Consumer Search with Observational Learning*

Daniel Garcia† Sandro Shelegia‡

October 6, 2015

Abstract

This paper studies observational learning in a consumer search environment. In our model, consumers observe the purchasing decision of a predecessor, and *emulate* by initiating their search at the firm she purchased from. Emulation induces a *social multiplier of demand* that leads to lower equilibrium price. As search costs increase, firms compete fiercely to attract consumers and the price converges to the marginal cost. The effect of emulation on prices is stronger when either: (i) the number of firms increases, (ii) consumers observe more predecessors, and (iii) prices are adjusted more frequently. We show that emulation is uniquely optimal when consumers' preferences are positively correlated, and the correlation leads to free-riding that further reduces prices for high search cost.

JEL Classification: D11, D83, L13

Keywords: Consumer Search; Observational Learning, Social Multiplier of Demand, Emulation.

---

*We thank Jan Eeckhout, Maarten Janssen, Marco Haan, Jose Luis Moraga-Gonzalez, Stephan Lauermann and Andrew Rhodes for their comments as well as audiences at the 6th Consumer Search and Switching Workshop at Indiana University, Jornadas de Economia Industrial in Barcelona, the MACCI Conference in Consumer Search in Bad Homburg, the Simposio de Analisis Economic in Palma de Mallorca, UPF, University of Bern, PUC Chile and Universidad Alberto Hurtado. Garcia gratefully acknowledges financial support by the Hardegg Foundation. All errors are to be attributed to the authors only.

†Department of Economics, University of Vienna. Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. Email: daniel.garcia@univie.ac.at.

‡Department of Economics and Business, Universitat Pompeu Fabra. Ramon Trias Fargas 25-27, Barcelona 08005, Spain. Email: sandro.shelegia@upf.edu.

1 Introduction

Observational learning has been the object of study of a large and important literature in economics since the seminal contributions of Banerjee (1992) and Bikhchandani et al. (1992). In the classical model, a sequence of individuals faces a simple decision problem under uncertainty and each individual observes the history of decisions of her predecessors. As argued by Banerjee (1992), this simple environment resembles the problem faced by consumers in markets where previous consumers' choices may be informative about the relative value of different products. Search markets, where consumers actively engage in costly activities to gather information about different alternatives, are prominent examples of such environments. Intuitively, consumers may *free ride* on the search effort of others, initiating their search in the firm their predecessors purchased from. In a sequential search environment, this form of *emulation* induces a *social multiplier of demand*, because consumers are more likely to purchase the good from the first firm they visit. A natural question to ask, then, is whether, and to what extent, are firm's prices affected by stronger social interactions in search markets.

To the best of our knowledge, no paper has studied this issue. We attempt to bridge this gap with a simple oligopoly model of search with heterogenous products in the spirit of Wolinsky (1986) and Anderson and Renault (1999) (henceforth ARW). In the simplest version of the model, a large number of consumers derive utility from consuming a single unit of a given good that comes in two different varieties, each sold by a different firm. Consumers are initially uninformed about their valuation for each variety or the price charged by each firm, but may learn these after engaging in costly sequential search. The search procedure is particularly simple: consumers visit one firm at *random* and buy there without further search if their surplus exceeds a certain cutoff. Otherwise, they visit both firms and buy from the highest surplus-offering firm. Firms take consumers' behavior as given and choose prices simultaneously to maximize expected profits.

We depart from this model in two ways. We allow valuations for each variety to be positively correlated across consumers. In addition, we inform each consumer of the purchasing decision of a single predecessor. Importantly, consumers do not observe whether their predecessors visited other firms nor the price they paid. Because of positive correlation, it is always optimal for consumers to initiate their search in the store their predecessor bought. We refer to such behavior as *emulation*. In order to understand the effect of emulation we first study the limit case where preferences are independent across consumers but consumers still emulate their predecessors (which amounts to a change in the tie-breaking rule used in ARW.).

