NET Institute*

www.NETinst.org

Working Paper #11-11

October 2011

Monopoly Pricing in the Presence of Social Learning

Bar Ifrach
Columbia Business School

Costis Maglaras
Columbia Business School

Marco Scarsini
LUISS

* The Networks, Electronic Commerce, and Telecommunications ("NET") Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, "virtual networks" comprised of computers that share the same technical standard or operating system, and on network issues in general.

Monopoly Pricing in the Presence of Social Learning

Bar Ifrach* Costis Maglaras† Marco Scarsini‡

November 2011

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. This analysis borrows from the literature on social learning and on pricing and revenue management.

Consumers follow a naive decision rule and, under some conditions, eventually learn the product's quality. Using mean-field approximation, the dynamics of this learning process are characterized for markets with high demand intensity. The relationship between the price and the speed of learning depends on the heterogeneity of quality preferences. Two pricing strategies are studied: a static price and a single price change. Properties of the optimal prices are derived. Numerical experiments suggest that pricing strategies that account for social learning may increase revenues considerably relative to strategies 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 on 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

*Columbia Business School, 4I Uris Hall, 3022 Broadway, NY, NY 10027. (bifrach14@gsb.columbia.edu). I thank the NET Institute (www.NETinst.org) for financial support.

†Columbia Business School, 409 Uris Hall, 3022 Broadway, NY, NY 10027. (c.maglaras@gsb.columbia.edu)

‡Dipartimento di Economia e Finanza, LUISS, Viale Romania 32, 00197 Roma, Italy. (marco.scarsini@luiss.it)

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

In broad terms the model assumes that consumers arrive to the market according to a Poisson process and sequentially face the decision of either purchasing a product with unknown quality, or choosing an outside option. Each consumer has a willingness-to-pay for the product that is a function of his idiosyncratic quality preference and the product's quality. These quality preferences are assumed to be independently and identically drawn from a know distribution. If the true product quality were known to consumers, this would give rise to a demand function that the monopolist could use as a basis of her pricing decision.

In our model, however, the quality is unknown, and consumers' estimates about it evolve according to a social learning mechanism. Specifically, motivated by the way information gets revealed and shared in practice, we study the following. Consumers who purchase the product report whether they "liked" or "disliked" the product, which corresponds to whether their ex-post utility was non-negative, or negative. Consumers do not report their quality preference, so a positive review may result from a high quality or high idiosyncratic quality preference (not necessarily both). A new consumer observes the information reported by a sample—or potentially all—of the consumers who made decisions prior to him, makes an inference about the product quality, and then makes his own purchase decision. This sequence of idiosyncratic purchase decisions affects the evolution of the observable information set, and as such the dynamics of the demand process over time. Optimizing the monopolist's pricing strategy requires detailed understanding of the learning dynamics and not just its asymptotic properties.

It is typical to assume that fully rational agents update their beliefs for the unknown quality of the product through a Bayesian analysis, but the extraordinary analytical and computational onus that this imposes on each agent may be hard to justify as a model of actual choice behavior. Instead, we will postulate a "naive" and fairly intuitive learning mechanism, where consumers do not see the sequence in which the decisions and the reports that they observe were made. Rather, they observe the fraction of consumers who decided to purchase the product up to that point, and out of them the fraction of buyers that liked the product and the fraction of buyers that disliked the product. They then find the quality estimate that would be consistent with the aggregate information that is available to them, and make their purchasing decision accordingly.

To motivate this structure, consider the establishment of a new hotel or resort. It is typically hard to evaluate the quality of such premises without first hand experience or word-of-mouth. This explains the huge impact online review websites such as Tripadvisor have had on the hospitality industry[1]. Assume the hotel is differentiated enough from competitors so that it could be considered

[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.

a monopoly in some category. For example, it may be the only hotel with a private beach in the area. Suppose it offers better services than what consumer think at first. Initially some consumers' idiosyncratic tastes would convince them to choose this hotel; perhaps they have strong preferences for the private beach. These consumers would recommend the hotel by posting a review online. This in turn increases future demand, as potential consumers learn that the hotel is better than they previously thought. 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 so to 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 some aspects of online reviews as well as the possible bounded rationality of the consumers. Second, in terms of analysis, by proposing a tractable 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 enough to be applicable in other related settings where the nature of information that consumers share and the learning mechanism are different. Third, in addressing some of the pricing questions faced by revenue maximizing sellers in such settings.

In more detail, the main findings of this paper are the following. First, the information reported by consumers is subject to a self-selection bias, since only consumers with a high enough quality preference purchase the product. If consumers ignore the self-selection bias, then they may not learn—even asymptotically—the true product quality. On the positive side, Proposition 3.6 shows that learning will eventually occur almost surely if consumers correct for this bias. Learning occurs for two types of social networks, the *basic network* where each consumer observes the entire sequence of past consumer decisions, and the *random network* where each agent independently observes each of his predecessors with some probability, but the size of the sample grows large as the number of agents increases to infinity. The speed of convergence, and, better yet, the learning trajectory over time is essential in capturing the tradeoff between consumer 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 4.2 shows that the asymptotic learning trajectory is characterized by a system of differential equations, and Proposition 4.4 derives its closed form solution. 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.

Third, we study the seller's pricing problem, starting with the effect of the price on the speed of the social learning process. Proposition 5.1 provides a crisp characterization of that effect, which depends on the generalized failure rate of the preference distribution, a concept that has been studied by Lariviere (2006). The seller's optimal price is one that balances the learning process with the discounted revenue objective, and is computed through an analysis of a pricing problem driven by the aforementioned deterministic ODE that characterizes the learning process. Intuitively, if learning is too slow with respect to the revenue discount rate, then the seller will price almost as if all consumers' purchasing decisions were based on their prior estimate of the product quality; if learning is fast relative to the discount rate of revenues, then the seller's price will approach the one that the monopolist would set if all consumers knew the true product quality. Proposition 5.4 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 6.2 shows, under general assumptions, that when initially consumers underestimate the true quality, the first price is lower than the second price. This pricing strategy can accelerate learning and increase revenues considerably. The numerical experiments of Section 6 suggest that a pricing strategy with two prices performs very well, and that the benefit of implementing more elaborate pricing strategies may be small.

1.1 Literature Review

This paper lies in the intersection of a strand 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 and spans the fields of economics, statistics, and engineering. Much of this work can be classified into two groups depending on the learning mechanism, which is either Bayesian or some type of "naive" or "ad-hoc" or "engineered/optimized" mechanism.

Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992) started the literature 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 (or lack thereof). 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. This will prevent any further learning and therefore may lead to herding on the bad decision.[2]

For social learning to be successful, with positive probability an agent would reverse the herd behavior of his predecessors. Smith and Sørensen (2000) show that this would be the case if agents' signals have unbounded strength. Goeree, Palfrey, and Rogers (2006) show that this achieved with enough heterogeneity in consumers' preferences. Our Assumption 3.4, which is key in proving learning, is similar in nature to that of Goeree et al. (2006).

Several papers have considered variations of the observational learning model with imperfect information that are related to our paper. Smith and Sørensen (2008) and Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) consider the case where agents observe only subsets of their predecessors. In the first, agents do not observe the sequence in which actions were made. Both papers show that observing a finite set of predecessors may suffice for learning when signals have unbounded strength. We, however, find that for the random network an increasing number of predecessors has to be observed for asymptotic learning to occur. These conclusions differ not because of the assumed bounded rationality of consumers in this paper, but due to the heterogeneity in consumers preferences. Since the types of consumers are not observed, the choices of any finite subset of predecessors may be driven by their idiosyncratic preference, and not by the underlying state of the world. Banerjee and Fudenberg (2004) consider a continuum of agents who interact repeatedly by a way of word-of-mouth (although it is not explicitly modeled). Here too a finite sample may suffice for learning.

Herrera and Hörner (2009) consider a case where agents can observe only one of two decisions of their predecessors. In the language of our model, their paper relaxes the assumption that consumers who did not purchase the product are observed. Instead, in their model 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 standard 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 naive learning mechanisms that employ simpler and perhaps more plausible learning protocols by each agent, which is also related to the growing body of engineering literature on sensor networks and decentralized algorithms. Ellison and Fudenberg (1993, 1995) consider settings in which consumers exchange information about their experienced utility and use naive 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. Golub

[2]In the nineties similar problems have been considered by several authors (see, among others Welch, 1992; Lee, 1993; Chamley and Gale, 1994; Vives, 1997) and the survey by Bikhchandani, Hirshleifer, and Welch (1998). A more recent survey by Acemoglu and Ozdaglar (2011) considers both Bayesian and non-Bayesian approaches, and many of the extensions to this principal model.

and Jackson (2010) and Jadbabaie, Molavi, Sandroni, and Tahbaz-Salehi (2011) consider naive learning mechanism where agents average their estimates or beliefs with those of their neighbors. Conditions for asymptotic learning are characterized. Contrary to these papers, we consider agents whose heterogeneous preferences directly influence their decision rule.[3]

The area of *revenue management* focuses, in part, on tactical problems of price optimization[4]. It is typical therein to assume that the consumer response is captured 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. The problem that we study in this paper can be viewed as one where the demand model is serially-correlated, and, specifically, evolves over time according to a social learning mechanism, which itself depends on the seller's policy.

Barring the social learning process, our problem resembles many problems in revenue management. This problem is well behaved if consumers' willingness-to-pay are drawn from a distribution with an increasing generalized failure rate. Lariviere (2006) discusses this condition, which is common in revenue management, and shows that it is related to the elasticity of demand. Interestingly, we find that the generalized failure rate is intimately related to the relationship between the price and the speed of the social learning process.

Mean-field approximation is widely used in revenue management to study complex dynamic demand systems and subsequently to optimize performance, for example see Gallego and van Ryzin (1994) and Bitran, Caldentey, and Vial (2005). These approximations explore the dynamics of the system as the market intensity, often captured by the demand rate, is increasing. We apply a result by Kurtz (1977/78), which is also used in Bitran et al. (2005), to derive the approximation.

The learning dynamics in our model give rise to a sales trajectory (or evolution of the underlying demand model) which, when properly interpreted, resemble 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, which is 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 considerations when agents are engaged in social learning or

[3]Related problems were studied in the engineering literature on aggregation of information over decentralized sensors. For example, Tsitsiklis (1984) and Tsitsiklis, Bertsekas, and Athans (1986) consider consensus via averaging of estimates, and Papastavrou and Athans (1992) considers a problem similar to the classic observational learning model.

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

[5]In particular, by considering a population with finite mass, and by simplifying the decision rule.

embedded in a social network. Bose, Orosel, Ottaviani, and Vesterlund (2006) consider a variation of the classic Bayesian model of social learning when a monopolist and agents are equally uninformed about the value of the good. The monopolist can change the price dynamically to extract revenue and to control the accumulation of information. Campbell (2009) studies the role of the pricing decision in the launching of a new product in a model of social interaction that builds on percolation theory. As such, the focus is on the connectivity of the network of consumers, which is closer to the analysis of targeted advertising and pricing than to our analysis, which considers more conservative pricing techniques.

Experience goods are goods whose quality can be determined only upon consumption. Several authors have studied pricing and social learning in the context of experience goods, for example see Bergemann and Välimäki (2000), Caminal and Vives (1999), and Vettas (1997). Almost all 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. They show that the expected price path of the new product is increasing when initially consumers underestimate the true quality. Our Proposition 6.2 is in accordance with their findings, even when competition is ignored.

Candogan, Bimpikis, and Ozdaglar (2011) study the optimal prices to offer a group of consumers embedded in a social network. Consumption of the product involves network externalities, so the optimal prices may differ between consumers depending on their social influence. In particular, the optimal prices are shown to be proportional to consumers' network centrality. Hartline, Mirrokni, and Sundararajan (2008) and Kempe, Kleinberg, and Tardos (2003) study a computational approach to the problem of targeted marketing, which attempts to promote the adoption of a product by an influential group of agents in order to accelerate word-of-mouth effects.

2 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 $(\underline{q}, \bar{q})$ with $\underline{q} > 0$. Consumers are heterogeneous; this is represented by a parameter α_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[6]. The utility derived from choosing the outside alternative is normalized to zero for all consumers.

For simplicity, we assume that the 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 $\bar{F}(x) := 1 - F(x)$, and assume that F has a differentiable density f with connected support either $[0, \bar{\alpha}]$ or $[0, \infty)$. The preference parameter can be interpreted as a premium that a consumer is willing to pay for a unit of quality. In this model, unlike most papers in social learning, consumers are heterogeneous and therefore even if the underlying state of the world, q, were known, not all of them would make the same decision. Namely, only those with preference parameters $\alpha_i \geq \alpha^* := p/q$ would decide to purchase the product, and the rest would choose the outside alternative.

The product is launched at time $t = 0$, and consumers arrive thereafter. Denote by t_i the random time consumer i enters the market and makes his purchasing decision. We assume that consumers' arrival process is Poisson with rate Λ. This arrival process is independent of product's quality and consumers' preference parameters. Consumers never reenter the market after making their purchasing decision, even if they chose the outside alternative. 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, $q_0 \in (\underline{q}, \bar{q})$. 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. We abstract away from the formation of this prior guess and assume that it is the same for all consumers.

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 report a review about their experience with the product. Reviews about the personal experience with the product take two values: 'like' and 'dislike'. A consumer reports that he likes the product if he purchased and his ex-post utility was nonnegative, taking into account the true quality of the product and his preference parameter. If the consumer's ex-post utility from purchasing was negative, then he reports that he dislikes the product. We view this binary report as a simplification of the star rating scales that are ubiquitous in online review systems. In addition, a consumer that 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. Since the value of a consumer's preference parameter is private, reviews are not fully informative. For example, a like could result from either high

[6] The functional form of the utility function does not play a big role in the analysis. We choose this one for convenience.

preference parameter or high quality (not necessarily both)[7]. We assume that reviews are truthful.

The information available to consumer i upon making his purchasing decision is denoted by I_i. This may include reviews made by all or a subset of his predecessors, and possibly knowledge about the order in which they acted. Each consumer uses his available information to decide whether to purchase the product or to choose the outside alternative. Before describing the flow of information and the 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,

$$\begin{aligned}\max_p \pi(p) &= \max_p \mathsf{E}\Big[\sum_{i=1}^{\infty} e^{-\delta t_i}\ p\mathbf{1}\{b_i(p)\}\Big] \\ &= \max_p \sum_{i=1}^{\infty} \mathsf{E}\Big[e^{-\delta t_i}\ p\mathsf{P}\big(b_i(p)|I_i\big)\Big], \qquad (1)\end{aligned}$$

where $b_i(p)$ is the event that consumer i purchases the product, $\delta > 0$ is the monopolist's discount factor, and the expectation is with respect to consumers' arrival times and quality preferences. Here the monopolist chooses a price once and for all (static price), and knows the true quality, the prior quality estimate, and the distribution of quality preferences. Section 6 considers two prices and 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. We discuss this optimization problem and its extensions in the sequel.

2.2 Naive Decision Rule

We consider two information structures. In the *basic social network* all predecessors are observed. In the *random social network* consumer $i+1$ observes a random subset of his predecessors, where each predecessor is drawn independently with probability ρ_i. We illustrate the available information and the decision rule on the basic social network. The decision rule generalizes to the random social network straightforwardly, and we comment about that when needed. We define the following quantities: $L(i) := \sum_{j=1}^{i} \mathbf{1}\{r_j = r^{\mathrm{l}}\}$ 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 by $l(i) := L(i)/i$, $d(i) := D(i)/i$,

[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.

