Monopoly Pricing in the Presence of Social Learning

Davide Crapis* Bar Ifrach† Costis Maglaras‡ Marco Scarsini§

Revision 2: December 6, 2015

Abstract

A monopolist offers a product to a market of consumers with heterogeneous quality preferences. Although initially uninformed about the product quality, they learn by observing past purchase decisions and reviews of other consumers. Our goal is to analyze the social learning mechanism and its effect on the seller's pricing decision. Consumers follow an intuitive, non-Bayesian decision rule. Under conditions that we identify, we show that consumers eventually learn the product's quality. We show how the learning trajectory can be approximated in settings with high demand intensity via a mean-field approximation that highlights the dynamics of this learning process, its dependence on the price, and the market heterogeneity with respect to quality preferences. Two pricing policies are studied: a static price, and one with a single price change. Finally, numerical experiments suggest that pricing policies that account for social learning may increase revenues considerably relative to policies that do not.

Keywords: social learning, information aggregation, bounded rationality, optimal pricing.

JEL Classification: D49, D83.

1 Introduction

Launching a new product involves uncertainty. Specifically, consumers may not initially know the true quality of the new product, but learn about it through some form of a social learning process, adjusting their estimates of its quality along the way, and making possible purchase decisions accordingly. The dynamics of this social learning process affect the market potential and realized sales trajectory over time. The seller's pricing policy can tactically accelerate or decelerate learning, which, in turn, affects sales at different points in time and the product's lifetime profitability.

*Columbia Business School, Columbia Univeristy (`dcrapis16@gsb.columbia.edu`).

†Airbnb, Inc. (`Barifrach@gmail.com`). This author thanks the NET Institute (www.NETinst.org) for financial support.

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

§Dipartimento di Economia e Finanza, LUISS (`marco.scarsini@luiss.it`). The work of this author is partially supported by PRIN 20103S5RN3 and IDG31300110. The author is a member of GNAMPA-INdAM.

This paper studies a monopolist's pricing decision in a market where quality estimates are evolving according to such a learning process.

Consumers arrive at the market according to a Poisson process and face the decision of either purchasing a product with unknown quality, or choosing an outside option. They differ in their base valuation for the observable attributes of the product, which, together with the product quality, determines their willingness-to-pay. These base valuation parameters are assumed to be independently and identically drawn from a known distribution. If consumers knew the true product quality, then the distribution of the base valuations would map directly into a willingness-to-pay (WtP) distribution and, in turn, into a demand function that the monopolist could use as a basis of her pricing decision.

In our model the quality is unknown, and consumers' prevailing estimate of the unknown quality evolves according to a social learning mechanism. Consumers who purchase the product experience its true quality plus some small quality disturbance, which is independent and identically distributed across purchasers. Purchasers report whether they "liked" or "disliked" the product, i.e., if their ex-post utility was positive or negative, respectively. Consumers do not report their base valuations, so a positive review may result from a high quality or high idiosyncratic quality preference. An arriving consumer observes the history of purchase decisions and reviews made prior to his arrival, combines this information with his prior quality estimate, infers the associated product quality, and makes his own purchase decision. The sequence of purchase decisions affects the evolution of the observable information set, and, as such, the dynamics of the market response over time. Optimizing the monopolist's pricing policy requires detailed understanding of the learning dynamics and not just its asymptotic properties.

It is typical to assume that fully rational agents (consumers) update their beliefs for the unknown quality of the product through a Bayesian analysis that takes into account the sequence of decisions and reviews, and accounts for the fact that each such decision was based on different information available at that time. This sequential update procedure introduces a formidable analytical and computational onus on each agent that may be hard to justify as a model of actual choice behavior. Instead, we postulate a non-Bayesian and fairly intuitive learning mechanism, where consumers *assume* that all prior decisions were based on the same information, and under this bounded rationality assumption, consumers pick the maximum likelihood estimate (MLE) of the quality level that would best explain the observed sequence of positive and negative reviews (non-purchase decisions are not observable). New reviews change the available information and the resulting MLE over time, and, of course, the rate at which consumers choose to purchase and later on submit new reviews about their experiences.

As a motivating example consider the launch of a new hotel. It is typically hard to evaluate the quality of such premises without first hand experience or word-of-mouth, which explains the importance that online review sites such as Tripadvisor have had on the hospitality industry.[1] Assume the hotel is sufficiently differentiated from its competitors to be considered a monopoly in some category; e.g., it may be the only hotel with a private beach in the area. Suppose it offers better services than what consumers think at first. Initially some consumers' idiosyncratic tastes would convince them to choose this hotel; perhaps they have strong preferences for having a private beach. These consumers would recommend the hotel by posting a review, which, in turn, increases future demand, as potential consumers learn that the hotel is better than previously thought.[2] The price charged by the hotel affects this learning process by controlling the number of guests who review the hotel and their degree of satisfaction. By accounting for the learning process the hotelier may be able to avoid a sluggish start and realize the establishment's full potential demand faster.

This paper strives to contribute in three ways. First, in terms of modeling, by specifying a social learning environment that tries to capture aspects of online reviews as well as the possible bounded rationality of consumers. Second, by proposing a tractable methodological framework, based on mean-field approximations, to study the learning dynamics and related price optimization questions in the presence of social learning. This approach is flexible and applicable in other related settings where the microstructure of the learning process and nature of information are different. Third, in addressing some of the pricing questions faced by revenue maximizing sellers in such settings.

Regarding the learning mechanism, the information reported by consumers is subject to a self-selection bias, since only consumers with a high enough base valuations purchase the product. The intuitive MLE procedure takes into account this crucial point, and, as we show in Section 3.2, the resulting quality estimate converges to the true product quality almost surely.

Detailed understanding of the learning trajectory is essential in optimizing the tradeoff between learning and the monopolist's discounted revenue objective. Second, Section 3.3 derives a mean-field (fluid model) asymptotic approximation for the learning dynamics motivated by settings where the rate of arrival of new consumers to the system grows large. Proposition 3 shows that the asymptotic learning trajectory is characterized by a system of differential equations. Restricting attention to uniformly distributed base valuations across consumers and focusing on the case where the markets prior quality estimate is below the true quality, Section 3.4 derives the closed form transient of the fraction of likes and dislikes over time, as well as that of the associated quality estimate. The transient dynamics imply that the instantaneous demand function evolves over time according to an ODE, which itself depends on the seller's price, i.e., it emerges endogenously through the interplay between consumer behavior and the seller's decisions. The solution of the mean-field model gives

[1] According to TripAdvisor 90% of hotel managers think that review websites are very important to their business and 81% monitor their reviews at least weekly.

[2] Many empirical papers found that positive consumer reviews increase sales. For example, Luca (2011) finds that a one star increase in the average consumer review on a popular review site (on a five star scale) translates to a 5-9% increase in sales for restaurants in Seattle, WA.

a crisp characterization of the dependence of the learning trajectory on the price, and specifically show that the time-to-learn decreases if the monopolist lowers her price. This result naturally exploits the suitability of mean field approximations to characterize transient behavior of discrete and stochastic systems. The paper illustrates that method in the context of the specific consumer learning model described above, however, the approach is fairly general and can be used to describe the transient learning dynamics under a broader set of micro consumer behavioral models, see Ifrach (2012, Sections 2.2 and 3.2).

Third, we study the seller's pricing problem under the assumption that the seller knows the true product quality, but that the consumers do not use the seller's price as a signal of quality. Section 4 studies the monopolist's problem of choosing the static price that optimizes her infinite horizon discounted revenues. Proposition 5 characterizes the optimal solution, which exists and is unique, and lies in the interval of two natural price points: (a) the optimal price assuming that consumers do not learn and always make purchase decisions based on their prior quality estimate; and (b) the optimal price in a setting where consumers knew the true quality all along. The learning transient and its speed in relation to the seller's discount factor determines the optimal price. Intuitively, if the learning transient is slow relative to the discounting of revenues, then she prices almost as if all consumers made purchasing decisions based on their prior on the quality; and, if learning is fast, then the seller's price will approach the one that the monopolist would set if all consumers knew the true product quality.

Lastly, Section 5.1 studies a model where the seller has some degree of dynamic pricing capability, namely she can change her price once, at a time of her choosing. In this case the monopolist may sacrifice short term revenues in order to influence the social learning process in the desired direction and capitalize on that after changing the price. Proposition 6 shows that when consumers initially underestimate the true quality, the first period price is lower than the second period one. This policy accelerates learning and increases revenues considerably. The numerical experiments of Section 5.2 suggest that a pricing policy with two prices performs very well, and that the benefit of implementing more elaborate pricing policies may be small.

We conclude this section with a brief literature review. The *social learning* literature is fairly broad. Much of this work can be classified into two groups depending on the learning mechanism, which is either Bayesian or non-Bayesian. Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992) are standard references in economics on observational learning where each agent observes a signal and the decisions of the agents who made a decision before him, but not their consequent satisfaction (in fact preferences are homogeneous). Agents are rational and update their beliefs in a Bayesian way. They show that at some point all agents will ignore their own signals and base their decisions only on the observed behavior of the previous agents, which will prevent further learning and may lead to herding on the bad decision.

For social learning to be successful, an agent must be able to reverse the herd behavior of his predecessors. Smith and Sørensen (2000) show that this is the case if agents' signals have unbounded strength. Goeree, Palfrey, and Rogers (2006) show that this is achieved with enough heterogeneity in consumers' preferences. Our Assumption 1, which is key in proving learning, is similar in nature to that of Goeree et al. (2006).[3] Social learning has been studied in great generality by Arieli and Mueller-Frank (2014).

Several papers have considered variations of the observational learning model with imperfect information. Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) and Acemoglu, Bimpikis, and Ozdaglar (2014) greatly contribute to the understanding of the interplay between social learning and the structure of the social network. Acemoglu et al. (2011) identify conditions on the network under which social learning is successful and, alternatively, herding may prevail. Acemoglu et al. (2014) consider agents who can delay their decision in order to obtain information from others by utilizing their social network. Jadbabaie, Molavi, Sandroni, and Tahbaz-Salehi (2012) consider a model where consumers communicate over a social network and update their information in a non-Bayesian way. They provide conditions for learning to occur in this setting. Herrera and Hörner (2013) consider a case where agents can observe only one of two decisions of their predecessors, which in the language of our model means that the number of no purchase decisions is not observed. Instead, consumers know the time of their arrival, which is associated with the number of predecessors who chose the unobservable option. They show that this relaxation does not change the asymptotic learning result of Smith and Sørensen (2000).

There is a growing literature in economics that studies non-Bayesian learning mechanisms that employ simpler and perhaps more plausible learning protocols. Ellison and Fudenberg (1993, 1995) consider settings in which consumers exchange information about their experienced utility and use simple decision rules to choose between actions. The nature of word-of-mouth in our paper is similar, although we consider reviews and not utilities directly.

A few papers in the operations management literature have considered social learning. In Debo and Veeraraghavan (2009) consumers observe private signals about the unknown value of the service and decide whether or not to join a queue, where congestion conveys information about the value of the service. Debo, Parlour, and Rajan (2012) study a server who chooses her service rate to signal quality, again in a queueing context. Related applications in inventory systems and retailing explore how stock outs or observed inventory positions may also signal product quality. The mean field approach of this paper may be applicable in studying transient learning phenomena in these operational settings.

Some recent papers have considered models of social learning in the presence of consumer

[3]See the surveys by Bikhchandani, Hirshleifer, and Welch (1998), and, more recently, by Acemoglu and Ozdaglar (2011) for many extensions to this model.

reviews. Ifrach, Maglaras, and Scarsini (2015) study a Bayesian model where both the quality of the product and the reviews can assume only two possible values and they provide conditions for learning. Besbes and Scarsini (2015) deal with a model where customers only observe the sample mean of past reviews, and show under which conditions customers can recover the true quality of the product based on the feedback they observe. They use stochastic approximation techniques to obtain their results. In our model the decision rule is not fully rational, yet consumers do account for the self-selection bias in their predecessors review, unlike other models that studied consumer reviews (e.g., Li and Hitt (2008)). Lafky (2014) experimentally deals with the fundamental issue of why people rate products and which biases arise in the behavior of reviewers.

The area of *revenue management* focuses, in part, on tactical problems of price optimization.[4] It is typical therein to capture consumer response through some form of a demand function, and to strive to optimize the seller's pricing policy—static or dynamic—so as to maximize her profitability. An important strand of literature in this area, which will not be reviewed here, considers an exogenous unobservable demand function—stationary or time-varying—and designs pricing policies under which the seller jointly learns the demand and optimizes revenues. In contrast to this literature, we study how a demand process is formed when a new product is introduced, and where consumer opinions evolve dynamically based on a social learning process. The resulting non-stationary demand process is endogenous to the pricing policy, and their interplay is characterized to optimize revenues.

Mean-field approximations have been used extensively in revenue management; perhaps the first reference in that area is Gallego and van Ryzin (1994). More broadly, the use of mean-field approximations that rely on an appropriate application of the functional strong law of large numbers to study the transient behavior of stochastic processes has a fairly broad literature that we will not review here. The particular result we will employ, due to Kurtz (1977/78), was originally derived for studying the asymptotic behavior of Markov Chain models with process-dependent transition parameters, used to analyze diffusion and epidemic systems.

The learning dynamics in our model give rise to a sales trajectory which, when properly interpreted, resembles the famous Bass diffusion model, see Bass (2004).[5] Contrary to the Bass model that specifies up front a differential equation governing social dynamics, we start with a micro model of agents' behavior and characterize its limit as the number of agents grows large. This limit —given by a differential equation as well—induces a macro level model of social dynamics. The application of mean-field approximation to our model bridges the literature on social learning and that on social dynamics by filling the gap between the detailed micro level model of agent behavior, and the subsequent macro level model of aggregate dynamics.

[4] Talluri and van Ryzin (2005) provides a good overview of that work.

[5] In particular, by considering a population with finite mass, and by simplifying consumers' decisions.

Several papers have studied pricing when agents are engaged in social learning or embedded in a social network. Bose, Orosel, Ottaviani, and Vesterlund (2006) consider pricing in the classic Bayesian observational learning model when a monopolist and agents are equally uninformed about the value of the good. Campbell (2013) studies the role of pricing in the launching of a new product in a model of social interaction that builds on percolation theory, where the latter focuses on dynamic pricing. Candogan, Bimpikis, and Ozdaglar (2012) consider optimal pricing strategies of a monopolist selling a product to consumers who are embedded in a social network and experience externalities in consumption. Strategic behavior of firms and consumers in the presence of social learning has been studied by Papanastasiou, Bakshi, and Savva (2014); Papanastasiou and Savva (2014), where in particular study a two period problem and study the effect of the firm's pricing policy on consumer purchase decisions as well as the impact of early adaptor reviews on downstream demand.

Also related is the literature on pricing of experience goods, whose quality can be determined only upon consumption; see, e.g., Bergemann and Välimäki (1997) and Vettas (1997). Most of these papers consider consumers that are homogenous ex-ante, i.e., before consuming the good. Bergemann and Välimäki (1997) consider a duopoly and heterogeneous consumers on a line who report their experienced utility, and show that the expected price path for the new product is increasing when consumers initially underestimate the quality; our Proposition 6 is consistent with their findings.

2 Model

2.1 The Monopolist's Pricing Problem

A sequence of consumers, indexed by $i = 1, 2, \ldots$, sequentially decide whether to purchase a newly launched good or service (henceforth, the product), or choose an outside alternative. The intrinsic quality of the product, denoted with q, is initially unknown and can take values in the interval $[q_{\min}, q_{\max}]$ with $q_{\min} \geq 0$. The quality experienced by consumer i, if he chooses to buy the product, is subject to a random disturbance ε_i and given by $q_i := q + \varepsilon_i$. This quality shock reflects variability in service levels (e.g., waiting times), production defects or exogenous factors influencing the way the product is consumed (e.g., weather).

Consumers are heterogeneous; this is represented by a parameter α_i that determines consumer i's base valuation, e.g., that would correspond to the observable attributes of the product. His utility from consuming the product is

$$u_i = \alpha_i + q_i - p,$$

where p is the price charged by the monopolist, which, for the time being, we assume to be fixed.[6] The utility derived from choosing the outside alternative is normalized to zero for all consumers.

Preference parameters, $\{\alpha_i\}_{i=1}^{\infty}$, are i.i.d. random variables drawn from a known distribution function F. We denote the corresponding survival function by $\bar{F}(\cdot) := 1 - F(\cdot)$, and assume that F has a differentiable density f, which is uniformly bounded by some constant $f_{\max}$ and has connected support $[\alpha_{\min}, \alpha_{\max}]$, or $[0, \infty)$. The α_i can be interpreted as an idiosyncratic premium that consumer i is willing to pay for the product. The failure rate of the quality preference distribution is defined as $h(x) := f(x)/\bar{F}(x)$. Throughout this paper we will assume that F has an increasing failure rate (IFR), that is, h is strictly increasing for all $x \geq 0$.

The quality disturbances are short-lived; they are i.i.d. random variables with mean zero and independent of the underlying quality, as well as of the preference parameters. To simplify the analysis, we assume that ε_i follows a symmetric, two-point distribution: specifically, ε_i takes the values $\{-\bar{\varepsilon}, \bar{\varepsilon}\}$ with equal probabilities 0.5. It is natural to think that the quality disturbances ε_i are small relative to the magnitude of the unknown quality q. Moreover, both q and the ε_i's are expressed in the units of the consumer's utility, e.g., in dollars.

Heterogeneity in terms of the α_i's implies that even if the product quality, q, was known, not all consumers would make the same decision: only those with $\alpha_i \geq \alpha^* := p - q$ would purchase the product, assuming that they are risk neutral with respect to quality disturbances. Equivalently, only consumers with WtP $\alpha + q \geq p$ would purchase; the distribution of α gives rise to a WtP distribution $\alpha + q$ for the product.

The product is launched at time $t = 0$, and consumers arrive thereafter according to a Poisson process with rate Λ, independent of the product's quality and consumers' preference parameters. Denote by t_i the random time consumer i enters the market and makes his purchasing decision. Consumer i does not re-enter the market regardless of his decision at t_i; this assumption is reasonable if the time horizon under consideration is not too long.

Consumers initially have some common prior conjecture on the quality of the product, $q_0 \in [q_{\min}, q_{\max}]$. This prior conjecture could be the expected value of some prior distribution of the quality, or could simply be consumers' best guess given the product's marketing campaign and previous encounters with the monopolist in other categories.

