p),$$

where

$$D(\pi, p) = \bar{F}_\Theta(\theta(\pi, p)) = \bar{F}_{\Theta + \pi H + (1-\pi)L}(p) \tag{5.1}$$

is the demand function. Having defined W, we can rewrite the seller's objective (2.4) as

$$\mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t W(\pi_{t-1}, p_t)\right].$$

The next assumption is standard in the revenue management literature (see, e.g., Lariviere and Porteus, 2001).

Assumption 5.1. The revenue function $W(\pi, p)$ has a unique global maximum in p for all $\pi \in [0, 1]$.

Lariviere (2006) shows that W has a unique global maximizer if the generalized failure rate of $\Theta + \pi H + (1 - \pi)L$ is increasing, where the generalized failure rate of X computed at x is $x f_X(x)/\bar{F}_X(x)$.

Let $p^*(\pi) := \arg\max_{p \in \mathbb{R}_+} W(\pi, p)$ be the optimal price and $W^*(\pi) := W(\pi, p^*(\pi))$ the maximal revenue.

Lemma 5.2. *The function $W^*(\pi)$ is convex in π.*

This result is the counterpart of Bose et al. (2006, Proposition 3) that showed convexity of the revenue function for a model with homogeneous preferences and learning from signals. Lemma 5.2 is key in proving our next result.

Theorem 5.3. *The expected discounted revenue of the seller under social learning is greater than the expected discounted revenue when consumers do not engage in social learning, i.e.,*

$$\sup_{\phi \in \Phi} \mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t W(\pi_t, \phi(\pi_t))\right] \geq \mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t W^*(\pi(h_t))\right] \geq \mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t W^*(\pi_0)\right] = \frac{W^*(\pi_0)}{1-\beta}.$$

Theorem 5.3 shows that social learning is beneficial for the seller *ex-ante*, before knowing the true quality of the product. *Ex-post* the seller loses when the quality is low (with probability $1-\pi_0$) and gains when it is high (with probability π_0). To gain intuition, compare the maximal revenue extracted from consumer 1, $W^*(\pi_0)$, and the expected maximal revenue extracted from consumer 2, $\mathsf{E}[W^*(\pi_1)]$. Combining Proposition 3.3, Lemma 5.2, and Jensen's inequality, we can see that $\mathsf{E}[W^*(\pi_1)] \geq W^*(\mathsf{E}[\pi_1]) = W^*(\pi)$. Thus, on average the seller extracts more revenue from the second consumer than from the first, due to social learning. This argument can be repeated to establish Theorem 5.3.

This result formalizes and demonstrates the claim of Bose et al. (2006) who argue that social learning benefits the seller in a setting with signals. We point out here that our result does not depend on our particular learning model; it would hold for any learning process with the same preference model as long as Proposition 3.3 holds. In particular, it could be adapted to a setting where consumers are heterogeneous in preferences and learn from signals.

5.2 Optimal pricing, myopic pricing and speeding up learning

Social learning benefits the seller. We consider next how the seller can leverage her pricing control to account for and optimally affect the social learning process. We first study whether the seller can speed up social learning by altering the price to control the dispersion of posteriors, building on the discussion preceding Proposition 4.4. We explore this under the following assumption.

Assumption 5.4. Consumer types are exponentially distributed with mean $1/\lambda$, i.e., $\Theta_t \sim \exp(\lambda)$.

Proposition 5.5. *Under Assumption 5.4, for any history h_{t-1} and for all prices p and p' such that $p > p' \geq \mathsf{E}_{\pi_{t-1}}[Q] = \pi_{t-1}H + (1-\pi_{t-1})L$, we have that $\pi(h_{t-1}, p, R_t) \leq_{\text{cx}} \pi(h_{t-1}, p', R_t)$.*

Proposition 5.5 shows that when types are exponentially distributed, the learning speed increases as the seller lowers the price. A price reduction by the seller has a dual effect on the learning process. First, the probability of purchase and of a review submission increases, and this supports faster learning. Second, the surplus of each purchasing consumer increases, and this can have a delaying effect on social learning, because  becomes rare, even if the quality is low, weakening the strength of the signal the review conveys. The latter is also the reason why Proposition 5.5 restricts attention to the prices where not all consumers purchase, e.g., $p' \geq \mathsf{E}[Q]$. When all consumers purchase, an additional reduction in the price only changes consumers' surplus and does not promote learning.

Next, we consider the seller's optimal pricing strategy ϕ^*. Recall that under this strategy the seller can adjust the price dynamically after each review. We focus on equilibrium pricing policies that are Markov in the belief, thus, with some abuse of notation, we let the pricing decision depend on beliefs instead of histories, $\phi : [0,1] \to \mathbb{R}_+$. This assumption is innocuous; the belief process $\{\pi_t\}_{t\geq 1}$ is Markov, and one can show that for any optimal non-Markov strategy there exists a Markov strategy yielding the same revenue, see Bertsekas and Shreve (2007, Proposition 9.1).

Building on that and given the buyers' purchase strategy, we can write the seller's optimization problem as a dynamic program with value function $v(\pi, \phi)$ being the total expected revenue from implementing pricing policy ϕ starting from belief π, namely,

$$v(\pi_0, \phi) := \mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t W(\pi_t, \phi(\pi_t))\right].$$

The equilibrium policy is the optimal policy

$$\phi^* = \operatorname{argmax} v(\pi_0, \phi)$$

and $v^*(\pi) = v(\pi, \phi^*)$. It is convenient to also consider the corresponding Bellman equation

$$v^*(\pi) = \max_{p\in\mathbb{R}_+} \left\{W(\pi, p) + \beta \mathsf{E}[v^*(\pi_{t+1}) | (\pi_t, p_{t+1}) = (\pi, p)]\right\}. \tag{5.2}$$

This dynamic program does not exhibit a simple solution, but nevertheless leads to interesting structural insights regarding the optimal pricing policy and its connections to social learning. We compare the equilibrium pricing strategy to a *myopic* pricing policy $\phi^{\text{myopic}}(\cdot)$ that maximizes in each period the instantaneous revenue, given the belief, but does not account for social learning; it *responds* to variations in the beliefs, but does not seek to *affect* them.

Proposition 5.6. *Let Assumption 5.4 hold. Then*

$$\phi^*(\pi) \leq \phi^{\text{myopic}}(\pi) \textit{ for all } \pi \in [0,1].$$

The seller accounts for the social learning process by lowering the price that, in turn, increases the likelihood of purchase. In the exponential case the greater likelihood of purchase is not offset by the resulting increase in consumer surplus that may, as argued above, blur the review signal.

It is interesting to compare this result with the dynamic pricing results obtained in Bose et al. (2006, Proposition 4, in particular). They find that the equilibrium pricing policy may charge a higher price than the myopic one in order to speed up social learning. Bose et al. (2006) consider a model of where consumers observe private signals and purchase decisions, not reviews. To speed learning in this setting, the seller chooses a price under which only consumers with high signals buy the product, and this is achieved by charging a high price.

The sharp distinction in the structure of the optimal pricing strategy demonstrates that the two informational models of social learning—the one with private information signals and the one with consumer reviews—are inherently different and call for separate treatment when addressed from the operational perspective.

6 Unobservable review sequence

Most review sites saliently display the average review or the cumulative number of positive and negative reviews, e.g., the average star rating on Amazon or the number of 👍 and 👎 on youtube. Information about the sequence in which these reviews were made is available on some sites, but obtaining this information requires additional effort from consumers. Therefore, it is likely that most of them focus their inference on the cumulative number of reviews, a result supported empirically in a number of papers, e.g., Luca (2011).

With that in mind, in this section we study an alternative informational structure where consumers observe the cumulative number of liking and disliking reviews, but not the sequence of reviews that generated that result nor the sequence of prices paid by the reviewers. Namely, we will assume that consumer t observes information $\hat{h}_{t-1}^{+} = (U_{t-1}, D_{t-1}, p_t)$, where

$$U_t = \#\{R_i = \text{👍}, i \in \{1, \dots, t\}\},$$
$$D_t = \#\{R_i = \text{👎}, i \in \{1, \dots, t\}\},$$

where p_t is the price offered to consumer t. We also restrict the seller into selecting a static price p at time $t = 0$, i.e., $p_t = p_0$ for all times t, which is a realistic assumption in many settings[5].

The sequence $\{\hat{h}_t^{+}\}_{t \geq 1}$ is not a filtration, and moreover we do not obtain a result equivalent to Proposition 3.3, i.e., $\mathsf{E}[\pi_{t+1} | \hat{h}_t^{+}] \neq \pi_t = \pi(\hat{h}_{t-1}^{+})$. In the absence of the sequence information,

[5] If prices were dynamic but unobservable by consumers, it would add a significant layer of complexity by having consumers form expectations about the possible price paths.

the posterior belief given $\hat{h}_t^+$ is the *expected* posterior over all review sequences consistent with $\hat{h}_t^+$. Formally, define

$$\widehat{\mathscr{H}}^+(\hat{h}_t^+) := \left\{ h_t^+ \,\middle|\, \sum_{i=1}^{t} \mathbb{1}_{\{R_i = \text{👍}\}} = U_t, \sum_{i=1}^{t} \mathbb{1}_{\{R_i = \text{👎}\}} = D_t, p_i = p \text{ for } i = 1, \ldots, t \right\}$$

to be the set of histories that are consistent with information $\hat{h}_t^+$. The posterior is given by

$$\pi(\hat{h}_t^+) = \mathsf{P}(Q = H | \hat{h}_t^+) = \sum_{h^+ \in \widehat{\mathscr{H}}^+(\hat{h}_t^+)} \mathsf{P}(Q = H | h^+) \mathsf{P}(h^+ | \hat{h}_t^+) = \sum_{h^+ \in \widehat{\mathscr{H}}^+(\hat{h}_t^+)} \pi(h^+) \mathsf{P}(h^+ | \hat{h}_t^+). \quad (6.1)$$

This posterior computation is significantly more involved than the already complex Bayesian inference of the previous section; consumers have to compute the probability of each consistent sequence of reviews, its corresponding posterior, and finally take the average over these histories. The added computational complexity is huge; it is proportional to the number of unique elements in $\widehat{\mathscr{H}}^+(\hat{h}_t^+)$ which is given by the binomial coefficient indexed by $(U_t + D_t)$ and U_t. Ifrach et al. (2012) show that if consumers do not observe the review sequence but follow a non-Bayesian and far simpler ("naive") inference protocol, then they asymptotically learn the unknown product quality.