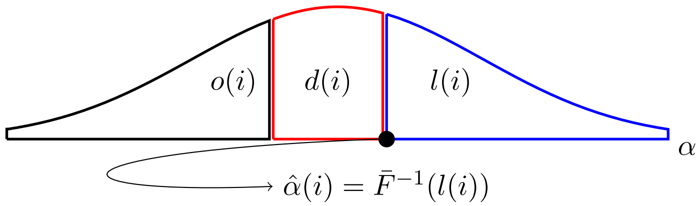


Figure 1: Illustration of naive estimation

and $o(i) := O(i)/i$ the corresponding fractions. The information available to consumer $i+1$ before making his decision is $I_{i+1} = (L(i), D(i), O(i))$, where $O(i)$ could be omitted if consumer $i+1$ knows that he is the $(i+1)$-st consumer.

We introduce a naive decision rule that is used by consumers to decide whether or not to purchase the product, followed by a discussion of its various elements. This decision rule is composed of two parts: (a) consumer i uses his available information to form a quality estimate $\hat{q}(i)$, and (b) he purchases the product if and only if his estimated utility is nonpositive $\alpha_i \hat{q}(i) - p \geq 0$.

In broad terms, the naive estimation part is described as follows. Recall that consumers do not observe the sequence in which reviews were submitted (alternatively, they ignore it to simplify their decision). The first naive approximation a consumer makes is to assume 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 $\hat{q} > q$). The second naive approximation a consumer makes is to assume that the empirical distribution of quality preferences among predecessors coincides with F. These two steps are illustrated on the 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 $\hat{\alpha}(i) := \bar{F}^{-1}(l(i))$ (see Figure 1) is the one that leaves the consumer with zero utility, since predecessors with higher quality preferences liked the product and those with lower quality preferences disliked it. Consumers' quality estimate is the one that rationalizes $\hat{\alpha}(i)$ based on the above mentioned naive approximations, i.e., $\hat{\alpha}(i)\hat{q}(i) - p = 0$ or $\hat{q}(i) = p/\bar{F}^{-1}(l(i))$. Put differently, consumers estimate $\alpha^* = \bar{F}(p/q)$ from observations of the type "$\alpha_j < \alpha^*$" (dislike) and "$\alpha_j \geq \alpha^*$" (like) for $j \leq i$. Generally, this estimation procedure should become successful as i grows large, however in our case these observations are censored. That is, likes and dislikes are reported condition on the decision of the predecessor to purchase the product and therefore they are subject to a *self-selection bias*. In particular, when $\hat{q} < q$ dislikes will not be reported. Consumers will naively account for this bias by employing a *correction term* as will be discussed in the sequel.

In more detail, the formation of the quality estimator—part (a)—includes two steps in the following order: (a.1) consumers account for the prior quality estimate, and for the noise in their available information when the number of observations is small; (a.2) consumers employ the above-mentioned estimation procedure while correcting for the self-selection bias.

To account for the prior quality estimate and for the noise in $l(i)$ and $d(i)$ when the number of predecessors observed is small, in step (a.1) consumer $i+1$ computes the *weighted* fractions of likes and of dislikes,

$$l_w(i) := \text{proj}_{[\underline{l},\bar{l}]}\left(\frac{w}{w+i}l_0 + \frac{i}{w+i}l(i)\right), \tag{2}$$

and

$$d_w(i) := \text{proj}_{[\underline{d},\bar{d}]}\left(\frac{w}{w+i}d_0 + \frac{i}{w+i}d(i)\right). \tag{3}$$

The operator $\text{proj}_{[\underline{x},\bar{x}]}(x)$ projects x to the interval $[\underline{x},\bar{x}]$, and the quantity $w > 0$ is the weight given to the prior quality estimate, as will be detailed later. It will be shown that for large i the weighted fractions $l_w(i)$ and $d_w(i)$ are close to $l(i)$ and $d(i)$. In step (a.2) the consumer naively corrects for the self-selection bias by inflating the weighted fraction of likes by a quantity $\psi(i) \geq 0$, which depends on the information available to him. This corresponds to moving $\hat{\alpha}(i)$ in Figure 1 to the left. Following the naive estimation procedure, the quality estimate is

$$\hat{q}(i+1) := \text{proj}_{[\underline{q},\bar{q}]}\left(p/\bar{F}^{-1}\left(l_w(i)+\psi(i)\right)\right), \tag{4}$$

where the projection to $[\underline{q},\bar{q}]$ ensures that it is consistent with the known range of q.

After estimating the quality of the product, consumer $i+1$ purchases the product if his estimated utility from purchasing, taking into account his quality preference, is higher than the utility from the outside alternative, that is if $\alpha_{i+1}\hat{q}(i+1) - p \geq 0$.

Definition 2.1. 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 (4). Take $l(i) = d(i) = 0$ if consumer $i+1$ observes no reviews (which could happen in the random network).

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

We now discuss different aspects of the formation and interpretation of the decision rule.

Step (a.1): weighted fractions of likes and dislikes. Consumers weigh the information available to them with their prior quality estimate and project it to the appropriate interval. This is motivated by the following. First, in the early stages reviews do not convey much information, since the noise from the random heterogeneity dominates any information about the quality of the product. In the extreme case, after the first consumer's report, one of $l(1)$, $d(1)$ or $o(1)$ will have value one and the other two value zero. It is, therefore, reasonable that consumers do not forget their prior estimate immediately after observing a few reviews, but rather they forget it gradually as they observe more reviews. Second, consumers may simply anchor to the prior until enough information is available. To account for these, we assume that consumer $i+1$ modifies the fraction of likes by placing weight $w/(w+i)$ on $l_0 := \bar{F}(p/q_0)$, which is the fraction of likes corresponding to q_0. In addition, consumer $i+1$ places the same weight on $d_0 = 0$ when modifying the fraction of dislikes, where $d_0 = 0$ since if the prior were true, consumers would have made the right decision for them resulting in no dislikes[8].

Another outcome of the noise is that $l(i)$ and $d(i)$ may take values that do not agree with the long run behavior of the learning process. Define $\underline{l} := \bar{F}(p/\underline{q})$ the fraction of consumers that would *always* purchase and like the product, and $\bar{l} := \bar{F}(p/\bar{q})$ the fraction of consumers that would *never* purchase the product. It follows by a law of large numbers that for a large i the fraction of likes will lie in the interval $[\underline{l}, \bar{l}]$ regardless of the true quality of the product. Similarly it follows that $d(i)$ will lie in $[0, \bar{d}]$ eventually, where $\bar{d} := 1 - \underline{l}$. To account for this, consumers project the fraction of likes and the fraction of dislikes to their corresponding intervals. Note that the weight on $w/(w+i)$ diminishes with i and the projection does not bind for large i, so these corrections are in effect in the short run only.

Step (a.2): correction term for self-selection bias $\psi(i)$. When the quality estimate is lower than the true quality estimate, only liking reviews are reported. This self-selection bias is in effect only when $l_w(i) < l^*$. It is assumed that consumers naively account for that by adding the quantity $\psi(i) \geq 0$ to $l_w(i)$ as in (4). This correction term is a function of $l_w(i)$ and $d_w(i)$. Proposition 3.2 will show that ignoring the self-selection bias may result in biased quality estimated even asymptotically.

We now discuss a few more aspects of the model.

Information structure. The information structure and decision rule have a particularly appealing property: decisions depend on the history of reports only through the number of consumers who liked and disliked the product, and are independent of the sequence in which these reports were

[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 conjugate prior, that induces a linear structure of estimators (see Diaconis and Ylvisaker, 1979). As Raiffa and Schlaifer (1968, Chapter 3) show the weight on the prior estimate can be interpreted as the cardinality of a fictitious sample of observations similar to the ones given in the real sample.

given. We allow for two interpretations for this property, first that the sequence of submission times of reviews is unobservable, and second is that consumers base their decisions on some aggregate measure of their observed information to simplify their decision making. The setting where one observes only the aggregate number of likes and dislikes resembles the information available on many online retailers and review aggregators like Amazon and TripAdvisor. This informational structure usually includes a review scale (typically the 5-star scale) and textual reviews. Consumers observe the number of reviews given to each bin in the scale without the submission time, except for reviews that include text, and even in that case gathering the submission times requires some effort. The information about the number of non-buyers is typically not available. As was discussed in the literature review, Herrera and Hörner (2009) relax the assumption that this number is observable. A similar relaxation could be applied to our model.

Random network. The decision rule generalizes to the random network case by simply considering the number of predecessors that liked, dislikes, and chose the outside alternative in the observed subset. To keep the interpretation of the weight given to the prior estimate as the equivalent of a fictitious sample of size w, consumer i places weight $\rho_i w$ instead of w on the prior estimate.

Pricing problem. In the pricing problem (1) the monopolist chooses the price upfront once and for all, and is assumed to know the true quality of the product, the distribution of quality preferences, and consumers' prior quality estimate. This is reasonable given the setup of the model, as he could test the attractiveness of the product on a test group. A discounted revenue criterion is considered to explore the ability and the incentive of the monopolist to affect the speed of the social learning process via the price.

Price as a signal. It is assumed that consumers are blind to the information on the quality of the product that is conveyed by the pricing decision. Fully rational consumers will regard the price as a signal for the quality, and consequentially the monopolist should take that into account. Since we do not consider a fully rational and Bayesian setting, we disregard this aspect of the pricing decision. Nevertheless, this could be an interesting extension.

Lack of endogenous timing decision. The assumption that consumers make a once and for all purchasing decision upon arrival may be restrictive, as some agents would elect to time their decision based on the availability of information, see Acemoglu, Bimpikis, and Ozdaglar (2010) for an example. In such cases we may see consumers with higher quality preferences purchasing first, which could have an effect on the monopolist's optimal pricing decision. We leave this to future research.

3 Asymptotic Learning

This section considers the asymptotic properties of the decision rule. It is of interest as most of the literature on social learning deals with the asymptotic properties of the learning process, and in particular with whether agents learn the true state of the world eventually. Asymptotic learning entails that eventually all consumers make the right decisions; all consumers with $\alpha_i \geq \alpha^*$, and only those, purchase the product. Subsequently, the fraction of consumers that like the product will converge to $l^* = \bar{F}(\alpha^*)$, the fraction of dislikes will go to zero, and the reminder of consumers will choose the outside alternative.

Definition 3.1. *Asymptotic learning* occurs for some decision rule carried out by consumers if $(l(i), d(i), o(i)) \to (l^*, 0, 1 - l^*)$ almost surely.

We are interested in whether asymptotic learning in the sense of Definition 3.1 occurs under the decision rule.

3.1 No Correction for Self-Selection Bias ($\psi = 0$)

It turns out that the answer to this question depends on the way in which consumers correct for the self-selection bias. The next proposition shows that asymptotic learning does not occur if consumers do not correct for the self-selection bias.

Proposition 3.2. *Asymptotic learning in the sense of Definition 3.1 fails if $\psi(l_w, d_w) = 0$ for all l_w , d_w.*

All proofs are relegated to the appendix. To understand this result, suppose that $l_w(i) \approx l(i)$. When $\psi = 0$ we have $\hat{q}(i+1) \approx p/\bar{F}^{-1}(l(i))$. If $\hat{q}(i+1) < q$, any purchase would result in a like, and this happens with probability $l(i)$. Therefore, $l(i)$, and consequently $\hat{q}(i+1)$, become self-confirming. Thus, learning may stop on an underestimating quality level. This conclusion is not symmetric; learning will not stop on an overestimating quality level, since in such a case dislikes will accumulate and as a result the quality estimate would lower.

To further understand this result, consider the probability that consumer $i+1$ will like the product given $(l(i), d(i))$[9],

$$\begin{aligned}
\mathsf{P}\big(r_{i+1} = r^1 | (l(i), d(i))\big) &= \mathsf{P}\big(\alpha_{i+1}\hat{q}(i+1) - p \geq 0, \alpha_{i+1}q - p \geq 0 | (l(i), d(i))\big) \\
&= \mathsf{P}\big(\alpha_{i+1} \geq \max\left(p/q, p/\hat{q}(i+1)\right)\big) \\
&= \min\left(\bar{F}(p/q), \bar{F}(p/\hat{q}(i+1))\right) \\
&= \min\big(l^*, \hat{l}(i)\big),
\end{aligned} \tag{5}$$

[9]Since $l(i) + d(i) + o(i) = 1$ we often omit $o(i)$.

where $\hat{l}(i) := \bar{F}(p/\hat{q}(i+1)) = \mathrm{proj}_{[\underline{l},\bar{l}]}(l_w(i) + \psi(i))$ is the probability that consumer $i+1$ will purchase the product. For completeness, the probability that $i+1$ will dislike is

$$\begin{aligned} \mathsf{P}\big(r_{i+1} = r^{\mathrm{d}} | (l(i), d(i))\big) &= \mathsf{P}\big(\alpha_{i+1}\hat{q}(i+1) - p \geq 0, \alpha_{i+1}q - p < 0 | (l(i), d(i))\big) \\ &= \big(\hat{l}(i) - l^*\big)^+, \end{aligned} \tag{6}$$

where $(x)^+ = \max(0, x)$. One can think of the difference $\mathsf{P}\big(r_{i+1} = r^{\mathrm{l}} | (l(i), d(i))\big) - l(i)$ as the drift in the learning process. For $l(i)$ to converge to l^*, there should be negative drift when $l(i) > l^*$, i.e., $\hat{q}(i+1) > q$, and positive drift in the other case, i.e., $\hat{q}(i+1) < q$. Assuming again that $l_w(i)$ is close to $l(i)$ and considering (5) with $\psi = 0$, we can see that the drift is not positive when $\hat{q}(i+1) < q$ since in this case $\min(l^*, \hat{l}(i)) - l(i) = \hat{l}(i) - l(i) = \tilde{l}(i) - l(i) \approx 0$.

We can relate this negative result to the notion of herding in social learning. Without correcting for the self-selection bias, consumers act in a way that confirms with the behavior of their predecessors. This hinders accumulation of new information and prevents learning.

3.2 Correcting for the Self-Selection Bias

Proposition 3.2 shows that correcting for the self-selection bias is necessary for asymptotic learning to occur. In order to serve its purpose, the correction term should be positive when $\hat{q} < q$, and otherwise it should be close to zero. A clear sign for $\hat{q} < q$ is that the fraction of disliking reviews is small, since such reviews may be reported only if $\hat{q}(i) > q$ for some i's. Therefore, qualitatively the correction term $\psi(i) = \psi(l_w(i), d_w(i))$ should be positive when $d_w(i)$ is small, and should decrease with $d_w(i)$ as it appears that $\hat{q} > q$. Correcting for the self-selection bias only when $d_w(i) = 0$ will not suffice, since in the early stages of the learning process the quality estimator varies a lot due to the randomness in the quality preferences. Thus, a few disliking reviews may be reported early on, but soon after that learning may stop on an underestimating level. However, $d_w(i)$ will remain positive for all i, even though it decreases to zero.

We assume the following functional form for the correction term,

Assumption 3.3. The correction term $\psi(l_w, d_w) = l_w \phi(d_w/l_w)$, where $\phi(x) = (\phi_0 - \phi_1 x)^+$ with $\phi_0, \phi_1 > 0$.

The assumed correction term is indeed decreasing in the fraction of dislikes, and it is strictly positive when the fraction of dislikes is small[10]. In addition, it has two more properties that make it appealing and also simplify the analysis of the pricing problem:

[10]Propositions 3.6, 4.1 and 4.2 would follow for a larger class of correction terms. For example, any decreasing, Lipshitz continuous ϕ with $\phi(0) > 0$ would suffice.

