Monopoly Pricing in the Presence of Social Learning

Bar Ifrach
Stanford University
Management Science and Engineering

Costis Maglaras
Columbia University
Columbia Business School

Marco Scarsini
LUISS
Dipartimento di Economia e Finanza
Monopoly Pricing in the Presence of Social Learning

Bar Ifrach*  Costis Maglaras†  Marco Scarsini‡

August 14, 2012

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 and, under some conditions, 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: learning, information aggregation, bounded rationality, pricing, 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. This paper studies a monopolist’s pricing decision in a market where quality estimates are evolving according to such a learning process.

*Management Science and Engineering, Stanford University (bifrach@stanford.edu). I thank 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)
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 quality preference that, together with the product quality, determines their willingness-to-pay. These quality preference 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 quality preference parameters 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’ estimates about it evolve according to a social learning mechanism. Consumers who purchase the product report whether they “liked” or “disliked” it, if their ex-post utility was positive or negative, respectively. Consumers do not report their quality preference, 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. Under this bounded rationality assumption, decisions follow a cutoff rule, and a quality estimate can be extracted. Subsequently, new information gets released and the prevailing quality estimate evolves dynamically over time.

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

---

\(^1\)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.
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 quality preference purchase the product. In Section 3 we show that if consumers ignore the self-selection bias, then they may not learn—even asymptotically—the true product quality. Learning will eventually occur almost surely if consumers correct for this bias. Detailed understanding of the learning trajectory is essential in optimizing the tradeoff between learning and the monopolist’s discounted revenue objective.

Second, we derive 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 2 shows that the asymptotic learning trajectory is characterized by a system of differential equations, and Proposition 6 derives its closed form solution. 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. Learning is fast if initially consumers overestimate the true quality of the product. It is much slower, due to the self-selection bias, when initially they underestimate the true quality. The solution of the mean-field model gives a crisp characterization of the dependence of the learning trajectory on the 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. Proposition 7 characterizes the effect of the price on the speed and transient of the social learning process; it depends on the generalized failure rate

\(^2\)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.
of the preference distribution, a concept that has been studied by Lariviere (2006). Taking these
dynamics into account, the seller’s objective is to maximize her discounted revenues. 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. Proposition 8 shows that the optimal static price in the presence of
social learning is between these two extreme points, naturally approaching the “full information”
price when the learning is fast.

Lastly, we study 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 9 shows that, under general assumptions,
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 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. This paper lies in the intersection of
some strands of work in economics, revenue management, engineering, and computer science. The
economic literature focuses on social learning and herding behavior, the revenue management papers
use dynamic models to study tactical price optimization questions, the articles in engineering and
computer science deal with decentralized learning, sensor networks, consensus propagation, and
message passing, as well as pricing and advertisement optimization.

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 2, which is key in proving learning, is
similar in nature to that of Goeree et al. (2006).\textsuperscript{3}

\textsuperscript{3}See the surveys by Bikhchandani, Hirshleifer, and Welch (1998), and, more recently, by Acemoglu and Ozdaglar
Several papers have considered variations of the observational learning model with imperfect information. Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) and Acemoglu, Bimpikis, and Ozdaglar (2010) 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. (2010) consider agents who can delay their decision in order to obtain information from others by utilizing their social network. Herrera and Hörner (2009) 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). A similar approach could be used to relax this assumption in our model.

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.

The area of revenue management focuses, in part, on tactical problems of price optimization. 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 (2011) for many extensions to this model.

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

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 (2009, 2012) 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 (2011) consider optimal pricing strategies of a monopolist selling a product to consumers who are embedded in a social network and experience externalities in consumption. 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 9 is consistent with their findings.

---

\(^5\)In particular, by considering a population with finite mass, and by simplifying consumers' decisions.
Model

2.1 The Monopolist’s Pricing Problem

A sequence of consumers, indexed by $i = 1, 2, \ldots$, sequentially decide between purchasing a newly launched good or service (henceforth, the product), or choosing an outside alternative. The quality of the product, $q$, is initially unknown and can take values in the interval $(q_{\text{min}}, q_{\text{max}})$ with $q_{\text{min}} > 0$. Consumers are heterogeneous; this is represented by a parameter $\alpha_i$ that determines consumer $i$’s willingness to pay for quality. His utility from consuming the product is

$$u_i = \alpha_i q - p,$$

where $q$ is the true quality of the product, and $p$ is the price charged by the monopolist. 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 $f$ with connected support either $[0, \bar{a}]$ or $[0, \infty)$. The preference parameter can be interpreted as a premium that a consumer is willing to pay per unit of quality. Heterogeneity in terms of the $\alpha_i$’s implies that even if the product quality, $q$, were known, not all consumers would make the same decision: only those with $\alpha_i \geq \alpha^* := p/q$ would purchase the product. Equivalently, only consumers with WtP $\alpha q$ greater or equal to $p$ would purchase; the distribution of $\alpha$ 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 $\Lambda$, 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 on the quality of the product, $q_0 \in (q_{\text{min}}, q_{\text{max}})$. 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. Consumers report their purchasing decisions and, if they decide to purchase the product, they also truthfully report a review about their experience with the product that takes two values: ‘like’ and ‘dislike’. A consumer who purchases reports that he likes the product if his ex-post utility

---

$^6$The functional form of the utility function does not play a big role in the analysis.
was nonnegative, taking into account the true quality of the product and his preference parameter; he reports he dislikes if his ex-post utility was negative. This binary report is a simplification of the star rating scales that are ubiquitous in online review systems. In addition, a consumer who chooses the outside alternative reports so. We denote consumer $i$'s review by $r_i$ which can take the values $r^l$, $r^d$, or $r^o$ if he purchased and liked the product, purchased and disliked the product, or chose the outside alternative, respectively. Consumers do not report their preference parameter, and as such reviews are not fully informative. For example, a ‘like’ could result from either a high preference parameter or from the product being of high quality (not necessarily both)\(^7\).

We define the following quantities: $L(i) := \sum_{j=1}^{i} 1\{r_j = r^l\}$ is the number of consumers who purchased and liked the product out of the first $i$ consumers, and, similarly, $D(i)$ and $O(i)$ are the number of consumers who purchased and disliked the product and the number of consumers who chose the outside alternative out of the first $i$ consumers, respectively. We denote $l(i) := L(i)/i$, $d(i) := D(i)/i$, and $o(i) := O(i)/i$ the corresponding fractions of consumers who ‘liked’, ‘disliked’ and ‘did not purchase’ up to consumer $i$. The information available to consumer $i$ before making his decision is $I_i = (L(i-1), D(i-1), O(i-1))$, where $O(i-1)$ can be omitted if consumer $i$ knows that he is the $(i)$-th consumer. 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 maximize her discounted expected revenue, $\pi(p)$, as follows,

$$\max_p \pi(p) = \max_p \mathbb{E}\left[\sum_{i=1}^{\infty} e^{-\delta t_i} p I\{r_i(p) \neq r^o\}\right] = \max_p \sum_{i=1}^{\infty} \mathbb{E}\left[e^{-\delta t_i} p P(r_i(p) \neq r^o | I_i)\right], \quad (1)$$

where $\delta > 0$ is the monopolist’s discount factor, and the expectation is with respect to consumers’ arrival times and quality preferences. Here the monopolist selects a static price, and it is assumed that she knows the true quality, the prior quality estimate, and the distribution of quality preferences. Section 5 considers a problem where the seller can select two prices as well as the optimal time to switch between them. Expression (1) 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 dynamics of the social learning process must be understood in order to solve for the optimal price.

\(^7\)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.
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, followed by a discussion of several key points. It is composed of two parts: (a) consumer $i$ 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 nonpositive $\alpha \hat{q}(i) - p \geq 0$.

Consumers form their quality estimate following an intuitive bounded rationality mechanism. First, while the sequence of reviews $(r_1, r_2, \ldots)$ carries information about each reviewer’s potential quality estimate and the evolution of that estimate over time, we assume that consumers either do not observe this sequence or do not process that information in their quality estimation procedure due to its inherent intractability in most practical settings. Second, we postulate that each consumer assumes that all his predecessors made decisions based on the same information, i.e., based on the same quality estimate $\hat{q}$. This means that the predecessors with the highest quality preferences bought and liked the product ($\alpha \geq \max(\alpha^*, p/\hat{q})$), those with the lowest quality preferences chose the outside alternative ($\alpha < p/\hat{q}$), and those with quality preferences in the middle bought and disliked the product ($p/q = \alpha^* > \alpha \geq p/\hat{q}$, if $\hat{q} > q$). This is illustrated on a distribution of quality preferences in Figure 1, i.e., a mass $l(i)$ of consumers is placed in the right tail, and mass $o(i)$ of consumers is placed in the left tail, and the reminder, a mass $d(i)$, is placed in the middle.

Following this logic, the quality preference value $\hat{\alpha}(i) := \hat{F}^{-1}(l(i))$ (see Figure 1) is the quality preference parameter of the marginal consumer, such that all consumers with $\alpha \geq \hat{\alpha}(i)$ purchased and ‘liked’ the product, and all consumers with $\alpha < \hat{\alpha}(i)$ either ‘disliked’ or did not purchase the product. This marginal parameter gives rise to a quality estimate $\hat{q}(i)$ such that $\hat{\alpha}(i) \hat{q}(i) = p$, which implies that $\hat{q}(i) = p/\hat{F}^{-1}(l(i))$. The observed information and the associated decision rule are subject to censoring or self-selection bias in a setting where $\hat{\alpha}(i)$ is computed without having observed any ‘dislikes’, i.e., when the middle region in Figure 1 is empty. In that setting, the consumer can infer that all predecessors with $\alpha \geq \hat{\alpha}(i)$ ‘liked’ the product, but cannot conclude that consumers with lower $\alpha$ values would have ‘disliked’ it. Recognizing that fact, consumers will naively account for this bias via a correction term that will slightly decrease $\hat{\alpha}(i)$ to reflect that more consumers may have ‘liked’ the product had they purchased it; this will lead to some experimentation from consumers close to the marginal value $\hat{\alpha}(i)$ when the observation is censored.
Moreover, consumers will also place some weight on their prior quality estimate. This weight will diminish as reviews accumulate.

