Pricing Models in the Presence of Informational and Social Externalities

Davide Crapis

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

under the Executive Committee

of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2016

ABSTRACT

Pricing Models in the Presence of Informational and Social Externalities

Davide Crapis

This thesis studies three game theoretic models of pricing, in which a seller is interested in optimally pricing and allocating her product or service to a market of agents, in order to maximize her revenue. These markets feature a large number of self-interested agents, who are generally heterogeneous with respect to some payoff relevant feature, e.g., willingness to pay when agents are consumers or private cost when agents are firms. Agents strategically interact with one another, and their actions affect other agents' payoffs, either directly through social influence or competition, or indirectly through a review system. The seller has demand uncertainty and strives to optimize expected discounted revenues. I use either a mean-field approximation or a continuum of agents assumption to reduce the complexity of the problems and provide crisp characterizations of system aggregates and equilibrium policies.

Chapter 2 considers the problem of an information provider who sells information products, such as demand forecasts, to a market of firms that compete with one another in a downstream market. We propose a general model that subsumes both price and quantity competition as special cases. We characterize the optimal selling strategy and find that it depends on both mode and intensity of competition. Several important extensions to heterogeneous production costs, information quality discrimination, and information leakage through aggregate actions are studied.

Chapter 3 considers the problem of optimally extracting a stream of revenues from a sequence of consumers, who have heterogeneous willingness to pay and uncertainty about the quality of the product being sold. Using an intuitive maximum likelihood procedure, we characterize the solution of consumers' quality estimation problem. Then, we use a mean-field approximation to characterize

the transient dynamics of quality estimates and demand. These allow us to simplify and solve the monopolist's problem, and ultimately provide a characterization of her optimal pricing policy.

Chapter 4 considers the problem of a seller who is interested in dynamically pricing her product when consumers' utility is influenced by the mass of consumers that have purchased in the past. Two scenarios are studied, one in which the monopolist has commitment power and one in which she does not. We characterize the optimal selling strategy under both scenarios and derive comparisons on equilibrium prices and demands. Our main result is a characterization of the value of price commitment as a function of the social influence level in the market.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my co-advisors, Professor Costis Maglaras and Professor Alireza Tahbaz-Salehi, for their continuous support, for challenging me to think critically about research questions, and for the many hours they spent in sharing and transferring their knowledge and expertise. In particular, I am grateful to Professor Maglaras for introducing me to the fields of Social Learning and Revenue Management, a process that started with the summer project in my first year at Columbia and resulted in the publication of Chapter 3 of this thesis. Working with him has been an invaluable learning experience through which I improved my technical skills, my ability to collaborate, and my ability to attack difficult problems, sometimes failing, but always learning and improving. I am extremely grateful to Professor Tahbaz-Salehi, who has introduced me to the microeconomics of information transmission and has taught me to appreciate and value academic rigor. Our joint work is condensed in Chapter 2 of this thesis, which is also the result of a collaboration with Professor Kostas Bimpikis. I am grateful to both of them for pushing me over my boundaries and teaching me to always challenge and improve upon my results. For these reasons and for the extra effort he put in advising me remotely I would like to thank Professor Bimpikis.

I would like to thank Bar Ifrach and Professor Marco Scarsini for their friendship and many stimulating discussions which greatly contributed to the completion of Chapter 3 of this thesis. I am very thankful to Professor Scarsini for his continuos support, especially in the early years of my life at Columbia. I am also thankful to Lijian Lu, for being a true friend and for his contribution to Chapter 4 of this thesis. I am thankful to Professor Bob Phillips and Professor Gabriel Weintraub for following my research work and being on my thesis committee. More generally, I am thankful to the entire faculty of the DRO division at Columbia Business School, for being an inspirational group of scholars and for creating a stimulating learning environment for Ph.D. students.

This thesis marks the end of a long journey, which I shared with my family and many friends. I am grateful to my beloved parents, Luciano and Patrizia, for being always there, and to my brother and friend Niccolo' for being an unlimited source of optimism. Along the way, I met my beloved girlfriend Lucia, I am thankful to her for supporting me in the hardest times and for sharing the most joyous moments in the best times, and I am very excited for the new journey we are embarking on. Thanks to Enrico, Caner, and the friends with whom I spent so many days at Columbia, thanks to Cinar and the friends that often worked with me in the division, thanks to Sacha and all the American friends for our amazing times together, thanks to Fabio and the friends from Milan for calling and coming to visit, thanks to Agostino and my family in Toronto for their generosity, thanks to Francesco and my friends from Lamezia for being family. I am very grateful to have all these amazing people in my life.

Dedicated to the memory of my grandfather, Giuseppe Ferrise.

Chapter 1

Introduction

Pricing and revenue optimization is a growing and expanding area in the field of operations research. It comprises a set of models and tools that allow a seller, business or individual, to scientifically approach the problem of setting prices, which is the interface between her product or service and the market. Initially developed to solve the classic problems of airline, hotel, and retail pricing, these models have recently evolved to address novel questions that arise in the internet economy where there is a constant release of information about product characteristics and consumer preferences and choices. New areas of application include high-frequency auction mechanisms in on-line advertising, customized pricing, and more broadly dynamic pricing in online markets. Recent advances in mobile technologies, opened yet another novel and promising area of application which is pricing in two-sided markets. In these types of markets there is usually a company that sets prices, or provides prices recommendation to sellers, and the exchange happens between independent buyers and sellers on the platform. In this context pricing is the mean of both demand and supply management. New models that will be developed will likely build on the latter set of models on dynamic pricing in online markets, to which part of this thesis contributes.[1]

Microeconomics is a discipline that starts by modeling the behavior of individual agents who interact in a market, and builds up to make statements and predictions about aggregate economic

[1]For an overview of this field and a treatment of its typical problems see Talluri and van Ryzin [2005] and Phillips [2005].

outcomes. Game theory is the subset of models that deal with situations in which agents strategically interact, i.e., situations in which the actions of an agent directly affect the payoff of other agents. For example, a monopolistic seller changing her present or future price, or two firms who compete by lowering their prices in order to attract more demand. Game theoretic models have been applied extensively in recent years to online markets. On the practical side, algorithmic game theory drives most of the online advertising industry. On the theoretical side, games of incomplete information and games with a large number of agents have been a prominent area of research, because they allow to understand drivers and outcomes of agents' decision in online markets. In particular, performing *mean-field approximations* or building models with a *continuum of agents* are common practices that dramatically simplify the analysis. These models abstract from reality since they imply that a single agent has a negligible impact on other agents' payoffs, however they are very powerful because of their computational tractability and they allow to draw insightful predictions about system behavior in large and complex markets. Finally, I would argue that these models are plausible for studying online markets, where there are many agents and their individual impact on system aggregate is small.

This thesis draws tools from and strives to contribute to the literature in both game theoretic models and revenue optimization. Although the three problems being studied are different in nature, there are at least few common themes that characterize all the models in this thesis that is worth emphasizing. There is a seller that interacts with a large market of agents, consumers or other businesses, and seeks to optimally allocate and/or price its product or service. Agents' payoffs are affected by the actions of other agents, either directly through social externalities or competition, or indirectly through some information release mechanism. Agents' heterogeneity, and the ability of the seller to optimize over it, plays a key role in determining optimal pricing and equilibrium strategies. Finally, we always use either a mean-field approximation or a continuum of agents assumption to reduce aggregate uncertainty and generate crisp insights on the structure of optimal policies for the seller. In the remainder of this chapter I provide a more detailed introduction to the content of each chapter.

Chapter 2 is motivated by the growing interest in markets for information and studies the problem of optimal sale of information, such as demand forecasts, to a set of competing firms. We show that the nature and intensity of competition among the information provider's potential customers have a first-order impact on her optimal selling strategy and profits. More specifically, our analysis illustrates that the value the provider can extract from her customers is largely determined by the trade-off between (i) the direct (positive) effect that more precise information has on their profit as it enables them to make more informed decisions and (ii) the strategic effects that may arise due to the fact that the information provider's customers may interact with one another in other markets. We present our findings in an environment that features a monopolistic information provider who can sell informative signals to a set of firms that compete with one another in a downstream market. We find that, when firms compete by setting prices, it is always optimal for the information provider to sell her most informative signal to the entire market of firms. This is a consequence of the fact that in this case firms' actions are strategic complements. On the other hand, when firms compete by setting quantities, it might be optimal for the provider to sell to a smaller fraction of firms at higher prices. Moreover, we find that in the presence of cost heterogeneity, if the providers excludes some firms from the sale, she always chooses the most inefficient firms. An important extension is the scenario in which private information can leak though aggregate actions, in this case we characterize the impact of the intensity of leakage on the provider's equilibrium strategy and profits.

Taken together, our findings provide a step towards understanding the intricacies involved in markets for information. Unlike traditional markets for physical goods, it is relatively inexpensive to offer a diverse menu of information products that differ in their precision and pricing. Our results highlight that the value that a given buyer can extract from procuring such products depends not only on the product's characteristics (such as its price and precision), but also on the seller's market share and the environment in which her customers interact. Our modeling framework provides several qualitative insights in how an information provider may optimally incorporate the characteristics of such strategic interactions into her selling strategy.

Chapter 3 is motivated by the widespread adoption of online review systems by consumers in a number of industries, online retail, food and hospitality to name a few. We study the optimal pricing problem of a monopolist that launches a product for sale to a stream of consumers that arrive on the market with heterogeneous willingness to pay and are uncertain about the product quality. Upon purchasing the product, consumers leave a binary like/dislike review based on their experienced utility net of price. Future consumers use the public review information in order to form a quality estimate and make their purchase decision. Our first result is a characterization of the quality estimate of consumers. We propose an intuitive maximum likelihood procedure that consumers use to make inference about the product quality from the observed amount of likes and dislikes. First, we characterize the maximum-likelihood estimator and study its asymptotic properties. Then, we characterize the learning transient based on a mean-field approximation. Finally, we solve for the pricing policy that maximizes the seller's discounted revenue.

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

Chapter 4 is motivated by the empirically observed fact that, in markets where consumers are connected and can observe the adoption patterns of other consumers, the popularity and appeal of products depends on the mass of consumers that purchase in the early stages of the product life-cycle. We study the problem of a seller that offers a product for sale to a large market of consumers and can set different prices for two periods, an introductory period and a mature period. Consumers are heterogeneous in their private valuation for the product and their overall willingness to pay is affected by the mass of consumers that purchased before them. Two different scenarios are studied, one in which the seller has commitment power and one in which she does not.

We find that the optimal price path is generally increasing, i.e., the seller offers a lower price in the introductory period and a higher price in the mature period. Moreover, equilibrium prices vary with the level of social influence in the market, and can be very different depending on whether the seller has commitment power or not. We characterize the equilibrium thresholds that consumers use to make purchase decisions based on their private valuations, and we find that, under both scenarios, these thresholds are such that the consumers with lower valuations never buy the product when the social influence level is low and everybody buys when the social influence level is high. The most important result provides a characterization of the value of price commitment as a function of the social influence level.

Chapter 2

Information Sale and Competition

2.1 Introduction

Recent advances in information technology have streamlined the process of mining, aggregating, and processing high volume data about economic activity. Arguably, it is widely believed that the availability of more accurate information about the business environment and market conditions can be hugely beneficial to firms across a wide variety of industries. For example, in a cross-industry study, Brynjolfsson *et al.* [2011] document that firms that emphasize data-driven decision making and invest heavily in information technology outperform their peers by a wide margin.

Such a realization has in turn led to a sizable demand for Business-to-Business information services. Several firms ranging from Nielsen N.V. to Thomson-Reuters and IRI Worldwide have built their business models around collecting, customizing, and selling information products to other market participants. For example, the market research firm IRI Woldwide offers its clients a variety of consumer, shopper, and retail market analyses focused on the consumer packaged goods industry, whereas the Economist Intelligence Unit sells industry-wide market analysis reports.[1]

---

[1]IRI offers an array of information products at different price points. For example, the Basic "Market Advantage Solution" includes a summary of industry sales and a detailed analysis of pricing strategies employed by a firm's competitors. The Premium "Market Advantage Solution", on the other hand, provides a more in-depth analysis of sales and competitors' pricing strategies along with more specialized analytics services. The Basic product is priced around $10,000 whereas the price for the Premium offering can range between $100,000 and $500,000. Similarly, the Economist Intelligence Unit offers several information products for a variety of industries. For example, their basic product for telecommunications industry provides a general overview of the industry and is priced at $205, whereas

We present our main findings in the context of an environment that involves a monopolistic information provider who can sell potentially informative signals to a collection of firms that compete with one another in a downstream market. More specifically, we assume that the customer firms face demand uncertainty and that the provider is endowed with a private signal that is (partially) informative about the actual demand realization, thus creating potential gains from trade. Crucially for our argument — and in line with the observation that many real-world information providers offer a variety of information products of varying qualities — we allow for a setting in which the provider can offer information products that are potentially less precise than her private information. In other words, the provider can potentially distort the informativeness of the signal at her disposal by reducing its accuracy.

As our main result, we show that the optimal selling strategy of the provider is largely dependent on the nature and intensity of competition among its potential customers in the downstream market. More specifically, we first show that when firms engage in price competition (Bertrand), the provider finds it optimal to sell her signal with no distortion to the entire set of firms. This is due to the fact that in a Bertrand market, firms' actions are strategic complements and hence, each firm's marginal benefit of procuring a more accurate signal is increasing in the fraction of its competitors that purchase the provider's information product. Therefore, the provider would obtain maximal profits by flooding the market with highly precise signals.

The situation, however, can be dramatically different if the information provider's customers compete with one another in quantities (Cournot). For such a downstream market, we show that the provider may no longer find it optimal to sell an undistorted version of her signal to all firms. Rather, she may find it optimal to either (i) reduce the quality of her information product by selling a signal of a lower precision than the one she possesses; (ii) strategically limit her market share by excluding a subset of its customers from the sale; or finally (iii) employ both strategies simultaneously by reducing the quality *and* quantity of the products offered. This is due to the fact that in a Cournot market, firms' actions are strategic substitutes, which leads to

a more specialized report that includes finer information on network and pricing strategies costs \$2,950. They also offer customized reports at prices that are negotiated on a case-by-case basis.

the emergence of two opposing effects. On the one hand, obtaining additional information about demand directly benefits firms as they can better align their production decisions with underlying market conditions. On the other hand, however, the provider's signal can also serve as a correlating device among its customers' equilibrium actions. In particular, providing the information product to an extra firm can only increase the correlation in the firms' production decisions, an outcome that reduces each firm's profit and hence, can adversely affect the provider's bottom line. Therefore, when downstream competition is intense enough (for example, when firms' products are sufficiently substitutable), this latter, strategic channel would dominate the positive effect of reducing demand uncertainty, implying that the information provider would be better off by restricting the quantity and/or quality of the information products that it offers its customers. Interestingly, unlike in Bertrand competition, the provider's profits in a Cournot market are decreasing in the intensity of competition and may end up being significantly lower than in the absence of any competition among the downstream firms.

Finally, we discuss a number of extensions to our benchmark setup. First, we let the provider offer a menu of information products with potentially different precisions and at different prices. We provide an explicit characterization of the optimal selling strategy as a function of the nature and intensity of competition. We show that when firms compete in quantities and offer substitutable products, there is a continuum of selling strategies that lead to the same equilibrium profits for the provider. In addition, we explore the implications of firm heterogeneity for the provider's selling strategy. More precisely, we consider a setting in which firms differ in their production costs and show that it is optimal for the provider to sell higher precision information products (at higher prices) to the more efficient firms, i.e., the optimal menu features information products with precisions that are decreasing in the firms' production costs.

Taken together, these findings provide a step towards understanding the intricacies involved in markets for information. Unlike traditional markets for physical goods, it is relatively inexpensive to offer a diverse menu of information products that differ in their precision and pricing. Our results highlight that the value that a given buyer can extract from procuring such products depends not only on the product's characteristics (such as its price and precision), but also on the seller's market share and the environment in which her customers interact. Our modeling framework provides several qualitative insights in how an information provider may optimally incorporate the characteristics of such strategic interactions into her selling strategy.

2.1.1 Related literature

Our paper is related to the extensive literature that studies firms' strategic considerations in sharing information with one another in oligopolistic markets. For example, Vives [1984], Gal-Or [1985], Li [1985], and Raith [1996] provide conditions under which firms find it optimal to share their private information about market conditions with their competitors.

Relatedly, a more recent collection of papers, such as Li [2002], Li and Zhang [2008], Shin and Tunca [2010], Shamir [2012], Ha and Tong [2008], Ha *et al.* [2011] and Shang *et al.* [forthcoming], studies information sharing incentives in vertical supply chains. For instance, Shamir and Shin [2013] determine conditions under which firms can credibly share their demand forecasts with one another, whereas Cui *et al.* [forthcoming] provide a theoretical and empirical assessment of the value of information sharing in two-stage supply chains. This literature, for the most part, focuses on firms' information sharing strategies according to which each firm decides whether or not to disclose its information in full to other firms. In contrast, we consider a setting in which a *third-party* decides not only the price but also the *accuracy* of the information product(s) she makes available to a set of competing firms. This allows for richer equilibrium outcomes that highlight the interplay between the nature of competition, the optimal selling strategy, and the profits for the information provider.

Our work is also related to the growing theoretical literature on the social and equilibrium value of public information. Morris and Shin [2002] illustrate that public disclosure of information regarding a payoff-relevant parameter may adversely affect social welfare as it may crowd out agents' reliance on their private information. Angeletos and Pavan [2007] extend this framework and provide a complete taxonomy of conditions under which private and public signals are efficiently utilized in equilibrium. Relatedly, Bergemann and Morris [2013] study games of incomplete information with the goal of providing equilibrium predictions that are robust to all possible information structures. Their analysis illustrates that information disclosure policies that involve a partial sharing of a firm's private information may lead to higher equilibrium payoffs.

The question of optimal sale of information has also been studied in the context of trading in financial markets. For example, Admati and Pfleiderer [1990] consider a monopolistic seller of information interacting with a set of traders. They argue that if market prices aggregate agents' private signals, agents may find it optimal to free-ride on the information revealed via prices (as opposed to purchasing signals from the seller), thus diluting the equilibrium value of information. In contrast, our paper focuses on a different type of inter-firm strategic interaction, as firms in our framework cannot free-ride on the information generated by the actions of other firms.

Finally, our work is related to the more recent work of Bergemann and Bonatti [forthcoming], who explore selling information in the form of cookies in the context of online advertising as well as Xiang and Sarvary [2013] who consider a market for information with competition on both the demand and supply sides of the market. In a similar application context, Babaioff *et al.* [2012] study the design of optimal mechanisms for a data provider to sell information to a single buyer and provide conditions under which a single round of communication is sufficient for profit maximization.

Outline of the Paper: The rest of the paper is organized as follows. Section 2.2 contains the model and shows that our framework nests Bertrand and Cournot competition as special cases. We provide a characterization of the equilibrium in Section 2.3, where we show that the nature of downstream competition has a first-order impact on the monopolist's optimal information selling strategy. In Section 2.4, we generalize our setting by allowing the monopolist to discriminate both on prices and on the quality of information offered to its customers. In Section 2.5 we allow for some degree of information leakage through aggregate actions and study the effect of leakage on the seller's equilibrium strategy and profits. Section 2.6 and Section 2.7 present two extensions, for the case of heterogeneous firm and for the discrete version of our model with a finite number of firms. Section 2.8 concludes. All the proofs are presented in the Appendix.

2.2 Model

Firms: Consider an economy consisting of a unit mass of firms indexed by $i \in [0,1]$ that compete with one another in a downstream market. Each firm i takes an action $a_i \in \mathbb{R}$ in order to maximize its profit

$$\pi(a_i, A, \theta) = \gamma_0 a_i \theta + \gamma_1 a_i A - \frac{\gamma_2}{2} a_i^2, \tag{2.1}$$

where $A = \int_0^1 a_i di$ denotes the aggregate action taken by the firms, $\theta \in \mathbb{R}$ is an unknown payoff-relevant parameter, and $\{\gamma_0, \gamma_1, \gamma_2\}$ are some exogenously given constants. Depending on the context, action a_i may represent the quantity sold or the price set by firm i. As we will show in Subsection 2.2.2, the above framework nests Cournot and Bertrand competitions as special cases. For the time being, however, we find it more convenient to work with the general setup above without taking a specific position on the mode of competition.

The unknown parameter θ is randomly drawn by nature before firms choose their actions. As we will discuss in the following subsections, this parameter can represent the intercept of the (inverse) demand curve in the downstream market. All firms hold a common prior belief on θ, which for simplicity we assume to be the (improper) uniform distribution over the real line.[2] Even though firms do not know the realization of θ, each firm i observes a noisy private signal

$$x_i = \theta + \epsilon_i \quad , \quad \epsilon_i \sim N(0, 1/\kappa_x),$$

with κ_x capturing the signals' precision. The noise terms ϵ_i are independently distributed across firms.

Given firm i's profit function in (2.1), we let

$$\beta = -\frac{\partial^2 \pi}{\partial a \partial A} \Big/ \frac{\partial^2 \pi}{\partial a^2} = \frac{\gamma_1}{\gamma_2}, \tag{2.2}$$

[2]More formally, suppose that θ is distributed according to a Gaussian distribution with mean 0 and variance σ_θ. By letting $\sigma_\theta \to \infty$, we obtain a distribution with full support over $(-\infty, \infty)$ that, in the limit, assigns the same probability to all intervals of identical Lebesgue measures.

denote the degree of strategic complementarity in firms' actions. Note that $\beta > 0$ corresponds to an economy in which firms' actions are strategic complements: the benefit of taking a higher action to firm i increases the higher the actions of other firms are. In contrast, when $\beta < 0$, firms face a game of strategic substitutes, where i's incentives for taking a higher action decrease with the aggregate action A. Finally, $\beta = 0$ corresponds to a market in which firms face no strategic interactions.

Throughout the paper, we assume that $\gamma_2 > \max\{2\gamma_1, 0\}$. This assumption, which implies that $\beta \in (-\infty, 1/2)$, is made to guarantee that firm i's profit is strictly concave in a_i and that i's marginal profit is more sensitive to its own action a_i than to the aggregate action A.

Information Provider: In addition to the competing firms, the economy contains a monopolist who possesses some private information about the realization of the unknown parameter θ that it can potentially sell to the firms before they take their actions. The provider has access to a private signal z with precision κ_z given by

$$z = \theta + \zeta \quad , \quad \zeta \sim N(0, 1/\kappa_z),$$

where the noise term ζ is independent of ϵ_i's. Given that our main focus is on the market for information, we assume that this signal has no intrinsic value to the provider and that she can only benefit from the signal by selling it to the firms.

The key feature of our model is that the provider has control over both the "quantity" and "quality" of information sold to the firms: the information provider not only chooses the set of firms $I \subseteq [0, 1]$ that she decides to trade with, but can also choose the precision of the signal offered to the firms. More specifically, she offers a signal

$$s_i = z + \xi_i \quad , \quad \xi_i \sim N(0, 1/\kappa_\xi),$$

to firm $i \in I$ at price p_i, where ξ_i is independent from z and $1/\kappa_\xi$ captures the variance of the noise introduced by the provider into s_i. This specification thus captures the idea that the provider can control the quality of the information sold to the firms: by choosing a smaller κ_ξ, the provider can “damage” the signals offered to the firms.[3] Throughout the paper, we refer to s_i as the *market signal* sold to firm i.

In general, the noise added to different firms’ signals by the provider may be correlated with one another. To capture this idea formally, we assume that in addition to their precision κ_ξ, the provider can also determine the correlation between different firms’ market signals by setting $\rho_\xi = \mathrm{corr}(\xi_i, \xi_j) \in [0,1]$. Our specification thus accommodates situations in which the information provider offers identical or conditionally independent signals to any subset of the firms as special cases.

Putting the above together, the market signal s_i offered to firm $i \in I$ can be rewritten as

$$s_i = \theta + \eta_i \quad , \quad \eta_i \sim N(0, 1/\kappa_s) \quad \text{and} \quad \mathrm{corr}(\eta_i, \eta_j) = \rho,$$

where $\kappa_s = (1/\kappa_z + 1/\kappa_\xi)^{-1}$ is the signal’s precision and $\rho = (\kappa_\xi + \rho_\xi \kappa_z)/(\kappa_\xi + \kappa_z)$. Note that by construction, signals sold by the provider cannot be more precise than the information she possesses; that is, $\kappa_s \leq \kappa_z$.

We remark that given firms’ *ex ante* symmetry, we can assume, without loss of generality, that $I = [0, \lambda]$, where $\lambda \in [0,1]$ captures the fraction of firms that the information provider decides to trade with. Also note that even though we assume that the seller chooses the fraction of firms she wants to trade with before offering them her information products, as we show in Section 2.4, our setting is isomorphic to an environment in which the provider announces the features of her product(s) — i.e., price and precision — and firms subsequently decide whether to purchase the products.

Finally, with some abuse of terminology, we refer to the firms who purchase the market signal s_i as *informed firms*, whereas firms that were denied the signal or decided not to purchase it from the information provider are simply referred to as being *uninformed*.

[3] Note that in our baseline setting, the provider offers a signal of the same precision to all firms $i \in I$; that is, κ_ξ is independent of i. We relax this assumption in Section 2.4 and show that all our insights are robust to this assumption.

2.2.1 Contracts and Equilibrium

Once the seller's and the firms' private signals are realized, the former has the option to sell potentially informative signals about θ to the latter. To capture this idea formally, we assume that the information provider makes a take-it-or-leave it offer $(\kappa_\xi, p_i, \rho_\xi)$ to a fraction λ of the firms, where κ_ξ captures the quality of the market signal offered to firm i and p_i is the corresponding firm-specific price.

Following the seller's offer, each firm $i \in [0, \lambda]$ then decides whether to accept $(b_i = 1)$ or reject $(b_i = 0)$ its corresponding offer. This stage is then followed by the competition subgame between the firms in which they choose their actions a_i. Note that whereas the strategy of an uninformed firm i is a mapping from its private signal x_i to an action, the strategy of an informed firm maps the pair (x_i, s_i) to an action.

Given this setup, we have the following natural solution concept:

Definition 2.2.1. An equilibrium consists of a fraction λ of firms that receive an offer, a set of individual prices $\{p_i\}_{i \in [0,\lambda]}$, market signal precision κ_ξ and correlation ρ_ξ chosen by the information provider, acceptance/rejection decisions $b_i \in \{0, 1\}$ for each firm, and firm-specific strategies such that:

(i) The information provider maximizes her profits.

(ii) Firm $i \in [0, \lambda]$ accepts the provider's offer only if it is individually rational to do so, taking the acceptance/rejection decisions of other firms as given.

(iii) Once the provider's offers are accepted or rejected, the firms' actions constitute a Bayes-Nash equilibrium of the competition subgame.

2.2.2 Examples

As already mentioned, Cournot and Bertrand competition can be derived as special cases of our general framework above. This feature of the model enables us to provide a comparison of the optimal information selling strategies in markets with different modes and intensities of competition.

The following simple examples illustrate how in the presence of linear demand functions, various forms of competition can induce quadratic profit functions in the form of Equation (2.1). We will use these examples in the subsequent sections to discuss the implications of our results for the optimal trading strategies of the information provider.

Example 2.2.1 (Cournot competition)**.** Consider a market in which firms sell a possibly differentiated product to a downstream market and compete by setting quantities. Firm i faces an inverse demand function given by

$$r_i = \gamma_0\theta - (1-\delta)Q - \delta q_i, \tag{2.3}$$

where q_i is the quantity sold by firm i, $Q = \int_0^1 q_i di$ is the aggregate quantity sold to the downstream market, and θ captures the intercept of the (inverse) demand curve. In this setting, $\delta \in [0,1]$ represents the degree of product differentiation among firms, as a smaller δ corresponds to a more homogenous set of products.[4] Assuming that firms' marginal cost of production is zero, it is then immediate that their profit function $\pi_i = r_i q_i$ is simply a special case of our framework in (2.1), with action a_i representing the quantity sold by firm i.

Note that in this environment, the degree of strategic complementarity defined in (2.2) is equal to $\beta = (\delta-1)/2\delta < 0$, thus implying that firms face a game of strategic substitutes. Parameter β also captures the intensity of competition between the firms. In particular, given that β is increasing in δ, a larger β corresponds to a market in which products are more differentiated. In the extreme case that $\beta \to 0$, the products are no longer substitutes and each firm essentially becomes a monopolist in its own market. At the other extreme, as $\beta \to -\infty$, the products become perfect substitutes and the oligopoly converges to a perfectly competitive market.

Example 2.2.2 (Bertrand competition)**.** Next, consider a market in which firms compete in prices and face a linear demand function given by

$$q_i = \gamma_0\theta + (\phi - 1)R - \phi r_i,$$

[4] See Myatt and Wallace [2015] for microfoundations for this demand system.

where r_i is the price set by firm i and $R = \int_0^1 r_i di$ is the average price in the market. Note that this demand system can be obtained by inverting (2.3) and setting $\phi = 1/\delta > 1$. Once again, it is immediate that firm i's profit function $\pi_i = r_i q_i$ would coincide with (2.1), where action a_i now represents the price set by firm i. Furthermore, it is straightforward to verify that, in this environment, $\beta = (\phi - 1)/2\phi > 0$, thus implying that the competition game between the firms exhibits strategic complementarities, the degree of which is increasing in ϕ.

Example 2.2.3. Once again consider the Cournot competition setting described in Example 2.2.1, but instead suppose that firms produce homogeneous products, i.e., $\delta = 0$, and have quadratic production costs given by $c(q_i) = q_i^2/2$. The profit of firm i is then given by

$$\pi(q_i, Q, \theta) = \gamma_0 q_i \theta - q_i Q - \frac{1}{2} q_i^2,$$

which again fits within our general framework.

We conclude this section by remarking that even though, for the sake of tractability and exposition, we focus on an environment consisting of a continuum of firms, as we show in Subsection 2.7, our results and qualitative insights carry over to a setting with finitely many firms.

2.3 Optimal Sale of Information

In this section, we present our main results and characterize the information provider's optimal information selling strategy. Our results show that the seller's strategy is highly sensitive to the mode and intensity of competition in the downstream market as expressed by β.

2.3.1 Competition Subgame

We start our analysis by studying the equilibrium in the competition subgame between the firms once the contracts offered by the information provider are accepted or rejected. Without loss of generality, let $[0, \ell]$ denote the set of firms who accept the seller's offer, where, clearly, $\ell \leq \lambda$. Our first result generalizes the results of Angeletos and Pavan [2007] and provides a characterization of the firms' equilibrium strategies in the competition subgame.

Proposition 2.3.1. *The competition subgame between the firms has a unique Bayes-Nash equilibrium in linear strategies. Furthermore, the equilibrium strategies of the firms are given by*

$$a_i = \begin{cases} \alpha[(1-\omega)x_i + \omega s_i] & \text{if } i \in [0, \ell] \\ \alpha x_i & \text{if } i \in [\ell, 1] \end{cases},$$

where

$$\omega = \frac{\kappa_s}{(1-\beta\ell\rho)\kappa_x + \kappa_s}$$

and $\alpha = \gamma_0/(\gamma_2 - \gamma_1)$.

Proposition 2.3.1 states that the equilibrium action of an informed firm is a weighted sum of its original private signal and the signal it obtains from the information provider. More importantly, however, it shows that the weights firm i assigns to its two signals not only depend on their relative precisions, but also on the fraction of informed firms, ℓ, as well as correlation ρ in the market signals. In particular, the equilibrium weight that each informed firm assigns to the market signal s_i is increasing in the degree of strategic complementarities β, regardless of the values of ρ and ℓ.[5] This is due to the fact that in the presence of stronger strategic complementarities, firms have stronger incentives to coordinate with one another, and as a result, rely more heavily on their market signals, which can function as (imperfect) coordination devices. On the other hand, in the absence of strategic considerations (i.e., when $\beta = 0$), the optimal strategy of all firms would be independent of ℓ and ρ, making the weight assigned to each signal proportional to its relative precision.

Relatedly, Proposition 2.3.1 also establishes that for a given positive (negative) β, the equilibrium weight that informed firms assign to their market signals is increasing (decreasing) in ℓ and

[5]Recall that $\rho \geq 0$.

ρ. To see the intuition underlying this observation, suppose that $\beta > 0$.[6] In such an economy, firms face a game of strategic complements, as for example would be the case if they compete *à la* Bertrand as in Example 2.2.2. Given that firms value coordinating their actions with one another, a given informed firm i assigns a higher weight to its market signal — above and beyond what its relative precision would justify — the more other firms base their own decisions on the signal sold by the information provider (i.e., higher ℓ) and the more informative s_i is about the signals of other firms (i.e., higher ρ).

With Proposition 2.3.1 at hand, in the remainder of this section, we turn to the the seller's problem and characterize her optimal information selling strategy as a function of the mode and intensity of competition in the downstream market. In order to present our results in the most transparent manner, we study Bertrand and Cournot competition separately.

2.3.2 Bertrand Competition

First, consider the case in which firms compete with one another *à la* Bertrand. As already mentioned in Example 2.2.2, such a market corresponds to a special case of our general framework with $\beta > 0$. Also, recall that the information provider needs to choose the fraction of firms with whom she trades (λ), the precision of the signal offered to the firms (κ_s) and the correlation induced in the noise terms (ρ_ξ). We have the following result:

Proposition 2.3.2. *If $\beta > 0$, the information provider sells her signal without any distortions to all firms; that is, $\kappa_s^* = \kappa_z$ and $\lambda^* = 1$. Furthermore, the provider's expected profit is given by*

$$\Pi^* = \alpha^2 \left(\frac{\gamma_2}{2}\right)\left(\frac{\kappa_z}{\kappa_x}\right)\frac{\kappa_z + \kappa_x}{[(1-\beta)\kappa_x + \kappa_z]^2}. \tag{2.4}$$

The above result thus establishes that under Bertrand competition, it is always optimal for the provider to sell her signal z to the entire set of firms without any additional noise. To understand the intuition underlying this result, recall that in a Bertrand market, the firms' actions are strategic complements: setting a lower price becomes more attractive the lower the prices of other competing

[6] The argument for $\beta < 0$ is identical.

firms are. Such strategic complementarities induce a strong coordination motive among the firms. Therefore, providing the market signal to an additional marginal firm, not only increases the profits of the seller directly (via sales to that new marginal firm), but also increases the surplus of all other firms who have already acquired the signal. This extra surplus can thus be appropriated by the seller via higher prices, leading to even higher profits. Consequently, the information provider always finds it optimal to sell to the entire market of firms. An identical argument then shows that the provider would not distort the signal either: sharing a more precise signal with a new firm increases the value of the market signal to the rest of the informed firms.

Proposition 2.3.2 also characterizes the expected profit of the seller. From (2.4), it is easy to verify that Π^* is increasing in the quality of information available to the monopolist (κ_z), but is decreasing in the precision of the firms' private signals (κ_x). The intuition underlying these observations is simple. Given that the information provider always has the option to reduce the precision of the signals it offers to the firms, her profits can never decrease by having access to a more precise signal. On the other hand, however, the extra benefit of the market signal to the firms is lower the more informed they are to begin with, thus reducing the provider's expected profits.

More importantly, however, (2.4) also shows that the monopolist's expected profit increases in the degree of strategic complementarities β. Recall from Example 2.2.2 that $\beta = (\phi - 1)/2\phi$, where $1/\phi = \delta$ is the degree of product differentiation among the firms. Therefore, increasing β is essentially equivalent to a lower degree of product differentiation, and hence, more intense competition. Thus, as β increases, coordination becomes more important to the firms, increasing the value of the seller's signal which in turn leads to higher expected profits.

As a final remark, note that since it is never optimal for the information provider to add noise to the signals, the correlation $\rho_\xi = \mathrm{corr}(\xi_i, \xi_j)$ is immaterial for her profits.

2.3.3 Cournot Competition

We next focus on the case in which firms compete with one another *à la* Cournot. Recall from Example 2.2.1 that such a market is a special case of our general setup with $\beta < 0$. In this case, firms choose quantities and their actions are strategic substitutes. Note that, unlike the case of

Bertrand competition, firms no longer value coordination *per se*. The following two propositions provide a characterization of the optimal information selling strategy of the monopolist as a function of the degree of strategic substitutability among the actions of downstream firms.

Proposition 2.3.3. *If $-(1+\kappa_z/\kappa_x) \leq \beta < 0$, the information provider sells her signal without any distortions to all firms; that is, $\kappa_s^* = \kappa_z$ and $\lambda^* = 1$. Furthermore, the provider's expected profit is given by*

$$\Pi^* = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_z}{\kappa_x}\right) \frac{\kappa_z + \kappa_x}{[(1-\beta)\kappa_x + \kappa_z]^2}. \tag{2.5}$$

Thus, in a Cournot market with a weak enough intensity of competition, the seller finds it optimal to follow the same strategy as in a Bertrand market: sell an undistorted version of her signal to the entire set of firms. The intuition underlying this result is straightforward: acquiring information about the demand intercept (θ) allows each firm i to better match its supply decision to the underlying demand and as a consequence, to increase its profit. The monopolist can then appropriate the increase in i's sales by demanding a higher price in exchange for the signal. Therefore, the provider is always better off by making the most precise version of her signal available to all firms i.

Even though the seller follows the same strategy as in the Bertrand market, comparing expressions (2.4) and (2.5) implies that her expected profit is lower under Cournot competition ($\beta < 0$). This is due to the fact that unlike Bertrand competition, firms do not have an incentive to coordinate their actions, undermining the role of the market signal as a coordination device.