The information transmission in our model is often called *word-of-mouth* communication. A consumer i who purchased the product, truthfully reports a review about his experience with the product, denoted by r_i that takes two values: 'like', denoted by r^{L} and 'dislike', denoted by r^{D}. A consumer who purchases the product reports that he likes it if his ex-post utility was nonnegative,

[6] The functional form of the utility function does not play a big role in the subsequent analysis. For example, another tractable alternative would be a vertically differentiated market in which utility takes the form $u_i = \alpha_i \cdot q_i - p$.

taking into account the unknown quality and quality disturbance, as well as his preference parameter; he reports that he dislikes it if his ex-post utility was negative. Consumers report neither their preference parameter nor the quality disturbance they faced and, as such, reviews are not fully informative. For example, a 'like' could result from a high preference parameter, the product being of high quality or a positive quality shock (not necessarily all).[7] This binary report is a simplification of the star rating scales of online review systems. Consumers who did not purchase the product do not report a review and are not observed. We will denote their decision by $r_i = r^{\mathrm{O}}$.

We make the following assumptions on the set of feasible prices.

Assumption 1. The price p charged by the monopolist belongs to $[p_{\min}, p_{\max}]$, where

(i) $p_{\max}$ is such that $\bar{F}(p_{\max} - q_{\min} + \bar{\varepsilon}) > 0$. (Equivalently $p_{\max} < \alpha_{\max} + q_{\min} - \bar{\varepsilon}$.)

(ii) $p_{\min}$ is such that $\bar{F}(p_{\min} - q_{\max} + \bar{\varepsilon}) < 1$. (Equivalently $p_{\min} > \alpha_{\min} + q_{\max} - \bar{\varepsilon}$.)

Assumption 1(i) implies that, even at the lowest possible quality level, there will always be some consumers who choose to buy the product (this follows from $p_{\max} < \alpha_{\max} + q_{\min}$), and moreover, at least some of these consumers will like the product—the latter ensures that new information about q will enter the learning process; if this assumption is violated and $p_{\max} > \alpha_{\max} + q_{\min} - \bar{\varepsilon}$, then at $q = q_{\min}$ all buyers with a negative shock would dislike and all buyers with a positive shock would like. Assumption 1(i) is similar to the "unbounded belief assumption" often used in Bayesian social learning in the sense that it implies that some new information will enter the system over time, which will ultimately allow the market to learn the unknown product quality. Assumption 1(ii), states that there are always some low-WtP consumers who will dislike the product if they get a negative disturbance realization. It is easy to verify that both conditions are satisfied if the support of α's is sufficiently wide relative to the unknown quality $[q_{\min}, q_{\max}]$ and the magnitude of the subsequent quality disturbances $\bar{\varepsilon}$ is small.

We define the following quantities: $L_i := \sum_{j=1}^{i-1} \mathbf{1}\left\{r_j = r^{\mathrm{L}}\right\}$ is the number of consumers who purchase and like the product out of the first $i-1$ consumers, and, similarly, D_i is the number of consumers who purchase and dislike the product. The information available to consumer i before making his decision is

$$I_i = (L_i, D_i). \tag{1}$$

The index 'i' itself is not observable. Before describing the evolution of information and consumers' decision rule, we introduce the monopolist's pricing problem, which is the main focus of this paper. The monopolist seeks to choose a static price p to maximize her discounted expected revenue, $R(p)$,

[7] This assumption is motivated by the fairly anonymous reviews that one may get online today. One possible extension would consider a model where consumers gather two sets of information, one from a process like the one above, and the other from a smaller set of their "friends" whose quality preferences are known with higher accuracy.

as follows,

$$\max_p R(p) = \max_p \mathsf{E}\left[\sum_{i=1}^{\infty} e^{-\delta t_i}\ p\mathbf{1}\left\{r_i(p) \neq r^{\mathrm{O}}\right\}\right] = \max_p \sum_{i=1}^{\infty} \mathsf{E}\left[e^{-\delta t_i}\ p\mathsf{P}\left(r_i(p) \neq r^{\mathrm{O}} \mid I_i\right)\right], \quad (2)$$

where $\delta > 0$ is the monopolist's discount factor, and the expectation is with respect to consumers' arrival times, the idiosyncratic quality preferences α_i's, and the sequence of quality disturbances ε_i's. The monopolist is assumed to know the true quality, the prior quality estimate, as well as the distribution of quality preferences and disturbances. Expression (2) reveals the complexity of the pricing problem in the presence of social learning. Consumers' purchasing decisions influence future revenues through the information available to successors. As such, the dynamic of the social learning process must be understood in order to solve for the optimal price. Section 5 considers a problem where the seller can select two prices as well as the optimal time to switch between them.

2.2 Decision Rule

We introduce a plausible non-Bayesian decision rule that consumers are assumed to employ to decide whether to purchase the product. It is composed of two parts: consumer i (a) uses his available information to form a quality estimate $\hat{q}_i$, and (b) purchases the product if and only if his estimated utility is non-negative $\alpha_i + \hat{q}_i - p \geq 0$.

In broad terms, consumers try to answer the following question: given the observed number of likes and dislikes, the distribution of idiosyncratic quality preferences, and the distribution of the quality shocks that affect the experienced quality, what value of intrinsic quality best explains the observed data assuming that *all* past purchasers made decisions based on the same quality estimate? The crucial simplification is that consumers disregard the fact that reviews have been submitted sequentially, and that the information available to the respective purchasers was itself evolving over time. Review information is typically aggregated in the form we postulate, but review aggregator sites often allow users to expand the information set and view the sequence and timestamps of the various reviews. Accessing this information is, however, cumbersome, and using this detailed information is computationally hard (perhaps implausible). Our simplifying behavioral assumption is a form of bounded rationality on the consumers' regard. Disregarding the sequence of reviews and processing the aggregated number of likes and dislikes, consumers are assumed to invoke a maximum likelihood estimation (MLE) procedure to compute their quality estimate.

Under the assumption that a consumer was using the correct value for q, the probability of a 'like' conditional on a purchase is

$$
\begin{aligned}
&\mathsf{P}(\text{consumer } j \text{ likes} \mid \text{consumer } j \text{ buys}, \hat{q}_j = q) \\
&\quad = \frac{\mathsf{P}(\alpha_j + q_j - p \geq 0, \alpha_j + q - p \geq 0)}{\mathsf{P}(\alpha_j + q - p \geq 0)} \\
&\quad = \frac{.5\mathsf{P}(\alpha_j + q + \bar{\varepsilon} - p \geq 0, \alpha_j + q - p \geq 0)}{\mathsf{P}(\alpha_j + q - p \geq 0)} + \frac{.5\mathsf{P}(\alpha_j + q - \bar{\varepsilon} - p \geq 0, \alpha_j + q - p \geq 0)}{\mathsf{P}(\alpha_j + q - p \geq 0)} \\
&\quad = .5 + .5\frac{\bar{F}(p - q + \bar{\varepsilon})}{\bar{F}(p - q)} \\
&\quad = .5 + .5G(p - q)
\end{aligned}
$$

where the second equality follows from Bayes' rule and

$$G(x) := \bar{F}(x + \bar{\varepsilon})/\bar{F}(x). \tag{3}$$

Similarly, the probability of observing a dislike conditional on a purchase is

$$
\begin{aligned}
&\mathsf{P}(\text{consumer } j \text{ dislikes} \mid \text{consumer } j \text{ buys}, \hat{q}_j = q) \\
&\quad = 1 - \mathsf{P}(\text{consumer } j \text{ likes} \mid \text{consumer } j \text{ buys}, \hat{q} = q) = .5 - .5G(p - q).
\end{aligned}
$$

The likelihood of observing (L_i, D_i) likes and dislikes under the assumption that all consumers were acting under the same quality estimate is

$$\mathscr{L}_i(q) = (.5 + .5G(p - q))^{L_i}(.5 - .5G(p - q))^{D_i}.$$

Next, consumers introduce the effect of their prior quality estimate into their learning mechanism. In order to do this, the prior quality estimator q_0 must be transformed into a number of fictitious reviews, L_0 and D_0, that are consistent with q_0 under our maximum likelihood learning mechanism. We assume that the total weight assigned to the prior estimator will be the one that is equivalent to the expected number of positive and negative reviews over a length of time of w time units. One can also think of $1/w$ as the standard error of the prior quality estimate q_0, i.e., the longer the accumulation period of prior information the more certain the consumers are about their prior.

With that in mind, we define L_0 and D_0 to be the expected number of like and dislike fictitious reviews under the assumption that the quality prior q_0 is equal to the true product quality as follows:

$$L_0 = w\Lambda\mathsf{P}(\text{customer } i \text{ buys \& likes} \mid \hat{q}_i = q = q_0) = .5w\Lambda\left[\bar{F}(p - q_0) + \bar{F}(p - q_0 + \bar{\varepsilon})\right], \tag{4}$$

and

$$D_0 = w\Lambda \mathsf{P}(\text{customer } i \text{ buys \& dislikes} \mid \hat{q}_i = q = q_0) = .5w\Lambda \left[\bar{F}(p - q_0) - \bar{F}(p - q_0 + \bar{\varepsilon})\right]. \quad (5)$$

Incorporating the effect of the prior quality estimate, consumers will pick the quality estimate $\hat{q}_i$ in the interval $[q_{\min}, q_{\max}]$ so as to maximize the *weighted* likelihood function defined by

$$\mathscr{L}_i^w(q) = (.5 + .5G(p - q))^{L_0 + L_i}(.5 - .5G(p - q))^{D_0 + D_i}. \quad (6)$$

It is useful to spell out the probability that consumer i will like, dislike or not purchase the product when his quality estimate, $\hat{q}_i$, is different from q. Consumer i reports a positive review if he buys the product ($\alpha_i + \hat{q}_i - p \geq 0$) and has a positive experience ($\alpha_i + q_i - p \geq 0$), where in the later we have to account for the disturbance ε_i.

$$\begin{aligned}\mathsf{P}(r_i = r^{\mathrm{L}}) &= \mathsf{P}(\alpha_i + \hat{q}_i - p \geq 0, \alpha_i + q_i - p \geq 0) \\ &= \mathsf{P}(\alpha_1 \geq p - \min(q_i, \hat{q}_i)) \\ &= .5\bar{F}(p - \min(q - \bar{\varepsilon}, \hat{q}_i)) + .5\bar{F}(p - \min(q + \bar{\varepsilon}, \hat{q}_i)).\end{aligned}$$

Similarly, the probability of a dislike is

$$\begin{aligned}\mathsf{P}(r_i = r^{\mathrm{D}}) &= \mathsf{P}(\alpha_i + \hat{q}_i - p \geq 0, \alpha_i + q_i - p < 0) \\ &= \bar{F}(p - \hat{q}_i) - .5\bar{F}(p - \min(q - \bar{\varepsilon}, \hat{q}_i)) - .5\bar{F}(p - \min(q + \bar{\varepsilon}, \hat{q}_i))\end{aligned}$$

and the probability of no purchase is given by

$$\Pr(r_i = r^{\mathrm{O}}) = \mathsf{P}(\alpha_i + \hat{q}_i - p < 0) = F(p - \hat{q}_i).$$

We finish this section with few brief comments.

Price as a signal. The seller's price conveys information about the product quality, but we assume that consumers do not adjust their quality estimate in response to that information; likewise the monopolist does not need to take that consideration into account.

Prior weight. All consumers assign the same weight to their common prior quality estimate. The weight w assigned to the prior is a measure of confidence in the prior estimate in the absence of other new information. This is analogous to the precision of prior beliefs in a Bayesian setting. As time goes by and more reviews accumulate in the system, consumers place increasingly more confidence in the review information versus their prior information, which is reflected into the fact that the effect of L_0 and D_0 on the weighted likelihood (6) becomes negligible as L_i and D_i grow.

No purchases. The MLE procedure described above can also be used to study social learning in a model where consumers are informed on the number of previous consumers who decided *not to purchase*. In that context, it is not hard to show that the quality estimator has similar properties to the ones of the estimator that we then characterize. This is not surprising, since statistical estimates can only improve with more information, but it is important because it shows the robustness of our estimation procedure.

Consumer learning. Different information models and micro models of consumer behavior could be considered. For example, consumers may only observe reviews from a random sample of their predecessors, which grows large in an appropriate sense; or, consumers may weigh their predecessors' reviews such that later reviews are more influential than earlier ones. The latter could also be done by the review site that acts as an information aggregator; see Ifrach (2012, Sections 2.2 and 3.2).

3 Asymptotic learning and the associated learning transient

In this section we will establish that consumers eventually learn the true quality of the product, and subsequently approximate the learning transient via the solution of an ordinary differential equation derived as a mean-field limit in a large market.

3.1 The consumer's MLE problem

Our first result characterizes the maximum-likelihood (MLE) quality estimate used by consumer i. First, we define

$$l_i := \frac{L_0 + L_i}{B_i}, \quad d_i := \frac{D_0 + D_i}{B_i} \quad \text{and} \quad B_i := L_0 + L_i + D_0 + D_i, \tag{7}$$

where l_i denotes the prevailing fraction of likes and d_i denotes the prevailing fraction of dislikes for consumer i, then we can state the following proposition.

Proposition 1. *The MLE quality estimator*

$$\hat{q}_i = \operatorname{argmax} \left\{ \mathscr{L}_i^w(q) \ : \ q_{\min} \leq q \leq q_{\max} \right\}, \tag{8}$$

is unique and given by

$$\hat{q}_i = \begin{cases} \text{proj}_{[q_{\min}, q_{\max}]}(q^*) & \text{if } l_i > d_i, \\ q_{\min} & \text{if } l_i \le d_i, \end{cases} \tag{9}$$

where q^* *solves the following equation*

$$G(p - q^*) = l_i - d_i = 2l_i - 1. \tag{10}$$

(All proofs can be found in the Appendix.) The maximum likelihood estimate $\hat{q}_i$ has appealing properties. Equation (10) shows that it depends on the data only through the fraction of like reviews. Moreover, by Lemma 1 (in the Appendix) we know that the function $G(p-q)$ is increasing in q, which implies that the estimator is increasing in the fraction of like reviews, as one would expect. Lemma 1 also implies that $G(p - q)$ is invertible for every $q \in [q_{\min}, q_{\max}]$, which means that (10) defines a one-to-one mapping between l_i and $\hat{q}_i$. We will exploit this observation in the subsequent analysis.

3.2 Asymptotic learning

The stochastic learning process converges in the sense that the quality estimate $\hat{q}_i$ converges to the true quality q almost surely. Assumption 1 implies that new information continues to enter the system since some consumers will always choose to purchase and as a result review the product. This, in turn, ultimately guarantees that learning is achieved.

Proposition 2. *Consider the sequential learning process described in the previous section, where consumer* i *estimates the prevailing quality* $\hat{q}_i$ *through* (9)*. Then,* $\hat{q}_i \to q$ *as* $i \to \infty$ *almost surely.*

The above result serves as a sanity check that learning is achieved under the proposed decision rule; this result addresses the question that underlies most of the literature on social learning of whether agents eventually learn the true state of the world. The ultimate goal of this paper is to study the pricing question described in the previous section for which one needs to have a more explicit characterization of the learning transient for the underlying stochastic learning process. This is intractable, however, as is in almost all social learning models in the literature, both Bayesian and non-Bayesian.[8] Our approach is to approximate the learning transient through a set of intuitive and tractable ordinary differential equations.

[8] Typical results establish that learning is achieved (or not) as $i \to \infty$, and in some cases the rate of convergence; see, e.g., Acemoglu, Dahleh, Lobel, and Ozdaglar (2009) for a characterization of the rate of convergence of Bayesian social learning for some social networks.

3.3 Approximation of learning dynamics in a large market

The proposed approximation is relevant in large market settings, and will be justified through an asymptotic argument as the arrival rate of consumers making purchase decisions grows large, rescaling processes so that the time scale within which information gets released and learning evolves is the one of interest. The mean-field or fluid model approximation yields a tractable characterization of the learning dynamics and provides insight on their dependence on the micro model of consumer learning behavior and other problem primitives, including the seller's price. We comment at the end of this section on the generality of this approach.

We consider a sequence of systems indexed by n. In the n-th system consumers' arrival process is Poisson with rate $\Lambda^n := n\Lambda$. The state variables of the n-th system at time t is given by $X^n(t) := \big(L^n(t), D^n(t)\big)$, where $L^n(t)$ is the number of consumers who report like by time t in the n-th system, and $D^n(t)$ is defined analogously. The superscript n indicates the dependence on the arrival rate. Denote the *scaled* state variable $\bar{X}^n(t) := X^n(t)/n$ and similarly for $\bar{L}^n(t)$ and $\bar{D}^n(t)$. This state variable comprises the information available to the first consumer arriving after time t. We will keep the prior initialization weight w constant, so the number of reviews associated with the prior quality estimate will stay proportional to the fictitious number of purchasers that would flow through the system over a fixed time window, e.g., a week.[9]

We carry the notation from the previous section with the necessary adjustments. Specifically, with some abuse of notation, in the n-th system we have from (7) that

$$l^n := \frac{L_0^n + L^n}{B^n} = \frac{\bar{L}_0^n + \bar{L}^n}{\bar{B}^n} =: \bar{l}^n,$$

where $B^n := L_0^n + L^n + D_0^n + D^n$ and $\bar{B}^n := B^n/n$. Similarly, $d^n := (D_0 + D^n)/B^n = (\bar{D}_0 + \bar{D}^n)/\bar{B}^n =: \bar{d}^n$. The fractions of likes and dislikes are independent of n, conditional on $\bar{X}^n$. Similarly, $\hat{q}^n(X^n(t))$ is directly defined through (9) and (10). We also note that $\hat{q}^n(\bar{X}^n(t)) = \hat{q}^n(X^n(t))$ through the normalized definitions of $\bar{l}^n$ and $\bar{d}^n$, and, moreover, that the mapping $\hat{q}^n$ itself does not depend on n, that is, the same quality estimation procedure is applied throughout the scaling that we consider, and simply evaluated at the appropriate state $X^n(t)$.

Building on the above we define the functions γ^{L} and γ^{D} such that

$$\gamma^{\mathrm{L}}(\bar{X}^n) := \mathsf{P}\left(r_i = r^{\mathrm{L}} \mid I_i = X^n\right) = .5\left[\bar{F}\big(p - \min(q - \bar{\varepsilon}, \hat{q}(\bar{X}^n))\big) + \bar{F}\big(p - \min(q + \bar{\varepsilon}, \hat{q}(\bar{X}^n))\big)\right],$$

with the interpretation that $\gamma^{\mathrm{L}}(\bar{X}^n)$ is the probability that a consumer who observes information

[9] It is possible to scale w^n differently, of course, in which case we would need to apply the corresponding time change in the $\bar{X}^n(t)$ process. The above assumption simplifies the transient analysis without affecting, however, the resulting structure and insights.