In more detail, assuming that the prior quality estimate \( q_0 \) was correct, consumers would expect a fraction of \( l_0 := \bar{F}(p/q_0) \) consumers to ‘like’ the product and a fraction \( d_0 := 0 \) of consumers to ‘dislike’ it. Consumers observe \( l(i) \) and \( d(i) \) and form weighted estimates of these fractions,

\[
\begin{bmatrix}
\hat{l}_w(i) \\
\hat{d}_w(i)
\end{bmatrix} := \frac{w}{w + i} \begin{bmatrix}
l_0 \\
d_0
\end{bmatrix} + \frac{i}{w + i} \begin{bmatrix}
l(i) \\
d(i)
\end{bmatrix},
\]

where \( w > 0 \) is the weight assigned to the prior quality estimate, whose effect diminishes as the information available by predecessors grows\(^8\). As mentioned above, the consumer will adjust the above observation when \( d_w(i) \) is small. Mathematically we can capture this by either perturbing the critical fractile \( \hat{\alpha}(i) \) that corresponds to \( l_w(i) \) by some small amount, or by perturbing \( l_w(i) \) itself. We adopt the latter for purposes of analytic tractability, and set:

\[
\hat{q}(i + 1) := \text{proj}_{[q_{\min}, q_{\max}]} \left( \frac{p}{\bar{F}^{-1}(l_w(i) + \psi(i))} \right),
\]

where the projection to \( [q_{\min}, q_{\max}] \) ensures that it is consistent with the known range of \( q \) (the operator \( \text{proj}_{[x_{\min}, x_{\max}]}(x) \) projects \( x \) to the interval \( [x_{\min}, x_{\max}] \)), and \( \psi(i) := \psi(l_w(i), d_w(i)) \) is a correction term that satisfies the following assumption.

**Assumption 1.** The function \( \psi : \mathbb{R}_+^2 \to \mathbb{R}, \psi(l, d) \) is nonincreasing in \( d \), nondecreasing in \( l \), and Lipschitz continuous in \( (l, d) \). In addition, \( \psi(l, 0) > 0 \) for all \( l \in [0, 1] \).

The monotonicity condition is natural given that the correction term should encourage experimentation when there are few ‘dislikes’ and should not discourage it as the number of ‘likes’ increases. We allow for the possibility that the correction term takes negative values, this could be the case if, e.g., consumers reduce their quality estimate when observing many ‘dislikes’.

After estimating its quality, \( \hat{q}(i + 1) \), consumer \( i + 1 \) purchases the product if his estimated utility from purchasing is non-negative, i.e., \( \alpha_{i+1}\hat{q}(i + 1) - p \geq 0 \).

---

\(^8\)This procedure is reminiscent of the linear credibility estimators used in actuarial science (see, e.g., Bühlmann and Gisler, 2005) and of Bayesian updating with a conjugate prior, that induces a linear structure of estimators (see Diaconis and Ylvisaker, 1979).
Definition 1. The decision rule for consumer $i + 1$ can be described as follows:

1. Use the observed information $I_{i+1} = (L(i), D(i), O(i))$ to form an estimate of the quality of the product according to (3).

2. Purchase the product if the estimated utility is non-negative, $\alpha_{i+1} \hat{q}(i + 1) - p \geq 0$.

Our decision rule satisfies the following useful characterization. Consider the probability that consumer $i + 1$ will like the product given $(l(i), d(i))$

$$P(r_{i+1} = r^d | (l(i), d(i))) = P(\alpha_{i+1} \hat{q}(i + 1) - p \geq 0, \alpha_{i+1} q - p \geq 0)$$

$$= P(\alpha_{i+1} \geq \max (p/q, p/\hat{q}(i + 1)))$$

$$= \min (\hat{F}(p/q), \hat{F}(p/\hat{q}(i + 1)))$$

$$= \min (l^*, \hat{l}(i)), \quad (4)$$

where $\hat{l}(i) := \hat{F}(p/\hat{q}(i + 1)) = \text{proj}_{[l_{\min}, l_{\max}]}(l_w(i) + \psi(i))$ is the probability that consumer $i + 1$ will purchase the product, and $l_{\max} := \hat{F}(p/q_{\max}), l_{\min} := \hat{F}(p/q_{\min})$. The projection to $[l_{\min}, l_{\max}]$ ensures that $\hat{q}(i + 1) \in [q_{\min}, q_{\max}]$. Let $(x)^+ = \max(0, x)$. The probability that $i + 1$ will dislike is

$$P(r_{i+1} = r^d | (l(i), d(i))) = P(\alpha_{i+1} \hat{q}(i + 1) - p \geq 0, \alpha_{i+1} q - p < 0) = (\hat{l}(i) - l^*)^+. \quad (5)$$

Equations (4)-(5) are simple and intuitive algebraic expressions that could be of use in the analysis of (1) as long as one has a good analytical handle for $\hat{l}(i)$ and its evolution as a function of $i$ (not just $i \to \infty$). This will be the emphasis of the next section.

We conclude this section with a few brief remarks on some aspects of the model.

**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.

**Information.** We assume that consumers make decisions based on aggregate information about the fraction of predecessors who ‘liked’, ‘disliked’ and did not purchase the product. In most practical settings, such as online retailers and review aggregators like Amazon and TripAdvisor, the latter is not available. Herrera and Hörner (2009) relaxed this assumption in their work, and a similar approach could be applied to our model.

**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 Learning Trajectory

This section establishes the asymptotic properties and the transient evolution of the social learning process. The first is of interest 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 second is necessary for the analysis of the monopolist’s pricing problem (1), as well as for the evaluation of other operational considerations that are not discussed in this paper, such as capacity planning, the timing of marketing strategies and others.

Direct characterization of the learning transient for the underlying stochastic learning process is intractable, as in almost all social learning models in the literature, both Bayesian and non-Bayesian. We will proceed to first derive an approximation for the transient behavior of the learning dynamics in the form of a tractable system of ordinary differential equations. This 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. Second, we establish sufficient conditions under which the market asymptotically learns the true quality of the product in a sense to be made precise later on. To build intuition, we begin by characterizing the limiting behavior of the mean-field model and show that it depends on whether or not consumers correct for the self-selection bias. We then go back to the stochastic model of §2 and show that the same conditions imply asymptotic learning there.

3.1 Large Market Setting

We consider a sequence of systems indexed by $n$. In the $n$-th system consumers’ arrival process is Poisson with rate $\Lambda^n := n\bar{\Lambda}$. The state variables of the $n$-th system at time $t$ is given by $X^n(t) := (L^n(t), D^n(t), O^n(t))$, where $L^n(t)$ is the number of consumers who reported like by time

\(^9\text{See Acemoglu, Dahleh, Lobel, and Ozdaglar (2009) for a characterization of the rate of convergence of Bayesian social learning for some social networks.}\)
\( t \) in the \( n \)-th system, and \( D^n(t) \) and \( O^n(t) \) are 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 \( \bar{L}^n(t), \bar{D}^n(t), \) and \( \bar{O}^n(t) \). This state variable comprises the information available to the first consumer arriving after time \( t \). To keep things consistent with the underlying system, we will also scale the weight placed on the prior quality estimate in the decision rule, which will be denoted by \( \bar{w} := n\bar{w} \). Given the acceleration of the consumer arrival rate, this scaling essentially guarantees that the weight assigned to the prior corresponds to a fictitious number of consumers that flows through the system over a fixed time window, e.g., a week.\(^{10}\)

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 (2) that

\[
l^n_w(X^n) := \frac{w^n}{w^n + S^n} l_0 + \frac{S^n}{w^n + S^n} \frac{L^n}{w + S^n} = \frac{\bar{w}l_0 + \bar{L}^n}{\bar{w} + S^n} =: l_w(\bar{X}^n),
\]

where \( S^n := L^n + D^n + O^n \) and \( \bar{S}^n := S^n/n \). Similarly, \( d^n_w(X^n) := (\bar{w}d_0 + \bar{D}^n)/(\bar{w} + \bar{S}^n) =: d_w(\bar{X}^n) \) and \( \bar{l}^n(X^n) := \text{proj}_{[t_{\min}, t_{\max}]}(l_w(\bar{X}^n) + \psi(l_w(\bar{X}^n), d_w(\bar{X}^n))) := \bar{l}(\bar{X}^n) \). Hence, we define the functions \( \gamma^1, \gamma^4, \) and \( \gamma^6 \) such that \( \gamma^k(\bar{X}^n) := P \left( r_i = r^k | I_i = X^n \right) \), with the interpretation that \( \gamma^k(\bar{X}^n) \) is the probability that a consumer who observes information \( X^n \) reports \( k \in \{1, 4, 6\} \), for ‘like’, ‘dislike’, and ‘outside alternative’. Using the above expressions we get that

\[
\gamma^1(\bar{X}^n) = \min \left( l^*, \bar{l}(\bar{X}^n) \right), \quad \gamma^4(\bar{X}^n) = \left( \bar{l}(\bar{X}^n) - l^* \right)^+ \quad \text{and} \quad \gamma^6(\bar{X}^n) = 1 - \gamma^1(\bar{X}^n) - \gamma^4(\bar{X}^n). \tag{7}
\]

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 \( N := (N^1, N^4, N^6) \) be a vector of independent Poisson processes with rate 1. Then,

\[
\bar{L}^n(t) = \frac{1}{n} N^1 \left( \Lambda^n \int_0^t \gamma^1(\bar{X}^n(s))ds \right),
\]

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

\[
\bar{X}^n(t) = \frac{1}{n} N \left( \Lambda^n \int_0^t \gamma(\bar{X}^n(s))ds \right), \tag{8}
\]

where \( \gamma := (\gamma^1, \gamma^4, \gamma^6) \) was defined in (7).

If the rate processes inside the expressions (8) did not depend on the state \( \bar{X}^n(t) \) itself, then a straight-forward application of the functional strong law of large numbers for the Poisson process

\(^{10}\)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.
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 $\alpha$, 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 1 that derives a deterministic limiting characterization for the system behavior as $n$ grows large using Kurtz (1977/78, Theorem 2.2).

**Proposition 1.** For every $t > 0$, 
\[
\lim_{n \to \infty} \sup_{s \leq t} |\bar{X}^n(s) - \bar{X}(s)| = 0 \quad \text{a.s.,}
\]
where $\bar{X}(t) = (\bar{L}(t), \bar{D}(t), \bar{O}(t))$ is deterministic and satisfies the integral equation,
\[
\bar{X}(t) = \bar{\Lambda} \int_0^t \gamma(\bar{X}(s))ds.
\] (9)

To better understand (9) consider the expression for the scaled number of likes,
\[
\bar{L}(t) = \bar{\Lambda} \int_0^t \gamma^l(\bar{X}(s))ds = \bar{\Lambda} \int_0^t \mathbb{P}(r_s = r^l|I_s = \bar{X}(s))ds.
\] (10)

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 $\bar{S}(t) := \bar{L}(t) + \bar{D}(t) + \bar{O}(t) = \bar{\Lambda} t$ since $\gamma^l(\bar{X}(t)) + \gamma^d(\bar{X}(t)) + \gamma^o(\bar{X}(t)) = 1$ for all $t$. It is convenient to derive from (9) the expressions for $(\bar{l}(t), \bar{d}(t))$, the counterparts of the stochastic $l_w(i)$ and $d_w(i)$ in the limiting (fluid) model, since these quantities determine the decision of an arriving consumer. Namely, from (6) we have
\[
\bar{l}(t) := l_w(\bar{X}(t)) = \frac{wd_0 + \bar{L}(t)}{\bar{w} + t\bar{\Lambda}} \quad \text{and} \quad \bar{d}(t) := d_w(\bar{X}(t)) = \frac{\bar{D}(t)}{\bar{w} + t\bar{\Lambda}},
\] (11)
where the expression for $\bar{d}(t)$ follows for $d_0 = 0$.