Despite the increased complexity, we can repeatedly apply Proposition 4.1 on the summands in (6.1) to obtain bounds for the posterior. Define

$$\underline{h}^+(u, d, p) := (\underbrace{p, \text{👎}, \ldots, p, \text{👎}}_{d}, \underbrace{p, \text{👍}, \ldots, p, \text{👍}}_{u}, p)$$

and

$$\overline{h}^+(u, d, p) := (\underbrace{p, \text{👍}, \ldots, p, \text{👍}}_{u}, \underbrace{p, \text{👎}, \ldots, p, \text{👎}}_{d}, p).$$

Corollary 6.1. *Let f_ε be log-concave. If $\hat{h}^+ = (u, d, p)$, then*

$$\pi(\underline{h}^+(u, d, p)) \le \pi(\hat{h}^+) \le \pi(\overline{h}^+(u, d, p)).$$

Corollary 6.1 is key in proving asymptotic learning in this setting. The next result—the main one of this section—shows that asymptotic learning occurs for a range of prices even when the sequence of reviews is not observed.

Proposition 6.2. *Suppose that Assumption 3.1 holds and, in addition, that Θ_t and L are nonnegative for all t. There exists a price $\tilde{p} > L$ such that for all $p < \tilde{p}$ we have $\pi(\hat{h}_t) \to \mathbb{1}_{\{Q=H\}}$ with probability* 1.

Apart form its practical significance in establishing that consumers eventually learn the unknown product quality from cumulative review information, this result is of interest in its own right and its derivation and bounding approach may be of interest in related settings.

A Proofs

Given random variables X, its failure rate is denoted by λ_X and its reverse failure rate by ρ_X, that is,

$$\lambda_X(t) = \frac{f_X(t)}{\bar{F}_X(t)},$$
$$\rho_X(t) = \frac{f_X(t)}{F_X(t)}.$$

Proofs of Section 3

Notice that

$$\mathsf{P}(R_t = r|\pi(h_t) = \pi, Q = q) = \begin{cases} \int_{\theta(\pi,p)}^{\infty} \bar{F}_\varepsilon(p_t - q - x)\ \mathrm{d}F_\Theta(x) & \text{for } r = \text{👍}, & \text{(A.1a)} \\ \int_{\theta(\pi,p)}^{\infty} F_\varepsilon(p_t - q - x)\ \mathrm{d}F_\Theta(x) & \text{for } r = \text{👎}, & \text{(A.1b)} \\ F_\Theta(\theta(\pi, p_t)) & \text{for } r = \text{✗}. & \text{(A.1c)} \end{cases}$$

where $\theta(\pi, p)$ is defined as in (3.3). We can therefore define the shorthand notation

$$G(r, \pi, q, p) := \mathsf{P}(R_t = r|\pi_t = \pi, p_t = p, Q = q) \tag{A.2}$$

and

$$G(r, \pi, p) := \pi G(r, \pi, q, p) + (1 - \pi)G(r, \pi, H, p). \tag{A.3}$$

Proof of Proposition 3.2. It follows from Assumption 3.1(b) that $0 \in \mathscr{P}(0)$. Note that $\mathscr{P}(0) \subset \mathscr{P}(\pi)$ for all π, and so $0 \in \mathscr{P}(\pi)$. By continuity of $\bar{F}_\Theta$ it follows that there exists a price $0 < \delta \in \mathscr{P}(0) \subset \mathscr{P}(\pi)$ for all π, where δ is potentially very small. Consider any equilibrium pricing strategy ϕ and suppose that at some history h_{t-1} the seller chooses a price $p_t \notin \mathscr{P}(\pi_{t-1})$. Clearly, the revenue in that period is zero and $h_t = h_{t-1}$, since a new review cannot possibly be submitted. Therefore, $p_{t+1} = p(h_t) = p(h_{t-1}) = p_t = 0$, and similarly for $t+2, t+3, \ldots$. We conclude that the seller's long run revenue starting from h_{t-1} is 0. However, the seller can charge δ at t and obtain a positive revenue, which contradicts the assumption that $p_t \notin \mathscr{P}(\pi_{t-1})$. □

Lemma A.1 illustrates some properties of the Bayesian updating,

Lemma A.1. (a) *For all $\pi \in [0, 1]$ and $p \in \Phi(\pi)$ we have $G(\text{👍}, \pi, H, p) > G(\text{👍}, \pi, L, p)$.*

(b) *For all* $\pi \in [0,1]$ *and* $p \in \Phi(\pi)$ *we have* $G(\text{👎}, \pi, H, p) < G(\text{👎}, \pi, L, p)$.

(c) *There exist* $\underline{G}, \overline{G} \in (0,1)$ *such that* $\underline{G} \le G(r, \pi, q, p) \le \overline{G}$ *for all* $r \in \{\text{👍}, \text{👎}, \text{✗}\}$, $q \in \{H, L\}$, $\pi \in [0,1]$.

(d) *For all* $q \in \{H, L\}$, $\pi \in [0,1]$ *and price* p *we have* $G(\text{👎}, \pi, q, p) + G(\text{👍}, \pi, q, p) = \bar{F}(\theta(\pi, p))$.

(e) *For any history* h_{t-1}, *we have* $\pi(h_{t-1}, \text{✗}) = \pi(h_{t-1})$.

(f) *Whenever* $\pi(h_{t-1}) \in (0,1)$ *we have*

$$\pi(h_{t-1}, p_t, \text{👎}) < \pi(h_{t-1}) < \pi(h_{t-1}, p_t, \text{👍})$$

for all prices $p_t \in \Phi(\pi_{t-1})$.

Proof. (a) Since $\bar{F}_\varepsilon$ is nonincreasing, we have

$$\begin{aligned} G(\text{👍}, \pi, H, p) &= \int_{\theta(\pi,p)}^{\infty} \bar{F}_\varepsilon(p - H - x)\, \mathrm{d}F_\Theta(x) \\ &> \int_{\theta(\pi,p)}^{\infty} \bar{F}_\varepsilon(p - L - x)\, \mathrm{d}F_\Theta(x) \\ &= G(\text{👍}, \pi, L, p). \end{aligned}$$

(b) Since F_ε is nondecreasing, we have

$$\begin{aligned} G(\text{👎}, \pi, H, p) &= \int_{\theta(\pi,p)}^{\infty} F_\varepsilon(p - H - x)\, \mathrm{d}F_\Theta(x) \\ &< \int_{\theta(\pi,p)}^{\infty} F_\varepsilon(p - L - x)\, \mathrm{d}F_\Theta(x) \\ &= G(\text{👎}, \pi, L, p). \end{aligned}$$

(c) This follows from Assumption 3.1, since there exists a fraction of consumers that would always choose to buy the product, and a different fraction that would always choose not to buy, and since the support of ε is large enough.

(d) Just add (A.1a) and (A.1b) and consider that, given π_t the probability of buying is independent of Q.

(e) In general, by Bayes' rule,

$$\pi(h_{t-1}, r) = \frac{\mathsf{P}(R_t = r | h_{t-1}, Q = H)\pi(h_{t-1})}{\mathsf{P}(R_t = r | h_{t-1}, Q = H)\pi(h_{t-1}) + \mathsf{P}(R_t = r | h_{t-1}, Q = L)(1 - \pi(h_{t-1}))}. \quad \text{(A.4)}$$

Hence

$$\pi(h_{t-1}, \text{✗}) = \frac{G(\text{✗}, \pi(h_{t-1}), H, p)\pi(h_{t-1})}{G(\text{✗}, \pi_t, H, p)\pi(h_{t-1}) + G(\text{✗}, \pi(h_{t-1}), L, p)(1 - \pi(h_{t-1}))} = \pi(h_{t-1}),$$

since $G(\text{✗}, \pi, q, p) = F_\Theta(\theta(\pi, p))$ for all $q \in \{H, L\}$.

(f)

$$\begin{aligned}
\pi(h_{t-1}, \text{👎}) &= \frac{G(\text{👎}, \pi(h_{t-1}), H, p)\pi(h_{t-1})}{G(\text{👎}, \pi(h_{t-1}), H, p)\pi(h_{t-1}) + G(\text{👎}, \pi(h_{t-1}), L, p)(1 - \pi(h_{t-1}))} \\
&< \frac{G(\text{👎}, \pi(h_{t-1}), H, p)\pi(h_{t-1})}{G(\text{👎}, \pi(h_{t-1}), H, p)\pi(h_{t-1}) + G(\text{👎}, \pi(h_{t-1}), H, p)(1 - \pi(h_{t-1}))} \\
&= \pi(h_{t-1}) \\
&= \frac{G(\text{👍}, \pi(h_{t-1}), H, p)\pi(h_{t-1})}{G(\text{👍}, \pi(h_{t-1}), H, p)\pi(h_{t-1}) + G(\text{👍}, \pi(h_{t-1}), H, p)(1 - \pi(h_{t-1}))} \\
&< \frac{G(\text{👍}, \pi(h_{t-1}), H, p)\pi(h_{t-1})}{G(\text{👍}, \pi(h_{t-1}), H, p)\pi(h_{t-1}) + G(\text{👍}, \pi_t, L, p)(1 - \pi(h_{t-1}))} \\
&= \pi(h_{t-1}, \text{👍}).
\end{aligned}$$

where the first inequality follows from (b) and the second from (a). □

Proof of Proposition 3.3. For $t \geq 1$ call

$$h_t^{\text{full}} := (p_1, R_1, \ldots, p_t, R_t) \quad \text{and} \quad h_t^{\text{full+}} := (p_1, R_1, \ldots, p_t, R_t, p_{t+1})$$

the full histories including ✗ reviews. As before $\pi(h_t^{\text{full}}) = \pi(h_t^{\text{full+}})$. Define

$$\pi_t^{\text{full}} = \pi(h_t^{\text{full}}).$$

We have

$$\mathsf{E}[\pi_{t+1}|h_t] = \mathsf{E}[\pi_{t+1}^{\text{full}}|h_t^{\text{full}}] = \pi_t^{\text{full}} = \pi_t,$$

where the first and last equality follow from A.1(e) and the second from fact that π_t^{full} is a Doob martingale. □

The following lemma is needed to prove Proposition 3.4.