X^n reports r^{L}. Similarly,

$$\gamma^{\mathrm{D}}(\bar{X}^n) := \bar{F}\big(p - \hat{q}(\bar{X}^n)\big) - \gamma^{\mathrm{L}}(\bar{X}^n).$$

Note that the above expressions imply that γ^{L} and γ^{D} are independent of n.

With this notation in mind, we use a Poisson thinning argument to express the scaled state variables as a Poisson processes with time dependent rates. Let $N := (N^{\mathrm{L}}, N^{\mathrm{D}})$ be a vector of independent Poisson processes with rate 1. Then,

$$\bar{L}^n(t) = \frac{1}{n} N^{\mathrm{L}} \left(\Lambda^n \int_0^t \gamma^{\mathrm{L}}(\bar{X}^n(s)) \, \mathrm{d}s \right),$$

and similarly for $\bar{D}^n$. The following shorthand notation is convenient,

$$\bar{X}^n(t) = \frac{1}{n} N \left(\Lambda^n \int_0^t \gamma(\bar{X}^n(s)) \, \mathrm{d}s \right), \tag{11}$$

where $\gamma := (\gamma^{\mathrm{L}}, \gamma^{\mathrm{D}})$. The dependence of the state-dependent rate functions γ^{L} and γ^{D} on the state $\bar{X}^n(t)$ enters through the quality estimate $\hat{q}(\bar{X}^n(t))$.

If the rate processes inside the expressions (11) did not depend on the state $\bar{X}^n(t)$ itself, then a straightforward application of the functional strong law of large numbers for the Poisson process would yield a deterministic limit for $\bar{X}^n(t)$ as n grew large. Our model is a bit more complex, but because the evolution of the state $\bar{X}^n(t)$ depends on the decisions made by all predecessors, one would expect it to vary slowly relative to the increasing number of consumers arriving at any given point in time. Intuitively, considering a short time interval $[t, t+\Delta]$, one would expect that the large pool of heterogeneous consumers arriving in that interval, each with a different quality parameter α, and making decisions based on similar information given by $\bar{X}^n(s)$ for some $s \in [t, t+\Delta]$, would lead to a deterministic but state-dependent evolution of $\bar{X}^n$ for large n; effectively, the stochastic nature of the decisions due to consumer heterogeneity is "averaged out" in such a setting.

This argument is made precise in Proposition 3 that derives a deterministic limiting characterization for the system behavior as n grows large using Kurtz (1977/78, Theorem 2.2) through a sample path analysis based on a strong approximation argument and a subsequent application of Gronwall's inequality.

Proposition 3. *For every $t > 0$,*

$$\lim_{n\to\infty} \sup_{s \le t} |\bar{X}^n(s) - \bar{X}(s)| = 0 \quad a.s.,$$

where $\bar{X}(t) = (\bar{L}(t), \bar{D}(t))$ *is deterministic and satisfies the integral equation,*

$$\bar{X}(t) = \Lambda \int_0^t \gamma(\bar{X}(s)) \, \mathrm{d}s. \tag{12}$$

To better understand (12) consider the expression for the scaled number of likes,

$$\bar{L}(t) = \Lambda \int_0^t \gamma^{\mathrm{L}}(\bar{X}(s)) \, \mathrm{d}s = \Lambda \int_0^t \mathsf{P}\big(r_s = r^{\mathrm{L}} \mid I_s = \bar{X}(s)\big) \, \mathrm{d}s.$$

This means that the scaled number of 'likes' at t is the sum over the mass of consumers who report a 'like' in each $s \leq t$, and this mass depends on past reviews via $\bar{X}(\cdot)$. It follows that the scaled number of consumers that arrive by time t is Λt and that the number of people that purchased the product and submitted a report is $\bar{B}(t) := \bar{L}_0 + \bar{D}_0 + \bar{L}(t) + \bar{D}(t)$. It is convenient to derive from (12) the expressions for $(\bar{l}(t), \bar{d}(t))$ in the limiting (fluid) model, since these quantities determine the decision of an arriving consumer:

$$\bar{l}(t) := l(\bar{X}(t)) = \frac{\bar{L}_0 + \bar{L}(t)}{\bar{B}(t)} \quad \text{and} \quad \bar{d}(t) := d(\bar{X}(t)) = \frac{\bar{D}_0 + \bar{D}(t)}{\bar{B}(t)} = 1 - \bar{l}(t). \tag{13}$$

From the definition of $(\gamma^{\mathrm{L}}, \gamma^{\mathrm{D}})$, (10) and (13), it follows that $(\bar{L}, \bar{D})$ is absolutely continuous and therefore differentiable almost everywhere. We refer to time t where $(\bar{L}, \bar{D})$ is differentiable as *regular*. At regular points t, $(\bar{L}, \bar{D})$ satisfies the differential equations:

$$\dot{\bar{L}}(t) = .5\Lambda \left[\bar{F}(p - \min(q - \bar{\varepsilon}, \hat{q}_t)) + \bar{F}(p - \min(q + \bar{\varepsilon}, \hat{q}_t))\right], \tag{14}$$

and

$$\begin{aligned}\dot{\bar{D}}(t) &= \Lambda \Big[\bar{F}(p - \hat{q}_t) - .5\left[\bar{F}(p - \min(q - \bar{\varepsilon}, \hat{q}_t)) + \bar{F}(p - \min(q + \bar{\varepsilon}, \hat{q}_t))\right]\Big] \\ &= \Lambda \bar{F}(p - \hat{q}_t) - \dot{\bar{L}}(t),\end{aligned} \tag{15}$$

where $\hat{q}_t$ is the maximum-likelihood estimator defined in (10) and evaluated at $(\bar{l}(t), \bar{d}(t))$.

Finally, as mentioned in the introduction, the approach of employing a mean field approximation to characterize the transient of the social learning process can be used to study additional micro learning models in other settings of interest. One key characteristic that underlies this approach is that each individual consumer has a diminishing influence on the others, and as such on the aggregate behavior, as the size of the population scales. This condition typically holds when agents decisions depend on system aggregates. This is related to the literature on the diffusion of products, innovation, and epidemics, often called social dynamics, that focuses on the evolution of system aggregates, such as the fraction of adopters. The approach described above allows one to determine how the structure of the micro model of consumer behavior affects the aggregate learning dynamics.

3.4 Transient learning dynamics: uniformly distributed valuations

In the remainder of the paper we will assume that $\alpha \sim U[0, \bar{\alpha}]$[10] and, without loss of generality, we will normalize $q_{\min} = 0$. This allows us to simplify the ODEs (14) and (15) as follows. At regular times t, we have

$$\dot{\bar{L}}(t) = \Lambda \left(\frac{\bar{\alpha} + .5 \min(q - \bar{\varepsilon}, \hat{q}_t) + .5 \min(q + \bar{\varepsilon}, \hat{q}_t) - p}{\bar{\alpha}} \right) \quad \text{and} \quad \dot{\bar{D}}(t) = \Lambda \left(\frac{\bar{\alpha} + \hat{q}_t - p}{\bar{\alpha}} \right) - \dot{\bar{L}}(t), \tag{16}$$

and the quality estimator can now be written as

$$\hat{q}_t \; = \; p - \bar{\alpha} + \frac{\bar{\varepsilon}}{2(1 - \bar{l}(t))} \; = \; p - \bar{\alpha} + \frac{\bar{\varepsilon}}{2} \left(1 + \frac{\bar{L}_0 + \bar{L}(t)}{\bar{D}_0 + \bar{D}(t)} \right). \tag{17}$$

The subsequent analysis of the paper will primarily focus on the learning transient when the prior estimate q_0 initially underestimates the true quality of the product, i.e., $q_0 < q$, and, moreover, focus on the portion of the learning transient over which $\hat{q}_t < q - \bar{\varepsilon}$ (we refer to this as "phase 1"). When $\bar{\varepsilon}$ is small, this first phase of the learning process is the most important to understand. [11]

Underestimating prior ($q_0 < q$); phase 1 of learning $\hat{q}_t < q - \bar{\varepsilon}$. At times where the prevailing quality estimate is such that $\hat{q}_t < q - \bar{\varepsilon}$, the consumers who purchase with $\bar{\alpha} + \hat{q}_t - p \geq 0$ are guaranteed to have a positive ex-post utility realization since $\bar{\alpha} + q - \bar{\varepsilon} - p \geq \bar{\alpha} + \hat{q}_t - p \geq 0$. As a result only like reviews will be submitted as long as $\hat{q}_t < q - \bar{\varepsilon}$. In this case, the ODEs (16) can be solved in closed form and the solutions can then be used together with (17) to characterize the learning trajectory for $\hat{q}_t$. We do this in Proposition 4, which is the main result of this section. Before presenting the result, we formally define the *time-to-learn*

$$\tau := \inf\{t : t \geq 0, |q - \hat{q}_t| \leq \bar{\varepsilon}\}.$$

This is the time it takes $\hat{q}_t$ to reach within $\bar{\varepsilon}$ of q and it measures the duration of the learning phase.

Proposition 4. *Consider the ODEs for the learning dynamics given in (16) and assume that $q_0 < q$. Then, for $t \leq \tau$,*

$$\hat{q}_t = p - \bar{\alpha} + (\bar{\alpha} + q_0 - p) \exp\left(\frac{t}{w} \right). \tag{18}$$

[10] It follows that $\bar{F}(x) = 1 - x/\bar{\alpha}$ for all $x \in [0, \bar{\alpha}]$, and that $G(x) = (\bar{\alpha} - x - \bar{\varepsilon})/(\bar{\alpha} - x)$ for all $x \in [0, \bar{\alpha} - \bar{\varepsilon}]$.

[11] The Appendix studies the case where the prior overestimates the true quality $q_0 > q$, and also shows how to approximate the evolution of the learning ODEs at times where $q - \bar{\varepsilon} < \hat{q}_t < q + \bar{\varepsilon}$ for the case where $\bar{\varepsilon}$ is small. Moreover, note that equation (17) corresponds to the solution of equation (10) when $\bar{l}(t) \leq 1 - \bar{\varepsilon}/2\bar{\alpha}$, otherwise we have a different solution. We only consider the above solution since in the relevant cases $\bar{l}(t)$ is never too close to 1.

Moreover,

$$\tau = w \log\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right). \tag{19}$$

Proposition 4 characterizes the learning transient in the underestimating case. In particular, expression (18) describes the learning trajectory of the quality estimate $\hat{q}_t$ for all $t \le \tau$, and expression (19) characterizes the time-to-learn as a function of the relevant model parameters, i.e., the market heterogeneity $\bar{\alpha}$, the distance $q - q_0$ of prior quality from true quality, and the price.

First, expression (18) shows that $\hat{q}_t$ starts at the prior estimate q_0 at time 0 and it increases monotonically to reach $q - \bar{\varepsilon}$ at time τ. Moreover, the lower the prior weight w, the faster the learning trajectory, i.e., when consumers place less weight on their prior estimate, they are more sensitive to the review information and as a result the quality estimator is updated faster.

Furthermore, note that expression (19) can be rewritten as

$$w \log\left(1 + \frac{q - q_0 - \bar{\varepsilon}}{\bar{\alpha} + q_0 - p}\right),$$

which allows us to make some interesting comparative statics observations on the learning transient. In particular, note that the time-to-learn τ is decreasing in the maximum (or equivalently the range of the) base valuation $\bar{\alpha}$, because if consumers have higher valuations, more consumers choose to buy (and review) the product and thus information accumulates faster. Moreover, τ is increasing in $(q - q_0)$, i.e., the more severely consumers underestimate quality, the longer it takes to learn q. Finally, note that the time-to-learn is scaled by the prior weight w, which is a natural time scale if we think about learning as the process of accumulating enough information to overcome the bias of the prior estimate q_0.

Before moving to the revenue maximization problem, we make one last important observation that relates the learning process to the monopolist's pricing decision.

Corollary 1. *The time-to-learn τ is increasing in p.*

This result can be explained as follows. Let

$$d_t(p) := \bar{F}(p - \hat{q}_t) = \frac{\bar{\alpha} + \hat{q}_t - p}{\bar{\alpha}},$$

which denotes the *instantaneous demand function* at time t in our large market setting. At any given $t \le \tau$, a price increase affects the instantaneous demand function through two channels: first, a direct channel, a higher price means a lower instantaneous demand at time t; second, through $\hat{q}_t$, a higher price means a lower $\hat{q}_t$ at time t.[12] Therefore, by increasing the price, the monopolist

[12]This can easily be verified by differentiating (18) with respect to p and noting that, for all $0 < t \le \tau$, $\partial\hat{q}_t/\partial p = 1 - \exp(t/w) < 0$.

effectively decreases the rate at which consumers are buying (and reviewing) the product, thus slowing down learning. Finally, using the characterization of $\hat{q}_t$ from (18) we can rewrite $d_t(p)$, for all $t \leq \tau$, as

$$d_t(p) = \left(\frac{\bar{\alpha} + q_0 - p}{\bar{\alpha}} \right) \exp\left(\frac{t}{w} \right).$$

The instantaneous demand function thus takes the linear form $d_t(p) = a_t - b_t \cdot p$ with $a_t = (1 + q_0/\bar{\alpha}) \exp(t/w)$ and $b_t = (1/\bar{\alpha}) \exp(t/w)$. Note that the instantaneous demand is positive for all $p \in [p_{\min}, p_{\max}]$, which follows directly from Assumption 1, and that both the slope b_t and the intercept a_t are increasing in t. Thus, when consumers initially underestimate quality ($q_0 < q$), the instantaneous demand and consequently the instantaneous revenue are increasing with t. Similarly, note that as time passes and the quality estimate increases, the monopolist can achieve the same instantaneous demand with a higher price, thus generating more revenue. In the following sections, we will further elaborate on these insights and we will study the pricing strategy of the monopolist.

The overestimating case $q_0 > q$ and the analysis of the ODEs after time τ in the case of a small quality disturbance $\bar{\varepsilon}$ are briefly reviewed in the Appendix. In both cases the transient is more complicated and its solution cannot be written in closed form, however, numerical solutions are very simple to attain and one can still establish useful structural properties, such as the fact that the quality estimate monotonically converges to the true value q from below (above) in the underestimating (overestimating) case.

4 Static Price Analysis

In this section we solve the monopolist's problem of choosing a static price to maximize her revenue as given in (2). Following the analysis of the previous section, the stochastic learning trajectory is replaced by its deterministic mean field approximation. This enables us to solve an otherwise intractable problem. The next two sections focus on the price optimization problem, for the case in which consumers initially underestimate quality through their prior, i.e., $q_0 < q$. Adapting by the mean-field approximation, we can write the seller's discounted revenue as

$$\bar{R}(p) = \Lambda \int_0^\infty e^{-\delta t} \pi_t(p) \text{ dt} = \Lambda \left(\int_0^\tau e^{-\delta t} \pi_t(p) \text{ dt} + \int_\tau^\infty e^{-\delta t} \pi_t(p) \text{ dt} \right)$$

where $\pi_t(p) = p \cdot d_t(p)$ denotes the *instantaneous revenue function* at time t, and τ is the time-to-learn that was defined in the previous section. We assume that once the learning process has converged to $\hat{q}_\tau = q - \bar{\varepsilon}$, revenues are accrued from then on according to $\hat{q}_t = q$ for $t \geq \tau$. This is safe in our setting since $\bar{\varepsilon}$ is a small quantity, and it leads to the following revenue function

$$\tilde{R}(p) = \Lambda \left(\int_0^\tau e^{-\delta t} \pi_t(p) \, \mathrm{d}t + \int_\tau^\infty e^{-\delta t} \pi_\infty(p) \, \mathrm{d}t \right), \tag{20}$$

where $\pi_\infty(p)$ denotes the instantaneous revenue at the true quality ($\hat{q}_\infty = q$).[13] The monopolist's revenue maximization problem can be written as

$$\text{maximize} \left\{ \tilde{R}(p) \ : \ p_{\min} \leq p \leq p_{\max} \right\}. \tag{21}$$

Before stating our main result, we define

$$p^{\mathrm{m}}(q_0) := \underset{p \in [p_{\min}, p_{\max}]}{\operatorname{argmax}} \left\{ \pi_0(p) \right\} \quad \text{and} \quad p^{\mathrm{m}}(q) := \underset{p \in [p_{\min}, p_{\max}]}{\operatorname{argmax}} \left\{ \pi_\infty(p) \right\},$$

which are the static monopoly prices at q_0 and q respectively.[14] The following proposition characterizes the optimal monopoly price in the presence of social learning.

Proposition 5. *Consider the case $q < q_0$. For $\bar{\varepsilon}$ sufficiently small, the monopolist revenue optimization problem (21) has a unique optimal solution $p^* = p^*(\delta, w)$ that satisfies the following:*

(a) $p^* \in [p^{\mathrm{m}}(q_0), p^{\mathrm{m}}(q)]$.

(b) $p^*(\delta, w) \to p^{\mathrm{m}}(q)$ *as* $\delta w \to 0$ *and* $p^*(\delta, w) \to p^{\mathrm{m}}(q_0)$ *as* $\delta w \to \infty$.

Proposition 5 characterizes the (unique) solution of the monopolist's revenue maximization problem. In particular, Part (a) states that the optimal price is straddled between two natural end points: the price that a monopolist would charge if consumers did not engage in social learning and based their purchase decisions only on their prior estimate q_0; the price that a monopolist would charge if consumers were fully informed of the product quality q.

Part (b) highlights the importance of the monopolist's patience level δ and the weight w, that consumers attach to their prior, on the optimal price with social learning. In particular, if the monopolist is very patient ($\delta \approx 0$), then the optimal price is close to the price under full information, since in this case the learning transient is short relative to the extent of revenue discounting. However, if the monopolist is very impatient ($\delta \gg 0$), then she finds it optimal to significantly decrease her price, in the limit learning is not important and the monopolist prices as if $\hat{q}_t = q_0$ for all $t \geq 0$. Finally, note that the prior weight w is a natural time unit that determines the learning speed and the effect of discounting on revenues.

[13] Lemma 4 in the Appendix provides an intuitive characterization of the revenue function (20) and establishes a bound on $|\tilde{R}(p) - \bar{R}(p)|$, which is of order $\bar{\varepsilon}$.