### 3.2 Learning Dynamics and Asymptotic Learning

The next proposition derives the ODEs governing the dynamics of $(\bar{l}, \bar{d})$ defined in (11). From (11), note 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 \textit{regular}.

**Proposition 2.** At regular points \( t \), \((\bar{l}, \bar{d})\) satisfies the differential equation:

\[
\begin{bmatrix}
\dot{\bar{l}}(t) \\
\dot{\bar{d}}(t)
\end{bmatrix} = \frac{\Lambda}{\bar{w} + t\Lambda} \left[ \min(l^*, \bar{l}(t)) - \bar{l}(t) \right],
\]

with \((\bar{l}(0), \bar{d}(0)) = (l_0, 0)\), where

\[
\bar{l}(t) := \text{proj}_{[\bar{l}_{\text{min}}, \bar{l}_{\text{max}}]}(\bar{l}(t) + \psi(\bar{l}(t), \bar{d}(t)))
\]

is the probability that a consumer would purchase the product at time \( t \).

These ODEs are intuitive. The rate of consumers who purchase at time \( t \) is \( \Lambda \bar{l}(t) \), of which \( \Lambda \min(l^*, \bar{l}(t)) \) would like the product. Subtracting out \( \bar{l}(t) \) and normalizing by the mass that arrived by that time, \( \bar{w} + t\Lambda \), gives the desired expression for \( \dot{\bar{l}}(t) \). A similar argument follows for dislikes.

Next we build on Proposition 2 to study asymptotic learning. Focusing on (12), it is convenient to think about asymptotic learning by means of its outcome. Intuitively, if learning occurs, eventually consumers with preference parameters greater or equal to \( \alpha^* \), and only them, will purchase the product. As a result, the fraction of likes will converge, in some appropriate sense, to \( l^* \), the fraction of dislikes will converge to 0, and the reminder will choose the outside alternative. The following assumption ensures that the price is such that there will always be some consumers who choose to buy the product even at the lowest possible quality level.

**Assumption 2** (Price is not prohibitive). The price charged by the monopolist is not greater than \( p_{\text{max}} \), and at \( p_{\text{max}} \), \( l_{\text{min}} = \bar{F}(p_{\text{max}}/q_{\text{min}}) > 0 \).

This assumption is clearly satisfied when the support of \( \alpha \) is unbounded. This is similar to the unbounded belief assumption often used in Bayesian social learning. Both imply that new information will enter the system, which is the main requirement for learning to take place.

We consider the asymptotic behavior of system (12) in two cases. First, when the self-selection bias is not accounted for and the experimentation function \( \psi \) is such that \( \psi(l, d) = 0 \) for all \((l, d)\). Second, when consumers correct for the self selection bias using a function \( \psi \) that satisfies Assumption 1. For the latter it is convenient to define, for a given price and quality of the product, the quantity \( d^* := \min(x^*, l_{\text{max}} - l^*) \), where \( x^* \) is the unique solution of \( \psi(l^*, x^*) = x^* \).

**Proposition 3.** Consider system (12)-(13) under any price satisfying Assumption 2.

1. If \( \psi(l, d) = 0 \) for all \((l, d)\), then \( \lim_{t \to \infty}(\bar{l}(t), \bar{d}(t)) = (\min(l_0, l^*), 0) \).
2. If Assumption 1 holds, then \( \lim_{t \to \infty} (\bar{l}(t), \bar{d}(t)) = (l^*, d^*) \). In addition, for any \( \epsilon > 0 \) there exists a \( \psi \) function satisfying Assumption 1 for which \( d^* < \epsilon \).

We say that consumers overestimate the quality if \( q_0 > q \) and underestimate the quality if \( q_0 < q \). Focusing on the corresponding fraction of ‘likes’, a high prior would imply that initially too many consumers purchase the product because \( l_0 > l^* \), while a low prior would correspond to too few people purchasing initially \( l_0 < l^* \).

Without experimentation the market never learns the correct quality of the product in the sense that the fraction of people purchasing and liking the product does not converge to \( l^* \) when \( l_0 < l^* \). In the underestimating case and under the assumption that \( \psi = 0 \), all consumers who buy the product like it, but not enough buy. Moreover, the number of ‘dislikes’ is zero, which can be used as an indication for the self-selection bias. In the overestimating case too many buy initially and there is no bias in the quality estimate. Therefore, in this case learning is achieved in the fluid model in the sense that \( \lim_{t \to \infty} (\bar{l}(t), \bar{d}(t)) = (l^*, 0) \).

Intuitively, when consumers correct for the self-selection bias, there is positive drift in \( \bar{l} \) if the fraction of dislikes is small, thus pushing \( \lim \inf_{t \to \infty} \bar{l}(t) \geq l^* \). In addition, \( \lim \sup_{t \to \infty} \bar{l}(t) \leq l^* \) because this is the maximal fraction of consumers that can like the product. The small fraction of dislikes \( d^* \) is the result of the correction of the self selection bias. When \( \bar{l} \) is at \( l^* \) and the number of dislikes is low, consumers who experiment end up disliking the product, leading to an increasing fraction of dislikes and discouraging any further experimentation. The limiting fraction \( d^* \) is the one that balances these two opposing drift terms and can be made arbitrarily small. One possible extension, not pursued in this paper in the interest of space, would be to allow for a non-stationary correction term \( \psi(l, d; t) \) that decays at an appropriate rate such that learning occurs but the number of dislikes in the limit is vanishingly small.

Finally, we complete this subsection by stating without a proof an analogous result to Proposition 3 for the stochastic model described in the previous section; details can be found in Ifrach (2012, Subsection 2.3.2). Our results rely on the so called ODE method for analyzing the asymptotic behavior of stochastic approximation schemes, as outlined in Kushner and Yin (2003), and makes specific use of their Theorem 2.3.

**Proposition 4.** Consider the model described in §2 and the policy specified in Definition 1 under Assumption 2.

1. Suppose that \( \psi(l, d) = 0 \) for all \( (l, d) \). Then, for any \( \epsilon \) small enough, there exists a \( \zeta(\epsilon) > 0 \) such that \( P(|l(i) - l^*| > \epsilon) > \zeta(\epsilon) \) for all \( i \geq 1 \). Consequently, asymptotic learning fails in the stochastic model.

2. Suppose \( \psi \) satisfies Assumption 1 and, moreover, that \( \psi(l, d) = l \phi(d/l) \) for some function \( \phi \).
Then, asymptotic learning is achieved in the sense that \((l(i), d(i)) \to (l^*, d^*)\) almost surely as \(i \to \infty\). In addition, for any \(\epsilon > 0\) there exists a function \(\psi\) such that \(d^* < \epsilon\).

Under Assumption 1, Propositions 3 and 4 establish that consumers acting under the specified decision rule eventually learn the quality of the product and make the right decision, except for a sliver of experimenters in the limit. In turn, the seller’s pricing decision cannot affect whether consumers will eventually learn, but rather she can affect the speed of the learning trajectory.

### 3.3 Learning Trajectory

The learning trajectory depends on whether consumers initially overestimate or underestimate the quality of the product as given by the next proposition.

**Proposition 5.** Suppose Assumptions 1 and 2 hold. Then, system (12) evolves as follows:

(i) If \(q_0 < q\) (underestimate), then \(\bar{l}\) is increasing and \(\bar{d}\) is nondecreasing;

(ii) if \(q_0 = q\), then \(\bar{l}(t)\) is constant;

(iii) if \(q_0 > q\) (overestimate), then there exist times \(t_l, t_d > 0\) (possibly infinite) such that \(\bar{l}\) is decreasing for \(t \in [0, t_l]\), and \(\bar{d}\) is increasing for \(t \in [0, t_d]\). If \(\psi \geq 0\) then \(t_l = \infty\).

Broadly speaking, in the underestimating case too few buy initially. The fraction of likes gradually increases as consumers experiment and like the product. In the overestimating case too many buy at first and the fraction of likes decreases over time.

For a general nonlinear \(\psi\) function the system of ODEs in (12) can be solved numerically. If \(\psi\) is linear, (12) simplifies to a system of (piecewise) non-homogeneous first order linear ODEs that can now be solved analytically. To avoid boundary cases, we impose conditions for which the projection on \([l_{min}, l_{max}]\) is never binding.

**Assumption 3.**

1. The function that controls the correction term is linear, \(\psi(l, d) = \psi_0 l - \psi_1 d\) with \(\psi_0, \psi_1 \geq 0\).

2. The price charged by the monopolist is no less than \(p_{min}\), where \(\max(\bar{F}(p_{min}/q_0)(1+\psi_0), \bar{F}(p_{min}/q)(1+x^*)) \leq l_{max}\).

The next proposition derives the learning trajectory that will be the basis of the monopolist’s pricing problem.

**Proposition 6.** Suppose Assumptions 1-3 hold. The solution of system (12) follows these cases.
1. If $l_0(1 + \psi_0) < l^*$ let

$$T := \left( \frac{\bar{w}}{\Lambda} \left( \frac{1}{1 + \psi_0 \frac{l^*}{l_0}} \right)^{1/\psi_0} - 1 \right)^+.$$  \hspace{1cm} (14)

Then for $t \leq T$

$$\begin{bmatrix} \bar{l}(t) \\ \bar{d}(t) \end{bmatrix} = l_0 \left( \frac{\bar{w} + t \Lambda}{\bar{w}} \right)^{\psi_0} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$  \hspace{1cm} (15)

and for $t > T$

$$\begin{bmatrix} \bar{l}(t) \\ \bar{d}(t) \end{bmatrix} = l^* \begin{bmatrix} 1 \\ x^* \end{bmatrix} - l^* \frac{\bar{w} + T \Lambda \psi_0}{\bar{w}} \begin{bmatrix} \tilde{\psi} \\ 1 \end{bmatrix} + l^* x^* \left( \frac{\bar{w} + T \Lambda}{\bar{w}} \right)^{1 + \psi_1} \begin{bmatrix} 0 \\ 1/\psi_1 \end{bmatrix},$$  \hspace{1cm} (16)

where $\tilde{\psi} := \psi_1/(1 + \psi_0)$.