(i) consumers normalize the fraction of dislikes by the fraction of likes, for example 2% dislikes compared to 10% likes calls for a smaller correction term than 2% dislikes compared to 40% likes; and

(ii) given the normalized magnitude of dislikes, the correction term increases with the popularity of the product, for example the correction term will be twice larger for 2% dislikes and 40% likes than for 1% dislikes and 20% likes.

Before analyzing the asymptotic behavior of the decision rule we add two assumptions.

Assumption 3.4 (Price is not prohibitive)**.** The price charged by the monopolist is not greater than $\bar{p}$ with $\underline{l} = \bar{F}(\bar{p}/\underline{q}) > 0$.

This condition says that even for the lowest possible quality of the product and the highest possible price, some consumers will like to purchase it. It will be clearly satisfied when the support of α_i is unbounded. This assumption is similar to the unbounded belief assumption that is often used in Bayesian social learning. Both imply that new information will enter the system (reviews in our case, actions in the Bayesian case), which is the main requirement for learning to take place.

Assumption 3.5 (Structure of the random social network)**.** For the random network ρ_i is such that $i\rho_i = \Omega(i^\kappa)$, for some $0 < \kappa \leq 1$, and $\rho_i > 0$ for all $i \geq 1$.

Recall that in the random network, for every predecessor j of i, the probability that i observes j is ρ_i (independent of j). The condition ensures that as i grows large, consumers observe enough information. In particular, consumers have to sample enough predecessors (note that the average number of predecessors sampled is $i\rho_i$). If we consider the basic network as a special case of the random network with $\rho_i = 1$ for all $i \geq 0$, then this condition is satisfied. It is interesting to compare this requirement with social networks that permit learning in the Bayesian setting, for example see Acemoglu et al. (2011). In contrast with some of the results there, in our case consumers must observe an increasing number of predecessors. This is mostly due to the heterogeneity in consumers' tastes, and not due the degree of rationality of the learning rule. If tastes are heterogeneous and not identifiable by successors, any finite set of predecessors cannot be fully informative since the predecessors' tastes may dictate their decisions and not the underlying quality level.

Denote by $x^* := \phi_0/(1+\phi_1)$ the unique solution to $\phi(x) = x$, and let $d^* := \min(x^* l^*, \bar{l} - l^*)$. The following result reveals the asymptotic behavior of the decision rule.

Proposition 3.6. *Under Assumptions 3.3, 3.4, and 3.5,* $(l(i), d(i), o(i)) \to (l^*, d^*, 1 - l^* - d^*)$. *For any* $\epsilon > 0$ *there exists a function* ϕ *such that* $d^* < \epsilon$*.*

Proposition 3.6 shows that asymptotic learning in the sense of Definition 3.1 is not achieved, unless $d^* = 0$. This is never the case, since the properties of ϕ show that x^* is always positive, although it could be arbitrarily small depending on ϕ. The proof relays on Kushner and Yin (2003, Theorem 2.1, Chapter 5), who use ordinary differential equations (ODEs) to identify limits of certain continuous time stochastic processes. This method was also used in the context of social dynamics in Ceyhan, Mousavi, and Saberi (2009). In essence, the proof shows that there is always drift in the system in the appropriate direction. For large i $(l(i), d(i)) \approx (l_w(i), d_w(i))$, so from (5) we have

$$\mathsf{P}\big(r_{i+1} = r^1 | (l(i), d(i))\big) - l(i) = \min\big(l^* - l(i), \hat{l}(i) - l(i)\big) \approx \min\big(l^* - l(i), l(i)\phi\big(d(i)/l(i)\big)\big),$$

where the projection is ignored for simplicity, so $\hat{l}(i) - l(i) \approx l(i)(1 + \phi(d(i)/l(i))) - l(i) = l(i)\phi(d(i)/l(i))$. This drift is always negative for $l(i) > l^*$ and positive for $l(i) < l^*$ as long as $d(i)$ is small, which will be the case eventually if $l(i) < l^*$.

To understand why $d(i) \to d^*$, note that consumers keep correcting for the self-selection bias even when $l_w(i)$ is close to l^*. But, when they are close enough, some consumers will end up disliking the product. Since the correction term is decreasing in the fraction of dislikes, more dislikes means a lower correction term, which then brings about fewer dislikes and so on. A fraction d^* of dislikes balances these effects, and thus prevails in the limit of the learning process. To see this, note that a fraction $l^* + l^*\phi(d^*/l^*)$ of consumers will purchase the product in the limit of the learning process. However, only a fraction l^* will like it. Therefore, a fraction $l^*\phi(d^*/l^*)$ of consumers will purchase the product and dislike it, so $l^*\phi(d^*/l^*) = d^*$, or $d^* = x^* l^*$[11].

In principle we could consider nonstationary correction terms $\{\phi_i\}_{i=1}^{\infty}$ that decrease over time to a limiting function $\underline{\phi}$ with corresponding $x_i^* \downarrow \underline{x}^*$. This would reflect the fact that the self-selection bias is resolved over time as $\hat{q}(i)$ approaches q. In this case the fraction of dislikes in equilibrium is $\underline{x}^* l^*$, which could be made arbitrarily small. We say in the reminder of the paper that asymptotic learning is achieved in spite of the persistent fraction of dislikes in equilibrium, with the understanding that d^* is small. We focus on stationary correction functions for simplicity.

We find the existence of a positive fraction of dislikes in asymptotic learning enlightening, since in practice one usually sees a persistent sliver of disliking reviews, even for very popular products. There are many potential explanations for this phenomenon like defects, impulsive shopping, etc. We think that the correction for the self-selection bias on the knife edge may be another explanation.

4 Mean-Field Approximation

The previous section establishes the asymptotic properties of the social learning process. This, however, does not alleviate the analysis of the monopolist's pricing problem (1). Since the proba-

[11] If, $l^* + l^*\phi(x^*) > \bar{l}$ only a fraction $\bar{l}$ will purchase the product, and $d^* = \bar{l} - l^*$.

bility that consumer i will purchase the product depends on the information available to him, (1) involves taking an expectation over all possible histories of the form $r_1, r_2, \ldots$. This renders the problem of choosing the optimal price intractable. In order to solve (1) the learning trajectory and its dependency on the price have to be understood.

Instead of applying numerical methods, we investigate the limit of a sequence of systems with increasing arrival rate of consumers. The learning trajectory of this limiting large scale model is governed by a system of ordinary differential equations that gives rise to a deterministic path for the weighted fractions of likes and dislikes, which in turn determines the decisions of consumers. This system can be solved rather simply for the basic network, and for some cases of the random network. In the reminder of the paper we restrict the attention to the basic network. We comment about the derivation of the mean-field approximation for the random network in Subsection 4.1.

Simulations show that this deterministic learning trajectory approximates sample paths from the original stochastic model very well. As such, we focus our analysis on this deterministic learning trajectory to draw conclusions on the rate of convergence, and to analyze the problem of finding the price that maximizes the expected discounted revenues. Subsection 4.1 constructs the large scale model, motivates the converges result, and establishes it. Subsection 4.2 specifies the learning trajectory and Subsection 4.3 demonstrates the quality of the approximation via simulations. Readers who are not interested in the derivation of the mean-field approximation may go directly to Subsection 4.2.

4.1 Large Market Setting

We consider a sequence of systems indexed by n. In the n-th system consumers' arrival process is Poisson with rate $\Lambda^n := n\bar{\Lambda}$, and the weight placed on the prior quality estimate is $w^n := n\bar{w}$. Recall the interpretation of w as the number of consumers arriving during some time window. The state variables of the n-th system at time t is given by $X^n(t) := \big(L^n(t), D^n(t), O^n(t)\big)$, where $L^n(t)$ is the number of consumers that reported like by time 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. We carry the notation from the previous sections with the necessary adjustments. For example, from (5) we have

$$l_w^n(X^n) := \frac{w^n l_0 + L^n}{w^n + S^n} = \frac{\bar{w} l_0 + \bar{L}^n}{\bar{w} + \bar{S}^n} =: l_w(\bar{X}^n), \tag{7}$$

where $S^n := L^n + D^n + O^n$ and $\bar{S}^n := S^n/n$, and analogously for d_w^n, $\hat{l}^n$, and the corresponding d_w, and $\hat{l}$. Hence, with some abuse of notation we define the functions γ^{l}, γ^{d}, and γ^{o} such that

$$\gamma^s(\bar{X}^n) := \mathsf{P}\left(r_i = r^{\mathrm{s}} | I_i = X^n\right),$$

with the interpretation that $\gamma^{\mathrm{s}}(\bar{X}^n)$ is the probability that a consumer who observes information X^n reports $\mathrm{s} = \mathrm{l}, \mathrm{d}, \mathrm{o}$, for like, dislike, and outside alternative. For example, from (5) and (7) we have

$$\gamma^{\mathrm{l}}(\bar{X}^n) := \mathsf{P}\big(r_i = r^{\mathrm{l}} | I_i = X^n\big) = \min\left(l^*, \hat{l}^n(X^n)\right) = \min\left(l^*, \hat{l}(\bar{X}^n)\right) \tag{8}$$

With this notation in mind, we use a Poisson thinning argument to express the scaled state variable as a Poisson process with time dependent rate,

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

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

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

where $N := (N^{\mathrm{l}}, N^{\mathrm{d}}, N^{\mathrm{o}})$ is a vector of independent Poisson processes with rate 1, and $\gamma := (\gamma^{\mathrm{l}}, \gamma^{\mathrm{d}}, \gamma^{\mathrm{o}})$. Put differently, we break the arrival process to three groups, like, dislike, and outside alternative, according to the probabilities given by γ, that in turn depend on $\bar{X}^n(t)$.

As n grows large, the arrival process becomes deterministic by the functional strong law of large numbers for Poisson processes. Therefore, it is reasonable that the path of $\bar{X}^n(t)$ also converges to its mean. We apply Kurtz (1977/78, Theorem 2.2) to obtain this result.

Proposition 4.1. *For every $t > 0$,*

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

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

$$\bar{X}(t) = \bar{\Lambda} \int_0^t \gamma(\bar{X}(s))\mathrm{d}s. \tag{9}$$

To better understand (9) consider the expression for the scaled number of likes,

$$\bar{L}(t) = \bar{\Lambda} \int_0^t \gamma^1(\bar{X}(s))\mathrm{d}s = \bar{\Lambda} \int_0^t \mathsf{P}\big(r_s = r^1 | I_s = \bar{X}(s)\big)\mathrm{d}s.$$

This means that the scaled number of likes at t is the sum over the mass of consumers that report like in each $s \le t$, and this mass depends on past reviews via $\bar{X}(\cdot)$. It is convenient to derive from (9) the expressions for $(\tilde{l}(t), \tilde{d}(t))$, the counterparts of the stochastic $l_w(i)$ and $l_w(i)$ in the limiting (fluid) model, since these quantities determine the decision of an arriving consumer. Namely, if we denote $\bar{S}(t) := \bar{L}(t) + \bar{D}(t) + \bar{O}(t)$, then from (7) we have

$$\tilde{l}(t) := \mathrm{proj}_{[\underline{l},\bar{l}]}\left(\frac{\bar{w}}{\bar{w} + \bar{S}(t)} l_0 + \frac{\bar{S}(t)}{\bar{w} + \bar{S}(t)} \frac{\bar{L}(t)}{\bar{S}(t)}\right) = \mathrm{proj}_{[\underline{l},\bar{l}]}\left(\frac{\bar{w} l_0 + \bar{L}(t)}{\bar{w} + t\bar{\Lambda}}\right),$$

where the second equality follows from $\bar{S}(t) = t\Lambda$. Similarly,

$$\tilde{d}(t) := \mathrm{proj}_{[\underline{d},\bar{d}]}\left(\frac{\bar{D}(t)}{\bar{w} + t\bar{\Lambda}}\right).$$

The next subsection derives the ODEs that govern $(\tilde{l}(t), \tilde{d}(t))$ for the basic network[12].

4.2 Learning Trajectory

The next proposition derives the ODEs controlling the mean-field approximation for (l_w, d_w) denoted by $(\tilde{l}, \tilde{d})$.

Proposition 4.2. *For the basic network,*

$$\begin{bmatrix} \dot{\tilde{l}}(t) \\ \dot{\tilde{d}}(t) \end{bmatrix} = \frac{\bar{\Lambda}}{\bar{w} + t\bar{\Lambda}} \begin{bmatrix} \min\big(l^*, \hat{l}(t)\big) - \tilde{l}(t) \\ \big(\hat{l}(t) - l^*\big)^+ - \tilde{d}(t) \end{bmatrix}, \tag{10}$$

almost everywhere, where $\hat{l}(t) := \mathrm{proj}_{[\underline{l},\bar{l}]}\big(\tilde{l}(t)(1 + \phi(\tilde{d}(t)/\tilde{l}(t)))\big)$ *is the probability that a consumer would purchase the product at time* t*. The initial conditions are* $\tilde{l}(0) = l_0$ *and* $\tilde{d}(0) = 0$*. Furthermore,* $\tilde{l}$ *is decreasing if* $l_0 < l^*$*, increasing if* $l_0 > l^*$*, and constant if* $l_0 = l^*$*.*

These ODEs are intuitive. A rate $\bar{\Lambda}\hat{l}(t)$ of consumers would purchase at time t. Out of them

[12] The limiting learning trajectory for the random network may be similar or different from the one for the basic network, depending on the sampling probabilities. Let $\rho^n(S)$ be the probability of sampling each predecessor by the $(S+1)$-st consumer. The expected number of predecessors seen by a consumer arriving at t is $\Lambda^n t \rho^n(\Lambda^n t)$, which could range from linear in n to constant in n, depending on the scale of ρ^n. Under general conditions Proposition 4.1 will hold for all these cases, however the limiting trajectory may be different than that of the basic network. If the number of predecessors sampled at time t grows no slower than $\sqrt{n}$, the limiting trajectories coincide. If it grows slower than that, the limiting trajectories may differ.

$\bar{\Lambda}\min(l^*, \hat{l}(t))$ would like the product. So, $\bar{\Lambda}\Big[\min(l^*, \hat{l}(t)) - \tilde{l}(t)\Big]$ is the rate of change in the *number* of likes at time t. To obtain the rate of change of the *fraction*, we normalize this quantity by the mass that arrived by that time, $\bar{w} + t\bar{\Lambda}$. A similar argument follows for dislikes. For a general nonlinear ϕ function, this system of ODEs can be solved numerically. The linearity of ϕ simplifies (10) to a system of nonhomogeneous first order liner ODEs that can be easily solved. To avoid boundary cases, the next assumption requires that the projection on $[\underline{l}, \bar{l}]$ is not binding.

Assumption 4.3. The price charged by the monopolist is no less than $\underline{p}$, where $\max(\bar{F}(\underline{p}/q_0)(1 + \phi_0), \bar{F}(\underline{p}/q)(1 + x^*)) \leq \bar{l}$.