Because emulation is at the core of our results (to be discussed shortly) it is important to note that emulation is consistent with a recent body of evidence in economics and marketing. For instance, Miller and Mobarak (2015) conduct a marketing intervention in Bangladesh to assess how stove adoption is affected by social learning. They show that some households are more likely to adopt the new technology if they are informed of the purchasing decision of an important member of their community. In another field experiment, Cai et al. (2009) show that restaurant-goers are more likely to order those goods that are presented to them as more popular. Similarly, Moretti (2011) analyzes the movie market where an unexpected increase in the first-week's box office revenue has a persistent and significant effect on future attendance. Finally, Zhang (2010) shows that patients waiting for a kidney draw negative quality inferences from earlier refusals in the queue, thus becoming more inclined towards refusal themselves.[1]

We show that emulation drastically changes pricing in a search model, even when consumers' preference are uncorrelated. The mechanism is as follows. In a search model, each individual consumer is more likely to start her search in a given firm the higher its market share, linking individual demand to aggregate demand (conditional on prices). Because of this *social multiplier*, firms lower prices not only to retain incoming consumers (as in the standard model) but also to increase the number of consumers they attract through emulation. We show that the resulting (symmetric) equilibrium price is a fraction of the price in ARW. We further show that this fraction depends on the magnitude of the social multiplier, it is decreasing in search costs and it converges to zero as the cost of search increases. Higher search cost exacerbates the impact of emulation because the share of *captive* consumers (those who never search) increases. In the standard model, firms have monopoly power on this segment so that, as the proportion of these consumers increases, prices tend to increase. In our model, however, the likelihood that one such consumer visits a given firm depends on the firm's share among those consumers who actively search. Since firms cannot price-discriminate between the two groups, as the proportion of consumers who stop at the first firm increases, firms engage on increasingly fierce fight for searchers, leading to Bertrand-like competition and eventually to prices that can be as low as the marginal cost.

We then extend this result in three important dimensions. First, we consider the role of competition. Armstrong and Zhou (2011) shows that the order in which consumers search has a vanishing effect in prices as the number of firms grow, provided that this order is exogenously given. In our model, however, regardless of the number of firms, the same comparative statics obtain. Furthermore, the ratio of the price in the ARW model to the price in our model increases with the number of firms, suggesting that the effect we uncover is stronger in more competitive markets. Second, we enrich the model by enlarging the set of predecessors that a new consumer observes. This has two main effects. It puts further downward pressure on prices since small price deviations trigger larger changes in the number of first visits, leading to a bigger social multiplier of demand. Thus, as long as pure strategy equilibria exist, prices are lower. Importantly, however, the

[1] For a recent discussion on the huge literature on social learning see Möbius and Rosenblat (2014).

only candidate to a pure-strategy equilibrium price converges to zero *before* the share of searchers vanishes. This is inconsistent with optimality since firms retain (some) monopoly power in the segment of those consumers who visit both firms. Therefore, pure strategy equilibrium fails to exist in environments with sufficiently rich observational learning.[2] Third, we show that if firms can adjust their prices frequently, the symmetric stationary equilibrium price is even lower than in the baseline model where firms set prices only once. This is because a reduction in the price today leads to an increase in future market shares that translates into profits at future (higher) prices. This gives further incentives to undercut and results in even lower stationary prices.

Once we understand the effect of (pure) emulation on pricing, we move to a richer (but less tractable) model where emulation is uniquely optimal because consumers' valuations are positively correlated. More precisely, we assume that the valuations that consumers derive from the variety offered by each firm are drawn from one of two distributions, one of which (High) stochastically dominates the other (Low). Assuming that neither firms nor consumers observe which distribution is realized, we can focus on the effect of learning on consumer behavior and prices.[3] Note that observing a predecessor buying a given variety leads to an upwards update of its distribution of valuations and a downwards update in the rival's distribution. In addition, consumers' searching strategies change due to the informational content of their predecessors' purchase. In particular, consumers are now willing to accept a lower surplus from their first visit, thus reducing search effort. We term this the *free-riding* effect of observational learning.

The total effect of learning on equilibrium prices is not straightforward. On the one hand, as explained above, less search leads to higher competition because the social multiplier becomes more important. On the other hand, a price deviation triggers a change of beliefs that may overcome this effect. If a firm deviates to a higher price, the proportion of sales in each demand state changes, thus changing the posterior belief that consumers hold about the other firm's distribution. Since the elasticity of demand is lower for the firm with a higher distribution of valuations, consumers become more pessimistic about its rival's distribution the higher the price of the firm they visit. Hence, the surplus (net of the price) they demand for buying right away decreases in the price, reducing the elasticity and increasing prices. We show that for the uniform case the first effect is stronger for relatively small search costs while for sufficiently high search costs the second effect dominates and prices may be even lower than in the model without correlation of preferences. Indeed, we show that, for any distribution satisfying a certain likelihood ratio

[2] As we discuss in detail later, this result is closely related to the mixed strategy equilibrium described by Armstrong and Zhou (2011), who assume that consumers observe prices before they embark on their first search.