[14] Note that Assumption 1(i) implies that the constraints in the definition of $p^{\mathrm{m}}(q)$ are never binding and $p^{\mathrm{m}}(q) = (\bar{\alpha} + q)/2$. Moreover, Assumption 1(ii) implies that the constraints in the definition of $p^{\mathrm{m}}(q_0)$ are not binding if and only if $(\bar{\alpha} + q_0)/2 > q_{\max} - \bar{\varepsilon}$, which is always true for reasonable values of $\bar{\alpha}$ and q_0.

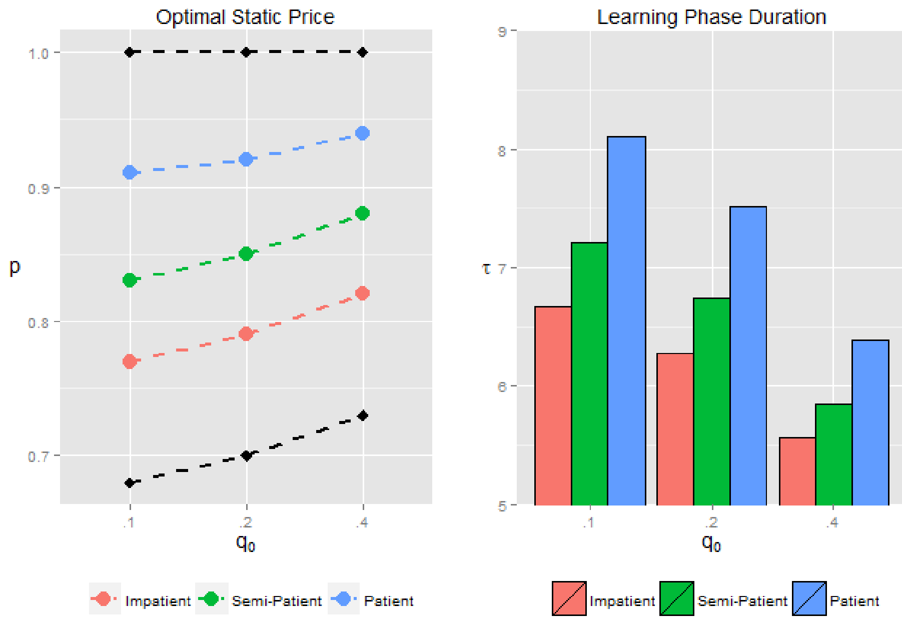

Figure 1: Optimal Static Price and Learning Phase Duration.

In what follows, we numerically illustrate the solution to (21) for different model parameters and derive some observations regarding comparative statics. We consider the underestimating case ($q_0 < q$) with a demand rate of 10 potential consumers per week. The most important parameters in the pricing problem are the monopolist's discount factor δ, the error in consumers prior estimate q_0 relative to the true quality, and the maximum base valuation $\bar{\alpha}$. As already noted, the learning transient also scales proportionally to the weight w attached to the prior estimate.

We consider three different prior estimates $q_0 \in \{.40, .20, .10\}$, with a prior weight of $w = 10$. The true quality is $q = 2$, and we set the small quality disturbance term $\bar{\varepsilon}$ to 5% of the true quality. The monopolist is either patient, semi-patient, or impatient, corresponding to annualized discount rates $\delta \in \{2.5\%, 7.5\%, 15\%\}$. We fix $\bar{\alpha} = 4$ and we think that this is a reasonable value for this parameter, which corresponds to a maximum quality premium of $2q$.[15]

The *left plot* in Figure 1 highlights how the optimal price p^* varies with the prior q_0 and the monopolist's patience level. The monopoly price under full information $p^{\mathrm{m}}(q)$, which is normalized to 1, and the monopolist price at q_0 are also plotted in black. In line with our theoretical result, we see that the optimal price with social learning is always between the static monopoly price at q_0 and the static monopoly price at q. Moreover, the price p^* is closer to the static monopoly price under full information when the monopolist is more patient and consumers' prior estimate q_0 is

[15]Note that if $\bar{\alpha} \gg q$ then learning becomes less important.

closer to q. On the contrary, when the monopolist is impatient, her optimal price is always closer to the static monopoly price at q_0. The *right plot* in Figure 1 reports the learning phase duration τ^* for different values of q_0 and different patience levels. It always takes 5.5 to 8.5 weeks for the quality estimate to get $\bar{\varepsilon}$-close to q. This observation is not surprising, since the facts that $w = 10$ and $q - \bar{\varepsilon} = 1.9$ imply that τ^* scales with

$$10 \log \left(\frac{\bar{\alpha} + 1.9 - p^*}{\bar{\alpha} + q_0 - p^*} \right),$$

however, the numerical results highlight that it always takes significantly longer to learn q when the monopolist is more patient. For the parameter values considered, learning q when the monopolist is patient always takes $15\% - 25\%$ longer than when the monopolist is impatient.

5 Two Price Analysis

Social learning implies a time varying demand process. As such, the ability to modify the price over time is valuable. Indeed, it is common for sellers to modify the prices of their products in proximity to their launching, for example by setting a low introductory price. Many factors and considerations, possibly separate from social learning, can support such pricing policies. A few examples include learning-by-doing, demand estimation, and endogenous timing of the purchasing decision (consumers with high valuations purchase first). These considerations are not part of our study which exclusively focuses on the impact of social learning on the dynamics of the pricing decision, and highlights the appeal of the tractable mean field approximation of the learning phenomenon to analyze the otherwise complex revenue optimization problem. For concreteness we focus on a two period pricing problem.

5.1 Optimal Prices

Consider the situation in which the monopolist can adjust her price once. She sets an initial price p_0 until time s, then p_1, and she can optimally choose (p_0, p_1, s) to maximize her discounted revenue objective. In this setting, we will show that the monopolist may choose to sacrifice short-term revenue to optimally speed up learning.

At the time of the price change consumers aggregate all information into a new prior $q_1 := \hat{q}_s$, i.e., the new prior equals the prevailing quality estimate at the time of the price change. Thus,

$$q_1 = p_0 - \bar{\alpha} + (\bar{\alpha} + q_0 - p_0) \exp \left(\frac{s}{w} \right),$$

moreover consumers use the following weight for the new prior,

$$w_1 = w + \Lambda \int_0^s \bar{F}(p - \hat{q}_t)\ \mathrm{d}t = w \left[1 + \Lambda \left(\exp\left(\frac{s}{w}\right) - 1\right) \frac{\bar{\alpha} + q_0 - p_0}{\bar{\alpha}}\right].$$

The (q_1, w_1) specification incorporates the fact that the reviews before time s were under a different price point. From time s onward the problem is analogous to the single price version studied in the previous section. In particular, after time s the learning process evolves according to equation (18), with initial condition q_1, price p_1, and prior weight w_1. The expected discounted revenue of the monopolist is given by

$$\bar{R}(p_0, p_1, s) = \Lambda \left(\int_0^s e^{-\delta t} \pi_t(p_0)\ \mathrm{d}t + \int_s^\infty e^{-\delta t} \pi_t(p_1)\ \mathrm{d}t \right).$$

As in the static price case, we study the situation in which once the learning process has converged to $\hat{q}_t = q - \bar{\varepsilon}$, revenues are accrued from then on according to $\hat{q}_t = q$. In this setting, it is without loss of generality to focus the attention on policies such that $s \leq \tau$.[16] Which leads to the following revenue function

$$\tilde{R}(p_0, p_1, s) = \Lambda \left(\int_0^s e^{-\delta t} \pi_t(p_0)\ \mathrm{d}t + \int_s^\tau e^{-\delta t} \pi_t(p_1)\ \mathrm{d}t + \int_\tau^\infty e^{-\delta t} \pi_\infty(p_1)\ \mathrm{d}t \right). \tag{22}$$

This setting is the natural extension to the static price case, indeed if $p_0 = p_1 = p$ one can easily verify that the revenue function as well as the learning process would be identical to the static price case. The monopolist solves the following optimization problem

$$\begin{aligned} \text{maximize} \quad & \tilde{R}(p_0, p_1, s) \\ \text{s.t.} \quad & p_0, p_1 \in [p_{\min}, p_{\max}] \\ & s \leq \tau, \end{aligned} \tag{23}$$

and the following proposition provides a characterization of the optimal pricing policy.

Proposition 6. *Consider the case $q_0 < q$ and assume that $\bar{\varepsilon}$ is sufficiently small. Let (p_0^*, p_1^*, s^*) be the optimal solution to (23). Then, the optimal prices (p_0^*, p_1^*) are such that $p_0^* \leq p_1^*$ and $p_1^* \in [p^{\mathrm{m}}(q_1), p^{\mathrm{m}}(q)]$.*

Proposition 6 states that, when consumers underestimate quality, the optimal price p_1^* is always between the static monopoly price $p^{\mathrm{m}}(q_1)$ and monopoly price under full information, $p^{\mathrm{m}}(q)$. Moreover, the optimal price path is increasing. This supports the intuition that the monopolist

[16] A formal argument is provided in the Appendix.

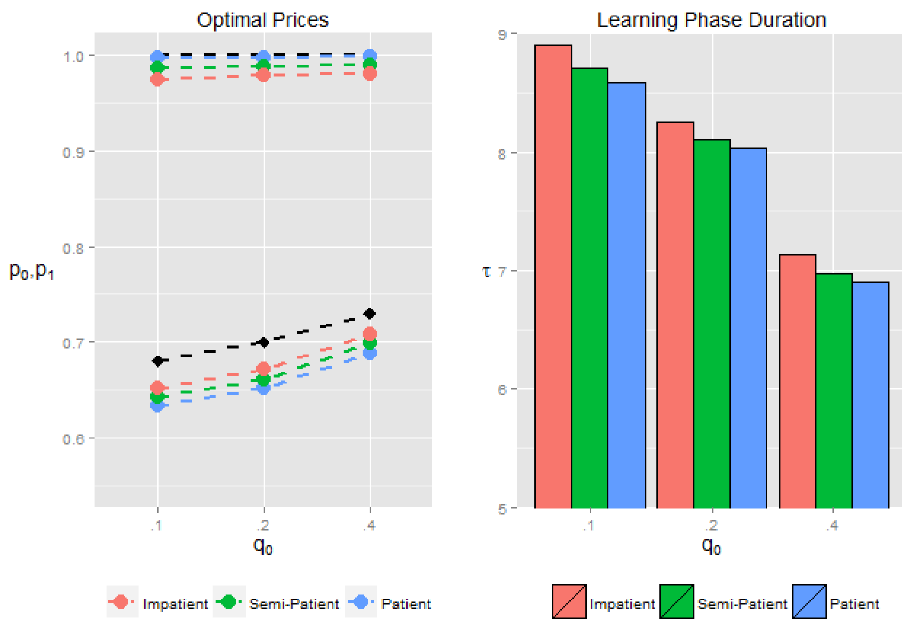

Figure 2: Optimal Prices and Learning Phase Duration.

has an incentive to lower the initial price in order to speed up learning.

Figure 2 displays numerical solutions to (23) for different model parameters, which provide some additional insights on the optimal pricing policy.[17] The *left plot* in Figure 2 shows the optimal prices p_1^* (above) and p_0^* (below) for different priors and different monopolist's patience levels. We clearly see that, in the two-price case, the monopolist may find it optimal to initially price below $p^{\mathrm{m}}(q_0)$ in order to speed up learning, and then switch to a price which is very close to the full information monopoly price $p^{\mathrm{m}}(q)$ in order to extract more revenues. Moreover, if the monopolist is more patient, i.e., δ is small, then the first period price p_0^* is lower. This has interesting implications for the speed of learning: the *right plot* in Figure 2 shows that a patient monopolist, who is willing to sacrifice initial revenues by under-pricing more aggressively, achieves faster learning than an impatient one.

5.2 Revenue Comparison of Pricing Policies

In this section we numerically compare the revenue performance of the optimal *static* price policy and the optimal *two* prices policy. Our measure of revenue performance for a given policy is the %-gap between the total revenue attained by using that policy and the total revenue that the monopolist would attain in an ideal scenario in which consumers know q and the monopolist

[17] The choice of parameter values for the numerical experiments is the same as the one described in Section 4.

	q_0	static	two	prior	true
Patient	**.40**	4.25	3.13	9.86	5.30
	.20	6.18	4.05	12.37	7.25
	.10	7.19	4.67	13.72	8.42
Semi-Patient	**.40**	12.05	9.21	14.66	14.05
	.20	15.31	11.61	18.13	18.72
	.10	17.06	12.91	19.97	21.44
Impatient	**.40**	19.32	16.29	20.41	23.64
	.20	23.78	20.18	24.83	30.50
	.10	26.09	22.22	27.12	34.32

Table 1: %-gap in revenues relative to full information scenario.

charges the monopoly price under full information.[18] We also report revenue performances for two benchmark policies: *prior* and *true*. For these cases, revenues are computed under the assumption that consumers follow the learning process specified in our model, but the monopolist does not take it into account and she charges the static price $p^{\mathrm{m}}(q_0)$ and the static price $p^{\mathrm{m}}(q)$ respectively.

The first two columns in Table 1 show the revenue performance of the *static* price and of the *two* prices policies respectively. The optimal two-period pricing policy performs consistently better than the optimal static price, and the relative revenue improvement becomes more significant as the seller's discount factor increases. In the last two columns of Table 1 we report the revenue performance of the *prior* and *true* policies respectively. By comparing the static price policy to these two benchmark policies we can appreciate the effectiveness of taking social learning into account when devising an optimal pricing policy.

Acknowledgments

The authors thank the Area Editor, the Associate Editor and the referees for their constructive comments that helped improve this paper. The authors also thank Kostas Bimpikis and Alireza Tahbaz-Salehi for insightful discussions, and Stefano Vaccari for his careful reading of the paper and his corrections.

References

Acemoglu, D., Bimpikis, K., and Ozdaglar, A. (2014) Dynamics of information exchange in endogenous social networks. *Theor. Econ.* **9**, 41–97.

[18]Note that when consumers underestimate quality, the latter provides an upper bound on the revenue that can be attained in our model by any policy.

URL http://dx.doi.org/10.3982/TE1204.

Acemoglu, D., Dahleh, M. A., Lobel, I., and Ozdaglar, A. (2009) Rate of convergence of learning in social networks. In *Proceedings of the American Control Conference 2009*, 2825–2830.

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

Acemoglu, D. and Ozdaglar, A. (2011) Opinion dynamics and learning in social networks. *Dynam. Games Appl.* **1**, 3–49.

Arieli, I. and Mueller-Frank, M. (2014) Learning what matters. Technical Report 2466903, SSRN.
URL http://ssrn.com/abstract=2466903.

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

Bass, F. M. (2004) Comments on "A new product growth for model consumer durables": the Bass model. *Management Sci.* **50**, 1833–1840.

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

Besbes, O. and Scarsini, M. (2015) On information distortions in online ratings. Technical Report 13-36, Columbia Business School Research.
URL http://ssrn.com/abstract=2266053.

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.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998) Learning from the behavior of others: conformity, fads, and informational cascades. *J. Econom. Perspect.* **12**, 151–170.

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

Boyd, S. and Vandenberghe, L. (2004) *Convex Optimization*. Cambridge University Press, Cambridge.
URL http://dx.doi.org/10.1017/CBO9780511804441.

Campbell, A. (2013) Word-of-mouth communication and percolation in social networks. *American Economic Review* **103**, 2466–2498.
URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.103.6.2466.

Candogan, O., Bimpikis, K., and Ozdaglar, A. (2012) Optimal pricing in networks with externalities. *Oper. Res.* **60**, 883–905.
URL http://dx.doi.org/10.1287/opre.1120.1066.

Debo, L. and Veeraraghavan, S. (2009) Joining longer queues: information externalities in queue choice. *Manufacturing Service Oper. Management* **11**, 543–562.

Debo, L. G., Parlour, C. A., and Rajan, U. (2012) Signaling quality via queues. *Management Sci.* **58**, 876–891.

Ellison, G. and Fudenberg, D. (1993) Rules of thumb for social learning. *J. Polit. Econ.* **101**, 612–643.

Ellison, G. and Fudenberg, D. (1995) Word-of-mouth communication and social learning. *Quart. J. Econom.* **110**, 93–125.

Gallego, G. and van Ryzin, G. (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Sci.* **40**, 999–1020.

Glynn, P. W. (1990) Diffusion approximations. In *Stochastic Models*, volume 2 of *Handbooks Oper. Res. Management Sci.*, 145–198. North-Holland, Amsterdam.
URL http://dx.doi.org/10.1016/S0927-0507(05)80168-9.

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 Econom. Behav.* **80**, 131–146.
URL http://dx.doi.org/10.1016/j.geb.2012.12.006.

Ifrach, B. (2012) Market Dynamics with Many Agents: Applications in Social Learning and Competition. Ph.D. thesis, Decision, Risk, and Operations, Columbia University.

Ifrach, B., Maglaras, C., and Scarsini, M. (2015) Bayesian social learning from consumer reviews. Technical Report 2293158, SSRN eLibrary.
URL http://ssrn.com/abstract=2293158.

Jadbabaie, A., Molavi, P., Sandroni, A., and Tahbaz-Salehi, A. (2012) Non-Bayesian social learning. *Games Econom. Behav.* **76**, 210–225.
URL http://dx.doi.org/10.1016/j.geb.2012.06.001.

Khalil, H. (2002) *Nonlinear Systems*. Prentice Hall, third edition.

Kurtz, T. G. (1977/78) Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl.* **6**, 223–240.

Kushner, H. J. and Yin, G. G. (2003) *Stochastic Approximation and Recursive Algorithms and Applications*. Springer-Verlag, New York, second edition.

Lafky, J. (2014) Why do people rate? Theory and evidence on online ratings. *Games Econom. Behav.* **87**, 554–570.
URL http://dx.doi.org/10.1016/j.geb.2014.02.008.

Li, X. and Hitt, L. M. (2008) Self-selection and information role of online product reviews. *Inform. Syst. Res.* **19**, 456–474.
URL http://isr.journal.informs.org/content/19/4/456.abstract.

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

Papanastasiou, Y., Bakshi, N., and Savva, N. (2014) Scarcity strategies under quasi-Bayesian social learning. Working paper, London Business School.

Papanastasiou, Y. and Savva, N. (2014) Dynamic pricing in the presence of social learning and strategic consumers. To appear in Management Science.

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

Talluri, K. T. and van Ryzin, G. J. (2005) *The Theory and Practice of Revenue Management*. Springer, New York.

Vettas, N. (1997) On the informational role of quantities: durable goods and consumers' word-of-mouth communication. *Internat. Econ. Rev.* **38**, 915–944.