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

Proposition 4.4. *Under Assumption 4.3,*

1. *If $l_0(1 + \phi_0) < l^*$, then for $t \leq T$,*

$$\begin{bmatrix} \tilde{l}(t) \\ \tilde{d}(t) \end{bmatrix} = l_0\Big(\frac{\bar{w} + t\bar{\Lambda}}{\bar{w}}\Big)^{\phi_0} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{11}$$

and for $t > T$

$$\begin{bmatrix} \tilde{l}(t) \\ \tilde{d}(t) \end{bmatrix} = l^* \begin{bmatrix} 1 \\ x^* \end{bmatrix} - l^* \frac{\bar{w} + T\bar{\Lambda}}{\bar{w} + t\bar{\Lambda}} \frac{\phi_0}{\phi_1} \begin{bmatrix} \tilde{\phi} \\ 1 \end{bmatrix} + l^* x^* \Big(\frac{\bar{w} + T\bar{\Lambda}}{\bar{w} + t\bar{\Lambda}}\Big)^{1+\phi_1} \begin{bmatrix} 0 \\ 1/\phi_1 \end{bmatrix}, \tag{12}$$

where $\tilde{\phi} := \phi_1/(1 + \phi_0)$ and

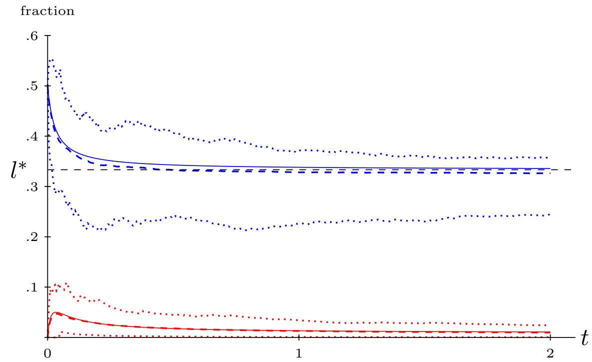
$$T := \frac{\bar{w}}{\bar{\Lambda}}\Big[\Big(\frac{1}{1 + \phi_0}\frac{l^*}{l_0}\Big)^{1/\phi_0} - 1\Big]. \tag{13}$$

2. *If $l_0(1 + \phi_0) \geq l^*$, then*

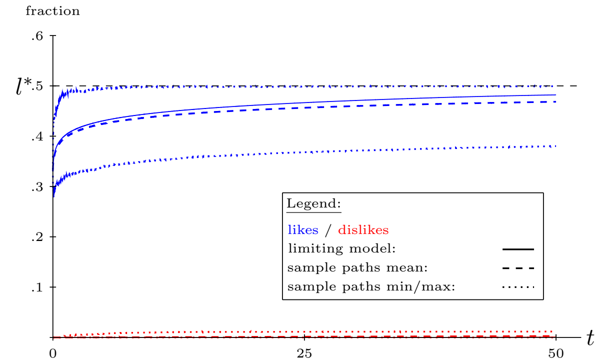
$$\begin{bmatrix} \tilde{l}(t) \\ \tilde{d}(t) \end{bmatrix} = l^* \begin{bmatrix} 1 \\ x^* \end{bmatrix} + (l_0 - l_*)\frac{\bar{w}}{\bar{w} + t\bar{\Lambda}} \begin{bmatrix} 1 \\ \tilde{\phi}^{-1} \end{bmatrix} + \Big(\frac{\bar{w}}{\bar{w} + t\bar{\Lambda}}\Big)^{1+\phi_1} \frac{1}{\phi_1} \begin{bmatrix} 0 \\ l^*(1 + x^*) - l_0(1 + \phi_0) \end{bmatrix}. \tag{14}$$

We call $q_0 < q$ the *underestimating* case and $q_0 > q$ the *overestimating* case. The limiting model differs considerably between the two cases. In the underestimating case typically $l_0(1 + \phi_0) < l^*$[13]. The speed of learning depends on the aggressiveness with which consumers correct the self-selection

[13]If q_0 is close to q from below, or if the price is very low, then it can be that $l_0(1 + \phi_0) \geq l^*$ while $q_0 < q$. This case is of little interest to us since social learning is almost achieved from the very beginning.



(a) Overestimating ($q_0 > q$)



(b) Underestimating ($q_0 < q$)

Figure 2: The blue lines correspond to $l_w(t)$ and the red lines to $d_w(t)$. Model parameters: $\alpha_i \sim U[0, 1]$, $q, q_0 \in \{.3, .4\}$, $p = 0.2$, $\bar{\Lambda} = 1000$, $\bar{w} = 30$, $\phi_0 = 0.05$, and $\phi_1 = 1$. Note the different scales of the horizontal axes. Scaling parameters: $n = 1$ in overestimating and $n = 10$ in underestimating (both with 100 sample paths).

bias as captured by ϕ_0. We suppose that consumers are careful in applying this correction in order to avoid overshooting. Hence it is assumed that ϕ_0 is close to zero and so learning is slow in the underestimating case. In particular, it is proportional to $(l^*/l_0)^{1/\phi_0}$. In this case time T is the first time that everyone that should buy the product indeed buys it, and this leads to a kink in the learning trajectory. In the overestimating case $l_0(1 + \phi_0) \geq l^*$, and social learning is quite fast, as $\tilde{l}$ drops to l^* in a rate of $w/(w + t)$.

The difference in the rates of convergence between these cases is intuitive. The monopolist cannot fool consumers for too 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 correct for the self-selection bias, learning may be slow.

4.3 Simulations

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, $\tilde{l}(t)$ 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, even for a relatively small market ($n = 10$). Figure 2a shows a typical overestimating case in which the $\tilde{l}(t)$ 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, even for a small market ($n = 1$). 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 (note the different time scale in both figures).

5 Static Price

In this section we study the effect of the price chosen by the monopolist on the social learning process. We then proceed to solve the monopolist's problem of choosing a price once and for all in order 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. It is assumed that the monopolist knows the true quality of the product, however the signal conveyed by the price is ignored (see discussion in Subsection 2.2). Section 6 extends the results obtained here to the case of two prices.

5.1 Price and Speed of Learning

We begin by studying the effect of the price on the speed of learning. For a given price of the product, define the *ϵ-time-to-learn* to be $\tilde{T}(\epsilon) := \min\{t \big| \, |\tilde{l}(t)/l^* - 1| \leq \epsilon\}$. Recall that $\tilde{l}(t)$ is the fraction of consumers that like the product at time t weighted by the prior quality estimate, and l^* is the fraction of likes once learning is achieved. Thus, $T(\epsilon)$ is the time to be ϵ away from learning the true quality of the product, measured by fraction of consumers. Note that the first time to see dislikes in the underestimating case is $T = \tilde{T}(\phi_0/(1+\phi_0))$. 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 ϵ-time-to-learn changes with the price.

Proposition 5.1. *The ϵ-time-to-learn is increasing in the price if either the prior underestimates and F is IGFR or if the prior overestimates and F is DGFR. The ϵ-time-to-learn is decreasing in the price if either the prior underestimates and F is DGFR or if the prior overestimates and F is IGFR.*

This proposition suggests that the effect of price on learning is controlled by the distribution of quality preferences and by the relationship between the quality and its prior estimate. The condition on the distribution of the quality preferences is related to its tail, as IGFR distributions have faster decaying tails than Pareto distribution and DGFR have slower decaying tails than Pareto distribution. These conditions determine the effect of price in this way. Suppose that

$q_0 < q$. Then, the social learning process starts with a fraction l_0 of consumers who buy and like the product and ends with a fraction $l^* > l_0$ of consumers who buy and like the product. Increasing the price lowers both l_0 and l^*. If the tail of the preference distribution decays fast, l_0 decreases faster than l^* (in relative terms), so learning takes longer. Conversely, for a slow decaying tail, l^* decreases faster than l_0 and the conclusion is reversed. In the overestimating case $l_0 > l^*$, and the conclusions are reversed again; learning accelerates with price for fast decaying tails, and learning decelerates with price for slow decaying tails.

Denote the *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[14], which in our case entails that the ϵ-time-to-learn increases with the price in the underestimating case, and decreasing with the price in the overestimating case. This condition is important for the two-prices problem of Section 6.

Suppose that we are interested only in the effect of the price on learning, and ignore its effect on the revenue accumulation. Then the monopolist would like to speed up learning if consumers initially underestimate the quality of the product, otherwise slowing down learning would be preferred. It follows from Proposition 5.1 that this can be achieved by lowering the price, i.e., learning will slow down if consumers overestimate the quality, and learning will speed up if consumers underestimate the quality. In the DGFR case, which is uncommon, increasing the price would achieve this in both cases.

5.2 Optimal Static Price

With Proposition 4.4, the seller's discounted revenues in (1) simplify to

$$\pi(p) = p \int_0^\infty e^{-\delta t} \; \bar{F}(p/\hat{q}(t))\mathrm{d}t = p \int_0^\infty e^{-\delta t} \; \Big[(1+\phi_0)\tilde{l}(t) - \phi_1 \tilde{d}(t)\Big]\mathrm{d}t,$$

where for notational clarity we rescale time to make $\bar{\Lambda} = 1$. We denote by p^* the price that maximizes π in the interval $[\underline{p}, \bar{p}]$, i.e., $p^* = \arg\max_{p\in[\underline{p},\bar{p}]} \pi(p)$. The following assumption is common in the literature.

Assumption 5.2. The *revenue function*, $R(p, q)$, is unimodal in p for all q.

With this assumption we can define $p(q) := \max_{p\in\mathbb{R}_+} R(p, q)$, which equates the marginal rev-

[14] It is satisfied by many distributions including exponential, uniform, normal, lognormal, and Pareto (weakly).

enue function, $R'(p,q)$, to zero. The following assumption focuses the pricing problem on the most relevant cases.

Assumption 5.3. 1. The price interval is such that, $\bar{p} \geq \max(p(q), p(q_0))$ and, $0 < \underline{p} \leq \min(p(q), p(q_0))$.

2. In the underestimating case ($q_0 < q$), $\bar{F}(p/q_0)(1+\phi_0) \leq \bar{F}(p/q)$ for all $p \in [\underline{p}, \bar{p}]$.

Assumption 5.3.1 ensures that the price interval is not too small, so that prices that are likely to be in the region of p^* are included in the interval. Assumption 5.3.2 can be interpreted as a condition on the difference between q and q_0 in the underestimating case, as a condition on the magnitude of correction of the self-selection, or as a condition on the price. In all three cases it requires that learning does not occur immediately in the underestimating case. This restricts the analysis to what we think is the interesting case.

We consider the first order conditions of the unconstrained optimization problem, and then check that the conditions of Assumption 5.3 are not violated. For the underestimating case the first order condition is

$$\frac{\partial \pi(p)}{\partial p} = R'(p, q_0)s_1 + R'(p,q)s_2 - [G(p/q_0) - G(p/q)]s_3 = 0, \tag{15}$$

and for the overestimating case it is

$$\frac{\partial \pi(p)}{\partial p} = R'(p, q_0)h_1 + R'(p,q)h_2 = 0, \tag{16}$$

where $\{s_j\}_{j=1}^3$ and $\{h_j\}_{j=1}^2$ are positive and depend on the learning trajectory and on the price (see the proof of Proposition 5.4 for details). These first order conditions have the nice property that they are a mixture of the marginal revenue functions under the prior estimate and under the true quality. In addition, the underestimating case 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 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 5.4. *In the underestimating case $p(q_0) \leq p^* \leq p(q)$, and in the overestimating case $p(q_0) \geq p^* \geq p(q)$.*

This result follows from the first order conditions. Since

$$\underline{p} \leq \min(p(q), p(q_0)) \leq p^* \leq \max(p(q), p(q_0)) \leq \bar{p},$$

the solution to the first order conditions does not violate Assumption 5.3, and this is also the optimal solution to the constrained problem. The result is intuitive, in both the underestimating and the overestimating cases consumers' quality estimate goes from q_0 to q over time. For each estimated quality level along the way a price between $p(q_0)$ and $p(q)$ will be optimal. The optimal price that accounts for the dynamics of social learning remains in this interval. In Section 6 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 static pricing problem is generally not concave, so computation may require a line search, which Proposition 5.4 limits to a relatively small interval. Nevertheless, concavity holds under certain conditions, as we show in the next lemma.

Proposition 5.5. *Assume that the revenue function $R(p, q)$ is concave in p and that F is IGFR, then $\pi(p)$ is concave for all $p \in [p(q_0), p(q)]$ in the underestimating case and $p \in [p(q), p(q_0)]$ in the overestimating case.*

It is not surprising that concavity of R with respect to price is necessary for concavity of π, as this condition is necessary even if the quality was known.

6 Dynamic Pricing

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, either related or unrelated to social learning, can support such pricing strategies. A few examples include learning-by-doing, demand estimation, seasonality, and endogenous timing of the purchasing decision (consumers with high valuations purchase first). These considerations may agree or conflict with the optimal price path we find here. We hope that future research will highlight their interplay. For simplicity, we focus the analysis on one price change.

6.1 Optimal Price Path

Consider the case where the monopolist can adjust her price once. In this case the first price has a dual role, in addition to extracting short term revenue, it controls the speed of the learning process. Therefore, the monopolist may sacrifice revenue in the short run in order to push the learning process in the desired direction.

The social learning procedure has to be adapted to account for the price change, since consumers' reviews depend on the price they were charged. To see this, suppose that the price goes up at some point in time. A consumer that reported like before the price change may report dislike after the price change, as the higher price decreases his net utility from purchase. Thus, we need to specify the information consumers have post-change about the first price and the reviews made under it. Given this information, the post-change decision rule has to be determined. The pre-change decision rule and information structure follow that of the single price model.

We assume the information structure and decision rule that follows most naturally from the static price model. At the time of the price change consumers aggregate all the information conveyed by the reviews posted up to that point to a new prior quality estimate. After the price change, consumers consider only reviews that were made post-change, which are weighted together with the new prior quality estimate that aggregates the information from reviews made pre-change. In addition, the fraction of dislikes, if different than zero, is reinterpreted based on the new price in a way that keeps the normalized fraction of dislikes (d/l) unaffected by the price change. This will give rise to two update equations that will determine the new prior estimates l_1 and d_1. Then at the time of the price change, τ, the learning process would restart with a new single price, and with (l_1, d_1) replacing (l_0, d_0) as prior estimates.

Denote by p_0 and by p_1 the first price and the second price, respectively. The social learning process starts with p_0 and evolves as described in Proposition 4.4 until time τ. For clarity of notation we denote this process up to time τ by $(\tilde{l}_0(t), \tilde{d}_0(t))$. After τ the social learning process restarts, this time with (l_1, d_1) as prior estimates that receive weight $w + \tau$, corresponding to the initial weight and to the mass of reviews that formed those estimates. The initial conditions for $(\tilde{l}_1(t), \tilde{d}_1(t))$, the social learning trajectory after τ, are determined by these prior estimates (l_1, d_1).