Interestingly, the predictions of Propositions 2.3.2 and 2.3.3 no longer hold if the intensity at which downstream firms compete with one another in a Cournot market is high. We have the following result:

Proposition 2.3.4. *If $\beta < -(1+\kappa_z/\kappa_x)$, the information provider maximizes her expected profit by following any information selling strategy that is a solution to the following equation:*

$$(\kappa_z + \beta\lambda^*\kappa_s^*)\kappa_x + \kappa_z\kappa_s^* = 0. \tag{2.6}$$

Furthermore, her expected profit is given by

$$\Pi^* = -\alpha^2 \left(\frac{\gamma_2}{2}\right) \frac{\kappa_z}{4\beta\kappa_x^2}. \tag{2.7}$$

The key observation here is that the pair $\kappa_s^* = \kappa_z$ and $\lambda^* = 1$ does not satisfy (2.6), leading to the following corollary:

Corollary 2.3.1. *Suppose that $\beta < -(1 + \kappa_z/\kappa_x)$. Then, either $\kappa_s^* < \kappa_z$ or $\lambda^* < 1$.*

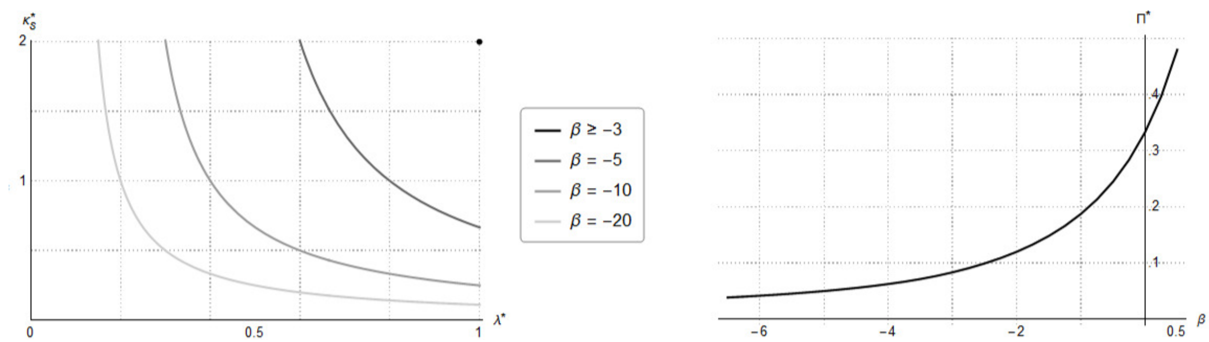
Therefore, when firms compete with one another *à la* Cournot and offer goods that are strong substitutes — corresponding to a large enough negative β — it is optimal for the seller to distort the information ($\kappa_s^* < \kappa_z$) and/or exclude a fraction of the firms from the sale ($\lambda^* < 1$).

To see the intuition underlying the above result, recall that in a Cournot market, firms' actions are strategic substitutes, i.e., increasing a firm's supply leads to higher marginal profit the lower the supply decisions of its competitors are. Therefore, providing the market signal to an additional firm i affects its profit through two distinct channel. On the one hand, a more precise market signal enables i to better match its supply to the realized demand. On the other hand, however, making such a signal available to i increases the correlation in the firms' actions, as now i's action would be more correlated with the market parameter θ. The presence of this second effect implies that the strategic value of the seller's signal to firm i and consequently i's willingness to pay for it are decreasing in the fraction of firms that accept the provider's offer. When the competition among the firms is sufficiently intense (i.e., the goods they offer are sufficiently substitutable), this strategic effect would dominate the first effect, thus making it profitable for the information provider to restrict her offer to a strict subset of the firms ($\lambda^* < 1$).

By Proposition 2.3.4, an alternative optimal strategy for the monopolist would be to distort the information she sells to the market. In fact, as Equation (2.6) suggests, the fraction λ of the firms that the monopolist trades with and the precision κ_s of the signal offered to the firms are substitutes: as the monopolist increases her market share, she finds it optimal to increasingly distort the signals.

Note that the information provider's expected profit decreases in the degree of strategic substi-

Figure 2.1: Optimal selling strategy for different levels of β (left); Equilibrium profit as a function of β (right). We use the following set of parameters for this example: $\alpha = \gamma_2 = 1$ for the firms' payoff functions and $\kappa_x = 1, \kappa_z = 2$ for the signal precisions of the firms' private signals and the provider's information respectively.



tutability ($|\beta|$) of the firms' actions. This is a consequence of the fact that the strategic value of the seller's signal and consequently a firm's willingness to pay for it decrease as the market becomes more competitive. This is in contrast with the case of Bertrand competition where the seller's expected profit increases with the intensity of competition among the firms as they have a stronger incentive to purchase the market signal and coordinate their actions.

We also remark that regardless of the value of β and the strategy adopted by the information provider, she never has an incentive to introduce correlation into market signals, i.e., it is always optimal to set $\rho_\xi^* = 0$. Increasing the correlation in the signals provided to downstream firms would invariably increase the correlation among their actions and lead to lower profits for the seller.

Finally, note that the threshold $-(1 + \kappa_z/\kappa_x)$ at which the seller finds it optimal to limit her market share and/or strategically distort the market signal is decreasing in the ratio κ_z/κ_x, implying that the more informed the information provider is relative to her customers, the more likely it is that she will be able to fully exploit her informational advantage by selling it to the entire market of firms without distortion.

Figure 2.1 illustrates the optimal selling strategy and the equilibrium profit of the information provider for the following set of parameters: $\alpha = \gamma_2 = 1$, $\kappa_x = 1$, and $\kappa_z = 2$. It turns out that for these parameters the threshold at which the seller finds it optimal to strategically distort the

	$\beta = 0$	$\beta = -3$	$\beta = -5$	$\beta = -10$	$\beta = -20$
Π^*_β / Π^*_0	1	.250	.150	.075	.038
$\Pi^{\textbf{no-dist}}_\beta / \Pi^*_0$	1	.250	.141	.053	.017
Increase in Profits (%)	0%	0%	6.67%	40.83%	120.42%

Table 2.1: Comparing profits under the optimal information selling strategy with selling the provider's signal undistorted to the entire market.

market signal is equal to -3. Note that as the *left plot* highlights for values of β greater than the threshold, the provider sets the precision of the market signal to $\kappa_s^* = 2$, i.e., she does not distort the information she has at her disposal, and does not exclude any firms from the sale ($\lambda^* = 1$). On the other hand, for values of β lower than the threshold, the seller finds it optimal to distort the information she provides to the market and limit her market share (in particular, her optimal selling strategy is given by Equation (2.6)). The *right plot* illustrates how the provider's profit varies with the intensity of competition. Note that the seller is always better off when firms view their actions as strategic complements ($\beta > 0$) as opposed to strategic substitutes.

We conclude this section by exploring the extent to which an information provider can increase her profits by strategically distorting the information she provides to her downstream customers and/or limiting her market share. In Table 2.1 we compare the profits for a provider that optimally sells her information to the firms (Π^*_β) with the profits for a provider that sells the information she has at her disposal as is to the entire market ($\Pi^{\text{no-dist}}_\beta$). We benchmark Π^*_β and $\Pi^{\text{no-dist}}_\beta$ against the profits for a provider that follows her optimal strategy in the absence of competition, i.e., when $\beta = 0$. The first two rows of the table clearly highlight the effect of intensifying the competition among the firms on the providers's profits. Furthermore and quite importantly, as we report in the third row of the table the provider earns significantly higher profits under competition when she distorts her market signal and/or limits her market share — the increase in her profits by following the optimal strategy characterized in Proposition 2.3.4 ranges from 6.67% to 120.42% as the extent to which firms view their actions as strategic substitutes increases.

2.4 Information Quality Discrimination

In our baseline model presented in Section 2.2 and analyzed in Section 2.3, we assumed that the information provider can only offer a single product to the entire market, in the sense that she offers a market signal of the same precision to all firms. In this section, we relax this assumption by allowing the seller to offer signals that potentially differ in both price and precision. Formally, the information provider makes a take-it-or-leave-it offer to each firm $i \in [0,1]$, specifying the signal precision κ_{si} and price p_i. Note that as in our earlier setting, the seller cannot offer a signal of a higher precision than her own private signal, that is, $\kappa_{si} \leq \kappa_z$ for all i. Furthermore, it is immediate to verify that the baseline model of Section 2.2 is a special case of this more general model, as the monopolist can simply exclude firm i by either charging $p_i = \infty$ or providing a completely uninformative signal with precision $\kappa_{si} = 0$. The following result, which generalizes Propositions 2.3.2–2.3.4, shows that all our earlier insights remain valid under this more general specification.

Proposition 2.4.1. *The optimal information selling strategy for the information provider is given as follows:*

(a) If $\beta \geq -(1 + \kappa_z/\kappa_x)$, the information provider offers an undistorted version of her signal to all firms at price

$$p^* = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_z}{\kappa_x}\right) \frac{\kappa_z + \kappa_x}{[(1-\beta)\kappa_x + \kappa_z]^2}.$$

(b) If $\beta < -(1 + \kappa_z/\kappa_x)$, the information provider offers a market signal of precision κ_{si}^* to firm i at price p_i^*, where $\{\kappa_{si}^*\}_{i \in [0,1]}$ solve

$$\int_0^1 \frac{\kappa_{si}^*}{\kappa_x + \kappa_{si}^*} di = -\frac{\kappa_z}{\beta \kappa_x} \tag{2.8}$$

and

$$p_i^* = \alpha^2 \left(\frac{\gamma_2}{2}\right) \frac{\kappa_{si}^*}{4\kappa_x(\kappa_x + \kappa_{si}^*)}. \tag{2.9}$$

Statement (a) of the above result shows that the information provider offers an undistorted version of her signal to all firms in the downstream market if either they compete *à la* Bertrand, or alternatively, if the intensity of the Cournot competition is not strong enough. In this sense, this result generalizes Propositions 2.3.2 and 2.3.3, establishing that the seller has no incentive to discriminate among the firms in either price or information quality.

Statement (b) of Proposition 2.4.1 considers the setting in which firms' actions are strong strategic substitutes, for example, when they compete *à la* Cournot and produce goods that are highly substitutable. Consistent with the discussion in Subsection 2.3.3, this result shows that the information provider finds it optimal to either distort the signals sold to the downstream firms or strategically restrict her market share. In particular, it is easy to verify that $\kappa_{si}^* = \kappa_z$ for all i does not satisfy the optimality condition (2.8). The intuition underlying this result parallels those behind Proposition 2.3.4 and Corollary 2.3.1: providing high quality signals to all firms increases the induced correlation in their actions, which in turn reduces their profit when their actions are strong strategic substitutes. Thus, the monopolist would be better off by limiting its market share or reducing the quality of the signals sold to the firms. Note, however, that the optimal strategy of the information provider is not unique. Rather, any signal precision profile $\{\kappa_{si}^*\}$ that satisfies (2.8) would lead to the same expected profit. Nevertheless, irrespective of the strategy chosen by the monopolist, her incentive to lower the precision of the market signals increases as firms' actions become stronger strategic substitutes. In particular, as $\beta \to -\infty$, the downside of coordination among firms that trade with the monopolist is so strong that essentially no trade takes place in equilibrium: the information provider offers a completely uninformative signal $\kappa_{si}^* \to 0$ to all firms at price $p_i^* \to 0$.

Example 2.4.1 (Selling two products)**.** Consider a Cournot market in which $\beta < -(1 + \kappa_z/\kappa_x)$. Suppose that the information provider can offer two distinct information products: a premium product of precision $\bar{\kappa}_s$ at price $\bar{p}$ and an inferior product of precision $\underline{\kappa}_s < \bar{\kappa}_s$ at price $\underline{p}$. Let $\bar{\lambda}$ and $\underline{\lambda}$ denote the fraction of firms offered the premium and inferior products, respectively, where by construction $\bar{\lambda} + \underline{\lambda} \leq 1$. Condition (2.8) implies that it is optimal for the seller to design her information products such that

$$\bar{\lambda}\left(\frac{\bar{\kappa}_s}{\kappa_x + \bar{\kappa}_s}\right) + \underline{\lambda}\left(\frac{\underline{\kappa}_s}{\kappa_x + \underline{\kappa}_s}\right) = -\frac{\kappa_z}{\beta\kappa_x}.$$

The above equation highlights the trade-off between information quality and quantity faced by the information provider in designing her menu of products. In particular, increasing the precision $\bar{\kappa}_s$ of the premium product requires either a reduction in its supply $\bar{\lambda}$, or alternatively, a reduction in the precision or the supply of the inferior product.

Note that, by selling a premium product the information provider is placing the well-informed firms at an advantage vis-à-vis their less-informed competitors. This enables her to charge the former set of firms a higher price. In fact, as equation (2.9) highlights, $\bar{p} > \underline{p}$.

We end this discussion by remarking that the ability to discriminate on information quality does not offer the seller any advantage compared to our benchmark model of Sections 2.2 and 2.3. In particular, equation (2.8) always has a solution such that $\kappa_{si} = \kappa_s$ for a fraction λ of the firms and $\kappa_{si} = 0$ for the rest. In other words, offering two information products, one with non-zero precision at a strictly positive price and another with zero precision at zero price, is sufficient for the seller to maximize her expected profit.

2.5 Information Leakage

Firms' actions typically reflect the payoff-relevant information they have at their disposal. Thus, a seller of information may need to take into account the dilution in the value of information to the competing firms due to its leakage through their actions. This section explores an extension of our benchmark setting that directly incorporates information leakage through the firms' actions. In particular, in addition to its private and market signals x_i and s_i, respectively, firm i also has access to signal S_i which centered around the aggregate action that is in the market. The precision of this signal about the aggregate action can be viewed as a measure of information leakage.

To illustrate clearly the effect leakage on the information provider's optimal selling strategy and profits, we focus on equilibria in which all firms purchase the provider's information, i.e., $\lambda = 1$. Note that this is without loss of generality since there always exists a profit maximizing selling strategy that induces such equilibrium behavior. Moreover, we assume that $\rho_\xi = 0$, i.e., the provider sells information signals that are independent conditional on the realization of state θ (this is, again, without any loss of generality as we establish in Section 2.3). Thus, the information provider optimizes over the precision κ_s of her information product. Moreover, in addition to signals x_i and s_i, firm i observes signal S_i that takes the following form

$$S_i = A + \nu_i \qquad , \qquad \nu_i \sim N(0, 1/\kappa_\nu),$$

where the noise terms ν_i's are independently distributed across firms and the precision κ_ν measures the extent of information leakage in the market. In particular, when $\kappa_\nu = \infty$ then signal S_i is perfectly informative of the aggregate action A whereas when $\kappa_\nu = 0$ then S_i does not convey any payoff-relevant information to firm i, i.e., there is no information leakage, and the setting is equivalent to our benchmark model.

We extend the firms' strategy space by allowing them to condition their actions on firm S_i. In particular, firms specify a *supply function* that depends on the information they have at their disposal, i.e., signals x_i, s_i, and S_i. In other words, firm i's action is a map from the signal space to the space of supply functions, i.e., a function $a_i(x_i, s_i, S_i)$, as opposed to a scalar as in our benchmark model.[7] Given the strategies of all firms $i \in [0,1]$ the aggregate action satisfies $A = \int_0^1 a(x_i, s_i, S_i)\, di$. The equation has a unique solution for a linear-quadratic framework with Gaussian noise. We denote the unique solution by $\hat{A}\big(\{a(x_i, s_i, S_i)\}_{i\in[0,1]}\big)$.[8] The profit of firm i for

[7] Our approach builds on Vives [2011] that studies competition among firms that possess private payoff-relevant information. Using supply functions and having firms condition on a (potentially noisy) signal about the aggregate action, allows us to directly incorporate information leakage into our benchmark model and study how the provider's optimal selling strategy and profits vary as a function of the extent of leakage in a market.

[8] This solution, which we characterize in the proof of Proposition 2.5.1, depends on state θ, the provider's noise ζ, and the coefficients of the firm's equilibrium strategies.

any given realization of signals is given as

$$\pi_i = \gamma_0 a(x_i, s_i, S_i)\theta + \gamma_1 a(x_i, s_i, S_i)A - \frac{\gamma_2}{2} a(x_i, s_i, S_i)^2,$$

where $S_i = A + \nu_i$ and $A = \hat{A}\big(\{a(x_j, s_j, S_i)\}_{j\in[0,1]}\big)$. Given signals (x_i, s_i, S_i), firm i determines its equilibrium strategy $a_i = a(x_i, s_i, S_i)$ so as maximizes its expected profit $\mathbb{E}[\pi_i | x_i, s_i, S_i]$ in the competition subgame, taking the strategies of other firms as given.

The following proposition summarizes our findings regarding the effect of information leakage on the provider's optimal selling strategy and equilibrium profits.

Proposition 2.5.1. *Let Π^* denote the equilibrium profits of the information provider in the presence of leakage. For sufficiently small κ_ν, we obtain that*

(a) $\dfrac{\partial \Pi^*}{\partial \kappa_\nu} < 0$ *for all* $\beta \in (-\infty, 1/2)$.

(b) There exists $-(1 + \kappa_z/\kappa_x) < \bar{\beta} < 0$ *such that* $\kappa_s^* < \kappa_z$ *for all* $\beta \in \big[-(1 + \kappa_z/\kappa_x), \bar{\beta}\big)$.

Part (a) of Proposition 2.5.1 establishes that the monopolist's equilibrium profits decrease in the presence of information leakage. This is true regardless of the value of β, i.e., whether actions are strategic complements or substitutes (Bertrand or Cournot competition). As one would expect, when firms are able to (partially) infer their competitors' information by observing a signal about their aggregate action, the value of the provider's information decreases and, thus, her profits go down.

More importantly, part (b) of Proposition 2.5.1 establishes that the range of β's for which the information provider finds it optimal to distort her information in the presence of leakage is wider than when firms determine their actions based solely on signals x_i and s_i. The provider's incentives to distort the information she sells to the downstream market grow stronger as her ability to extract surplus from the firms when increasing the precision of signal s_i is hindered by information leakage. That said, even in the presence of leakage, distorting her information is never optimal when firms's actions are strategic complements, i.e., as in Bertrand competition. [9]

[9] Note that this does not imply that the optimal selling strategy is the same as in the absence of leakage. In particular, the provided is forced to sell her information at a lower price.

Figure 2.2: The provider's equilibrium profits (left) and her optimal selling strategy (right) as functions of β for different levels of information leakage. We use the following set of parameters for this example: $\alpha = \gamma_2 = 1$ for the firms' payoff functions and $\kappa_x = 1, \kappa_z = 2$ for the signal precisions of the firms' private signals and the provider's information respectively. We plot the provider's profits and the precision of the signal she sells to the downstream market (κ_s^*) as a function of β for three levels of information leakage $\kappa_\nu = 0$ (no leakage), $\kappa_\nu = 1$, and $\kappa_\nu = 10$.

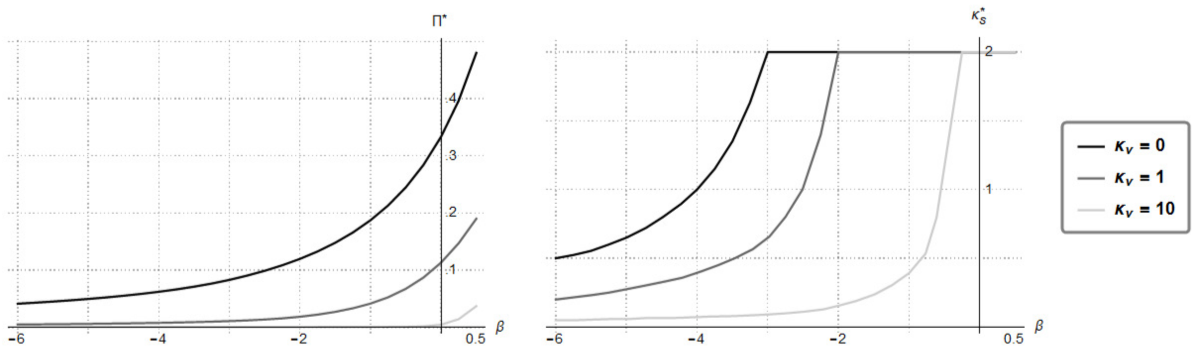


Figure 2.2 illustrates the equilibrium profits and the provider's optimal selling strategy for different levels of leakage. As can be clearly seen in the left plot shows the provider's profits are decreasing in the level of information leakage in the market irrespective of the value of β. The right plot explores the impact of leakage on the provider's optimal selling strategy. In particular, she finds it optimal to distort the information she sells to the downstream market for a wider range of β's than in the absence of leakage. In addition and as already mention above, distortion is never optimal in the presence of strategic complementarities ($\beta > 0$) irrespective of the level of leakage.

Table 2.2 presents our numerical solutions for different level of leakage at different values of β. We report both the equilibrium profits and the optimal precision in three different scenarios, our baseline where there is no leakage ($\kappa_\nu = 0$) as well as two scenarios with low and high leakage intensity respectively. These allow us to appreciate how severely leakage decreases the equilibrium profits. In particular, comparing the equilibrium profits when the level of leakage is high to the equilibrium profits in the absence of leakage, we see that at $\beta = 1/3$ the equilibrium profit with high leakage is only 5% of the corresponding profit when there is no leakage. Moreover, when $\beta < 0$ the negative effect of leakage on profits is more severe the higher the level of strategic substitutability,

	Leakage Level	$\beta = -4$	$\beta = -3$	$\beta = -2$	$\beta = -1$	$\beta = 0$	$\beta = 1/3$
	$\kappa_\nu = 0$	.0625	.0833	.1200	.1875	.3333	.4218
Equilibrium Profits	$\kappa_\nu = 1$	.0082	.0014	.0189	.0420	.1134	.1606
	$\kappa_\nu = 10$	.0002	.0003	.0005	.0008	.0050	.0209
	$\kappa_\nu = 0$	1	2	2	2	2	2
Optimal Precision	$\kappa_\nu = 1$	.40	.65	2	2	2	2
	$\kappa_\nu = 10$	.08	.12	.16	.36	2	2

Table 2.2: Equilibrium profit and optimal precision for different levels of information leakage.

when $\beta = -4$ we see that the equilibrium profits under high leakage are extremely low, less than .5% of the corresponding profits in the absence of leakage. Moreover, the second part of the table highlights the impact on information leakage on the optimal selling strategy when β is negative, for example, when $\beta = -2$ it is optimal to sell a signal with full precision when leakage is absent or low, but when leakage is high the seller dramatically decreases the signal precision to only 8% of her best signal precision.

2.6 Heterogeneous Firms

Our analysis thus far focused on an environment consisting of a continuum of homogenous firms. In this section, we discuss how our results are affected by introducing heterogeneity among the firms (in terms of their production costs). Once again, consider the environment presented in Example 2.2.1, where firms compete with one another in quantities. We generalize this setting in two dimensions, by (i) allowing heterogeneity in firms' production costs and (ii) introducing a transaction cost borne by the information provider whenever she trades with a downstream firm.

More specifically, we assume that downstream firms are heterogeneous with respect to their costs of production: firm i faces a quadratic production cost of $C_i(q_i) = c_i q_i^2/2$, where q_i is the quantity produced by i and $c_i > 0$. The firm's profit is thus given by

$$\pi_i(q_i, Q, \theta) = \gamma_0 q_i \theta + \gamma_1 q_i Q - \frac{1}{2} c_i q_i^2, \tag{2.10}$$

where Q denotes the aggregate quantity in the market and $\gamma_1 < 0$ is some constant. Note that even though the above expression is similar to (2.1), the extent of strategic complementarities can no longer be captured by a single parameter β, as now firms face potentially heterogeneous production costs.

As for transaction costs, we assume that the seller incurs a cost equal to $\upsilon\kappa_{si}$ whenever she sells a signal of precision κ_{si} to firm i, where $\upsilon > 0$. This cost can, for example, capture the idea that the firm cannot provide verifiable and/or credible information to its customers at no cost. Rather, it needs to spend resources to ensure its customer that the market signal is indeed as informative as claimed. Alternatively, it can be thought of as the cost associated with customizing the provider's information to meet the customer's informational needs. As in Section 2.4, we allow the seller to discriminate along both signal precision and price. We have the following result:

Proposition 2.6.1. *There exist $\overline{\upsilon} > \underline{\upsilon}$ such that*

(a) if $\upsilon > \overline{\upsilon}$, the information provider does not transact with any of the firms; that is, $\kappa^*_{si} = 0$ for all i.

(b) if $\upsilon < \underline{\upsilon}$, the information provider sells her signal with no distortion to all firms; that is, $\kappa^*_{si} = \kappa_z$ for all i.

(c) For any $\upsilon \in (\underline{\upsilon}, \overline{\upsilon})$, then there exist c^* such that

$$\kappa^*_{si} = \begin{cases} 0 & \text{if } c_i > c^* \\ \kappa_z & \text{if } c_i < \frac{\kappa_x^2}{(\kappa_x+\kappa_z)^2} c^* \\ \kappa_x\left(\sqrt{c^*/c_i} - 1\right) & \text{otherwise.} \end{cases}$$

The above result thus establishes that the information provider finds it optimal to follow an information selling strategy that involves offering a signal to firm i with a precision that is decreasing in the firm's cost c_i, i.e., the provider sells higher quality signals to more efficient firms. Formally, κ^*_{si} is always non-increasing in c_i. However, note that this does not mean that the monopolist sells her best available information to all firms, even when transactions are costless. Rather, due to the presence of strategic interactions between downstream firms (and in line with our earlier results), the provider may either sell distorted signals to some firms or simply even exclude them by offering non-informative signals $\kappa_{si}^* = 0$ altogether. Thus, Proposition 2.6.1 generalizes Propositions 2.3.4 and 2.4.1 to the case in which firms face heterogeneous production costs.

We end this discussion by remarking that, depending on the parameter values, the threshold $\underline{v}$ in the above result may be negative, thus ruling out the case in which the information provider sells an undistorted signal to all firms. In fact, as the proof of the proposition highlights, $\underline{v} < 0$ whenever

$$\int_0^1 \frac{1}{c_i} di < -\frac{1}{\gamma_1}(1 + \kappa_z/\kappa_x),$$

which reduces to the condition of Proposition 2.3.4 when firms face identical production costs.

2.6.1 Cost dispersion and optimal information selling strategy

This subsection considers a setting in which the firms that compete in the downstream market can be of one of two types that differ in their production costs. In particular, type $i \in \{1, 2\}$ firms have production costs that take the form $C_i(q_i) = c_i q_i^2/2$ and the two types have equal mass. We let

$$\frac{1}{c_1} = \frac{1}{c} + \delta \qquad \text{and} \qquad \frac{1}{c_2} = \frac{1}{c} - \delta,$$

for some $\delta > 0$. Note that since the two types have equal mass, c is equal to the average cost coefficient in the market and δ can be viewed as a measure of dispersion in costs. The following proposition characterizes the effect of increasing the dispersion (increasing δ) on the optimal selling strategy.

Proposition 2.6.2. *Let κ_{s1}^* and κ_{s2}^* be the optimal signal precisions offered to firms of type 1 and type 2. Then, for any $\delta < 1/\left(c\sqrt{2}\right)$ we have:*

$$\frac{\partial \kappa_{s1}^*}{\partial \delta} \geq 0 \qquad \textit{and} \qquad \frac{\partial \kappa_{s2}^*}{\partial \delta} \leq 0.$$

	$\delta = .5$	$\delta = 1$	$\delta = 2$	$\delta = 4$
κ_{s1}^*	1.334	1.505	1.819	2
κ_{s2}^*	.976	.789	.409	0
Profits	1.067	1.076	1.114	1.251

Table 2.3: Optimal information selling strategy and equilibrium profits as a function of the cost dispersion between the two types of firms. For this example, we use the following set of parameters: $\kappa_x = 1,\ \kappa_z = 2,\ c = 1/6,\ \gamma_1 = 3/5,$ and $\gamma_0 = 10.$

Proposition 2.6.2 establishes that as the cost dispersion among the downstream firms increases, the provider finds it optimal to sell increasingly more accurate signals to the efficient type while decreasing the accuracy of the signals she sells to the type that has high production costs.

Table 2.3 reports a set of numerical results that shed additional light on the effect of cost dispersion on the provider's information selling strategy and equilibrium profits. In particular, when the dispersion between the firms' production costs is sufficiently high, the monopolist may find it optimal to exclude the less efficient type altogether from the information sale. Moreover, the provider's profits are increasing in the dispersion (although she may be selling to a smaller subset of firms).

2.7 Finite Number of Firms

To simplify the exposition and allow for a tractable analysis, most of the paper focused on an environment with a continuum of firms. In this section, we show that our qualitative insights regarding the monopolist's optimal information selling strategy carry over to a market consisting of finitely many firms. In particular, we focus on a Cournot oligopoly with a finite set of firms $N = \{1, ..., n\}$ which compete with one another in quantities. The inverse demand function in the market is given by

$$r = \gamma_0 \theta + \gamma_1 Q,$$

where $Q = \frac{1}{n}\sum_{i\in N} q_i$ is the average quantity, q_i is the quantity produced by firm i, r denotes the market price, and $\gamma_1 < 0$ is some constant. We assume that firm i faces quadratic production costs given by $c(q_i) = \gamma_2 q_i^2/2$. Therefore, firm i's profit can be expressed as

$$\pi_i = \gamma_0 q_i \theta + \frac{n-1}{n}\gamma_1 q_i Q_{-i} - \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right) q_i^2,$$

where $Q_{-i} = \frac{1}{n-1}\sum_{j\neq i} q_j$. Note that firm i's profit function has the same form as the one when a continuum of firms compete. Finally, the degree of strategic substitutability among firms' actions is given by

$$\beta_n = -\frac{\partial^2 \pi_i}{\partial q_i \partial Q_{-i}} / \frac{\partial^2 \pi_i}{\partial q_i^2} = \frac{n-1}{n}\frac{\gamma_1}{\gamma_2 - 2\gamma_1/n}.$$

As in the environment with a continuum of firms, we assume that each firm i observes a noisy private signal x_i about the realization of θ and that the information provider can offer a market signal s_i to firm i. We denote with K the set of firms that the information provider trades with, where $|K| = k \leq n$. Lemma 2.9.5 in the Appendix provides a complete characterization of the equilibrium for the competition subgame for any k, which can be viewed as the discrete analog of Proposition 2.3.1 in Section 2.3. Finally, for the remainder of this section, we assume that the provider sells an information signal to all firms, i.e., $K = N$, and we focus on characterizing the signal's precision that maximizes the provider's profit (note that as we argued in Section 2.3 there always exists an optimal information selling strategy that involves selling a signal to all firms). We obtain the following characterization for the optimal signal precision:

Proposition 2.7.1. *The optimal information selling strategy is given as follows:*

(a) If $\beta_n \geq -(1 + \kappa_z/\kappa_x)$, the information provider offers an undistorted version of her signal, i.e.,$\kappa_s^* = \kappa_z$, to all k firms.

(b) If $\beta_n < -(1 + \kappa_z/\kappa_x)$, the information provider offers a signal of lower precision, i.e.,

$$\kappa_s^* = -\frac{\kappa_z}{\beta_n + \kappa_z/\kappa_x} < \kappa_z.$$

Furthermore, the seller's expected profit is given as

$$\Pi^* = \begin{cases} n\ \alpha_n^2 \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right)\left(\frac{\kappa_z}{\kappa_x}\right)\frac{\kappa_z + \kappa_x}{[(1-\beta_n)\kappa_x + \kappa_z]^2} & \textit{if } \beta_n \geq -(1+\kappa_z/\kappa_x) \\ \\ n\ \alpha_n^2 \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right)\frac{\kappa_z}{-4\beta_n \kappa_x^2} & \textit{otherwise} \end{cases},$$

where $\alpha_n = \gamma_0 / \left(c - \frac{n+1}{n}\gamma_1\right)$.

Proposition 2.7.1 establishes that the insights underlying our main results remain unchanged when the downstream market is composed of a finite number of firms. Additionally, it is straightforward to verify that as n grows to infinity we recover the results of Section 2.3.3, both in terms of the optimal strategy as well as in terms of the profits for the information provider.[10] Finally, since $\beta_n \downarrow \beta$ we obtain that the set of markets for which the monopolist chooses to distort is increasing in n.

2.8 Conclusions

This paper considers the problem of selling information to a set of firms that compete in a downstream market. We establish that both the information provider's optimal selling strategy as well as her profits depend critically on the environment in which its cutomers operate. In particular, our results highlight that the extent of strategic substitutability and complementarity in the latter's actions has a first-order impact on the former's optimal strategy: when the firms' actions are strategic complements, the provider finds it optimal to sell an undistorted version of her information to the entire market of firms, whereas if their actions are strategic substitutes, the optimal strategy involves offering an inferior information product, and/or limiting the supply of information.

Our results are largely driven by the following trade-off: on the one hand, information about market conditions, e.g., demand realization, has always a direct positive effect on firms' profits as

[10]Note that as $n \to \infty$ expected profits diverge, however average profits $\frac{1}{n}\Pi^*$ converge to the result we obtained for a continuum of firms.

they can better align their actions with the underlying environment. On the other hand, however, in the presence of strategic substitutability among the firms, the provider's signal may have an additional (adverse) effect by increasing the correlation between the firms' actions. It turns out that this latter effect may dominate the former when firms' view their actions as strong strategic substitutes, in which case the provider finds it optimal to degrade the quality of her information products and/or exclude a subset of the firms from the sale.

We showcase the implications of our results in the context of Bertrand and Cournot competition thus complementing the extensive prior literature in operations management that explores vertical and horizontal information sharing in a supply chain. We extend our findings to the case when firms differ in their production costs and establish that the optimal selling strategy involves offering several information products with varying precisions and at different prices. In particular, we show that in equilibrium, the information provider offers more precise signals to the more efficient firms at higher prices in order to maximize her profit.

2.9 Proofs

With the exception of our results in Section 2.6, firms in our model are otherwise *ex ante* symmetric. Therefore, unless otherwise noted, we assume without loss of generality that the price offered by the provider to the firms is non-decreasing in the firms' index; that is, $p_i \geq p_j$ for $i > j$. Given that excluding a firm i from trade is equivalent to offering a price $p_i = \infty$, the above assumption also implies that the set of firms that are offered a contract by the provider is of the form $[0, \lambda]$ for some $\lambda \in [0, 1]$.

Let ℓ denote the fraction of firms who accept the provider's offer. In view of the above assumption, it is immediate that

$$\ell = \sup\{i \in [0, \lambda] : b_i = 1\},$$

and that $b_i = 1$ for all $i \leq \ell$.

Proof of Proposition 2.3.1

The first-order condition for firm i's problem with respect to action a_i is given by

$$\mathbb{E}\left[\frac{\partial}{\partial a_i}\pi(a_i, A, \theta)\middle|\mathcal{I}_i\right] = 0,$$

where $\mathcal{I}_i = \{x_i\}$ if $i \in [\ell, 1]$ and $\mathcal{I}_i = \{x_i, s_i\}$ if $i \in [0, \ell]$. Consequently,

$$a_i = \mathbb{E}\left[\beta A + (1-\beta)\alpha\theta | \mathcal{I}_i\right],$$

where $\beta = \gamma_1/\gamma_2$ is the degree of strategic complementarity in the downstream market as defined in (2.2) and $\alpha = \gamma_0/(\gamma_2 - \gamma_1)$. Thus, the firms' equilibrium actions are given by

$$a_i = \begin{cases} \mathbb{E}\left[\beta A + (1-\beta)\alpha\theta | x_i\right] & \forall i \in [\ell, 1], \\ \mathbb{E}\left[\beta A + (1-\beta)\alpha\theta | x_i, s_i\right] & \forall i \in [0, \ell]. \end{cases}$$

Noticing that $\mathbb{E}[\theta|x_i]$ is linear in x_i and $\mathbb{E}[\theta|x_i, s_i]$ is linear in x_i and s_i, we conjecture that equilibrium strategies are linear functions of x_i and s_i and then verify our hypothesis. In particular, we conjecture that

$$a_i = \begin{cases} c_0 x_i & \forall i \in [\ell, 1] \\ c_1 x_i + c_2 s_i & \forall i \in [0, \ell] \end{cases},$$

for some constants $c_0, c_1, c_2 \in \mathbb{R}$.