A Appendix

Throughout this appendix, given a real number x we denote its orthogonal projection onto a closed real interval $[a,b]$ as $\text{proj}_{[a,b]}(x)$, given a vector $y = (y^1, \ldots, y^k)$ we define $|y| = \|y\|_1 = \sum_{j=1}^{k} |y^j|$, and given a function $y(t)$ of time, $\dot{y}(t)$ denotes its derivative.

Proofs of Section 3

Lemma 1. *The function $G(x)$ defined in (3) is non-increasing for all $x \leq \alpha_{\max}$, and is strictly decreasing for all $x \in [\alpha_{\min} - \bar{\varepsilon}, \alpha_{\max} - \bar{\varepsilon}]$. Equivalently, $G(p-q)$ is non-decreasing in q for all $q \geq p - \alpha_{\max}$ and is strictly increasing in q for all $q \in [q_{\min}, q_{\max}]$.*

Proof. First note that $G(x)$ is a well-defined function if and only if $x \leq \alpha_{\max}$. If $x < \alpha_{\min} - \bar{\varepsilon}$ then $G(x) = 1$. If $x \in [\alpha_{\min} - \bar{\varepsilon}, \alpha_{\min})$ then $G(x)$ is strictly decreasing because in this case $G(x) = \bar{F}(x + \bar{\varepsilon})$ and

$$\frac{\mathrm{d}G(x)}{\mathrm{d}x} = -f(x + \bar{\varepsilon}) < 0.$$

If $x \in [\alpha_{\min}, \alpha_{\max} - \bar{\varepsilon})$ we have

$$\frac{\mathrm{d}G(x)}{\mathrm{d}x} = \frac{\bar{F}(x + \bar{\varepsilon})}{\bar{F}(x)} \big(h(x) - h(x + \bar{\varepsilon})\big) < 0,$$

where the strict inequality follows from the assumption that α is IFR, or equivalently $h(x) - h(x + \bar{\varepsilon}) < 0$. Finally, if $x \geq \alpha_{\max} - \bar{\varepsilon}$, then $G(x) = 0$. Thus proving that $G(x)$ is non-increasing for all $x \leq \alpha_{\max}$ and strictly decreasing for all $x \in [\alpha_{\min} - \bar{\varepsilon}, \alpha_{\max} - \bar{\varepsilon}]$. □

A direct consequence of the above lemma is that $G(p-q)$ is invertible for every $q \in [p - \alpha_{\max} + \bar{\varepsilon}, p - \alpha_{\min} + \bar{\varepsilon}]$. This follows from the fact that G is invertible wherever it is strictly monotone.

Proof of Proposition 1. Taking logs of the weighted likelihood function in (6) yields

$$\log(\mathscr{L}_i^w(q)) = (L_0 + L_i)\log(.5 + .5G(p-q)) + (D_0 + D_i)\log(.5 - .5G(p-q)),$$

and differentiating the log-likelihood with respect to q we obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}q}\log(\mathscr{L}_i^w(q)) &= \frac{L_0 + L_i}{1 + G(p-q)} G'(p-q) - \frac{D_0 + D_i}{1 - G(p-q)} G'(p-q) \\ &= \left(\frac{L_0 + L_i + D_0 + D_i}{(1 + G(p-q))(1 - G(p-q))}\right) G'(p-q)(l_i - d_i - G(p-q)). \end{aligned} \tag{24}$$

Lemma 1 and Assumption 1 imply that $0 < G(p-q) < 1$ for all $q \in [q_{\min}, q_{\max}]$ and moreover that $G(p-q)$ is strictly increasing in q for all $q \in [q_{\min}, q_{\max}]$. We will now use these observations and (24) to construct the unique optimal solution to problem (8).

First, note that $0 < G(p-q) < 1$ implies that the denominator in (24) is always positive. Since $G(p-q)$ is strictly increasing in q, then $l_i - d_i - G(p-q)$ is strictly decreasing in q. If

$l_i - d_i - G(p - q_{\min}) \le 0$, then

$$\frac{\mathrm{d}}{\mathrm{d}q} \log(\mathscr{L}_i^w(q)) \le 0 \quad \text{for all } q \in [q_{\min}, q_{\max}],$$

thus $\hat{q}_i = q_{\min}$ maximizes the log-likelihood. If $l_i - d_i - G(p - q') > 0$ for some $q' \in [q_{\min}, q_{\max}]$, then

$$\frac{\mathrm{d}}{\mathrm{d}q} \log(\mathscr{L}_i^w(q)) > 0 \quad \text{for all } q < q'.$$

This implies that the quality that maximizes the log-likelihood is the solution to

$$\frac{\mathrm{d}}{\mathrm{d}q} \log(\mathscr{L}_i^w(q)) = 0 \quad \Longleftrightarrow \quad G(p - q) = l_i - d_i,$$

given by q^* if $l_i - d_i - G(p - q_{\max}) \le 0$, or by $q_{\max}$ if $l_i - d_i - G(p - q_{\max}) > 0$. Note that since $G(p - q)$ is strictly increasing for all $q \in [q_{\min}, q_{\max}]$ then $\hat{q}_i = q^*$ is always unique in $[q_{\min}, q_{\max}]$.

Summarizing the above conditions, the quality estimate that maximizes the log-likelihood can be defined as follows: if $l_i \le d_i$ then $l_i - d_i - G(p - q_{\min}) \le 0$ and thus $\hat{q}_i = q_{\min}$; otherwise, if $l_i > d_i$ then $\hat{q}_i = \mathrm{proj}_{[q_{\min}, q_{\max}]}(q^*)$. To complete the proof, note that $d_i = 1 - l_i$ implies $l_i - d_i = 2l_i - 1$. □

The following lemma is instrumental in the proof of Proposition 2.

Lemma 2. *Suppose that Assumption 1(i) holds, then $\sum_{i=1}^{\infty} (B_i)^{-2} < \infty$ almost surely.*

Proof. The proof will proceed as follows. First, we rewrite the process $\{B_i, i = 1, 2, \dots\}$ in a more convenient form $\{B_0 + X_i, i = 1, 2, \dots\}$, where $B_0 := L_0 + D_0$ is the total number of initial reviews associated to the prior q_0, defined in (4) and (5), and X_i is an appropriately defined sequence of random variables. Then, we bound from below the process $\{X_i, i = 1, 2, \dots\}$ with a process $\{Y_i, i = 1, 2, \dots\}$, which is more tractable for the purpose of the analysis. Finally, we use the Strong Approximation Theorem Glynn (1990, Theorem 5) to show that $\sum_{i=1}^{\infty} (B_0 + Y_i)^{-2} < \infty$ almost surely and complete the proof.

Note that Assumption 1(i) implies that for all i there exists an $\eta > 0$ such that for all admissible p the following is true:

$$\mathsf{P}(i\text{-th customer buys} \mid I_{i-1}, p) \ge \mathsf{P}(i\text{-th customer buys} \mid q_{\min}, p) \ge 2\eta,$$

where I_i is defined as in (1). Thus, we can rewrite B_i in the form $B_i = B_0 + \sum_{j=1}^{i} \chi_j(\eta_j)$, where χ_j is a Bernoulli random variable with success (i.e., purchase) probability $\eta_j > \eta$ for all j, and η_j depends on I_j and the price p. Let $X_i = \sum_{j=1}^{i} \chi_j$ for all $i = 1, 2, \dots$.

Next, define the random variables $\xi_j = \chi_j(\eta_j) \cdot \upsilon(\eta/\eta_j)$, where the random variable $\upsilon(\eta/\eta_j)$ is Bernoulli with success probability η/η_j, independent of χ_j. That is, ξ_j is a random sample of the customers that purchased. It is easy to verify that the distribution ξ_j is Bernoulli with success probability η and that ξ_j is independent of ξ_k for all $j \neq k$. Let $Y_i = \sum_{j=1}^{i} \xi_j$ and note that, by construction, $Y_i + B_0 \leq X_i + B_0 = B_i$ for all $i = 1, 2, \ldots$.

Finally, the Strong Approximation Theorem Glynn (1990, Theorem 5) implies that there exist a probability space that supports a standard Brownian motion W and a sequence Y_i' such that $\{Y_i' : i \geq 1\} \overset{\mathscr{D}}{=} \{Y_i : i \geq 1\}$, and

$$Y_i' = i \cdot \eta + \sigma_\eta W(i) + O(\log\ i) \quad \text{a.s.},$$

and $W(\cdot)$ is a standard Brownian motion. (The symbol $\overset{\mathscr{D}}{=}$ denotes equality in distribution.) In the sequel we write Y_i instead of Y_i'. Rewriting the above expression we have that

$$Y_i = (i) \cdot \left(\eta + \sigma_\eta \frac{W(i)}{i} + O\left(\frac{\log\ i}{i}\right)\right) \quad \text{a.s.}$$

From the strong law of large numbers for the standard Brownian motion we know that $W(i)/i \to 0$ a.s., which implies that, for any $\varepsilon > 0$, there exists a constant $M_1 > 0$ such that $W(i)/i < \varepsilon$ for all $i > M_1$ and almost all sample paths. Similarly, there exists a constant $M_2 > 0$ such that the error term $O(\log(i)/i) < M_2$ for all $i = 1, \ldots$. It follows that there exists a constant $M_3 > 0$ such that

$$\sum_{i=1}^{\infty} (B_0 + Y_i)^{-2} = \sum_{i \leq M_1} (B_0 + Y_i)^{-2} + \sum_{i > M_1} (B_0 + Y_i)^{-2} < M_1 \cdot B_0^{-2} + \sum_{i > M_1} \frac{M_3}{i^2} < \infty \quad \text{a.s.}$$

Noting that $B_i \leq B_0 + Y_i$ for all $i = 1, 2, \ldots$ implies that $\sum_{i=1}^{\infty} (B_i)^{-2} \leq \sum_{i=1}^{\infty} (B_0 + Y_i)^{-2}$ completes the proof. □

Proof of Proposition 2. We study the evolution of l_i, and relate it to $\hat{q}_i$ using (10). It is easy to verify that l_i evolves according to the stochastic recursion

$$\begin{aligned} l_i &= \text{proj}_{[l_{\min}, l_{\max}]} \left[l_{i-1} + (B_i)^{-1} \Big(L_0 + L_{i-1} + \mathbf{1}\left\{ r_{i-1} = r^{\text{L}} \right\} - (L_0 + L_{i-1}) B_i / B_{i-1} \Big) \right] \\ &= \text{proj}_{[l_{\min}, l_{\max}]} \left[l_{i-1} + (B_i)^{-1} \Big((1 - l_{i-1}) \mathbf{1}\left\{ r_{i-1} = r^{\text{L}} \right\} - l_{i-1} \mathbf{1}\left\{ r_{i-1} = r^{\text{D}} \right\} \Big) \right], \end{aligned}$$

where $l_{\min} = 0.5(1 + G(p - q_{\min}))$ and $l_{\max} = 0.5(1 + G(p - q_{\max}))$. Setting $Y_i = (1 - l_i) \mathbf{1}\left\{ r_i = r^{\text{L}} \right\} - l_i \mathbf{1}\left\{ r_i = r^{\text{D}} \right\}$, the iterative process can be rewritten as

$$l_i = \text{proj}_{[l_{\min}, l_{\max}]} \left[l_{i-1} + (B_i)^{-1} Y_{i-1} \right],$$

which is equivalent to the recursion

$$l_i = l_{i-1} + (B_i)^{-1} Y_{i-1} + (B_i)^{-1} Z_{i-1},$$

where the *projection* term $Z_{i-1} := l_i - l_{i-1} - (B_i)^{-1} Y_{i-1}$.

This stochastic recursion process belongs to the class of processes studied in Kushner and Yin (2003). We next show that the assumptions of (Kushner and Yin, 2003, Chapter 5, Theorem 2.1) hold and then identify the equilibrium point of the process using a Lyapunov function. For this purpose, it is useful to define the maximum likelihood quality estimate as a function of the fraction of likes

$$\hat{q}(l_i) := p - G^{-1}(2l_i - 1),$$

and note that Lemma 1 implies that $G^{-1}(2l_i - 1)$ is well-defined for all $l_i \in [l_{\min}, l_{\max}]$.

Assumption (A.2.1). By subadditivity of the absolute value, it follws that

$$|Y_i| \le (1 - l_i)\mathbf{1}\left\{r_i = r^{\mathrm{L}}\right\} + l_i \mathbf{1}\left\{r_i = r^{\mathrm{D}}\right\} \quad \forall i,$$

and thus

$$\begin{aligned} |Y_i|^2 &\le (1 - l_i)^2 \mathbf{1}\left\{r_i = r^{\mathrm{L}}\right\} + l_i^2 \mathbf{1}\left\{r_i = r^{\mathrm{D}}\right\} + 2(1 - l_i) l_i \mathbf{1}\left\{r_i = r^{\mathrm{L}}\right\} \mathbf{1}\left\{r_i = r^{\mathrm{D}}\right\} \\ &= (1 - l_i)^2 \mathbf{1}\left\{r_i = r^{\mathrm{L}}\right\} + l_i^2 \mathbf{1}\left\{r_i = r^{\mathrm{D}}\right\} \le 1 \quad \forall i, \end{aligned}$$

since $\left\{r_i = r^{\mathrm{L}}\right\}$ and $\left\{r_i = r^{\mathrm{D}}\right\}$ are mutually exclusive. It then follows that $\sup_i E|Y_i|^2 \le 1 < \infty$.

Assumption (A.2.2). We have that

$$\begin{aligned} \mathsf{E}[Y_i \mid l_0, Y_j, j < i] &= (1 - l_i)\mathsf{P}(r_i = r^{\mathrm{L}}) - l_i \mathsf{P}(r_i = r^{\mathrm{D}}) \\ &= .5\Big[\bar{F}\big(p - \min(q + \bar{\varepsilon}, \hat{q}(l_i))\big) + \bar{F}\big(p - \min(q - \bar{\varepsilon}, \hat{q}(l_i)))\big)\Big] - l_i \bar{F}\big(p - \hat{q}(l_i)\big) \\ &= .5\Big[\bar{F}\big(p - \min(\hat{q}(l_i), q + \bar{\varepsilon})\big) + \bar{F}\big(p - \min(\hat{q}(l_i), q - \bar{\varepsilon})\big) - \bar{F}\big(p - \hat{q}(l_i) + \bar{\varepsilon}\big) - \bar{F}\big(p - \hat{q}(l_i)\big)\Big] \end{aligned}$$

where the last equality follows by substituting

$$l_i = .5(1 + G(p - \hat{q}(l_i))) = .5\left(1 + \frac{\bar{F}(p - \hat{q}(l_i) + \bar{\varepsilon})}{\bar{F}(p - \hat{q}(l_i))}\right).$$

Thus, we can define the function

$$g(l_i) := \mathsf{E}[Y_i \mid l_0, Y_j, j < i], \tag{25}$$

which is measurable since $F(\cdot)$ is measurable. Finally, note that the above derivation implies that finite difference bias terms $\beta_i = 0, \forall i$, which follows from the fact that the distribution function F is known.

Assumption (A.2.3) follows from the fact that F is continuous.

Assumption (A.2.4) is shown in Lemma 2.

Assumption (A.2.5) is immediate since $\beta_i = 0, \forall i$, as the proof of Assumption (A.2.2) shows.

Before applying the theorem it is useful to further decompose l_i as follows. First note that $Y_i := g(l_i) + \bar{M}_i$, where the function $g(l_i)$ is the drift function that we defined in (25) and $\bar{M}_i$ is a martingale difference noise given by

$$\bar{M}_i = (1 - l_i)\mathbf{1}\left\{r_i = r^{\mathrm{L}}\right\} - l_i\mathbf{1}\left\{r_i = r^{\mathrm{D}}\right\} - g(l_i).$$

Then, it is straightforward to see that

$$l_i = l_{i-1} + (B_i)^{-1}g(l_{i-1}) + (B_i)^{-1}Z_{i-1} + (B_i)^{-1}\bar{M}_{i-1}.$$

Now we can apply (Kushner and Yin, 2003, Chapter 5, Theorem 2.1), to conclude that l_i converges almost surely to the set of locally asymptotically stable points of the ODE $\dot{l} = g(l)$ that we denote with S. We next show that the ODE has a unique locally asymptotically stable point denoted by $l^* := .5(1 + G(p - q))$. For that purpose we define the candidate Lyapunov function $V(l) = (l - l^*)^2$. We need to show that $\dot{V}(l) = \nabla V(l)g(l) < 0$ for all $l \in [0, 1]/\{l^*\}$ and $\dot{V}(l^*) = 0$. See Khalil (2002) for details on Lyapunov stability. Thus, we have to show that $g(l) > (<)0$ when $l < (>)l^*$ (or equivalently when $\hat{q}(l) < (>)q$).

Case 1: $l < l^*$ **(or equivalently** $\hat{q}(l) < q$**).** In this case $\min(\hat{q}(l), q + \bar{\varepsilon}) = \hat{q}(l)$ and $g(l) = .5\big[\bar{F}(p - \min(\hat{q}(l), q - \bar{\varepsilon})) - \bar{F}(p - \hat{q}(l) + \bar{\varepsilon})\big]$. In addition, $-\min(\hat{q}(l), q - \bar{\varepsilon}) \geq -\hat{q}(l) + \bar{\varepsilon}$. If $\hat{q}(l) \leq q - \bar{\varepsilon}$, then $g(l) = .5\big[\bar{F}(p - \hat{q}(l)) - \bar{F}(p - \hat{q}(l) + \bar{\varepsilon})\big] > 0$, since $\bar{F}$ is a decreasing function. If $q > \hat{q}(l) > q - \bar{\varepsilon}$, then $g(l) = .5\left[\bar{F}(p - q + \bar{\varepsilon}) - \bar{F}(p - \hat{q}(l) + \bar{\varepsilon})\right] > 0$, since $\hat{q}(l) < q$. We conclude that $g(l) > 0$ in this case.

Case 2: $l > l^*$ **(or equivalently** $\hat{q}(l) > q$**).** In this case g simplifies to

$$g(l) = .5\left[\bar{F}\big(p - \min(\hat{q}(l), q + \bar{\varepsilon})\big) + \bar{F}\big(p - q + \bar{\varepsilon}\big) - \bar{F}\big(p - \hat{q}(l) + \bar{\varepsilon}\big) - \bar{F}\big(p - \hat{q}(l)\big)\right],$$

which can be shown to be negative, using $-\min(\hat{q}(l), q + \bar{\varepsilon}) > -\hat{q}(l)$ and that $\bar{F}$ is decreasing.