Following (4) the estimated quality at time τ, without accounting for the self-selection bias, is $\hat{q}_0 := p_0/\bar{F}^{-1}(\tilde{l}_0(\tau))$. This quality estimate is translated to l_1 in this way,

$$l_1 := \bar{F}(p_1/\hat{q}_0) = \bar{F}\Big(\frac{p_1}{p_0}\bar{F}^{-1}(\tilde{l}_0(\tau))\Big). \tag{17}$$

Consumers do not account for the self-selection bias here to avoid duplication, as it will be accounted for in the decision rule. This equation reveals only part of the information about the system before τ, since it does not take into account the fraction of dislikes, which is related to the magnitude of the self-selection bias. To keep the normalized fraction of dislikes unaffected by the price change, d_1 is given by

$$d_1 := l_1 \frac{\tilde{d}_0(\tau)}{\tilde{l}_0(\tau)}. \tag{18}$$

With the specification of (l_1, d_1), the learning path after τ can be obtained similarly to Proposition

4.4. The two-price profit function is then

$$\pi(p_0,p_1,\tau) = p_0\int_0^{\tau} e^{-\delta t}\big[(1+\phi_0)\tilde{l}_0(t)-\phi_1\tilde{d}_0(t)\big]\mathrm{d}t + p_1\int_{\tau}^{\infty} e^{-\delta t}\big[(1+\phi_0)\tilde{l}_1(t)-\phi_1\tilde{d}_1(t)\big]\mathrm{d}t. \quad (19)$$

We focus on the underestimating case, since there learning is relatively slow and so the first price may have a big impact on the learning process and consequently on the revenue. Recall that in the underestimating case with a static price, T is the first time that dislikes start to appear, which implies that learning was almost achieved ($|\tilde{l}(T)-l^*| \le \phi_0$ which is assumed to be small). Therefore, if $\tau > T$, setting $p_1 = p(q)$, the price that corresponds to the true quality, will be close to optimal, which leaves two effective controls only. Denote by $\bar{T}$ the first time that dislikes appear in the two price case. We concentrate on the most interesting case where $\tau < \bar{T}$, so that all controls are effective. We call this the *regular underestimating* case.

Three regions comprise the learning trajectory in this case: the learning phase before τ, the learning phase after τ and finally the post-learning phase. 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.

An inspection of the pricing problem after time τ reveals that this problem is identical to the static price problem with prior quality $\hat{q}_0 = p_0/\bar{F}(\tilde{l}_0(\tau)) < q$ that receives weight $w+\tau$. The optimal choice of p_0 balances between the short run revenue and its effect on the social learning process. The second effect is summarized by $\hat{q}_0$, the estimated quality at the time of the price change. The first price expedites learning if $\hat{q}_0$ is increasing in p_0, and it delays learning if the converse holds. This effect is summarized in the next lemma.

Lemma 6.1. *The estimated quality at the time of the price change, $\hat{q}_0$, is decreasing (increasing) in p_0 if F is IGFR (DGFR).*

This result complements Proposition 5.1 that shows the relationship between monotonicity of the GFR and the effect of price on the speed of learning. The next proposition analyzes the optimal price path.

Proposition 6.2. *The revenue function of the regular underestimating case is*

$$\begin{aligned}\pi(p_0,p_1,\tau) = {} & p_0\int_0^{\tau} e^{-\delta t} l_0\Big[(1+\phi_0)(\frac{w+t}{w})^{\phi_0}\Big]\mathrm{d}t + p_1\int_{\tau}^{\bar{T}} e^{-\delta t} l_1\Big[(1+\phi_0)(\frac{w+t}{w+\tau})^{\phi_0}\Big]\mathrm{d}t \\ & + p_1\int_{\bar{T}}^{\infty} e^{-\delta t} l_1^*\Big[1+x^*-x^*\Big(\frac{w+T}{w+t}\Big)^{1+\phi_1}\Big]\mathrm{d}t, \end{aligned} \quad (20)$$

where $l_1^* := \bar{F}(p_1/q)$*. Assume* F *is IGFR. Let* p_0^**,* p_1^**, and* τ^* *be the optimal controls of the problem*

$$\max_{\substack{p_0,p_1\in[\underline{p},\bar{p}]\\ \tau\le\bar{T}}} \pi(p_0,p_1,\tau),$$

then $p_0^* < p_1^*$ *and* $p_1^* \in [p(\hat{q}_0), p(q)]$*.*

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.

6.2 Comparison of Pricing Strategies

In this subsection we numerically test the pricing strategies of Sections 5 and 6. 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 strategy the longer it takes consumers to learn the true quality, relative to the monopolist's discount rate of revenues.

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 $\bar{T}$ that are measures of the length of the learning phase in each of these cases. In addition, the revenues of different pricing strategies 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 in the two prices case. The optimal second price in the two prices case is in all cases higher than the optimal static price, and it is very close to $p^{\text{true}} = p(q) = .2$. 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^{true} in the second stage to maximize revenues when consumers learn the true quality.

We report the percentage increase in revenues of the optimal static price over charging p^{true} and

		static price		two prices				% increase in profits			
		p^*	T	p_0^*	p_1^*	τ^*	$\bar{T}$	$\frac{\text{static}}{\text{true}}$	$\frac{\text{static}}{\text{prior}}$	$\frac{\text{two}}{\text{static}}$	$\frac{\text{full}}{\text{two}}$
patient	$q_0 = .35$	0.20	1.81	0.18	0.20	0.50	1.30	0	2	0	0
	$q_0 = .30$	0.19	123.32	0.12	0.20	2.00	9.63	1	5	1	0
	$q_0 = .25$	0.14	396.73	0.10	0.20	7.50	72.85	50	5	10	1
semi-patient	$q_0 = .35$	0.20	1.80	0.18	0.20	0.50	1.31	0	2	0	0
	$q_0 = .30$	0.18	81.07	0.12	0.20	2.00	9.63	2	5	2	0
	$q_0 = .25$	0.14	230.62	0.10	0.19	7.50	61.44	52	3	12	1
im-patient	$q_0 = .35$	0.20	1.80	0.18	0.20	0.50	1.31	0	2	0	0
	$q_0 = .30$	0.17	37.84	0.11	0.20	1.00	8.12	5	4	3	1
	$q_0 = .25$	0.13	138.65	0.10	0.19	7.50	52.30	56	1	13	3

Table 1: Comparison of pricing strategies and revenues.

$p^{\text{prior}} = p(q_0) = q_0/2$, see $\frac{\text{static}}{\text{true}}$ and $\frac{\text{static}}{\text{prior}}$ in the table, respectively. In the patient case and when the prior estimate is close to the true quality, the static price wins the monopolist a few percentage points over these suboptimal prices. If, however, learning takes a long time, the improvement over p^{true} 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 the cases where learning is slow, the two-prices strategy performs at least 10% better than the static price, see $\frac{\text{two}}{\text{static}}$ in the table. In other cases the improvement is smaller to nonexistent.

Lastly, we compare the performance of the two-prices strategy against the *full information* case, a scenario when consumers *know* the true quality, and it is optimal to charge p^{true}, see $\frac{\text{full}}{\text{two}}$ in the table. Interestingly, the two-prices strategy recovers almost all the revenue of the full information case, with a loss of 3% at most. We can think of this quantity as the cost to the monopolist of the underestimation in consumers' prior estimate. By accounting for this in designing the pricing strategy, the monopolist is able to recover almost the entire loss, even with two prices only.

7 Conclusions

Motivated by the increasing importance of social review systems in consumers' decisions, this paper introduces a model of social learning. The contribution of this paper is a social learning model with heterogeneous preferences, whose asymptotic learning properties and learning dynamics are characterized. This is achieved by a fluid model that bridges the literatures on social learning and on social dynamics. We hope that this framework will be used to study other social learning dynamics in settings with many agents. Our analysis of the pricing problems contributes to the relatively sparse literature in pricing and social learning by: (a) deriving a crisp characterization of the relationship between the price and the speed of learning, (b) deriving structural results for the optimal pricing strategies with one and two prices, (c) evaluating the performance of these pricing strategies in numerical experiments.

The assumption that consumers observe predecessors that chose not to purchase the product is perhaps the most limiting assumption in the model. We discuss one approach to relax it in Section 2. We conjecture that it could also be done by including more informative reviews. This would enable consumers to learn something about the predecessor's private quality preference.

Several interesting directions of future analysis arise. One natural extension concerns endogenous timing of the purchasing decision. Naturally, consumers who elect to postpone the purchasing decision will observe more information, yet waiting might be costly. It would be interesting to introduce this new dimension to the consumers decision rule. Another possible extension would consider a joint learning framework, where consumers learn the quality of the product and the monopolist learns the distribution of preference parameters. This is related to the literature in revenue management that balances demand estimation and revenue extraction. In addition, one could consider the role of price as a signal for the quality. This would be more naturally implemented in a Bayesian setting.

References

ACEMOGLU, D., BIMPIKIS, K., and OZDAGLAR, A. (2010) Dynamics of information exchange in endogenous social networks. Working Paper 16410, National Bureau of Economic Research.

ACEMOGLU, D., DAHLEH, M. A., LOBEL, I., and OZDAGLAR, A. (2011) Bayesian learning in social networks. *Rev. Econom. Stud.* Forthcoming.

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

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

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

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

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

Bergemann, D. and Välimäki, J. (2000) Experimentation in markets. *Rev. Econom. Stud.* **67**, 213–234.

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

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

Bitran, G., Caldentey, R., and Vial, R. (2005) Pricing policies for perishable products with demand substitution. Mimeo.

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

Bühlmann, H. and Gisler, A. (2005) *A Course in Credibility Theory and its Applications.* Springer-Verlag, Berlin.

Caminal, R. and Vives, X. (1999) Price dynamics and consumer learning. *J. Econom. & Management Strat.* **8**, 95–131.

Campbell, A. (2009) Tell your friends! Word of mouth and percolation in social networks. Mimeo.

Candogan, O., Bimpikis, K., and Ozdaglar, A. E. (2011) Optimal pricing in networks with externalities. Eprint, arXiv:1101.5617.

Ceyhan, S., Mousavi, M., and Saberi, A. (2009) Social infuence and evolution of market share. In *Workshop on the Economics of Networked Systems (NetEcon 2009).*

Chamley, C. and Gale, D. (1994) Information revelation and strategic delay in a model of investment. *Econometrica* **62**, 1065–1085.

Diaconis, P. and Ylvisaker, D. (1979) Conjugate priors for exponential families. *Ann. Statist.* **7**, 269–281.

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

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

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

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

Golub, B. and Jackson, M. O. (2010) Naïve learning in social networks and the wisdom of crowds. *Amer. Econ. J. Microeconomics* **2**, 112–149.

Hartline, J., Mirrokni, V., and Sundararajan, M. (2008) Optimal marketing strategies over social networks. In *Proceeding of the 17th international conference on World Wide Web*, WWW '08, 189–198. ACM, New York, NY, USA.

Herrera, H. and Hörner, J. (2009) Biased social learning. Technical report, SSRN eLibrary. URL `http://ssrn.com/paper=1499221`.

Jadbabaie, A., Molavi, P., Sandroni, A., and Tahbaz-Salehi, A. (2011) 'Non-Bayesian Social Learning', Third Version. *SSRN eLibrary* .

Kempe, D., Kleinberg, J., and Tardos, E. (2003) Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '03, 137–146. ACM, New York, NY, USA.

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

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

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

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

Lee, I. H. (1993) On the convergence of informational cascades. *J. Econom. Theory* **61**, 395–411.

Papastavrou, J. and Athans, M. (1992) Distributed detection by a large team of sensors in tandem. *Aerospace and Electronic Systems, IEEE Transactions on* **28**, 639 –653.

Raiffa, H. and Schlaifer, R. (1968) *Applied Statistical Decision Theory*. The M.I.T. Press, Cambridge, Mass.-London.

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

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

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

Tsitsiklis, J. (1984) Problems in Decentralized Decision Making and Computation. Ph.D. thesis, Laboratory for Information and Decision Systems, MIT.

Tsitsiklis, J., Bertsekas, D., and Athans, M. (1986) Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *Automatic Control, IEEE Transactions on* **31**, 803 – 812.

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

Vives, X. (1997) Learning from others: a welfare analysis. *Games Econom. Behav.* **20**, 177–200.

Welch, I. (1992) Sequential sales, learning, and cascades. *J. Finance* **47**, 695–732.

A Proofs

A.1 Preliminaries

The following notation will be used in several proofs. Denote by $\theta = (l, d) \in [0,1]^2$. With some abuse of notation, define

$$l_w(\theta) := \mathrm{proj}_{[\underline{l},\bar{l}]}(l), \tag{21}$$

$$d_w(\theta) := \mathrm{proj}_{[\underline{d},\bar{d}]}(d), \text{ and} \tag{22}$$

$$\hat{l}(\theta) := \mathrm{proj}_{[\underline{l},\bar{l}]}\big(l_w(\theta)(1 + \phi(d_w(\theta)/l_w(\theta)))\big). \tag{23}$$

which are to the weighted fractions of likes and dislikes, and the corresponding probability that a consumer will purchase the product. It is also useful to define the equivalent with i consumers,

$$l_w(\theta, i) := l_w\big(((wl_0 + il)/(w+i)\big), \tag{24}$$

$$d_w(\theta, i) := d_w\big(id/(w+i)\big), \text{ and} \tag{25}$$

$$\hat{l}(\theta, i) := \mathrm{proj}_{[\underline{l},\bar{l}]}\big(l_w(\theta, i)(1 + \phi(d_w(\theta, i)/l_w(\theta, i)))\big). \tag{26}$$

Throughout this appendix for a vector $x = (x^1, \ldots, x^k)$ we denote by $|x| = \|x\|_1 = \sum_{j=1}^k |x^j|$, and for a univariate function $x(t)$ we denote by $x'(t)$ the derivative with respect the variable t. To reduce the notational burden we rescale time such that $\Lambda = \bar{\Lambda} = 1$. The first lemma is used repeatedly in different proofs.

Lemma A.1. *1.* $|\min(x,y)-\min(z,y)| \le |x-y|$ *and* $|\max(x,y)-\max(z,y)| \le |x-z|$.

2. There exist a $\hat{\Gamma}<\infty$ *such that* $|\hat{l}(\theta)-\hat{l}(\theta')| \le \hat{\Gamma}|\theta-\theta'|$.

3. For $x_1, x_2, y_1, y_2 \ge 0$ *and* $z>0$,

$$\Big|\frac{x_1}{z+x_1+y_1}-\frac{x_2}{z+x_2+y_2}\Big| \le \frac{1}{z}[|x_1-x_2|+|y_1-y_2|].$$

Proof. 1. Minimum operator: If $x,z=y$ or $x,z\le y$ this holds trivially. If $x\le y$ and $z\ge y$ then $|\min(x,y)-\min(z,y)|=|x-y|=y-x\le z-x=|x-z|$. For the maximum operator take $-x$, $-y$, and $-z$.

2. Let $\theta=(l,d)$ and $\theta'=(l',d')$. From the triangular inequality, and from Lemma A.1.1,

$$\begin{aligned}|\phi(d_w(\theta)/l_w(\theta)-\phi(d_w(\theta')/l_w(\theta'))| &\le \phi_1|d_w(\theta)/l_w(\theta)-d_w(\theta')/l_w(\theta')|\\ &=\phi_1|d_w(\theta)l_w(\theta')-d_w(\theta')l_w(\theta)|/l_w(\theta)l_w(\theta')\\ &\le(\phi_1/\underline{l}^2)|d_w(\theta)l_w(\theta')-(d_w(\theta')-d_w(\theta)+d_w(\theta))l_w(\theta)|\\ &\le(\phi_1/\underline{l}^2)\bar{d}|l_w(\theta')-l_w(\theta)|+\frac{1}{\underline{l}^2}\bar{l}|d_w(\theta')-d_2(\theta)|\\ &\le(\phi_1\max(\bar{d},\bar{l})/\underline{l}^2)|\theta-\theta'|.\end{aligned}$$

Now,

$$\begin{aligned}|\hat{l}(\theta)-\hat{l}(\theta')| &\le |l_w(\theta)-l_w(\theta')|+l_w(\theta)|\phi(d_w(\theta)/l_w(\theta)-\phi(d_w(\theta')/l_w(\theta'))|\\ &\quad+\phi(d_w(\theta')/l_w(\theta'))|l_w(\theta)-l_w(\theta')|\\ &\le(1+\phi_0)|l-l'|+(\phi_1\max(\bar{d},\bar{l})/\underline{l}^2)|\theta-\theta'|\\ &\le\hat{\Gamma}|\theta-\theta'|,\end{aligned}$$

where the first inequality follows from the triangular inequality, and the second follows from Lemma A.1.1, $|\phi(x)|\le\phi_0$ and $\hat{\Gamma}:=1+\phi_0+\phi_1\max(\bar{d},\bar{l})/\underline{l}^2$.