Replacing the candidate equilibrium strategy of an uninformed firm $i \in (\ell, 1]$ in its first-order condition yields

$$\begin{aligned} c_0 x_i &= \mathbb{E}\left[\beta\left(\int_0^\ell c_1 x_j + c_2 s_j dj + \int_\ell^1 c_0 x_j dj\right) + (1-\beta)\alpha\theta \Big| x_i\right] \\ &= [\beta\ell(c_1 + c_2) + \beta(1-\ell)c_0 + (1-\beta)\alpha]x_i, \end{aligned}$$

where we are using the fact that $\mathbb{E}[\theta|x_i] = \mathbb{E}[x_j|x_i] = \mathbb{E}[s_j|x_i] = x_i$. Consequently, the equilibrium strategy coefficients must satisfy $c_0 = \beta\ell(c_1+c_2)+\beta(1-\ell)c_0+(1-\beta)\alpha$ for any admissible $\ell \in [0, 1]$, which implies

$$c_0 = \alpha \quad \text{and} \quad c_1 + c_2 = \alpha. \tag{2.11}$$

On the other hand, replacing the candidate equilibrium strategy of an informed firm $i \in [0, \ell]$ in its first-order condition yields

$$c_1 x_i + (\alpha - c_1)s_i = \mathbb{E}\left[\beta\left(\int_0^\ell c_1 x_j + (\alpha - c_1)s_j dj + \int_\ell^1 \alpha x_j dj\right) + (1-\beta)\alpha\theta \Big| x_i, s_i\right], \tag{2.12}$$

where we use the expressions for the coefficients we derived in (2.11). We can now use the above expression to solve for c_1. To this end, note that let

$$\mathbb{E}[\theta|x_i, s_i] = \mathbb{E}[x_j|x_i, s_i] = \delta_1 x_i + (1-\delta_1)s_i$$
$$\mathbb{E}[s_j|x_i, s_i] = \delta_1(1-\rho)x_i + [1-\delta_1(1-\rho)]s_i,$$

where $\delta_1 = \kappa_x/(\kappa_x+\kappa_s)$. Consequently, we can rewrite (2.12) as

$$\begin{aligned} c_1 x_i + (\alpha - c_1)s_i =& [\beta\ell c_1\delta_1 + \beta\ell(\alpha-c_1)\delta_1(1-\rho) + \beta(1-\ell)\alpha\delta_1 + (1-\beta)\alpha\delta_1]x_i \\ &+ [\beta\ell c_1(1-\delta_1) + \beta\ell(\alpha-c_1)(1-\delta_1(1-\rho)) + \beta(1-\ell)\alpha(1-\delta_1) + (1-\beta)\alpha(1-\delta_1)]s_i, \end{aligned}$$

and conclude that the equilibrium coefficient c_1 must satisfy

$$c_1 = \beta\ell c_1\delta_1\rho + \alpha\delta_1(1-\beta\ell\rho).$$

Solving for c_1 thus implies that

$$c_1 = \alpha\frac{(1-\beta\ell\rho)\kappa_x}{(1-\beta\ell\rho)\kappa_x+\kappa_s},$$

and hence,

$$c_2 = \alpha - c_1 = \alpha\frac{\kappa_s}{(1-\beta\ell\rho)\kappa_x+\kappa_s}.$$

Combining the above, we conclude that firms' actions at equilibrium are given by

$$a_i = \begin{cases} \alpha x_i & \forall i \in [\ell, 1] \\ \alpha\frac{(1-\beta\ell\rho)\kappa_x}{(1-\beta\ell\rho)\kappa_x+\kappa_s}x_i + \alpha\frac{\kappa_s}{(1-\beta\ell\rho)\kappa_x+\kappa_s}s_i & \forall i \in [0, \ell] \end{cases},$$

completing the proof. $\square$

Two Auxiliary Lemmas

We state and prove two lemmas that we use in the remainder of the appendix. The first lemma characterizes the expected surplus of an informed firm, whereas the second lemma shows that, for any given λ, the provider always finds it optimal to charge a constant price to all firms $i \in [0, \lambda]$.

Lemma 2.9.1. *The expected surplus of each firm from buying the market signal is given by*

$$\Delta(\ell, \kappa_s, \rho, \kappa_x) = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_s}{\kappa_x}\right) \frac{\kappa_s + \kappa_x}{[(1-\beta\ell\rho)\kappa_x + \kappa_s]^2}, \tag{2.13}$$

where ℓ denotes the fraction of informed firms.

Proof. Let $a_i^1 := \alpha \dfrac{\kappa_s s_i + (1-\beta\ell\rho)\kappa_x x_i}{\kappa_s + (1-\beta\ell\rho)\kappa_x}$ denote the equilibrium action of an informed firm and let $a_i^0 := \alpha x_i$ denote the equilibrium action of an uninformed firm. Recall that ℓ denotes the fraction of informed firms, and thus the aggregate equilibrium action is $A = \int_0^\ell a_i^1 di + \int_\ell^1 a_i^0 di$. By replacing the equilibrium actions in the expressions for the firms' payoffs and then taking the expectations conditional on θ, we get

$$\mathbb{E}\left[\pi(a_i^1, A, \theta)|\theta\right] = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left[\theta^2 + \frac{2\beta\ell\rho\kappa_s}{[(1-\beta\ell\rho)\kappa_x + \kappa_s]^2} - \frac{(1-\beta\ell\rho)^2\kappa_x + \kappa_s}{[(1-\beta\ell\rho)\kappa_x + \kappa_s]^2}\right], \tag{2.14}$$

and

$$\mathbb{E}\left[\pi(a_i^0, A, \theta)|\theta\right] = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left[\theta^2 - \frac{1}{\kappa_x}\right]. \tag{2.15}$$

Next note that we can use the two conditional expectations (2.14) and (2.15) to compute the (unconditional) expectation for a firm's surplus given by

$$\Delta := \mathbb{E}\left[\pi(a_i^1, A, \theta)\right] - \mathbb{E}\left[\pi(a_i^0, A, \theta)\right].$$

Applying the law of total expectation yields

$$\Delta = \alpha^2 \left(\frac{\gamma_2}{2}\right)\left(\frac{\kappa_s}{\kappa_x}\right)\frac{\kappa_s + \kappa_x}{[(1-\beta\ell\rho)\kappa_x + \kappa_s]^2},$$

which completes the proof of the lemma. □

Lemma 2.9.2. *The provider sets $p_i = p^*(\lambda)$ for all $i \in [0, \lambda]$, where $p^*(\lambda)$ is equal to the expected equilibrium surplus of an informed firm when the fraction of informed firms is λ. Furthermore, $p^*(\lambda)$ is such that all firms that receive the provider's offer accept in equilibrium, thus $\ell = \lambda$.*

Proof. Consider the simultaneous game of accepting/rejecting the provider's offer. Recall that in such game each firm $i \in [0, \lambda]$ accepts the offer if her expected surplus is bigger than her individual price p_i while taking the decisions of the rest of the firms as given.

We suppose that a fraction $\ell \in [0, \lambda]$ of firms has accepted the provider's offer, and we write the optimal decision of each firm $i \in [0, \lambda]$ as a function of firm's i individual price. We have

$$b_i(p_i) = \begin{cases} 1 & \text{if} \quad \Delta(\ell) > p_i \\ 0 & \text{if} \quad \Delta(\ell) < p_i \\ \in \{0,1\} & \text{if} \quad \Delta(\ell) = p_i \end{cases},$$

where $\Delta(\ell)$ is given by equation (2.13) and denotes the expected surplus of an informed firm when a fraction ℓ is informed.

We can write the provider's optimization problem as follows

$$\begin{aligned} \max_{\{p_i\}_{i\in[0,\lambda]}} \quad & \int_0^\lambda p_i b_i(p_i)\,di \\ s.t. \quad & b_i(p_i) = \begin{cases} 1 & \text{if} \quad \Delta(\ell) > p_i \\ 0 & \text{if} \quad \Delta(\ell) < p_i \\ \in \{0,1\} & \text{if} \quad \Delta(\ell) = p_i \end{cases}, \quad \forall i \in [0,\lambda]. \end{aligned} \tag{2.16}$$

Before solving for the provider's optimal selling strategy, we rewrite the set of constraints (2.16) as

$$\begin{cases} \sup\limits_{i\in[0,\lambda]} p_i \leq \Delta(\lambda) & \text{if} \quad \ell = \lambda \\ \inf\limits_{i\in[0,\lambda]} p_i \geq \Delta(0) & \text{if} \quad \ell = 0 \\ \int_0^\lambda \mathbb{I}_{\{p_i<\Delta(\ell)\}} di \leq \ell \leq \int_0^\lambda \mathbb{I}_{\{p_i\leq\Delta(\ell)\}} di & \text{if} \quad \ell \in (0,\lambda) \end{cases},$$

Recall that without loss of generality the pricing schedule $p : [0, \lambda] \to \mathbb{R}_+$ is non-decreasing, thus we can further simplify the set of constraints as

$$\begin{cases} p_\lambda \leq \Delta(\lambda) & \text{if} \quad \ell = \lambda & (2.17\text{a}) \\ p_0 \geq \Delta(0) & \text{if} \quad \ell = 0 & (2.17\text{b}) \\ p_\ell \leq \Delta(\ell) \quad \text{and} \quad p_{\ell+} \geq \Delta(\ell) & \text{if} \quad \ell \in (0,\lambda). & (2.17\text{c}) \end{cases}$$

The proof proceeds by showing that for any equilibrium of the subgame that results from a fraction ℓ of the firms accepting the provider's offer, there exists an optimal pricing schedule such that $p_i = \Delta(\ell)$ for all $i \leq \ell$ and $p_i = \infty$ for all $i > \ell$. There are three cases to consider.
First, for case (2.17a), the problem simplifies to

$$\max_{\{p_i\}_{i\in[0,\lambda]}} \quad \int_0^\lambda p_i di$$
$$s.t. \quad p_\lambda \leq \Delta(\lambda).$$

In this case a fraction $\ell = \lambda$ of firms accepts and as we show below it is optimal for the provider to set $p_i = \Delta(\lambda)$ for all $i \in [0, \lambda]$. Suppose, for the sake of contradiction, that p is optimal but $u := \sup\{i \in [0, \lambda] : p_i < \Delta(\lambda)\} \geq 0$. If $u = 0$, then we have $p_i = \Delta(\lambda)$ except for a set of measure 0, so this case is immaterial. If $u > 0$, the maintained assumption that p is non-decreasing implies that

$$p_i < \Delta(\lambda), \ \forall i < u \quad \text{and} \quad p_i = \Delta(\lambda), \ \forall i \geq u.$$

This implies that we can construct pricing schedule p' such that

$$p_i < p'_i \leq \Delta(\lambda),\ \forall i < u \quad \text{and} \quad p'_i = p_i,\ \forall i \geq u,$$

that is feasible and achieves a higher objective value. Thus, it must be that $p_i = \Delta(\lambda)$ for all $i \leq \lambda$. For case (2.17b), $\ell = 0$ and the objective function is always equal to 0. Thus, p can be chosen such that $p_i = \infty$ for all $i \in [0, \lambda]$.

Finally, for case (2.17c), the problem simplifies to

$$\begin{aligned} \max_{\{p_i\}_{i\in[0,\lambda]}} \quad & \int_0^\ell p_i di \\ s.t. \quad & p_\ell \leq \Delta(\ell) \\ & p_{\ell+} \geq \Delta(\ell). \end{aligned}$$

First, we show that the provider can always set $p_i = \infty, \forall i > \ell$. Note that the individual price of each firm $i > \ell$ does not affect the objective function of the provider. This implies that all feasible solutions p that differ only on $(\ell, \lambda]$ attain the same objective value, so it is without loss of generality to focus on solutions that are such that $p_i = \infty$ for all $i > \ell$. Next, we show that $p_i = \Delta(\ell), \forall i \leq \ell$. Suppose, for the sake of contradiction, that p is optimal but $u := \sup\{i \in [0, \ell] : p_i < \Delta(\ell)\} \geq 0$. If $u = 0$ we have $p_i = \Delta(\ell)$, except for a set of measure 0. If $u > 0$, the assumption that p is non-decreasing implies that

$$p_i < \Delta(\ell),\ \forall i < u \quad \text{and} \quad p_i = \Delta(\ell),\ \forall i \geq u,$$

which in turn implies that we can construct a pricing schedule p'' such that

$$p_i < p''_i \leq \Delta(\ell),\ \forall i < u \quad \text{and} \quad p''_i = p_i,\ \forall i \geq u,$$

that is feasible and achieves a higher objective value. Thus, it must be that $p_i = \Delta(\ell)$ for all $i \leq \ell$.

Thus, there always exists an optimal pricing schedule such that $p_i = \Delta(\ell)$ for all $i \leq \ell$ and

$p_i = \infty$ for all $i > \ell$, which implies that only a fraction ℓ of firms accepts the provider's offer and the latter's optimal profit is $\ell \cdot \Delta(\ell)$. Without loss of generality the provider sets $\lambda = \ell$ and $p_i = \Delta(\lambda)$ for all $i \in [0, \lambda]$. Thus, all firms accept her offer in equilibrium and the provider's profit is given by $\lambda \cdot \Delta(\lambda)$. Setting $p^*(\lambda) = \Delta(\lambda)$ completes the proof. $\square$

Proof of Proposition 2.3.2

By Lemma 2.9.2, the provider's problem simplifies to choosing λ, κ_y and ρ in order to maximize the expected profit $\Pi := \lambda \cdot p^*(\lambda, \kappa_s, \rho, \kappa_x) = \lambda \cdot \Delta(\lambda, \kappa_s, \rho, \kappa_x)$, subject to the constraints imposed by the information structure. Replacing the expected surplus (2.13) into the objective function yields

$$\Pi(\lambda, \kappa_s, \rho, \kappa_x) = \lambda\ \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_s}{\kappa_x}\right) \frac{\kappa_s + \kappa_x}{[(1 - \beta\lambda\rho)\kappa_x + \kappa_s]^2}, \tag{2.18}$$

and thus the provider's problem can be rewritten as

$$\begin{aligned} \max_{\rho, \kappa_s, \lambda} \quad & \Pi(\lambda, \kappa_s, \rho, \kappa_x) \\ s.t. \quad & \frac{\kappa_s}{\kappa_z} \leq \rho \leq 1 \\ & \kappa_s \leq \kappa_z \\ & 0 \leq \lambda \leq 1. \end{aligned} \tag{2.19}$$

Note that the partial derivative of Π with respect to ρ, i.e.,

$$\frac{\partial \Pi}{\partial \rho} = \lambda\ \alpha^2 \gamma_2 \frac{\beta\lambda\kappa_s(\kappa_x + \kappa_s)}{[(1 - \beta\lambda\rho)\kappa_x + \kappa_s]^3}, \tag{2.20}$$

is positive for $\beta \in (0, 1/2)$, which implies that $\rho^* = 1$. Replacing this into (2.18) and differentiating with respect to λ yields

$$\frac{\partial \Pi}{\partial \lambda} = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_s}{\kappa_x}\right) \frac{(\kappa_x + \kappa_x)[(1 + \beta\lambda)\kappa_x + \kappa_s]}{[(1 - \beta\lambda)\kappa_x + \kappa_s]^3}. \tag{2.21}$$

Similarly, the partial derivative with respect to κ_s is given by

$$\frac{\partial \Pi}{\partial \kappa_s} = \lambda\ \alpha^2 \left(\frac{\gamma_2}{2}\right) \frac{(1-\beta\lambda)\kappa_x + (1-2\beta\lambda)\kappa_s}{[(1-\beta\lambda)\kappa_x + \kappa_s]^3}. \tag{2.22}$$

In addition, note that (2.21) and (2.22) are positive for $\beta \in (0, 1/2)$, so the provider finds it optimal to set $\lambda^* = 1$ and $\kappa_s^* = \kappa_z$. Replacing ρ^*, λ^* and κ_s^* into (2.18) yields

$$\Pi^* = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_z}{\kappa_x}\right) \frac{\kappa_z + \kappa_x}{[(1-\beta)\kappa_x + \kappa_z]^2}, \tag{2.23}$$

which completes the proof. □

Proof of Proposition 2.3.3

Consider the provider's expected profit (2.18) and her profit-maximization problem (2.19), and let $-(1 + \kappa_z/\kappa_x) \le \beta < 0$. In this case, the partial derivative of Π with respect to ρ given in (2.20) is negative, which implies that the provider finds it optimal to set the level of correlation to its minimum, i.e., $\rho_\xi^* = 0$ or $\rho^* = \kappa_s/\kappa_z$. Replacing this into (2.18) and differentiating with respect to κ_s yields

$$\frac{\partial \Pi}{\partial \kappa_s} = \lambda\ \alpha^2 \left(\frac{\gamma_2}{2}\right) \frac{(1+\beta\lambda\kappa_s/\kappa_z)\kappa_x + \kappa_s}{[(1-\beta\lambda\kappa_s/\kappa_z)\kappa_x + \kappa_s]^3}, \tag{2.24}$$

while differentiating with respect to λ yields

$$\frac{\partial \Pi}{\partial \lambda} = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_s}{\kappa_x}\right) \frac{(\kappa_x + \kappa_s)[(1+\beta\lambda\kappa_s/\kappa_z)\kappa_x + \kappa_s]}{[(1-\beta\lambda\kappa_s/\kappa_z)\kappa_x + \kappa_s]^3}. \tag{2.25}$$

The assumption on β implies that (2.24) and (2.25) are positive, which results in $\lambda^* = 1$ and $\kappa_s^* = \kappa_z$. The proof follows by replacing the optimal values for ρ^*, λ^*, and κ_s^* into expression (2.18). □

Proof of Proposition 2.3.4

Consider the provider's expected profit (2.18) and her profit-maximization problem (2.19), and let $\beta < -(1 + \kappa_z/\kappa_x)$. First, note that in this case (2.20) is negative, so the provider finds it optimal to set $\rho^* = \kappa_s/\kappa_z$. Replacing this into (2.18) and differentiating with respect to κ_s and λ we again obtain (2.24) and (2.25) respectively. Both (2.24) and (2.25) are equal to 0 if and only if

$$(\kappa_z + \beta\lambda\kappa_s)\kappa_x + \kappa_z\kappa_s = 0. \tag{2.26}$$

Moreover, Π is unimodal in both κ_s and λ, which implies that the set of optimal allocations (κ_s^*, λ^*) is given by the solutions to (2.26).

Finally, replacing $\kappa_s^*, \rho^* = \kappa_s^*/\kappa_z$ and $\lambda^* = \dfrac{(\kappa_x + \kappa_s^*)\kappa_z}{-\beta\kappa_x\kappa_s^*}$ into (2.18) yields

$$\Pi^* = -\alpha^2 \left(\frac{\gamma_2}{2}\right) \frac{\kappa_z}{4\beta\kappa_x^2},$$

which completes the proof. □

Proof of Proposition 2.4.1

We solve the game by backward induction, i.e., first, we characterize the firms' equilibrium actions in the competition subgame that results from a (subset) of them obtaining the provider's information signal; then, we solve for their acceptance/rejection decisions; and, finally, we turn to the provider's problem and complete the proofs of parts (a) and (b) of the proposition.

Recall that the provider possesses a signal $z = \theta + \zeta$, with $\zeta \sim N(0, 1/\kappa_z)$, and offers to firm $i \in [0, 1]$ a signal $s_i = z + \xi_i$ with $\xi_i \sim N(0, \kappa_{\xi i})$. Without loss of generality we assume that the provider does not add any correlation to the signal she sells, i.e., $\mathrm{corr}(\xi_i, \xi_j) = 0$. The market signal s_i offered to firm $i \in [0, 1]$ can be rewritten as

$$s_i = \theta + \eta_i, \quad \eta_i \sim N(0, 1/\kappa_{si}),$$

where $\kappa_s = (1/\kappa_z + 1/\kappa_{\xi i})^{-1}$ and $\mathrm{Cov}(s_i, s_j) = 1/\kappa_z$.

We have the following auxiliary lemma.

Lemma 2.9.3. *The competition subgame between the firms has a unique Bayes-Nash equilibrium in linear strategies, given by*

$$a(\kappa_{si}, \kappa_{\mathbf{s-i}}) = \alpha[(1-\omega_i)x_i + \omega_i s_i] \quad \textit{for all} \quad i \in [0,1],$$

where

$$\omega_i = \left(\frac{\kappa_{si}}{\kappa_x + \kappa_{si}}\right) \Big/ \left(1 - \beta\frac{\kappa_x}{\kappa_z}\int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di\right)$$

and $\alpha = \gamma_0/(\gamma_2 - \gamma_1)$.

Proof. The first-order optimality condition of firm i with respect to action a_i implies that in equilibrium

$$a_i = \mathbb{E}[\beta A + (1-\beta)\alpha\theta | x_i, s_i]. \tag{2.27}$$

Assume that each firm $i \in [0,1]$ uses a linear strategy $c_i x_i + h_i s_i$, for some constants $c_i, h_i \in \mathbb{R}$. Then, we can rewrite the equilibrium condition (2.27) as

$$c_i x_i + h_i s_i = \mathbb{E}\left[\beta \int_0^1 (c_j x_j + h_j s_j) dj + (1-\beta)\alpha\theta \middle| x_i, s_i\right]. \tag{2.28}$$

Using equations

$$\mathbb{E}[s_j | x_i, s_i] = \frac{\kappa_x(1 - \kappa_{si}/\kappa_z)}{\kappa_x + \kappa_{si}} x_i + \frac{\kappa_{si}(1 + \kappa_x/\kappa_z)}{\kappa_x + \kappa_{si}} s_i,$$

and

$$\mathbb{E}[\theta | x_i, s_i] = \mathbb{E}[x_j | x_i, s_i] = \frac{\kappa_x}{\kappa_x + \kappa_{si}} x_i + \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} s_i,$$

which are obtained using the formula for the conditional expectation of Gaussian random vectors, we can rewrite (2.28) as

$$
\begin{aligned}
c_i x_i + h_i s_i =& \beta \left[\left(\frac{\kappa_x}{\kappa_x + \kappa_{si}} x_i + \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} s_i \right) \int_0^1 c_j dj + \left(\frac{\kappa_x (1 - \kappa_{si}/\kappa_z)}{\kappa_x + \kappa_{si}} x_i + \frac{\kappa_{si}(1 + \kappa_x/\kappa_z)}{\kappa_x + \kappa_{si}} s_i \right) \int_0^1 h_j dj \right] \\
&+ (1-\beta)\alpha \left(\frac{\kappa_x}{\kappa_x + \kappa_{si}} x_i + \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} s_i \right).
\end{aligned}
$$

Note that the equilibrium coefficients $(c_i, , h_i)$ for $i \in [0,1]$, must solve the following sets of equations

$$
c_i = \beta \frac{\kappa_x}{\kappa_x + \kappa_{si}} \int_0^1 c_j dj + \beta \frac{\kappa_x (1 - \kappa_{si}/\kappa_z)}{\kappa_x + \kappa_{si}} \int_0^1 h_j dj + (1-\beta)\alpha \frac{\kappa_x}{\kappa_x + \kappa_{si}} \quad \forall i \in [0,1], \tag{2.29}
$$

and

$$
h_i = \beta \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} \int_0^1 c_j dj + \beta \frac{\kappa_{si} (1 + \kappa_x/\kappa_z)}{\kappa_x + \kappa_{si}} \int_0^1 h_j dj + (1-\beta)\alpha \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} \quad \forall i \in [0,1]. \tag{2.30}
$$

Integrating with respect to di over $[0,1]$ in (2.29) and (2.30) yields a linear-system of two equations, which implies that

$$
\int_0^1 c_i di = \alpha \left(1 - \left(1 + \beta \frac{\kappa_x}{\kappa_z} \right) \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di \right) \Big/ \left(1 - \beta \frac{\kappa_x}{\kappa_z} \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di \right)
$$

and

$$
\int_0^1 h_i di = \alpha \left(\int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di \right) \Big/ \left(1 - \beta \frac{\kappa_x}{\kappa_z} \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di \right).
$$

Replacing the above expressions into (2.29) and (2.30) yields

$$
c_i = \alpha \frac{\kappa_x}{\kappa_x + \kappa_{si}} \left(1 - \beta \frac{\kappa_x + \kappa_{si}}{\kappa_z} \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di \right) \Big/ \left(1 - \beta \frac{\kappa_x}{\kappa_z} \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di \right)
$$

and

$$
h_i = \alpha \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} \Big/ \left(1 - \beta \frac{\kappa_x}{\kappa_z} \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di \right).
$$

Finally, noting that $c_i + h_i = \alpha$ and setting $h_i = \alpha\omega_i$ completes the proof. □

The next step in our analysis involves studying the firms' acceptance/rejection decisions that precede the competition subgame. We restrict attention to subgame perfect equilibria in which all firms accept the provider's offers. This is without loss of generality, since the case in which there is a firm i that rejects the provider's offer is surplus-equivalent to the case in which the provider offers a signal of precision $\kappa_{si} = 0$ at price $p_i = 0$ to firm i, and firm i accepts the offer.

The equilibrium acceptance/rejection decisions can be characterized as follows. Each firm $i \in [0,1]$ accepts the provider's offer if

$$\Delta_i = \mathbb{E}[\pi(a(\kappa_{si}, \kappa_{\mathbf{s-i}}))] - \mathbb{E}[\pi(a(0, \kappa_{\mathbf{s-i}}))] \geq p_i,$$

i.e., if price p_i is lower than the expected surplus of firm i. Thus, it is optimal for the provider to offer $p_i = \Delta_i$ for all $i \in [0,1]$.

Using the equilibrium characterization from Lemma 2.9.3, we can compute the expected surplus Δ_i of firm i, which, in turn, is equal to price p_i, i.e.,

$$p_i = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_{si}}{\kappa_x + \kappa_{si}}\right) \left(\frac{1}{\kappa_x}\right) \Big/ \left[1 - \beta \frac{\kappa_x}{\kappa_z} \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di\right]^2. \tag{2.31}$$

The provider's expected equilibrium profit is given by

$$\Pi(\kappa_{\mathbf{s}}, \beta) = \int_0^1 p_i di = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{1}{\kappa_x} \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di\right) \Big/ \left[1 - \beta \frac{\kappa_x}{\kappa_z} \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di\right]^2, \tag{2.32}$$

and her problem can now be simply written as

$$\begin{aligned} \max_{\{\kappa_{si}\}_{i \in [0,1]}} \quad & \Pi(\kappa_{\mathbf{s}}, \beta) \\ s.t. \quad & 0 \leq \kappa_{si} \leq \kappa_z \quad \forall i \in [0,1]. \end{aligned} \tag{2.33}$$

The following lemma allows us to further simplify the optimization problem above and characterize the set of optimal solutions.

Lemma 2.9.4. *The objective function of problem (2.33) depends on* $\{\kappa_{si}\}_{i\in[0,1]}$ *only through a constant*

$$D = \int_0^1 \frac{\kappa_{si}}{\kappa_x + \kappa_{si}} di. \tag{2.34}$$

Furthermore, for any optimal solution $\{\kappa_{si}^*\}_{i\in[0,1]}$ *of problem (2.33), there exist a constant solution* $\bar{\kappa}_s$ *that is feasible and achieves the same objective value of* $\{\kappa_{si}^*\}_{i\in[0,1]}$*.*

Proof. The first statement follows directly from expression (2.32). For the second statement, let $\{\kappa_{si}^*\}_{i\in[0,1]}$ be an optimal solution of problem (2.33), with corresponding $D^* = \int_0^1 \frac{\kappa_{si}^*}{\kappa_x + \kappa_{si}^*} di$. Define constant $\bar{\kappa}_s$ as $\bar{\kappa}_s := \frac{D^*\kappa_x}{1-D^*}$. Note that $\frac{\bar{\kappa}_s}{\kappa_x + \bar{\kappa}_s} = D^*$, which implies that $\bar{\kappa}_s$ achieves the same objective value as $\{\kappa_{si}^*\}_{i\in[0,1]}$. Finally, we need to verify that $\bar{\kappa}_s$ is feasible. By the feasibility of $\{\kappa_{si}^*\}_{i\in[0,1]}$, i.e., $0 \le \kappa_{si}^* \le \kappa_z$ for all $i \in [0,1]$, it follows that $0 \le D^* \le \frac{\kappa_z}{\kappa_x + \kappa_z}$ and thus $0 \le \bar{\kappa}_s \le \kappa_z$. This implies that the constant $\bar{\kappa}_s$ is feasible and it achieves the maximum objective value, which completes the proof. □

Lemma 2.9.4 allows us to solve a simplified problem, in which the provider offers a signal of precision κ_s to all firms $i \in [0,1]$. Furthermore, using the optimal value for κ_s together with equation (2.34) allows us to characterize the set of optimal solutions for the original problem (2.33). In particular, replacing κ_s for κ_{si} in problem (2.33), the provider's problem simplifies to

$$\begin{aligned} \max_{\kappa_s} \quad & \Pi = \alpha^2 \left(\frac{\gamma_2}{2}\right)\left(\frac{\kappa_s}{\kappa_x}\right) \frac{\kappa_s + \kappa_x}{[(1-\beta\kappa_s/\kappa_z)\kappa_x + \kappa_s]^2} \\ s.t. \quad & 0 \le \kappa_s \le \kappa_z. \end{aligned} \tag{2.35}$$

Proof of part (a): Let $\beta \ge -(1+\kappa_z/\kappa_x)$ and consider the simplified problem (2.35). Differentiating the objective with respect to κ_s yields

$$\frac{\partial \Pi}{\partial \kappa_s} = \alpha^2 \left(\frac{\gamma_2}{2}\right) \frac{(1+\beta\kappa_s/\kappa_z)\kappa_x + \kappa_s}{[(1-\beta\kappa_s/\kappa_z)\kappa_x + \kappa_s]^3}. \tag{2.36}$$

The assumption on β implies that (2.36) is positive, which means that it is optimal to set $\kappa_s^* = \kappa_z$.

By Lemma 2.9.4, this implies that any solution $\{\kappa_{si}^*\}_{i\in[0,1]}$ to problem (2.33) that is feasible and such that

$$\int_0^1 \frac{\kappa_{si}^*}{\kappa_x + \kappa_{si}^*} di = \frac{\kappa_z}{\kappa_x + \kappa_z},$$

is an optimal solution. Thus, problem (2.33) has a unique optimal solution in this case, i.e.,$\kappa_{si}^* = \kappa_z, \forall i \in [0,1]$. Replacing this solution into (2.31) we obtain

$$p_i^* = \alpha^2 \left(\frac{\gamma_2}{2}\right) \left(\frac{\kappa_z}{\kappa_x}\right) \frac{\kappa_z + \kappa_x}{[(1-\beta)\kappa_x + \kappa_z]^2} = p^*.$$

Proof of part (b): Let $\beta < -(1 + \kappa_z/\kappa_x)$ and consider problem (2.35). In this case, the partial derivative given in (2.36) evaluated at $\kappa_s^* = \kappa_z$ is negative, so the provider is better off by offering noisy signals to the firms. Solving for the optimal κ_s using a firm's first-order optimality condition yields

$$\kappa_s^* = \frac{\kappa_x}{-(1 + \beta\kappa_x/\kappa_z)} < \kappa_z.$$

By Lemma 2.9.4, this implies that any solution $\{\kappa_{si}^*\}_{i\in[0,1]}$ to problem (2.33) that is feasible and such that

$$\int_0^1 \frac{\kappa_{si}^*}{\kappa_x + \kappa_{si}^*} di = -\frac{\kappa_z}{\beta\kappa_x}, \tag{2.37}$$

is an optimal solution. Finally, replacing (2.37) into (2.31) yields

$$p_i^* = \alpha^2 \left(\frac{\gamma_2}{2}\right) \frac{\kappa_{si}^*}{4(\kappa_x + \kappa_{si}^*)\kappa_x}.$$

$\square$

Proof of Proposition 2.5.1

Throughout this proof we rescale firms' profit so that $\alpha = \gamma_0/(\gamma_2 - \gamma_1) = 1$, this is without loss of generality and it simplifies the notation. We conjecture that equilibrium strategies are linear in x_i, s_i, and S_i, and then we verify our hypothesis. In particular, we conjecture that

$$a(x_i, s_i, S_i) = b_1 x_i + b_2 S_i + b_3 s_i.$$

By definition $A = \int_0^1 a(x_i, s_i, S_i)\, di$, we thus have $A = b_1\theta + b_2 A + b_3 z$ which implies that

$$A = \frac{b_1\theta + b_3 z}{1 - b_2}.$$

Using the above equation for A, firm's i expected profit simplifies to

$$\begin{aligned}\mathbb{E}[\pi_i|x_i, s_i, S_i] &= \gamma_0 a_i \mathbb{E}[\theta|x_i, s_i, S_i] + \gamma_1 a_i \frac{b_1\mathbb{E}[\theta|x_i, s_i, S_i] + b_3\mathbb{E}[z|x_i, s_i, S_i]}{1 - b_2} - \frac{\gamma_2}{2} a_i^2 \\ &= a_i \mathbb{E}[\theta|x_i, s_i, S_i]\left(\gamma_0 + \frac{\gamma_1 b_1}{1 - b_2}\right) + a_i \mathbb{E}[z|x_i, s_i, S_i]\left(\frac{\gamma_1 b_3}{1 - b_2}\right) - \frac{\gamma_2}{2} a_i^2.\end{aligned}$$

Taking the first-order condition with respect to a_i, $\frac{\partial}{\partial a_i}\mathbb{E}[\pi_i|x_i, s_i, S_i] = 0$, and recalling that $z = \zeta + \theta$, we can express the equilibrium action of firm i as

$$a_i = \left[(1-\beta) + \beta\frac{b_1 + b_3}{1 - b_2}\right] E[\theta|x_i, s_i, S_i] + \beta\left(\frac{b_3}{1 - b_2}\right) E[\zeta|x_i, s_i, S_i]. \tag{2.38}$$

Before proceeding, we make a change of variable, setting $\kappa_s = t\kappa_z$ with $t \in [0, 1]$.

The conditional expectations in (2.38) are given respectively by

$$\begin{aligned}E[\theta|x_i, s_i, S_i] \quad &= \quad \kappa_x \frac{(1 - b_2)^2\kappa_z + (1 - t)b_3^2\kappa_\nu}{D} x_i + \kappa_\nu \frac{(1 - b_2)\,[b_1 + (1 - t)b_3]\,\kappa_z}{D} S_i + \\ &\quad + t\kappa_z \frac{(1 - b_2)^2\kappa_z - b_1 b_3 \kappa_\nu}{D} z_i, \end{aligned} \tag{2.39}$$

and

$$
\begin{aligned}
E[\zeta|x_i,s_i,S_i] &= -\kappa_x\frac{(1-b_2)^2 t\kappa_z + b_3(b_1+b_3)(1-t)\kappa_\nu}{D}x_i + \kappa_\nu\frac{(1-b_2)[b_3(1-t)\kappa_x - b_1 t\kappa_z]}{D}S_i + \\
&\quad + t\kappa_z\frac{b_1(b_1+b_3)\kappa_\nu + (1-b_2)^2\kappa_x}{D}z_i, \qquad (2.40)
\end{aligned}
$$

where

$$
D = \kappa_\nu\left[b_1^2\kappa_z + b_3(1-t)(b_3(\kappa_x+\kappa_z) + 2b_1\kappa_z)\right] + \kappa_z(1-b_2)^2(t\kappa_z+\kappa_x),
$$

which are obtained using the formula for the conditional expectation of Gaussian random vectors. Substituting (2.39) and (2.40) into (2.38), and identifying b_1, b_2, b_3, yields the following system of equations, which the equilibrium coefficients must satisfy:

$$
\begin{aligned}
b_1 &= \kappa_x\frac{(1-\beta)b_3^2(1-t)\kappa_\nu + (1-b_2)[\beta b_1 + (1-\beta)(1-b_2) + \beta b_3(1-t)]\kappa_z}{D} \qquad (2.41)\\
b_2 &= \kappa_\nu\frac{b_1\left\{[\beta b_1 + (1-\beta)(1-b_2)] + \beta b_3^2(1-t) + b_3(1-t)[2\beta b_1 + (1-\beta)(1-b_2)]\right\}\kappa_z + \beta b_3^2(1-t)\kappa_x}{D} \\
& \qquad (2.42)\\
b_3 &= t\kappa_z\frac{(1-b_2)[\beta b_1 + (1-\beta)(1-b_2)]\kappa_z + b_3[\beta(1-b_2)(\kappa_x+\kappa_z) - (1-\beta)b_1\kappa_\nu]}{D}. \qquad (2.43)
\end{aligned}
$$

Let $a_i^1 := b_1 x_i + b_2 S_i + b_3 s_i$ denote the equilibrium action of firm i and let $a_i^0 := \tilde{b}_1 x_i + \tilde{b}_2 S_i$ denote the action that firm i takes if she deviates from the equilibrium path and does not observe s_i, her information set is thus (x_i, S_i) and the coefficients $\tilde{b}_1, \tilde{b}_2$ can be characterized following the same procedure used for b_1, b_2, b_3. We have

$$
\begin{aligned}
\tilde{b}_1 &= \kappa_x\frac{(1-\beta)b_3^2\kappa_\nu + (1-b_2)[\beta b_1 + (1-\beta)(1-b_2) + \beta b_3]\kappa_z}{\kappa_x\left[b_3^2\kappa_\nu + (1-b_2)^2\kappa_z\right] + (b_1+b_3)^2\kappa_\nu\kappa_z}, \qquad (2.44)\\
\tilde{b}_2 &= \kappa_\nu\frac{\beta b_3^2\kappa_x + (b_1+b_3)[\beta b_1 + (1-\beta)(1-b_2) + \beta b_3]\kappa_z}{\kappa_x\left[b_3^2\kappa_\nu + (1-b_2)^2\kappa_z\right] + (b_1+b_3)^2\kappa_\nu\kappa_z}, \qquad (2.45)
\end{aligned}
$$

note that this coefficients must depend on b_1, b_2, b_3, since they are derived under the assumption that firms i deviates, while all other firms are playing their equilibrium strategy. Using a_i^1 and a_i^0

we can characterize the expected profits of a firm that observes (x_i, s_i, S_i) as

$$E[\pi_i(a_i^1, A)] = \gamma_2 \left[\frac{1}{2\kappa_\theta} + (\beta - 1/2) \left(\frac{1}{\kappa_z} \right) \frac{b_3^2}{(b_1 + b_3)^2} - \frac{b_1^2}{2\kappa_x} - \frac{b_2^2}{2\kappa_\nu} - \left(\frac{1-t}{t\kappa_z} \right) \frac{b_3^2}{2} \right],$$

and the expected profit of a firm that does not observe s_i as

$$E[\pi_i(a_i^0, A)] = \gamma_2 \left[\frac{1}{2\kappa_\theta} + (\beta - \tilde{b}_2/2)\, \tilde{b}_2 \left(\frac{1}{\kappa_z} \right) \frac{b_3^2}{(b_1 + b_3)^2} - \frac{\tilde{b}_1^2}{2\kappa_x} - \frac{\tilde{b}_2^2}{2\kappa_\nu} \right].$$

The equilibrium surplus of a firm is now $\Delta := E[\pi_i(a_i^1, A)] - E[\pi_i(a_i^0, A)]$, in Lemma 2.9.2 we proved that in the simultaneous acceptance/rejection game the monopolist offers to each firm a price $p_i = \Delta$, and all firms accept the offer. Thus, the expected equilibrium profit of the monopolist is $\Pi = \int_0^1 p_i \, di = \Delta$, i.e.

$$\Pi(t, \kappa_\nu, \beta) = \gamma_2 \left\{ \left[(1 - \tilde{b}_2)\beta - (1 - \tilde{b}_2^2)/2 \right] \left(\frac{1}{\kappa_z} \right) \frac{b_3^2}{(b_1 + b_3)^2} - \frac{b_1^2 - \tilde{b}_1^2}{2\kappa_x} - \frac{b_2^2 - \tilde{b}_2^2}{2\kappa_\nu} - \left(\frac{1-t}{t\kappa_z} \right) \frac{b_3^2}{2} \right\}. \tag{2.46}$$

Proof of part (a): Recall the change of variable $\kappa_s = t\kappa_z$, the provider needs to chose $t \in [0, 1]$. Clearly it is never optimal to set $t \leq 0$, with this observation the provider problem simplifies to: maximize $\Pi(t, \kappa_\nu, \beta)$ subject to $t \leq 1$. Define

$$\Pi^*(\kappa_\nu) := \max_t \ \Pi(t, \kappa_\nu, \beta) \quad s.t. \quad t \leq 1,$$

we are interested in characterizing how the maximum profit changes when we introduce some leakage, i.e.,we want to sign

$$\left. \frac{\partial \Pi^*}{\partial \kappa_\nu}(\kappa_\nu) \right|_{\kappa_\nu = 0}.$$

We will use the envelope theorem for constrained optimization problems. In particular, the Lagrangean associated to the provider's problem is

$$L(t, \mu, \kappa_\nu, \beta) = \Pi(t, \kappa_\nu, \beta) + \mu(1 - t),$$

by the envelope theorem we have that

$$\frac{\partial \Pi^*}{\partial \kappa_\nu}(\kappa_\nu) = \frac{\partial L}{\partial \kappa_\nu}(t^*, \mu^*, \kappa_\nu, \beta)\bigg|_{t^*=t^*(\kappa_\nu), \mu^*=\mu^*(\kappa_\nu)} = \frac{\partial \Pi}{\partial \kappa_\nu}(t^*, \kappa_\nu, \beta)\bigg|_{t^*=t^*(\kappa_\nu)}, \tag{2.47}$$

where the second equality holds because the constraint itself does not depend on κ_ν.