2. If $l_0(1 + \psi_0) \geq l^*$, then

$$\begin{bmatrix} \bar{l}(t) \\ \bar{d}(t) \end{bmatrix} = l^* \begin{bmatrix} 1 \\ x^* \end{bmatrix} + (l_0 - l^*) \frac{\bar{w}}{\bar{w} + t \Lambda} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \left( \frac{\bar{w}}{\bar{w} + t \Lambda} \right)^{1 + \psi_1} \begin{bmatrix} 0 \\ 1/\psi_1 \end{bmatrix} \begin{bmatrix} l^*(1 + x^*) - l_0(1 + \psi_0) \end{bmatrix}.$$  \hspace{1cm} (17)

The cases above almost coincide with the underestimating and overestimating cases. When $l_0(1 + \psi_0) < l^*$ consumers underestimate the quality, while when consumers overestimate the quality it is always the case that $l_0(1 + \psi_0) > l^*$. Therefore, with abuse of terminology we refer to these cases as the underestimating and overestimating cases, respectively\(^{11}\). The speed of learning depends on the aggressiveness with which consumers correct the self-selection bias. This is captured by the parameter $\psi_0$ since $\psi(l, 0) = \psi_0 l$. We suppose that consumers are careful in applying this correction in order to avoid overshooting. Hence it is assumed that $\psi_0$ is close to zero and so learning is slow in the underestimating case. In this case time $T$ is the first time that everyone who should buy the product indeed buys it, and this leads to a kink in the learning trajectory. In the overestimating case $l_0(1 + \psi_0) \geq l^*$, and social learning is quite fast, as $\bar{l}$ drops to $l^*$ at a rate of $w/(w + t)$.

The difference in the rates of convergence between these cases is intuitive. The monopolist cannot fool consumers for long that the quality of the product is high since the disappointed consumers will complain. Yet, if too few buy the product initially and consumers are slow to

\(^{11}\)If $q_0$ is close to $q$ from below, or if the price is very low, then it can be that $l_0(1 + \psi_0) \geq l^*$ while $q_0 < q$. This case is of little interest to us since social learning is almost achieved from the very beginning.
Figure 2: The blue lines correspond to \( l_\omega(t) \) (asymptote to \( l^* \)) and the red lines to \( d_\omega(t) \) (asymptote to \( x^*l^* \), close to zero). Model parameters: \( \alpha_i \sim U[0,1] \), \( q = 0.4 \), \( q_0 \in \{0.325, 0.5\} \), \( p = 0.24 \), \( \Lambda = 1000 \), \( \bar{w} = 30 \), \( \psi_0 = 0.05 \), and \( \psi_1 = 1 \). Note the different scales of the horizontal axes. Scaling parameters: \( n = 1 \) in overestimating and \( n = 10 \) in underestimating (both with 100 sample paths).

correct for the self-selection bias, learning may be slow. Subsection 4.1 formalizes the notion of the speed of learning and discusses the effect of price on this speed.

Before moving to the pricing problem, we illustrate the fit of the mean-field approximation by comparing it to simulated sample paths of the stochastic model. Figure 2b shows a typical case where the prior quality underestimates the true quality. As such, \( \bar{l}(t) \) converges to \( l^* \) from below, and the weighted fraction of dislikes starts from zero and converges to \( x^*l^* \). We can see that the mean of 100 sample paths essentially coincides with the mean-field approximation, and that all sample paths lie in a band around this deterministic approximation. Figure 2a shows a typical overestimating case in which \( \bar{l}(t) \) converges to \( l^* \) from above. The weighted fraction of dislikes peaks early since too many consumers are tempted to purchase the product due to the high prior estimate. Here too the mean of the sample paths is very close to the mean-field approximation, and all sample paths lie in a band around this deterministic approximation. These simulations illustrate that learning is much faster in the overestimating case, where the weighted fractions are close to their limits after less than one time unit, versus about 50 in the underestimating case.

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.

4 Static Price Analysis

The ODEs specified in (12)-(13) and further simplified in (14)-(17) provides a simple characterization of the evolution of the market conditions, which is endogenously determined via the learning mechanism and the seller’s price. This simplifies the evolution of (4)-(5) and allows us to approximate (1) via a simpler and tractable problem. Specifically, this section characterizes the effect of the price on the speed of learning, and subsequently solves the monopolist’s problem of choosing a static price to maximize her profits as given in (1). Following the analysis of the previous section, the stochastic learning trajectory is replaced by its deterministic mean field approximation. This enables to solve an otherwise intractable problem. For the remainder of the paper we will assume that Assumptions 1 - 3 hold. Notice that the analysis would go through, with some modifications, even without Assumption 1, that is, when asymptotic learning is not achieved.

4.1 Price and Speed of Learning

For a given price, define the \( \epsilon \)-time-to-learn to be \( T(\epsilon) := \min\{t \mid \bar{l}(s)/l^*-1 \leq \epsilon \text{ for all } s \geq t\} \). Recall that \( \bar{l}(t) \) is the cumulative fraction of consumers who like the product at time \( t \) (weighted by the prior quality estimate), and \( l^* \) is the fraction of likes once learning is achieved. Note that the first time to see dislikes in the underestimating case is \( T = T(\psi_0/(1 + \psi_0)) \), since at \( T \) we have \( \check{q}(T) = q \), and substituting for \( \check{q} \) and \( \psi \) we see that \( \bar{l}(T)(1 + \psi_0) = l^* \). Denote the generalized failure rate (henceforth GFR) by \( G(x) := xf(x)/\bar{F}(x) \), where \( f(x)/\bar{F}(x) \) is the failure rate. A distribution is increasing (decreasing) GFR if \( G(x) \) is increasing (decreasing) in \( x \) (henceforth IGFR and DGFR, respectively). The next proposition determines how the \( \epsilon \)-time-to-learn changes with the price.

**Proposition 7.** The \( \epsilon \)-time-to-learn is increasing (decreasing) in the price if \( F \) is IGFR (DGFR).

This proposition suggests that the effect of price on learning depends on the tail properties of the distribution of quality preferences; IGFR distributions have faster decaying tails than Pareto distribution and DGFR have slower decaying tails than Pareto distribution. Consider the underestimating case, where \( T(\epsilon) \) depends on the price via the ratio \( l^*/l_0 \). In the IGFR case, i.e., where the tail of the quality preference distribution decays fast, as price increases, \( l_0 \) decreases faster than \( l^* \) because it is further out in the tail of the distribution. As a result, \( l^*/l_0 \) and subsequently the \( \epsilon \)-time-to-learn increases in the price. The tail condition arises because of the way learning is achieved through experimentation of the marginal consumers whose preferences are close to the critical value \( \check{\alpha} \) (cf. Figure 1) from right to left. The speed of learning will therefore depend on the
tail behavior of the market, which is related to the degree of heterogeneity. The opposite argument holds true in the DGFR case, and analogous explanations hold for the overestimating case.\(^{12}\)

Denote the instantaneous revenue function by \(R(p, q) := p\bar{F}(p/q)\). This is the expected revenue extracted from a single consumer who believes the quality of the product is \(q\). In our setting IGFR (together with \(\lim_{x\to\infty} G(x) > 1\)) is a sufficient condition for \(R\) to be unimodal in price. Conversely, if \(F\) is DGFR, then it is always beneficial to increase the price, and an optimal price does not exist. Furthermore, if we think of \(\bar{F}(p/q)\) as the demand function, then \(G(p)\) is the price elasticity of demand (see Lariviere (2006) for more details). As such, IFGR is a standard assumption in the field of revenue management\(^{13}\), which in our case entails that the \(\epsilon\)-time-to-learn increases with the price. This condition plays a big role in the two-price problem of Section 5.

Suppose that we are interested only in the effect of the price on learning, and ignore its effect on revenue accumulation. Then the monopolist would like to speed up learning when consumers initially underestimate the quality of the product and to slow it down learning otherwise. It follows from Proposition 7 that this can be achieved by lowering the price in the underestimating case and increasing it in the overestimating case, as is intuitive. When \(F\) is DGFR, which is less common, this would be achieved by moving the price in the opposite direction.

4.2 Optimal Static Price

Adapting by the mean-field approximation, the seller’s discounted revenue in (1) simplifies to

\[
\pi(p) = p\bar{\Lambda} \int_0^\infty e^{-\delta t} \bar{F}(p/\hat{q}(t))dt = p\bar{\Lambda} \int_0^\infty e^{-\delta t} \left[(1 + \psi_0)\bar{l}(t) - \psi_1\bar{d}(t)\right]dt, \tag{18}
\]

where \((\bar{l}, \bar{d})\) evolve according to (14)-(17), which themselves depend on the price, discounting is in continuous time, and the second equality follows by substituting for \(\hat{q}\) and \(\psi\). The seller’s revenue optimization problem (1) reduces to \(\max_{p \in [p_{\min}, p_{\max}]} \pi(p)\), and we denote by \(\mathcal{P}^*\) the set of optimal solutions with typical element \(p^*\). The following assumption is common in the literature.

**Assumption 4.** The instantaneous revenue function, \(R(p, q)\), is unimodal in \(p\) for all \(q\).

With this assumption we can define \(p(q) := \arg \max_{p \in \mathbb{R}_+} R(p, q)\), which equates the marginal instantaneous revenue function (w.r.t. price), \(R'(p, q)\), to zero. Consider the first order conditions

\(^{12}\)Alternatively, one could instead study the time after which all newly arriving consumers make decisions that are \(\epsilon\)-away from those obtained when learning has been achieved, e.g., when the probability of buying satisfies \((1 - \epsilon)\bar{l}^* \leq \bar{l}(s) \leq (1 + \epsilon)(\bar{l}^* + \bar{d})\), i.e., where the right fraction of consumers are purchasing, most are satisfied, and only a few, if any, are dissatisfied by their decision. This criterion will be achieved sooner than \(\bar{T}(\epsilon)\) as it only considers the decisions of newly arriving consumers that act on the most up-to-date information, in contrast to the above definition of \(\bar{T}(\epsilon)\) that focuses on the average review among all consumers, including the early buyers who may have been acting under incorrect information. In any case, similar structural properties can be obtained under the new criterion where once again the time-to-learn would depend on whether market heterogeneity is IGFR or DGFR.