3. From the triangular inequality,

$$\begin{aligned}\Big|\frac{x_1}{z+x_1+y_1}-\frac{x_2}{z+x_2+y_2}\Big| &\le \Big|\frac{x_1}{z+x_1+y_1}-\frac{x_1}{z+x_1+y_2}\Big|+\Big|\frac{x_1}{z+x_1+y_2}-\frac{x_2}{z+x_2+y_2}\Big|\\ &=\frac{x_1}{w+x_1+y_1}\Big|\frac{y_1-y_2}{w+x_1+y_2}\Big|+\frac{w+y_2}{w+x_1+y_2}\Big|\frac{x_1-x_2}{w+x_2+y_2}\Big|\\ &\le\frac{1}{w}[|x_1-x_2|+|y_1-y_2|].\end{aligned}$$

□

A.2 Proofs

Proof of Proposition 3.2. The proof relays on the derivation of the probability of reporting a liking review in (5). For convenience suppose $\underline{l} = 0$, $\bar{l} = 1$ and $l_0 < l^*$. Then,

$$\begin{aligned}
\mathsf{E}[l(i+1)] &= \mathsf{E}[\mathsf{E}[l(i+1)|l(i)]] \\
&= \frac{1}{i+1}\mathsf{E}[il(i) + \mathsf{E}[\mathbf{1}\{r_{i+1} = r^{\mathrm{l}}\}|l(i)]] \\
&= \frac{1}{i+1}\mathsf{E}[il(i) + \min(l^*, (wl_0 + il(i))/(w+i))] \\
&\leq \frac{1}{i+1}\mathsf{E}[il(i) + (wl_0 + il(i))/(w+i)],
\end{aligned}$$

where the third inequality follows by assumption and by (5). Note that $\mathsf{E}[l(1)] = l_0$ (recall $l(0) = 0$). It follows that

$$\mathsf{E}[l(2)] = \frac{1}{2}\mathsf{E}[l(1) + (wl_0 + l(1))/(w+1)] = l_0$$

and by iterating this further we get $\mathsf{E}[l(i)] = l_0$ for all $i \geq 1$. Since $l_0 < l^*$ and $l(i) \in [0,1]$ this completes the proof, since if $l(i) \to l^*$ then the mean would have converged to l^* as well. We comment that for large i the weight of l_0 vanishes over i, so with positive probability $l(i) < l^*$ even for $l_0 > l^*$ and the result follows. Similarly, the assumption on $\underline{l}$ and $\bar{l}$ can be removed. $\square$

Lemma A.2 (Theorem 2.3. of Kushner and Yin (2003))**.** *Let $\{\theta(i) : i \geq 0\}$ in $H = \prod_{j=1}^{r}[a_j, b_j]$ be a multivariate stochastic process satisfying,*

$$\theta(i+i) = \theta(i) + \epsilon_i Y(i), \tag{27}$$

where $Y(i) = M(i) + \beta(i) + g(\theta_i, i)$ is composed of a martingale difference noise $M(i) = Y(i) - \mathsf{E}[Y(i)|\theta_0, Y(j), j < i]$, a random variable $\beta(i)$, and a measurable function of $g(\cdot, i)$. Assume that,

1. *$\sum_i \mathsf{E}|Y(i)|^2 < \infty$.*

2. *$g(\cdot, i)$ are continuous uniformly in i, and for each $\theta \in H$*

$$\lim_{i\to\infty} \left\| \sum_{j=i}^{m(t_i+t)} (g(\theta, j) - g(\theta))/(j+1) \right\|_2 = 0, \tag{28}$$

   where $m(x) := \{i | t_i \leq t < t_{i+1}\}$.

3. *$\epsilon_i \geq 0$, $\sum_i \epsilon_i = \infty$, and $\sum_i \epsilon_i^2 < \infty$.*

4. *$\beta(i) \to 0$ w.p. 1.*

If $A_H \subset H$ is locally asymptotically stable in the sense of Lyapunov for ODE $\dot{\theta}(t) := \partial\theta(t)/\partial t = g(\theta(t))$ and $\theta(i)$ is in some compact set in the domain of attraction of A_H infinitely often with probability at least ν, then $\theta(i) \to A_H$ with at least probability ν.

Proof of Proposition 3.6. We begin by proving the second statement of the proposition. Note that $x^* = \phi_0/(1+\phi_1)$. Fix $\epsilon > 0$ and choose $\phi_1 > (-1+\phi_0/\epsilon)^+$.

The proof of the first statement relays on Kushner and Yin (2003, Theorem 2.3, Chapter 5), which is repeated for convenience in Lemma A.2. We start by writing the processes $l(i)$ and $d(i)$ as iterative processes with martingale difference noise and a drift term. Then we use the theorem to show that these processes converge to a set containing solutions of a certain ODE. We characterize the unique solution of this ODE on the subset $C = \{(l,d) \in [0,1]^2 \big|\ l+d \le 1\}$ using a Lyapunov function. For reference on stability of autonomous systems (Lyapunov theorem) see Khalil (2002). We consider the basic network first and then comment on the random network case.

Denote by $\theta(i) = (l(i), d(i)) \in [0,1]^2$ the state of the system. If follows that,

$$l(i+1) = l(i) + \frac{1}{i+1}\delta M_1(i) + \frac{1}{i+1} g_1(\theta(i), i), \tag{29}$$

where $\delta M_1(i) := \mathbf{1}\{r_{i+1} = r^{\mathrm{l}}\} - \mathsf{E}[\mathbf{1}\{r_{i+1} = r^{\mathrm{l}}\}|\theta(i)]$ is a martingale difference noise, and

$$g_1(\theta, i) := \mathsf{E}[\mathbf{1}\{r_{i+1} = r^{\mathrm{l}}\}|\theta(i)] - l = \min(l^*, \hat{l}(\theta, i)) - l, \tag{30}$$

by (2) and (5). Similarly,

$$d(i+1) = d(i) + \frac{1}{i+1}\delta M_2(i) + \frac{1}{i+1} g_2(\theta(i), i), \tag{31}$$

where $\delta M_2(i) := \mathbf{1}\{r_{i+1} = r^{\mathrm{d}}\} - \mathsf{E}[\mathbf{1}\{r_{i+1} = r^{\mathrm{d}}\}|\theta(i)]$ is a martingale difference noise, and

$$g_2(\theta, i) := \mathsf{E}[\mathbf{1}\{r_{i+1} = r^{\mathrm{d}}\}|\theta(i)] - d = (\hat{l}(\theta, i) - l^*)^+ - d \tag{32}$$

by (3) and (6). It is convenient to use the notation of Kushner and Yin (2003), $Y_j(i) = g_j(\theta(i), i) + \delta M_j(i)$ for $j \in \{1,2\}$ with the vector notation $Y(i) = (Y_1(i), Y_2(i))$, $g(\theta, i) = (g_1(\theta, i), g_2(\theta, i)$, and $\delta M(i) = (\delta M_1(i), \delta M_2(i))$. A close inspection of the drift functions $g(\theta, i)$ shows that they converge to a function $g(\theta)$ that govern the drift of the process in the limit as $i \to \infty$. Thus, we define

$$g(\theta) = \begin{bmatrix} g_1(\theta) \\ g_2(\theta) \end{bmatrix} = \begin{bmatrix} \min(l^*, \hat{l}(\theta)) - l \\ (\hat{l}(\theta) - l^*)^+ - d \end{bmatrix}. \tag{33}$$

We now verify the conditions of the theorem:

1. $\sup_i \mathsf{E}|Y(i)|^2 \leq 4 < \infty$ since $|g(\cdot; i)| < 1$ and $|\delta M| \leq 1$.

2. $\mathsf{E}[Y(i)|\theta(i)] = g(\theta, i)$ by construction of the martingale difference noise.

3. $g(\theta)$ is continuous since ϕ is continuous and $\underline{l} > 0$ (Assumptions 3.4 and 3.3).

4. $\sum_{i\geq 1} \frac{1}{i+1} = \infty$ and $\sum_{i\geq 1}(\frac{1}{i+1})^2 < \infty$.

5. Lemma A.3 will show that,

$$\lim_{i\to\infty} \Big\| \sum_{j=i}^{m(t_i+t)} (g(\theta, j) - g(\theta))/(j+1) \Big\|_2 = 0, \tag{34}$$

   where $m(x) := \{i | t_i \leq t < t_{i+1}\}$.

We apply Theorem 2.3 of Kushner and Yin (2003) to conclude that $\theta(i)$ converges almost surely to the set of limit trajectories of the mean limit ODE,

$$\dot{\theta} = g(\theta). \tag{35}$$

Denote by A the set of solutions to (35). Therefore, we need to show that $A = \{\theta^*\}$, where $\theta^* := (l^*, d^*)$.

Define the candidate Lyaponuv function,

$$V(\theta) = B\frac{1}{2}(l - l^*)^2 + K\frac{1}{2}(d - x^* l + z^*)^2,$$

where $B := 1+B_1^2+K(1+\phi_0)$ with $B_1 := \big(l^*(1+x^*)-\underline{l}(1+\phi_1)\big)/\big((1+x^*)^2(l^*-\underline{l})\big)$, $K := 4/(1+x^*)^2$, and $z^* := ((1+x^*)l^* - \bar{l})^+$. In the reminder of the proof assume that d^* is interior, so that $z^* = 0$, i.e., assume

$$l^*(1+x^*) \leq \bar{l}.$$

The boundary case with $z^* > 0$ follows immediately. The gradient of V is,

$$\nabla V(\theta) = \big(B(l - l^*) - x^* K(d - x^* l), K(d - x^* l)\big).$$

We show that $\dot{V}(\theta) = \nabla V(\theta)^\top g(\theta) < 0$ for every $\theta \in C$ and $\dot{V}(\theta^*) = 0$. This will verify that θ^* is the unique solution of (35) in C by the theorem of Lyapunov. We consider the cases $\hat{l}(\theta) \geq l^*$ and $\hat{l}(\theta) < l^*$ for which

$$g(\theta) = \begin{cases} (l^* - l, \hat{l}(\theta) - l^* - d), & \text{if } \hat{l}(\theta) \geq l^* \\ (\hat{l}(\theta) - l, -d), & \text{if } \hat{l}(\theta) < l^*. \end{cases}$$

Case a: $\hat{l}(\theta) \geq l^*$**.**

$$\begin{aligned}\dot{V}(\theta) &= -B(l^*-l)^2 - x^*K(d-x^*l)(l^*-l) + K(d-x^*l)(\hat{l}(\theta)-l^*-d)\\ &= -B(l^*-l)^2 - K(d-x^*l)^2 - K(d-x^*l)(l^*(1+x^*)-\hat{l}(\theta)).\end{aligned}$$

Since the first two terms are non-positive, we focus the attention on the third term.

Case a.1: $d \leq x^*l$**.** Suppose $l \leq \bar{l}$. Then $\phi(d_w(\theta)/l_w(\theta)) \geq x^*$. If $l(1+x^*) \leq \bar{l}$, then the third term is $-K(d-x^*l)(l^*(1+x^*)-\hat{l}(\theta)) < -(1+x^*)K(d-x^*l)(l^*-l) = -2\sqrt{K}(d-x^*l)(l^*-l)$, and $\dot{V}(\theta) < -(l^*-l+K(d-x^*l))^2 - (B-1)(l^*-l)^2 < 0$ with equality only if $\theta = \theta^*$. If, however, $l(1+x^*) > \bar{l}$, then the third term is $-K(d-x^*l)(l^*(1+x^*)-\hat{l}(\theta)) \leq -K(d-x^*l)(l^*(1+x^*)-\bar{l})$, which is negative by assumption that d^* is interior.

Suppose $l > \bar{l}$ then the third term is $-K(d-x^*l)(l^*(1+x^*)-\bar{l})$ which is again nonpositive by assumption that d^* is interior.

Case a.2: $d > x^*l$**.** Suppose $d \leq \bar{d}$ and $l \geq \underline{l}$. Then, $\phi(d_w(\theta)/l_w(\theta)) < x^*$. The third term is

$$\begin{aligned}-K(d-x^*l)\big(l^*(1+x^*)-\hat{l}(\theta)\big) &< -K(d-x^*l)\big(l^*(1+x^*)-\mathrm{proj}_{[\underline{l},\bar{l}]}(l(1+x^*))\big)\\ &< -(1+x^*)K(d-x^*l)(l^*-l),\end{aligned}$$

and we can again complete a sum of perfect squares to show $\dot{V}(\theta) < 0$.

Suppose $l < \underline{l}$ (this includes the case $d > \bar{d}$ since $l + d \leq 1$). The third term is less than $-(1+x^*)K(d-x^*l)(l^*-\underline{l}(1+\phi_0)/(1+x^*)) \leq -(1+x^*)KB_1(d-x^*l)(l^*-\underline{l}) < -(1+x^*)KB_1(d-x^*l)(l^*-l)$, and the a perfect square can be completed to show $\dot{V}(\theta) < 0$.

Case b: $\hat{l}(\theta) < l^*$**.** This implies $l < l^* < \bar{l}$ since $\hat{l}(\theta) \geq l$ for all θ. Therefore, the time derivative of V is

$$\dot{V}(\theta) = B(l-l^*)(-\hat{l}(\theta)+l) - K(d-x^*l)(x^*\hat{l}(\theta)-x^*l+d).$$

Note that the first term is nonpositive.

Case b.1: $d \geq x^*l$**.** Here $\dot{V}(\theta) < 0$, since $\hat{l}(\theta) \geq l$ for all $l \leq \bar{l}$.

Case b.2: $d < x^* l$**.**

$$\begin{aligned}\dot V(\theta) &< B(l-l^*)(\hat l(\theta)-l) - K(x^*l)(x^*\hat l(\theta))\\ &\le -x^*l\big(B(\hat l(\theta)-l) - Kx^*\hat l(\theta)\big)\\ &\le x^*l\big(B(\max(\underline{l},l)(1+x^*)-l) - Kx^*\max(\underline{l},l)(1+\phi_0)\big)\\ &\le (x^*)^2 l\max(\underline{l},l)\big(B-K(1+\phi_0)\big)\\ &< 0,\end{aligned}$$

where the first inequality follows from $-K(d-x^*l)^2 < 0$ and $-Kdx^*\hat l(\theta) < 0$. To see the second inequality note that since $l \le \bar l$ we have $\phi(d_w(\theta)/l_w(\theta)) \ge \phi(d/l) > x^*$. So, $l^* > \hat l(\theta) \ge l(1+x^*)$, and $l - l^* > -x^*l$. The third inequality follows from the same inequality, from $l < \bar l$ and from $\phi \le \phi_0$. The last step follows by construction of B and K.

We conclude that in all cases $\dot V(\theta) < 0$ for $\theta \in C$ expect for $\dot V(\theta^*) = 0$. Therefore, $A = \{\theta^*\}$ is locally asymptotically stable for $\theta \in C$. Since C is visited infinitely often w.p.1 $\theta(i) \to \theta^*$ w.p.1.