First, differentiating (2.46) with respect to κ_ν yields

$$\frac{\partial \Pi}{\partial \kappa_\nu}(t, \kappa_\nu, \beta) = \gamma_2 \Big(K_0 + K_1 + K_2 + K_3 + K_4 + K_5 + K_6 \Big), \tag{2.48}$$

where

$$\begin{aligned}
K_0 &= \frac{\tilde{b}_2 \tilde{b}_2^{\kappa_\nu} - \tilde{b}_2^{\kappa_\nu} \beta}{(b_1 + b_3)^2 \kappa_z} b_3^2 \\
K_1 &= \frac{2(1 - \tilde{b}_2)\beta - (1 - \tilde{b}_2^2)}{(b_1 + b_3)^2 \kappa_z} b_3 b_3^{\kappa_\nu} \\
K_2 &= \frac{2(1 - \tilde{b}_2)\beta - (1 - \tilde{b}_2^2)}{(b_1 + b_3)^3 \kappa_z} b_3^2 (b_1^{\kappa_\nu} + b_3^{\kappa_\nu}) \\
K_3 &= -\frac{b_1 b_1^{\kappa_\nu} - \tilde{b}_1 \tilde{b}_1^{\kappa_\nu}}{\kappa_x} \\
K_4 &= -\frac{b_2 b_2^{\kappa_\nu} - \tilde{b}_2 \tilde{b}_2^{\kappa_\nu}}{\kappa_\nu} \\
K_5 &= -\frac{1 - t}{t \kappa_z} b_3 b_3^{\kappa_\nu} \\
K_6 &= \frac{b_2^2 - \tilde{b}_2^2}{2\kappa_\nu^2},
\end{aligned}$$

and where

$$b_j^{\kappa_\nu} := \frac{\partial}{\partial \kappa_\nu} b_j(t, \kappa_\nu), \quad j = 1, 2, 3 \qquad \text{and} \qquad \tilde{b}_k^{\kappa_\nu} := \frac{\partial}{\partial \kappa_\nu} \tilde{b}_k(t, \kappa_\nu), \quad k = 1, 2. \tag{2.49}$$

Next, we need to evaluate (2.48) at $\kappa_\nu = 0$, to do so we first compute and evaluate all the coefficients at $\kappa_\nu = 0$ and then we replace them in (2.48).

Evaluating equations (2.41)-(2.45) at $\kappa_\nu = 0$ yields

$$b_1(t,0) = \frac{(1-\beta t)\kappa_x}{t\kappa_z + (1-\beta t)\kappa_x}, \qquad b_2(t,0) = 0, \qquad b_3(t,0) = \frac{t\kappa_z}{t\kappa_z + (1-\beta t)\kappa_x},$$
$$\tilde{b}_1(t,0) = 1, \qquad \tilde{b}_2(t,0) = 0.$$

Differentiating (2.41)-(2.45) with respect to κ_ν, and then evaluating at $\kappa_\nu = 0$ yields

$$\begin{aligned}
b_1^{\kappa_\nu}(t,0) &= -\frac{\kappa_x\left[(1-\beta t)^2\kappa_x + (1-t)t\kappa_z\right]^2}{\left[(1-\beta t)\kappa_x + t\kappa_z\right]^4},\\
b_2^{\kappa_\nu}(t,0) &= \frac{(1-\beta t)^2\kappa_x + (1-t)t\kappa_z}{\left[(1-\beta t)\kappa_x + t\kappa_z\right]^2},\\
b_3^{\kappa_\nu}(t,0) &= -\frac{t\kappa_z\left[\kappa_x + (1-2\beta)t\kappa_x + t\kappa_z\right]\left[(1-t)t\kappa_z + (1-\beta t)^2\kappa_x\right]}{\left[(1-\beta t)\kappa_x + t\kappa_z\right]^4},\\
\tilde{b}_1^{\kappa_\nu}(t,0) &= -\frac{t^2\kappa_z^2 + (1-\beta t)^2\kappa_x^2 + t(2-\beta t)\kappa_x\kappa_z}{\kappa_x\left[(1-\beta t)\kappa_x + t\kappa_z\right]^2},\\
\tilde{b}_2^{\kappa_\nu}(t,0) &= \frac{t^2\kappa_z^2 + (1-\beta t)^2\kappa_x^2 + t(2-\beta t)\kappa_x\kappa_z}{\kappa_x\left[(1-\beta t)\kappa_x + t\kappa_z\right]^2}.
\end{aligned}$$

Replacing the above sets of coefficients into Equation (2.48) and simplifying, yields the following expression

$$\left.\frac{\partial \Pi}{\partial \kappa_\nu}(t,\kappa_\nu,\beta)\right|_{\kappa_\nu=0} = \gamma_2 \frac{c_5 t^5\kappa_z^5 + c_4 t^4\kappa_z^4\kappa_x + c_3 t^3\kappa_z^3\kappa_x^2 + c_2 t^2\kappa_z^2\kappa_x^3 + c_1 t\kappa_z\kappa_x^4}{2\kappa_x^2\left[t\kappa_z + (1-\beta t)\kappa_x\right]^5} \tag{2.50}$$

where

$$c_5 = -1, \qquad c_4 = 3\beta t - 5, \qquad c_3 = -[(5\beta^2 - 1)t - 14\beta + 2]t - 9,$$
$$c_2 = -7 + t\{-4 + 19\beta + t[1 + (4+t)\beta - (15+4t)\beta^2 + 5t\beta^3]\},$$
$$c_1 = 2(\beta t - 1)^2((2\beta - 1)t - 1).$$

Now we can proceed to the last step to sign $\left.\frac{\partial \Pi^*}{\partial \kappa_\nu}(\kappa_\nu)\right|_{\kappa_\nu=0}$, there are two relevant cases.

Case 1: $\beta \in (-\infty, 0)$**.** Recall that $t \in [0,1]$, and consider Equation (2.50). It is easy to see that denominator is always positive, we next show that the numerator is always negative. Note that coefficients c_5, c_4 and c_1 are always negative. Moreover, we have that $(5\beta^2 - 1)t - 14\beta + 2 > -14\beta + 1 > 0$, which implies that c_3 is negative for all $\beta < 0$ and $t \in [0,1]$. Finally, noting that $1 + (4+t)\beta - (15+4t)\beta^2 + 5t\beta^3 < 1$ it is easy to verify that c_2 is negative. This establishes that $\left.\frac{\partial \Pi}{\partial \kappa_\nu}(t, \kappa_\nu, \beta)\right|_{\kappa_\nu = 0} < 0$ for all $\beta < 0$ and $t \in [0,1]$. In particular,

$$\left.\frac{\partial \Pi}{\partial \kappa_\nu}(t^*, \kappa_\nu, \beta)\right|_{t^* = t^*(\kappa_\nu), \kappa_\nu = 0} < 0 \qquad \forall \beta < 0.$$

Case 2: $\beta \in [0, 1/2)$**.** Recall that when $\kappa_\nu = 0$ we have $t^* = 1$ for all β in this interval. We next substitute $t^* = 1$ in Equation (2.50) and show that it is negative. Note that the denominator simplifies to $2\kappa_x^2 \left[\kappa_z + (1-\beta)\kappa_x\right]^5 < 0$, and the numerator simplifies to

$$-\kappa_z \left[\kappa_z^4 + (5 - 3\beta)\kappa_x\kappa_z^3 + (10 - 14\beta + 5\beta^2)\kappa_x^2\kappa_z^2 + (1-\beta)(10 - 14\beta + 5\beta^2)\kappa_x^3\kappa_z + 4(1-\beta)^3\kappa_x^4\right].$$

Noting that $10 - 14\beta + 5\beta^2$ is positive, since it is a convex quadratic with negative determinant, it is easy to see that the numerator is always negative. This establishes that

$$\left.\frac{\partial \Pi}{\partial \kappa_\nu}(t^*, \kappa_\nu, \beta)\right|_{t^* = t^*(\kappa_\nu), \kappa_\nu = 0} < 0 \qquad \forall \beta \in [0, 1/2).$$

Combining the results of Case 1 and Case 2, we can now apply the envelope theorem, and use Equation (2.47) to conclude that

$$\left.\frac{\partial \Pi^*}{\partial \kappa_\nu}(\kappa_\nu)\right|_{\kappa_\nu = 0} < 0 \qquad \forall \beta \in (-\infty, 1/2),$$

which completes the proof of part (a).

Proof of part (b): Let $\hat{\beta} = -(1 + \kappa_z/\kappa_x)$. When $\kappa_\nu = 0$, we know from Proposition 2.3.3 that $\hat{\beta}$ is exactly the level of β where the provider's constraint $\kappa_s \le \kappa_z$ becomes non-binding, i.e.,$\kappa_s^* = \kappa_z$

is the unconstrained optimum at $\beta = \hat{\beta}$. Adapting to the change of variable, we have

$$\arg\max_{t}\ \Pi(t, 0, \hat{\beta}) = 1,$$

and

$$\frac{\partial}{\partial t}\Pi(t, 0, \hat{\beta})\bigg|_{t=1} = 0. \tag{2.51}$$

We want to show that $\exists\ \bar{K}$ such that $\forall\ \kappa_\nu \leq \bar{K}$, $\exists\ \bar{T}$ such that $\forall\ t \in [1-\bar{T}, 1)$ the following holds

$$\Pi(t, \kappa_\nu, \hat{\beta}) > \Pi(1, \kappa_\nu, \hat{\beta}),$$

which means that when there is some leakage it is optimal for the seller to distort her signal at $\beta = \hat{\beta}$, and then we will use the continuity of Π to establish the final result.

In order to prove the above, let us define

$$g(\kappa_\nu, \beta) = \frac{\partial}{\partial t}\Pi(t, \kappa_\nu, \beta)\bigg|_{t=1},$$

and show that $\dfrac{\partial g}{\partial \kappa_\nu}(0, \hat{\beta}) < 0$. For a generic β, we have

$$\frac{\partial g}{\partial \kappa_\nu}(0, \beta) \quad = \quad \gamma_2\Big(G_0 + G_1 + G_2 + G_3 + G_4 + G_5 + G_6 + G_7 + G_8 + G_9\Big), \tag{2.52}$$

where

$$
\begin{aligned}
G_0 =& \left.\frac{\tilde{b}_2\tilde{b}_2^{\kappa_\nu} - \beta\tilde{b}_2^{\kappa_\nu}}{\kappa_z\left(b_1+b_3\right)^2} 2b_3^t b_3\right|_{t=1,\kappa_\nu=0}\\
G_1 =& \left.\frac{\tilde{b}_2\tilde{b}_2^{t} - \beta\tilde{b}_2^{t}}{\kappa_z\left(b_1+b_3\right)^2} 2b_3^{\kappa_\nu} b_3\right|_{t=1,\kappa_\nu=0}\\
G_2 =& \left.\frac{\tilde{b}_2^{\kappa_\nu} + \tilde{b}_2^{t} - \beta\tilde{b}_2^{t,\kappa_\nu} + \tilde{b}_2\tilde{b}_2^{t,\kappa_\nu}}{\kappa_z\left(b_1+b_3\right)^2} b_3^2\right|_{t=1,\kappa_\nu=0}\\
G_3 =& -\left.\frac{\left(b_1^{\kappa_\nu} + b_3^{\kappa_\nu}\right)\left(\tilde{b}_2\tilde{b}_2^{t} - \beta\tilde{b}_2^{t}\right)}{\kappa_z\left(b_1+b_3\right)^3} 2b_3^2\right|_{t=1,\kappa_\nu=0}\\
G_4 =& \left.\frac{2\beta(1-\tilde{b}_2) + (\tilde{b}_2^2 - 1)}{\kappa_z\left(b_1+b_3\right)^2}\left(b_3^{\kappa_\nu} b_3^t + b_3^{t,\kappa_\nu} b_3\right)\right|_{t=1,\kappa_\nu=0}\\
G_5 =& -\left.\frac{2\beta(1-\tilde{b}_2) + (\tilde{b}_2^2 - 1)}{\kappa_z\left(b_1+b_3\right)^3}\left[\left(b_1^{t,\kappa_\nu} + b_3^{t,\kappa_\nu}\right) b_3^2 + 2\left(b_1^{\kappa_\nu} + b_3^{\kappa_\nu}\right) b_3^t b_3\right]\right|_{t=1,\kappa_\nu=0}\\
G_6 =& -\left.\frac{b_1^{\kappa_\nu} b_1^t - \tilde{b}_1^{\kappa_\nu}\tilde{b}_1^t + b_1 b_1^{t,\kappa_\nu} - \tilde{b}_1\tilde{b}_1^{t,\kappa_\nu}}{\kappa_x}\right|_{t=1,\kappa_\nu=0}\\
G_7 =& -\left.\frac{b_2^{\kappa_\nu} b_2^t - \tilde{b}_2^{\kappa_\nu}\tilde{b}_2^t + b_2 b_2^{t,\kappa_\nu} - \tilde{b}_2\tilde{b}_2^{t,\kappa_\nu}}{\kappa_\nu}\right|_{t=1,\kappa_\nu=0}\\
G_8 =& \left.\frac{b_2 b_2^t - \tilde{b}_2\tilde{b}_2^t}{\kappa_\nu^2}\right|_{t=1,\kappa_\nu=0}\\
G_9 =& \left.\frac{b_3^{\kappa_\nu} b_3}{\kappa_z}\right|_{t=1,\kappa_\nu=0},
\end{aligned}
$$

and where

$$
\begin{aligned}
b_j^t &:= \frac{\partial}{\partial t} b_j(t,\kappa_\nu), \qquad b_j^{t,\kappa_\nu} := \frac{\partial^2}{\partial\kappa_\nu\partial t} b_j(t,\kappa_\nu), \qquad \text{for} \quad j = 1,2,3\ ,\\
\tilde{b}_k^t &:= \frac{\partial}{\partial t}\tilde{b}_k(t,\kappa_\nu), \qquad \tilde{b}_k^{t,\kappa_\nu} := \frac{\partial^2}{\partial\kappa_\nu\partial t}\tilde{b}_k(t,\kappa_\nu), \qquad \text{for} \quad k = 1,2\ ,\\
&,
\end{aligned}
$$

and the remaining coefficient were defined in (2.49). The first step in the computation of the second-derivative above is to evaluate all the coefficients at $t = 1$ and $\kappa_\nu = 0$, which is done as follows.

Evaluating (2.41)-(2.45) at $(t = 1, \kappa_\nu = 0)$, and solving the resulting system yields

$$b_1(1,0) = \frac{(1-\beta)\kappa_x}{\kappa_z + (1-\beta)\kappa_x}, \qquad b_2(1,0) = 0, \qquad b_3(1,0) = \frac{\kappa_z}{\kappa_z + (1-\beta)\kappa_x}, \qquad \tilde{b}_1(1,0) = 0, \qquad \tilde{b}_2(1,0) = 0.$$

Differentiating (2.41)-(2.45) with respect to κ_ν, evaluating at $(t = 1, \kappa_\nu = 0)$, and solving the resulting system yields

$$\begin{aligned}
b_1^{\kappa_\nu}(1,0) &= -\frac{(1-\beta)^4\kappa_x^3}{[(1-\beta)\kappa_x + \kappa_z]^4}, \\
b_2^{\kappa_\nu}(1,0) &= \frac{(1-\beta)^2\kappa_x}{[(1-\beta)\kappa_x + \kappa_z]^2}, \\
b_3^{\kappa_\nu}(1,0) &= -\frac{(1-\beta)^2\,[2(1-\beta)\kappa_x + \kappa_z]\,\kappa_x\kappa_z}{[(1-\beta)\kappa_x + \kappa_z]^4}, \\
\tilde{b}_1^{\kappa_\nu}(1,0) &= -\frac{(1-\beta)^2\kappa_x^2 + (2-\beta)\kappa_x\kappa_z + \kappa_z^2}{\kappa_x\,[(1-\beta)\kappa_x + \kappa_z]^2}, \\
\tilde{b}_2^{\kappa_\nu}(1,0) &= \frac{(1-\beta)^2\kappa_x^2 + (2-\beta)\kappa_x\kappa_z + \kappa_z^2}{\kappa_x\,[(1-\beta)\kappa_x + \kappa_z]^2}.
\end{aligned}$$

Differentiating (2.41)-(2.45) with respect to t, evaluating at $(t = 1, \kappa_\nu = 0)$, and solving the resulting system yields

$$\begin{aligned}
&b_1^t(1,0) = -\frac{\kappa_x\kappa_z}{[(1-\beta)\kappa_x + \kappa_z]^2}, \qquad b_2^t(1,0) = 0, \qquad b_3^t(1,0) = \frac{\kappa_x\kappa_z}{[(1-\beta)\kappa_x + \kappa_z]^2}, \\
&\tilde{b}_1^t(1,0) = 0, \qquad \tilde{b}_2^t(1,0) = 0.
\end{aligned}$$

Finally, differentiating (2.41)-(2.45) with respect to both t and κ_ν, evaluating at $(t = 1, \kappa_\nu = 0)$, and solving the resulting system yields

$$\begin{aligned}
b_1^{t,\kappa_\nu}(1,0) &= \frac{2(1-\beta)^2\left[3(1-\beta)\kappa_x+\kappa_z\right]}{\left[(1-\beta)\kappa_x+\kappa_z\right]^5}\kappa_x^2\kappa_z,\\
b_2^{t,\kappa_\nu}(1,0) &= -\frac{3(1-\beta)\kappa_x+\kappa_z}{\left[(1-\beta)\kappa_x+\kappa_z\right]^3}\kappa_z,\\
b_3^{t,\kappa_\nu}(1,0) &= \frac{-3(1-\beta)^3\kappa_x^3+5(1-\beta)^2\kappa_x^2\kappa_z+5(1-\beta)\kappa_x\kappa_z^2+\kappa_z^3}{\left[(1-\beta)\kappa_x+\kappa_z\right]^5}\kappa_z,\\
\tilde{b}_1^{t,\kappa_\nu}(1,0) &= -\frac{2\beta\kappa_x\kappa_z}{\left[(1-\beta)\kappa_x+\kappa_z\right]^3},\\
\tilde{b}_2^{t,\kappa_\nu}(1,0) &= \frac{2\beta\kappa_x\kappa_z}{\left[(1-\beta)\kappa_x+\kappa_z\right]^3}.
\end{aligned}$$

We now substitute the coefficients in the Equation (2.52), and simplify it, to get

$$\frac{\partial g}{\partial\kappa_\nu}(0,\beta) = -\gamma_2\frac{\kappa_z\left[(1-\beta)^2\left(\beta^2+\beta+2\right)\kappa_x^3+5(1-\beta)^2\kappa_x^2\kappa_z+(4\beta^2-9\beta+4)\kappa_x\kappa_z^2+(1-2\beta)\kappa_z^3\right]}{\left[(1-\beta)\kappa_x+\kappa_z\right]^6}.$$

It is easy to verify that the above equation is strictly negative for all $\beta\in(-\infty,1/2)$, and in particular it is strictly negative at $\hat{\beta}=-(1+\kappa_z/\kappa_x)$. Moreover, noting that $g(0,\hat{\beta})=0$ holds, since it is a restatement of (2.51), we can conclude that $\exists\bar{K}$ such that $\forall\kappa_\nu<\bar{K}$ we have $g(\kappa_\nu,\hat{\beta})<0$. This in turn implies that $\exists\,\bar{T}$ such that $\forall\, t\in[1-\bar{T},1)$ we have

$$\Pi(t,\kappa_\nu,\hat{\beta}) > \Pi(1,\kappa_\nu,\hat{\beta}).$$

Thus, at $\beta=\hat{\beta}$ it is optimal for the seller to set $t^*<1$, i.e.,$\kappa_s^*<\kappa_z$.

To complete the proof, note that by continuity of $\Pi(t,\kappa_\nu,\beta)$ with respect to β, there exists $\bar{\beta}\in(0,\hat{\beta})$ such that $\Pi(t,\kappa_\nu,\hat{\beta})-\Pi(1,\kappa_\nu,\hat{\beta})>0$ for all $\beta\in[\hat{\beta},\bar{\beta})$, and thus $\kappa_s^*<\kappa_z$ for all $\beta\in[\hat{\beta},\bar{\beta})$.

$\square$

Proof of Proposition 2.6.1

Following the same steps as in the proof of Proposition 2.4.1 we can derive the equilibrium quantity decisions of each firm i in the competition subgame. These, in turn, allow us to compute the equilibrium surplus of a firm i that observes a signal of precision κ_{si} as

$$\Delta_i = \mathbb{E}[\pi(q(\kappa_{si}, \kappa_{\mathbf{s-i}}))] - \mathbb{E}[\pi(q(0, \kappa_{\mathbf{s-i}}))] = K\left(\frac{\kappa_{si}/c_i}{\kappa_x + \kappa_{si}}\right)\Big/\kappa_x(1-\gamma_1 D\kappa_x/\kappa_z)^2,$$

where $K = \dfrac{\gamma_0^2/2}{\left(1-\gamma_1\int_0^1 \frac{1}{c_i}di\right)^2}$ and $D = \displaystyle\int_0^1 \frac{\kappa_{si}/c_i}{\kappa_x+\kappa_{si}}di.$

Moreover, firm $i \in [0,1]$ accepts the provider's offer $\{p_i, \kappa_{si}\}$ if and only if $\Delta_i \geq p_i$. Thus, it is optimal for the provider to offer $p_i = \Delta_i$ for all $i \in [0,1]$ and leave no surplus to the firms.

The provider's equilibrium profit is then given by

$$\Pi(\kappa_{\mathbf{s}}, D) = \int_0^1 (p_i - v\kappa_{si})di = K\frac{D}{\kappa_x(1-\gamma_1 D\kappa_x/\kappa_z)^2} - v\int_0^1 \kappa_{si}di,$$

and her problem can be written as

$$\begin{aligned}
\max_{\{\kappa_{si}\}_{i\in[0,1]}} \quad & \Pi(\kappa_{\mathbf{s}}, D) \\
s.t. \quad & D = \int_0^1 \frac{\kappa_{si}/c_i}{\kappa_x+\kappa_{si}}di \\
& 0 \leq \kappa_{si} \leq \kappa_z \quad \forall i \in [0,1].
\end{aligned}$$

Differentiating the objective with respect to κ_{si} yields

$$\frac{d\Pi}{d\kappa_{si}} = \frac{\partial\Pi}{\partial D}\cdot\frac{\partial D}{\partial\kappa_{si}} + \frac{\partial\Pi}{\partial\kappa_{si}} = K\cdot\frac{1+\gamma_1 D\kappa_x/\kappa_z}{c_i(\kappa_x+\kappa_{si})^2(1-\gamma_1 D\kappa_x/\kappa_z)^3} - v, \tag{2.53}$$

for all $i \in [0,1]$.

Proof of Part (a): Let $\overline{v} = K/(c_{\min}\kappa_x^2)$. Suppose that $v > \overline{v}$ and consider the profile of precisions

$\{\kappa_{si}=0\}_{i\in[0,1]}$, with corresponding $D=0$. Evaluating (2.53) at the profile above yields

$$\frac{d\Pi}{d\kappa_{si}}=\frac{K}{c_i\kappa_x^2}-v,$$

for all $i\in[0,1]$. Our assumption on v then implies that, at $\{\kappa_{si}=0\}_{i\in[0,1]}$, we have $\frac{d\Pi}{d\kappa_{si}}<0$ for all $i\in[0,1]$, thus the provider can only increase her profit by decreasing κ_{si} for some i. Then the non-negativity constraint on κ_{si} imply that the optimal solution is to set $\kappa_{si}^*=0$ for all $i\in[0,1]$.

Proof of Part (b): Let $\underline{v}=K\frac{\kappa_z+\left(1+\gamma_1\int_0^1\frac{1}{c_i}di\right)\kappa_x}{c_{\max}\left[\kappa_z+\left(1-\gamma_1\int_0^1\frac{1}{c_i}di\right)\kappa_x\right]^3}$. Suppose that $v<\underline{v}$ and consider the profile of precisions $\{\kappa_{si}=\kappa_z\}_{i\in[0,1]}$, with corresponding $D=\frac{\kappa_z}{\kappa_x+\kappa_z}\int_0^1\frac{1}{c_i}di$. Evaluating (2.53) at the profile above yields

$$\frac{d\Pi}{d\kappa_{si}}=K\frac{\kappa_z+\left(1+\gamma_1\int_0^1\frac{1}{c_i}di\right)\kappa_x}{c_i\left[\kappa_z+\left(1-\gamma_1\int_0^1\frac{1}{c_i}di\right)\kappa_x\right]^3}-v,$$

for all $i\in[0,1]$. Our assumption on v then implies that, at $\{\kappa_{si}=\kappa_z\}_{i\in[0,1]}$, we have $\frac{d\Pi}{d\kappa_{si}}>0$ for all $i\in[0,1]$, thus the provider can only increase her profit by increasing κ_{si} for some i. Then, the constraint $\kappa_{si}\leq\kappa_z\ \forall i\in[0,1]$, imply that the optimal solution is to set $\kappa_{si}^*=\kappa_z$ for all $i\in[0,1]$.

Proof of Part (c): Consider now the remaining case in which v takes an intermediate value. Let κ_{si}^* be the optimal solution, with corresponding $D^*=\int_0^1\frac{\kappa_{si}^*/c_i}{\kappa_x+\kappa_{si}^*}di$. Substituting the optimal solution into equation (2.53) yields

$$\frac{d\Pi}{d\kappa_{si}}=K\frac{1+\gamma_1D^*\kappa_x/\kappa_z}{c_i(\kappa_x+\kappa_{si}^*)^2(1-\gamma_1D^*\kappa_x/\kappa_z)^3}-v, \tag{2.54}$$

for all $i\in[0,1]$.

Let $c^* = \dfrac{K(1+\gamma_1 D^* \kappa_x/\kappa_z)}{v\ \kappa_x^2(1-\gamma_1 D^* \kappa_x/\kappa_z)^3}$. If $c_i > c^*$ we have

$$\frac{d\Pi}{d\kappa_{si}} < \left(\frac{\kappa_x^2}{(\kappa_x+\kappa_{si}^*)^2} - 1\right) v,$$

which implies that $d\Pi/d\kappa_{si} < 0$ for all $\kappa_{si}^* \geq 0$, thus it is optimal to set $\kappa_{si}^* = 0$.

If $c_i < \dfrac{\kappa_x^2}{(\kappa_x+\kappa_z)^2} c^*$ we have

$$\frac{d\Pi}{d\kappa_{si}} > \left[\left(\frac{\kappa_x+\kappa_z}{\kappa_x+\kappa_{si}^*}\right)^2 - 1\right] v,$$

which implies that $d\Pi/d\kappa_{si} > 0$ for all $\kappa_{si}^* \in [0, \kappa_z]$, thus it is optimal to set $\kappa_{si}^* = \kappa_z$ in this case.

Finally, for intermediate values of c_i the optimal κ_{si}^* must be a solution to $d\Pi/d\kappa_{si} = 0$, thus

$$\kappa_{si}^* = \sqrt{\frac{K(1+\gamma_1 D^* \kappa_x/\kappa_z)}{c_i v(1-\gamma_1 D^* \kappa_x/\kappa_z)^3}} - \kappa_x = \kappa_x \left(\sqrt{c^*/c_i} - 1\right).$$

□

Proof of Proposition 2.6.2

Consider the optimal thresholds $\overline{v}$ and $\underline{v}$ characterized in the Proof of Proposition 2.6.1, adapting to the two-type environment we have

$$\overline{v} = \left(\frac{1}{c} + \delta\right)\frac{K}{\kappa_x^2} \qquad \text{and} \qquad \underline{v} = \left(\frac{1}{c} - \delta\right) K \frac{\kappa_z + (1+\gamma_1/c)\,\kappa_x}{[\kappa_z + (1-\gamma_1/c)\,\kappa_x]^3},$$

where $K = \dfrac{\gamma_0^2/2}{(1-\gamma_1/c)^2}$. It is easy to see that when $v > \overline{v}$ or $v < \underline{v}$ the optimal precisions are not affected by a marginal increase in dispersion. For intermediate values of v there are four relevant cases. Before proving the result for each of the cases, recall from Proposition 2.6.1 that the optimal selling strategy has the following structure

$$\kappa_{sk}^* = \begin{cases} 0 & \text{if } \frac{1}{c_k} < \frac{1}{c^*} \\ \kappa_z & \text{if } \frac{(\kappa_x+\kappa_z)^2}{\kappa_x^2 c^*} < \frac{1}{c_k} \ , \\ \kappa_x \left(\sqrt{c^*/c_k} - 1\right) & \text{otherwise} \end{cases} \qquad k = 1, 2 \ , \tag{2.55}$$

where

$$c^* = \frac{K(1 + \gamma_1 D^* \kappa_x/\kappa_z)}{v \ \kappa_x^2 (1 - \gamma_1 D^* \kappa_x/\kappa_z)^3},$$

which is always positive when $v \in (\underline{v}, \overline{v})$, and

$$D^* = \frac{1}{2}\left(\frac{1}{c} + \delta\right)\frac{\kappa_{s1}^*}{\kappa_x + \kappa_{s1}^*} + \frac{1}{2}\left(\frac{1}{c} - \delta\right)\frac{\kappa_{s2}^*}{\kappa_x + \kappa_{s2}^*}. \tag{2.56}$$

Moreover, differentiating c^* with respect to δ yields $\dfrac{\partial c^*}{\partial \delta} = \dfrac{\partial c^*}{\partial D^*}\left(\dfrac{\partial D^*}{\partial \delta} + \dfrac{\partial D^*}{\partial c^*}\dfrac{\partial c^*}{\partial \delta}\right)$, which implies that

$$\frac{\partial c^*}{\partial \delta} = \frac{\partial c^*}{\partial D^*}\frac{\partial D^*}{\partial \delta} \Big/ \left(1 - \frac{\partial c^*}{\partial D^*}\frac{\partial D^*}{\partial c^*}\right). \tag{2.57}$$

Note that $\dfrac{\partial c^*}{\partial D^*} = \dfrac{2\gamma_1 K(2 + \gamma_1 D^* \kappa_x/\kappa_z)}{v \ \kappa_x \kappa_z (1 - \gamma_1 D^* \kappa_x/\kappa_z)^4}$ is always negative since $\gamma_1 < 0$ and $2 + \gamma_1 D^* \kappa_x/\kappa_z > 0$. We now proceed with the proof of the four relevant cases.

Case (i): $(1/c + \delta), \ (1/c - \delta) \in \left[\dfrac{1}{c^*}, \dfrac{(\kappa_x + \kappa_z)^2}{\kappa_x^2 c^*}\right]$. The optimal precision takes intermediate values for both firm types, replacing the optimal precisions from (2.55) into (2.56) yields

$$D^* = \frac{1}{c} - \frac{1}{\sqrt{c^*}}\left(\frac{1}{2}\sqrt{\frac{1}{c} + \delta} + \frac{1}{2}\sqrt{\frac{1}{c} - \delta}\right).$$

Note that $\dfrac{\partial D^*}{\partial \delta} > 0$ by concavity of the square root, therefore we can verify from (2.57) that $\dfrac{\partial c^*}{\partial \delta} < 0$.

Differentiating the optimal precision of type 2 firms with respect to δ yields

$$\begin{aligned}\frac{\partial \kappa_{s2}^*}{\partial \delta} &= \frac{\partial}{\partial \delta}\left[\kappa_x \sqrt{c^*\left(\frac{1}{c}-\delta\right)} - \kappa_x\right] \\ &= \frac{\kappa_x}{2\sqrt{c^*\left(\frac{1}{c}-\delta\right)}}\left[-c^* + \frac{\partial c^*}{\partial \delta}\left(\frac{1}{c}-\delta\right)\right],\end{aligned}$$

which is always negative since $\dfrac{\partial c^*}{\partial \delta} < 0$. For type 1 firms we have

$$\begin{aligned}\frac{\partial \kappa_{s1}^*}{\partial \delta} &= \frac{\kappa_x}{2\sqrt{c^*\left(\frac{1}{c}+\delta\right)}}\left[c^* + \frac{\partial c^*}{\partial \delta}\left(\frac{1}{c}+\delta\right)\right] \\ &= \frac{\kappa_x c^*}{2\sqrt{c^*\left(\frac{1}{c}+\delta\right)}}\left(1 + \frac{\frac{\partial c^*}{\partial D^*}\sqrt{\frac{1}{c}+\delta}\left(\sqrt{\frac{1}{c}+\delta} - \sqrt{\frac{1}{c}-\delta}\right)}{\sqrt{\frac{1}{c}-\delta}\left[4c^{*3/2} - \frac{\partial c^*}{\partial D^*}\left(\sqrt{\frac{1}{c}+\delta} + \sqrt{\frac{1}{c}-\delta}\right)\right]}\right),\end{aligned}$$

which is positive if and only if

$$\left(\delta - \sqrt{\frac{1}{c}+\delta}\,\sqrt{\frac{1}{c}-\delta}\right)\frac{\partial c^*}{\partial D^*} + 2\sqrt{\frac{1}{c}+\delta}\; c^{*3/2} > 0,$$

which is always positive if $\delta < 1/(c\sqrt{2})$.

Case (ii): $(1/c+\delta) \in \left[\dfrac{1}{c^*}, \dfrac{(\kappa_x+\kappa_z)^2}{\kappa_x^2 c^*}\right]$ and $(1/c-\delta) < \dfrac{1}{c^*}$. In this case, κ_{s1}^* takes intermediate values and $\kappa_{s2}^* = 0$, replacing the optimal precisions into (2.56) yields

$$D^* = \frac{1}{c} - \frac{\sqrt{\frac{1}{c}+\delta}}{\sqrt{c^*}}.$$

Note that $\dfrac{\partial D^*}{\partial \delta} < 0$, which implies that $\dfrac{\partial c^*}{\partial \delta} > 0$ and therefore

$$\frac{\partial \kappa_{s1}^*}{\partial \delta} = \frac{\kappa_x}{2\sqrt{c^*\left(\frac{1}{c}+\delta\right)}}\left[c^* + \frac{\partial c^*}{\partial \delta}\left(\frac{1}{c}+\delta\right)\right] > 0,$$

For type 2 firms we have $\dfrac{\partial \kappa_{s2}^*}{\partial \delta} = 0$, which follows from $\kappa_{s2}^* = 0$ and continuity of the optimal threshold with respect to δ.