We conclude that $\dot{V}(l) < 0$ at all points $l \neq l^*$. It is easy to verify through the above expressions that at $l = l^*$ and $\hat{q}(l) = q$ we get that $\dot{V}(l^*) = 0$. Also, by construction $V(l^*) = 0$, which establishes that $S = \{l^*\}$ and that $l_i \to l^*$ almost surely. Applying the continuous mapping theorem, we get that $\hat{q}_i \to q$ almost surely, which completes the proof. $\square$

Lemma 3. (a) *For all $x, y, z \in \mathbb{R}$ we have*

$$|\min(x, y) - \min(z, y)| \leq |x - z| \quad \text{and} \quad |\max(x, y) - \max(z, y)| \leq |x - z|.$$

(b) *For $x_1, x_2 \geq a > 0$ and $y_1, y_2 \geq b > 0$*

$$\Big|\frac{x_1}{x_1 + y_1} - \frac{x_2}{x_2 + y_2}\Big| \leq \frac{1}{a}(|x_1 - x_2| + |y_1 - y_2|).$$

Proof. (a) Minimum operator: If $x, z \geq y$ or $x, z \leq y$ this holds trivially. If $x \leq y$ and $z \geq y$ then $|\min(x, y) - \min(z, y)| = |x - y| = y - x \leq z - x = |x - z|$. For the maximum operator take $-x$, $-y$, and $-z$.

(b) From the triangular inequality,

$$\begin{aligned}
\Big|\frac{x_1}{x_1 + y_1} - \frac{x_2}{x_2 + y_2}\Big| &\leq \Big|\frac{x_1}{x_1 + y_1} - \frac{x_1}{x_1 + y_2}\Big| + \Big|\frac{x_1}{x_1 + y_2} - \frac{x_2}{x_2 + y_2}\Big| \\
&= \frac{x_1}{x_1 + y_1}\Big|\frac{y_1 - y_2}{x_1 + y_2}\Big| + \frac{y_2}{x_1 + y_2}\Big|\frac{x_1 - x_2}{x_2 + y_2}\Big| \\
&\leq \frac{1}{a}(|x_1 - x_2| + |y_1 - y_2|).
\end{aligned}$$

$\square$

Proof of Proposition 3. Throughout this proof, to reduce the notational burden and without loss of generality, we rescale time such that $\Lambda = 1$. We verify the conditions of Theorem 2.2 of Kurtz (1977/78). First we note that $X^n(t) \in \mathbb{Z}_+^2$, $\bar{X}^n(t) = X^n(t)/n \in \{k/n \mid k \in \mathbb{Z}_+^2\}$ as required. To satisfy the conditions of the theorem we validate the construction (12) and then show that the following inequalities hold

$$\gamma(x) \leq \Gamma_1(1 + |x|) \quad \text{and} \quad |\gamma(x) - \gamma(y)| \leq \Gamma_2|x - y| \tag{26}$$

for $x, y \in \mathbb{R}^2$ and $x, y \geq [L_0, D_0]$ componentwise, and for some finite constants Γ_1 and Γ_2.

The integral form of $\bar{X}^n(t)$ in (12) follows from Poisson arrivals and Poisson thinning of the standard Poisson process N. For example, $\bar{L}^n(t)$ can be written in the form,

$$\begin{aligned}\bar{L}^n(t) &= \frac{1}{n}\int_0^t \mathbf{1}\left\{r_s = r^{\mathrm{L}} \mid \bar{X}^n(s)\right\}\, \mathrm{d}A^n(s) \\ &= \frac{1}{n}N^{\mathrm{L}}\left(\int_0^t \mathsf{P}\left(r_s = r^{\mathrm{L}} \mid \bar{X}^n(s)\right)\, \mathrm{d}s\right) \\ &= \frac{1}{n}N^{\mathrm{L}}\left(\int_0^t \gamma^{\mathrm{L}}\left(\bar{X}^n(s)\right)\, \mathrm{d}s\right),\end{aligned}$$

where A^n is a Poisson process with rate n and, with some abuse of notation, r_s is a review given by a consumer arriving at time s. The second equality follows by splitting the Poisson process into likes, dislikes, and outside options; the probability with which an arriving consumer submits one of these reviews depends on his quality preference and on his observable information $X^n(s)$. The Poisson thinning property guarantees that the process that counts only those consumers who like the product is still Poisson with rate proportional to the probability of liking the product. Similarly, this can be shown for $D^n(t)$.

Finally, focusing on the rate conditions required for Theorem 2.2 of Kurtz (1977/78). The first inequality in (26) holds for $\Gamma_1 = 1$ since γ^k are probabilities for $k = \mathrm{L}, \mathrm{D}$. We derive the last inequality there for γ^{L} in two steps. First, we observe that $\gamma^{\mathrm{L}}(X^n(t))$ depends on $X^n(t)$ through $\hat{q}(X^n(t))$. It follows from Lemma 3(a) and the fact that the density of α is uniformly bounded by $f_{\max}$ that γ^{L} is Lipschitz continuous in $q \in [q_{\min}, q_{\max}]$. Second, we show that $\hat{q}(x)$ is Lipschitz continuous in $x = X^n(t)$. Lemma 3(b) for $a = L_0$ and $b = D_0$ establishes that q^*, defined in (10), is Lipschitz continuous in $x = X^n(t)$. The projection operator of q^* onto $[q_{\min}, q_{\max}]$ preserves the Lipschitz property. Similarly, one can show that γ^{D} is Lipschitz continuous, which establishes that γ is Lipschitz. Finally, invoking Theorem 2.2 of Kurtz (1977/78) we complete the proof. □

Proof of Proposition 4. The ODEs (16) reduce to

$$\dot{\bar{L}}(t) = \Lambda\left(\frac{\bar{\alpha} + \hat{q}_t - p}{\bar{\alpha}}\right) \quad \text{and} \quad \dot{\bar{D}}(t) = 0,$$

noting that $\bar{D}(t) = 0$ and plugging (17) into the expression for $\dot{\bar{L}}(t)$ yields a linear ODE for $\bar{L}(t)$, i.e., $\dot{\bar{L}}(t) = \frac{\Lambda\bar{\varepsilon}}{2\bar{\alpha}}(1 + \bar{L}_0/\bar{D}_0 + \bar{L}(t)/\bar{D}_0)$. This equation is readily solvable in closed form, and the particular solution with initial condition $\bar{L}_0$ is given by

$$\bar{L}(t) = (\bar{L}_0 + \bar{D}_0)\left[\exp\left(\frac{\Lambda\bar{\varepsilon}}{2\bar{\alpha}\bar{D}_0}t\right) - 1\right]. \tag{27}$$

The trajectory for $\hat{q}_t$ can now be obtained by replacing (27) and $\bar{D}(t) = 0$ into (17), which yields

$$\hat{q}_t = p - \bar{\alpha} + \frac{\bar{\varepsilon}}{2}\left(\frac{\bar{L}_0 + \bar{D}_0}{\bar{D}_0}\right)\exp\left(\frac{\Lambda\bar{\varepsilon}}{2\bar{\alpha}\bar{D}_0}t\right),$$

and the time-to-learn can be calculated by setting $\hat{q}_\tau = q - \bar{\varepsilon}$ and then solving for τ, which yields

$$\tau = \frac{2\bar{\alpha}\bar{D}_0}{\Lambda\bar{\varepsilon}}\log\left(\frac{2(\bar{\alpha} + q - \bar{\varepsilon} - p)}{\bar{\varepsilon}}\left(\frac{\bar{D}_0}{\bar{L}_0 + \bar{D}_0}\right)\right).$$

Rewriting (4) and (5) for α uniformly distributed we get that

$$\bar{L}_0 = \frac{w\Lambda(\bar{\alpha} + q_0 - p - \bar{\varepsilon}/2)}{\bar{\alpha}} \quad \text{and} \quad \bar{D}_0 = \frac{w\Lambda\bar{\varepsilon}}{2\bar{\alpha}},$$

and plugging them in $\hat{q}_t$ and τ yields

$$\hat{q}_t = p - \bar{\alpha} + (\bar{\alpha} + q_0 - p)\exp\left(\frac{t}{w}\right) \quad \text{and} \quad \tau = w\log\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right).$$

□

In the sequel, we provide a sketch for the solution of the ODE's derived in the main body of the paper for the cases not considered in Proposition 4.

Overestimating prior ($q_0 > q$); phase 1 of learning $\hat{q}_t > q + \bar{\varepsilon}$. The following proposition characterizes the trajectories for $\bar{L}(t)$ and $\bar{D}(t)$ in this case. From these one can easily obtain the trajectory for $\hat{q}_t$ by using (17).

Proposition 7. *Consider the ODEs for the learning dynamics given in (16) and assume that $q_0 > q$. Then, for $t \leq \tau$,*

$$\bar{L}(t) = \Lambda\left(\frac{\bar{\alpha} + q - p}{\bar{\alpha}}\right)t, \tag{28}$$

and $\bar{D}(t)$ is defined implicitly by

$$\left(1 + \frac{\bar{L}(t) + \bar{D}(t)}{\bar{L}_0 + \bar{D}_0}\right)^{\bar{\alpha}+q-p}\left(1 + \frac{(\bar{\varepsilon}/2)\bar{L}(t) - (\bar{\alpha} + q - p)\bar{D}(t)}{(\bar{\varepsilon}/2)\bar{L}_0 - (\bar{\alpha} + q - p)\bar{D}_0}\right)^{\bar{\varepsilon}/2} = 1. \tag{29}$$

Proof. (Sketch only.) In this case the ODEs (16) reduce to

$$\dot{\bar{L}}(t) = \Lambda\left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}}\right) \quad \text{and} \quad \dot{\bar{D}}(t) = \frac{\Lambda}{\bar{\alpha}}\left[p-\bar{\alpha}+\frac{\bar{\varepsilon}}{2}\left(1+\frac{\bar{L}_0+\bar{L}(t)}{\bar{D}_0+\bar{D}(t)}\right)-q\right].$$

It can be easily verified that the solution for $\bar{L}(t)$ with initial conditions $\bar{L}_0$ is given by (28), and substituting it into the ODE for dislikes yields

$$\dot{\bar{D}}(t) = \frac{\Lambda}{\bar{\alpha}}\left[\frac{\bar{\varepsilon}}{2}+\frac{\bar{\varepsilon}}{2}\left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}}\right)\frac{\Lambda t}{\bar{D}_0+\bar{D}(t)}+\frac{\bar{\varepsilon}}{2}\frac{\bar{L}_0}{\bar{D}_0+\bar{D}(t)}-(\bar{\alpha}+q-p)\right].$$

Which is an ODE of the form

$$\dot{\bar{D}}(t) = a + b\frac{t}{\bar{D}_0+\bar{D}(t)} + c\frac{1}{\bar{D}_0+\bar{D}(t)},$$

where

$$a = \frac{\Lambda}{\bar{\alpha}}\left(p-\bar{\alpha}-q+\frac{\bar{\varepsilon}}{2}\right), \quad b = \frac{\Lambda^2\bar{\varepsilon}}{2\bar{\alpha}^2}(\bar{\alpha}+q-p), \quad c = \frac{\Lambda\bar{\varepsilon}\bar{L}_0}{2\bar{\alpha}}.$$

The above ODE may be converted into an equation with separable variables by setting

$$Z(t) = \frac{bt+c}{\bar{D}_0+\bar{D}(t)}, \quad \text{that is,} \quad \bar{D}(t) = \frac{bt+c}{Z(t)} - \bar{D}_0.$$

In fact, we have

$$\dot{\bar{D}}(t) = \frac{b}{Z(t)} - \frac{bt+c}{Z(t)^2}\dot{Z}(t), \quad \text{which implies that} \quad \dot{Z}(t) = (bZ(t)-aZ(t)^2-Z(t)^3)\left(\frac{1}{bt+c}\right).$$

This equation for $Z(t)$ is solvable, which, in turn, leads to the solution for $\bar{D}(t)$ given in (29). □

Phase 2 of learning $q-\bar{\varepsilon} < \hat{q}_t < q+\bar{\varepsilon}$. In this case the ODEs are less tractable, nonetheless we will establish a useful structural property for $\hat{q}_t$ in Proposition 8, and then provide an approximation of the $\hat{q}_t$ trajectory when $\bar{\varepsilon}$ is small.

Proposition 8. *Assume that $q-\bar{\varepsilon} < \hat{q}_{t_0} < q+\bar{\varepsilon}$ for some $t_0 > 0$. Then, $\hat{q}_t \to q$ as $t \to \infty$. Moreover, if $\hat{q}_{t_0} < q$ then $\hat{q}_t$ is strictly monotonically increasing for all $t \geq t_0$, otherwise, if $\hat{q}_{t_0} > q$ then $\hat{q}_t$ is strictly monotonically decreasing for all $t \geq t_0$.*

Proof. (Sketch only.) The ODEs (16) reduce to

$$\dot{\bar{L}}(t) = \frac{\Lambda\bar{\varepsilon}}{4\bar{\alpha}}\left(1 + \frac{\bar{L}_0 + \bar{L}(t)}{\bar{D}_0 + \bar{D}(t)}\right) + \frac{\Lambda}{2\bar{\alpha}}(\bar{\alpha} + q - \bar{\varepsilon} - p), \tag{30}$$

$$\dot{\bar{D}}(t) = \frac{\Lambda\bar{\varepsilon}}{4\bar{\alpha}}\left(1 + \frac{\bar{L}_0 + \bar{L}(t)}{\bar{D}_0 + \bar{D}(t)}\right) - \frac{\Lambda}{2\bar{\alpha}}(\bar{\alpha} + q - \bar{\varepsilon} - p). \tag{31}$$

It follows from Assumption 1 that $p < \bar{\alpha} - \bar{\varepsilon}$, which implies that $\dot{\bar{L}}(t) > 0$ for all t, moreover, using equation (17) for $\hat{q}_t$, it is easy to see that $\dot{\bar{D}}(t) > 0$ if and only if $\hat{q}_t > q - \bar{\varepsilon}$. These, coupled with the fact that $L(t_0)$ and $D(t_0)$ are strictly positive, imply that $\bar{L}(t)$ and $\bar{D}(t)$ are positive and increasing for all $t \geq t_0$. Using $\bar{l}(t) = \bar{L}_0 + \bar{L}(t)/(\bar{L}_0 + \bar{L}(t) + \bar{D}_0 + \bar{D}(t))$, we can now write the ODE for the fraction of likes $\bar{l}(t)$ as

$$\dot{\bar{l}}(t) = \frac{1}{\bar{L}_0 + \bar{L}(t) + \bar{D}_0 + \bar{D}(t)}\left[(1 - \bar{l}(t))\dot{\bar{L}}(t) - \bar{l}(t)\dot{\bar{D}}(t)\right], \tag{32}$$

and noting that $\bar{L}_0 + \bar{L}(t) + \bar{D}_0 + \bar{D}(t) > 0$ for all $t \geq t_0$, we have that the steady state l^* must be such that (32), evaluated at l^*, is equal to 0. Noting that $1 + (\bar{L}_0 + \bar{L}(t))/(\bar{D}_0 + \bar{D}(t)) = 1/(1 - \bar{l}(t))$ and replacing (30) and (31) into condition (32) we find the unique steady state

$$l^* = \frac{\bar{\alpha} + q - p - \bar{\varepsilon}/2}{\bar{\alpha} + q - p}.$$

It is easy to verify that $\dot{\bar{l}}(t) > 0$ if and only if $\bar{l}(t) < l^*$. Recalling that $\hat{q}_t = p - \bar{\alpha} + \dfrac{\bar{\varepsilon}}{2(1 - \bar{l}(t))}$ and replacing in l^* for $\bar{l}(t)$ we can readily verify that quality in steady state is equal to q. Moreover, since $\hat{q}_t$ is increasing in $\bar{l}(t)$, it follows from $\dot{\bar{l}}(t) > 0 \Leftrightarrow \bar{l}(t) < l^*$ that $\mathrm{d}\hat{q}_t/\mathrm{d}t > 0 \Leftrightarrow \hat{q}_t < q$. Thus, if $\hat{q}_{t_0} < q$ then $\hat{q}_t$ is strictly monotonically increasing in t, otherwise it is strictly monotonically decreasing. □

When $\bar{\varepsilon}$ is small we can approximate the trajectories of $\bar{L}(t)$ and $\bar{D}(t)$ as follows. Note that the initial conditions for the system of ODEs defined by (30) and (31), when $\hat{q}_\tau = q - \bar{\varepsilon}$, are given by $(t_0 = \tau, \bar{L}_\tau, \bar{D}_\tau, \dot{\bar{L}}_\tau, \dot{\bar{D}}_\tau)$ where

$$\bar{L}_\tau = w\frac{\Lambda}{\bar{\alpha}}(\bar{\alpha} + q - 2\bar{\varepsilon} - p), \quad \bar{D}_\tau = w\frac{\Lambda}{2\bar{\alpha}}\bar{\varepsilon}, \quad \dot{\bar{L}}_\tau = \frac{\Lambda}{\bar{\alpha}}(\bar{\alpha} + q - \bar{\varepsilon} - p), \quad \dot{\bar{D}}_\tau = 0. \tag{33}$$

First, note that we can write

$$\bar{L}(t) = \bar{L}_\tau + \int_\tau^{\tau+t} \dot{\bar{L}}(s)\ \mathrm{d}s = \bar{L}_\tau + \dot{\bar{L}}_\tau t + \int_\tau^{\tau+t} (\dot{\bar{L}}(s) - \dot{\bar{L}}_\tau)\ \mathrm{d}s,$$
$$\bar{D}(t) = \bar{D}_0 + \int_\tau^{\tau+t} \dot{\bar{D}}(s)\ \mathrm{d}s = \bar{D}_0 + \int_\tau^{\tau+t} \dot{\bar{D}}(s)\ \mathrm{d}s.$$

From (30) and (33) it is easy to see that $\dot{\bar{L}}(s) - \dot{\bar{L}}_\tau = \dot{\bar{D}}(s)$, and (31) and (33) imply that

$$\dot{\bar{D}}(s) = \frac{\Lambda\bar{\varepsilon}}{4\bar{\alpha}}\left(\frac{\bar{L}(s)}{\bar{D}(s)} - \frac{\bar{L}_\tau}{\bar{D}_\tau}\right).$$

We already established that $\dot{\bar{D}}(s) > 0$ if and only if $\hat{q}_s > q - \bar{\varepsilon}$, it follows from the definition of τ and Proposition 8 that $\dot{\bar{D}}(s) > 0$ for all $s > \tau$. Using this property, it is easy to verify that

$$\dot{\bar{D}}(s) < \frac{\Lambda\bar{\varepsilon}}{4\bar{\alpha}}\left(\frac{\bar{L}(s) - \bar{L}_\tau}{\bar{D}_\tau}\right) \leq \frac{\Lambda\bar{\varepsilon}}{4\bar{\alpha}}\left(\frac{\Lambda(s-\tau)}{\bar{D}_\tau}\right) = O(\bar{\varepsilon}).$$

Setting $\chi(t) = \int_\tau^{\tau+t} \dot{\bar{D}}(s)\ \mathrm{d}s$, we can rewrite $\bar{L}(t), \bar{D}(t)$ as follows

$$\bar{L}(t) = \bar{L}_\tau + \dot{\bar{L}}_\tau t + \chi(t) \quad \text{and} \quad \bar{D}(t) = \bar{D}_0 + \chi(t),$$

which allow us to express the ratio for the number of likes to dislikes as

$$\frac{\bar{L}(t)}{\bar{D}(t)} \;=\; \frac{\bar{L}_\tau + \dot{\bar{L}}_\tau t + \chi(t)}{\bar{D}_0}\left(\frac{\chi(t)}{\bar{D}_0 + \chi(t)}\right) \;=\; \frac{\bar{L}_\tau + \dot{\bar{L}}_\tau t}{\bar{D}_0} - \frac{\chi(t)}{\bar{D}_0}\left(\frac{\bar{L}_\tau + \dot{\bar{L}}_\tau t}{\bar{D}_0} - 1\right) + \xi + \xi',$$

where

$$\xi = \frac{\bar{L}_\tau + \dot{\bar{L}}_\tau t + \chi(t)}{\bar{D}_0}\left(\frac{\chi(t)}{\bar{D}_0} - \frac{\chi(t)}{\bar{D}_0 + \chi(t)}\right) = O(\bar{\varepsilon}) \quad \text{and} \quad \xi' = -\left(\frac{\chi(t)}{\bar{D}_0}\right)^2 = O(\bar{\varepsilon}^2).$$

When $\bar{\varepsilon}$ is small, the ratio of likes to dislikes can be approximated by

$$\frac{\bar{L}(t)}{\bar{D}(t)} = \frac{\bar{L}_\tau}{\bar{D}_0} + \frac{\dot{\bar{L}}_\tau}{\bar{D}_0}t - \frac{\bar{L}_\tau - \bar{D}_0}{\bar{D}_0^2}\chi(t) - \frac{\dot{\bar{L}}_\tau}{\bar{D}_0^2}\chi(t)t.$$

Substituting the above ratio into equation (31), and differentiating twice one obtains a third order, non-homogeneous linear ODE, that is solvable in terms of the matrix exponential (not in closed form due to the non-homogeneous coefficients).