Random network. When consumers sample predecessors the proof has to be modified to account for the randomness in consumers' observations. Given $\theta(i)$ and i, define by $L^s(i)$, $D^s(i)$, $O^s(i)$ the number of likes, dislikes, and no-purchases observed by consumer i. Note that $L^s(i)$ is a binomial random variable with probability ρ_i and $L(i)$ trials (analogously for dislikes and no-purchases). Denote by $l^s(i) = L^s(i)/S^s(i)$ and by $d^s(i) := D^s(i)/S^s(i)$ the sampled fraction of likes and of dislikes, respectively, where $S^s(i) = L^s(i) + D^s(i) + O^s(i)$ (take $l^s(i) = d^s(i) = 0$ if no predecessor is sampled). Define the shorthand notation $\theta^s(i) = (l^s(i), d^s(i))$. By (5),

$$\begin{aligned}\mathsf{E}\big[\mathbf{1}\{r_{i+1} = r^{\mathrm{l}}\}|\theta(i)\big] &= \mathsf{E}\big[\mathsf{E}\big[\mathbf{1}\{r_{i+1} = r^{\mathrm{l}}\}|L^s(i), D^s(i), O^s(i)\big]|\theta(i)\big]\\ &= \mathsf{E}\big[\min(l^*, \hat l(\theta^s(i), i)|\theta(i)\big],\end{aligned}$$

and similarly for dislikes,

$$\mathsf{E}\big[\mathbf{1}\{r_{i+1} = r^{\mathrm{d}}\}|\theta(i)\big] = \mathsf{E}\big[(\hat l(\theta^s(i) - l^*)^+|\theta(i)\big].$$

We modify (29) and (31) to account for the random samples,

$$\theta(i+1) = \theta(i) + \frac{1}{i+1}\delta M(i) + \beta(i) + \frac{1}{i+1}g(\theta(i); i),$$

where

$$\beta(i) = \begin{bmatrix} \beta_1(i) \\ \beta_2(i) \end{bmatrix} := \begin{bmatrix} \mathsf{E}[\mathbf{1}\{r_{i+1} = r^{\mathrm{l}}\}|\theta(i)] - \min(l^*, \hat{l}(\theta(i))) \\ \mathsf{E}[\mathbf{1}\{r_{i+1} = r^{\mathrm{d}}\}|\theta(i)] - (\hat{l}(\theta(i)) - l^*)^+ \end{bmatrix}.$$

In other words $\beta_1(i)$ is the difference between the probability that consumer i will like the product in the random network, and the probability that consumer i will like the product in the basic network, and analogously $\beta_2(i)$ is the difference in probabilities of a dislike under the different networks. By Kushner and Yin (2003, Theorem 2.3, Chapter 5) it suffices to show that $\beta_j(i) \to 0$ for $j = 1, 2$ to conclude that $\theta(i) \to \theta^*$. Under Assumption 3.5 it can be shown that the sampled fractions of likes and dislikes are close to their corresponding fractions in the population for large i. Coupled with the Lipschitz continuity of $\hat{l}$ (Lemma A.1.2), it can be shown that $\beta \to 0$ almost surely. □

Lemma A.3. *Let $g(\theta) = (g_1(\theta), g_2(\theta))$ and $g(\theta, i) = (g_1(\theta, i), g_2(\theta, i))$ as defined in* (30),(32)*, and* (33)*. The following holds for every $\theta \in [0, 1]^2$, and every $t > 0$,*

$$\lim_{i\to\infty} \Big\| \sum_{j=i}^{m(t_i+t)} (g(\theta, j) - g(\theta))/(j + 1) \Big\|_2 = 0, \tag{36}$$

where $m(x) := \{i|t_i \le t < t_{i+1}\}$.

Proof. Note that

$$g(\theta, i) = g\bigg(\frac{wl_0 + il}{w + i}, \frac{id}{w + i}\bigg).$$

It follows that,

$$\begin{aligned} \Big|g_1(\theta, i) - g_1(\theta)\Big| &= \Bigg|\min\Bigg(l^*, \hat{l}\bigg(\frac{wl_0 + il}{w + i}, \frac{id}{w + i}\bigg)\Bigg) - l - \min(l^*, \hat{l}(\theta) + l\Bigg| \\ &\le \hat{\Gamma}\Bigg[\bigg|\frac{wl_0 + il}{w + i} - l\bigg| + \bigg|\frac{id}{w + i} - d\bigg|\Bigg] \\ &\le 2\hat{\Gamma}\frac{w}{w + i}, \end{aligned}$$

where the first inequality follows from Lemma A.1.2, and the second inequality from $0 \le d, l_0, l \le 1$.

Similarly, we conclude that $|g_2(\theta; i) - g_2(\theta)| \le 2\hat{\Gamma}/(w+i)$, and

$$\begin{aligned}
\left\| \sum_{j=i}^{m(t_i+t)} (g(\theta,j) - g(\theta))/(j+1) \right\|_2 &\le \sum_{j=i}^{\infty} \left\| g(\theta,j) - g(\theta) \right\|_2 /(j+1) \\
&\le 2\hat{\Gamma} \sum_{j=i}^{\infty} \frac{1}{(w+j)(1+j)} \\
&\le 2\hat{\Gamma} \sum_{j=i}^{\infty} \frac{1}{(\min(w,1)+j)^2} \\
&\le 2\hat{\Gamma} \frac{1}{w-1+i},
\end{aligned}$$

where the first inequality follows from the triangular inequality, and the last inequality follows from an integral bound. □

Lemma A.4 (Theorem 2.2 of Kurtz (1977/78))**.** *Let $\bar{X}^n(t)$ be a Markov chain with state space $\{k/n | k \in \mathbb{Z}^v\}$ for $n \ge 1$. Assume there exist nonnegative functions γ and $\{\gamma^n\}_{n\ge 1}$ from $\mathbb{R}^v$ to $\mathbb{R}^v$ such that for each $x, y \in \mathbb{R}^v$,*

$$\gamma^n(x) \le \Gamma_1(1+|x|), \tag{37}$$

$$|\gamma^n(x) - \gamma(x)| \le \frac{\Gamma_2}{n}(1+|x|), \tag{38}$$

and,

$$|\gamma(x) - \gamma(y)| \le \Gamma_3 |x-y|, \tag{39}$$

for some finite positive constants Γ_1, Γ_2, and Γ_3. Further assume that $\bar{X}(t)$ satisfies,

$$\bar{X}^n(t) = \bar{X}^n(0) + N\left(\int_0^t \gamma^n(\bar{X}^n(s))\mathrm{d}s\right)/n, \tag{40}$$

where N is a vector of v independent Poisson processes with rate 1, $\bar{X}^n(0)$ is independent of N, and $\bar{X}^n(0) \to \bar{X}(0)$. Then for every $\bar{t} > 0$,

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

where $\bar{X}(t)$ is the deterministic limit,

$$\bar{X}(t) = \bar{X}(0) + \int_0^t \gamma(\bar{X}(s))\mathrm{d}s.$$

Proof of Proposition 4.1. We verify the conditions of Lemma A.4 for the basic network. Since $X^n(t) \in \mathbb{Z}_+^3$, $\bar{X}^n(t) = X^n(t)/n \in \{k/n | k \in \mathbb{Z}_+^3\}$. The integral form of $\bar{X}^n(t)$ in (40) follows from Poisson arrivals and Poisson thinning. For example, $\bar{L}^n(t)$ can be written in the form,

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

where A^n is a Poisson process with rate n and, with some abuse of notation, r_s is a review given by a consumer arriving at time s. The second equality follows from splitting the Poisson process to likes, dislikes, and outside options, based on the probability that an arriving consumer would submit each type of review, which depends on his quality preference and on his observable information $\bar{X}^n(s)$. The property of Poisson processes known as Poisson thinning 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 $O^n(t)$. Condition (37) holds for $\Gamma_1 = 1$ since γ^{s} are probabilities for s = l, d, o, hence $0 \leq \gamma \leq 1$. Condition (38) holds since $\gamma^n = \gamma$ (see (8)). We derive condition (39) for γ^{l},

$$\begin{aligned}|\gamma^{\mathrm{l}}(\bar{X}_1) - \gamma^{\mathrm{l}}(\bar{X}_2)| &\leq \left|\hat{l}\left(\frac{\bar{w}l_0 + \bar{L}_1}{\bar{w} + \bar{S}_1}, \frac{\bar{D}_1}{\bar{w} + \bar{S}_1}\right) - \hat{l}\left(\frac{\bar{w}l_0 + \bar{L}_2}{\bar{w} + \bar{S}_2}, \frac{\bar{D}_2}{\bar{w} + \bar{S}_2}\right)\right| \\ &\leq \hat{\Gamma}\left[\left|\frac{\bar{w}l_0 + \bar{L}_1}{\bar{w} + \bar{S}_1} - \frac{\bar{w}l_0 + \bar{L}_2}{\bar{w} + \bar{S}_2}\right| + \left|\frac{\bar{D}_1}{\bar{w} + \bar{S}_1} - \frac{\bar{D}_2}{\bar{w} + \bar{S}_2}\right|\right] \\ &\leq \hat{\Gamma}\left[wl_0\left|\frac{1}{\bar{w} + \bar{S}_1} - \frac{1}{\bar{w} + \bar{S}_2}\right| + \left|\frac{\bar{L}_1}{\bar{w} + \bar{S}_1} - \frac{\bar{L}_2}{\bar{w} + \bar{S}_2}\right| + \left|\frac{\bar{D}_1}{\bar{w} + \bar{S}_1} - \frac{\bar{D}_2}{\bar{w} + \bar{S}_2}\right|\right] \\ &\leq \hat{\Gamma}\Bigg[|\bar{S}_1 - \bar{S}_2| + \frac{1}{\bar{w}}\big[|\bar{L}_1 - \bar{L}_2| + |\bar{D}_1 + \bar{O}_1 - \bar{D}_2 - \bar{O}_2| \\ &\quad + |\bar{D}_1 - \bar{D}_2| + |\bar{L}_1 + \bar{O}_1 - \bar{L}_2 - \bar{O}_2|\big]\Bigg] \\ &\leq \frac{(\bar{w} + 2)}{\bar{w}}\hat{\Gamma}\big[|\bar{X}_1 - \bar{X}_2|\big],\end{aligned}$$

where the first and second inequalities follow from Lemmas A.1.1-2, the third and last inequalities follow from the triangular inequality, and the forth inequality follows from Lemma A.1.3 since $\bar{S}_j = \bar{L}_j + \bar{D}_j + \bar{O}_j$ for $j = 1, 2$. Similarly, one can show that γ^{d} and γ^{o} are Lipschitz continuous (recall that $\gamma^{\mathrm{l}}(\bar{X}) + \gamma^{\mathrm{d}}(\bar{X}) + \gamma^{\mathrm{o}}(\bar{X}) = 1$). Finally, in our case $\bar{X}^n(0) = 0$ for all $n \geq 1$, which completes the proof for the basic network case. □

Proof of Proposition 4.2. We first do this for the case where $\tilde{l}(t) \in [\underline{l}, \bar{l}]$ and $\tilde{d}(t) \in [\underline{d}, \bar{d}]$ for all $t \geq 0$, so that some of the projections can be omitted. Then we show that the solution remains in $C := [\underline{l}, \bar{l}] \times [\underline{d}, \bar{d}]$. When the projections are dropped

$$
\begin{aligned}
\dot{\tilde{l}}(t) &= \frac{\bar{L}'(t)}{\bar{w} + \bar{\Lambda} t} - \frac{\bar{\Lambda}}{\bar{w} + t\bar{\Lambda}} \frac{\bar{w} l_0 + \bar{L}(t)}{\bar{w} + \bar{\Lambda} t} \\
&= \frac{\bar{\Lambda}}{\bar{w} + \bar{\Lambda} t} \big[\gamma^{\mathrm{l}}(\bar{X}(t)) - \tilde{l}(t) \big] \\
&= \frac{\bar{\Lambda}}{\bar{w} + \bar{\Lambda} t} \big[\min(l^*, \hat{l}(t)) - \tilde{l}(t) \big],
\end{aligned}
$$

where $\bar{L}'(t) = \gamma^{\mathrm{l}}(\bar{X}(t))$ follows from (9). The derivative exists almost everywhere since $\bar{X}(t)$ is absolutely continuous and thus so is $\bar{L}(t)$. To see this observe that $\big|\bar{X}(t) - \bar{X}(s)\big| \leq \bar{\Lambda}\big|t - s\big|$ since $0 \leq \gamma \leq 1$. The initial condition is $\tilde{l}(0) = l_0$ since $\bar{L}(0) = 0$. Similarly,

$$
\dot{\tilde{d}}(t) = \frac{\bar{\Lambda}}{\bar{w} + \bar{\Lambda} t} \big[(\hat{l}(t) - l^*)^+ - \tilde{d}(t) \big]
$$

by substituting for the definition of γ^{d} with initial condition $\tilde{d}(0) = 0$ since $\bar{D}(0) = 0$. It remains to show that $(\tilde{l}(t), \tilde{d}(t))$, as defined by these ODEs, remains in C for all $t \geq 0$. First note that the initial conditions are in C. In addition that for $\tilde{l}(t) \geq l^*$, $\dot{\tilde{l}}(t) \leq 0$, for $\tilde{l}(t) < l^*$, $\dot{\tilde{l}}(t) \geq 0$. Finally, $\tilde{l}$ is continuous and converges to $l^* \in C$, which shows that $\tilde{l}(t) \in [\underline{l}, \bar{l}]$ for all $t \geq 0$. Moreover, $0 \leq \tilde{d}(t) \leq 1 - \tilde{l}(t) \leq 1 - \underline{l} = \bar{d}$, which completes the proof. $\square$

Proof of Proposition 4.4. If $l_0(1 + \phi_0) < l^*$ then $\dot{\tilde{d}}(0) = \big((l_0(1 + \phi_0) - l^*)^+ - \tilde{d}(0)\big)/w = 0$ and subsequently $\tilde{d}(t) = 0$ for t small. Therefore, for small t, $\dot{\tilde{l}}(t) = \tilde{l}(t)\phi_0/(w + t)$ with the solution $\tilde{l}(t) = l_0((w + t)/w)^{\phi_0}$. At time T (see (13)) this is no longer the case since $\tilde{l}(T)(1 + \phi_0) = l^*$, and dislikes appear for the first time. We first guess that after time T $\bar{l} \geq \tilde{l}(t)(1 + \phi(\tilde{d}(t)/\tilde{l}(t))) \geq l^*$ and then check that is indeed the case. This is a system of first order nonhomogeneous liner ODE and the solution is given in (12). It is then easy to check that indeed $\bar{l} \geq \tilde{l}(t)(1 + \phi(\tilde{d}(t)/\tilde{l}(t))) \geq l^*$.