Case (iii): $(1/c-\delta)\in\left[\frac{1}{c^*},\frac{(\kappa_x+\kappa_z)^2}{\kappa_x^2c^*}\right]$ and $(1/c+\delta)>\frac{(\kappa_x+\kappa_z)^2}{\kappa_x^2c^*}$. In this case, κ_{s2}^* takes intermediate values and $\kappa_{s1}^*=\kappa_z$, replacing the optimal precisions into (2.56) yields

$$D^*=\frac{1}{2}\left(\frac{1}{c}-\frac{\sqrt{\frac{1}{c}-\delta}}{\sqrt{c^*}}\right)+\frac{1}{2}\left(\frac{\kappa_z}{\kappa_x+\kappa_z}\right)\left(\frac{1}{c}+\delta\right).$$

Note that $\frac{\partial D^*}{\partial\delta}<0$, which implies that $\frac{\partial c^*}{\partial\delta}>0$ and therefore

$$\frac{\partial\kappa_{s2}^*}{\partial\delta}=\frac{\kappa_x}{2\sqrt{c^*\left(\frac{1}{c}-\delta\right)}}\left[-c^*+\frac{\partial c^*}{\partial\delta}\left(\frac{1}{c}-\delta\right)\right]<0,$$

For type 1 firms we have $\frac{\partial\kappa_{s1}^*}{\partial\delta}=0$.

Case (iv): $(1/c-\delta)<\frac{1}{c^*}$ and $(1/c+\delta)>\frac{(\kappa_x+\kappa_z)^2}{\kappa_x^2c^*}$. In this case, $\kappa_{s2}^*=0$ and $\kappa_{s1}^*=\kappa_z$, and $\frac{\partial\kappa_{s1}^*}{\partial\delta}=\frac{\partial\kappa_{s2}^*}{\partial\delta}=0$.

Combining the above cases, we have that $\delta<1/(c\sqrt{2})$ implies $\frac{\partial\kappa_{s1}^*}{\partial\delta}\geq 0$ and $\frac{\partial\kappa_{s2}^*}{\partial\delta}\leq 0$, thus proving the result.

□

Proof of Proposition 2.7.1

We begin by stating a lemma which is the discrete analogue of Proposition 2.3.1 presented in Section 2.3. The proof of the lemma follows similar arguments and is therefore omitted.

Lemma 2.9.5. *The competition subgame between the firms has a unique Bayes-Nash equilibrium in linear strategies. Furthermore, the equilibrium quantities of the firms are given by*

$$q_i=\begin{cases}\alpha_n[(1-\omega_{k,n})x_i+\omega_{k,n}s_i] & \textit{if } i\in K\\ \alpha_n x_i & \textit{if } i\in N\setminus K\end{cases},$$

where

$$\omega_{k,n} = \frac{\kappa_s}{\left(1 - \beta_n \frac{k-1}{n-1}\rho\right)\kappa_x + \kappa_s},$$

and $\alpha_n = \gamma_0 / \left(\gamma_2 - \frac{n+1}{n}\gamma_1\right)$.

Using the above lemma we can characterize the expected equilibrium profits of an uninformed and of an informed firm, respectively as

$$\mathbb{E}\left[\pi^1|\theta\right] = \alpha_n^2 \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right)\left[\theta^2 + \frac{2\beta_n\rho\kappa_s}{[(1-\beta_n\rho)\kappa_x + \kappa_s]^2} - \frac{(1-\beta_n\rho)^2\kappa_x + \kappa_s}{[(1-\beta_n\rho)\kappa_x + \kappa_s]^2}\right],$$

and

$$\mathbb{E}\left[\pi^0|\theta\right] = \alpha_n^2 \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right)\left[\theta^2 - \frac{1}{\kappa_x}\right],$$

where we used the assumption that $k = n$. The provider extracts all surplus generated by the information she sells. Thus, we can use the law of total expectation to characterize the provider's expected profit as

$$\Pi(\kappa_s, \rho, \kappa_x) = n \;\left(\mathbb{E}\left[\pi^1\right] - \mathbb{E}\left[\pi^0\right]\right) = n\;\alpha_n^2\left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right)\left(\frac{\kappa_s}{\kappa_x}\right)\frac{\kappa_s + \kappa_x}{[(1-\beta_n\rho)\kappa_x + \kappa_s]^2}.$$

Thus, the provider's optimization problem is

$$\begin{aligned} \max_{\rho,\kappa_s} \quad & \Pi(\kappa_s, \rho, \kappa_x) \\ s.t. \quad & \frac{\kappa_s}{\kappa_z} \le \rho \le 1 \\ & \kappa_s \le \kappa_z. \end{aligned}$$

Differentiating the objective with respect to ρ yields

$$\frac{\partial \Pi}{\partial \rho} = 2n\ \alpha_n^2 \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right) \frac{\beta_n \kappa_s(\kappa_x + \kappa_s)}{[(1-\beta_n \rho)\kappa_x + \kappa_s]^3}, \tag{2.58}$$

which is always negative since $\beta_n < 0$. Thus, it is optimal for the provider to set $\rho^* = \kappa_s/\kappa_z$. Replacing ρ^* into the objective and then differentiating with respect to κ_s yields

$$\frac{\partial \Pi}{\partial \kappa_s} = n\ \alpha_n^2 \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right) \frac{(1 + \beta_n \kappa_s/\kappa_z)\kappa_x + \kappa_s}{[(1-\beta_n \lambda \kappa_s/\kappa_z)\kappa_x + \kappa_s]^3}.$$

When $\beta_n \geq -(1 + \kappa_z/\kappa_x)$, the above expression is always positive and it is optimal to set $\kappa_s^* = \kappa_z$. Otherwise, the optimal precision is given by the solution to

$$(1 + \beta_n \kappa_s/\kappa_z)\kappa_x + \kappa_s = 0,$$

which implies that $\kappa_s^* = -\kappa_z/(\beta_n + \kappa_z/\kappa_x)$. Substituting the optimal strategy into the objective, we can easily characterize the optimal profits as

$$\Pi^* = \begin{cases} n\ \alpha_n^2 \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right) \left(\frac{\kappa_z}{\kappa_x}\right) \frac{\kappa_z + \kappa_x}{[(1-\beta_n)\kappa_x + \kappa_z]^2} & \text{if } \beta_n \geq -(1 + \kappa_z/\kappa_x) \\ \\ n\ \alpha_n^2 \left(\frac{\gamma_2}{2} - \frac{\gamma_1}{n}\right) \frac{\kappa_z}{-4\beta_n \kappa_x^2} & \text{otherwise} \end{cases},$$

thus completing the proof. □

2.10 Measure Theoretic Framework

In Appendix 2.10, we provide formal conditions for an “exact law of large numbers” to hold in our context. Our discussion largely builds on the results in Sun [2006] and Sun and Zhang [2009]. We show that, within the appropriate measure theoretic framework, our environment with a continuum of firms retains the most important measure theoretic properties of an environment with a finite number of firms. The main results are a Fubini property, which allows us to exchange the order of integration, and a strong law of large numbers for pairwise independent random variables, which allows us to work with aggregate uncertainty. In what follows we will first introduce formally our measure theoretic framework, which entails a definition of an extension of the Lebesgue unit interval as our set of firms, and then state formally the properties mentioned above.

Consider a measure space $([0,1],\mathcal{L},m)$ of *firms*, where $\mathcal{L}$ is the σ-algebra of Lebesgue measurable sets and m is the Lebesgue measure, and a probability space capturing the *uncertainty* in the model $(\Omega,\mathcal{F},P)$, where Ω is the sample space, $\mathcal{F}$ is the σ-algebra of events, and P is a probability measure. As per the discussion in Sun [2006], our objective is to define an extension $([0,1],\mathcal{I},M)$ of the measure space $([0,1],\mathcal{L},m)$ such that $\mathcal{L}\subseteq\mathcal{I}$ and the restriction of M on $\mathcal{L}$ coincides with m.

To this end, consider the Cartesian product $[0,1]\times\Omega$, and endow it with the product σ-algebra $\mathcal{I}\otimes\mathcal{F}:=\sigma(\mathcal{R})$, which is the σ-algebra generated by the class of measurable rectangles

$$\mathcal{R}:=\{I\times F \mid I\in\mathcal{I} \text{ and } F\in\mathcal{F}\}.$$

In addition, define the product measure $M\otimes P$ with the property that

$$M\otimes P(I\times F)=M(I)\cdot P(F),\ \forall\ I\in\mathcal{I},\ F\in\mathcal{F}.$$

Since M and P are probability measures, and thus σ-finite, the desired property is straightforward from the product measure theorem (see Billingsley [2008]). After having defined the product space $([0,1]\times\Omega,\mathcal{I}\otimes\mathcal{F},M\otimes P)$, the next step is to define a *Fubini extension*. Before proceeding with the formal definition, we introduce some additional notation. In particular, for a process $f: [0,1]\times\Omega\to\mathbb{R}$ we let f_i denote the random variable $f(i,\cdot):\Omega\to\mathbb{R}$ and f_ω denote the random variable $f(\cdot,\omega):[0,1]\to\mathbb{R}$.

Next, we define an extension $([0,1]\times\Omega,\mathcal{I}\boxtimes\mathcal{F},M\boxtimes P)$ of the product space described above, such that $\mathcal{I}\otimes\mathcal{F}\subseteq\mathcal{I}\boxtimes\mathcal{F}$ and the restriction of $M\boxtimes P$ on $\mathcal{I}\otimes\mathcal{F}$ coincides with $M\otimes P$. The extension $([0,1]\times\Omega,\mathcal{I}\boxtimes\mathcal{F},M\boxtimes P)$ satisfies the following Fubini-type property: for any $M\boxtimes P$-integrable process f on $([0,1]\times\Omega,\mathcal{I}\boxtimes\mathcal{F})$, we have that f_i is $P-integrable$ and f_ω is $M-integrable$. In addition,

$$\int_{\Omega} f_i \; dP \quad is\ M-integrable, \qquad \text{and} \qquad \int_{[0,1]} f_\omega \; dM \quad is\ P-integrable.$$

Finally,

$$\int_{[0,1]\times\Omega} f(i,\omega)\; d(M\boxtimes P)(i,\omega) = \int_{[0,1]} \left(\int_{\Omega} f_i(\omega)\; dP(\omega)\right)\; dM(i) = \int_{\Omega} \left(\int_{[0,1]} f_\omega(i)\; dM(i)\right)\; dP(\omega).$$

Throughout the paper, and with special reference to the derivations presented in Appendix 2.9, we assume that all relevant quantities lie in the Fubini extension of the product probability space constructed above. For additional details we refer the interested reader to Sun [2006] that provides a detailed treatment of the construction and establishes that there are Fubini extensions in which one can construct processes with pairwise independent random variables taking any distribution and Sun and Zhang [2009] that show that this result holds for the case in which the index space is an extension of the Lebesgue unit interval. Moreover, we require any process f on $([0,1]\times\Omega, \mathcal{I}\boxtimes\mathcal{F})$ to be measurable, which is consistent with the information structure specified throughout the paper. This assumption, coupled with the fact that $M\boxtimes P$ is a probability measure, implies that f is $M\boxtimes P$-integrable. Thus, the Fubini-type property above holds for any random variable in the paper. Showing existence of conditional expectations is immediate.

The other important relation that we use in the paper is a law of large numbers for stochastic processes depending on a continuous parameter. Sun [2006] proves such a law, which he calls an *exact law of large numbers*, in the framework of a Fubini extension. In the following lemma, we adapt and present a key auxiliary result stemming from his work.

Lemma 2.10.1. *Consider a process $f : [0,1]\times\Omega \to \mathbb{R}$ and assume it is square integrable with respect to $M\boxtimes P$. If the random variables $\{f_i\}_{i\in[0,1]}$ are uncorrelated, then for any set $I\in\mathcal{I}$ such that $M(I)>0$, we have*

$$\int_{I} f(i,\omega)\; dM(i) = \int_{I\times\Omega} f(i,\omega)\; d(M\boxtimes P)(i,\omega) \quad for\ P-almost\ all\ \omega\in\Omega.$$

This lemma essentially states that the sample average of a random variable over any set of firms I with positive measure is exactly equal to its expectation. We use this result in the derivations of Appendix 2.9 and, specifically to claim that $\int_{[0,1]} \epsilon(i,\omega)\ dM(i) = \int_{[0,1]} \xi(i,\omega)\ dM(i) = 0$. We should clarify that we slightly abuse notation throughout the paper to simplify the exposition of our analysis and results. In particular, we write $\int_0^1 f_i\ di$ instead of $\int_{[0,1]} f_i\ dM(i)$.

Chapter 3

Monopoly Pricing in the Presence of Social Learning

3.1 Introduction

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

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

(WtP) distribution and, in turn, into a demand function that the monopolist could use as a basis of her pricing decision.

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

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

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

Regarding the learning mechanism, the information reported by consumers is subject to a self-selection bias, since only consumers with a high enough base valuations purchase the product. The intuitive MLE procedure takes into account this crucial point, and, as we show in Section 3.3.2, the resulting quality estimate converges to the true product quality almost surely.

Detailed understanding of the learning trajectory is essential in optimizing the tradeoff between learning and the monopolist's discounted revenue objective. Second, Section 3.3.3 derives a mean-field (fluid model) asymptotic approximation for the learning dynamics motivated by settings where the rate of arrival of new consumers to the system grows large. Proposition 3.3.3 shows that the asymptotic learning trajectory is characterized by a system of differential equations. Restricting attention to uniformly distributed base valuations across consumers and focusing on the case where the markets prior quality estimate is below the true quality, Section 3.3.4 derives the closed form transient of the fraction of likes and dislikes over time, as well as that of the associated quality estimate. The transient dynamics imply that the instantaneous demand function evolves over time according to an ODE, which itself depends on the seller's price, i.e., it emerges endogenously through the interplay between consumer behavior and the seller's decisions. The solution of the

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

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

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

Third, we study the seller's pricing problem under the assumption that the seller knows the true product quality, but that the consumers do not use the seller's price as a signal of quality. Section 3.4 studies the monopolist's problem of choosing the static price that optimizes her infinite horizon discounted revenues. Proposition 3.4.1 characterizes the optimal solution, which exists and is unique, and lies in the interval of two natural price points: (a) the optimal price assuming that consumers do not learn and always make purchase decisions based on their prior quality estimate; and (b) the optimal price in a setting where consumers knew the true quality all along. The learning transient and its speed in relation to the seller's discount factor determines the optimal price. Intuitively, if the learning transient is slow relative to the discounting of revenues, then she prices almost as if all consumers made purchasing decisions based on their prior on the quality; and, if learning is fast, then the seller's price will approach the one that the monopolist would set if all consumers knew the true product quality.

Lastly, Section 3.5.1 studies a model where the seller has some degree of dynamic pricing capability, namely she can change her price once, at a time of her choosing. In this case the monopolist may sacrifice short term revenues in order to influence the social learning process in the desired direction and capitalize on that after changing the price. Proposition 3.5.1 shows that when consumers initially underestimate the true quality, the first period price is lower than the second period one. This policy accelerates learning and increases revenues considerably. The numerical experiments of Section 3.5.2 suggest that a pricing policy with two prices performs very well, and that the benefit of implementing more elaborate pricing policies may be small.

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

For social learning to be successful, an agent must be able to reverse the herd behavior of his predecessors. Smith and Sørensen [2000] show that this is the case if agents' signals have unbounded strength. Goeree *et al.* [2006] show that this is achieved with enough heterogeneity in consumers' preferences. Our Assumption 3.2.1, which is key in proving learning, is similar in nature to that of Goeree *et al.* [2006].[3] Social learning has been studied in great generality by Arieli and Mueller-Frank [2014].

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

[3] See the surveys by Bikhchandani *et al.* [1998], and, more recently, by Acemoglu and Ozdaglar [2011] for many extensions to this model.

the asymptotic learning result of Smith and Sørensen [2000].

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

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

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

Mean-field approximations have been used extensively in the area of revenue management[4]; perhaps the first reference in that area is Gallego and van Ryzin [1994]. More broadly, the use of mean-field approximations that rely on an appropriate application of the functional strong law of large numbers to study the transient behavior of stochastic processes has a fairly broad literature

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

that we will not review here. The particular result we will employ, due to Kurtz [1978], was originally derived for studying the asymptotic behavior of Markov Chain models with process-dependent transition parameters, used to analyze diffusion and epidemic systems.

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

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

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

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

increasing when consumers initially underestimate the quality; our Proposition 3.5.1 is consistent with their findings.

3.2 Model

3.2.1 The Monopolist's Pricing Problem

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

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

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

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

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

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

failure rate (IFR), that is, h is strictly increasing for all $x \geq 0$.

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

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

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

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

The information transmission in our model is often called *word-of-mouth* communication. A consumer i who purchased the product, truthfully reports a review about his experience with the product, denoted by r_i that takes two values: 'like', denoted by r^{L} and 'dislike', denoted by r^{D}. A consumer who purchases the product reports that he likes it if his ex-post utility was nonnegative, taking into account the unknown quality and quality disturbance, as well as his preference parameter; he reports that he dislikes it if his ex-post utility was negative. Consumers report neither their preference parameter nor the quality disturbance they faced and, as such, reviews are not fully informative. For example, a 'like' could result from a high preference parameter, the product being of high quality or a positive quality shock (not necessarily all).[7] This binary report is a simplification of the star rating scales of online review systems. Consumers who did not purchase the product do not report a review and are not observed. We will denote their decision by $r_i = r^{\mathrm{O}}$.

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

Assumption 3.2.1. *The price p charged by the monopolist belongs to $[p_{\min}, p_{\max}]$, where*

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

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

Assumption 3.2.1–*(i)* implies that, even at the lowest possible quality level, there will always be some consumers who choose to buy the product (this follows from $p_{\max} < \alpha_{\max} + q_{\min}$), and moreover, at least some of these consumers will like the product—the latter ensures that new information about q will enter the learning process; if this assumption is violated and $p_{\max} > \alpha_{\max} + q_{\min} - \bar{\varepsilon}$, then at $q = q_{\min}$ all buyers with a negative shock would dislike and all buyers with a positive shock would like. Assumption 3.2.1–*(i)* is similar to the "unbounded belief assumption" often used in Bayesian social learning in the sense that it implies that some new information will enter the system over time, which will ultimately allow the market to learn the unknown product quality. Assumption 3.2.1–*(ii)*, states that there are always some low-WtP consumers who will dislike the product if they get a negative disturbance realization. It is easy to verify that both conditions are satisfied if the support of α's is sufficiently wide relative to the unknown quality $[q_{\min}, q_{\max}]$ and the magnitude of the subsequent quality disturbances $\bar{\varepsilon}$ is small.

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

$$I_i = (L_i, D_i). \tag{3.1}$$

[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.

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

$$\max_p R(p) = \max_p \mathsf{E}\left[\sum_{i=1}^{\infty} e^{-\delta t_i} \; p\mathbf{1}\left\{r_i(p) \neq r^{\mathrm{O}}\right\}\right] = \max_p \sum_{i=1}^{\infty} \mathsf{E}\left[e^{-\delta t_i} \; p\mathsf{P}\left(r_i(p) \neq r^{\mathrm{O}} \mid I_i\right)\right], \quad (3.2)$$

where $\delta > 0$ is the monopolist's discount factor, and the expectation is with respect to consumers' arrival times, the idiosyncratic quality preferences α_i's, and the sequence of quality disturbances ε_i's. The monopolist is assumed to know the true quality, the prior quality estimate, as well as the distribution of quality preferences and disturbances. Expression (3.2) reveals the complexity of the pricing problem in the presence of social learning. Consumers' purchasing decisions influence future revenues through the information available to successors. As such, the dynamic of the social learning process must be understood in order to solve for the optimal price. Section 3.5 considers a problem where the seller can select two prices as well as the optimal time to switch between them.

3.2.2 Decision Rule

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

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

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

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

where the second equality follows from Bayes' rule and

$$G(x) := \bar{F}(x + \bar{\varepsilon})/\bar{F}(x). \tag{3.3}$$

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

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

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

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

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

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

$$L_0 = w\Lambda \mathsf{P}(\text{customer } i \text{ buys \& likes} \mid \hat{q}_i = q = q_0) = .5w\Lambda\left[\bar{F}(p-q_0) + \bar{F}(p-q_0+\bar{\varepsilon})\right], \quad (3.4)$$

and

$$D_0 = w\Lambda \mathsf{P}(\text{customer } i \text{ buys \& dislikes} \mid \hat{q}_i = q = q_0) = .5w\Lambda\left[\bar{F}(p-q_0) - \bar{F}(p-q_0+\bar{\varepsilon})\right]. \quad (3.5)$$

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

$$\mathcal{L}_i^w(q) = (.5 + .5G(p-q))^{L_0+L_i}(.5 - .5G(p-q))^{D_0+D_i}. \quad (3.6)$$

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

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

Similarly, the probability of a dislike is

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

and the probability of no purchase is given by

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

We finish this section with few brief comments.

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

Prior weight. All consumers assign the same weight to their common prior quality estimate. The weight assigned to the prior w is assumed to be constant over time, that is, consumers arriving later in time still assign the same weight to the prior, but due to the accumulation of review information, these consumers end up slowly "forgetting" their prior estimate as it becomes less significant in the MLE procedure. The weight w is a measure of inertia of the learning process, or measure of confidence in the prior estimate in the absence of other new information. As time goes by and more reviews accumulate in the system, consumers place increasingly more confidence in the review information versus their prior information, which is reflected into the fact that the effect of L_0 and

D_0 on the weighted likelihood (3.6) becomes negligible as L_i and D_i grow.

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

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

3.3 Asymptotic learning and the associated learning transient

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

3.3.1 The consumer's MLE problem

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

$$l_i := \frac{L_0 + L_i}{B_i}, \quad d_i := \frac{D_0 + D_i}{B_i} \quad \text{and} \quad B_i := L_0 + L_i + D_0 + D_i, \tag{3.7}$$

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

Proposition 3.3.1. *The MLE quality estimator*

$$\hat{q}_i = \operatorname{argmax}\left\{\mathcal{L}_i^w(q) \ : \ q_{\min} \le q \le q_{\max}\right\}, \tag{3.8}$$

is unique and given by

$$\hat{q}_i = \begin{cases} \operatorname{proj}_{[q_{\min}, q_{\max}]}(q^*) & \text{if } l_i > d_i, \\ q_{\min} & \text{if } l_i \le d_i, \end{cases} \tag{3.9}$$

where q^* solves the following equation

$$G(p - q^*) = l_i - d_i = 2l_i - 1. \tag{3.10}$$

(All proofs can be found in the Appendix.) The maximum likelihood estimate $\hat{q}_i$ has appealing properties. Equation (3.10) shows that it depends on the data only through the fraction of like reviews. Moreover, by Lemma 3.6.1 (in the Appendix) we know that the function $G(p - q)$ is increasing in q, which implies that the estimator is increasing in the fraction of like reviews, as one would expect. Lemma 3.6.1 also implies that $G(p-q)$ is invertible for every $q \in [q_{\min}, q_{\max}]$, which means that (3.10) defines a one-to-one mapping between l_i and $\hat{q}_i$. We will exploit this observation in the subsequent analysis.

3.3.2 Asymptotic learning

The stochastic learning process converges in the sense that the quality estimate $\hat{q}_i$ converges to the true quality q almost surely. Assumption 3.2.1 implies that new information continues to enter the system since some consumers will always choose to purchase and as a result review the product. This, in turn, ultimately guarantees that learning is achieved.

Proposition 3.3.2. *Consider the sequential learning process described in the previous section, where consumer i estimates the prevailing quality $\hat{q}_i$ through* (3.9)*. Then, $\hat{q}_i \to q$ as $i \to \infty$ almost surely.*

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

3.3.3 Approximation of learning dynamics in a large market

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

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

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

flow through the system over a fixed time window, e.g., a week.[9]

We carry the notation from the previous section with the necessary adjustments. Specifically, with some abuse of notation, in the n-th system we have from (3.7) that

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

where $B^n := L_0^n + L^n + D_0^n + D^n$ and $\bar{B}^n := B^n/n$. Similarly, $d^n := (D_0 + D^n)/B^n = (\bar{D}_0 + \bar{D}^n)/\bar{B}^n =: \bar{d}^n$. The fractions of likes and dislikes are independent of n, conditional on $\bar{X}^n$. Similarly, $\hat{q}^n(X^n(t))$ is directly defined through (3.9) and (3.10). We also note that $\hat{q}^n(\bar{X}^n(t)) = \hat{q}^n(X^n(t))$ through the normalized definitions of $\bar{l}^n$ and $\bar{d}^n$, and, moreover, that the mapping $\hat{q}^n$ itself does not depend on n, that is, the same quality estimation procedure is applied throughout the scaling that we consider, and simply evaluated at the appropriate state $X^n(t)$.

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

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

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

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

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

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

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

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

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

$$\bar{X}^n(t) = \frac{1}{n} N \left(\Lambda^n \int_0^t \gamma(\bar{X}^n(s)) \, \mathrm{d}s \right), \tag{3.11}$$

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

If the rate processes inside the expressions (3.11) did not depend on the state $\bar{X}^n(t)$ itself, then a straightforward application of the functional strong law of large numbers for the Poisson process would yield a deterministic limit for $\bar{X}^n(t)$ as n grew large. Our model is a bit more complex, but because the evolution of the state $\bar{X}^n(t)$ depends on the decisions made by all predecessors, one would expect it to vary slowly relative to the increasing number of consumers arriving at any given point in time. Intuitively, considering a short time interval $[t, t+\Delta]$, one would expect that the large pool of heterogeneous consumers arriving in that interval, each with a different quality parameter α, and making decisions based on similar information given by $\bar{X}^n(s)$ for some $s \in [t, t+\Delta]$, would lead to a deterministic but state-dependent evolution of $\bar{X}^n$ for large n; effectively, the stochastic nature of the decisions due to consumer heterogeneity is "averaged out" in such a setting.

This argument is made precise in Proposition 3.3.3 that derives a deterministic limiting characterization for the system behavior as n grows large using Kurtz [1978, Theorem 2.2] through a sample path analysis based on a strong approximation argument and a subsequent application of Gronwall's inequality.

Proposition 3.3.3. *For every $t > 0$,*

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

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

$$\bar{X}(t) = \Lambda \int_0^t \gamma(\bar{X}(s)) \, \mathrm{d}s. \tag{3.12}$$

To better understand (3.12) consider the expression for the scaled number of likes,

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

This means that the scaled number of 'likes' at t is the sum over the mass of consumers who report a 'like' in each $s \leq t$, and this mass depends on past reviews via $\bar{X}(\cdot)$. It follows that the scaled number of consumers that arrive by time t is Λt and that the number of people that purchased the product and submitted a report is $\bar{B}(t) := \bar{L}_0 + \bar{D}_0 + \bar{L}(t) + \bar{D}(t)$. It is convenient to derive from (3.12) the expressions for $(\bar{l}(t), \bar{d}(t))$ in the limiting (fluid) model, since these quantities determine the decision of an arriving consumer:

$$\bar{l}(t) := l(\bar{X}(t)) = \frac{\bar{L}_0 + \bar{L}(t)}{\bar{B}(t)} \quad \text{and} \quad \bar{d}(t) := d(\bar{X}(t)) = \frac{\bar{D}_0 + \bar{D}(t)}{\bar{B}(t)} = 1 - \bar{l}(t). \tag{3.13}$$

From the definition of $(\gamma^{\mathrm{L}}, \gamma^{\mathrm{D}})$, (3.10) and (3.13), it follows that $(\bar{L}, \bar{D})$ is absolutely continuous and therefore differentiable almost everywhere. We refer to time t where $(\bar{L}, \bar{D})$ is differentiable as *regular*. At regular points t, $(\bar{L}, \bar{D})$ satisfies the differential equations:

$$\dot{\bar{L}}(t) = .5\Lambda \left[\bar{F}(p - \min(q - \bar{\varepsilon}, \hat{q}_t)) + \bar{F}(p - \min(q + \bar{\varepsilon}, \hat{q}_t))\right], \tag{3.14}$$

and

$$\begin{aligned} \dot{\bar{D}}(t) &= \Lambda \left[\bar{F}(p - \hat{q}_t) - .5\left[\bar{F}(p - \min(q - \bar{\varepsilon}, \hat{q}_t)) + \bar{F}(p - \min(q + \bar{\varepsilon}, \hat{q}_t))\right]\right] \\ &= \Lambda \bar{F}(p - \hat{q}_t) - \dot{\bar{L}}(t), \end{aligned} \tag{3.15}$$

where $\hat{q}_t$ is the maximum-likelihood estimator defined in (3.10) and evaluated at $(\bar{l}(t), \bar{d}(t))$.

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

3.3.4 Transient learning dynamics: uniformly distributed valuations

In the remainder of the paper we will assume that $\alpha \sim U[0, \bar{\alpha}]$[10] and, without loss of generality, we will normalize $q_{\min} = 0$. This allows us to simplify the ODEs (3.14) and (3.15) as follows. At regular times t, we have

$$\dot{\bar{L}}(t) = \Lambda\left(\frac{\bar{\alpha} + .5\min(q-\bar{\varepsilon}, \hat{q}_t) + .5\min(q+\bar{\varepsilon}, \hat{q}_t) - p}{\bar{\alpha}}\right) \quad \text{and} \quad \dot{\bar{D}}(t) = \Lambda\left(\frac{\bar{\alpha} + \hat{q}_t - p}{\bar{\alpha}}\right) - \dot{\bar{L}}(t), \tag{3.16}$$

and the quality estimator can now be written as

$$\hat{q}_t \;=\; p - \bar{\alpha} + \frac{\bar{\varepsilon}}{2(1-\bar{l}(t))} \;=\; p - \bar{\alpha} + \frac{\bar{\varepsilon}}{2}\left(1 + \frac{\bar{L}_0 + \bar{L}(t)}{\bar{D}_0 + \bar{D}(t)}\right). \tag{3.17}$$

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

Underestimating prior ($q_0 < q$); phase 1 of learning $\hat{q}_t < q - \bar{\varepsilon}$. At times where the prevailing quality estimate is such that $\hat{q}_t < q - \bar{\varepsilon}$, the consumers who purchase with $\bar{\alpha} + \hat{q}_t - p \geq 0$ are guaranteed to have a positive ex-post utility realization since $\bar{\alpha} + q - \bar{\varepsilon} - p \geq \bar{\alpha} + \hat{q}_t - p \geq 0$. As

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

[11] The Appendix studies the case where the prior overestimates the true quality $q_0 > q$, and also shows how to approximate the evolution of the learning ODEs at times where $q - \bar{\varepsilon} < \hat{q}_t < q + \bar{\varepsilon}$ for the case where $\bar{\varepsilon}$ is small. Moreover, note that equation (3.17) corresponds to the solution of equation (3.10) when $\bar{l}(t) \leq 1 - \bar{\varepsilon}/2\bar{\alpha}$, otherwise we have a different solution. We only consider the above solution since in the relevant cases $\bar{l}(t)$ is never too close to 1.

a result only like reviews will be submitted as long as $\hat{q}_t < q - \bar{\varepsilon}$. In this case, the ODEs (3.16) can be solved in closed form and the solutions can then be used together with (3.17) to characterize the learning trajectory for $\hat{q}_t$. We do this in Proposition 3.3.4, which is the main result of this section. Before presenting the result, we formally define the *time-to-learn*

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

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

Proposition 3.3.4. *Consider the ODEs for the learning dynamics given in* (3.16) *and assume that* $q_0 < q$. *Then, for* $t \leq \tau$,

$$\hat{q}_t = p - \bar{\alpha} + (\bar{\alpha} + q_0 - p) \exp\left(\frac{t}{w}\right). \tag{3.18}$$

Moreover,

$$\tau = w \log\left(\frac{\bar{\alpha} + q - \bar{\varepsilon} - p}{\bar{\alpha} + q_0 - p}\right). \tag{3.19}$$

Proposition 3.3.4 characterizes the learning transient in the underestimating case. In particular, expression (3.18) describes the learning trajectory of the quality estimate $\hat{q}_t$ for all $t \leq \tau$, and expression (3.19) characterizes the time-to-learn as a function of the relevant model parameters, i.e., the market heterogeneity $\bar{\alpha}$, the distance $q - q_0$ of prior quality from true quality, and the price.

First, expression (3.18) shows that $\hat{q}_t$ starts at the prior estimate q_0 at time 0 and it increases monotonically to reach $q - \bar{\varepsilon}$ at time τ. Moreover, the lower the prior weight w, the faster the learning trajectory, i.e., when consumers place less weight on their prior estimate, they are more sensitive to the review information and as a result the quality estimator is updated faster.

Furthermore, note that expression (3.19) can be rewritten as

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

which allows us to make some interesting comparative statics observations on the learning transient. In particular, note that the time-to-learn τ is decreasing in the maximum (or equivalently the range of the) base valuation $\bar{\alpha}$, because if consumers have higher valuations, more consumers choose to buy (and review) the product and thus information accumulates faster. Moreover, τ is increasing in $(q - q_0)$, i.e., the more severely consumers underestimate quality, the longer it takes to learn q. Finally, the expression for τ highlights that the speed of the learning transient is proportional to w, which is a natural time scale if we think of learning as the process of accumulating enough information to overcome the bias of the prior estimate q_0.[12]

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

Corollary 3.3.1. *The time-to-learn τ is increasing in p.*

This result can be explained as follows. Let

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

which denotes the *instantaneous demand function* at time t in our large market setting. At any given $t \leq \tau$, a price increase affects the instantaneous demand function through two channels: first, a direct channel, a higher price means a lower instantaneous demand at time t; second, through $\hat{q}_t$, a higher price means a lower $\hat{q}_t$ at time t.[13] Therefore, by increasing the price, the monopolist effectively decreases the rate at which consumers are buying (and reviewing) the product, thus slowing down learning. Finally, using the characterization of $\hat{q}_t$ from (3.18) we can rewrite $d_t(p)$,

[12] Note that the specific way in which the ODEs for the transient analysis depend on w is determined by the assumptions that we make on w in the large market approximation. In particular, recall our assumptions that as the arrival rate of consumers scales, the prior weight itself scales, and likewise the reviews corresponding to the prior L_0^n, D_0^n, all scale proportionally to the scaling constant n. This assumption is desirable because it implies that the ensuing transient of the quality estimate evolves on the natural time-scale of the system, e.g., if we measure time in days, then the prevailing quality estimate also evolves in the time frame of days. We could have assumed that the prior weight scales, for example, with order of $\sqrt{n}$ reviews, in which case we would have obtained the same ODE characterization after an appropriate rescaling of time. To understand how the analysis relates to the study of a system of original interest, consider an example where $\hat{\Lambda} = 1000$ consumers per day, and $\hat{w} = 100$ reviews. These parameters are then embedded in a sequence of systems of growing scale, say $\Lambda^n = 10 \cdot n$ and $w^n = 1 \cdot n$, which we then study asymptotically to obtain tractable characterizations of their evolution; the 100^{th} system in that sequence is the original system we wanted to analyze. We could have just as well defined a different sequence, for example $\Lambda^n = 10 \cdot n$ and $w^n = 10 \cdot \sqrt{n}$. This modeling choice affects the downstream scaling of the system processes as briefly explained above, but not the results.

[13] This can easily be verified by differentiating (3.18) with respect to p and noting that, for all $0 < t \leq \tau$, $\partial \hat{q}_t / \partial p = 1 - \exp\left(t/w\right) < 0$.

for all $t \leq \tau$, as

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

The instantaneous demand function thus takes the linear form $d_t(p) = a_t - b_t \cdot p$ with $a_t = (1 + q_0/\bar{\alpha}) \exp(t/w)$ and $b_t = (1/\bar{\alpha}) \exp(t/w)$. Note that the instantaneous demand is positive for all $p \in [p_{\min}, p_{\max}]$, which follows directly from Assumption 3.2.1, and that both the slope b_t and the intercept a_t are increasing in t. Thus, when consumers initially underestimate quality ($q_0 < q$), the instantaneous demand and consequently the instantaneous revenue are increasing with t. Similarly, note that as time passes and the quality estimate increases, the monopolist can achieve the same instantaneous demand with a higher price, thus generating more revenue. In the following sections, we will further elaborate on these insights and we will study the pricing strategy of the monopolist.

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

3.4 Static Price Analysis

In this section we solve the monopolist's problem of choosing a static price to maximize her revenue as given in (3.2). Following the analysis of the previous section, the stochastic learning trajectory is replaced by its deterministic mean field approximation. This enables us to solve an otherwise intractable problem. The next two sections focus on the price optimization problem, for the case in which consumers initially underestimate quality through their prior, i.e., $q_0 < q$. Adapting by the mean-field approximation, we can write the seller's discounted revenue as

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

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

$$\tilde{R}(p) = \Lambda \left(\int_0^\tau e^{-\delta t} \pi_t(p) \; \mathrm{d}t + \int_\tau^\infty e^{-\delta t} \pi_\infty(p) \; \mathrm{d}t \right), \tag{3.20}$$

where $\pi_\infty(p)$ denotes the instantaneous revenue at the true quality ($\hat{q}_\infty = q$).[14] The monopolist's revenue maximization problem can be written as

$$\text{maximize} \left\{ \tilde{R}(p) \;:\; p_{\min} \leq p \leq p_{\max} \right\}. \tag{3.21}$$

Before stating our main result, we define

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

which are the static monopoly prices at q_0 and q respectively.[15] The following proposition characterizes the optimal monopoly price in the presence of social learning.