\(^{13}\)It is satisfied by many distributions including exponential, uniform, normal, and lognormal.
of the unconstrained optimization problem, i.e., ignoring the price interval. For the first case in Proposition 6 (typically underestimating) the first order condition is

$$\frac{\partial \bar{\pi}(p)}{\partial p} = R'(p, q_0)s_1 + R'(p, q)s_2 - [G(p/q_0) - G(p/q)]s_3 = 0,$$

and for the second case there (typically overestimating) it is

$$\frac{\partial \bar{\pi}(p)}{\partial p} = R'(p, q_0)h_1 + R'(p, q)h_2 = 0,$$

where the constants \(\{s_j\}_{j=1}^3\) and \(\{h_j\}_{j=1}^2\) are positive and depend on the learning trajectory and on the price through equations (14)-(17) (see the proof of Proposition 8 for details). These first order conditions have the nice property that they are a mixture of the marginal instantaneous revenue functions under the prior estimate and under the true quality. In addition, (19) includes a third term whose sign depends on whether the price speeds up or slows down learning. As \(\{s_j\}_{j=1}^3\) and \(\{h_j\}_{j=1}^2\) are non-analytical integrals, obtaining a closed form for \(p^*\) is not possible, even if the marginal instantaneous revenue function, \(R'\), is invertible. It is, however, interesting to compare \(p^*\) to the optimal price that would be charged if the true quality were known upfront, and to the optimal price that would be charged if consumers did not engage in social learning and consequently used only the prior estimate to make decisions.

**Proposition 8.** Suppose Assumption 4 holds and that the price interval is such that \(p(q), p(q_0) \in [p_{\min}, p_{\max}]\). In the underestimating case \(p(q_0) \leq p^* \leq p(q)\) for all \(p^* \in \mathcal{P}^*\), and in the overestimating case \(p(q_0) \geq p^* \geq p(q)\) for all \(p^* \in \mathcal{P}^*\).

Note that given the earlier result regarding the effect of price on the speed of learning, one could envision that, e.g., in the underestimating IGFR case the seller may choose to price below \(p(q_0)\) to speed up learning as this may accelerate learning in a way that would overcompensate for the lower price charged. The above result shows that this is never optimal when a static price is selected by the seller. In Section 5 we will show that when the monopolist has the freedom to choose more than one price, she may choose a price that lies outside this interval to accelerate learning.

The numerical experiments of Subsection 5.2 provide some insight on how the position of \(p^*\) in the interval \([p(q_0), p(q)]\) trades off learning and revenue extraction. For example, a patient seller would price closer to \(p(q)\) than an impatient seller, since she would put a higher weight on the future where the quality estimate is closer to its limit.
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 case where the monopolist can adjust her price once. She charges an initial price $p_0$ up to time $\tau$, a price $p_1$ thereafter, and she is free to choose $\{p_0, p_1, \tau\}$ to maximize her discounted revenue objective. The first price now has a dual role, in addition to extracting short term revenue, it controls the speed of the learning process. Therefore, the monopolist may sacrifice short-term revenue to optimally affect the learning process in the desired direction.

Reviews depend on the price charged at the time of purchase, and as such we need to specify how the available information and the corresponding decision rule will evolve in a setting with dynamic pricing, and specifically before and after the price change. We adopt the simplest assumption that would be consistent with the preceding analysis, where at the time of the price change consumers aggregate all the information conveyed by past reviews into a new prior quality estimate, which we denote by $q_1$. From then on consumers consider only new reviews that were made post the price change, taking into account the new prior quality estimate $q_1$. In more detail, for $t \in [0, \tau)$ the social learning process starts with the prior quality estimate $q_0$ and price $p_0$ and evolves as described in Proposition 6. We denote the associate state variables by $(\tilde{l}_0(t), \tilde{d}_0(t))$. At $t = \tau$, the past information is summarized into a new prior quality estimate, $q_1$, that is equal to the prevailing quality estimate given all review available up to time $\tau$ under the price $p_0$. Specifically,

$$q_1 := p_0 / \tilde{F}^{-1}(\tilde{l}_0(\tau)),$$

and consumers will use as weight $w + \Lambda \tau$ assigned to the new prior estimate in the update routine (2). The learning processes for $t \geq \tau$ are denoted by $(\tilde{l}_1(t), \tilde{d}_1(t))$. Their transient evolution is captured by the ODEs (12) and the initial condition $(\tilde{l}_1(\tau), \tilde{d}_1(\tau)) = (l_1, d_1)$ with $l_1 := \tilde{F}(p_1/q_1)$.
and $d_1 = 0$. The associated discounted revenue objective is

$$
\bar{\pi}(p_0, p_1, \tau) = p_0 \bar{\Lambda} \int_0^\tau e^{-\delta t} [(1 + \psi_0) \bar{l}_0(t) - \psi_1 \bar{d}_0(t)] \, dt + p_1 \bar{\Lambda} \int_{\tau}^{\infty} e^{-\delta t} [(1 + \psi_0) \bar{l}_1(t) - \psi_1 \bar{d}_1(t)] \, dt.
$$

(21)

Applying Proposition 6, we obtain a deterministic learning trajectory that can be analyzed to gain insight on optimal pricing policies. Following a similar construction, multiple price changes can be considered as well. We demonstrate the above by studying the two price problem of the underestimating case, where learning is slow and adjusting the price can be impactful.

We restrict attention to policies under which ‘dislikes’ first appear after the price change; this pins down the appropriate learning trajectory in a regime that is most natural in our setting. Three regions comprise the learning trajectory in this case: the learning phase before $\tau$, the learning phase after $\tau$ and finally the post-learning phase. Denoting by $\bar{\tau}$ the first time ‘dislikes’ would appear under $p_0$, we require that $\tau \leq \bar{\tau}$. This is a natural extension to the static price case, since if we take $p_0 = p_1 = p$ the revenue and the learning process would be identical to the ones in the static price case with price $p$. After the price change at time $\tau$, the downstream pricing problem is identical to the static one studied in the previous section starting with prior quality $q_1 = p_0/\bar{F}^{-1}(\bar{l}_0(\tau)) < q$.

The optimal choice of $p_0$ balances between the short run revenue and its effect on the social learning process; the latter is summarized by $q_1$, the estimated quality at the time of the price change. Increasing the first price expedites learning if $q_1$ is increasing in $p_0$, and it delays learning if the converse holds. This effect is summarized in the next lemma.

**Lemma 1.** The estimated quality at the time of the price change, $q_1$, is decreasing (increasing) in $p_0$ if $F$ is IGFR (DGFR).

This result complements Proposition 7 that shows the relationship between monotonicity of the GFR and the effect of price on the speed of learning.

**Proposition 9.** The discounted revenue objective of the regular underestimating case is

$$
\bar{\pi}(p_0, p_1, \tau) = p_0 \bar{\Lambda} \int_0^\tau e^{-\delta t} [(1 + \psi_0) \left(\frac{w + t}{w}\right) \psi_0] \, dt + p_1 \bar{\Lambda} \int_{\tau}^{\infty} e^{-\delta t} \left[1 + x^* - x^* \left(\frac{w + T}{w + t}\right) \frac{1 + \psi_1}{1 + \psi_0}\right] \, dt,
$$

(22)

where $l_1^* := \bar{F}(p_1/q)$. Assume $F$ is IGFR, and that $p(q_0), p(q) \in [p_{\text{min}}, p_{\text{max}}]$. Let $p_0^*, p_1^*$, and $\tau^*$ be the optimal controls of the problem

$$
\max_{p_0, p_1 \in [p_{\text{min}}, p_{\text{max}}]} \bar{\pi}(p_0, p_1, \tau),
$$

\[ \tau \leq \bar{\tau} \]
then $p_0^* < p_1^*$ and $p_1^* \in [p_1(q_1), p(q_1)]$.

This result shows that an increasing price path is optimal when consumers underestimate the quality of the product. By setting a lower price initially, the monopolist expedites social learning, and consequently increases her revenue in later periods. The benefit from expediting social learning can be so high that the monopolist would initially price below the optimal price corresponding to the prior estimate, as will be shown in the next subsection.

5.2 Comparison of Pricing Policies

We numerically test the pricing policies of Sections 4 and 5. We consider the underestimating case of Figure 2b (expect that here $\bar{w} = 250$) with a demand rate of 1000 potential consumers per week. The two most important parameters in the pricing problem are the monopolist’s discount factor and the error in consumers’ prior estimate relative to the true quality. Considerations pertaining to social learning are more central to the pricing decision in settings where the duration of the learning process is long when compared to the time scale of the monopolist’s discounting.

Three different prior estimates, $q_0 \in \{.35, .3, .25\}$, of the true quality $q = .4$ are considered, for which learning is fast, medium, and slow, respectively. The monopolist is either patient, semi-patient, or impatient, corresponding to annual discount rates of 5%, 10% and 25%. Nine different cases in the intersection of these parameters are tested. For each one of them the optimal static price, and the optimal two prices and switching time are computed. We also report $T$ and $\bar{\tau}$ that are measures of the length of the learning phase in each of these cases. In addition, the revenues of different pricing policies are compared.

Table 1 shows that the optimal static price is decreasing in the patience of the monopolist, and in the time it takes consumers to learn (or increases in $q_0$). This is to be expected given the discussion above. The optimal static price is always higher than the optimal first price and lower

<table>
<thead>
<tr>
<th></th>
<th>static price</th>
<th>two prices</th>
<th>% increase in profits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p^*/p(q)$</td>
<td>$p_0^*/p(q)$</td>
<td>$p_1^*/p(q)$</td>
</tr>
<tr>
<td><strong>patient</strong></td>
<td>$q_0 = .35$</td>
<td>0.98</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td>$q_0 = .30$</td>
<td>0.93</td>
<td>81.07</td>
</tr>
<tr>
<td></td>
<td>$q_0 = .25$</td>
<td>0.73</td>
<td>396.73</td>
</tr>
<tr>
<td><strong>semi-patient</strong></td>
<td>$q_0 = .35$</td>
<td>0.98</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td>$q_0 = .30$</td>
<td>0.93</td>
<td>81.07</td>
</tr>
<tr>
<td></td>
<td>$q_0 = .25$</td>
<td>0.70</td>
<td>230.62</td>
</tr>
<tr>
<td><strong>im-patient</strong></td>
<td>$q_0 = .35$</td>
<td>0.97</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td>$q_0 = .30$</td>
<td>0.88</td>
<td>37.84</td>
</tr>
<tr>
<td></td>
<td>$q_0 = .25$</td>
<td>0.68</td>
<td>138.65</td>
</tr>
</tbody>
</table>

Table 1: Comparison of pricing policies and revenues.
than the second in the two prices case. The low first price allows the monopolist to speed up the learning process considerably in a very short time; the optimal switching time ranges from half a week to less than two months. This allows the monopolists to choose a price that is close to \( p(q) \) in the second stage to maximize revenues when consumers learn the true quality.