If $l_0(1 + \phi_0) < l^*$ then $\dot{\tilde{d}}(0) > 0$. We guess that $\bar{l} \geq \tilde{l}(t)(1 + \phi(\tilde{d}(t)/\tilde{l}(t))) \geq l^*$, solve the system of ODEs and verify that the condition holds. $\square$

Proof of Proposition 5.4. Consider the first order conditions (15) and (16) with

$$
s_1 := \int_0^T (1 + \phi_0) e^{-\delta t} \Big(\frac{w + t}{w} \Big)^{\phi_0} \mathrm{d}t,
$$

$$
s_2 := \int_T^\infty e^{-\delta t} \Big[1 + x^* - x^* \Big(\frac{w + T}{w + t} \Big)^{1 + \phi_1} \Big] \mathrm{d}t,
$$

$$s_3 := \bar{F}(p/q) \int_T^\infty e^{-\delta t} \Big(\frac{w+T}{w+t}\Big)^{1+\phi_1} \mathrm{d}t,$$

$$h_1 := \int_0^\infty e^{-\delta t} (1+\phi_0) \Big(\frac{w+t}{w}\Big)^{-1-\phi_1} \mathrm{d}t,$$

and

$$h_2 := \int_0^\infty e^{-\delta t} (1+x^*) \Big(1 - \Big(\frac{w+t}{w}\Big)^{-1-\phi_1}\Big) \mathrm{d}t.$$

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

First we show that $q < q'$ implies $p(q) \le p(q')$. Take $q < q'$ and assume by contradiction $p(q) > p(q')$. Note that $R'(p,q) = \bar{F}(p/q) + (p/q)f(p/q)$. By unimodality, $0 > R'(p(q), q') = R'(p(q)q/q', q) > R'(p(q), q) = 0$, where the second inequality follows from $p(q)q/q' < p(q)$. This is a contradiction, hence $p(q) \le p(q')$.

In the overestimating case, $R'(p', q_0)$ and $R'(p', q)$ are both positive for all $p' < p(q)$, and are both negative for all $p' > p(q_0)$. Since $h_1 > 0$ and $h_2 > 0$, we conclude that the optimal price must lie in the specified interval. In the underestimating case for $p' < p(q_0)$ we have

$$\begin{aligned}
\frac{\partial \pi(p)}{\partial p}\bigg|_{p=p'} &= R'(p', q_0)s_1 + R'(p', q)s_2 + \bar{F}(p/q)[r(p/q_0) - r(p/q)])s_3 \\
&> \big[R'(p', q) - F(p/q)[G(p/q_0) - G(p/q)]\big]\big[s_2 - s_3\big] \\
&\ge \big[R'(p', q) - F(p/q)[1 - G(p/q)]\big]\big[s_2 - s_3\big] \\
&= 0,
\end{aligned}$$

where the first inequality follows since $R'(p', q_0) > 0$, which also implies that $G(p'/q_0) < 1$. This together with the fact that $s_2 \ge s_3$ justifies the second inequality. For $p' > p(q)$

$$\begin{aligned}
\frac{\partial \pi(p)}{\partial p}\bigg|_{p=p'} &< \big[R'(p', q) - F(p/q)[G(p/q_0) - G(p/q)]\big]\big[s_2 - s_3\big] \\
&\le \big[R'(p', q) - F(p/q)[1 - r(p/q)]\big]\big[s_2(t) - s_3(t)\big] \\
&= 0,
\end{aligned}$$

where the first inequality follows since $R'(p', q_0) < 0$, which also implies that $G(p'/q_0) > 1$, and that justifies the second inequality. Therefore, in this case too, the optimal price lies in the specified interval. Since $\underline{p} \le \min(p(q), p(q_0)) \le p^* \le \max(p(q), p(q_0)) \le \bar{p}$, the global optimum is also the local optimum in the specified price interval. $\square$

Proof of Lemma 5.5. Denote by $R''(p,q) := \partial^2/\partial p^2 R(p,q) = -2f(p/q)/q - f'(p/q)p/q^2$, which is nonpositive by assumption. In the overestimating case

$$\frac{\partial^2 \pi(p)}{\partial p^2} = R''(p, q_0)h_1 + R''(p, q)h_2 < 0,$$

since $h_1, h_2 > 0$. In the underestimating case,

$$\begin{aligned}
\frac{\partial^2 \pi(p)}{\partial p^2} &= R''(p, q_0)s_1 + R''(p, q)s_2 \\
&\quad + \Big[f(p/q)G(p/q_0)/q - \bar{F}(p/q)G'(p/q_0)/q_0 + f(p/q)/q + f'(p,q)p/q^2\Big]s_3(t) \\
&\quad + \int_T^\infty e^{-\delta t}\Big[-x^* R'(p,q) - \bar{F}(p/q)[G(p/q_0) - G(p/q)]\Big]\frac{\partial T}{\partial p}(1+\phi_1)\frac{(w+T)^{\phi}2}{(w+t)^{1+\phi_1}}\mathrm{d}t \\
&\quad + e^{-\delta T}\Big[R'(p, q_0)\bar{F}(p/q)/\bar{F}(p/q_0) - R'(p,q) + \bar{F}(p/q)G(p/q_0) - (p/q)f(p/q)\Big] \\
&\leq \Big[R''(p,q) + f(p/q)G(p/q_0)/q - \bar{F}(p/q)G'(p/q_0)/q_0 + f(p/q)/q + f'(p,q)p/q^2\Big]s_3(t),
\end{aligned}$$

where the first equality follows by substituting for T in s_2 and s_3, and $G'(x) = \mathrm{d}G(x)/\mathrm{d}x$. The inequality follows from nonpositivity of R'', the fact that $s_2 > s_3 \geq 0$, from nonnegativity of $\partial T/\partial p$ under IGFR, and from the fact the last term after the equality equals zero. It suffices to focus on the terms in brackets,

$$\begin{aligned}
&R''(p,q) + f(p/q)G(p/q_0)/q - \bar{F}(p/q)G'(p/q_0)/q_0 + f(p/q)/q + f'(p,q)p/q^2 \\
&= -f(p/q)/q + f(p/q)G(p/q_0)/q - \bar{F}(p/q)G'(p/q_0)/q_0 \\
&\leq f(p/q)/q\Big[-1 + G(p/q_0) - G'(p/q_0)p/q_0\Big] \\
&= f(p/q)/q\Big[-1 + G(p/q_0) - p/q_0\frac{f(p/q_0) + (p/q_0)f'(p/q_0)}{\bar{F}(p/q_0)} - (G(p/q_0))^2\Big] \\
&\leq f(p/q)/q\Big[-1 + G(p/q_0) + G(p/q_0) - (G(p/q_0))^2\Big] \\
&= -f(p/q)/q\Big[G(p/q_0) - 1\Big]^2 \\
&\leq 0,
\end{aligned}$$

where the first inequality follows from $G' \geq 0$ by assumption IGFR and since $R'(p,q) = \bar{F}(p/q) - (p/q)f(p/q) \geq 0$ since $p \in [p(q_0), p(q)]$ by assumption. The second equality follows by substituting for $G'(p/q_0)$, and the second inequality follows from concavity, $-f(p/q_0) - (p/q_0)f'(p/q_0) \leq f(p/q_0)$. □

Proof of Lemma 6.1. Note that,

$$\frac{\partial \bar{F}^{-1}(\tilde{l}_0(\tau))}{\partial p_0} = \frac{f(p_0/q_0)/q_0}{f(\bar{F}^{-1}(\tilde{l}_0(\tau)))}\left(\frac{w+t}{w}\right)^{\phi_0},$$

since whenever a function, h, and its inverse are differentiable, $\mathrm{d}h^{-1}(x)/\mathrm{d}x = -1/h'(h^{-1}(x))$. Therefore,

$$\begin{aligned}\frac{\partial \hat{q}_0}{\partial p_0} &= C(p_0)\Big[\bar{F}^{-1}(\tilde{l}_0(\tau))f(\bar{F}^{-1}(\tilde{l}_0(\tau))) - f(p_0/q_0)(p_0/q_0)\big(\frac{w+t}{w}\big)^{\phi_0}\Big] \\ &= C(p_0)\Big[\bar{F}(\bar{F}^{-1}(\tilde{l}_0(\tau)))G(\bar{F}^{-1}(\tilde{l}_0(\tau))) - \bar{F}(p_0/q_0)\big(\frac{w+t}{w}\big)^{\phi_0}G(p_0/q_0)\Big] \\ &= C(p_0)\tilde{l}_0(\tau)\Big[G(\bar{F}^{-1}(\tilde{l}_0(\tau))) - G(p_0/q_0)\Big],\end{aligned}$$

where $C(p_0) = (\bar{F}^{-1}(\tilde{l}_0(\tau)))^{-2}/f(\bar{F}^{-1}(\tilde{l}_0(\tau))) > 0$, and G is the generalized failure rate. The second equality above follows by dividing and multiplying the first term by $\bar{F}(\bar{F}^{-1}(\tilde{l}_0(\tau)))$, so it becomes $\bar{F}(x)xf(x)/\bar{F}(x) = \bar{F}(x)G(x)$, where $x = \bar{F}^{-1}(\tilde{l}_0(\tau))$. The second term follows similarly. Since $\bar{F}^{-1}(\tilde{l}_0(\tau)) < p_0/q_0$ for $\tau > 0$, and IGFR is assumed, we conclude that $\hat{q}_0$ is decreasing in p_0. □

Lemma A.5. *If F is IGFR, then $\partial\bar{T}_1/\partial\tau \le (\ge) 0$ if $p_0 \le (\ge) p_1$.*

$$\frac{\partial \bar{T}_1}{\partial \tau} = \Big(\frac{1}{1+\phi_0}\frac{l_1^*}{l_1}\Big)^{1/\phi_0}\Big[1 - G\Big(\frac{p_1}{p_0}\bar{F}^{-1}(\tilde{l}_0(\tau))\Big)\Big/G\Big(\bar{F}^{-1}(\tilde{l}_0(\tau))\Big)\Big]$$

Proof. First compute

$$\frac{\partial l_1}{\partial \tau} = l_0\frac{p_1}{p_0}\phi_0\Big(\frac{(w+t)^{\phi_0-1}}{w^{\phi_0}}\Big)f\Big(\frac{p_1}{p_0}\bar{F}^{-1}(\tilde{l}_0(\tau))\Big)\Big/f\Big(\bar{F}^{-1}(\tilde{l}_0(\tau))\Big),$$

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

$$\begin{aligned}\frac{\partial \bar{T}_1}{\partial \tau} &= \Big(\frac{1}{1+\phi_0}\frac{l_1^*}{l_1}\Big)^{1/\phi_0}\Big[1 - \frac{w+\tau}{\phi_0 l_1}\frac{\partial l_1}{\partial \tau}\Big] \\ &= \Big(\frac{1}{1+\phi_0}\frac{l_1^*}{l_1}\Big)^{1/\phi_0}\Big[1 - \frac{\frac{p_1}{p_0}\bar{F}^{-1}(\tilde{l}_0(\tau))f\Big(\frac{p_1}{p_0}\bar{F}^{-1}(\tilde{l}_0(\tau))\Big)}{l_1}\frac{\bar{F}(\bar{F}^{-1}(\tilde{l}_0(\tau)))}{\bar{F}^{-1}(\tilde{l}_0(\tau))f(\bar{F}^{-1}(\tilde{l}_0(\tau)))}\Big] \\ &= \Big(\frac{1}{1+\phi_0}\frac{l_1^*}{l_1}\Big)^{1/\phi_0}\Big[1 - G\Big(\frac{p_1}{p_0}\bar{F}^{-1}(\tilde{l}_0(\tau))\Big)\Big/G\Big(\bar{F}^{-1}(\tilde{l}_0(\tau))\Big)\Big].\end{aligned}$$

Since F is IGFR, then G is nondecreasing and the result follows. □

Proof of Proposition 6.2. Define the first time dislikes would appear under p_0,

$$T_0 = w\Big(\frac{1}{1+\phi_0}\frac{l_0^*}{l_0}\Big)^{1/\phi_0} - w.$$

This is identical to T in the static price. In order for $\tau \le \bar{T}$, the constraint $\tau \le T_0$ surely must be imposed. Define by T_1 the first time dislikes would appear under p_1, assuming they did not appear at time τ,

$$T_1 = (w+\tau)\Big(\frac{1}{1+\phi_0}\frac{l_1^*}{l_1}\Big)^{1/\phi_0} - w.$$

Therefore, $\bar{T} = \max(\tau, T_1)$. The optimization problem assumes that $\bar{T} = T_1$. It is later shown that this is indeed the case. In addition, the constraints on the prices are ignored, except for $p_0 \geq \underline{p}$, and their validity is later verified. The proof establishes the result for the unconstraint case and the $p_0^* = \underline{p}$ case, and then shows by contradiction that $p_0^* \geq p_1^*$ is not possible when the constraint $\tau \leq T_0$ is binding. The associated Lagrangian is,

$$\mathscr{L} = \pi(p_0, p_1, \tau) + \mu_1[T_0 - \tau] + \mu_2[p_0 - \underline{p}],$$

where $\mu_1, \mu_2 \geq 0$. The revenue after τ follows the same structure of the static pricing problem with initial weight $w+\tau$ and prior quality estimate $\hat{q}_0$. Moreover, p_1 does not interact with the constraints. Thus, we conclude that $p_1^* \in [p(\hat{q}_0), p(q)]$. In the reminder of this proof we will use the notation c_k to denote a nonnegative quantity (which may depend on the controls) unless otherwise mentioned. Computation shows that

$$\frac{\partial \mathscr{L}}{\partial p_0} = R'(p_0, q_0)c_1 + \frac{\partial l_1}{\partial p_0}c_2 - \frac{\partial T_1}{\partial p_0}c_3 + \mu_1\frac{\partial T_0}{\partial p_0} + \mu_2.$$

Recall that $l_1 = \bar{F}(p_1/\hat{q}_0)$ and that $\partial\hat{q}_0/\partial p_0 < 0$ (Lemma 6.1). 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 constrain $p_0 \geq \underline{p}$ is binding, then $p_0^* = \underline{p} < p(q_0)$ by Assumption 5.3. So, in both cases $p_0^* < p(q_0) < p_1^*$. Thus, we need to check that $p_0^* \geq p_1^*$ is not possible when $\tau^* = T_0$. The first order condition for τ is

$$\frac{\partial \mathscr{L}}{\partial \tau} = e^{-\delta\tau}(1+\phi_0)[p_0\tilde{l}_0(\tau) - p_1 l_1] - c_5 - c_6\frac{\partial T_1}{\partial \tau} - \mu_1.$$

Suppose that $p_0^* \geq p_1^*$, then $p_1^* l_1 = R(p_1^*, \hat{q}_0) = R(p_1^*, p_0^*/\bar{F}^{-1}(\tilde{l}_0(\tau))) > R(p_0^*, p_0^*/\bar{F}^{-1}(\tilde{l}_0(\tau))) = p_0^*\bar{F}(\bar{F}^{-1}(\tilde{l}_0(\tau))) = p_0^*\tilde{l}_0(\tau)$. Also from Lemma A.5, $\partial T_1/\partial\tau \geq 0$. We conclude that $\partial\mathscr{L}/\partial\tau < 0$, 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, $\underline{p} \leq p_0^* < p_1^* \leq p(q) < \bar{p}$. To conclude the proof we show that $T_1 = \bar{T}$, by showing $T_1 \geq \tau$. Note that $\bar{F}(p/q)/\bar{F}(p/q')$ is increasing in p for $q > q'$ and F IGFR. Therefore,

$$\begin{aligned}
\tau &\leq w\Big(\frac{1}{1+\phi_0}\frac{\bar{F}(p_0^*/q)}{\bar{F}(p_0^*/\hat{q}_0)}\Big)^{1/\phi_0} - w \\
&\leq (w+\tau)\Big(\frac{1}{1+\phi_0}\frac{\bar{F}(p_1^*/q)}{\bar{F}(p_1^*/\hat{q}_0)}\Big)^{1/\phi_0} - w \\
&= T_1,
\end{aligned}$$

which concludes the proof. □