Proposition 3.4.1. *Consider the case $q < q_0$. For $\bar{\varepsilon}$ sufficiently small, the monopolist revenue optimization problem* (3.21) *has a unique optimal solution $p^* = p^*(\delta, w)$ that satisfies the following:*

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

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

Proposition 3.4.1 characterizes the (unique) solution of the monopolist's revenue maximization problem. In particular, Part 3.4.1 states that the optimal price is straddled between two natural

[14] Lemma 3.6.4 in the Appendix provides an intuitive characterization of the revenue function (3.20) and establishes a bound on $|\tilde{R}(p) - \bar{R}(p)|$, which is of order $\bar{\varepsilon}$.

[15] Note that Assumption 3.2.1**??** implies that the constraints in the definition of $p^{\mathrm{m}}(q)$ are never binding and $p^{\mathrm{m}}(q) = (\bar{\alpha} + q)/2$. Moreover, Assumption 3.2.1**??** implies that the constraints in the definition of $p^{\mathrm{m}}(q_0)$ are not binding if and only if $(\bar{\alpha} + q_0)/2 > q_{\max} - \bar{\varepsilon}$, which is always true for reasonable values of $\bar{\alpha}$ and q_0.

end points: the price that a monopolist would charge if consumers did not engage in social learning and based their purchase decisions only on their prior estimate q_0; the price that a monopolist would charge if consumers were fully informed of the product quality q.

Part 3.4.1 highlights the importance of the monopolist's patience level δ and the weight w, that consumers attach to their prior, on the optimal price with social learning. In particular, if the monopolist is very patient ($\delta \approx 0$), then the optimal price is close to the price under full information, since in this case the learning transient is short relative to the extent of revenue discounting. However, if the monopolist is very impatient ($\delta \gg 0$), then she finds it optimal to significantly decrease her price, in the limit learning is not important and the monopolist prices as if $\hat{q}_t = q_0$ for all $t \geq 0$. Finally, note that the prior weight w is a natural time unit that determines the learning speed and the effect of discounting on revenues.

In what follows, we numerically illustrate the solution to (3.21) for different model parameters and derive some observations regarding comparative statics. We consider the underestimating case ($q_0 < q$) with a demand rate of 10 potential consumers per week. The most important parameters in the pricing problem are the monopolist's discount factor δ, the error in consumers prior estimate q_0 relative to the true quality, and the maximum base valuation $\bar{\alpha}$. As already noted, the learning transient also scales proportionally to the weight w attached to the prior estimate.

We consider three different prior estimates $q_0 \in \{.40, .20, .10\}$, with a prior weight of $w = 10$. The true quality is $q = 2$, and we set the small quality disturbance term $\bar{\varepsilon}$ to 5% of the true quality. The monopolist is either patient, semi-patient, or impatient, corresponding to annualized discount rates $\delta \in \{2.5\%, 7.5\%, 15\%\}$. We fix $\bar{\alpha} = 4$ and we think that this is a reasonable value for this parameter, which corresponds to a maximum quality premium of $2q$.[16]

The *left plot* in Figure 3.1 highlights how the optimal price p^* varies with the prior q_0 and the monopolist's patience level. The monopoly price under full information $p^{\mathrm{m}}(q)$, which is normalized to 1, and the monopolist price at q_0 are also plotted in black. In line with our theoretical result, we see that the optimal price with social learning is always between the static monopoly price at q_0 and the static monopoly price at q. Moreover, the price p^* is closer to the static monopoly price

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

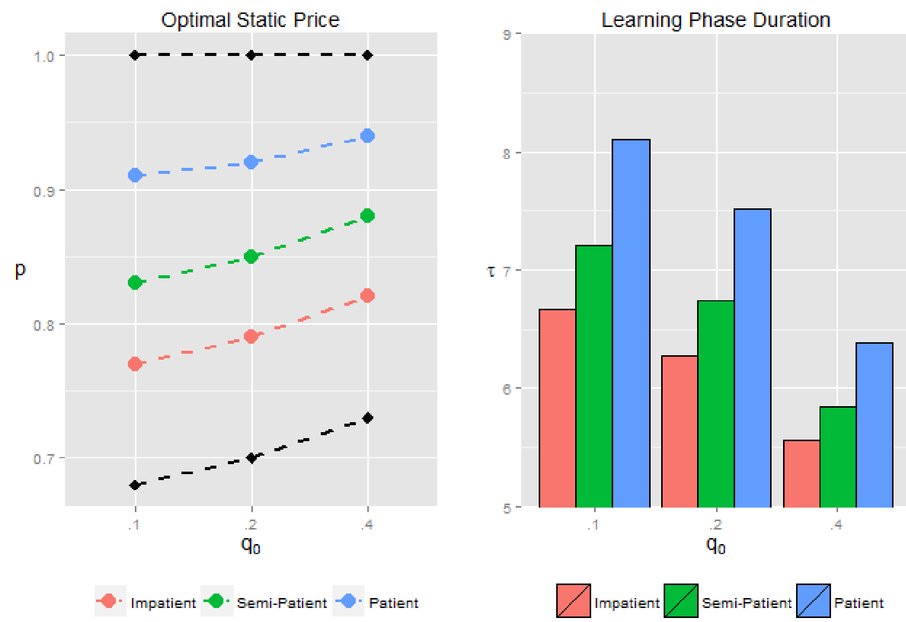


Figure 3.1: Optimal Static Price and Learning Phase Duration.

under full information when the monopolist is more patient and consumers' prior estimate q_0 is closer to q. On the contrary, when the monopolist is impatient, her optimal price is always closer to the static monopoly price at q_0. The *right plot* in Figure 3.1 reports the learning phase duration τ^* for different values of q_0 and different patience levels. It always takes 5.5 to 8.5 weeks for the quality estimate to get $\bar{\varepsilon}$-close to q. This observation is not surprising, since the facts that $w = 10$ and $q - \bar{\varepsilon} = 1.9$ imply that τ^* scales with

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

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

3.5 Two Price Analysis

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

3.5.1 Optimal Prices

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

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

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

moreover consumers use the following weight for the new prior,

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

The (q_1, w_1) specification incorporates the fact that the reviews before time s were under a different price point. From time s onward the problem is analogous to the single price version studied in the previous section. In particular, after time s the learning process evolves according to equation (3.18), with initial condition q_1, price p_1, and prior weight w_1. The expected discounted revenue of the monopolist is given by

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

As in the static price case, we study the situation in which once the learning process has converged to $\hat{q}_t = q - \bar{\varepsilon}$, revenues are accrued from then on according to $\hat{q}_t = q$. In this setting, it is without loss of generality to focus the attention on policies such that $s \leq \tau$.[17] Which leads to the following revenue function

$$\tilde{R}(p_0, p_1, s) = \Lambda\left(\int_0^s e^{-\delta t}\pi_t(p_0)\ \mathrm{d}t + \int_s^\tau e^{-\delta t}\pi_t(p_1)\ \mathrm{d}t + \int_\tau^\infty e^{-\delta t}\pi_\infty(p_1)\ \mathrm{d}t\right). \tag{3.22}$$

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

$$\begin{aligned} \text{maximize} \quad & \tilde{R}(p_0, p_1, s) \\ \text{s.t.} \quad & p_0, p_1 \in [p_{\min}, p_{\max}] \\ & s \leq \tau, \end{aligned} \tag{3.23}$$

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

Proposition 3.5.1. *Consider the case $q_0 < q$ and assume that $\bar{\varepsilon}$ is sufficiently small. Let (p_0^*, p_1^*, s^*) be the optimal solution to* (3.23). *Then, the optimal prices (p_0^*, p_1^*) are such that $p_0^* \leq p_1^*$ and $p_1^* \in [p^{\mathrm{m}}(q_1), p^{\mathrm{m}}(q)]$.*

Proposition 3.5.1 states that, when consumers underestimate quality, the optimal price p_1^* is always between the static monopoly price $p^{\mathrm{m}}(q_1)$ and monopoly price under full information, $p^{\mathrm{m}}(q)$.

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

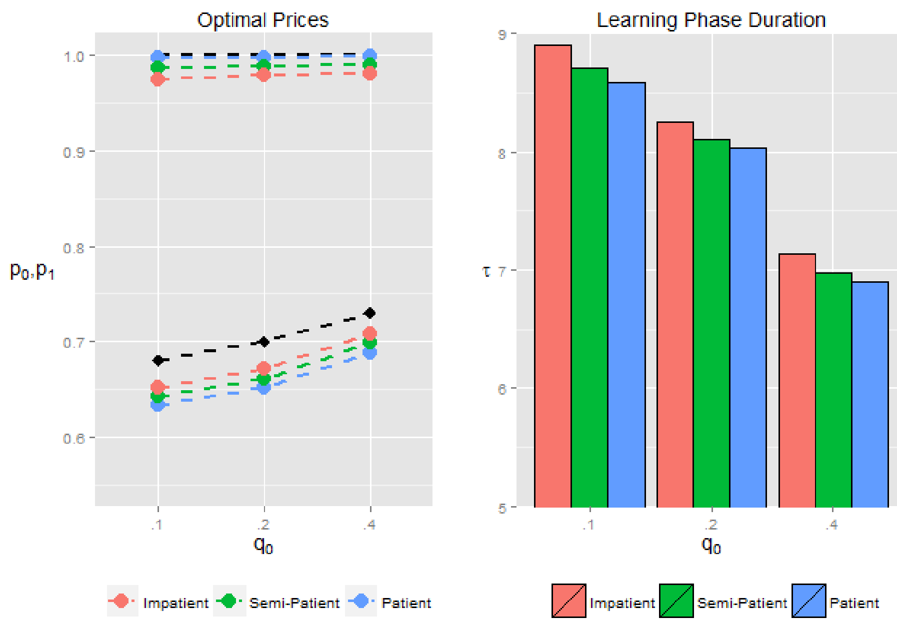


Figure 3.2: Optimal Prices and Learning Phase Duration.

Moreover, the optimal price path is increasing. This supports the intuition that the monopolist has an incentive to lower the initial price in order to speed up learning.

Figure 3.2 displays numerical solutions to (3.23) for different model parameters, which provide some additional insights on the optimal pricing policy.[18] The *left plot* in Figure 3.2 shows the optimal prices p_1^* (above) and p_0^* (below) for different priors and different monopolist's patience levels. We clearly see that, in the two-price case, the monopolist may find it optimal to initially price below $p^{\mathrm{m}}(q_0)$ in order to speed up learning, and then switch to a price which is very close to the full information monopoly price $p^{\mathrm{m}}(q)$ in order to extract more revenues. Moreover, if the monopolist is more patient, i.e., δ is small, then the first period price p_0^* is lower. This has interesting implications for the speed of learning: the *right plot* in Figure 3.2 shows that a patient monopolist, who is willing to sacrifice initial revenues by under-pricing more aggressively, achieves faster learning than an impatient one.

[18]The choice of parameter values for the numerical experiments is the same as the one described in Section 3.4.

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

Table 3.1: %-gap in revenues relative to full information scenario.

3.5.2 Revenue Comparison of Pricing Policies

In this section we numerically compare the revenue performance of the optimal *static* price policy and the optimal *two* prices policy. Our measure of revenue performance for a given policy is the %-gap between the total revenue attained by using that policy and the total revenue that the monopolist would attain in an ideal scenario in which consumers know q and the monopolist charges the monopoly price under full information.[19] We also report revenue performances for two benchmark policies: *prior* and *true*. For these cases, revenues are computed under the assumption that consumers follow the learning process specified in our model, but the monopolist does not take it into account and she charges the static price $p^{\mathrm{m}}(q_0)$ and the static price $p^{\mathrm{m}}(q)$ respectively.

The first two columns in Table 3.1 show the revenue performance of the *static* price and of the *two* prices policies respectively. The optimal two-period pricing policy performs consistently better than the optimal static price, and the relative revenue improvement becomes more significant as the seller's discount factor increases. In the last two columns of Table 3.1 we report the revenue performance of the *prior* and *true* policies respectively. By comparing the static price policy to these two benchmark policies we can appreciate the effectiveness of taking social learning into account when devising an optimal pricing policy.

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

3.6 Proofs

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

Proofs of Section 3.3

Lemma 3.6.1. *The function $G(x)$ defined in (3.3) is non-increasing for all $x \leq \alpha_{\max}$, and is strictly decreasing for all $x \in [\alpha_{\min} - \bar{\varepsilon}, \alpha_{\max} - \bar{\varepsilon}]$. Equivalently, $G(p-q)$ is non-decreasing in q for all $q \geq p - \alpha_{\max}$ and is strictly increasing in q for all $q \in [q_{\min}, q_{\max}]$.*

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

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

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

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

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

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

Proof of Proposition 3.3.1. Taking logs of the weighted likelihood function in (3.6) yields

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

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

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}q}\log(\mathcal{L}_i^w(q)) &= \frac{L_0+L_i}{1+G(p-q)}G'(p-q) - \frac{D_0+D_i}{1-G(p-q)}G'(p-q)\\ &= \left(\frac{L_0+L_i+D_0+D_i}{(1+G(p-q))(1-G(p-q))}\right)G'(p-q)(l_i-d_i-G(p-q)).\end{aligned} \tag{3.24}$$

Lemma 3.6.1 and Assumption 3.2.1 imply that $0 < G(p-q) < 1$ for all $q \in [q_{\min}, q_{\max}]$ and moreover that $G(p-q)$ is strictly increasing in q for all $q \in [q_{\min}, q_{\max}]$. We will now use these observations and (3.24) to construct the unique optimal solution to problem (3.8).

First, note that $0 < G(p-q) < 1$ implies that the denominator in (3.24) is always positive. Since $G(p-q)$ is strictly increasing in q, then $l_i - d_i - G(p-q)$ is strictly decreasing in q. If $l_i - d_i - G(p-q_{\min}) \leq 0$, then

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

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

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

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

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

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

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

The following lemma is instrumental in the proof of Proposition 3.3.2.

Lemma 3.6.2. *Suppose that Assumption 3.2.1*?? *holds, then* $\sum_{i=1}^{\infty}(B_i)^{-2} < \infty$ *almost surely.*

Proof. The proof will proceed as follows. First, we rewrite the process $\{B_i, i = 1, 2, \dots\}$ in a more convenient form $\{B_0 + X_i, i = 1, 2, \dots\}$, where $B_0 := L_0 + D_0$ is the total number of initial reviews associated to the prior q_0, defined in (3.4) and (3.5), and X_i is an appropriately defined sequence of random variables. Then, we bound from below the process $\{X_i, i = 1, 2, \dots\}$ with a process $\{Y_i, i = 1, 2, \dots\}$, which is more tractable for the purpose of the analysis. Finally, we use the Strong Approximation Theorem Glynn [1990, Theorem 5] to show that $\sum_{i=1}^{\infty}(B_0 + Y_i)^{-2} < \infty$ almost surely and complete the proof.

Note that Assumption 3.2.1**??** implies that for all i there exists an $\eta > 0$ such that for all admissible p the following is true:

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

where I_i is defined as in (3.1). Thus, we can rewrite B_i in the form $B_i = B_0 + \sum_{j=1}^{i} \chi_j(\eta_j)$, where χ_j is a Bernoulli random variable with success (i.e., purchase) probability $\eta_j > \eta$ for all j, and η_j depends on I_j and the price p. Let $X_i = \sum_{j=1}^{i} \chi_j$ for all $i = 1, 2, \dots$.

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

Finally, the Strong Approximation Theorem Glynn [1990, Theorem 5] implies that there exist a probability space that supports a standard Brownian motion W and a sequence Y'_i such that $\{Y'_i : i \geq 1\} \stackrel{\mathcal{D}}{=} \{Y_i : i \geq 1\}$, and

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

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

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

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

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

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

Proof of Proposition 3.3.2. We study the evolution of l_i, and relate it to $\hat{q}_i$ using (3.10). It is easy to verify that l_i evolves according to the stochastic recursion

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

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

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

which is equivalent to the recursion

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

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

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

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

and note that Lemma 3.6.1 implies that $G^{-1}(2l_i - 1)$ is well-defined for all $l_i \in [l_{\min}, l_{\max}]$.

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

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

and thus

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

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

Assumption (A.2.2). We have that

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

where the last equality follows by substituting

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

Thus, we can define the function

$$g(l_i) := \mathsf{E}[Y_i \mid l_0, Y_j, j < i], \tag{3.25}$$

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

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

Assumption (A.2.4) is shown in Lemma 3.6.2.

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

Before applying the theorem it is useful to further decompose l_i as follows. First note that $Y_i := g(l_i) + \bar{M}_i$, where the function $g(l_i)$ is the drift function that we defined in (3.25) and $\bar{M}_i$ is a martingale difference noise given by

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

Then, it is straightforward to see that

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

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

Case 1: $l < l^*$ **(or equivalently** $\hat{q}(l) < q$**).** In this case $\min(\hat{q}(l), q + \bar{\varepsilon}) = \hat{q}(l)$ and $g(l) = .5\big[\bar{F}(p - \min(\hat{q}(l), q - \bar{\varepsilon})) - \bar{F}(p - \hat{q}(l) + \bar{\varepsilon})\big]$. In addition, $-\min(\hat{q}(l), q - \bar{\varepsilon}) \geq -\hat{q}(l) + \bar{\varepsilon}$. If

$\hat{q}(l) \leq q - \bar{\varepsilon}$, then $g(l) = .5\big[\bar{F}(p - \hat{q}(l)) - \bar{F}(p - \hat{q}(l) + \bar{\varepsilon})\big] > 0$, since $\bar{F}$ is a decreasing function. If $q > \hat{q}(l) > q - \bar{\varepsilon}$, then $g(l) = .5\left[\bar{F}(p - q + \bar{\varepsilon}) - \bar{F}(p - \hat{q}(l) + \bar{\varepsilon})\right] > 0$, since $\hat{q}(l) < q$. We conclude that $g(l) > 0$ in this case.

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

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

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

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

Lemma 3.6.3. *[(a)]*

1. *For all* $x, y, z \in \mathbb{R}$ *we have*

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

2. *For* $x_1, x_2 \geq a > 0$ *and* $y_1, y_2 \geq b > 0$

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

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

2 From the triangular inequality,

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

□

Proof of Proposition 3.3.3. Throughout this proof, to reduce the notational burden and without loss of generality, we rescale time such that $\Lambda = 1$. We verify the conditions of Theorem 2.2 of Kurtz [1978]. First we note that $X^n(t) \in \mathbb{Z}_+^2$, $\bar{X}^n(t) = X^n(t)/n \in \{k/n \mid k \in \mathbb{Z}_+^2\}$ as required. To satisfy the conditions of the theorem we validate the construction (3.12) and then show that the following inequalities hold

$$\gamma(x) \le \Gamma_1(1+|x|) \quad \text{and} \quad |\gamma(x)-\gamma(y)| \le \Gamma_2|x-y| \tag{3.26}$$

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

The integral form of $\bar{X}^n(t)$ in (3.12) follows from Poisson arrivals and Poisson thinning of the standard Poisson process N. For example, $\bar{L}^n(t)$ can be written in the form,

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

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

Finally, focusing on the rate conditions required for Theorem 2.2 of Kurtz [1978]. The first inequality in (3.26) holds for $\Gamma_1 = 1$ since γ^k are probabilities for $k = \mathrm{L}, \mathrm{D}$. We derive the last inequality there for γ^{L} in two steps. First, we observe that $\gamma^{\mathrm{L}}(X^n(t))$ depends on $X^n(t)$ through $\hat{q}(X^n(t))$. It follows from Lemma 3.6.31 and the fact that the density of α is uniformly bounded by $f_{\max}$ that γ^{L} is Lipschitz continuous in $q \in [q_{\min}, q_{\max}]$. Second, we show that $\hat{q}(x)$ is Lipschitz continuous in $x = X^n(t)$. Lemma 3.6.32 for $a = L_0$ and $b = D_0$ establishes that q^*, defined in (3.10), is Lipschitz continuous in $x = X^n(t)$. The projection operator of q^* onto $[q_{\min}, q_{\max}]$ preserves the Lipschitz property. Similarly, one can show that γ^{D} is Lipschitz continuous, which establishes that γ is Lipschitz. Finally, invoking Theorem 2.2 of Kurtz [1978] we complete the proof. □

Proof of Proposition 3.3.4. The ODEs (3.16) reduce to

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

noting that $\bar{D}(t) = 0$ and plugging (3.17) into the expression for $\dot{\bar{L}}(t)$ yields a linear ODE for $\bar{L}(t)$, i.e., $\dot{\bar{L}}(t) = \dfrac{\Lambda\bar{\varepsilon}}{2\bar{\alpha}}(1 + \bar{L}_0/\bar{D}_0 + \bar{L}(t)/\bar{D}_0)$. This equation is readily solvable in closed form, and the particular solution with initial condition $\bar{L}_0$ is given by

$$\bar{L}(t) = (\bar{L}_0 + \bar{D}_0)\left[\exp\left(\frac{\Lambda\bar{\varepsilon}}{2\bar{\alpha}\bar{D}_0}t\right) - 1\right]. \tag{3.27}$$

The trajectory for $\hat{q}_t$ can now be obtained by replacing (3.27) and $\bar{D}(t) = 0$ into (3.17), which yields

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

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

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

Rewriting (3.4) and (3.5) for α uniformly distributed we get that

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

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

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

□

In the sequel, we provide a sketch for the solution of the ODE's derived in the main body of the paper for the cases not considered in Proposition 3.3.4.

Overestimating prior ($q_0 > q$); phase 1 of learning $\hat{q}_t > q + \bar{\varepsilon}$. The following proposition characterizes the trajectories for $\bar{L}(t)$ and $\bar{D}(t)$ in this case. From these one can easily obtain the trajectory for $\hat{q}_t$ by using (3.17).

Proposition 3.6.1. *Consider the ODEs for the learning dynamics given in* (3.16) *and assume that* $q_0 > q$. *Then, for* $t \le \tau$,

$$\bar{L}(t) = \Lambda\left(\frac{\bar{\alpha} + q - p}{\bar{\alpha}}\right)t, \tag{3.28}$$

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

$$\left(1 + \frac{\bar{L}(t) + \bar{D}(t)}{\bar{L}_0 + \bar{D}_0}\right)^{\bar{\alpha}+q-p}\left(1 + \frac{(\bar{\varepsilon}/2)\bar{L}(t) - (\bar{\alpha} + q - p)\bar{D}(t)}{(\bar{\varepsilon}/2)\bar{L}_0 - (\bar{\alpha} + q - p)\bar{D}_0}\right)^{\bar{\varepsilon}/2} = 1. \tag{3.29}$$

Proof. (Sketch only.) In this case the ODEs (3.16) reduce to

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

It can be easily verified that the solution for $\bar{L}(t)$ with initial conditions $\bar{L}_0$ is given by (3.28), and substituting it into the ODE for dislikes yields

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

Which is an ODE of the form

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

where

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

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

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

In fact, we have

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

This equation for $Z(t)$ is solvable, which, in turn, leads to the solution for $\bar{D}(t)$ given in (3.29). □

Phase 2 of learning $q-\bar{\varepsilon} < \hat{q}_t < q+\bar{\varepsilon}$. In this case the ODEs are less tractable, nonetheless we will establish a useful structural property for $\hat{q}_t$ in Proposition 3.6.2, and then provide an approximation of the $\hat{q}_t$ trajectory when $\bar{\varepsilon}$ is small.

Proposition 3.6.2. *Assume that $q-\bar{\varepsilon} < \hat{q}_{t_0} < q+\bar{\varepsilon}$ for some $t_0 > 0$. Then, $\hat{q}_t \to q$ as $t \to \infty$. Moreover, if $\hat{q}_{t_0} < q$ then $\hat{q}_t$ is strictly monotonically increasing for all $t \geq t_0$, otherwise, if $\hat{q}_{t_0} > q$ then $\hat{q}_t$ is strictly monotonically decreasing for all $t \geq t_0$.*

Proof. (Sketch only.) The ODEs (3.16) reduce to

$$\dot{\bar{L}}(t) = \frac{\Lambda\bar{\varepsilon}}{4\bar{\alpha}}\left(1 + \frac{\bar{L}_0 + \bar{L}(t)}{\bar{D}_0 + \bar{D}(t)}\right) + \frac{\Lambda}{2\bar{\alpha}}(\bar{\alpha} + q - \bar{\varepsilon} - p), \tag{3.30}$$

$$\dot{\bar{D}}(t) = \frac{\Lambda\bar{\varepsilon}}{4\bar{\alpha}}\left(1 + \frac{\bar{L}_0 + \bar{L}(t)}{\bar{D}_0 + \bar{D}(t)}\right) - \frac{\Lambda}{2\bar{\alpha}}(\bar{\alpha} + q - \bar{\varepsilon} - p). \tag{3.31}$$

It follows from Assumption 3.2.1 that $p < \bar{\alpha} - \bar{\varepsilon}$, which implies that $\dot{\bar{L}}(t) > 0$ for all t, moreover, using equation (3.17) for $\hat{q}_t$, it is easy to see that $\dot{\bar{D}}(t) > 0$ if and only if $\hat{q}_t > q - \bar{\varepsilon}$. These, coupled with the fact that $L(t_0)$ and $D(t_0)$ are strictly positive, imply that $\bar{L}(t)$ and $\bar{D}(t)$ are positive and increasing for all $t \geq t_0$. Using $\bar{l}(t) = \bar{L}_0 + \bar{L}(t)/(\bar{L}_0 + \bar{L}(t) + \bar{D}_0 + \bar{D}(t))$, we can now write the ODE for the fraction of likes $\bar{l}(t)$ as

$$\dot{\bar{l}}(t) = \frac{1}{\bar{L}_0 + \bar{L}(t) + \bar{D}_0 + \bar{D}(t)}\left[(1 - \bar{l}(t))\dot{\bar{L}}(t) - \bar{l}(t)\dot{\bar{D}}(t)\right], \tag{3.32}$$

and noting that $\bar{L}_0 + \bar{L}(t) + \bar{D}_0 + \bar{D}(t) > 0$ for all $t \geq t_0$, we have that the steady state l^* must be such that (3.32), evaluated at l^*, is equal to 0. Noting that $1+(\bar{L}_0+\bar{L}(t))/(\bar{D}_0+\bar{D}(t)) = 1/(1-\bar{l}(t))$ and replacing (3.30) and (3.31) into condition (3.32) we find the unique steady state

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

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

When $\bar{\varepsilon}$ is small we can approximate the trajectories of $\bar{L}(t)$ and $\bar{D}(t)$ as follows. Note that the initial conditions for the system of ODEs defined by (3.30) and (3.31), when $\hat{q}_\tau = q - \bar{\varepsilon}$, are given by $(t_0 = \tau, \bar{L}_\tau, \bar{D}_\tau, \dot{\bar{L}}_\tau, \dot{\bar{D}}_\tau)$ where

$$\bar{L}_\tau = w\frac{\Lambda}{\bar{\alpha}}(\bar{\alpha} + q - 2\bar{\varepsilon} - p), \quad \bar{D}_\tau = w\frac{\Lambda}{2\bar{\alpha}}\bar{\varepsilon}, \quad \dot{\bar{L}}_\tau = \frac{\Lambda}{\bar{\alpha}}(\bar{\alpha} + q - \bar{\varepsilon} - p), \quad \dot{\bar{D}}_\tau = 0. \tag{3.33}$$

First, note that we can write

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

From (3.30) and (3.33) it is easy to see that $\dot{\bar{L}}(s) - \dot{\bar{L}}_\tau = \dot{\bar{D}}(s)$, and (3.31) and (3.33) imply that

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

We already established that $\dot{\bar{D}}(s) > 0$ if and only if $\hat{q}_s > q - \bar{\varepsilon}$, it follows from the definition of τ and Proposition 3.6.2 that $\dot{\bar{D}}(s) > 0$ for all $s > \tau$. Using this property, it is easy to verify that

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

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

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

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

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

where

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

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

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

Substituting the above ratio into equation (3.31), and differentiating twice one obtains a third order, non-homogeneous linear ODE, that is solvable in terms of the matrix exponential (not in closed form due to the non-homogeneous coefficients).

Proofs of Section 3.4

Lemma 3.6.4. *The monopolist's revenue function* (3.20) *can be written as*

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

where

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

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

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

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

substituting into the first term in the right-hand side of (3.20) yields

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

Similarly, the second term in the right-hand side of (3.20) simplifies to

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

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

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

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

Proof of Proposition 3.4.1. The proof will proceed as follows. First we establish that for $\bar{\epsilon}$ sufficiently small the revenue maximization problem (3.21) admits a unique optimal solution. Then we establish its properties, proving in order Part 3.4.1 and Part 3.4.1.

Consider the expression for $\tilde{R}(p)$ given in Lemma 3.6.4. Differentiating with respect to p once, we get

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

and, differentiating twice, we get

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

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

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

where

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

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

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

It follows from Assumption 3.2.1–*(ii)* that the revenue function $\tilde{R}(p) = \tilde{R}(p,\bar{\varepsilon}) \in C^\infty(\bar{\varepsilon})$ for all $p \in [0, p_{\max}]$. This can be easily verified by noting that $p_{\max} < \bar{\alpha} - \bar{\varepsilon}$ implies that the quantity

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

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

$p \in [p_{\min}, p_{\max}]$. Evaluating second-derivative $\tilde{R}''(p, \bar{\varepsilon})$ at $\bar{\varepsilon} = 0$ yields

$$\tilde{R}''(p,0) = \Lambda \cdot \frac{2 - \left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}+q_0-p}\right)^{-\delta w}\left(\frac{2}{\delta w} + (q-q_0)\left(\frac{2\bar{\alpha}-2p-q+3q_0-\delta w(q-q_0)}{(\bar{\alpha}+q_0-p)^2} - \frac{\bar{\alpha}+q}{\bar{\alpha}+q-p}\cdot\frac{(q-q_0)(\delta w-1)}{(\bar{\alpha}+q_0-p)^2}\right)\right)}{\bar{\alpha}(1/w-\delta)}. \tag{3.34}$$

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

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

and the numerator of (3.34) is strictly increasing in this term, replacing $(\bar{\alpha}+q)/(\bar{\alpha}+q-p)$ with 1 we find that the numerator of (3.34) is strictly bigger than

$$2\left(1 - \left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}+q_0-p}\right)^{-\delta w}\left(\frac{1}{\delta w} + \frac{q-q_0}{\bar{\alpha}+q_0-p}\right)\right) \tag{3.35}$$

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

$$2\left(1 - \left(\frac{\bar{\alpha}+q-p}{\bar{\alpha}+q_0-p}\right)^{1-\delta w}\right). \tag{3.36}$$

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

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

This implies that (3.36) is strictly bigger than 0 and so is the numerator of (3.34). Thus, we conclude that $\tilde{R}''(p,0) < 0$ for all $p \in (0, \bar{\alpha}+q_0)$.

Case 2: $\delta w < 1$. The denominator in equation (3.34) is always positive. Moreover, since

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

and the numerator is strictly decreasing in this term, replacing $(\bar{\alpha}+q)/(\bar{\alpha}+q-p)$ with 1 we find that the numerator of (3.34) is strictly smaller than the quantity in (3.35). Moreover, since $\delta w < 1$, replacing $1/\delta w$ with 1 in (3.35) we find that the quantity in (3.35) is strictly smaller than the quantity in (3.36). Now note that $q > q_0$ and $\delta w < 1$ imply that

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

This implies that (3.36) is strictly smaller than 0 and so is the numerator of (3.34). Thus, we conclude that $\tilde{R}''(p,0) < 0$ for all $p \in (0, \bar{\alpha}+q_0)$.

Case 1 and Case 2 establish that $\tilde{R}''(p,0) < 0$ for all $p \in (0,\bar{\alpha}+q_0)$. Moreover, since $\tilde{R}''(p,\bar{\varepsilon})$ is continuous at $\bar{\varepsilon}=0$, there exists $\varepsilon' > 0$ such that $\tilde{R}''(p,\bar{\varepsilon}) < 0$ for all $\bar{\varepsilon} < \varepsilon'$. In particular, this implies that when $\bar{\varepsilon}$ is sufficiently small, $\tilde{R}(p)$ is strictly concave for all $p \in [p_{\min}, p_{\max}]$, and therefore Problem (3.21) admits a unique optimal solution.

Proof of Part 3.4.1. First, note that, since $\tilde{R}(p)$ is strictly concave, then $\tilde{R}'(p)$ is strictly decreasing for all $p \in [p_{\min}, p_{\max}]$.

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

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

since $q > q_0$, thus $p^* > p^{\mathrm{m}}(q_0)$. It follows from Assumption 3.2.1**??** that $p^{\mathrm{m}}(q) = (\bar{\alpha}+q)/2$, so evaluating the first-derivative at $(p^{\mathrm{m}}(q),0)$ we have

$$\tilde{R}'(p^{\mathrm{m}}(q),0) = \Lambda \cdot (q-q_0)\frac{1-\left(\dfrac{\bar{\alpha}+q}{\bar{\alpha}+2q_0-q}\right)^{1-\delta w}}{\bar{\alpha}(1/w-\delta)}. \tag{3.37}$$

Note that $q > q_0$ implies that

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

which in turn implies that when $\delta w > 1$ the denominator in (3.37) is always strictly negative and the numerator is always strictly positive. When $\delta w < 1$ the reverse inequalities hold for the denominator and numerator in (3.37), therefore $\tilde{R}'(p^{\mathrm{m}}(q), 0) < 0$. By continuity of $\tilde{R}'(p, \bar{\varepsilon})$ at $\bar{\varepsilon} = 0$, there exists $\varepsilon'' > 0$ such that $\tilde{R}'(p^{\mathrm{m}}(q_0), \bar{\varepsilon}) > 0$ and $\tilde{R}'(p^{\mathrm{m}}(q), \bar{\varepsilon}) < 0$ for all $\bar{\varepsilon} < \varepsilon''$. Thus proving that, for $\bar{\varepsilon}$ sufficiently small, we have $p^* \in [p^{\mathrm{m}}(q_0), p^{\mathrm{m}}(q)]$.

Proof of Part 3.4.1. Setting $\tilde{R}'(p) = 0$, we get the following first-order condition for the monopolist's problem

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

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

$$(1-\omega(p)) \cdot \pi_0'(p) + \omega(p) \cdot \pi_\infty'(p) + \xi(p) = 0 \tag{3.38}$$

where

$$\omega(p) = \frac{h_\infty(p)/h_0(p)}{1 + h_\infty(p)/h_0(p)} = \frac{(\bar{\alpha}+q_0-p)(1-\delta w)}{\bar{\alpha}+q_0-p+\delta w\left(q-\bar{\varepsilon}-q_0-(\bar{\alpha}+q_0-p)\left(\frac{\bar{\alpha}+q-\bar{\varepsilon}-p}{\bar{\alpha}+q_0-p}\right)^{\delta w}\right)} \tag{3.39}$$

and

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

Clearly, as $\delta \to 0$ or $w \to 0$ we have $\omega(p) \to 1$ and $\xi(p) \to 0$ for all $p \in (0, \bar{\alpha} + q_0)$, consequently the left-hand side of (3.38) goes to $\pi'_\infty(p)$. Finally, let $p^* = p^*(\delta, w)$ be the unique solution of (3.38) and note that by strict concavity $R''(p^*) < 0$, then p^* converges to the solution of $\pi'_\infty(p) = 0$, i.e. $p^*(\delta, w) \to p^{\mathrm{m}}(q)$ as $\delta \to 0$ or $w \to 0$. For the other case, as $\delta \to \infty$ or $w \to \infty$ we have $\omega(p) \to 0$ for all $p \in (0, \bar{\alpha} + q_0)$. This is easy to verify by first dividing numerator and denominator in (3.39) by δw, next noting that

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

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

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

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

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

then $\xi(p) \to 0$ as $\delta \to \infty$ or $w \to \infty$. Thus, the left-hand side of (3.38) goes to $\pi'_0(p)$ and p^* converges to the solution of $\pi'_0(p) = 0$, i.e. $p^*(\delta, w) \to (\bar{\alpha} + q_0)/2$ as $\delta \to \infty$ or $w \to \infty$, if $(\bar{\alpha} + q_0)/2 \geq p_{\min}$. Otherwise, if $p_{\min} > (\bar{\alpha} + q_0)/2$ then $p^*(\delta, w) \to p_{\min}$ as $\delta \to \infty$ or $w \to \infty$. Recalling that $p^{\mathrm{m}}(q) = \max\{p_{\min}, (\bar{\alpha} + q_0)/2\}$ completes the proof. $\square$

Proofs of Section 3.5

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

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

When $s \leq \tau$ the above equation reduces to (3.22), when $s > \tau$ the revenue function is given by

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

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

Before proving the proposition, we introduce the following definition

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

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

The following lemma is needed for the proof of Proposition 3.5.1.