We report the percentage increase in revenues under the optimal static price against charging \( p(q) \) and \( p(q_0) \), see \( \text{static}_\text{true} \) and \( \text{static}_\text{prior} \) in the table, respectively. In the patient case and when the prior estimate is close to the true quality, the static price increases revenues for the monopolist by a few percentage points over the two extreme pricing choices. If, however, learning takes longer because the initial prior is further away from the true quality, the improvement over \( p(q) \) is at least 50%, even for the patient monopolist. This sharp increase is due to the nonlinearity of the time-to-learn with respect to the prior estimate. In addition, in the cases where learning is slow, the two-period pricing policy performs at least 10% better than the optimized static price, see \( \text{two}_\text{static} \) in the table. In other cases the improvement is smaller, as one would expect depending on the mis-estimation error of the prior as well as discount factor of the seller.

Lastly, we compare the performance of the two-price policy against the full information case, a scenario when consumers know the true quality, and it is optimal to charge \( p(q) \), see \( \text{full}_\text{two} \) in the table. This provides an upper bound on the total discounted revenues that can be collected in this setting under any pricing policy, and the relative gap we report captures the cost to the monopolist of the fact that consumers initially underestimate the product quality. By accounting for this underestimation and consumer learning in designing her pricing policy, the monopolist is able to recover almost the entire gap, even with just two prices. Indeed, the two-prices policy captures almost all the revenue achievable under the full information case, with a gap of 3% or less. Figure 3 illustrates the increase in revenues from using optimized pricing that accounts for social learning over the naive approach of pricing as if consumers knew the true quality of the product from \( t = 0 \).

The optimal two-price policy almost achieves the upper bound for all values of \( q_0 \).
References


A Appendix

Throughout this appendix, given a vector \( x = (x^1, \ldots, x^k) \) we define \( |x| = \|x\|_1 = \sum_{j=1}^k |x^j| \), and given a function \( x(t) \) of time, \( \dot{x}(t) \) denotes its derivative. To reduce the notational burden we rescale time such that \( \Lambda = \bar{\Lambda} = 1 \). The next lemma is used in the proof of Proposition 1.

Lemma 2. (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 \( x_1, x_2, y_1, y_2 \geq 0 \) and \( z > 0 \) we have

\[
\left| \frac{x_1}{z + x_1 + y_1} - \frac{x_2}{z + x_2 + y_2} \right| \leq \frac{1}{z} (|x_1 - x_2| + |y_1 - y_2|).
\]

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 \(-x, -y, \text{and} -z\).

(b) From the triangular inequality,

\[
\left| \frac{x_1}{z + x_1 + y_1} - \frac{x_2}{z + x_2 + y_2} \right| \leq \left| \frac{x_1}{z + x_1 + y_1} - \frac{x_1}{z + x_1 + y_2} \right| + \left| \frac{x_1}{z + x_1 + y_2} - \frac{x_2}{z + x_2 + y_2} \right|
= \left| \frac{x_1}{z + x_1 + y_1} - \frac{y_1 - y_2}{z + x_1 + y_2} \right| + \left| \frac{z + y_2}{z + x_1 + y_2} - \frac{x_2}{z + x_2 + y_2} \right|
\leq \frac{1}{z} (|x_1 - x_2| + |y_1 - y_2|).
\]

Proof of Proposition 1. We verify the conditions of Theorem 2.2 of Kurtz (1977/78). First we note that \( X^n(t) \in \mathbb{Z}_+^3 \), \( \bar{X}^n(t) = X^n(t)/n \in \{k/n | k \in \mathbb{Z}_+^3 \} \) as required. To satisfy the conditions of the theorem we validate the construction (9) and then show that the following inequalities hold

\[
\gamma^n(x) \leq \Gamma_1 (1 + |x|), \quad |\gamma^n(x) - \gamma(x)| \leq \frac{\Gamma_2}{n} (1 + |x|), \quad \text{and} \quad |\gamma(x) - \gamma(y)| \leq \Gamma_3 |x - y| \quad (23)
\]

for all \( x, y \in \mathbb{R}^3 \) for some finite constants \( \Gamma_1, \Gamma_2, \text{and} \Gamma_3 \).

The integral form of \( \bar{X}^n(t) \) in (9) 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,

\[
\bar{L}^n(t) = \int_0^t \mathbf{1}\{r_s = r^1|\bar{X}^n(s)\} dA^n(s)/n = N^1 \left( \int_0^t P(r_s = r^1|\bar{X}^n(s)) ds \right)/n = N^1 \left( \int_0^t \gamma^1(\bar{X}^n(s)) ds \right)/n,
\]
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 \( D^n(t) \) and \( O^n(t) \). The first inequality in (23) holds for \( \Gamma_1 = 1 \) since \( \gamma^k \) are probabilities for \( k = l, d, o \). The second inequality in (23) holds since \( \gamma^n = \gamma \) (see (6)). We derive the last inequality there for \( \gamma_1 \) in two step. First, we have

\[
|\gamma^l(\bar{X}_1) - \gamma^l(\bar{X}_2)| \leq \left| \int \left( \frac{\bar{w}l_0 + \bar{L}_1}{\bar{w} + S_1} - \frac{1}{\bar{w} + S_2} \right) \right| - \int \left( \frac{\bar{w}l_0 + \bar{L}_2}{\bar{w} + S_1} - \frac{1}{\bar{w} + S_2} \right) \right| \leq \left| \frac{\bar{w}l_0 - \bar{w}l_0 + \bar{L}_2}{\bar{w} + S_1} - \frac{\bar{w}l_0 + \bar{L}_1}{\bar{w} + S_2} \right| + \hat{\Gamma} \left[ \left( \frac{\bar{w}l_0 + \bar{L}_1}{\bar{w} + S_1} - \frac{\bar{w}l_0 + \bar{L}_2}{\bar{w} + S_2} \right) \right],
\]

where both inequalities follow from Lemma 2.1, and the second inequality also uses the triangular inequality and \( \hat{\Gamma} \) is the Lipschitz constant of \( \psi \). We continue as follows

\[
|\gamma^l(\bar{X}_1) - \gamma^l(\bar{X}_2)| \leq (1 + \hat{\Gamma}) \left[ \frac{\bar{w}l_0}{\bar{w} + S_1} - \frac{L_1}{\bar{w} + S_2} + \frac{L_2}{\bar{w} + S_1} - \frac{\bar{L}_2}{\bar{w} + S_2} + \frac{\bar{D}_1}{\bar{w} + S_1} - \frac{\bar{D}_2}{\bar{w} + S_2} \right]
\]

where the first and last inequalities follow from the triangular inequality. The second inequality follows from Lemma 2.2 and from some algebra using \( S_j = \bar{L}_j + \bar{D}_j + \bar{O}_j \) for \( j = 1, 2 \). Similarly, one can show that \( \gamma^d \) and \( \gamma^o \) are Lipschitz continuous (recall that \( \gamma^l(\bar{X}) + \gamma^d(\bar{X}) + \gamma^o(\bar{X}) = 1 \)). Finally, in our case \( \bar{X}^n(0) = 0 \) for all \( n \geq 1 \), which completes the proof.

**Proof of Proposition 2.** First note that \( \bar{X}(t) \) is absolutely continuous, since \( 0 \leq \gamma \leq 1 \) and \(|\bar{X}(t) - \bar{X}(s)| \leq \bar{L}|t - s| \). Therefore, the derivative of \( \bar{X} \) exists almost everywhere. From (11) we have at regular points \( t \)

\[
\bar{l}(t) = \frac{\bar{L}(t)}{\bar{w} + \Lambda t} - \frac{\bar{\Lambda}}{\bar{w} + \bar{\Lambda} t} \frac{\bar{w}l_0 + \bar{L}(t)}{\bar{w} + \Lambda t} = \frac{\bar{\Lambda}}{\bar{w} + \bar{\Lambda} t} \left[ \gamma^l(\bar{X}(t)) - \bar{l}(t) \right] = \frac{\bar{\Lambda}}{\bar{w} + \bar{\Lambda} t} \left[ \min(t^*, \bar{l}(t)) - \bar{l}(t) \right],
\]

where the derivation follows from (10) with \( \bar{L}(t) = \bar{\Lambda} \gamma^l(\bar{X}(t)) \) and from (7). The initial condition is \( \bar{l}(0) = l_0 \) since \( \bar{L}(0) = 0 \). Similarly, we obtain \( \bar{d}(t) = \bar{\Lambda} \left[ (\bar{l}(t) - t^*)^+ - \bar{d}(t) \right] / (\bar{w} + \bar{\Lambda} t) \) using the
definition of $\gamma^d$ with initial condition $\bar{d}(0) = 0$, since $\bar{D}(0) = 0$.

Proof of Proposition 3. In this proof we begin with the initial conditions $(\bar{I}(0), \bar{d}(0)) = (l_0, 0)$ and follow the learning trajectory resulting from (12). The price is arbitrary in $\mathbb{R}_+$ and satisfies Assumption 2.

No experimentation ($\psi = 0$): Following (13) we have $\dot{\bar{I}}(t) = \bar{I}(t)$ (it is easy to show that the projection in $\dot{I}(t)$ does not bind). In the underestimating case $\bar{I}(0) = l_0 \leq l^*$, hence $\dot{\bar{I}}(t) = 0$ and $\bar{I}(t) = l_0$ for all $t$. In addition, $\dot{\bar{d}}(0) = \bar{d}(0) = 0$ and the fraction of dislikes is constant at 0 for all $t$. In the overestimating case $\dot{\bar{I}}(0) < 0$ and $\dot{\bar{d}}(0) > 0$. A simple analysis of the ODE (12) shows that $\bar{I}$ converges to $l^*$ from above. When $\bar{I} = l^*$ all buyers like the product, therefore as $t$ grows large, the fraction of dislikes converges to zero and the system converges to $(l^*, 0)$. This establishes $\lim_{t \to \infty}(\bar{I}(t), \bar{d}(t)) = (\min(l_0, l^*), 0)$.

Experimentation: This proof builds on the later Proposition 5. The analysis of the transient case there reveals that in both the underestimating and overestimating cases there exists a time $\tilde{t}$ such that $\dot{\bar{I}}(t) \geq l^*$ for all $t \geq \tilde{t}$. Therefore, the first equation in (12) becomes $\dot{\bar{I}}(t) = \bar{I}(t)/\bar{I}(t) + t\bar{A}$, and by solving this ODE we find that $\lim_{t \to \infty}\bar{I}(t) = l^*$. For $t$ large the ODE for the fraction of dislikes becomes $\dot{\bar{d}}(t) = \bar{A}(\psi(l^*, d(t)) - d(t))/\bar{d}(t)$ and it follows that $\lim_{t \to \infty}\bar{d}(t) = d^*$. To complete the proof note that for any $\epsilon > 0$ we can take, for example, $\psi(l, d) = \psi_0 l - \psi_1 d$ with $\psi_0, \psi_1 > 0$. It follows that $d^* = l^*\psi_0/(1 + \psi_1)$, and for $\psi_0$ small enough $d^* \leq \epsilon$.