Lemma A.2. *Define the function*

$$g(x, y, z) = \log\left(\frac{x}{y}\right)x + \log\left(\frac{z-x}{z-y}\right)(z - x). \tag{A.5}$$

Then, $0 < x \leq y < z < 1$ *implies* $g(x, y, z) \geq 0$ *with equality if and only if* $x = y$.

Proof. All variables in the proof are assumed to satisfy the condition in the statement of the lemma. We will show that g is strictly monotonically decreasing in x and that $g(x, x, z) = 0$. We first show that g is convex with respect to the first argument,

$$\frac{\partial g(x,y,z)}{\partial x} = \log\left(\frac{x}{y}\right) - \log\left(\frac{z-x}{z-y}\right),$$

and,

$$\frac{\partial^2 g(x,y,z)}{\partial x^2} = \frac{1}{x} + \frac{1}{z-x} > 0.$$

It is easy to see that

$$\left.\frac{\partial g(x,y,z)}{\partial x}\right|_{x=y} = 0$$

and that

$$\left.\frac{\partial g(x,y,z)}{\partial x}\right|_{x=x-\delta} < 0$$

for $\delta > 0$ small. Thus by convexity we have, $g(x, y, z) \geq g(y, y, z) = 0$, with equality if and only if $x = y$. □

Proof of Proposition 3.4. Following Proposition 3.3 we have

$$\pi_t = \pi_t^{\text{full}} = \mathsf{P}(Q = H | h_t^{\text{full}}) \to \pi_\infty,$$

using the martingale convergence theorem (see, for instance, Karlin and Taylor, 1975). Since $\pi_t \in [0, 1]$ for all $t \geq 0$, we further conclude that $\pi_\infty \in [0, 1]$, and that $\pi_0 = \mathsf{E}[\pi_t] = \mathsf{E}[\pi_\infty]$.

Recall the definitions of G in (A.2) and (A.3). We have

$$
\begin{aligned}
\mathsf{E}[\log \pi_t | \pi_{t-1}, Q = H] &= \log\left(\frac{G(\text{✗}, \pi_t, H, p)\pi_{t-1}}{G(\text{✗}, \pi_t, p)}\right) G(\text{✗}, \pi_t, H, p) \\
&\quad + \log\left(\frac{G(\text{👎}, \pi_t, H, p)\pi_{t-1}}{G(\text{👎}, \pi_t, p)}\right) G(\text{👎}, \pi_t, H, p) \\
&\quad + \log\left(\frac{G(\text{👍}, \pi_t, H, p)\pi_{t-1}}{G(\text{👍}, \pi_t, p)}\right) G(\text{👍}, \pi_t, H, p) \\
&= \log \pi_{t-1} \qquad \text{(A.6)} \\
&\quad + \log\left(\frac{G(\text{👎}, \pi_t, H, p)}{G(\text{👎}, \pi_t, p)}\right) G(\text{👎}, \pi_t, H, p) \\
&\quad + \log\left(\frac{G(\text{👍}, \pi_t, H, p)}{G(\text{👍}, \pi_t, p)}\right) G(\text{👍}, \pi_t, H, p) \\
&= \log \pi_{t-1} + g\left(G(\text{👎}, \pi_t, H, p), G(\text{👎}, \pi_t, p), F_\Theta(\theta(\pi_t))\right)
\end{aligned}
$$

where the second equality stems from

$$
\begin{aligned}
&G(\text{✗}, \pi_t, H, p) = G(\text{✗}, \pi_t, L, p) \\
&G(\text{✗}, \pi_t, H, p) + G(\text{👎}, \pi_t, H, p) + G(\text{👍}, \pi_t, H, p) = 1,
\end{aligned}
$$

and Lemma A.1(d). The function g was defined in (A.5).

We can show that there exists $\underline{G} > 0$ such that

$$\underline{G} \le G(\text{👎}, \pi_t, H, p) \le G(\text{👎}, \pi_t, p) < F_\Theta(\theta(\pi_t)) < 1.$$

Indeed, the first inequality follows from Lemma A.1(c); the second inequality stems from Lemma A.1(b) and (A.2); the third inequality follows from Lemma A.1(d) and (A.2); the last inequality is a consequence of Assumption 3.1(b).

Define

$$\gamma(\pi) := g\left(G(\text{👎}, \pi, H, p), G(\text{👎}, \pi, p), F_\Theta(\theta(\pi, p))\right). \tag{A.7}$$

Using Lemma A.1(c) and (A.2), we see that

$$G(\text{👎}, \pi_t, H, p) = G(\text{👎}, \pi_t, p)$$

if and only if $\pi = 1$. Therefore, by Lemma A.2,

$$
\begin{aligned}
\gamma(\pi) &> 0 \quad \text{for } \pi \in [0, 1), \\
\gamma(\pi) &= 0 \quad \text{for } \pi = 1.
\end{aligned}
$$

Assume now, by contradiction, that there exist $\delta, \eta > 0$ such that

$$\mathsf{P}(\pi_\infty < 1 - \eta | Q = H) > 2\delta.$$

Then there exists an integer T such that for all $t > T$

$$\mathsf{P}(\pi_t < 1 - \eta | Q = H) > \delta.$$

The function γ is continuous and strictly positive for all $\pi \in [0, 1 - \eta]$, therefore

$$\min_{\pi \in [0,1-\eta]} \gamma(\pi) =: \underline{\gamma} > 0.$$

Finally, using (A.6) iteratively, we get

$$\begin{aligned}
\mathsf{E}[\log \pi_{T+k} | Q = H] &= \log \pi_0 + \sum_{i=1}^{T+k} \mathsf{E}[\gamma(\pi_i) | Q = H] \\
&\geq \log \pi_0 + \sum_{i=T}^{T+k} \mathsf{E}[\gamma(\pi_i) | Q = H] \\
&\geq \log \pi_0 + \sum_{i=T}^{T+k} \mathsf{E}[\gamma(\pi_i) \mathbb{1}_{\{\pi_i \leq 1-\eta\}} | Q = H] \\
&\geq \log \pi_0 + k \underline{\gamma} \delta,
\end{aligned}$$

and we conclude that $\mathsf{E}[\log(\pi_{T+k}) | Q = H] > 0$ by taking k large enough, which contradicts the fact that $\log \pi_t \leq 0$. Therefore, $\mathsf{P}(\pi_\infty < 1 | Q = H) = 0$, or $\mathsf{P}(\pi_\infty = 1 | Q = H) = 1$.

The same argument can be used to prove that $\mathsf{P}(\pi_\infty = 0 | Q = L) = 1$. $\square$

Proofs of Section 4

Define

$$\Gamma(\pi) = \frac{\pi}{1 - \pi}, \tag{A.8}$$

$$\Lambda(r, \pi, p) = \frac{G(r, \pi, H, p)}{G(r, \pi, L, p)}, \tag{A.9}$$

Definition A.3. We say that the condition ILR (increasing likelihood ratio) holds if

$$\Lambda(r, \pi, p)$$

is nondecreasing in π for $r \in \{$ 👍, 👎 $\}$ for any price $p \in \mathscr{P}(\pi)$.

We next discuss some sufficient conditions for ILR, which will depend on the following definitions.

Definition A.4. (a) The distribution of a random variable X is IFR (increasing failure rate) if its failure rate is nondecreasing.

(b) The distribution of a random variable X is DRFR (decreasing reverse failure rate) if its reverse failure rate is nonincreasing.

Proposition A.5. *If the distribution of ε is both IFR and DRFR, then condition ILR holds.*

The proof of Proposition A.5 requires some properties of TP_2 (total positivity of order two), for which the reader is referred to Karlin (1968) and Karlin and Rinott (1980).

Proof of Proposition A.5. Notice that ε is IFR iff its survival function $\bar{F}_\varepsilon$ is log-concave. Write

$$\mathsf{P}(\Theta+\varepsilon > s, \Theta > t) = \int_t^\infty \bar{F}_\varepsilon(s-x)\ \mathrm{d}F_\Theta(x) = \int_{-\infty}^\infty \mathbb{1}_{\{(t,\infty)\}}(x)\bar{F}_\varepsilon(s-x)\ \mathrm{d}F_\Theta(x).$$

Notice that $\bar{F}_\varepsilon$ is log-concave iff $K(s,x) := \bar{F}_\varepsilon(s-x)$ it TP_2. Moreover $L(x,t) := \mathbb{1}_{\{(t,\infty)\}}(x)$ is TP_2. Therefore the convolution

$$\int K(s,x)L(x,t)\ \mathrm{d}F_\Theta(x) = \mathsf{P}(\Theta+\varepsilon > s, \Theta > t)$$

is TP_2. This implies that for $\pi_1 < \pi_2$ we have

$$\begin{aligned}&\mathsf{P}(\Theta+\varepsilon > p-H, \Theta > p-\theta(\pi_2))\mathsf{P}(\Theta+\varepsilon > p-L, \Theta > p-\theta(\pi_1)) \geq \\ &\qquad \mathsf{P}(\Theta+\varepsilon > p-H, \Theta > p-\theta(\pi_1))\mathsf{P}(\Theta+\varepsilon > p-L, \Theta > p-\theta(\pi_2)).\end{aligned} \tag{A.10}$$

Hence

$$G(\text{☞}, \pi_2, H, p)G(\text{☞}, \pi_1, L, p) \geq G(\text{☞}, \pi_1, H, p)G(\text{☞}, \pi_2, L, p),$$

that is,

$$\frac{G(\text{☞}, \pi, H, p)}{G(\text{☞}, \pi, L, p)}$$

is nondecreasing in π.

Next, notice that if ε is DRFR, then its distribution function F_ε is log-concave.