[3] This rules out signaling motives for prices. One possible interpretation of our model would then be a sub-market within a larger economy. Each sub-market draws its state independently and all social interactions occur within its limits.

condition, the equilibrium price converges to marginal cost for a smaller search cost than in the baseline model.

These results have important implications for our understanding of the impact of Information Technologies on market outcomes. On the one hand, search engines, price-comparison websites and similar internet platforms allow consumers to obtain information about prices and product characteristics at low cost. On the other hand, online social networks and consumer information aggregators provide information about other consumers' behavior and foster consumer emulation. For the classical search model, both innovations would translate into lower search costs and more competitive outcomes. According to our model, however, the combination of these two forces may increase or decrease prices, depending on the size of search costs, the strength of social interactions and the number of firms in the industry.

Literature Review

Several recent papers have analyzed consumer search with observational learning, but they all assume that prices are fixed exogenously. Kircher and Postlewaite (2008) study consumers who differ in their willingness to search among firms differ in quality. Although prices are fixed, firms may decide to offer a valuable service (an exogenously fixed discount) to any consumer who visits their store. Kircher and Postlewaite show that equilibria may arise where high-quality firms offer a discount to those consumers who search more actively and those who search less actively follow their advice. Hendricks et al. (2012) present a model of observational learning with multiple types in the spirit of Smith and Sørensen (2000), where each consumer has to decide between acquiring a costly signal about the quality of a single good, buying it right away, or not buying. They focus on the long-run dynamics of sales for high and low quality products and the possibility of bad herds arising.[4]

In the consumer search literature with price competition, the closest papers to ours are Armstrong et al. (2009) and Armstrong and Zhou (2011), who present a model of prominence in consumer search where one firm is sampled first by all consumers. In Armstrong et al. (2009) a given firm is made prominent exogenously. In the resulting equilibrium, the prominent firm charges a lower price than her rivals because her share of returning customers (who are typically less responsive to prices) is lower. One may view our framework as that of endogenous prominence, where the share of first visits depends on the price. Armstrong and Zhou (2011) study several different models that rationalize prominence. One such model is based on observable price competition where consumers rationally search the lowest-pricing firm first. Because demand is discontinuous in prices, the resulting equilibrium involves mixed strategies and has a property that higher search

[4] See also Ali (2014) and Mueller-Frank and Pai (forthcoming) for related models.

cost leads to (stochastically) lower prices. One may view observability of prices as an extreme example of observational learning where consumers observe market shares that are a sufficient statistic for the price ranking. We show that as long as this signal is sufficiently noisy, pure strategy equilibrium still exists.[5]

Our paper is also related to the literature on pricing in the presence of social learning. Galeotti (2010) studies a model of non-sequential search with homogenous goods whereby consumers exchange price information in a network. In a mixed-strategy equilibrium, consumers free-ride on others' search efforts. As a result, when search costs are low, an increase in interpersonal communication leads to higher rather than lower prices. Campbell (2013) and Chuhay (2010) analyze the impact of word-of-mouth communication on monopoly pricing and product design. Bose et al. (2006) study a dynamic interaction between a monopolist and a sequence of consumers with common valuation who observe each other's purchasing decisions. In a related effort, Kovac and Schmidt (2014) study a dynamic market where two firms offer a homogenous product and consumers learn prices from others. Since our focus is on competition and we abstract from dynamic issues, we view our work as complementary to this strand of the literature.[6]

Finally, a number of papers have studied the relationship between current market shares and future demand. Becker (1991) introduces aggregate demand as an argument of the individual demand function. Our model provides a simple micro-foundation for such an assumption and derives novel implications. Caminal and Vives (1996) studies a dynamic signaling game where new cohorts of consumers observe past market shares of an experience good, but not prices and try to infer quality from this information. Firms use prices to manipulate market shares and attract consumers. While their setting is different in that consumers do not search, we also find that firms use prices to attract consumers, but here consumers free-ride on their predecessors' efforts and due to search cost, herds form, while in Caminal and Vives these effects are mute.

The reminder of the paper is organized as follows. Section 2 introduces the full model with correlated consumer valuations. Section 3 focuses on the duopoly case with independent valuations. We then extend the model to oligopoly in Section 3, and to more complex emulation processes in Section 5. Section 6 re-introduces correlation in consumers' valuations, characterizes the equilibrium and derives comparative statics and welfare analysis. Section 7 shows that the model is robust to the introduction of dynamic pricing. Section 8 concludes. All proofs are contained in the Appendix.