Proof of Proposition 5. We prove this under the assumption that the projection on $\dot{I}$ does not bind. The other case follows similarly. Consider first the underestimating case. We begin with the initial conditions of the ODE and show that $\dot{\bar{I}}$ crosses $l^*$ at most once from below at some finite time $T$. We then verify the statement for time points smaller and greater than $T$. At $t = 0$, by (13), we have $\bar{I}(0) = l_0 + \psi(l_0, 0) > l_0 = \bar{I}(0)$. Therefore, $\dot{\bar{I}}(0) = 0$. Let $T$ be the first time that $\dot{\bar{I}}(t) \geq l^*$ ($T$ may be 0). It follows that $\bar{d}(t) = \dot{\bar{d}}(0) = 0$ for $t \in [0, T]$, and since $\psi(l, 0) > 0$ we see that $\bar{I}$ is increasing in $[0, T]$. By continuity of $(\bar{I}, \bar{d})$ (this follows from system (12) and the continuity of $\psi$), there exists $\epsilon > 0$ such that $\dot{\bar{I}}(t) > l^*$ and $\dot{\bar{d}}(t) > 0$ for all $t \in (T, T + \epsilon]$. An analysis of system (12) with the initial conditions above shows that as long as $\dot{\bar{I}}(t) \geq l^*$ we have $\dot{\bar{I}}(t) > 0$. Therefore, we now show that $\dot{\bar{I}}(t) \geq l^*$ for all finite $t$ and that $\bar{d}$ is nondecreasing. By contradiction suppose that $\dot{\bar{I}}(t') < l^*$ for some $t' > T$. Then by continuity of $(\bar{I}, \bar{d})$ there exists a time $t'' < t'$ such that $\dot{\bar{d}}(t'') = (\dot{\bar{I}}(t'') - l^*)^+ - \bar{d}(t'') = 0$. But then $\dot{\bar{I}}(t') \geq 0$ by the monotonicity properties of $\psi$, and $\dot{\bar{I}}(t)$ cannot decrease to $l^*$ in violation of the assumption. This shows that $\dot{\bar{I}}(t) \geq l^*$ for all finite $t$ and that $\bar{d}$ is nondecreasing. The case where $q = q_0$ follows similarly.

In the analysis of the overestimating case we again begin with the initial conditions. Since $\dot{\bar{I}}(0) > l^*$, we have $\dot{\bar{I}}(0) < 0$ and $\dot{\bar{d}}(0) > 0$. By continuity of $(\bar{I}, \bar{d})$ there exist $t_l, t_d > 0$ such that $\dot{\bar{I}}(t) > 0$ for $t \in [0, t_l]$ and $\dot{\bar{d}}(t) > 0$ for $t \in [0, t_d]$. If $\psi \geq 0$ then $\dot{\bar{I}}(t) \geq \bar{I}(t)$ for all $t$. It follows that $\bar{I}$ may become

31
negative only if there exists a time point \( t' < \infty \) such that \( \bar{l}(t') = l^* \) (see (12)). For \( [0, t'] \), i.e., while \( \bar{l} \geq l^* \), we can solve (12) for \( \bar{l} \). The analysis of the trajectory shows that \( \bar{l}(t) > l^* \) for any finite time, and in particular at \( t' \), showing that \( t_f = \infty \).

**Proof of Proposition 6.** The learning trajectory has kinks at points \( t \) where \( \bar{l}(t) = l^* \) (in fact there is one such point, \( T \)). Between these points the system is a non-homogenous first order linear ODE that can be solved in closed form with standard methods. In this proof we identify these regions and establish the initial condition governing each one of them.

**Case 1** \( l_0(1 + \psi_0) < l^* \): Here \( \ddot{\bar{l}}(0) = ((l_0(1 + \psi_0) - l^*)^+ - \bar{d}(0))/w = 0 \) and subsequently \( \ddot{\bar{l}}(t) = 0 \) for \( t \) small. Therefore, for small \( t \), \( \bar{l}(t) = \bar{l}(0)\psi_0/(w + t) \) with the solution \( \bar{l}(t) = l_0((w + t)/w)^{\psi_0} \).

This holds until time \( T \) (see (14)), since there \( \ddot{l}(T) = \bar{l}(T)(1 + \psi_0) = l^* \), and dislikes appear for the first time \( \dot{\bar{l}} > 0 \). We first guess that \( \ddot{l}(t) > l^* \) and that \( l_{\max} = \bar{l}(t)(1 + \phi(\dot{d}(t)/\dot{l}(t))) \geq l^* \) for all \( t > T \), and then check that it is indeed the case. The solution is given in (16), and it is then easy to verify that the two inequalities hold for all \( t > T \).

**Case 2** \( l_0(1 + \psi_0) \geq l^* \): Here \( \ddot{l}(0) > l^* \) and \( \ddot{d}(0) > 0 \). We guess that \( \ddot{l}(t) > l^* \) and \( l_{\max} = \bar{l}(t)(1 + \phi(\dot{d}(t)/\dot{l}(t))) \geq l^* \) for all \( t, \) solve the system of ODEs and verify that the conditions hold.

**Proof of Proposition 7.** First we note that \( \bar{l}(t) \), as given in Proposition 6, is increasing over time in the underestimating case and decreasing over time in the overestimating case, that \( \bar{l}(0) = l_0 \), and that \( \bar{l}(t) \rightarrow l^* \). Therefore, we conclude that \( T(\epsilon) = 0 \) for all \( \epsilon > |l_0/l^* - 1| \) and otherwise \( T(\epsilon) \) is the solution of \( |\bar{l}(T(\epsilon))/l_0 - 1| = \epsilon \). Breaking this absolute value and substituting for \( \bar{l}(t) \) we find that in the underestimating case \( \mathrm{sign}(\partial T(\epsilon)/\partial p) = \mathrm{sign}(\partial (l^*/l_0)/\partial p) = \mathrm{sign}(G(p/q_0) - G(p/q)) \), where the last step follows by differentiation by \( p \). Since \( p/q_0 > p/q \), we conclude that the \( \epsilon \)-time-to-learn is increasing in the price if \( F \) is IGFR, and decreasing if \( F \) is DGFR. In the overestimating case we have \( \mathrm{sign}(\partial T(\epsilon)/\partial p) = \mathrm{sign}(\partial (l_0/l^*)/\partial p) = \mathrm{sign}(G(p/q) - G(p/q_0)) \). Here \( p/q > p/q_0 \) and the same conclusions hold.

**Proof of Proposition 8.** Let \( \bar{p} := \max(p(q), p(q_0)) \) and \( \underline{p} := \min(p(q), p(q_0)) \) (recall that \( p(q) = \arg\max_p R(p,q) \)). The proof first establishes that \( \partial \bar{\pi}/\partial p \big|_{p=\bar{p}} > 0 \) for all \( \bar{p} < \underline{p} \), and \( \partial \bar{\pi}/\partial p \big|_{p=\bar{p}} < 0 \) for all \( \bar{p} > \bar{p} \). This shows that all optimal unconstrained prices \( p^* \) lie in \( [\underline{p}, \bar{p}] \). Since we assumed that \( p(q), p(q_0) \in [p_{\min}, p_{\max}] \), this also shows that all optimal constrained prices lie in this interval. We state without a proof that \( q < \tilde{q} \) implies \( p(q) \leq p(\tilde{q}) \). Thus we have for the overestimating case \( p^* \in [\underline{p}, \bar{p}] = [p(q), p(q_0)] \) and for the underestimating case \( p^* \in [\underline{p}, \bar{p}] = [p(q_0), p(q)] \), as in the statement of the proposition.

The reminder of the proof establishes the conditions on the signs of \( \partial \bar{\pi}/\partial p \) stated above. Terms
\( \{s_j\}_{j=1}^3 \) and \( \{h_j\}_{j=1}^2 \) in (19) and (20) are given by

\[
s_1 := \int_0^T (1 + \psi_0)e^{-\delta t}\left(\frac{w + t}{w}\right)^\psi_0 dt, \quad s_2 := \int_T^\infty e^{-\delta t}\left[1 + x^* - x^*(\frac{w + T}{w + t})^{1+\psi_1}\right] dt, \\
s_3 := \tilde{F}(p/q)\int_T^\infty e^{-\delta t}\left(\frac{w + T}{w + t}\right)^{1+\psi_1} dt, \quad h_1 := \int_0^\infty e^{-\delta t}(1 + \psi_0)(\frac{w + t}{w})^{-1-\psi_1} dt, \\
\text{and } h_2 := \int_0^\infty e^{-\delta t}(1 + x^*)(1 - (\frac{w + t}{w})^{-1-\psi_1}) dt.
\]

The derivation of \( \{s_j\}_{j=1}^3 \) follows by substituting for \( T, \) and \( \partial T/\partial p. \)

**Overestimating:** Note that \( h_1, h_2 > 0. \) For \( \tilde{p} < p(q) \) we have that \( R'(\tilde{p}, q_0) \) and \( R'(\tilde{p}, q) \) are both positive and thus \( \partial \pi(\tilde{p})/\partial p > 0. \) For all \( \tilde{p} < p(q) \) we have that \( R'(\tilde{p}, q_0) \) and \( R'(\tilde{p}, q) \) are both negative, thus \( \partial \pi(\tilde{p})/\partial p < 0. \)

**Underestimating:** Here the learning process may follow case (trajectory) 1 or 2 in Proposition 6 depending on whether \( l_0(1 + \psi_0) < l^* \) or not, or equivalently on whether \( T > 0 \) or \( T = 0. \) At prices where the trajectory changes from 1 to 2 or reversely, there may be a kink in the derivative, although continuity of \( \pi(p) \) is preserved for all \( p. \) We now show that for any price \( \tilde{p} < p(q_0), \) \( \partial \pi(\tilde{p})/\partial p > 0 \) under both trajectories. For trajectory 2 this follows from (20) similarly to the overestimating case. For trajectory 1 we have

\[
\frac{\partial \pi(p)}{\partial p} \bigg|_{p=\tilde{p}} > s_2 \tilde{F}(\tilde{p}/q)[1 - G(\tilde{p}/q)] + [G(\tilde{p}/q_0) - G(\tilde{p}/q)]s_3 \\
> [s_2 \tilde{F}(\tilde{p}/q) - s_3][1 - G(\tilde{p}/q)] \\
\geq 0,
\]

where the first inequality follows from \( R'(\tilde{p}, q_0) > 0 \) and from \( R'(p, q) = \tilde{F}(p/q)[1 - G(p/q)]. \) This also implies that \( G(\tilde{p}/q), G(\tilde{p}/q_0) < 1 \) and thus justifies the second inequality as \( s_3 \geq 0. \) Lastly, the last inequality follows from \( s_2 \tilde{F}(\tilde{p}/q) \geq s_3 \) that follows by comparing the integrands in (24). Finally, we show that for any \( \tilde{p} > p(q) \) we have \( \partial \pi(p)/\partial p < 0 \) under both trajectories. Trajectory 2 follows from (20) as in the overestimating case. For trajectory 1 one can follow the reverse steps to the ones taken in (25), while noting that now \( G(\tilde{p}/q_0), G(\tilde{p}/q) > 1. \) This concludes the proof.