Lemma 3.6.5. *The monopolist's revenue function* (3.22) *can be written as*

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

where

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

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

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

Following the argument of the proof of Lemma 3.6.4, we establish the desired result. □

Proof of Proposition 3.5.1. First we establish that $p_1^* \in [p^{\mathrm{m}}(q_1), p^{\mathrm{m}}(q)]$. Differentiating $\tilde{R}$ twice with respect to p_1 and then evaluating at $\bar{\varepsilon} = 0$ yields

$$\frac{\partial^2}{\partial p_1^2} \tilde{R}(p_0, p_1, s)\bigg|_{\bar{\varepsilon}=0} = \Lambda \cdot e^{-s\delta} \cdot \frac{2 - \left(\frac{\bar{\alpha} + q - p_1}{\bar{\alpha} + q_1 - p_1}\right)^{-\delta w_1}}{\bar{\alpha}(1/w_1 - \delta)} \cdot \frac{\frac{2}{\delta w_1} + (q - q_1)\left(\frac{2\bar{\alpha} - 2p_1 - q + 3q_1 - \delta w_1(q - q_1)}{(\bar{\alpha} + q_1 - p_1)^2} - \frac{\bar{\alpha} + q}{\bar{\alpha} + q - p_1} \cdot \frac{(q - q_1)(\delta w_1 - 1)}{(\bar{\alpha} + q_1 - p_1)^2}\right)}{\bar{\alpha}(1/w_1 - \delta)}. \tag{3.40}$$

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

Case 1: $\delta w > 1$. The denominator of the equation (3.40) is always negative. Since

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

and the numerator of (3.40) is strictly increasing in this term, replacing $(\bar{\alpha} + q)/(\bar{\alpha} + q - p_1)$ with

1 we find that the numerator of (3.40) is strictly bigger than

$$2\left(1-\left(\frac{\bar{\alpha}+q-p_1}{\bar{\alpha}+q_1-p_1}\right)^{-\delta w_1}\left(\frac{1}{\delta w_1}+\frac{q-q_1}{\bar{\alpha}+q_1-p_1}\right)\right) \tag{3.41}$$

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

$$2\left(1-\left(\frac{\bar{\alpha}+q-p_1}{\bar{\alpha}+q_1-p_1}\right)^{1-\delta w_1}\right). \tag{3.42}$$

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

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

This implies that (3.42) is strictly bigger than 0 and so is the numerator of (3.40). Thus, we conclude that (3.40) is strictly smaller than 0 for all $p_1 \in (0, \bar{\alpha}+q_1)$.

Case 2: $\delta w < 1$. The denominator in equation (3.40) is always positive. Since

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

and the numerator of (3.40) is strictly decreasing in this term, replacing $(\bar{\alpha}+q)/(\bar{\alpha}+q-p_1)$ with 1 we find that the numerator of (3.40) is strictly smaller than the quantity in (3.41). Moreover, since $\delta w_1 < 1$, replacing $1/\delta w_1$ with 1 in (3.41) we find that the quantity in (3.41) is strictly smaller than the quantity in (3.42). Now note that $q > q_1$ and $\delta w_1 > 1$ imply that

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

This implies that (3.42) is strictly smaller than 0 and so is the numerator of (3.40). Thus, we conclude that (3.40) is strictly less than 0 for all $p_1 \in (0, \bar{\alpha}+q_1)$.

Case 1 and Case 2 prove that

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

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

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

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

since $q > q_1$. It follows from Assumption 3.2.1**??** that $p^{\mathrm{m}}(q) = (\bar{\alpha} + q)/2$, so evaluating the first-derivative at $p_1 = p^{\mathrm{m}}(q)$ and $\bar{\varepsilon} = 0$ we have

$$\left.\frac{\partial}{\partial p_1}\tilde{R}(p_0, p^{\mathrm{m}}(q), s)\right|_{\bar{\varepsilon}=0} = \Lambda \cdot e^{-\delta s} \cdot (q - q_1)\frac{1 - \left(\frac{\bar{\alpha} + q}{\bar{\alpha} + 2q_1 - q}\right)^{1-\delta w_1}}{\bar{\alpha}(1/w_1 - \delta)}. \tag{3.43}$$

Note that $q > q_1$ implies that

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

which implies that when $\delta w_1 > 1$ the denominator in (3.43) is always strictly negative and the numerator is always strictly positive. When $\delta w_1 < 1$ the reverse inequalities hold for the denominator and numerator in (3.43). Therefore

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

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

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

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

Finally, to establish that $p_0^* \leq p_1^*$ we first obtain an equivalent problem (see Boyd and Vandenberghe [2004, Chapter 4, Section 4.1.3] for the formal definition of equivalent optimization problems) by making a change of variable. Noting that $s \leq \tau$ if and only if $s \leq \tau_0$, we replace the last constraint in the monopolist's optimization problem and set $\phi = s/\tau_0$, equivalently $s = \phi\tau_0$. We let the monopolist choose $\phi \in [0,1]$ instead of s. It is clear that the optimal solution (p_0^*, p_1^*, s^*) of the original problem, can be readily obtained from the optimal solution (p_0^*, p_1^*, ϕ^*) of the transformed problem and vice versa, thus the two problems are equivalent. Making the change of variable in the monopolist's objective (3.22) yields

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

and the associated monopolist's problem is

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

Let (p_0^*, p_1^*, ϕ^*) be the optimal solution to the above problem and suppose, by contradiction, that $p_0^* > p_1^*$. We will now construct a solution (p_0', p_1', ϕ') that strictly dominates (p_0^*, p_1^*, ϕ^*). Let $p_0' = p_1^* < p_0^*$ and note that $\tau_0(p_0') < \tau_0(p_0^*)$ since

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

Moreover, one can immediately verify that

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

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

$$\int_0^{\phi' \tau_0(p'_0)} e^{-\delta t} \pi_t(p'_0) \, \mathrm{d}t > \int_0^{\phi' \tau_0(p'_0)} e^{-\delta t} \pi_t(p^*_0) \, \mathrm{d}t,$$

since by construction $\hat{q}_t \leq q_1(p^*_0, \phi^*)$ and $\pi_t(p_0)$ is a strictly concave function of p_0 which is maximized at $p_0 = (\bar{\alpha} + \hat{q}_t)/2 \leq p'_0 < p^*_0$. Finally, note that our choice of p'_1 implies that the continuation value of the policy (p'_0, p'_1, ϕ') after $\phi' \tau_0(p'_1)$ is not smaller than the continuation value of the policy (p^*_0, p^*_1, ϕ^*). Thus, we reach the contradiction $\tilde{R}(p'_0, p'_1, \phi') > \tilde{R}(p^*_0, p^*_1, \phi^*)$. It must be $p^*_0 \leq p^*_1$. $\square$

Chapter 4

Dynamic Pricing, Social Influence, and Price Commitment

4.1 Introduction

The overall popularity of many types of products is influenced and amplified by the mass of consumers that purchase in the early release stages and spread the word about the product. The types of products we have in mind are cultural products, like books or movies, for which the urge to adopt a popular product is driven by preference for conformity or sharing with peers. This type of social influence is pervasive in today's economy, where information on the behavior of peers, or even socially distant consumers, is readily available at the touch of a finger. A seller that wants to launch a new product must therefore take the social influence channel into account when devising her optimal pricing strategy.

This chapter studies a model of a seller that is launching a new product to a large market of consumers. There are two periods in the selling season, an introductory period and a mature period. Consumers are on the market at the beginning of the selling season and remain on the market until they purchase the product or the selling season ends. Consumers are influenced by other consumers that purchased before them, i.e., their net utility at the time they are making a decision is increasing in the mass of consumers that have already purchased, we call the intensity of this effect the *level of social influence*. Motivated by the increasing sophistication of pricing in cultural and online markets, we want to study the optimal dynamic pricing policy of the seller. We consider two scenarios: (I) is a scenario in which the seller is able to commit upfront or pre-announce a sequence of prices for the entire selling season; (II) is a scenario in which the seller does not have such a commitment power and thus engages in "responsive pricing", which means that she can change her price at the beginning of the mature period.

Our main result compares the equilibrium profit that the seller attains in scenario (I) to the equilibrium profits she attains in scenario (II) as the level of social influence in the economy varies. We show that when social influence is low, then committing upfront to a sequence of prices always yields higher profits than not doing so, however, when social influence is high the profits in the two scenarios are equal and the value of price commitment for the seller is zero.

In our other results, we derive a complete characterization of the market equilibria in the two scenarios, in terms of seller's optimal pricing decisions, consumers' optimal purchasing decisions and equilibrium profits. Our analytic characterization of the optimal solutions allows us to drive interesting sensitivity analysis conclusions on prices and demands. In particular, we see that when social influence is present it is generally optimal to offer an introductory price in the first period and a higher price in the second period. Finally, we show that the difference between mature and introductory prices is always increasing in the level of social influence. Which means that the seller chooses to use the first period discounted price as a lever to generate hype about her product and then exploit with a higher price in the second period. The seller uses this lever more aggressively the higher the impact of early purchases on late ones.

Before moving to a survey of the literature we want to highlight the normative nature of our results. In this model, we try to specify plausible assumptions on how social influence affects consumers' decisions in markets for cultural products. Then, we derive results that dictate how a seller should optimally price her product, in the absence of other exogenous constraints. In reality, cultural markets have historically been subjected to various types of constraints. Orbach [2004] and Orbach and Einav [2007] show that pricing of motion pictures in movie theaters is affected by regulatory and legal constraints that severely limit the ability of a seller to engage in dynamic pricing and price differentiation. Similar constraints apply to other types of cultural products, such as books in physical bookstores. In these cases, our results may not reflect actual behavior of sellers. However, our model captures some of the most significant features of the effects of dynamic pricing in cultural markets when consumers behave strategically, and all our propositions should be regarded as normative statements. Finally, we note that new markets for cultural products, where books and movies are sold on line in different formats, have less regulatory constraints and one can easily conduct experiments and empirical tests on the value of social influence in cultural markets that may inform future extensions of this work.

4.1.1 Related Literature

The question of conditioning a pricing strategy on some social parameter has received increased attention in the operations research community, Candogan *et al.* [2012b] study static price discrimination when consumers are embedded in a social network, and characterize optimal individual prices as a function of consumers network positions. Our model is a two-period model with intertemporal externalities, where the choices of early consumers affect the payoffs of late consumers, and it is more related to Jing [2011] and Yu *et al.* [2015]. These papers study two-period dynamic pricing models, and although they frame the externality from the introductory period to the next in the form of social learning, they propose reduced form models of information transmission or social learning intensity, that generate similar dynamics to our model of social influence.

Social influence is considered a key driver of consumers' decisions in economics and related disciplines. In this chapter, we model the social influence channel in a similar fashion to Arthur [1989], i.e., we assume that a consumer utility is increasing in the installed base of a product when the consumer makes a purchase. This can be considered a reduced form model of product adoption, via word-of-mouth effects, in the style of Bass [1969], or as a boundedly rational model of network effects in which consumers only take into account past adoptions, see Arthur [1994]. In particular, we do not consider rational expectations over network effects as in Katz and Shapiro [1985]. Our modeling framework is motivated by recent experimental evidence on consumer choice in cultural markets, see Salganik *et al.* [2006], Salganik and Watts [2009], and Moretti [2011].

The question of the value of price commitment has been considered, first in the economics and then in the management science literature. Coase [1972], considers a monopolist that sells a good to a large market of consumers with heterogeneous valuations, and discusses how a monopolist should set the price in order to maximize profit. The optimal price should be decreasing over time, in order to extract more value from consumers with high valuations in the initial periods. Further work by Stokey [1979], Bulow [1982] and Besanko and Winston [1990] study more complex dynamic pricing models and quantify Coase's insights. Recent works in operations research tackle a similar problem with more refined models of price commitment, sometimes paired with capacity commitment, see Aviv *et al.* [2009], Su and Zhang [2008] and Besbes and Lobel [2015].

A recent stream of papers considers problems related to the value of price commitment in the presence of strategic consumers, which is the main question of our work. Aviv and Pazgal [2008] consider both a price commitment and a responsive pricing scenario and compare outcomes to a benchmark in which consumers are myopic. Similar questions are addressed by Liu and Zhang [2013] in the context of a two-firm competition game, and by the works of Cachon and Swinney [2009] and Papanastasiou and Savva [forthcoming], however, none of these works studies the value of price commitment in the presence of social influence. Our model is more related to the latter of the above works, who studies a two-period dynamic pricing model where the main results depend on an exogenous social learning intensity parameter. We study a different model, with an exogenous level of social influence, but we believe that our main result sheds more light on their contrasting result on the value of price commitment.

The rest of the chapter is organized as follows. Section 4.2 contains our model and formally introduces our measure of social influence and the seller's problem. In Section 4.3 we fully characterize the optimal prices of the seller when she can commit to a pricing policy upfront, and the equilibrium decisions of consumer. In Section 4.4 we characterize optimal prices and consumers' decisions when the seller does not have commitment power. Section 4.5 presents our main result on the value of price commitment for the seller and Section 4.6 concludes. All proofs are presented in the Appendix.

4.2 Model

Monopolist The market features a monopolist launching a new product, which is made available for purchase in two consecutive time periods $t = 1, 2$, at unit price p_t. We will consider two possible scenarios.

(I) Price Commitment. The monopolist commits to (p_1, p_2) at the beginning of period 1.

(II) Responsive Pricing. The monopolist does not have commitment power, she sets p_1 at the beginning of period 1 and she sets p_2 at the beginning of period 2.

Consumers There is a population of consumers whose size is normalized to one and consumers are indexed by $i \in [0, 1]$. The value of the product for a consumer i has two components: (i) a private value component or individual willingness to pay, and (ii) a social influence component.

Consumers are heterogeneous in their *private value* α^i for the product, $\{\alpha^i\}_{i\in[0,1]}$ are i.i.d. random draws from a known distribution $U[0, \bar{\alpha}]$. On top of their private value, a consumer product valuation is affected by the mass of consumers that have already bought the product at the time he chooses to purchase. In particular, the *social influence* effect goes from one period to the next, i.e., early consumers influence late consumers.

The net (undiscounted) utility of a consumer i that purchases the product at time t is

$$u_t^i = \alpha^i + \beta s_{t-1} - p_t, \qquad t = 1, 2 \ , \tag{4.1}$$

where $s_t \in [0, 1]$ denotes the fraction of consumers that purchase the product in period t, and since the new product is first available in period 1, we assume that $s_0 = 0$. The parameter $\beta > 0$ measures the absolute intensity of social influence, given this parameter and the commonly known distribution of private valuations, we define

$$\gamma = \frac{\partial u}{\partial s} \Big/ \sup\{\alpha_i\} = \frac{\beta}{\bar{\alpha}}, \tag{4.2}$$

which measures the intensity of social influence relative to consumers' maximum private valuation.

This is our metric for the *level of social influence*, and the main driver of our results.

Each consumer purchases at most one unit of the product, thus if a consumer decides to purchase in period 1 he will not be on the market in period 2. Moreover, we assume that a consumer that purchases in period t does not extract any utility from the product in period $t+1$. This assumption captures diminishing returns of consumption, which are common for cultural products like books or movies (see Varian [2000] and Rao [2015]).

Purchasing Decision Consumers are impatient and they discount their payoffs in period 2 by $\delta \in (0,1)$. Their purchasing decision can be spelled out as follows: (i) consumer i purchases the product in period 1 if his utility in the current period is higher than both the utility from not purchasing and the discounted utility in period 2, i.e., $u_1^i \geq \max\{0, \delta u_2^i\}$; (ii) if consumer i is still on the market in period 2, he purchases if $u_2^i \geq 0$. Equivalently, replacing consumer utilities (4.1) into the above inequalities and recalling our assumption $s_0 = 0$, we can express the above decision rule in terms of the consumers' private valuations as follows:

(i) consumer i purchases in period 1 if

$$\alpha_i \geq \max\left\{p_1, \frac{p_1 - \delta p_2 + \delta\beta s_1}{1-\delta}\right\} = \lambda_1; \tag{4.3}$$

(ii) if consumer i is still on the market in period 2, he purchases if

$$\alpha_i \geq p_2 - \beta s_1 = \lambda_2. \tag{4.4}$$

The key feature of this decision rule is its simple threshold form, there are two thresholds λ_1 and λ_2 that consumers use to make decisions based on their private valuations. In equilibrium, the thresholds are determined by solving a fixed point equation.

Monopolist Objective The monopolist is also impatient and discounts future profits by δ, i.e., we assume that consumers and seller have the same discount factor. The seller is interested is setting a sequence of prices (p_1, p_2) in order to maximize her discounted expected profit

$$\Pi(p_1, p_2) = p_1 D_1(p_1, p_2) + \delta p_2 D_2(p_1, p_2), \tag{4.5}$$

where $D_1(p_1, p_2)$ and $D_2(p_1, p_2)$ are the expected demands in periods 1 and 2 at prices (p_1, p_2). The monopolist objective is clearly the same in the two scenarios, whether she has commitment power or not, however the optimization problem she faces is different, as it can be its solution, which we characterize in the two following sections.

4.3 Price Commitment

In the scenario where the seller has price commitment, her problem is to optimally choose (p_1, p_2) in order to maximize (4.5) at the beginning of $t = 1$. Given (p_1, p_2), consumers compute the optimal thresholds (λ_1, λ_2) consistent with decision rules (4.3) and (4.4). In this section, we present our characterization of the optimal prices and thresholds that are a solution to the game described above. We also show how the equilibrium profits change as a function of the level of social influence (γ) and other relevant model parameters. Proposition 4.3.1 presents the results for the case in which γ is small.

Proposition 4.3.1. *If $\gamma < 2$ the equilibrium prices are such that $p_1^* \leq p_2^*$ and the equilibrium thresholds are such that*

$$1 > \lambda_1^* > \lambda_2^* > 0.$$

Moreover, the equilibrium profit for the seller is

$$\Pi^* = \frac{1 - \delta + \gamma\delta}{4 - (2 - \gamma)^2 \delta} \bar{\alpha}. \tag{4.6}$$

The above proposition states that, when $\gamma < 2$, it is always optimal for the seller to set an introductory price in period 1 that is lower than the price she sets for period 2. This means that, even in the presence of low levels of social influence, when the monopolist has commitment power she finds it optimal to set a higher price in the mature period and extract consumers' boosted valuations.

Moreover, Proposition 4.3.1 shows that when γ is low the seller always finds it optimal to price in such a way that she induces positive sales in both periods ($\lambda_1^* > \lambda_2^*$) and not all the potential demand is fulfilled ($\lambda_2^* > 0$). In particular, the seller prefers to exclude consumers with lower valuations by charging higher prices and sell only to consumers with higher private valuations, thus extracting more surplus from each consumer that purchases.

The equilibrium profit of the seller is increasing in the level of social influence γ. To see why this is the case note that a higher γ induces higher valuations in period 2 while leaving them unchanged in period 1. Moreover, the equilibrium profit is also increasing in the patience level δ. This is a consequence of the fact that, as both seller and consumers become more patient part of the sales shift to period 2, where the effect of social influence induces higher valuations for consumers. Note also that the equilibrium profit that the seller can extract is proportional to the consumers' maximum willingness to pay.

Our next result characterizes the equilibrium under price commitment when the social influence level is high.

Proposition 4.3.2. *If $\gamma \geq 2$ the equilibrium prices are such that $p_1^* \leq p_2^*$ and the equilibrium thresholds are such that*

$$\bar{\alpha}/2 = \lambda_1^* > \lambda_2^* = 0.$$

Moreover, the equilibrium profit for the seller is

$$\Pi^* = \bar{\alpha}(1 - \delta + \gamma\delta)/4. \tag{4.7}$$

Proposition 4.3.2 shows that when the seller has commitment power and $\gamma \geq 2$, it is again optimal to set a lower introductory price in period 1 and a higher price in period 2. In this case, since the level of social influence is high the seller has even more incentives to price low in period 1 to generate more demand in the current period and then price high in period 2 to capitalize on inflated valuations.

Moreover, the seller always sets prices so that all consumers buy by the end of period 2, i.e. she engages in exhaustive sales. In particular, when γ is high enough the seller finds it optimal to split the market, by selling to the half of the consumers with higher valuations ($\alpha_i > \bar{\alpha}/2$) in the introductory period and extracting inflated valuations from all remaining consumers in the mature period. Finally, note that the optimal profit in this case, (4.7), is always higher than the case when γ is low, (4.6), and it is still increasing in both γ and δ.

Before concluding this section, we remark its most important results. Propositions 4.3.1 and 4.3.2 together establish that, under price commitment, when $\gamma < 2$ the monopolist finds it optimal to exclude some consumers from the sale, while she sells to all consumers when $\gamma \geq 2$. Moreover, the seller always finds it optimal to set an increasing price path ($p_1^* \leq p_1^*$), and the equilibrium profit is continously increasing in the level of social influence.

4.3.1 Sensitivity Analysis of Equilibrium Prices and Demands

In this subsection we present and discuss a comprehensive result that summarizes the sensitivity of equilibrium prices and consumers' demands to the relevant model parameters, for all levels of social influence. In particular we are interested in how the above quantities vary with the social influence and patience levels. We have the following result.

Proposition 4.3.3. *Let $(p_1^*, p_2^*, \lambda_1^*, \lambda_2^*)$ be the equilibrium prices and thresholds under price commitment and let (D_1^*, D_2^*) be the corresponding equilibrium demands in periods 1 and 2. Then,*

(a) For all $\gamma > 0$, the optimal prices have the following properties:

$$(a.i) \quad \frac{\partial p_1^*}{\partial \gamma} \leq 0 \quad and \quad \frac{\partial p_2^*}{\partial \gamma} \geq 0; \qquad (a.ii) \quad \frac{\partial}{\partial \delta}(p_2^* - p_1^*) \geq 0.$$

(b) For all $\gamma > 0$, the equilibrium demands have the following properties:

$$(b.i) \quad D_1^* \geq D_2^*; \qquad (b.ii) \quad \frac{\partial D_1^*}{\partial \delta} \leq 0 \quad and \quad \frac{\partial D_2^*}{\partial \delta} \geq 0; \qquad (b.iii) \quad \frac{\partial}{\partial \gamma}(D_1^* - D_2^*) \leq 0.$$

Part (a) of the above proposition establishes two properties of the equilibrium prices. Property (a.i) says that the introductory price in period 1 is always decreasing in the level of social influence and the equilibrium price in period 2 is always increasing, clearly this implies that the difference between the two prices, or the introductory mark-down, is increasing in the intensity of social influence. Property (a.ii) says that the difference between the optimal price in period 2 and the optimal price in period 1 is increasing in the patience level, which is a consequence of the fact that with higher patience, the monopolist is more willing to sacrifice first period revenues to induce higher valuations in the second period and the consumers discount second period utility by a smaller factor.

Part (b) of the above proposition presents three properties of the equilibrium demands. Property (b.i) says that demand is always higher in the introductory period, this mirrors the fact that it is always optimal for the seller to set an increasing price path. Property (b.ii) states that the equilibrium demand in the introductory period is decreasing in the discount factor and the equilibrium demand in the mature period is increasing in the discount factor. Property (b.iii) states that the difference between period 1 demand and period 2 demand is non-increasing in the level of social influence. Note that when $\gamma < 2$ this difference is decreasing and when $\gamma \geq 2$ the demand is always split equally between the two periods.

4.4 Responsive Pricing

In the responsive pricing scenario, where there is no commitment power, the seller's expected profit in period 2 is $\pi_2 = p_2 D_2$ and her expected discounted profit in at the beginning of period 1 is $\pi_1 = p_1 D_1 + \delta\pi_2$. The game unfolds as follows: the seller chooses p_1 at the beginning of period 1 in order to maximize π_1, then observes how many consumers have purchased in the introductory period (λ_1) and chooses p_2 in the beginning of period 2 in order to maximize π_2, finally the remaining consumers make their optimal purchase decision in period 2. In this section, we characterize the equilibrium prices and thresholds in the subgame perfect equilibrium of the game described above, which we solve by backward induction. Moreover, we discuss how equilibrium profits are affected by social influence and patience levels.

The next proposition presents the equilibrium characterization for low levels of social influence, in this case equilibrium prices require a more detailed discussion, that we will present in Proposition 4.4.3. We have the following result.

Proposition 4.4.1. *If* $\gamma < \dfrac{5\delta - 2 + \sqrt{4 - 4\delta + 17\delta^2}}{4\delta}$ *the equilibrium thresholds are such that*

$$1 > \lambda_1^* > \lambda_2^* > 0.$$

Moreover, the equilibrium profit for the seller is

$$\Pi^* = \frac{(2-\delta)^2 + 4\gamma\delta}{2[(8-6\delta) + 8\gamma\delta - 2\gamma^2\delta]}\bar{\alpha}. \tag{4.8}$$

Proposition 4.4.1 states that, when γ is small enough, in equilibrium the seller sales in both periods ($\lambda_1^* > \lambda_2^*$) and not all the potential demand is fulfilled ($\lambda_2^* > 0$). The seller excludes consumers with lower valuations from the sale, and the qualitative insights are the same as in the case with price commitment. However, quantitatively the equilibrium prices, thresholds and profits are different. Note that the equilibrium profits are still increasing in all the parameters of interest, i.e., the social influence and patience levels and the maximum private valuation of consumers.

Our next result presents the equilibrium characterization for the cases in which the level of social influence is high.

Proposition 4.4.2. *If* $\gamma \geq \dfrac{5\delta - 2 + \sqrt{4 - 4\delta + 17\delta^2}}{4\delta}$ *the equilibrium prices are such that* $p_1^* \leq p_2^*$ *and the equilibrium thresholds are such that*

$$\bar{\alpha}/2 = \lambda_1^* > \lambda_2^* = 0.$$

Moreover, the equilibrium profit for the seller is

$$\Pi^* = \bar{\alpha}(1 - \delta + \gamma\delta)/4. \tag{4.9}$$

The above proposition states that, when γ is higher than a given threshold, it is always optimal for the seller to set a lower introductory price in period 1 and a higher price in period 2. This is in line with the optimal pricing strategy with price commitment and different from the responsive pricing with low social influence, as we will see in Proposition 4.4.3. Moreover, the seller sells to all consumers by the end of period 2 and sets prices in such a way that the demand splits evenly between period 1 and period 2.

Most importantly, note that when the social influence level is high enough, all the equilibrium quantities are both qualitatively and quantitatively the same as in the price commitment case. The intuition behind this result is that, when the equilibrium sales become exhaustive, the seller is able to sustain the same level of equilibrium prices that he would be willing to set if she had commitment power.

In the next subsection, we provide a comprehensive characterization of how equilibrium prices change as a function of the level of social influence in the responsive pricing scenario.

4.4.1 Equilibrium Prices and Social Influence

In Section 4.3, we saw that under price commitment the equilibrium price in period 1 is always lower than the equilibrium price period 2. We already anticipated that results can be much different in the absence of commitment power, and our next result establishes that with responsive pricing an increasing price schedule is not always optimal.

Proposition 4.4.3. *The seller's optimal prices (p_1^*, p_2^*) have the following properties.*

(i) There exist a threshold $0 < \tilde{\gamma} < \dfrac{5\delta - 2 + \sqrt{4 - 4\delta + 17\delta^2}}{4\delta}$ *such that*

$$p_1^* \geq p_2^* \quad if \quad \gamma \leq \tilde{\gamma} \qquad and \qquad p_1^* \leq p_2^* \quad if \quad \gamma > \tilde{\gamma}.$$

Moreover, $\tilde{\gamma} \to 0$ *as* $\delta \to 1$.

(ii) $\dfrac{\partial p_1^*}{\partial \gamma} \leq 0$ *and* $\dfrac{\partial p_2^*}{\partial \gamma} \geq 0$ *for all* $\gamma > 0$.

Property (i) of the above proposition focuses on the responsive pricing scenario with low levels of social influence, i.e., the equilibrium characterized in Proposition 4.4.1. This property states that there exist a threshold $\tilde{\gamma}$ such that when the level of social influence is lower than this threshold, then the equilibrium price in period 1 is higher than the equilibrium price in period 2. Otherwise, for high levels of social influence we get an increasing price scheme, which is qualitatively the same as in the price commitment scenario. Moreover, note that $\tilde{\gamma}$ approaches zero as δ approaches 1. Property (ii) of the above proposition states that the introductory price is always decreasing in the level of social influence and the price in the second period is always increasing. This holds for all γ, even when the level of social influence is very low and $p_1^* \geq p_2^*$.

4.5 Value of Price Commitment

In both scenarios analyzed above, the price commitment in Section 4.3 and the responsive pricing in Section 4.4, the seller was able to chose different prices in periods 1 and 2, the difference between the two scenarios is that in the first one we endow the seller with some commitment power. In this section, we ask what is the value of having commitment power to the seller. Our main result characterizes the *value of price commitment* as a function of the level of social influence. The value of price commitment is calculated as the equilibrium profit under price commitment minus the equilibrium profit under responsive pricing.

Let Π^*_{PC} be the equilibrium profit under price commitment, characterized in Propositions 4.3.1 and 4.3.2, and let Π^*_{RP} be the equilibrium profit under responsive pricing, characterized in Propositions 4.4.1 and 4.4.2. We are now ready to state our main result.

Theorem 4.5.1. *The value of price commitment for the seller is characterized as follows.*

(a) If $\gamma < 2$, the value of price commitment is strictly positive, i.e. $\Pi^*_{PC} - \Pi^*_{RP} > 0$.

(b) If $\gamma \geq 2$, the value of price commitment is zero, i.e. $\Pi^*_{PC} - \Pi^*_{RP} = 0$.

Theorem 4.5.1 establishes that, in a market with many heterogeneous consumers and social influence, commitment power can never decrease the seller's equilibrium profit. This is in line with most works in the literature that find that for dynamic monopoly pricing models having a commitment mechanism is generally valuable.

Statement (a) of the theorem states that when the level of social influence is low, the seller always achieves strictly higher profit under price commitment than under responsive pricing. More importantly, Statement (b) states that if the level of social influence is high enough, the ability to commit to a pricing schedule does not give the seller any additional profit with respect to the responsive pricing case.

We now provide some further explanations that allow to better appreciate why the above theorem holds. Recall that in our model the total demand is normalized to one, and in particular it is finite. When the level of social influence is high enough, it is optimal to set the price for period 2 in such a way that all demand is exhausted, and this is true irrespective of whether the seller has commitment power. Thus, in both price commitment and responsive pricing scenarios, the optimal period 2 price is the boundary solution that corresponds to exhaustive sales. Consequently, the period 1 prices, which are pinned down by the best responses for period 1, are equivalent, and the seller attains the same equilibrium profits in the two scenarios.

When the level of social influence is low, a seller with commitment power finds it optimal to set prices in such a way that the consumers with low valuations are excluded from the sale, i.e., the optimal price in period 2 is always an interior solution. However, when the seller does not have commitment power she finds it optimal to set a lower price in period 2. This hurts the seller in two ways, by decreasing period 1 profits, since a higher fraction of consumers purchases in period 2 at a lower price, and also by decreasing period 2 valuations, since the social influence component of utility is reduced by less consumers purchasing in period 1. As a result, in the absence of price commitment the seller makes strictly lower profits when the level of social influence is low enough.

4.6 Conclusion

This chapter studies a two-period pricing model with a seller that launches a new product to a large market of consumers and she is able to set different prices for each of the two periods. Consumers are heterogeneous in their private valuations, and their overall valuation for the product is also affected by how many consumers have already purchased, through a positive social influence channel with exogenous intensity level. We study two scenarios, one in which the seller has commitment power and one in which she does not. The optimal purchase decision of consumers always takes a simple threshold form. We characterize the equilibrium in the two scenarios in terms of equilibrium prices and thresholds and then compare the results.

In the presence of price commitment, we show that it is always optimal for the seller to set an increasing pricing schedule. We also show that it is always optimal to exclude some consumers from sale when the level of social influence is low and to sell to everyone when the level of social influence is high. In the responsive pricing scenario we have a similar result on consumers' equilibrium decisions but different results on pricing. In particular, the seller still excludes some consumers from sale when the level of social influence is below a given threshold, and sells to everyone when it is above. In the latter case the seller finds it optimal to set an increasing price path, however, when there is no price commitment and the level of social influence is very low the seller sets a decreasing price path.

Our main result concerns the characterization of the value of price commitment in this model as a function of the social influence level. We find that, when the level of social influence is below a given threshold, then price commitment is always strictly valuable. However, for high levels of social influence, a monopolist that engages in responsive pricing does equally well than one that has access to commitment mechanism.

4.7 Proofs

Recall that tail distribution of a $U[0, \bar{\alpha}]$ random variable is

$$\overline{F}(x) = \begin{cases} 1 & \text{if} \quad x \leq 0 \\ 1 - x/\bar{\alpha} & \text{if} \quad x \in [0, \bar{\alpha}] \\ 0 & \text{if} \quad x \geq \bar{\alpha} \end{cases} \cdot \tag{4.10}$$

Note that given consumers' threshold purchasing decision, the expected demands in period 1 and 2 can be written as

$$D_1 = \overline{F}(\lambda_1) \quad \text{and} \quad D_2 = \left[\overline{F}(\lambda_2) - \overline{F}(\lambda_1)\right]^+,$$

and the expected discounted profits (4.5) can be rewritten as

$$\Pi(p_1, p_2) = p_1 \overline{F}(\lambda_1) + \delta p_2 \left[\overline{F}(\lambda_2) - \overline{F}(\lambda_1)\right]^+. \tag{4.11}$$

The following lemma characterizes consumers' equilibrium thresholds as a function of the seller's prices (p_1, p_2) and we will use it for the proofs of propositions 4.3.1 and 4.3.2.