[5]See also Haan and Moraga-Gonzalez (2011), where prominence depends on advertising efforts and firm profits may decrease in search cost (although prices increase and consumer surplus decreases in the symmetric case).

[6]Dynamic pricing is also the focus of a recent literature on switching costs (e.g. Doganoglu (2010); Cabral (2013); Rhodes (2014)). See Section 7 for a detailed discussion. We can note here that observational learning leads to lower prices even when price is chosen once, firms cannot discriminate between consumers and there is a finite horizon.

2 The Model

Consider a market populated by a countably infinite number of consumers interested in purchasing a single unit of a differentiated good that comes in two varieties, each sold by one firm, 1 and 2. Consumers are initially uncertain about their valuation of each firm's product, but may acquire this information through sequential search. We assume that each visit has a cost of c in utility units and all consumers can recall their previously sampled varieties at no additional cost. For each firm utilities are drawn according to a cumulative distribution $G(u; s)$, where $s \in \mathcal{S}$ represents the realized state in each firm. We assume that states are independent across firms and unobserved by firms.[7] Conditional on the state, utility realizations are independent across consumers and firms. Notice that, for a given state all consumers draw their utility from the same distribution, thereby utilities are positively correlated across consumers. We denote by $q(s)$ be the unconditional probability that state s realized in any given firm.

We also impose some technical assumptions on the corresponding density function, $g(u; s)$. In particular, we assume that for any s, $g(u; s)$ is strictly positive and continuously differentiable on a closed finite support $[\underline{u}, \overline{u}]$, independent of s.[8] We further assume that $1 - G(u; s)$ is Log-concave on $[\underline{u}, \overline{u}]$. Finally, for most of the analysis we assume that the outside option is sufficiently bad so that every consumer buys one unit regardless of the observed prices.[9]

Consumers arrive to the market sequentially. In particular, we assume that consumer i arrives to the market in period i ($i = 1, 2, 3...$) and leaves the market before the next consumer arrives. All consumers (with the exception of the first one) are uncertain over their arrival time and hold a common prior $\nu(i)$ over it.[10] We introduce observational learning by allowing each individual consumer $i > 1$ to observe the purchasing decision of her predecessor $i - 1$. Importantly, and in line with the literature on social learning, i does not observe her predecessor's search behavior, utility realizations or prices. Given this information, the consumer decides which firm to visit first and finally decides where to buy. Consumer $i = 1$ has no predecessor and, thus, has no additional information. As in ARW, we assume that she chooses which firm to visit by flipping a fair coin.

Firms choose prices simultaneously at the beginning of the game in order to maximize average expected discounted profits. Let $x^i(p_1, p_2)$ be the probability that consumer i buys at firm 1 given the price pair (p_1, p_2). Then firm 1's expected discounted profit per

[7] This is the natural assumption if firms compete in many local markets and the state is drawn independently in each of them. Moreover, firms may be ex-ante uncertain of the quality of their products and we are mostly interested in the introductory phase.

[8] Most of our results could be extended to infinite supports, provided that a regularity condition on the Hazard Rate holds.

[9] See Appendix C for an extension where we introduce a relevant outside option.

[10] This assumption greatly simplifies the problem of the consumer and allows us to focus on stationary strategies. Monzón and Rapp (2014) show that the classical results in herding models survive even if consumers are uncertain about their arrival time.

consumer can be written as

$$\Pi_1(p_1, p_2) = p_1(1-\delta)\sum_{i=1}^{\infty}\delta^{i-1}x^i(p_1, p_2).$$

Notice that if $\delta = 0$, the pricing problem is the same as in ARW because the firm prices to the first consumer, who visits each firm with exogenous and equal probability.

In what follows we focus on the symmetric pure strategy (Perfect Bayesian) Equilibrium.

2.1 Consumer Behavior

Let p^* denote the symmetric equilibrium price. If firm 1 charges $p_1 \neq p^*$ while consumers expect the other firm to charge p^*, then a consumer who visits it first will search further if and only if her utility realization u_1 at firm 1 is below $\tilde{w} - p^* + p_1$ where $\tilde{w}$ solves

$$\int_{\tilde{w}}^{\bar{u}} (u - \tilde{w}) \sum_{s \in \mathcal{S}} q(s; \tilde{w}) g(u; s) du = c. \tag{1}$$

where $q(s; u)$ is the probability that the state in the other firm is s given that the consumer observed her predecessor buying in this firm and the utility realization was u. If no solution exists, then set $\tilde{w} = \underline{u}$.