Proofs of Section 4

Lemma 4. *The monopolist's revenue function* (20) *can be written as*

$$\tilde{R}(p) = \Lambda \left[h_0(p) \cdot \pi_0(p) + h_\infty(p) \cdot \pi_\infty(p)\right]$$

where

$$h_0(p) = \frac{\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{1-\delta w} - 1}{1/w - \delta} \quad \text{and} \quad h_\infty(p) = \frac{\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{-\delta w}}{\delta}.$$

Moreover, $\tilde{R}(p)$ *is such that* $|\tilde{R}(p) - \bar{R}(p)| \leq \bar{\varepsilon}[\Lambda(p/\bar{\alpha})h_\infty(p)]$.

Proof. First, recall that when $q_0 < q - \bar{\varepsilon}$ we have

$$\hat{q}_t = p - \bar{\alpha} + (\bar{\alpha} + q_0 - p) \exp\left(\frac{t}{w}\right) \quad \textit{and} \quad \tau = w \log\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right),$$

substituting into the first term in the right-hand side of (20) yields

$$\begin{aligned}
\int_0^\tau e^{-\delta t}\pi_t(p) \, \mathrm{d}t &= \int_0^\tau e^{-\delta t} \frac{(\bar{\alpha} + \hat{q}_t - p)p}{\bar{\alpha}} \, \mathrm{d}t \\
&= \frac{(\bar{\alpha} + q_0 - p)p}{\bar{\alpha}} \cdot \left[\frac{e^{t(1/w-\delta)}}{1/w - \delta}\right]_0^\tau \\
&= \frac{(\bar{\alpha} + q_0 - p)p}{\bar{\alpha}} \cdot \frac{\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{1-\delta w} - 1}{1/w - \delta} \\
&= \pi_0(p) \cdot h_0(p).
\end{aligned}$$

Similarly, the second term in the right-hand side of (20) simplifies to

$$\begin{aligned}
\int_\tau^\infty e^{-\delta t}\pi_\infty(p) \, \mathrm{d}t &= \int_\tau^\infty e^{-\delta t} \frac{(\bar{\alpha} + q - p)p}{\bar{\alpha}} \, \mathrm{d}t \\
&= \frac{(\bar{\alpha} + q - p)p}{\bar{\alpha}} \cdot \left[\frac{e^{-\delta t}}{\delta}\right]_\tau^\infty \\
&= \frac{(\bar{\alpha} + q - p)p}{\bar{\alpha}} \cdot \frac{\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{-\delta w}}{\delta} \\
&= \pi_\infty(p) \cdot h_\infty(p).
\end{aligned}$$

Thus we have established that $\tilde{R}(p) = \Lambda \left[h_0(p) \cdot \pi_0(p) + h_\infty(p) \cdot \pi_\infty(p)\right]$.

To establish the bound, note that $\hat{q}_t \geq q - \bar{\varepsilon}$ for all $t \geq \tau$ implies the following inequality

$$
\begin{aligned}
|\tilde{R}(p) - \bar{R}(p)| &= \Lambda\left(\frac{p}{\bar{\alpha}}\right) \int_{\tau}^{\infty} e^{-\delta t}(q - \hat{q}_t)\ \mathrm{d}t \\
&\leq \bar{\varepsilon} \cdot \Lambda\left(\frac{p}{\bar{\alpha}}\right) \int_{\tau}^{\infty} e^{-\delta t}\ \mathrm{d}t \\
&= \bar{\varepsilon} \cdot \Lambda\left(\frac{p}{\bar{\alpha}}\right) \frac{\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{-\delta w}}{\delta} = \bar{\varepsilon} \cdot \Lambda\left(\frac{p}{\bar{\alpha}}\right) h_{\infty}(p). \qquad \square
\end{aligned}
$$

Proof of Proposition 5. The proof will proceed as follows. First we establish that for $\bar{\epsilon}$ sufficiently small the revenue maximization problem (21) admits a unique optimal solution. Then we establish its properties, proving in order Part (a) and Part (b).

Consider the expression for $\tilde{R}(p)$ given in Lemma 4. Differentiating with respect to p once, we get

$$
\tilde{R}'(p) = \Lambda\left[h_0'(p) \cdot \pi_0(p) + h_0(p) \cdot \pi_0'(p) + h_{\infty}'(p) \cdot \pi_{\infty}(p) + h_{\infty}(p) \cdot \pi_{\infty}'(p)\right].
$$

and, differentiating twice, we get

$$
\begin{aligned}
\tilde{R}''(p) = \Lambda\big[&h_0''(p) \cdot \pi_0(p) + 2h_0'(p) \cdot \pi_0'(p) + h_0(p) \cdot \pi_0''(p) \\
&+ h_{\infty}''(p) \cdot \pi_{\infty}(p) + 2h_{\infty}'(p) \cdot \pi_{\infty}'(p) + h_{\infty}(p) \cdot \pi_{\infty}''\big].
\end{aligned}
$$

Noting that $\pi_0''(p) = \pi_{\infty}''(p) = -2/\bar{\alpha}$, the latter equation simplifies to

$$
\tilde{R}''(p) = \Lambda\left\{\left[h_0''(p) \cdot \pi_0(p) + h_{\infty}''(p) \cdot \pi_{\infty}(p)\right] + 2\left[h_0'(p) \cdot \pi_0'(p) + h_{\infty}'(p) \cdot \pi_{\infty}'(p)\right] - \frac{2}{\bar{\alpha}}\left[h_0(p) + h_{\infty}(p)\right]\right\},
$$

where

$$
\begin{aligned}
h_0(p) + h_{\infty}(p) &= \frac{\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{1-\delta w} - 1}{1/w - \delta} + \frac{\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{-\delta w}}{\delta}, \\
h_0'(p) \cdot \pi_0'(p) + h_{\infty}'(p) \cdot \pi_{\infty}'(p) &= \left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{-1-\delta w} \frac{w(q_0 - q + \bar{\varepsilon})[p(q - q_0) + \bar{\varepsilon}(\bar{\alpha} - 2p + q_0)]}{\bar{\alpha}(\bar{\alpha} + q_0 - p)^3}, \\
h_0''(p) \cdot \pi_0(p) + h_{\infty}''(p) \cdot \pi_{\infty}(p) &= \\
\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right)^{-\delta w} &\frac{w(q_0 - q + \bar{\varepsilon})[(p - q - \bar{\alpha})(q - q_0) + \bar{\varepsilon}(3(\alpha - p + q) - \delta w(q - q_0)) - \bar{\varepsilon}^2(2 - \delta w)]p}{\bar{\alpha}(\bar{\alpha} + q - \bar{\varepsilon} - p)^2(\bar{\alpha} + q_0 - p)^2}.
\end{aligned}
$$

It follows from Assumption 1(i) that the revenue function $\tilde{R}(p) = \tilde{R}(p, \bar{\varepsilon}) \in C^{\infty}(\bar{\varepsilon})$ for all $p \in$

$[0, p_{\max}]$. This can be easily verified by noting that $p_{\max} < \bar{\alpha} - \bar{\varepsilon}$ implies that the quantity

$$\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}$$

in h_0 and h_∞ is always bounded away from 0. The statement above implies that $\tilde{R}''(p, \bar{\varepsilon})$ is a continuous function of $\bar{\varepsilon}$ for all $p \in [p_{\min}, p_{\max}]$, in particular it is continuous at $\bar{\varepsilon} = 0$. We will now prove that, when $\bar{\varepsilon}$ is sufficiently small, the revenue function $\tilde{R}(p)$ is strictly concave for all $p \in [p_{\min}, p_{\max}]$. Evaluating second-derivative $\tilde{R}''(p, \bar{\varepsilon})$ at $\bar{\varepsilon} = 0$ yields

$$\tilde{R}''(p,0) = \Lambda \cdot \frac{2 - \left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}+q_0-p}\right)^{-\delta w}\left(\frac{2}{\delta w} + (q-q_0)\left(\frac{2\bar{\alpha}-2p-q+3q_0-\delta w(q-q_0)}{(\bar{\alpha}+q_0-p)^2} - \frac{\bar{\alpha}+q}{\bar{\alpha}+q-p}\cdot\frac{(q-q_0)(\delta w-1)}{(\bar{\alpha}+q_0-p)^2}\right)\right)}{\bar{\alpha}(1/w-\delta)}. \tag{34}$$

Case 1: $\delta w > 1$. The denominator of the above equation is always negative. Moreover, since

$$\frac{\bar{\alpha}+q}{\bar{\alpha}+q-p} > 1 \quad \text{for all } p \in (0, \bar{\alpha}+q_0)$$

and the numerator of (34) is strictly increasing in this term, replacing $(\bar{\alpha}+q)/(\bar{\alpha}+q-p)$ with 1 we find that the numerator of (34) is strictly bigger than

$$2\left(1 - \left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}+q_0-p}\right)^{-\delta w}\left(\frac{1}{\delta w} + \frac{q-q_0}{\bar{\alpha}+q_0-p}\right)\right) \tag{35}$$

and since $\delta w > 1$, replacing $1/\delta w$ with 1 we find that the above quantity is strictly bigger than

$$2\left(1 - \left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}+q_0-p}\right)^{1-\delta w}\right). \tag{36}$$

Now note that $q > q_0$ and $\delta w > 1$ imply that

$$\left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}+q_0-p}\right)^{1-\delta w} < 1 \quad \text{for all } p < \bar{\alpha}+q_0.$$

This implies that (36) is strictly bigger than 0 and so is the numerator of (34). Thus, we conclude that $\tilde{R}''(p,0) < 0$ for all $p \in (0, \bar{\alpha}+q_0)$.

Case 2: $\delta w < 1$. The denominator in equation (34) is always positive. Moreover, since

$$\frac{\bar{\alpha}+q}{\bar{\alpha}+q-p} > 1 \quad \text{for all } p \in (0, \bar{\alpha}+q_0)$$

and the numerator is strictly decreasing in this term, replacing $(\bar{\alpha}+q)/(\bar{\alpha}+q-p)$ with 1 we find that the numerator of (34) is strictly smaller than the quantity in (35). Moreover, since $\delta w < 1$, replacing $1/\delta w$ with 1 in (35) we find that the quantity in (35) is strictly smaller than the quantity in (36). Now note that $q > q_0$ and $\delta w < 1$ imply that

$$\left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}+q_0-p}\right)^{1-\delta w} > 1 \quad \text{for all } p < \bar{\alpha}+q_0.$$

This implies that (36) is strictly smaller than 0 and so is the numerator of (34). Thus, we conclude that $\tilde{R}''(p,0) < 0$ for all $p \in (0, \bar{\alpha}+q_0)$.

Case 1 and Case 2 establish that $\tilde{R}''(p,0) < 0$ for all $p \in (0, \bar{\alpha}+q_0)$. Moreover, since $\tilde{R}''(p,\bar{\varepsilon})$ is continuous at $\bar{\varepsilon} = 0$, there exists $\varepsilon' > 0$ such that $\tilde{R}''(p,\bar{\varepsilon}) < 0$ for all $\bar{\varepsilon} < \varepsilon'$. In particular, this implies that when $\bar{\varepsilon}$ is sufficiently small, $\tilde{R}(p)$ is strictly concave for all $p \in [p_{\min}, p_{\max}]$, and therefore Problem (21) admits a unique optimal solution.

Proof of Part (a). First, note that, since $\tilde{R}(p)$ is strictly concave, then $\tilde{R}'(p)$ is strictly decreasing for all $p \in [p_{\min}, p_{\max}]$.

Now, note that $\tilde{R}'(p,\bar{\varepsilon})$ is a continuous function of $\bar{\varepsilon}$ at $\bar{\varepsilon} = 0$ for all $p \in [0, p_{\max}]$. This follows directly from $\tilde{R}(p,\bar{\varepsilon}) \in C^{\infty}(\bar{\varepsilon})$ for all $p \in [0, p_{\max}]$, which was established above. We will now prove that $p^* \in \left[p^{\mathrm{m}}(q_0), p^{\mathrm{m}}(q)\right]$. By definition, $p^{\mathrm{m}}(q_0), p^{\mathrm{m}}(q) \in [p_{\min}, p_{\max}]$, and it is easy to verify that that $p^{\mathrm{m}}(q_0) = \max\{p_{\min}, (\bar{\alpha}+q_0)/2\}$. If $p^{\mathrm{m}}(q_0) = p_{\min}$ then clearly $p^* \geq p^{\mathrm{m}}(q_0)$. If $p^{\mathrm{m}}(q_0) = (\bar{\alpha}+q_0)/2$, then evaluating the first-derivative $\tilde{R}'(p,\bar{\varepsilon})$ at $(p^{\mathrm{m}}(q_0), 0)$ yields

$$\tilde{R}'(p^{\mathrm{m}}(q_0), 0) = \Lambda \cdot \frac{q-q_0}{\bar{\alpha}\delta}\left(2\frac{\bar{\alpha}+q}{\bar{\alpha}+q_0} - 1\right)^{-\delta w} > 0$$

since $q > q_0$, thus $p^* > p^{\mathrm{m}}(q_0)$. It follows from Assumption 1(i) that $p^{\mathrm{m}}(q) = (\bar{\alpha}+q)/2$, so evaluating the first-derivative at $(p^{\mathrm{m}}(q), 0)$ we have

$$\tilde{R}'(p^{\mathrm{m}}(q), 0) = \Lambda \cdot (q-q_0)\frac{1-\left(\dfrac{\bar{\alpha}+q}{\bar{\alpha}+2q_0-q}\right)^{1-\delta w}}{\bar{\alpha}(1/w-\delta)}. \tag{37}$$

Note that $q > q_0$ implies that

$$\left(\frac{\bar{\alpha}+q}{\bar{\alpha}+2q_0-q}\right)^{1-\delta w} > 1,$$

which in turn implies that when $\delta w > 1$ the denominator in (37) is always strictly negative and the numerator is always strictly positive. When $\delta w < 1$ the reverse inequalities hold for the denominator and numerator in (37), therefore $\tilde{R}'(p^{\mathrm{m}}(q), 0) < 0$. By continuity of $\tilde{R}'(p,\bar{\varepsilon})$ at $\bar{\varepsilon} = 0$, there exists $\varepsilon'' > 0$ such that $\tilde{R}'(p^{\mathrm{m}}(q_0), \bar{\varepsilon}) > 0$ and $\tilde{R}'(p^{\mathrm{m}}(q), \bar{\varepsilon}) < 0$ for all $\bar{\varepsilon} < \varepsilon''$. Thus proving that, for $\bar{\varepsilon}$ sufficiently small, we have $p^* \in [p^{\mathrm{m}}(q_0), p^{\mathrm{m}}(q)]$.