Write

$$\mathsf{P}(\Theta+\varepsilon \leq s, \Theta > t) = \int_t^\infty F_\varepsilon(s-x)\ \mathrm{d}F_\Theta(x) = \int_{-\infty}^\infty \mathbb{1}_{\{(t,\infty)\}}(x)F_\varepsilon(s-x)\ \mathrm{d}F_\Theta(x).$$

Notice that F_ε is log-concave iff $K(s,x) := F_\varepsilon(s-x)$ it TP_2. Moreover $L(x,t) := \mathbb{1}_{\{(t,\infty)\}}(x)$ is

TP_2. Therefore the convolution

$$\int K(s,x)L(x,t)\ \mathrm{d}F_\Theta(x) = \mathsf{P}(\Theta+\varepsilon \le s, \Theta > t)$$

is TP_2. This implies that for $\pi_1 < \pi_2$ we have

$$\begin{aligned}\mathsf{P}(\Theta+\varepsilon \le p-H, \Theta > p-\theta(\pi_2))\mathsf{P}(\Theta+\varepsilon \le p-L, \Theta > p-\theta(\pi_1)) \ge \\ \mathsf{P}(\Theta+\varepsilon \le p-H, \Theta > p-\theta(\pi_1))\mathsf{P}(\Theta+\varepsilon \le p-L, \Theta > p-\theta(\pi_2)).\end{aligned} \tag{A.11}$$

Hence

$$G(\Thumbsdown, \pi_2, H, p)G(\Thumbsdown, \pi_1, L, p) \ge G(\Thumbsdown, \pi_1, H, p)G(\Thumbsdown, \pi_2, L, p),$$

that is,

$$\frac{G(\Thumbsdown, \pi, H, p)}{G(\Thumbsdown, \pi, L, p)}$$

is nondecreasing in π, and therefore ILR holds. □

A stronger yet simpler sufficient condition is the following.

Corollary A.6. *If the density f_ε is log-concave, then ILR holds.*

Proof. If the density f_ε is log-concave, then both the distribution function F_ε and the survival function $\bar{F}_\varepsilon$ are log-concave, therefore the proof of Proposition A.5 can be applied. □

Corollary A.6 shows that ILR is a fairly natural assumption on ε given its interpretation as a mean zero noise around the product's quality. For example, ILR holds if ε is has a normal distribution or a Gumble distribution. We can now prove our result.

Proof of Proposition 4.1. We have

$$\begin{aligned}&\Gamma\left(\pi(h_{\Thumbsup,\Thumbsdown})\right) - \Gamma\left(\pi(h_{\Thumbsdown,\Thumbsup})\right) \\ &\quad = \Gamma(\pi_0)\left[\Lambda\left(\Thumbsup, \pi_0, p\right)\Lambda\left(\Thumbsdown, \pi(p,\Thumbsup), p'\right) - \Lambda\left(\Thumbsdown, \pi_0, p'\right)\Lambda\left(\Thumbsup, \pi(p',\Thumbsdown), p\right)\right]\end{aligned} \tag{A.12}$$

We know from Lemma A.1(f) that

$$\pi(p',\Thumbsdown) \le \pi_0 \le \pi(p,\Thumbsup).$$

Therefore, by the ILR property (Definition A.3),

$$\Lambda\left(\Thumbsup, \pi_0, p\right) \ge \Lambda\left(\Thumbsup, \pi(p',\Thumbsdown), p\right)$$

and

$$\Lambda\left(\Thumbsdown, \pi(,\Thumbsup), p'\right) \ge \Lambda\left(\Thumbsdown, \pi_0, p'\right),$$

which implies that the right hand side of (A.12) is nonnegative. Nonnegativity of the left hand side provides the result $\pi(h_{\text{👍},\text{👎}}) \geq \pi(h_{\text{👎},\text{👍}})$. Note that h' is summarized in π_0 in (A.12). Since the result holds for all π_0, it will hold for all prior histories h'. By monotonicity of the Bayesian update in the belief, the inequality is preserved after history h''. □

Proof of Proposition 4.2. Note that

$$\mathsf{P}\left(R_{t+1} = \text{👍}\,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\right) = G(\text{👍}, \pi, q, p)/\bar{F}_\Theta(\theta(\pi, p)$$

and

$$\mathsf{P}\left(R_{t+1} = \text{👎}\,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\right) = G(\text{👎}, \pi, q, p)/\bar{F}_\Theta(\theta(\pi, p).$$

For ease of notation, we omit the arguments of the cutoff function $\theta(\pi, p)$ in this proof and write θ instead. Consider the derivative

$$\begin{aligned}
&\frac{\partial \mathsf{P}\left(R_{t+1} = \text{👍}\,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\right)}{\partial \pi} \\
&\quad = (H - L)(\bar{F}_\Theta(\theta))^{-2}\left[\bar{F}_\varepsilon(p - q - \theta) f_\Theta(\theta)\bar{F}_\Theta(\theta) - f_\Theta(\theta)\int_{\theta(\pi,p)}^{\infty} \bar{F}_\varepsilon(p - q - x) f_\Theta(x)\mathrm{d}x\right] \\
&\quad = (H - L) f_\Theta(\theta)(\bar{F}_\Theta(\theta))^{-2}\left[\bar{F}_\varepsilon(p - q - \theta)\bar{F}_\Theta(\theta) - \int_{\theta}^{\infty} \bar{F}_\varepsilon(p - q - \theta) f_\Theta(x)\mathrm{d}x\right] \\
&\quad < (H - L) f_\Theta(\theta)\bar{F}_\varepsilon(p - q - \theta)(\bar{F}_\Theta(\theta))^{-2}\left[\bar{F}_\Theta(\theta) - \int_{\theta}^{\infty} f_\Theta(x)\mathrm{d}x\right] \\
&\quad = 0,
\end{aligned}$$

where $(H - L) = -\partial\theta(\pi, p)/\partial\pi$, the inequality follows from the fact that $\bar{F}_\Theta$ is decreasing in $x \in [\theta, \infty)$, and the final equality from the definition of survival function. Note that

$$\mathsf{P}\left(R_{t+1} = \text{👍}\,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\right) + \mathsf{P}\left(R_{t+1} = \text{👎}\,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\right) = 1,$$

and so

$$\begin{aligned}
&\frac{\partial \mathsf{P}\left(R_{t+1} = \text{👎}\,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\right)}{\partial \pi} \\
&\quad = -\frac{\partial \mathsf{P}\left(R_{t+1} = \text{👍}\,|(B_{t+1}, \pi_t, Q, p_{t+1}) = (1, \pi, q, p)\right)}{\partial \pi} \\
&\quad > 0.
\end{aligned}$$

□

Proof of Proposition 4.4. Without loss of generality we can assume $\Theta_t = 0$ a.s.. Suppose that $\theta(\pi_{t-1}, p) = p - \pi_{t-1}H - (1 - \pi_{t-1})L < 0$. Then consumer t will not purchase the product and $\pi_t^a = \pi_{t-1} = \pi_t^b$. Next consider $\theta(\pi_{t-1}, p) \geq 0$. Note that in this case all consumers buy the product, while $p - L > 0$ and $p - H < 0$. Adapting (A.1) and (A.4) we see that for $c \in \{a, b\}$

$$\pi_t^c = \begin{cases} \pi_{t-1}^c(p, \text{👍}) & \text{w.p. } \pi_{t-1}\bar{F}_{\varepsilon^c}(p-H) + (1-\pi_{t-1})\bar{F}_{\varepsilon^c}(p-L) \\ \pi_{t-1}^c(p, \text{👎}) & \text{w.p. } \pi_{t-1}F_{\varepsilon^c}(p-H) + (1-\pi_{t-1})F_{\varepsilon^c}(p-L), \end{cases}$$

where

$$\pi_{t-1}^c(p, \text{👍}) = \frac{\pi_{t-1}\bar{F}_{\varepsilon^c}(p-H)}{\pi_{t-1}\bar{F}_{\varepsilon^c}(p-H) + (1-\pi_{t-1})\bar{F}_{\varepsilon^c}(p-L)}$$

and

$$\pi_{t-1}^c(p, \text{👎}) = \frac{\pi_{t-1}F_{\varepsilon^c}(p-H)}{\pi_{t-1}F_{\varepsilon^c}(p-H) + (1-\pi_{t-1})F_{\varepsilon^c}(p-L)}.$$

In addition, recall from Proposition 3.3 that $\mathsf{E}[\pi_t^a] = \mathsf{E}[\pi_t^b] = \pi_{t-1}$. Given that the distributions of ε_t^a and ε_t^b are symmetric, if they have a single crossing, this must be at 0. Therefore

$$\begin{aligned} \bar{F}_{\varepsilon^a}(p-H) &\geq \bar{F}_{\varepsilon^b}(p-H), \qquad & \bar{F}_{\varepsilon^a}(p-L) &\leq \bar{F}_{\varepsilon^b}(p-L), \\ F_{\varepsilon^a}(p-H) &\leq F_{\varepsilon^b}(p-H), \qquad & F_{\varepsilon^a}(p-L) &\geq F_{\varepsilon^b}(p-L). \end{aligned}$$

This implies that $\pi_{t-1}^a(p, \text{👍}) \geq \pi_{t-1}^b(p, \text{👍})$ and $\pi_{t-1}^a(p, \text{👎}) \leq \pi_{t-1}^b(p, \text{👎})$, that is, π_t^a can be obtained from π_t^b via a mean-preserving spread. Hence, $\pi_t^b \leq_{\text{cx}} \pi_t^a$. □

Proofs of Section 5

Proof of Lemma 5.2. For ease of notation we drop the arguments of $\theta = \theta(\pi, p^*(\pi)) = p^*(\pi) - H\pi - (1-\pi)L$. The first and second order conditions of $\max_{p \in \mathbb{R}_+} W(\pi, p)$ are

$$\bar{F}_\Theta(\theta) - f_\Theta(\theta) = 0$$

and

$$-2f_\Theta(\theta) - f'_\Theta(\theta) < 0,$$

respectively. Using the envelope theorem we have

$$\frac{\partial W^*(\pi)}{\partial \pi} = p^*(\pi)(H-L)f_\Theta(\theta).$$

By differentiating W^* once more we see that

$$\frac{\partial^2 W^*(\pi)}{\partial \pi^2} = (H-L)\left[\frac{\partial p^*(\pi)}{\partial \pi} f_\Theta(\theta) + p^*(\pi) f'_\Theta(\theta)\left(\frac{\partial p^*(\pi)}{\partial \pi} - (H-L)\right)\right].$$