The consumer who arrives in period 1 has no a priori information to discriminate firms. She makes her first visit randomly and buys at the first firm if and only if $u_{i1} - p_1 > \hat{w} - p^*$ and searches the other firm otherwise, where $\hat{w}$ solves

$$\int_{\hat{w}}^{\bar{u}} (u - \hat{w}) \sum_{s \in \mathcal{S}} q(s) g(u; s) du = c. \tag{2}$$

Notice that $\hat{w}$ is the reservation utility of a consumer in the ARW model who faces the prior distribution $G(u) = \sum q(s)G(u; s)$.

All remaining consumers observe a predecessor buying from firm $j \in \{1, 2\}$ and expect the same price p^* at both firms before embarking on their first search. Because utilities are positively correlated across individuals for a given store, it is optimal for consumers to start their search from the store where their predecessors has bought. If at that store they observe a utility realization u_{ij} and a price p_j, they continue to search if $u_{ij} < \tilde{w}(p_j)$, where $\tilde{w}(p_j)$ solves

$$\int_{\tilde{w} - p_j + p^*}^{\bar{u}} (u - \tilde{w} + p_j - p^*) \sum_{s \in \mathcal{S}} q(s; \tilde{w}, p_j) g(u; s) du = c. \tag{3}$$

Notice that, as long as $q(s; u, p)$ is not constant, $\tilde{w}(p_j)$ will not be linear in p_j. That is, if a consumer observes a price p_j, she will buy outright only if she obtains a net surplus exceeding $\tilde{w}(p_j) - p_j$, which is not constant on p_j.

Therefore, there are three main differences between our model and ARW. First, consumers begin their search in the store their predecessor bought from. We will refer to this behavior as *emulation*. Second, consumers who observe that their predecessors bought in a given store update their beliefs about their prospects in other stores. Because consumers' preferences are positively correlated, they will become pessimistic about their prospects elsewhere, thus reducing their search effort. We will say that consumers *free ride* on others' search effort. Finally, consumers price-elasticity will be different since firms can affect learning by changing the price as shown by the non-linearity of $\tilde{w}(p_j)$.

In the next Section we will concentrate on the effects of *emulation* and defer the discussion of the other two forces to Section 7. The reason for this is twofold. First, the model with emulation but without free-riding is very tractable, which gives us a lot of mileage to consider interesting variations of the baseline model. Second, as we demonstrate in Section 7, correlated preferences either mitigate or exacerbate the effects of emulation but emulation remains the main force at work.

3 Independent Preferences

In order to shut down the effects of correlation, we assume in what follows that $G(u;s) = G(u)$ for all $s \in S$. We continue to assume, however, that consumers follow their predecessors when deciding which firm to visit first. As we shall see, in equilibrium, consumers are indifferent over which store to start their search on, and so we may take this assumption as a tie-break rule. In this sense, ARW can be considered as a special case of our model whereby consumers ignore the actions of their predecessor and choose randomly. This behavior, however, fails to satisfy trembling-hand-perfection refinement because, if firms may make pricing mistakes, then following the predecessor is strictly preferable. In addition, as we shall show, emulation is collectively optimal and, therefore, may be supported as a social norm among consumers.

We shall now denote by $\hat{w}$ the reservation utility of a consumer observing the equilibrium price. Since the distribution in the other state is known, the equation simplifies to

$$\int_{\hat{w}}^{\bar{u}} (u - \hat{w}) g(u) du = c.. \tag{4}$$

If no solution exists, then set $\hat{w} = \underline{u}$. Since predecessors' choices are not informative about unobserved prices or utilities, once at the first firm, consumers demand the same reservation utility $\hat{w} + p_j - p^*$, which is now linear in p_j.