**Proof of Lemma 1.** The proof shows that the sign of \( \partial q_1/\partial p_0 \) is negative under IGFR and positive under DGFR. Note that,

\[
\frac{\partial \tilde{F}^{-1}(\tilde{l}_0(\tau))}{\partial p_0} = \frac{f(p_0/q_0)/q_0}{f(F^{-1}(\tilde{l}_0(\tau))))}\left(\frac{w + t}{w}\right)^\psi_0,
\]
since whenever a function \( h \) and its inverse \( h^{-1} \) are differentiable, we have \( dh^{-1}(x)/dx = -1/h'(h^{-1}(x)). \)
Therefore,

\[
\frac{\partial q_1}{\partial p_0} = C(p_0) \left[ \bar{F}^{-1}(\bar{1}_0(\tau)) f(\bar{F}^{-1}(\bar{1}_0(\tau))) - f(p_0/q_0)(p_0/q_0) \left( \frac{w + t}{w} \right)^{\psi_0} \right] \\
= C(p_0) \left[ \bar{F}(\bar{F}^{-1}(\bar{1}_0(\tau))) G(\bar{F}^{-1}(\bar{1}_0(\tau))) - \bar{F}(p_0/q_0) \left( \frac{w + t}{w} \right)^{\psi_0} G(p_0/q_0) \right] \\
= C(p_0) \bar{1}_0(\tau) \left[ G(\bar{F}^{-1}(\bar{1}_0(\tau))) - G(p_0/q_0) \right],
\]

where \( C(p_0) = (\bar{F}^{-1}(\bar{1}_0(\tau)))^{-2}/f(\bar{F}^{-1}(\bar{1}_0(\tau))) > 0 \), and \( G \) is the generalized failure rate. The second equality above follows by dividing and multiplying the first term by \( \bar{F}(\bar{F}^{-1}(\bar{1}_0(\tau))) \), so it becomes \( \bar{F}(x)f(x)/\bar{F}(x) = \bar{F}(x)G(x) \), where \( x = \bar{F}^{-1}(\bar{1}_0(\tau)) \). The second term follows similarly. Since \( \bar{F}^{-1}(\bar{1}_0(\tau)) < p_0/q_0 \) for \( \tau > 0 \), the derivate is negative if \( F \) is IGFR and positive if \( F \) is DGFR.

**Lemma 3.** Suppose that \( F \) is IGFR. Then \( \partial \bar{T}_1/\partial \tau \leq 0 \) if \( p_0 \leq p_1 \), and \( \partial \bar{T}_1/\partial \tau \geq 0 \) if \( p_0 \geq p_1 \).

**Proof.** This proof studies the derivative of \( \bar{T}_1 \) with respect to \( \tau \). First compute

\[
\frac{\partial \bar{T}_1}{\partial \tau} = l_0 \frac{p_1}{p_0} \psi_0 \left( \frac{w + t}{w} \right)^{\psi_0-1} f \left( \frac{p_1}{p_0} \bar{F}^{-1}(\bar{1}_0(\tau)) \right) / f \left( \bar{F}^{-1}(\bar{1}_0(\tau)) \right),
\]

which follows from \( d\bar{F}^{-1}(x)/dx = -1/f(\bar{F}^{-1}(x)) \). We use this to compute

\[
\frac{\partial \bar{T}_1}{\partial \tau} = \left( \frac{1}{1 + \psi_0} \frac{l_1^*}{l_1} \right)^{1/\psi_0} \left[ 1 - \frac{w + \tau \partial \bar{1}_1}{\psi_0 l_1} \frac{\partial l_1}{\partial \tau} \right] \\
= \left( \frac{1}{1 + \psi_0} \frac{l_1^*}{l_1} \right)^{1/\psi_0} \left[ 1 - \frac{p_1}{p_0} \bar{F}^{-1}(\bar{1}_0(\tau)) f \left( \frac{p_1}{p_0} \bar{F}^{-1}(\bar{1}_0(\tau)) \right) \frac{\bar{F}(\bar{F}^{-1}(\bar{1}_0(\tau)))}{\bar{F}^{-1}(\bar{1}_0(\tau)) f(\bar{F}^{-1}(\bar{1}_0(\tau)))} \right] \\
= \left( \frac{1}{1 + \psi_0} \frac{l_1^*}{l_1} \right)^{1/\psi_0} \left[ 1 - G \left( \frac{p_1}{p_0} \bar{F}^{-1}(\bar{1}_0(\tau)) \right) / G \left( \bar{F}^{-1}(\bar{1}_0(\tau)) \right) \right].
\]

Since \( F \) is IGFR, \( G \) is nondecreasing. Thus, \( G \left( (p_1/p_0) \bar{F}^{-1}(\bar{1}_0(\tau)) \right) / G \left( \bar{F}^{-1}(\bar{1}_0(\tau)) \right) \geq 1 \) for \( p_1 > p_0 \), and the reverse for \( p_1 < p_0 \).

**Proof of Proposition 9.** This proof analyzes the Kuhn Tucker conditions of the optimization problem. The constraint \( \tau \leq \bar{\tau} \) requires some analysis. The quantity \( T_0 := w(l_0^*/l_0)^{1/\psi_0}(1 + \psi_0)^{-1/\psi_0} - w \) represents the first time dislikes would appear under \( p_0 \). This is identical to \( T \) in the static price. In order for \( \tau \leq \bar{\tau} \), the constraint \( \tau \leq T_0 \) surely must be imposed. Call \( T_1 \) the first time dislikes would appear under \( p_1 \), assuming they did not appear at time \( \tau \); \( T_1 := (w + \tau)(l_1^*/l_1)^{1/\psi_0}(1 + \psi_0)^{-1/\psi_0} - w \). Therefore, \( \bar{\tau} = \max(\tau, T_1) \). The optimization problem assumes that \( \bar{\tau} = T_1 \). It is later shown that this is indeed the case. In addition, the constraints on the prices are ignored, except for \( p_0 \geq p_{\min} \), and their validity is later verified. The proof establishes the result for the unconstraint case and the \( p_0^* = p_{\min} \) case, and then shows by contradiction that \( p_0^* > p_1^* \) is not possible when the constraint \( \tau \leq T_0 \) is binding.
The associated Lagrangian is,
\[
\mathcal{L} = \pi(p_0, p_1, \tau) + \mu_1[T_0 - \tau] + \mu_2[p_0 - p_{\text{min}}],
\]
where \(\mu_1, \mu_2 \geq 0\). The revenue after \(\tau\) follows the same structure of the static pricing problem with initial weight \(w + \tau\) and prior quality estimate \(q_1\). Moreover, \(p_1\) does not interact with the constraints. Thus, we conclude that \(p_1^* \in [p(q_1), p(q)]\). In the reminder of this proof we use the notation \(c_k\) to denote a nonnegative quantity (which may depend on the controls) unless otherwise mentioned. Computation shows that
\[
\frac{\partial \mathcal{L}}{\partial p_0} = R'(p_0, q_0)c_1 + \frac{\partial l_1}{\partial p_0}c_2 - \frac{\partial T_1}{\partial p_0}c_3 + \mu_1 \frac{\partial T_0}{\partial p_0} + \mu_2.
\]
Recall that \(l_1 = \bar{F}(p_1/q_1)\) and that \(\partial q_1/\partial p_0 < 0\) (Lemma 1). This shows \(\partial l_1/\partial p_0 \leq 0\). By the same token, \(\partial T_1/\partial p_0 = -c_4\partial l_1/\partial p_0 \geq 0\). Suppose that the constraints are not binding so that \(\mu_1 = \mu_2 = 0\), then it must be the case that \(R'(p_0, q_0) > 0\), and from unimodality \(p_0^* < p(q_0)\). If the constraint \(p_0 \geq p_{\text{min}}\) is binding, then \(p_0^* = p_{\text{min}} < p(q_0)\) by Assumption. So, in both cases \(p_0^* < p(q_0) < p(q_1) \leq p_1^*\). Thus, we need to check that \(p_0^* \geq p_1^*\) is not possible when \(\tau^* = T_0\). The first order condition for \(\tau\) is
\[
\frac{\partial \mathcal{L}}{\partial \tau} = e^{-\delta \tau}(1 + \psi_0)[p_0\bar{l}_0(\tau) - p_1l_1] - c_5 - c_6 \frac{\partial T_1}{\partial \tau} - \mu_1.
\]
Suppose that \(p_0^* \geq p_1^*\), then \(p_1^*l_1 = R(p_1^*, q_1) = R(p_1^*, p_0^*/\bar{F}^{-1}(\bar{l}_0(\tau))) > R(p_0^*, p_0^*/\bar{F}^{-1}(\bar{l}_0(\tau))) = p_0^*\bar{F}(\bar{F}^{-1}(\bar{l}_0(\tau))) = p_0^*\bar{l}_0(\tau)\). Also from Lemma 3, \(\partial T_1/\partial \tau \geq 0\). We conclude that \(\partial \mathcal{L}/\partial \tau < 0\), which contradict the Kuhn Tucker conditions, so \(p_0^* \geq p_1^*\) is not possible if \(\tau = T_0\).

It is easy to see that the omitted constraints on the prices are not binding, \(p_{\text{min}} \leq p_0^* < p_1^* \leq p(q) < p_{\text{max}}\). To conclude the proof we show that \(T_1 = \bar{\tau}\), by showing \(T_1 \geq \tau\). Note that \(\bar{F}(p/q)/\bar{F}(p/q')\) is increasing in \(p\) for \(q > q'\) and \(\bar{F}\) IGFR. Therefore,
\[
\tau \leq w\left(\frac{1}{1 + \psi_0} \frac{\bar{F}(p_0^*/q_1)}{\bar{F}(p_0^*/q)}\right)^{1/\psi_0} - w \leq (w + \tau)\left(\frac{1}{1 + \psi_0} \frac{\bar{F}(p_1^*/q_1)}{\bar{F}(p_1^*/q)}\right)^{1/\psi_0} - w = T_1,
\]
which concludes the proof.