By the implicit function theorem applied to the first order condition we have

$$\begin{aligned}\frac{\partial p^*(\pi)}{\partial \pi} &= -\frac{(H-L)\left(f_\Theta(\theta) + p^*(\pi) f'_\Theta(\theta)\right)}{-2f_\Theta(\theta) - p^*(\pi) f'_\Theta(\theta)} \\ &= \frac{(H-L)\left(f_\Theta(\theta) + p^*(\pi) f'_\Theta(\theta)\right)}{2f_\Theta(\theta) + p^*(\pi) f'_\Theta(\theta)} \\ &= H - L - \frac{(H-L) f_\Theta(\theta)}{2f_\Theta(\theta) + p^*(\pi) f'_\Theta(\theta)}.\end{aligned}$$

Putting the above together we obtain

$$\begin{aligned}\frac{\partial^2 W^*(\pi)}{\partial \pi^2} &= \frac{(H-L)^2}{2f_\Theta(\theta) + p^*(\pi) f'_\Theta(\theta)} \left[\left(f_\Theta(\theta) + p^*(\pi) f'_\Theta(\theta)\right) f_\Theta(\theta) - p^*(\pi) f'_\Theta(\theta) f_\Theta(\theta)\right] \\ &= \frac{(H-L)^2 (f_\Theta(\theta))^2}{2f_\Theta(\theta) + p^*(\pi) f'_\Theta(\theta)} \\ &> 0,\end{aligned}$$

where the inequality follows from the second order conditions. We conclude that $W^*(\pi)$ is convex.

□

Lemma A.7. *If f_Θ is differentiable and $f_\Theta(\theta(\pi, p)) > 0$, then*

$$\frac{\partial G(\text{👍}, \pi, q, p)}{\partial p} = \int_{\theta(\pi,p)}^{\infty} \bar{F}_\varepsilon(p - q - x) f'_\Theta(x) \, \mathrm{d}x \tag{A.13}$$

and

$$\frac{\partial G(\text{👍}, \pi, q, p)}{\partial p} = \int_{\theta(\pi,p)}^{\infty} F_\varepsilon(p - q - x) f'_\Theta(x) \, \mathrm{d}x. \tag{A.14}$$

Proof. Using Leibniz integral rule we obtain

$$\frac{\partial G(\text{👍}, \pi, q, p)}{\partial p} = -\int_{\theta(\pi,p)}^{\infty} f_\varepsilon(p - q - x) \, \mathrm{d}F_\Theta(x) - \bar{F}_\varepsilon(p - q - \theta(\pi, p)) f_\Theta(\theta(\pi, p)).$$

Integrating by parts the right hand side of (A.13) provides the result for 👍. The result for 👎 is obtained similarly. □

Proof of Theorem 5.3. From Proposition 3.3 we have that $\mathsf{E}[\pi_{t+1}|h_t] = \pi_t$. Therefore for all $t \geq 1$ we have that $\pi(h_t)$ is a dilation of $\pi(h_{t-1})$, hence $\mathsf{E}[\psi(\pi(h_t))] \geq \mathsf{E}[\psi(\pi(h_{t-1}))]$ for all convex functions ψ. A direct implication of Lemma 5.2 is that the the expected revenue extracted from consumer t is increasing in t, i.e., $\mathsf{E}[W^*(\pi(h_{t+1}))] \geq \mathsf{E}[W^*(\pi(h_t))]$. Iteration of the argument proves the result. □

Proof of Proposition 5.5. Bayes' rule gives

$$\pi(h_{t-1}p, r) = \frac{\pi_{t-1}G(r, \pi_{t-1}, H, p)}{\pi_{t-1}G(r, \pi_{t-1}, H, p) + (1 - \pi_{t-1})G(r, \pi_{t-1}, L, p)} = \frac{\pi_{t-1}G(r, \pi_{t-1}, H, p)}{G(r, \pi_{t-1}, p)}.$$

For $p \geq \mathsf{E}_\pi[Q]$ we have $\theta(\pi, p) > 0$. Thus, we can apply Lemma A.7 to obtain

$$\begin{aligned}
\frac{\partial \pi(h_{t-1}, p, \text{👍})}{\partial p} &= \frac{\pi_{t-1}(1-\pi_{t-1})}{(G(\text{👍}, \pi_{t-1}, p))^2}\Bigg[\frac{\partial G(\text{👍}, \pi_{t-1}, H, p)}{\partial p} G(\text{👍}, \pi_{t-1}, L, p) \\
&\quad - \frac{\partial G(\text{👍}, \pi_{t-1}, L, p)}{\partial p} G(\text{👍}, \pi_{t-1}, H, p)\Bigg] \\
&= \frac{\pi_{t-1}(1-\pi_{t-1})}{(G(\text{👍}, \pi_{t-1}, p))^2}\Bigg[\int_{\theta(\pi_{t-1},p)}^{\infty} \bar{F}_\varepsilon(p - H - x) f'_\Theta(x)\, \mathrm{d}x \int_{\theta(\pi_{t-1},p)}^{\infty} \bar{F}_\varepsilon(p - L - x)\, \mathrm{d}F_\Theta(x) \\
&\quad - \int_{\theta(\pi_{t-1},p)}^{\infty} \bar{F}_\varepsilon(p - L - x) f'_\Theta(x)\, \mathrm{d}x \int_{\theta(\pi_{t-1},p)}^{\infty} \bar{F}_\varepsilon(p - H - x)\, \mathrm{d}F_\Theta(x)\Bigg]. \qquad \text{(A.15)}
\end{aligned}$$

Similarly

$$\begin{aligned}
\frac{\partial \pi(h_{t-1}, \text{👎}, p)}{\partial p} &= \frac{\pi_{t-1}(1-\pi_{t-1})}{(G(\text{👎}, \pi_{t-1}, p))^2}\Bigg[\frac{\partial G(\text{👎}, \pi_{t-1}, H, p)}{\partial p} G(\text{👎}, \pi_{t-1}, L, p) \\
&\quad - \frac{\partial G(\text{👎}, \pi_{t-1}, L, p)}{\partial p} G(\text{👎}, \pi_{t-1}, H, p)\Bigg] \\
&= \frac{\pi_{t-1}(1-\pi_{t-1})}{(G(\text{👎}, \pi_{t-1}, p))^2}\Bigg[\int_{\theta(\pi_{t-1},p)}^{\infty} F_\varepsilon(p - H - x) f'_\Theta(x)\, \mathrm{d}x \int_{\theta(\pi_{t-1},p)}^{\infty} F_\varepsilon(p - L - x)\, \mathrm{d}F_\Theta(x) \\
&\quad - \int_{\theta(\pi_{t-1},p)}^{\infty} F_\varepsilon(p - L - x) f'_\Theta(x)\, \mathrm{d}x \int_{\theta(\pi_{t-1},p)}^{\infty} F_\varepsilon(p - H - x)\, \mathrm{d}F_\Theta(x)\Bigg]. \qquad \text{(A.16)}
\end{aligned}$$

If Θ is exponentially distributed, then $f'_\Theta(\cdot) = f_\Theta(\cdot)/\mathsf{E}[\Theta]$. In this case (A.15) and (A.16) equal zero, showing that the support of $\pi(h_{t-1}, p, r_t)$ is identical to the support of $\pi(h_{t-1}, p', r_t)$. Moreover, Lemma A.7 shows that $G(r, \pi_{t-1}, p) < G(r, \pi_{t-1}, p')$ for $r \in \{\text{👍}, \text{👎}\}$. Thus, $\pi(h_{t-1}, p, r_t)$ is obtained from $\pi(h_{t-1}, p', r_t)$ by moving mass from the center of the distribution to the extreme points $\pi(h_{t-1}, p, \text{👍})$ and $\pi(h_{t-1}, p, \text{👎})$ while the mean remains unchanged by Proposition 3.3. Therefore, $\pi(h_{t-1}, p', r_t)$ is a mean preserving spread of $\pi(h_{t-1}, p, r_t)$. □

Lemma A.8. *Under Assumption 5.4 the myopic pricing policy is given by*

$$\phi^{\text{myopic}}(\pi) = \max\left(\frac{1}{\lambda}, \pi H + (1-\pi)L\right) = \begin{cases} \dfrac{1}{\lambda} & \textit{if } \pi \leq \dfrac{1 - \lambda L}{\lambda(H - L)}, \\ \pi H + (1-\pi)L & \textit{if } \pi > \dfrac{1 - \lambda L}{\lambda(H - L)}. \end{cases}$$

Proof. When $\theta(\pi, \phi^{\text{myopic}}(\pi)) > 0$ it follows by the first order conditions that

$$\arg\max_p p\bar{F}_\Theta(p - (\pi H + (1-\pi)L)) = \frac{1}{\lambda}.$$

When $\theta(\pi, \phi^{\text{myopic}}(\pi)) = 0$, then all consumers buy, and it is optimal to charge the highest price under this constraint, $p = \pi H + (1-\pi)L$. Putting these together yields the result. □

The following lemma will be instrumental in comparing the two policies. Given an optimal dynamic pricing policy ϕ^*, we define the following shorthand notation

$$\theta^*(\pi) = \theta(\pi, \phi^*(\pi)). \tag{A.17}$$

Lemma A.9. *If Assumption 5.4 holds, then for every $\pi \in [0,1]$ such that $\phi(\pi) > \pi H + (1-\pi)L$, we have*

$$v^*(\pi) = \frac{1}{\lambda(1-\beta)}\bar{F}(\theta^*(\pi)). \tag{A.18}$$

Proof. It is implicit in the proof that the history is summarized in π. Under Assumption 5.4 we have $f'_\Theta(x) = -\lambda f_\Theta(x) = -\lambda^2 \bar{F}_\Theta(x)$.