Proof of Part (b). Setting $\tilde{R}'(p) = 0$, we get the following first-order condition for the monopolist's problem

$$h_0'(p) \cdot \pi_0(p) + h_0(p) \cdot \pi_0'(p) + h_\infty'(p) \cdot \pi_\infty(p) + h_\infty(p) \cdot \pi_\infty'(p) = 0.$$

Dividing both sides of the above equation by $h_0(p)$, and then dividing again by $1 + h_\infty(p)/h_0(p)$, the first-order condition can be rewritten as

$$(1 - \omega(p)) \cdot \pi_0'(p) + \omega(p) \cdot \pi_\infty'(p) + \xi(p) = 0 \tag{38}$$

where

$$\omega(p) = \frac{h_\infty(p)/h_0(p)}{1 + h_\infty(p)/h_0(p)} = \frac{(\bar{\alpha} + q_0 - p)(1 - \delta w)}{\bar{\alpha} + q_0 - p + \delta w \left(q - \bar{\varepsilon} - q_0 - (\bar{\alpha} + q_0 - p) \left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p} \right)^{\delta w} \right)} \tag{39}$$

and

$$\begin{aligned} \xi(p) &= \frac{(h_0'(p)/h_0(p)) \cdot \pi_0(p) + (h_\infty'(p)/h_0(p)) \cdot \pi_\infty(p)}{1 + h_\infty(p)/h_0(p)} \\ &= \bar{\varepsilon} \cdot \frac{p(q_0 - q + \bar{\varepsilon})(1 - \delta w)\delta w}{\bar{\alpha}(\bar{\alpha} + q - \bar{\varepsilon} - p) \left(\bar{\alpha} + q_0 - p + \delta w \left(q - \bar{\varepsilon} - q_0 - (\bar{\alpha} + q_0 - p) \left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p} \right)^{\delta w} \right) \right)}. \end{aligned}$$

Clearly, as $\delta \to 0$ or $w \to 0$ we have $\omega(p) \to 1$ and $\xi(p) \to 0$ for all $p \in (0, \bar{\alpha} + q_0)$, consequently the left-hand side of (38) goes to $\pi_\infty'(p)$. Finally, let $p^* = p^*(\delta, w)$ be the unique solution of (38) and note that by strict concavity $R''(p^*) < 0$, then p^* converges to the solution of $\pi_\infty'(p) = 0$, i.e. $p^*(\delta, w) \to p^{\mathrm{m}}(q)$ as $\delta \to 0$ or $w \to 0$. For the other case, as $\delta \to \infty$ or $w \to \infty$ we have $\omega(p) \to 0$ for all $p \in (0, \bar{\alpha} + q_0)$. This is easy to verify by first dividing numerator and denominator in (39) by δw, next noting that

$$\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p} > 1$$

since $q_0 < q - \bar{\varepsilon}$, and then taking the limit. Moreover, as $\delta \to \infty$ or $w \to \infty$ we have $\xi(p) \to 0$ for all $p \in (0, \bar{\alpha} + q_0)$, this can be verified as follows. Divide numerator and denominator by δw and note that

$$\xi(p) = \bar{\varepsilon} \cdot \frac{A + B \cdot \delta w}{\frac{C}{\delta w} + D \cdot \left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p} \right)^{\delta w}},$$

for the obvious choices of A, B, C and D. It is easy to see that since

$$\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p} > 1,$$

then $\xi(p) \to 0$ as $\delta \to \infty$ or $w \to \infty$. Thus, the left-hand side of (38) goes to $\pi_0'(p)$ and p^* converges to the solution of $\pi_0'(p) = 0$, i.e. $p^*(\delta, w) \to (\bar{\alpha} + q_0)/2$ as $\delta \to \infty$ or $w \to \infty$, if $(\bar{\alpha} + q_0)/2 \geq p_{\min}$. Otherwise, if $p_{\min} > (\bar{\alpha} + q_0)/2$ then $p^*(\delta, w) \to p_{\min}$ as $\delta \to \infty$ or $w \to \infty$. Recalling that $p^{\mathrm{m}}(q) = \max\{p_{\min}, (\bar{\alpha} + q_0)/2\}$ completes the proof. □

Proofs of Section 5

First, we argue that setting $s \leq \tau$ in the monopolist's problem is without loss of generality. We define formally the discounted revenue, for a generic $s > 0$, as

$$\tilde{R}(p_0, p_1, s) =$$
$$\Lambda \left(\int_0^{\min\{s,\tau\}} e^{-\delta t} \pi_t(p_0) \, \mathrm{d}t + \int_{\min\{s,\tau\}}^{\max\{s,\tau\}} e^{-\delta t} [\pi_t(p_1) \mathbf{1}\{s \leq \tau\} + \pi_\infty(p_0) \mathbf{1}\{s > \tau\}] \, \mathrm{d}t + \int_{\max\{s,\tau\}}^{\infty} e^{-\delta t} \pi_\infty(p_1) \, \mathrm{d}t \right).$$

When $s \leq \tau$ the above equation reduces to (22), when $s > \tau$ the revenue function is given by

$$\tilde{R}(p_0, p_1, s) = \Lambda \left(\int_0^{\tau} e^{-\delta t} \pi_t(p_0) \, \mathrm{d}t + \int_{\tau}^{s} e^{-\delta t} \pi_\infty(p_0) \, \mathrm{d}t + \int_{s}^{\infty} e^{-\delta t} \pi_\infty(p_1) \, \mathrm{d}t \right).$$

Suppose that $\tilde{R}(p_0^*, p_1^*, s^*)$ is optimal and $s^* > \tau$, clearly it must be that $p_1^* = \underset{p_1 \in [p_{\min}, p_{\max}]}{\operatorname{argmax}} \{\pi_\infty(p_1)\}$. But this implies that $\tilde{R}(p_0^*, p_1^*, \tau) \geq \tilde{R}(p_0^*, p_1^*, s^*)$, thus it is without loss of generality to consider only policies such that $s \leq \tau$.

Before proving the proposition, we introduce the following definition

$$\tau_k := \inf\{t : t \geq 0, |q - \hat{q}_t| \leq \bar{\varepsilon} \mid q_k, p_k\}, \quad k = 0, 1,$$

where τ_k denotes the time that the prevailing quality estimate reaches within $\bar{\varepsilon}$ from q, starting from a prior q_k and a price p_k. This is analogous to the simpler definition of $\tau = \inf\{t : t \geq 0, |q - \hat{q}_t| \leq \bar{\varepsilon}\}$ that was introduced previously, and it simplifies the exposition of the following proofs.

The following lemma is needed for the proof of Proposition 6.

Lemma 5. *The monopolist's revenue function* (22) *can be written as*

$$\tilde{R}(p_0, p_1, s) = \Lambda \left[h_0(s) \cdot \pi_0(p_0) + e^{-\delta s} \left[h_s(p_1) \cdot \pi_s(p_1) + h_\infty(p_1) \cdot \pi_\infty(p_1) \right] \right],$$

where

$$h_0(s) = \frac{e^{s(1/w-\delta)} - 1}{1/w - \delta}, \quad h_s(p_1) = \frac{\left(\dfrac{\bar{\alpha} + q - \bar{\varepsilon} - p_1}{\bar{\alpha} + q_1 - p_1}\right)^{1-\delta w_1} - 1}{1/w_1 - \delta}, \quad h_\infty(p_1) = \frac{\left(\dfrac{\bar{\alpha} + q - \bar{\varepsilon} - p_1}{\bar{\alpha} + q_1 - p_1}\right)^{-\delta w_1}}{\delta}.$$

Proof. (Sketch only.) First, note that $q_0 < q - \bar{\varepsilon}$ and that $s \leq \tau$ implies that $q_1 \leq q - \bar{\varepsilon}$. Thus, for all $t \leq \tau$ the quality estimate is given by

$$\hat{q}_t = \begin{cases} p_0 - \bar{\alpha} + (\bar{\alpha} + q_0 - p_0) \exp\left(\frac{t}{w}\right) & \text{if} \quad t < s \\ p_1 - \bar{\alpha} + (\bar{\alpha} + q_1 - p_1) \exp\left(\frac{t-s}{w_1}\right) & \text{if} \quad t \geq s \end{cases}.$$

Following the argument of the proof of Lemma 4, we establish the desired result. □

Proof of Proposition 6. First we establish that $p_1^* \in \left[p^{\mathrm{m}}(q_1), p^{\mathrm{m}}(q)\right]$. Differentiating $\tilde{R}$ twice with respect to p_1 and then evaluating at $\bar{\varepsilon} = 0$ yields

$$\left.\frac{\partial^2}{\partial p_1^2} \tilde{R}(p_0, p_1, s)\right|_{\bar{\varepsilon}=0} = \Lambda \cdot e^{-s\delta}.$$

$$\frac{2 - \left(\dfrac{\bar{\alpha} + q - p_1}{\bar{\alpha} + q_1 - p_1}\right)^{-\delta w_1} \left(\dfrac{2}{\delta w_1} + (q - q_1)\left(\dfrac{2\bar{\alpha} - 2p_1 - q + 3q_1 - \delta w_1(q - q_1)}{(\bar{\alpha} + q_1 - p_1)^2} - \dfrac{\bar{\alpha} + q}{\bar{\alpha} + q - p_1} \cdot \dfrac{(q - q_1)(\delta w_1 - 1)}{(\bar{\alpha} + q_1 - p_1)^2}\right)\right)}{\bar{\alpha}(1/w_1 - \delta)}. \tag{40}$$

The following two cases establish that the equation above is always negative.

Case 1: $\delta w > 1$. The denominator of the equation (40) is always negative. Since

$$\frac{\bar{\alpha} + q}{\bar{\alpha} + q - p_1} > 1 \quad \text{for all } p_1 \in (0, \bar{\alpha} + q_1)$$

and the numerator of (40) is strictly increasing in this term, replacing $(\bar{\alpha} + q)/(\bar{\alpha} + q - p_1)$ with 1 we find that the numerator of (40) is strictly bigger than

$$2\left(1 - \left(\frac{\bar{\alpha} + q - p_1}{\bar{\alpha} + q_1 - p_1}\right)^{-\delta w_1} \left(\frac{1}{\delta w_1} + \frac{q - q_1}{\bar{\alpha} + q_1 - p_1}\right)\right) \tag{41}$$

and since $\delta w_1 > 1$, replacing $1/\delta w_1$ with 1 we find that the above quantity is strictly bigger than

$$2\left(1 - \left(\frac{\bar{\alpha} + q - p_1}{\bar{\alpha} + q_1 - p_1}\right)^{1-\delta w_1}\right). \tag{42}$$

Now note that $q > q_1$ and $\delta w_1 > 1$ imply that

$$\left(\frac{\bar{\alpha}+q-p_1}{\bar{\alpha}+q_1-p_1}\right)^{1-\delta w_1} < 1 \quad \text{for all } p_1 < \bar{\alpha}+q_1.$$

This implies that (42) is strictly bigger than 0 and so is the numerator of (40). Thus, we conclude that (40) is strictly smaller than 0 for all $p_1 \in (0, \bar{\alpha}+q_1)$.

Case 2: $\delta w < 1$. The denominator in equation (40) is always positive. Since

$$\frac{\bar{\alpha}+q}{\bar{\alpha}+q-p_1} > 1 \quad \text{for all } p_1 \in (0, \bar{\alpha}+q_1)$$

and the numerator of (40) is strictly decreasing in this term, replacing $(\bar{\alpha}+q)/(\bar{\alpha}+q-p_1)$ with 1 we find that the numerator of (40) is strictly smaller than the quantity in (41). Moreover, since $\delta w_1 < 1$, replacing $1/\delta w_1$ with 1 in (41) we find that the quantity in (41) is strictly smaller than the quantity in (42). Now note that $q > q_1$ and $\delta w_1 > 1$ imply that

$$\left(\frac{\bar{\alpha}+q-p_1}{\bar{\alpha}+q_1-p_1}\right)^{1-\delta w_1} > 1 \quad \text{for all } p_1 < \bar{\alpha}+q_1.$$

This implies that (42) is strictly smaller than 0 and so is the numerator of (40). Thus, we conclude that (40) is strictly less than 0 for all $p_1 \in (0, \bar{\alpha}+q_1)$.

Case 1 and Case 2 prove that

$$\left.\frac{\partial^2}{\partial p_1^2}\tilde{R}(p_0, p_1, s)\right|_{\bar{\varepsilon}=0} < 0 \quad \text{for all feasible } s, p_0 \text{ and } p_1 \in (0, \bar{\alpha}+q_1).$$

Moreover, since $\frac{\partial^2}{\partial p_1^2}\tilde{R}(p_0, p_1, s)$ as a function of $\bar{\varepsilon}$, is continuous at $\bar{\varepsilon} = 0$ for all feasible s, p_0 and $p_1 \in [0, p_{\max}]$, it follows that, when $\bar{\varepsilon}$ is sufficiently small, $\tilde{R}(p_0, p_1, s)$ is strictly concave in p_1 for all feasible s, p_0 and $p_1 \in [p_{\min}, p_{\max}]$.

We next show that $p_1^* \in [p^{\mathrm{m}}(q_1), p^{\mathrm{m}}(q)]$. By definition, $p^{\mathrm{m}}(q_1), p^{\mathrm{m}}(q) \in [p_{\min}, p_{\max}]$, and it is easy to verify that that $p^{\mathrm{m}}(q_1) = \max\{p_{\min}, (\bar{\alpha}+q_1)/2\}$. If $p^{\mathrm{m}}(q_1) = p_{\min}$ then clearly $p^* \geq p^{\mathrm{m}}(q_1)$. If $p^{\mathrm{m}}(q_1) = (\bar{\alpha}+q_1)/2$, then evaluating the first-derivative at $p_1 = p^{\mathrm{m}}(q_1)$ and $\bar{\varepsilon} = 0$ yields

$$\left.\frac{\partial}{\partial p_1}\tilde{R}(p_0, p^{\mathrm{m}}(q_1), s)\right|_{\bar{\varepsilon}=0} = \Lambda \cdot e^{-\delta s} \cdot \frac{q-q_1}{\bar{\alpha}\delta}\left(2\frac{\bar{\alpha}+q}{\bar{\alpha}+q_1} - 1\right)^{-\delta w_1} > 0$$

since $q > q_1$. It follows from Assumption 1(i) that $p^{\mathrm{m}}(q) = (\bar{\alpha}+q)/2$, so evaluating the first-derivative at $p_1 = p^{\mathrm{m}}(q)$ and $\bar{\varepsilon} = 0$ we have

$$\frac{\partial}{\partial p_1}\tilde{R}(p_0, p^{\mathrm{m}}(q), s)\bigg|_{\bar{\varepsilon}=0} = \Lambda \cdot e^{-\delta s} \cdot (q - q_1)\frac{1 - \left(\dfrac{\bar{\alpha} + q}{\bar{\alpha} + 2q_1 - q}\right)^{1-\delta w_1}}{\bar{\alpha}(1/w_1 - \delta)}. \tag{43}$$

Note that $q > q_1$ implies that

$$\left(\frac{\bar{\alpha} + q}{\bar{\alpha} + 2q_1 - q}\right)^{1-\delta w_1} > 1,$$

which implies that when $\delta w_1 > 1$ the denominator in (43) is always strictly negative and the numerator is always strictly positive. When $\delta w_1 < 1$ the reverse inequalities hold for the denominator and numerator in (43). Therefore

$$\frac{\partial}{\partial p_1}\tilde{R}(p_0, p^{\mathrm{m}}(q), s)\bigg|_{\bar{\varepsilon}=0} < 0.$$

By continuity at $\bar{\varepsilon} = 0$ it follows that

$$\frac{\partial}{\partial p_1}\tilde{R}(p_0, p^{\mathrm{m}}(q_1), s) > 0 \quad \text{and} \frac{\partial}{\partial p_1}\tilde{R}(p_0, p^{\mathrm{m}}(q), s) < 0,$$

when $\bar{\varepsilon}$ is sufficiently small.

Finally, to establish that $p_0^* \leq p_1^*$ we first obtain an equivalent problem (see Boyd and Vandenberghe (2004, Chapter 4, Section 4.1.3) for the formal definition of equivalent optimization problems) by making a change of variable. Noting that $s \leq \tau$ if and only if $s \leq \tau_0$, we replace the last constraint in the monopolist's optimization problem and set $\phi = s/\tau_0$, equivalently $s = \phi\tau_0$. We let the monopolist choose $\phi \in [0, 1]$ instead of s. It is clear that the optimal solution (p_0^*, p_1^*, s^*) of the original problem, can be readily obtained from the optimal solution (p_0^*, p_1^*, ϕ^*) of the transformed problem and vice versa, thus the two problems are equivalent. Making the change of variable in the monopolist's objective (22) yields

$$\tilde{R}(p_0, p_1, \phi) = \Lambda\left(\int_0^{\phi\tau_0} e^{-\delta t}\pi_t(p_0)\ \mathrm{d}t + \int_{\phi\tau_0}^{\phi\tau_0+\tau_1} e^{-\delta t}\pi_t(p_1)\ \mathrm{d}t + \int_{\phi\tau_0+\tau_1}^{\infty} e^{-\delta t}\pi_\infty(p_1)\ \mathrm{d}t\right),$$

and the associated monopolist's problem is

$$\begin{aligned} \max \quad & \tilde{R}(p_0, p_1, \phi) \\ \text{s.t.} \quad & p_0, p_1 \in [p_{\min}, p_{\max}] \\ & \phi \in [0, 1]. \end{aligned}$$

Let (p_0^*, p_1^*, ϕ^*) be the optimal solution to the above problem and suppose, by contradiction, that

$p_0^* > p_1^*$. We will now construct a solution (p_0', p_1', ϕ') that strictly dominates (p_0^*, p_1^*, ϕ^*). Let $p_0' = p_1^* < p_0^*$ and note that $\tau_0(p_0') < \tau_0(p_0^*)$ since

$$\frac{\partial \tau_0}{\partial p_0} = \frac{w(q - \bar{\varepsilon} - q_0)}{(\bar{\alpha} + q - \bar{\varepsilon} - p_0)(\bar{\alpha} + q_0 - p_0)} > 0 \quad \text{for all } p_0 \in [p_{\min}, p_{\max}].$$

Moreover, one can immediately verify that

$$q_1(p_0, \phi) = p_0 - \bar{\alpha} + (\bar{\alpha} + q_0 - p_0)\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p_0}{\bar{\alpha} + q_0 - p_0}\right)^{\phi}$$

is decreasing in p_0 for all $\phi \in [0, 1]$, thus $q_1(p_0^*, \phi^*) \leq q_1(p_0', \phi^*)$. Set $\phi' \leq \phi^*$ to be such that $q_1(p_0^*, \phi^*) = q_1(p_0', \phi')$ and note that ϕ' is always feasible since ϕ^* is feasible. Finally, set $p_1' = \text{argmax}_{p_1}\{\tilde{R}(p_0', p_1, \phi')\}$. Now, consider the revenue function evaluated at the new solution and note that

$$\int_0^{\phi' \tau_0(p_0')} e^{-\delta t} \pi_t(p_0') \, \mathrm{d}t > \int_0^{\phi' \tau_0(p_0')} e^{-\delta t} \pi_t(p_0^*) \, \mathrm{d}t,$$

since by construction $\hat{q}_t \leq q_1(p_0^*, \phi^*)$ and $\pi_t(p_0)$ is a strictly concave function of p_0 which is maximized at $p_0 = (\bar{\alpha} + \hat{q}_t)/2 \leq p_0' < p_0^*$. Finally, note that our choice of p_1' implies that the continuation value of the policy (p_0', p_1', ϕ') after $\phi' \tau_0(p_1')$ is not smaller than the continuation value of the policy (p_0^*, p_1^*, ϕ^*). Thus, we reach the contradiction $\tilde{R}(p_0', p_1', \phi') > \tilde{R}(p_0^*, p_1^*, \phi^*)$. It must be $p_0^* \leq p_1^*$. $\square$