Lemma 4.7.1. *Let (p_1, p_2) be the prices set by the seller for period 1 and period 2, consumers' equilibrium thresholds are determined as follows.*

(a) If $p_1 \leq \dfrac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha}$, then

$$\lambda_1 = \frac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha} \quad \textit{and} \quad \lambda_2 = \frac{\beta p_1 - \bar{\alpha}(1-\delta)(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}.$$

(b) If $p_1 > \dfrac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha}$, then

$$\lambda_1 = p_1 \quad \textit{and} \quad \lambda_2 = p_2 - \beta(1 - p_1/\bar{\alpha}).$$

Proof. In equilibrium all consumers make their optimal purchase decision in period 1 according to

(4.3), since the private valuations are uniformly distributed it must be that $s_1 = 1 - \lambda_1/\bar{\alpha}$. Replace s_1 in (4.3) and suppose that

$$p_1 \leq \frac{p_1 - \delta p_2 + \delta\beta(1 - \lambda_1/\bar{\alpha})}{1-\delta},$$

then the equilibrium threshold in period 1 must solve

$$\lambda_1 = \frac{p_1 - \delta p_2 + \delta\beta(1 - \lambda_1/\bar{\alpha})}{1-\delta},$$

which implies that

$$\lambda_1 = \frac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha}. \tag{4.12}$$

Proof of part (a): If $p_1 \leq \frac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha}$ then it follows from (4.3) that (4.12) is the equilibrium threshold in period 1, and the equilibrium threshold in period 2 is obtained replacing $s_1 = 1 - \lambda_1/\bar{\alpha}$ into (4.4) as follows

$$\lambda_2 = p_2 - \beta\left(1 - \lambda_1/\bar{\alpha}\right) = \frac{\beta p_1 - \bar{\alpha}(1-\delta)(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta},$$

Proof of part (b): If $p_1 > \frac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha}$ it follows from (4.3) that $\lambda_1 = p_1$ and replacing into (4.4) yields $\lambda_2 = p_2 - \beta(1 - p_1/\bar{\alpha})$. $\square$

Proof of Proposition 4.3.1

Using the tail distribution $\overline{F}$ defined in (4.10) we can write the seller's problem as

$$\begin{aligned} \max_{p_1, p_2} \quad & p_1(1 - \lambda_1/\bar{\alpha}) + \delta p_2(\lambda_1 - \lambda_2)/\bar{\alpha} \\ s.t. \quad & \lambda_1 - \lambda_2 \geq 0 \\ & \bar{\alpha} \geq \lambda_1, \lambda_2 \geq 0 \\ & p_1, p_2 \geq 0. \end{aligned} \tag{4.13}$$

Suppose that $p_1 \leq \dfrac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha}$, we will later verify that this is always true in equilibrium. From Lemma 4.7.1 it follows that the optimal consumers' threshold for period 1 is

$$\lambda_1 = \frac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha}, \tag{4.14}$$

and the optimal threshold for period 2 is

$$\lambda_2 = \frac{\beta p_1 - \bar{\alpha}(1-\delta)(\beta - p_2)}{\bar{\alpha}(1-\delta) + \beta\delta}. \tag{4.15}$$

Replacing (4.14) and (4.15) into the objective function of Problem (4.13) yields

$$\Pi = \frac{-\bar{\alpha}p_1^2 + \bar{\alpha}^2(1-\delta)p_1 + (2\bar{\alpha} - \beta)\delta p_1 p_2 + \bar{\alpha}\beta\delta p_2 - \bar{\alpha}\delta p_2^2}{\bar{\alpha}[\bar{\alpha}(1-\delta) + \beta\delta]}, \tag{4.16}$$

and the associated Hessian is

$$\mathrm{H} = \frac{1}{\bar{\alpha}[\bar{\alpha}(1-\delta) + \beta\delta]} \begin{bmatrix} -2\bar{\alpha} & (2\bar{\alpha} - \beta)\delta \\ (2\bar{\alpha} - \beta)\delta & -2\bar{\alpha}\delta \end{bmatrix}.$$

Note that Π is jointly concave in (p_1, p_2) if and only if $\det(\mathrm{H}) \geq 0$, i.e.

$$4\bar{\alpha}^2 - (2\bar{\alpha} - \beta)^2\delta \geq 0,$$

which holds when $\beta/\bar{\alpha} \leq 2\left(1 + 1/\sqrt{\delta}\right)$ and it is always true in this case since $\gamma = \beta/\alpha < 2$. Thus, the optimal prices (p_1^*, p_2^*) must solve the first-order conditions $\dfrac{\partial \Pi}{\partial p_1} = 0$ and $\dfrac{\partial \Pi}{\partial p_2} = 0$, which are given by

$$\frac{-2\bar{\alpha}p_1 + \bar{\alpha}^2(1-\delta) + (2\bar{\alpha} - \beta)\delta p_2}{\bar{\alpha}[\bar{\alpha}(1-\delta) + \beta\delta]} = 0 \quad \text{and} \quad \frac{(2\bar{\alpha} - \beta)\delta p_1 + \bar{\alpha}\beta\delta - 2\bar{\alpha}\delta p_2}{\bar{\alpha}[\bar{\alpha}(1-\delta) + \beta\delta]} = 0.$$

It follows that

$$p_1^* = \frac{2\bar{\alpha}^2(1-\delta) + 2\bar{\alpha}\beta\delta - \beta^2\delta}{4\bar{\alpha}^2 - (2\bar{\alpha} - \beta)\delta}\bar{\alpha} = \frac{2(1-\delta) + 2\gamma\delta - \gamma^2\delta}{4 - (2-\gamma)^2\delta}\bar{\alpha}, \tag{4.17}$$

and

$$p_2^* = \frac{2\bar{\alpha}^2(1-\delta) + \bar{\alpha}\beta(1+\delta)}{4\bar{\alpha}^2 - (2\bar{\alpha} - \beta)\delta}\bar{\alpha} = \frac{2(1-\delta) + \gamma(1+\delta)}{4 - (2-\gamma)^2\delta}\bar{\alpha}. \tag{4.18}$$

Replacing (p_1^*, p_2^*) into (4.14) and (4.15) and simplifying yields

$$\lambda_1^* = \frac{2(1-\delta) + 3\gamma\delta - \gamma^2\delta}{4 - (2-\gamma)^2\delta}\bar{\alpha} \quad \text{and} \quad \lambda_2^* = \frac{(2-\gamma)[(1-\delta) + \gamma\delta]}{4 - (2-\gamma)^2\delta}\bar{\alpha}. \tag{4.19}$$

Moreover, it is easy to verify our assumption that $p_1^* \leq \frac{p_1^* + \delta(\beta - p_2^*)}{\bar{\alpha}(1-\delta) + \beta\delta}\bar{\alpha}$, note that the right-hand side is equal to λ_1^* and clearly $\lambda_1^* \geq p_1^*$, and it is easy to verify that all the constraints of Problem (4.13) hold under the condition $\gamma < 2$. In particular, note that $\lambda_2^* > 0$ if and only if $\gamma < 2$. To see that $\lambda_1^* > \lambda_2^*$ note that

$$\lambda_1^* - \lambda_2^* = \frac{\gamma}{4 - (2-\gamma)^2\delta}\bar{\alpha} > 0.$$

To see that $p_1^* \leq p_2^*$ note that, subtracting (4.17) from (4.18) yields

$$p_2^* - p_1^* = \frac{\gamma(1 - \delta + \gamma\delta)}{4 - (2-\gamma)^2\delta}\bar{\alpha}, \tag{4.20}$$

which is always positive. Finally, replacing (p_1^*, p_2^*) into (4.16) and simplifying yields

$$\Pi^* = \frac{(1-\delta) + \gamma\delta}{4 - (2-\gamma)^2\delta}\bar{\alpha}.$$

$\square$

Proof of Proposition 4.3.2

Consider the optimal solution of Problem (4.13) characterized in Proposition 4.3.1. In particular, note that the optimal consumers' threshold for period 2 becomes negative when $\gamma \geq 2$. In this case we need to solve the seller's problem with the constraint that $\lambda_2 \leq 0$, which using the tail distribution (4.10) can be written as

$$\begin{aligned} \max_{p_1,p_2} \quad & p_1(1-\lambda_1/\bar{\alpha}) + \delta p_2 \lambda_1/\bar{\alpha} \\ s.t. \quad & \lambda_1 - \lambda_2 \geq 0 \\ & \bar{\alpha} \geq \lambda_1 \geq 0 \\ & \lambda_2 \leq 0 \\ & p_1, p_2 \geq 0. \end{aligned} \tag{4.21}$$

Suppose that $p_1 \leq \dfrac{p_1 + \delta(\beta - p_2)}{\bar{\alpha}(1-\delta)+\beta\delta}\bar{\alpha}$, we will later verify that this is always true in equilibrium. From Proposition 4.7.1 it follows that the optimal consumers' thresholds for period 1 and period 2 are given by (4.14) and (4.15) respectively. Replacing (4.14) and (4.15) into the objective function of Problem (4.21) yields

$$\Pi = \frac{-p_1^2 + \bar{\alpha}(1-\delta)p_1 + 2\delta p_1 p_2 + \beta\delta^2 p_2 - \delta^2 p_2^2}{\bar{\alpha}(1-\delta)+\beta\delta}, \tag{4.22}$$

and the associated Hessian is given by

$$\mathrm{H} = \frac{1}{\bar{\alpha}(1-\delta)+\beta\delta}\begin{bmatrix} -2 & 2\delta \\ 2\delta & -2\delta^2 \end{bmatrix}.$$

Clearly the profit function is concave in p_2, solving the first-order condition $\dfrac{\partial \Pi}{\partial p_2} = 0$ yields the best response

$$\hat{p}_2(p_1) = \frac{\beta\delta + 2p_1}{2\delta}.$$

The constraint $\lambda_2 \leq 0$ imposes the following upper bound on p_2, i.e.

$$\bar{p}_2(p_1) = \frac{\beta[\bar{\alpha}(1-\delta) - 2p_1]}{\bar{\alpha}(1-\delta)},$$

note that $\hat{p}_2(p_1) \leq \bar{p}_2(p_1)$ if and only if

$$p_1 \leq \frac{\bar{\alpha}\beta(1-\delta)\delta}{2\bar{\alpha}(1-\delta) + 2\beta\delta} = \bar{p}_1.$$

If $p_1 \leq \bar{p}_1$, replacing $\hat{p}_2(p_1)$ into (4.22) yields

$$\Pi(p_1, \hat{p}_2(p_1)) = \frac{\beta^2\delta^2}{4[\bar{\alpha}(1-\delta) + \beta\delta]} + p_1,$$

thus clearly it is optimal to set $p_1 = \bar{p}_1$. Then it must be that $p_1 \geq \bar{p}_1$, which implies $\hat{p}_2(p_1) \geq \bar{p}_2(p_1)$ and the best response $\hat{p}_2(p_1)$ is not feasible. Replacing $\bar{p}_2(p_1)$ into (4.22) yields

$$\Pi(p_1, \bar{p}_2(p_1)) = \frac{[-p_1^2 + \bar{\alpha}(1-\delta)p_1][\bar{\alpha}(1-\delta) + \beta\delta]}{\alpha^2(1-\delta)^2}$$

which is a concave quadratic in p_1, thus solving the first-order condition to get p_1^* and then replacing p_1^* into $\bar{p}_2(p_1)$ to get p_2^* yields

$$p_1^* = \bar{\alpha}(1-\delta)/2 \quad \text{and} \quad p_2^* = \beta/2 = \bar{\alpha}\gamma/2 \tag{4.23}$$

and replacing (p_1^*, p_2^*) into (4.14) and (4.15) yields $\lambda_1^* = \bar{\alpha}/2 > p_1^*$, clearly $\lambda_2^* = 0$. It is easy to verify our assumption that $p_1^* \leq \lambda_1^*$, as well as that all constraints hold at the optimal solution. Further note that clearly $\lambda_1^* > \lambda_2^* = 0$, and

$$p_2^* - p_1^* = \bar{\alpha}(\gamma - 1 + \delta)/2, \tag{4.24}$$

which is always positive. To complete the proof note that replacing (p_1^*, p_2^*) into (4.22) yields $\Pi^* = \bar{\alpha}(1 - \delta + \gamma\delta)/4$. □

Proof of Proposition 4.3.3

Proof of Part (a): For property $(a.i)$, note that when $\gamma < 2$, differentiating p_1^* and p_2^* with respect to γ yields

$$\frac{\partial p_1^*}{\partial \gamma} = -\frac{2\bar{\alpha}\gamma\delta(2-2\delta+\gamma\delta)}{[4-(2-\gamma)^2\delta]^2} \leq 0 \quad \text{and} \quad \frac{\partial p_2^*}{\partial \gamma} = \frac{4-8\delta+4\delta^2+4\delta(1-\delta)\gamma+\delta(1+\delta)\gamma^2}{[4-(2-\gamma)^2\delta]^2}\bar{\alpha} \geq 0,$$

where the first inequality is immediate, and the second inequality follows from the fact that the numerator of $\partial p_2^*/\partial\delta$ is a convex quadratic in γ with determinant $\delta(-1+2\delta-\delta^2) \leq 0$ for all $\delta \in (0,1)$. When $\gamma \geq 2$, differentiating p_1^* and p_2^* with respect to γ yields

$$\frac{\partial p_1^*}{\partial \gamma} = 0 \quad \text{and} \quad \frac{\partial p_2^*}{\partial \gamma} = \bar{\alpha}/2 \geq 0.$$

For property $(a.ii)$, note that when $\gamma < 2$, differentiating (4.20) with respect to δ yields

$$\frac{\partial}{\partial \delta}(p_2^* - p_1^*) = \frac{\gamma^3\bar{\alpha}}{[4-(2-\gamma)^2\delta]^2} \geq 0.$$

When $\gamma \geq 2$, differentiating (4.24) with respect to δ yields $\dfrac{\partial}{\partial \delta}(p_2^* - p_1^*) = \bar{\alpha}/2 > 0$. Which completes the proof of Part (a).

Proof of Part (b): For property $(b.i)$, note that the equilibrium demands are given by $D_1^* = 1-\lambda_1^*/\bar{\alpha}$ and $D_2^* = (\lambda_1^* - \lambda_2^*)/\bar{\alpha}$. When $\gamma < 2$, using the optimal thresholds λ_1^* and λ_2^* from (4.19), we can compute the difference between period 1 and period 2 demands as

$$D_1^* - D_2^* = 1 - 2\lambda_1^*/\bar{\alpha} + \lambda_2^*/\bar{\alpha} = \frac{(2-\gamma)(1-\delta)}{4-(2-\gamma)^2\delta}, \tag{4.25}$$

which is always positive in this case. When $\gamma \geq 2$ the optimal thresholds are given by $\lambda_1^* = \bar{\alpha}/2$ and $\lambda_2^* = 0$, replacing them into the above expressions for D_1^* and D_2^* yields $D_1^* = D_2^* = 1/2$, and thus $D_1^* - D_2^* = 0$. For property $(b.ii)$, note that when $\gamma < 2$ we have

$$\frac{\partial D_1^*}{\partial \delta} = \frac{\partial}{\partial \delta}(1 - \lambda_1^*/\bar{\alpha}) = \frac{2\gamma(\gamma-2)}{[4-(2-\gamma)^2\delta]^2},$$

which is always negative in this case, and

$$\frac{\partial D_1^*}{\partial \delta} = \frac{\partial}{\partial \delta}(\lambda_1^* - \lambda_2^*)/\bar{\alpha} = \frac{\gamma(\gamma-2)^2}{[4-(2-\gamma)^2\delta]^2},$$

which is always positive. When $\gamma \geq 2$ we have $\dfrac{\partial D_1^*}{\partial \delta} = \dfrac{\partial D_2^*}{\partial \delta} = 0$. Finally, for property *(b.iii)*, note that when $\gamma < 2$ differentiating (4.25) with respect to γ yields

$$\frac{\partial}{\partial \gamma}(D_1^* - D_2^*) = -\frac{(1-\delta)[4+(2-\gamma)^2\delta]}{[4-(2-\gamma)^2\delta]^2} \leq 0,$$

and when $\gamma \geq 2$ we have $\dfrac{\partial}{\partial \gamma}(D_1^* - D_2^*) = 0$. □

Proof of Proposition 4.4.1

We solve the game by backward induction. Using the tail distribution $\overline{F}$ defined in (4.10), the seller's problem in period 2 can be written as

$$\begin{aligned} \max_{p_2} \quad & p_2(\lambda_1 - \lambda_2)/\bar{\alpha} \\ s.t. \quad & \lambda_1 - \lambda_2 \geq 0 \\ & \bar{\alpha} \geq \lambda_1, \lambda_2 \geq 0 \\ & p_2 \geq 0. \end{aligned} \tag{4.26}$$

The optimal consumers' threshold in period 2 is given by $\lambda_2 = p_2 - \beta(1 - \lambda_1/\bar{\alpha})$, and replacing it into the objective of Problem (4.26) yields

$$\pi_2 = \frac{-p_2^2 + [\beta + (1-\beta/\bar{\alpha})\lambda_1]p_2}{\bar{\alpha}},$$

which is a concave quadratic in p_2. Solving the first-order condition $\dfrac{\partial \pi_2}{\partial p_2} = 0$ yields the best response

$$\hat{p}_2(\lambda_1) = \frac{\bar{\alpha}\beta + (\bar{\alpha} - \beta)\lambda_1}{2\bar{\alpha}}.$$

Now we turn to the fist period decisions. The consumers' optimal threshold in period 1 must be a solution to the following fixed point equation

$$\lambda_1 = \frac{p_1 - \delta\hat{p}_2(\lambda_1) + \delta\beta(1 - \lambda_1/\bar{\alpha})}{1 - \delta},$$

solving for for λ_1 yields

$$\hat{\lambda}_1(p_1) = \frac{2p_1 + \beta\delta}{(2 - \delta)\bar{\alpha} + \beta\delta}\bar{\alpha}.$$

The monopolist problem in the first period is

$$\begin{aligned} \max_{p_1} \quad & p_1(1 - \lambda_1/\bar{\alpha}) + \delta p_1 \pi_2(\lambda_1) \\ s.t. \quad & \lambda_1 - \lambda_2 \geq 0 \\ & \bar{\alpha} \geq \lambda_1, \lambda_2 \geq 0 \\ & p_1 \geq 0 \end{aligned} \tag{4.27}$$

replacing $\hat{p}_2(\lambda_1)$ and then $\hat{\lambda}_1(p_1)$ into the objective of Problem (4.27) yields

$$\pi_1 = \frac{p_1^2[(3\delta - 4)\bar{\alpha}^2 - 4\bar{\alpha}\beta\delta + \beta^2\delta] + p_1[(2 - \delta)^2\bar{\alpha}^2 + \bar{\alpha}\beta\delta(4 - \delta)]\bar{\alpha} + \bar{\alpha}^2\beta^2\delta}{\bar{\alpha}[(2 - \delta)\bar{\alpha} + \beta\delta]^2}. \tag{4.28}$$

Note that π_1 is concave in p_1 if and only if

$$\frac{\partial^2 \pi_1}{\partial p_1^2} = 2\beta^2\delta - 8\beta\delta - (8 - 6\delta) \leq 0,$$

which is true when $\beta/\bar{\alpha} \leq (2 + \sqrt{1 + 4/\delta})$ and in particular it is always true in this case since $\gamma = \beta/\bar{\alpha} < \frac{5\delta - 2 + \sqrt{4 + \delta(17\delta - 4)}}{4\delta} < (2 + \sqrt{1 + 4/\delta})$. Thus, the optimal price for period 1 must solve the first-order condition $\frac{\partial \pi_1}{\partial p_1} = 0$, which implies that

$$p_1^* = \frac{(2 - \delta)^2\bar{\alpha}^2 + (4 - \delta)\bar{\alpha}\beta\delta - 2\beta^2\delta}{(8 - 6\delta)\bar{\alpha}^2 + 8\bar{\alpha}\beta\delta - 2\beta^2\delta}\bar{\alpha} = \frac{(2 - \delta)^2 + (4 - \delta)\gamma\delta - 2\gamma^2\delta}{(8 - 6\delta) + 8\gamma\delta - 2\gamma^2\delta}\bar{\alpha}. \tag{4.29}$$

Replacing p_1^* into $\hat{\lambda}_1(p_1)$ and simplifying yields

$$\lambda_1^* = \frac{(2-\delta)^2 + 3\gamma\delta - \gamma^2\delta}{(8-6\delta) + 8\gamma\delta - 2\gamma^2\delta} 2\bar{\alpha},$$

replacing λ_1^* into the best response $\hat{p}_2(\lambda_1)$ and simplifying yields

$$p_2^* = \frac{(2-\delta) + (2+\delta)\gamma}{(8-6\delta) + 8\gamma\delta - 2\gamma^2\delta} \bar{\alpha}, \tag{4.30}$$

and replacing λ_1^* and p_2^* into $\lambda_2 = p_2 - \beta(1 - \lambda_1/\bar{\alpha})$ yields

$$\lambda_2^* = \frac{(2-\delta)^2 - (2-5\delta)\gamma - 2\gamma^2\delta}{(8-6\delta) + 8\gamma\delta - 2\gamma^2\delta} \bar{\alpha}.$$

It is easy to verify that all constraints hold at the optimal solution, and in particular, note that $\lambda_2^* > 0$ if and only if $\gamma < \dfrac{5\delta - 2 + \sqrt{4 + \delta(17\delta - 4)}}{4\delta}$. Moreover, note that $\lambda_1^* > \lambda_2^*$ since

$$\lambda_1^* - \lambda_2^* = \frac{2 - \delta + \gamma(2+\delta)}{(8-6\delta) + 8\gamma\delta - 2\gamma^2\delta} \bar{\alpha} > 0.$$

Finally, note that the equilibrium profit can be computed by replacing p_1^* into (4.28) and simplifying to get

$$\Pi^* = \frac{(2-\delta)^2 + 4\gamma\delta}{2[(8-6\delta) + 8\gamma\delta - 2\gamma^2\delta]} \bar{\alpha},$$

thus completing the proof. □

Proof of Proposition 4.4.2

Consider the optimal consumers' threshold for period 2 characterized in Proposition 4.4.1, and note that the threshold becomes negative when $\gamma \geq \dfrac{5\delta - 2 + \sqrt{4 - 4\delta + 17\delta^2}}{4\delta}$. In this case, we need to solve the seller's problem with the constraint that $\lambda_2 \leq 0$. Using the tail distribution (4.10), the seller's problem in the second period can be written as

$$\begin{aligned} \max_{p_2} \quad & p_2\lambda_1/\bar{\alpha} \\ s.t. \quad & \lambda_1 - \lambda_2 \geq 0 \\ & \bar{\alpha} \geq \lambda_1 \geq 0 \\ & \lambda_2 \leq 0 \\ & p_2 \geq 0. \end{aligned} \tag{4.31}$$

The constraint $\lambda_2 \leq 0$ imposes the following upper bound on the period 2 price

$$p_2 \leq \beta(1 - \lambda_1/\bar{\alpha}) = \bar{p}_2(\lambda_1).$$

Note that the above condition must hold with equality at the optimal solution because the objective of Problem (4.31) is always increasing in p_2, which implies that the optimal consumers' threshold for period 1 must be a solution to the following fixed point equation

$$\lambda_1 = \frac{p_1 - \delta\bar{p}_2(\lambda_1) + \delta\beta(1 - \lambda_1/\bar{\alpha})}{1 - \delta},$$

solving for for λ_1 yields $\bar{\lambda}_1(p_1) = p_1/(1 - \delta)$.

The seller's problem in period 1 is

$$\begin{aligned} \max_{p_1} \quad & p_1(1 - \lambda_1/\bar{\alpha}) + \delta p_1\pi_2(\lambda_1) \\ s.t. \quad & \lambda_1 - \lambda_2 \geq 0 \\ & \bar{\alpha} \geq \lambda_1 \geq 0 \\ & \lambda_2 \leq 0 \\ & p_1 \geq 0. \end{aligned} \tag{4.32}$$

Replacing $\bar{p}_2(\lambda_1)$ and then $\bar{\lambda}_1(p_1)$ into the objective of Problem (4.32) yields

$$\pi_1 = \frac{\bar{\alpha}(1-\delta)+\beta\delta}{\bar{\alpha}(1-\delta)}p_1 - \frac{\bar{\alpha}(1-\delta)+\beta\delta}{\bar{\alpha}^2(1-\delta)^2}p_1^2, \tag{4.33}$$

which is a concave function of p_1. Solving the first-order condition yields

$$p_1^* = \bar{\alpha}(1-\delta)/2, \tag{4.34}$$

replacing p_1^* into $\bar{\lambda}_1(p_1)$ yields $\lambda_1^* = \bar{\alpha}/2 = \bar{\alpha}\gamma/2$, and replacing λ_1^* into $\bar{p}_2(\lambda_1)$ yields

$$p_2^* = \bar{\alpha}\gamma/2. \tag{4.35}$$

Clearly $\lambda_2^* = 0$, and it can be easily verified that all constraints hold at the optimal solution. Finally, note that clearly $\lambda_1^* > \lambda_2^*$ and that we compute the equilibrium profit by replacing p_1^* into (4.33) and simplifying, which yields $\Pi^* = \bar{\alpha}(1-\delta+\gamma\delta)/4$, thus completing the proof. □

Proof of Proposition 4.4.3

For property (i), note that when $\gamma < \dfrac{5\delta-2+\sqrt{4-4\delta+17\delta^2}}{4\delta}$ the optimal prices are given by (4.29) and (4.30), subtracting p_1^* from p_2^* yields

$$p_2^* - p_1^* = \frac{2-3\delta+\delta^2-\gamma(2-3\delta+\delta^2)-\gamma^2\delta}{(8-6\delta)+8\gamma\delta-2\gamma^2\delta}\bar{\alpha},$$

where the denominator is always positive and the numerator is positive when

$$\gamma > \frac{-2+3\delta-\delta^2+\sqrt{4+4\delta-11\delta^2+2\delta^3+\delta^4}}{4\delta} = \tilde{\gamma}(\delta),$$

thus $p_2^* - p_1^*$ is positive if $\gamma > \tilde{\gamma}(\delta)$ and it is negative otherwise. Clearly $\tilde{\gamma}(\delta) \to 0$ as $\delta \to 1$. When $\gamma \geq \dfrac{5\delta-2+\sqrt{4-4\delta+17\delta^2}}{4\delta}$ the optimal prices are given by (4.34) and (4.35), subtracting p_1^* from p_2^* yields $p_2^* - p_1^* = \bar{\alpha}(\gamma-1+\delta)/2$ which is always positive in this case. Thus completing the proof of property (i).

For property (ii), when $\gamma < \dfrac{5\delta - 2 + \sqrt{4 - 4\delta + 17\delta^2}}{4\delta}$ the optimal prices are given by (4.29) and (4.30), differentiating p_1^* with respect to δ yields

$$\frac{\partial p_1^*}{\partial \gamma} = \frac{-\delta^2 + \gamma(-8 + 4\delta + 2\delta^2) - \gamma^2\delta(4 + \delta)}{(8 - 6\delta) + 8\gamma\delta - 2\gamma^2\delta},$$

which is negative since the numerator is positive for all $\gamma < \dfrac{5\delta - 2 + \sqrt{4 - 4\delta + 17\delta^2}}{4\delta}$ and the numerator is always negative. Moreover, differentiating p_2^* with respect to δ yields

$$\frac{\partial p_2^*}{\partial \gamma} = \frac{8 - 10\delta + \delta^2 + \gamma\delta(4 - 2\delta) + \gamma^2\delta(2 + \delta)}{(8 - 6\delta) + 8\gamma\delta - 2\gamma^2\delta},$$

which is positive since the numerator is positive for all $\gamma < \dfrac{5\delta - 2 + \sqrt{4 - 4\delta + 17\delta^2}}{4\delta}$ and the numerator is always positive. When $\gamma \geq \dfrac{5\delta - 2 + \sqrt{4 - 4\delta + 17\delta^2}}{4\delta}$, differentiating the optimal prices from (4.34) and (4.35) with respect to γ yields

$$\frac{\partial p_1^*}{\partial \gamma} = 0 \quad \text{and} \quad \frac{\partial p_2^*}{\partial \gamma} = \bar{\alpha}/2 \geq 0,$$

thus completing the proof. □

Proof of Theorem 4.5.1

Define $\bar{\gamma}_{RP} = \dfrac{5\delta - 2 + \sqrt{4 + \delta(17\delta - 4)}}{4\delta}$, which is the threshold on γ from the no commitment case. Note that $\bar{\gamma}_{RP} < 2$ since

$$2 - \bar{\gamma}_{RP} = \frac{2 + 3\delta - \sqrt{4 + \delta(17\delta - 4)}}{4\delta},$$

which is positive if and only if $8\delta(2 - \delta) > 0$, which is clearly true for all $\delta \in (0, 1]$.

Proof of Part (a): If $\gamma < \bar{\gamma}_{RP}$ the equilibrium profit with price commitment is given by (4.6) and the equilibrium profit with no commitment is given by (4.8), subtracting Π_{RP}^* from Π_{PC}^* yields

$$\Pi_{PC}^* - \Pi_{RP}^* = \frac{\delta[2 + (2 - \gamma)\delta]^2}{4[4 - (2 - \gamma)^2\delta][(4 - 3\delta) + 4\gamma\delta - \gamma^2\delta]},$$

which is strictly positive for all $\delta \in (0,1]$. To see this, note that the numerator is always positive, and for the denominator, note that $4-(2-\gamma)^2\delta$ is clearly positive for all $\gamma < 2$ and $(4-3\delta)+4\gamma\delta-\gamma^2\delta$ is positive if $\gamma < 2+\sqrt{1+4/\delta}$, which true since $\bar{\gamma}_{RP} < 2$.

If $\gamma \in [\bar{\gamma}_{RP}, 2)$ the equilibrium profit with price commitment is given by (4.6) and the equilibrium profit with no commitment is given by (4.9), subtracting Π^*_{RP} from Π^*_{PC} yields

$$\Pi^*_{PC} - \Pi^*_{RP} = \frac{(2-\gamma)^2\delta[1+(1-\gamma)\delta]}{4[4-(2-\gamma)^2\delta]},$$

which is always positive in this case.

Proof of Part (b): It suffices to note that when $\gamma \geq 2$ the equilibrium profit with price commitment is given by (4.7) and the equilibrium profit with no commitment is given by (4.9), clearly $\Pi^*_{PC} = \Pi^*_{RP}$. This completes the proof of the theorem. □

Bibliography

Daron Acemoglu and Asuman Ozdaglar. Opinion dynamics and learning in social networks. *Dynam. Games Appl.*, 1:3–49, 2011.

Daron Acemoglu, Munther A. Dahleh, Ilan Lobel, and Asuman Ozdaglar. Rate of convergence of learning in social networks. In *Proceedings of the American Control Conference 2009*, pages 2825–2830, 2009.

Daron Acemoglu, Munther A. Dahleh, Ilan Lobel, and Asuman Ozdaglar. Bayesian learning in social networks. *Rev. Econom. Stud.*, 78(4):1201–1236, 2011.

Daron Acemoglu, Kostas Bimpikis, and Asuman Ozdaglar. Dynamics of information exchange in endogenous social networks. *Theor. Econ.*, 9(1):41–97, 2014.

Anat R. Admati and Paul Pfleiderer. Direct and indirect sale of information. *Econometrica*, pages 901–928, 1990.

George-Marios Angeletos and Alessandro Pavan. Efficient use of information and social value of information. *Econometrica*, 75(4):1103–1142, 2007.

Itai Arieli and Manuel Mueller-Frank. Learning what matters. Technical Report 2466903, SSRN, 2014.

W Brian Arthur. Competing technologies, increasing returns, and lock-in by historical events. *The economic journal*, 99(394):116–131, 1989.

W Brian Arthur. Inductive reasoning and bounded rationality. *The American economic review*, 84(2):406–411, 1994.

Yossi Aviv and Amit Pazgal. Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing & Service Operations Management*, 10(3):339–359, 2008.

Yossi Aviv, Yuri Levin, and Mikhail Nediak. Counteracting strategic consumer behavior in dynamic pricing systems. In *Consumer-Driven Demand and Operations Management Models*, pages 323–352. Springer, 2009.

Moshe Babaioff, Robert Kleinberg, and Renato Paes Leme. Optimal mechanisms for selling information. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 92–109. ACM, 2012.

Abhijit V. Banerjee. A simple model of herd behavior. *Quart. J. Econom.*, 107(3):797–817, 1992.

Frank Bass. A new product growth for model consumer durables. *Management Science*, 15(5):215–227, 1969.

Frank M. Bass. Comments on "a new product growth for model consumer durables": the Bass model. *Management Sci.*, 50(12):1833–1840, 2004.

Dirk Bergemann and Alessandro Bonatti. Selling cookies. *American Economic Journal: Microeconomics*, forthcoming.

Dirk Bergemann and Stephen Morris. Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308, 2013.

Dirk Bergemann and Juuso Välimäki. Market diffusion with two-sided learning. *RAND J. Econ.*, 28(4):773–795, 1997.

David Besanko and Wayne L Winston. Optimal price skimming by a monopolist facing rational consumers. *Management Science*, 36(5):555–567, 1990.

Omar Besbes and Ilan Lobel. Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science*, 61(1):92–110, 2015.

Omar Besbes and Marco Scarsini. On information distortions in online ratings. Technical Report 13-36, Columbia Business School Research, 2015.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change in informational cascades. *J. Polit. Econ.*, 100(5):992–1026, October 1992.

Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. Learning from the behavior of others: conformity, fads, and informational cascades. *J. Econom. Perspect.*, 12(3):151–170, 1998.

Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

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

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

Erik Brynjolfsson, Lorin Hitt, and Heekyung Kim. Strength in numbers: How does data-driven decision-making affect firm performance? Working paper, 2011.

Jeremy I Bulow. Durable-goods monopolists. *The Journal of Political Economy*, pages 314–332, 1982.

Gérard P Cachon and Robert Swinney. Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science*, 55(3):497–511, 2009.

Arthur Campbell. Word-of-mouth communication and percolation in social networks. *American Economic Review*, 103(6):2466–2498, 2013.

Ozan Candogan, Kostas Bimpikis, and Asuman Ozdaglar. Optimal pricing in networks with externalities. *Oper. Res.*, 60(4):883–905, 2012.

Ozan Candogan, Kostas Bimpikis, and Asuman Ozdaglar. Optimal pricing in networks with externalities. *Operations Research*, 60(4):883–905, 2012.

Ronald H Coase. Durability and monopoly. *Journal of Law and Economics*, 15:143, 1972.

Ruomeng Cui, Gad Allon, Achal Bassamboo, and Jan Albert Van Mieghem. Information sharing in supply chains: An empirical and theoretical valuation. *Management Science*, forthcoming.

Laurens Debo and Senthil Veeraraghavan. Joining longer queues: information externalities in queue choice. *Manufacturing Service Oper. Management*, 11(4):543–562, 2009.

Laurens G. Debo, Christine A. Parlour, and Uday Rajan. Signaling quality via queues. *Management Sci.*, 58(2):876–891, 2012.

Glenn Ellison and Drew Fudenberg. Rules of thumb for social learning. *J. Polit. Econ.*, 101(4):612–643, 1993.

Glenn Ellison and Drew Fudenberg. Word-of-mouth communication and social learning. *Quart. J. Econom.*, 110(1):93–125, 1995.

Esther Gal-Or. Information sharing in oligopoly. *Econometrica*, pages 329–343, 1985.

Guillermo Gallego and Garrett van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Sci.*, 40(8):999–1020, 1994.

Peter W. Glynn. Diffusion approximations. In *Stochastic Models*, volume 2 of *Handbooks Oper. Res. Management Sci.*, pages 145–198. North-Holland, Amsterdam, 1990.

Jacob K. Goeree, Thomas R. Palfrey, and Brian W. Rogers. Social learning with private and common values. *Econom. Theory*, 28:245–264, 2006.

Albert Y. Ha and Shilu Tong. Contracting and information sharing under supply chain competition. *Management science*, 54(4):701–715, 2008.

Albert Y. Ha, Shilu Tong, and Hongtao Zhang. Sharing demand information in competing supply chains with production diseconomies. *Management Science*, 57(3):566–581, 2011.

Helios Herrera and Johannes Hörner. Biased social learning. *Games Econom. Behav.*, 80:131–146, 2013.

Bar Ifrach, Costis Maglaras, and Marco Scarsini. Bayesian social learning from consumer reviews. Technical Report 2293158, SSRN eLibrary, 2015.

Bar Ifrach. *Market Dynamics with Many Agents: Applications in Social Learning and Competition.* PhD thesis, Decision, Risk, and Operations, Columbia University, 2012.

Ali Jadbabaie, Pooya Molavi, Alvaro Sandroni, and Alireza Tahbaz-Salehi. Non-Bayesian social learning. *Games Econom. Behav.*, 76(1):210–225, 2012.

Bing Jing. Social learning and dynamic pricing of durable goods. *Marketing Science*, 30(5):851–865, 2011.

Michael L Katz and Carl Shapiro. Network externalities, competition, and compatibility. *The American economic review*, 75(3):424–440, 1985.

Hassan Khalil. *Nonlinear Systems.* Prentice Hall, third edition, January 2002.

Thomas G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl.*, 6(3):223–240, 1978.

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

Jonathan Lafky. Why do people rate? Theory and evidence on online ratings. *Games Econom. Behav.*, 87:554–570, 2014.

Xinxin Li and Lorin M. Hitt. Self-selection and information role of online product reviews. *Inform. Syst. Res.*, 19(4):456–474, 2008.

Lode Li and Hongtao Zhang. Confidentiality and information sharing in supply chain coordination. *Management Science*, 54(8):1467–1481, 2008.

Lode Li. Cournot oligopoly with information sharing. *The RAND Journal of Economics*, pages 521–536, 1985.

Lode Li. Information sharing in a supply chain with horizontal competition. *Management Science*, 48(9):1196–1212, 2002.

Qian Liu and Dan Zhang. Dynamic pricing competition with strategic customers under vertical product differentiation. *Management Science*, 59(1):84–101, 2013.

Micael Luca. Reviews, reputation, and revenue: the case of yelp.com. Technical Report 12-016, Harvard Business School, 2011.

Enrico Moretti. Social learning and peer effects in consumption: Evidence from movie sales. *The Review of Economic Studies*, 78(1):356–393, 2011.

Stephen Morris and Hyun Song Shin. Social value of public information. *The American Economic Review*, 92(5):1521–1534, 2002.

David P. Myatt and Chris Wallace. Cournot competition and the social value of information. *Journal of Economic Theory*, 2015.

Barak Y Orbach and Liran Einav. Uniform prices for differentiated goods: The case of the movie-theater industry. *International Review of Law and Economics*, 27(2):129–153, 2007.

Barak Orbach. Antitrust and pricing in the motion picture industry. *Yale Journal on Regulation*, 21:317–366, 2004.

Yiangos Papanastasiou and Nicos Savva. Dynamic pricing in the presence of social learning and strategic consumers. To appear in Management Science, 2014.

Yiangos Papanastasiou and Nicos Savva. Dynamic pricing in the presence of social learning and strategic consumers. *Management Science*, forthcoming.

Yiangos Papanastasiou, Nitin Bakshi, and Nicos Savva. Scarcity strategies under quasi-bayesian social learning. working paper, London Business School, 2014.

Robert Lewis Phillips. *Pricing and revenue optimization*. Stanford University Press, 2005.

Michael Raith. A general model of information sharing in oligopoly. *Journal of Economic Theory*, 71:260–288, 1996.

Anita Rao. Online content pricing: Purchase and rental markets. *Marketing Science*, 34(3):430–451, 2015.

Matthew J Salganik and Duncan J Watts. Web-based experiments for the study of collective social dynamics in cultural markets. *Topics in Cognitive Science*, 1(3):439–468, 2009.

Matthew J Salganik, Peter Sheridan Dodds, and Duncan J Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *science*, 311(5762):854–856, 2006.

Noam Shamir and Hyoduk Shin. Public forecast information sharing in a market with competing supply chains. Working paper, 2013.

Noam Shamir. Strategic information sharing between competing retailers in a supply chain with endogenous wholesale price. *International Journal of Production Economics*, 136(2):352–365, 2012.

Weixin Shang, Albert Y. Ha, and Shilu Tong. Information sharing in a supply chain with a common retailer. *Management Science*, forthcoming.

Hyoduk Shin and Tunay I. Tunca. Do firms invest in forecasting efficiently? the effect of competition on demand forecast investments and supply chain coordination. *Operations Research*, 58(6):1592–1610, 2010.

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

Nancy L Stokey. Intertemporal price discrimination. *The Quarterly Journal of Economics*, pages 355–371, 1979.

Xuanming Su and Fuqiang Zhang. Strategic customer behavior, commitment, and supply chain performance. *Management Science*, 54(10):1759–1773, 2008.

Yeneng Sun and Yongchao Zhang. Individual risk and lebesgue extension without aggregate uncertainty. *Journal of Economic Theory*, 144(1):432–443, 2009.

Yeneng Sun. The exact law of large numbers via fubini extension and characterization of insurable risks. *Journal of Economic Theory*, 126(1):31–69, 2006.

Kalyan T. Talluri and Garrett J. van Ryzin. *The Theory and Practice of Revenue Management*. Springer, New York, 2005.

Hal R Varian. Buying, sharing and renting information goods. *The Journal of Industrial Economics*, 48(4):473–488, 2000.

Nikolaos Vettas. On the informational role of quantities: durable goods and consumers' word-of-mouth communication. *Internat. Econ. Rev.*, 38(4):915–944, 1997.

Xavier Vives. Duopoly information equilibrium: Cournot and Bertrand. *Journal of Economic Theory*, 34(1):71–94, 1984.

Xavier Vives. Strategic supply function competition with private information. *Econometrica*, 79(6):1919–1966, 2011.

Yi Xiang and Miklos Sarvary. Buying and selling information under competition. *Quantitative Marketing and Economics*, pages 1–31, 2013.

Man Yu, Laurens Debo, and Roman Kapuscinski. Strategic waiting for consumer-generated quality information: Dynamic pricing of new experience goods. *Management Science*, 2015.