Lemma A.7 shows that
$$\frac{\partial G(r,\pi,q,p)}{\partial p} = -\lambda G(r,\pi,q,p),$$

hence

$$\frac{\partial G(r,\pi,p)}{\partial p} = -\lambda G(r,\pi,p). \tag{A.19}$$

Moreover, from (A.15) and (A.16) we have $\partial\pi(r,p)/\partial p = 0$. Therefore, when $\phi^*(\pi) > \mathsf{E}_\pi[Q]$, the first order conditions in (5.2) become

$$\begin{aligned}
&\bar{F}_\Theta(\theta) + pf_\Theta(\theta)\\
&\quad + \beta\left[v^*(\pi(p,\text{👍}))\frac{\partial G(\text{👍},\pi,p)}{\partial p} + v^*(\pi(p,\text{✗}))\frac{\partial G(\text{✗},\pi,p)}{\partial p} + v^*(\pi(p,\text{👎}))\frac{\partial G(\text{👎},\pi,p)}{\partial p}\right]\\
&= \bar{F}_\Theta(\theta) + pf_\Theta(\theta)\\
&\quad + \beta\left[-\lambda v^*(\pi(p,\text{👍}))G(\text{👍},\pi,p) + v^*(\pi(p,\text{✗}))f_\Theta(\theta) - \lambda v^*(\pi(p,\text{👎}))G(\text{👎},\pi,p)\right]\\
&= \bar{F}_\Theta(\theta) + pf_\Theta(\theta) + \beta v^*(\pi)\left(f_\Theta(\theta) + \lambda F_\Theta(\theta)\right)\\
&\quad - \lambda\beta\left[v^*(\pi(p,\text{👍}))G(\text{👍},\pi,p) + v^*(\pi(p,\text{✗}))F_\Theta(\theta) + v^*(\pi(p,\text{👎}))G(\text{👎},\pi,p)\right]\\
&= \bar{F}_\Theta(\theta) + \beta\lambda v^*(\pi) - \lambda\left[p\bar{F}_\Theta(\theta) + \beta\mathsf{E}[v^*(\pi_t)|\pi_{t-1} = \pi, p_t = p]\right] \qquad (\text{A.20})\\
&= 0
\end{aligned}$$

for $\theta = \theta^*(\pi)$ and $p = \phi^*(\pi)$, where we use (A.19) in the first equality. In the second equality we add and subtract $\lambda F_\Theta(\theta)$ and we use Lemma A.1(e), namely, $\pi(\text{✗}, p) = \pi$; in the third equality we use $f_\Theta(\theta) + \lambda F_\Theta(\theta) = \lambda$.

Substituting (A.20) in

$$v^*(\pi) = \phi(\pi)\bar{F}_\Theta(\theta^*(\pi)) + \beta \mathsf{E}[v^*(\pi_t)|\pi_{t-1} = \pi, p_t = \phi(\pi)]$$

we obtain

$$\bar{F}_\Theta(\theta^*(\pi)) + (\beta - 1)\lambda v^*(\pi) = 0$$

that yields (A.18). □

Proof of Proposition 5.6. Define the expected discounted revenue from charging the myopic price

$$v^{\text{myopic}}(\pi_0) = \mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t W(\pi_t, \phi^{\text{myopic}}(\pi_t))\right] = \mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t \phi^{\text{myopic}}(\pi_t)\bar{F}_\Theta(\theta(\pi_t, \phi^{\text{myopic}}(\pi_t)))\right]$$

We find a lower bound for v^{myopic} and then use the optimality of v^* to prove our result.

Define

$$\tilde{\pi} = \frac{\lambda^{-1} - L}{H - L}.$$

Then from Lemma A.8 we have

$$\begin{aligned} v^{\text{myopic}}(\pi_0) &= \mathsf{E}\left[\sum_{t=0}^{\infty} \beta^t \left[\mathbb{1}_{\{\pi_t \le \tilde{\pi}\}} \lambda^{-1} \bar{F}_\Theta\left(\lambda^{-1} - \mathsf{E}_{\pi_t}[Q]\right) + \mathbb{1}_{\{\pi_t > \tilde{\pi}\}} (\mathsf{E}_{\pi_t}[Q]) \bar{F}_\Theta(0)\right]\right] \qquad \text{(A.21)} \\ &= \sum_{t=0}^{\infty} \beta^t \mathsf{E}\left[\max\left(\lambda^{-1}\bar{F}_\Theta\left(\lambda^{-1} - \mathsf{E}_{\pi_t}[Q]\right), \mathsf{E}_{\pi_t}[Q]\right)\right], \end{aligned}$$

where the equalities follow from Lemma A.8, as explained below.

When $\pi < \tilde{\pi}$, price λ^{-1} is myopically optimal, hence $W(\pi, \lambda^{-1}) \ge W(\pi, \mathsf{E}_\pi[Q])$. Similarly, when $\pi > \tilde{\pi}$, price $\mathsf{E}_\pi[Q]$ is optimal, hence $W(\pi, \mathsf{E}_\pi[Q]) \ge W(\pi, \lambda^{-1})$. Therefore

$$\lambda^{-1}\bar{F}_\Theta\left(\lambda^{-1} - \mathsf{E}_\pi[Q]\right) > \mathsf{E}_\pi[Q]$$

when $\pi < \tilde{\pi}$, the reverse inequality holds for $\pi > \tilde{\pi}$ and the two quantities are equal for $\pi = \tilde{\pi}$.

Using the convexity of the maximum and of $\bar{F}_\Theta$ we have

$$\begin{aligned}
v^{\text{myopic}}(\pi_0) &\geq \sum_{t=0}^{\infty} \beta^t \max\left(\mathsf{E}\left[\lambda^{-1}\bar{F}_\Theta\left(\lambda^{-1} - \pi_t H - (1-\pi_t)L\right)\right], \mathsf{E}\left[\pi_t H + (1-\pi_t)L\right]\right) \\
&\geq \sum_{t=0}^{\infty} \beta^t \max\left(\lambda^{-1}\bar{F}_\Theta\left(\lambda^{-1} - \mathsf{E}\left[\pi_t H - (1-\pi_t)L\right]\right), \mathsf{E}\left[\pi_t H + (1-\pi_t)L\right]\right) \\
&= \frac{1}{1-\beta} \max\left(\lambda^{-1}\bar{F}_\Theta\left(\lambda^{-1} - \pi_0 H - (1-\pi_0)L\right), \pi_0 H + (1-\pi_0)L\right),
\end{aligned}$$

where the last equality follows by the tower property of expectations applied to Proposition 3.3.

For posteriors for which $\phi^*(\pi) \leq \mathsf{E}_\pi[Q]$ we have

$$\phi^*(\pi) \leq \mathsf{E}_\pi[Q] \leq \max(\lambda^{-1}, \mathsf{E}_\pi[Q]) = \phi^{\text{myopic}}(\pi).$$

Thus we need to consider posteriors where $\phi^*(\pi) > \mathsf{E}_\pi[Q]$, for which Lemma A.9 holds. When $\pi \leq \tilde{\pi}$ the following inequality holds

$$\begin{aligned}
\frac{1}{\lambda(1-\beta)}\lambda^{-1}\bar{F}_\Theta\left(\phi^*(\pi) - \mathsf{E}_\pi[Q])L\right) &= v^*(\pi) \\
&\geq v^{\text{myopic}}(\pi) \\
&\geq \frac{1}{\lambda(1-\beta)}\lambda^{-1}\bar{F}_\Theta\left(\lambda^{-1} - \mathsf{E}_\pi[Q]\right)
\end{aligned}$$

from which we conclude that

$$\bar{F}_\Theta\left(\phi^*(\pi) - \mathsf{E}_\pi[Q]\right) \geq \bar{F}_\Theta\left(\lambda^{-1} - \mathsf{E}_\pi[Q]\right).$$

Nonincreasingness of $\bar{F}_\Theta$ implies $\phi^*(\pi) \leq \lambda^{-1} = \phi^{\text{myopic}}(\pi)$ for $\pi \leq \tilde{\pi}$. The inequality is strict for $\pi \in (0,1)$, since $\bar{F}_\Theta$ is strictly convex.

When $\pi > \tilde{\pi}$, i.e., when $\lambda(\pi H + (1-\pi)L) > 1$, we have

$$\begin{aligned}
\frac{1}{\lambda(1-\beta)}\lambda^{-1}\bar{F}_\Theta\left(\phi^*(\pi) - \mathsf{E}_\pi[Q]\right) &= v^*(\pi) \\
&\geq v^{\text{myopic}}(\pi) \\
&\geq \frac{1}{1-\beta}\mathsf{E}_\pi[Q].
\end{aligned}$$

This is equivalent to

$$\bar{F}_\Theta\left(\phi^*(\pi) - \mathsf{E}_\pi[Q]\right) \geq \lambda\mathsf{E}_\pi[Q] > 1,$$

which is impossible since $\bar{F}_\Theta$ cannot be larger than 1. Thus $\phi^*(\pi) > \mathsf{E}_\pi[Q]$ is not possible for $\pi > \tilde{\pi}$. Therefore $\phi^*(\pi) \leq \mathsf{E}_\pi[Q] = \phi^{\text{myopic}}$ for $\pi > \tilde{\pi}$. Hence we always have $\phi^*(\pi) \leq \phi^{\text{myopic}}$. □

Proofs of Section 6

Proof of Corollary 6.1. For every $h^+ \in \widehat{\mathscr{H}}(\hat{h}^+)$, swap adjacent pairs in the following way: $(p, \text{👎}, p, \text{👍})$ to $(p, \text{👍}, p, \text{👎})$ to obtain $\overline{h}^+(u,d,p)$. By Proposition 4.1 each iteration weakly increases the belief associated with the history: $\pi(h^+) \leq \pi(\overline{h}^+(u,d,p))$. Similarly, sequentially swap adjacent pairs in the reverse direction to obtain a lower bound $\pi(\underline{h}^+(u,d,p)) \leq \pi(h^+)$.

These bounds hold for all summands in (6.1), since they share the same number of 👍, 👎 and price. Therefore (6.1) provides the result. ☐

Recall the definition of the odds ratio Λ in (A.9) and define

$$\widehat{\Lambda}(\pi,u,d,p) = \left(\Lambda(\text{👍},\pi,p)\right)^u \left(\Lambda(\text{👎},\pi,p)\right)^d. \tag{A.22}$$

Lemma A.10. *Suppose that Assumption 3.1 and the ILR property hold. Let $\hat{h}^+ = (u,d,p)$. Then*

$$\frac{\Gamma(\pi_0)\widehat{\Lambda}(0,u,d,p)}{1+\Gamma(\pi_0)\widehat{\Lambda}(0,u,d,p)} \leq \pi(\hat{h}) \leq \frac{\Gamma(\pi_0)\widehat{\Lambda}(1,u,d,p)}{1+\Gamma(\pi_0)\widehat{\Lambda}(1,u,d,p)}. \tag{A.23}$$

Proof. From (A.8) we get $\pi = \Gamma/(1+\Gamma)$. Therefore by Lemma A.10 we have

$$\pi(\hat{h}^+) \geq \frac{\Gamma(\pi_0)\prod_{k=1}^d \Lambda\left(\text{👎},\pi(0,k-1),p\right)\prod_{k=1}^u \Lambda\left(\text{👍},\pi(k-1,0),p\right)}{1+\Gamma(\pi_0)\prod_{k=1}^d \Lambda\left(\text{👎},\pi(0,k-1),p\right)\prod_{k=1}^u \Lambda\left(\text{👍},\pi(k-1,0),p\right)} \tag{A.24}$$

By the IRL assumption we have that $\Lambda(r,\pi,p)$ is weakly increasing in π. Noting that for $c>0$ the function $cx/(1+cx)$ is increasing in x, when $x>0$, we can repeatedly decrease each of the posteriors in the right hand size of (A.24) to obtain

$$\begin{aligned}
\pi(\hat{h}) &\geq \frac{\Gamma(\pi_0)\prod_{k=1}^d \Lambda\left(\text{👎},0,p\right)\prod_{k=1}^u \Lambda\left(\text{👍},0,p\right)}{1+\Gamma(\pi_0)\prod_{k=1}^d \Lambda\left(\text{👎},0,p\right)\prod_{k=1}^u \Lambda\left(\text{👍},0,p\right)} \\
&= \frac{\Gamma(\pi_0)\left(\Lambda\left(\text{👎},0,p\right)\right)^d\left(\Lambda\left(\text{👍},0,p\right)\right)^u}{1+\Gamma(\pi_0)\left(\Lambda\left(\text{👎},0,p\right)\right)^d\left(\Lambda\left(\text{👍},0,p\right)\right)^u} \\
&= \frac{\Gamma(\pi_0)\widehat{\Lambda}(0,u,d,p)}{1+\Gamma(\pi_0)\widehat{\Lambda}(0,u,d,p)}.
\end{aligned}$$

A similar argument can be used to obtain the upper bound. ☐

Consider the inequalities

$$\mathsf{P}(R_t = \text{👍} | \pi(h_t) = \pi, Q = H) \log \left(\frac{\mathsf{P}(R_t = \text{👍} | \pi(h_t) = 0, Q = H)}{\mathsf{P}(R_t = \text{👍} | \pi(h_t) = 0, Q = L)} \right)$$
$$+ \mathsf{P}(R_t = \text{👎} | \pi(h_t) = \pi, Q = H) \log \left(\frac{\mathsf{P}(R_t = \text{👎} | \pi(h_t) = 0, Q = H)}{\mathsf{P}(R_t = \text{👎} | \pi(h_t) = 0, Q = L)} \right) > 0 \tag{A.25}$$

and

$$\mathsf{P}(R_t = \text{👍} | \pi(h_t) = \pi, Q = H) \log \left(\frac{\mathsf{P}(R_t = \text{👍} | \pi(h_t) = 1, Q = H)}{\mathsf{P}(R_t = \text{👍} | \pi(h_t) = 1, Q = L)} \right)$$
$$+ \mathsf{P}(R_t = \text{👎} | \pi(h_t) = \pi, Q = H) \log \left(\frac{\mathsf{P}(R_t = \text{👎} | \pi(h_t) = 1, Q = H)}{\mathsf{P}(R_t = \text{👎} | \pi(h_t) = 1, Q = L)} \right) < 0. \tag{A.26}$$

Lemma A.11. *There exists a price $\tilde{p} > L$ such that for all $\pi \in [0,1]$ and $p \leq \tilde{p}$ the following inequalities hold:*

$$G(\text{👍}, \pi, H, p) \log \big(\Lambda(\text{👍}, 0, p)\big) + G(\text{👎}, \pi, H, p) \log \big(\Lambda(\text{👎}, 0, p)\big) > 0, \tag{A.27}$$
$$G(\text{👍}, \pi, L, p) \log \big(\Lambda(\text{👍}, 1, p)\big) + G(\text{👎}, \pi, L, p) \log \big(\Lambda(\text{👎}, 1, p)\big) < 0. \tag{A.28}$$

Proof. Note for $p \leq L$ we have $\theta(\pi, p) < 0$ for all $\pi \in [0,1]$. Consequently all consumers buy and

$$G(r, \pi, q, p) = G(r, 0, q, p) = G(r, 1, q, p)$$

for all $\pi \in [0,1]$, $r \in \{\text{👎}, \text{👍}, \text{✗}\}$, $q \in \{H, L\}$. Therefore

$$\begin{aligned} &G(\text{👍}, \pi, H, p) \log \big(\Lambda(\text{👍}, 0, p)\big) + G(\text{👎}, \pi, H, p) \log \big(\Lambda(\text{👎}, 0, p)\big) \\ &\quad = \gamma \left(G(\text{👎}, 0, H, p), G(\text{👎}, 0, L, p), \bar{F}_\Theta(\theta(0, p)) \right) \\ &\quad > 0 \end{aligned}$$

and

$$\begin{aligned} &G(\text{👍}, \pi, L, p) \log \big(\Lambda(\text{👍}, 1, p)\big) + G(\text{👎}, \pi, L, p) \log \big(\Lambda(\text{👎}, 1, p)\big) \\ &\quad = -\gamma \left(G(\text{👍}, 1, L, p), G(\text{👍}, 1, H, p), \bar{F}_\Theta(\theta(1, p)) \right) \\ &\quad < 0, \end{aligned}$$

for all $\pi \in [0,1]$, where γ is as in (A.7) and the inequalities follow from Lemma A.5.

Since all the above functions are continuous in p, we have that there exists a price $\tilde{p} > L$ such that these inequalities hold for all $p \leq \tilde{p}$. ∎

Lemma A.12. *Let* $\{Z_t\}_{t\in\mathbb{N}_+}$ *be a sequence of binary random variables such that for some* $\underline{\rho} > 0$

$$\mathsf{P}(Z_t = 1|Z_1,\ldots,Z_{t-1}) = 1 - \mathsf{P}(Z_t = 0|Z_1,\ldots,Z_{t-1}) \geq \underline{\rho} \quad a.s..$$

Then for all $\rho < \underline{\rho}$, *there exists* $\underline{\alpha} > 0$ *such that*

$$\mathsf{P}\left(\sum_{i=1}^{t} Z_i < \rho t\right) \leq \mathrm{e}^{-\underline{\alpha} t} \quad \textit{for all } t \geq 1.$$

Analogously if for some $\overline{\rho} < 1$

$$\mathsf{P}(Z_t = 1|Z_1,\ldots,Z_{t-1}) \leq \overline{\rho} \quad a.s.,$$

then for all $\rho > \overline{\rho}$, *there exists* $\overline{\alpha} > 0$ *such that*

$$\mathsf{P}\left(\sum_{i=1}^{t} Z_i > \rho t\right) \leq \mathrm{e}^{-\overline{\alpha} t} \quad \textit{for all } t \geq 1.$$

Proof. Let $\{\underline{Z}_t\}_{t\in\mathbb{N}_+}$ and $\{\overline{Z}_t\}_{t\in\mathbb{N}_+}$ be two sequences of Bernoulli random variables with parameters $\underline{\rho}$ and $\overline{\rho}$, respectively. Then, using a result by Veinott (1965) (see Shaked and Shanthikumar, 2007, Theorem 6.B.3), we have that

$$(\underline{Z}_1,\ldots,\underline{Z}_t) \leq_{\mathrm{st}} (Z_1,\ldots,Z_t) \leq_{\mathrm{st}} (\overline{Z}_1,\ldots,\overline{Z}_t),$$

for all $t \geq 1$.

Therefore using the Chernoff-Hoeffding bound (Hoeffding, 1963) we obtain that for every $\rho \leq \underline{\rho}$

$$\mathsf{P}\left(\sum_{i=1}^{t} Z_i < \rho t\right) \leq \mathsf{P}\left(\sum_{i=1}^{t} \underline{Z}_i < \rho t\right) \leq \mathrm{e}^{-\underline{\alpha} t}$$

for some $\underline{\alpha} > 0$. Similarly for every $\rho \geq \overline{\rho}$

$$\mathsf{P}\left(\sum_{i=1}^{t} Z_i > \rho t\right) \leq \mathsf{P}\left(\sum_{i=1}^{t} \overline{Z}_i < \rho t\right) \leq \mathrm{e}^{-\overline{\alpha} t}$$

For some $\overline{\alpha} > 0$. □

Proof of Proposition 6.2. We use Corollary 6.1 and show that, conditionally on $Q = H$, the odds $\Gamma(\pi(\hat{h}_t^+))$ diverge almost surely to $+\infty$ and, conditionally on $Q = L$, the odds they converge to 0 almost surely. In particular we show that $\widehat{\Lambda}(0, U_t, D_t, p) \to \infty$ and $\widehat{\Lambda}(1, U_t, D_t, p) \to 0$. This, together with the bounds in (A.23), implies that $\pi(\hat{h}_t^+) \to \mathbb{1}_{\{Q=H\}}$ a.s.

We begin by studying the asymptotic behavior of the lower bound. Call

$$X_t := \mathbb{1}_{\{R_t=\text{👍}\}} \log\left(\Lambda\left(\text{👍}, 0, p\right)\right) + \mathbb{1}_{\{R_t=\text{👎}\}} \log\left(\Lambda\left(\text{👎}, 0, p\right)\right).$$

Call

$$\bar{B}_t := \{B_i = 1, i \in \{1, \dots, t\}\}.$$

Then

$$\log\left(\widehat{\Lambda}(0, U_t, D_t, p)\right) = \sum_{i\in\bar{B}_t} X_i.$$

Thus, we can write

$$\mathsf{P}\left(\log\left(\widehat{\Lambda}(0, U_t, D_t, p)\big|Q = H\right) < M\right) =$$
$$\sum_{k=1}^{t} \mathsf{P}\left(\sum_{i\in\bar{B}_t} X_i < M\Big|\operatorname{card}(\bar{B}_t) = k, Q = H\right) \mathsf{P}\left(\operatorname{card}(\bar{B}_t) = k\big|Q = H\right). \quad \text{(A.29)}$$

For some $\alpha \in (0, 1)$ we can bound this probability as follows

$$\begin{aligned}
\mathsf{P}\left(\log\left(\widehat{\Lambda}(0, U_t, D_t, p)\right) < M\big|Q = H\right) \le &\sum_{k=0}^{\lfloor\alpha t\rfloor-1} \mathsf{P}\left(\operatorname{card}(\bar{B}_t) = k\big|Q = H\right) \\
&+ \sum_{k=\lfloor\alpha t\rfloor}^{t} \mathsf{P}\left(\sum_{i\in\bar{B}_t} X_i < M\Big|\operatorname{card}(\bar{B}_t) = k, Q = H\right) \quad \text{(A.30)} \\
= &\,\mathsf{P}\left(\sum_{i=1}^{t} \mathbb{1}_{\{B_i=1\}} < \lfloor\alpha t\rfloor\big|Q = H\right) \\
&+ \sum_{k=\lfloor\alpha t\rfloor}^{t} \mathsf{P}\left(\sum_{i\in\bar{B}_t} \widetilde{X}_i < \widetilde{M}_k\Big|\operatorname{card}(\bar{B}_t) = k, Q = H\right),
\end{aligned}$$

where

$$\widetilde{X}_i := \frac{X_i - \log\left(\Lambda\left(\text{👎}, 0, p\right)\right)}{\log\left(\Lambda\left(\text{👍}, 0, p\right)\right) - \log\left(\Lambda\left(\text{👎}, 0, p\right)\right)},$$

$$\widetilde{M}_k := \frac{M_k - k\log\left(\Lambda\left(\text{👎}, 0, p\right)\right)}{\log\left(\Lambda\left(\text{👍}, 0, p\right)\right) - \log\left(\Lambda\left(\text{👎}, 0, p\right)\right)}.$$

Notice that $\widetilde{X}_i = \mathbb{1}_{\{R_i=\text{👍}\}}$.

From Assumption 3.1(a) we have

$$0 < \min_{\pi\in[0,1]} \bar{F}_\Theta(\theta(\pi,p)) \leq \mathsf{P}(B_t = 1 | B_1 = 1, \ldots, B_{t-1} = 1) \leq \max_{\pi\in[0,1]} \bar{F}_\Theta(\theta(\pi,p)) \leq 1$$

and

$$0 < \min_{\pi\in[0,1]} \frac{G(\text{👍}, \pi, H, p)}{\bar{F}_\Theta(\theta(\pi,p))} \leq \mathsf{P}(\widetilde{X}_t = 1 | \widetilde{X}_1, \ldots, \widetilde{X}_{t-1}) \leq \max_{\pi\in[0,1]} \frac{G(\text{👍}, \pi, H, p)}{\bar{F}_\Theta(\theta(\pi,p))} < 1$$

almost surely, for all $t \geq 1$. The above maxima and minima exist since all the functions of π are continuous and defined on $[0,1]$.

Take $\alpha < \min_{\pi\in[0,1]} \bar{F}_\Theta(\theta(\pi,p)) = \bar{F}_\Theta(\theta(0,p))$ and apply Lemma A.12 to conclude that

$$\mathsf{P}\left(\sum_{i=1}^{t} \mathbb{1}_{\{B_i=1\}} < \lfloor \alpha t \rfloor \big| Q = H\right) \leq \mathrm{e}^{-\underline{\alpha} t}$$

for some $\underline{\alpha} > 0$.

In order to apply Lemma A.12 to

$$\mathsf{P}\left(\sum_{i\in\bar{B}_t} \widetilde{X}_i < \widetilde{M}_k \big| \operatorname{card}(\bar{B}_t) = k, Q = H\right)$$

for $k \geq \lfloor \alpha t \rfloor$, we must verify that

$$\widetilde{M}_k \leq k\underline{\rho} \tag{A.31}$$

with the shorthand notation

$$\underline{\rho} := \min_{\pi\in[0,1]} \frac{G(\text{👍}, \pi, H, p)}{\bar{F}_\Theta(\theta(\pi,p))}.$$

Noting that the denominator in the definition of $\widetilde{M}_k$ is positive, the inequality (A.31) becomes

$$\frac{M}{k} \leq \frac{M}{\lfloor \alpha t \rfloor} < \log\left(\Lambda(\text{👍}, 0, p)\right) \underline{\rho} + \log\left(\Lambda(\text{👎}, 0, p)\right)(1 - \underline{\rho}).$$

The right hand side is positive by our assumption on the price using Lemma A.11 and thus for any M the inequality is satisfied for t large enough.

Thus, we have concluded that for any $M > 0$ there exists t_M such that for all $t > t_M$

$$\mathsf{P}\left(\log\left(\widehat{\Lambda}(0, U_t, D_t, p)\right) < M \big| Q = H\right) \leq K\,\mathrm{e}^{-\bar{\alpha} t}$$

for some positive constants $K, \bar{\alpha}$. As a consequence, for all $M > 0$

$$\sum_{t=1}^{\infty} \mathsf{P}\left(\log\left(\widehat{\Lambda}(0, U_t, D_t, p)\right) < M \big| Q = H\right) \leq \infty.$$

Applying the first Borel-Cantelli Lemma, we conclude that $\log\left(\widehat{\Lambda}(0, U_t, D_t, p)\right) \to \infty$ a.s. when $Q = H$ and $p \leq \tilde{p}$, where $\tilde{p}$ satisfied the conditions in Lemma A.11.

This implies

$$\frac{\widehat{\Lambda}(0, U_t, D_t, p)}{1 + \widehat{\Lambda}(0, U_t, D_t, p)} \to 0$$

and therefore, by (A.23), that $\pi(\hat{h}_t) \to 1$ a.s., conditionally on $Q = H$.

The case $Q = L$ can be proved along the same line. □

References

Acemoglu, D., Bimpikis, K., and Ozdaglar, A. (2013) Dynamics of information exchange in endogenous social networks. *Theor. Econom.* Forthcoming.

Acemoglu, D., Dahleh, M. A., Lobel, I., and Ozdaglar, A. (2011) Bayesian learning in social networks. *Rev. Econom. Stud.* **78**, 1201–1236.

Banerjee, A. and Fudenberg, D. (2004) Word-of-mouth learning. *Games Econom. Behav.* **46**, 1–22.

Banerjee, A. V. (1992) A simple model of herd behavior. *Quart. J. Econom.* **107**, 797–817.

Bergemann, D. and Välimäki, J. (1997) Market diffusion with two-sided learning. *RAND J. Econ.* **28**, pp. 773–795.

Bertsekas, D. P. and Shreve, S. E. (2007) *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992) A theory of fads, fashion, custom, and cultural change in informational cascades. *J. Polit. Econ.* **100**, 992–1026.

Bose, S., Orosel, G., Ottaviani, M., and Vesterlund, L. (2006) Dynamic monopoly pricing and herding. *RAND J. Econ.* **37**, 910–928.

Candogan, O., Bimpikis, K., and Ozdaglar, A. (2012) Optimal pricing in networks with externalities. *Oper. Res.* **60**, 883–905.

Goeree, J. K., Palfrey, T. R., and Rogers, B. W. (2006) Social learning with private and common values. *Econom. Theory* **28**, 245–264.

Herrera, H. and Hörner, J. (2013) Biased social learning. *Games and Economic Behavior* **80**, 131–146.

Hoeffding, W. (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30.

Hu, N., Pavlou, P. A., and Zhang, J. (2006) Can online reviews reveal a product's true quality?: empirical findings and analytical modeling of online word-of-mouth communication. In *Proceedings of the 7th ACM conference on Electronic commerce*, EC '06, 324–330. ACM, New York, NY, USA. URL http://doi.acm.org/10.1145/1134707.1134743.

Ifrach, B., Maglaras, C., and Scarsini, M. (2012) Monopoly pricing in the presence of social learning. Technical Report 1957924, SSRN eLibrary. URL http://dx.doi.org/10.2139/ssrn.1957924.

Karlin, S. (1968) *Total Positivity. Vol. I.* Stanford University Press, Stanford, Calif.

Karlin, S. and Rinott, Y. (1980) Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10**, 467–498.

Karlin, S. and Taylor, H. M. (1975) *A First Course in Stochastic Processes.* Academic Press, second edition.

Lariviere, M. A. (2006) A note on probability distributions with increasing generalized failure rates. *Oper. Res.* **54**, 602–604.

Lariviere, M. A. and Porteus, E. L. (2001) Selling to the newsvendor: an analysis of price-only contracts. *Manuf. Serv. Oper. Manag.* **3**, 293–305.

Li, X. and Hitt, L. M. (2008) Self-selection and information role of online product reviews. *Inform. Syst. Res.* **19**, 456–474.

Luca, M. (2011) Reviews, reputation, and revenue: the case of Yelp.com. Technical Report 12-016, Harvard Business School.

Osborne, M. J. and Rubinstein, A. (1994) *A Course in Game Theory.* MIT Press, Cambridge, MA.

Shaked, M. and Shanthikumar, J. G. (2007) *Stochastic Orders.* Springer, New York. URL http://dx.doi.org/10.1007/978-0-387-34675-5.

Smith, L. and Sørensen, P. (2000) Pathological outcomes of observational learning. *Econometrica* **68**, 371–398.

Smith, L. and Sørensen, P. (2008) Rational social learning with random sampling. Mimeo.

Talwar, A., Jurca, R., and Faltings, B. (2007) Understanding user behavior in online feedback reporting. In *Proceedings of the 8th ACM conference on Electronic commerce*, EC '07, 134–142. ACM, New York, NY, USA. URL http://doi.acm.org/10.1145/1250910.1250931.

Veinott, Jr., A. F. (1965) Optimal policy in a dynamic, single product, non-stationary inventory model with several demand classes. *Operations Res.* **13**, 761–778.