3.1 Equilibrium

In order to write firm 1's market share in period i (i.e the probability a consumer buys from firm 1 in the period) let us first introduce two objects, M_1 and M_2 as the probability that a consumer who makes her first visit to firm 1 and 2, respectively, purchases at firm 1. These probabilities can be written as

$$M_1(p_1, p^*) = (1 - G(\hat{w} + p_1 - p^*)) + \int_{\underline{u}}^{\hat{w}+p_1-p^*} G(u - p_1 + p^*)g(u)du \tag{5}$$

$$M_2(p_1, p^*) = G(\hat{w})(1 - G(\hat{w} + p_1 - p^*)) + \int_{\underline{u}}^{\hat{w}+p_1-p^*} G(u - p_1 + p^*)g(u)du. \tag{6}$$

The expression for $M_1(p_1, p^*)$ includes consumers who buy outright after a utility draw above $\hat{w} + p_1 - p^*$ (the first term) and consumers who continue to firm 2 but come back (the second term). $M_2(p_1, p^*)$ is composed of those consumers who visit firm 2, draw a utility below $\hat{w}$ and go on to firm 1. Of these, consumers with utility draws at firm 1 above $\hat{w} + p_1 - p^*$ all buy from firm 1 (the first term), whereas all remaining consumers have to compare prices and utility levels (the second term). First visits are valuable because some consumers buy outright, i.e., $M_1(p_1, p^*) \geq M_2(p_1, p^*)$.

Using this notation, and suppressing terms in brackets while noting that x^i and M_j depend on prices, the probability that consumer i (who observes the purchasing decision of consumer $i-1$) purchases from firm 1 can be written as

$$x^i = x^{i-1}M_1 + (1 - x^{i-1})M_2 = M_2 + x^{i-1}\big(M_1 - M_2\big), \tag{7}$$

where $x^0 = 1/2$ since the first consumer visits either firm with equal probability. Consumer i first visits firm 1 with probability x^{i-1} and buys there with probability M_1. She may also buy at firm 1 after having visited firm 2 first, which happens with probability M_2. The last expression clearly shows the value of being the first to be visited. Firm 1's market share is M_2 plus the extra consumers it gains because they have visited it first, and thus bought with an increased probability $M_1 - M_2 > 0$.

We use the recursion in (7) to obtain the closed form expression for x_i for $i \geq 2$ as

$$x^i = \frac{(1 - M_1 - M_2)(M_1 - M_2)^i + 2M_2}{2(1 - (M_1 - M_2))}. \tag{8}$$

We can now use (8) to compute the discounted per consumer market share (demand) of firm 1 that can be written as

$$Q_1(p_1) = (1 - \delta)\sum_{i=0}^{\infty} \delta^{i-1}x_i = \frac{M_1 + M_2 - \delta(M_1 - M_2)}{2(1 - \delta(M_1 - M_2)).} \tag{9}$$

The expected discounted demand is a combination of the demand of early consumers (whose visiting probabilities are closer to random) and the demand of all the later arriving consumers, who are allocated according to the stationary market shares. The difference between these groups stems from those consumers whose utility realizations are such that, regardless of the firm they visit first, they buy there. If these consumers do not observe previous purchases, they buy in each firm with probability 1/2. If they observe a predecessor, however, they buy at each firm with the probability that a searcher buys at that firm (since they are following them, or following others who follow them, etc.) Thus, their price elasticity is higher the later they arrive in the queue.

Notice also that this profit function includes the demand function of Anderson and Renault (1999) as a special case. In particular, when $\delta \to 0$, this becomes

$$Q_1(p_1) = \frac{1}{2}(M_1 + M_2), \tag{10}$$

the familiar demand equation in the ARW model. Recall that in ARW each firm is equally likely to be visited first by a consumer. Half of them make the first visit to firm 1 and purchase with probability M_1, while the other half make the first visit to firm 2, and purchase from firm 1 with probability M_2.

At the other extreme where $\delta \to 1$, initial individuals whose visiting probabilities are not equal to the stationary ones are no longer relevant and we obtain

$$Q_1(p_1) = \frac{M_2}{1 - (M_1 - M_2)}. \tag{11}$$

To gain some intuition for this expression, notice that in this model there are two types of consumers: those who buy in the store they visit first (captives) and those who buy in the store that offers them the highest surplus (searchers). The demand of the first group is driven by the choice of the most recent searcher in the queue. Thus, demand of both searchers and captives is entirely determined by the firm's share of searchers. This share is the probability that a consumer buys from that firm if she does not visit it first (M_2) divided by the total mass of searchers, which is given by $1 - (M_1 - M_2)$. To see this, recall that $M_1 - M_2$ captures those consumers who buy in firm 1 only if they visit it first, i.e. consumers who always buy in the first store they visit. Notice also that in a symmetric equilibrium we have $M_1 + M_2 = 1$ and thus the demand is 1/2 regardless of δ.

The following definition, inspired by the literature on social interactions (e.g. Manski (1993), Glaeser et al. (2003)) turns out to play a crucial role in our analysis.

Definition 1. *The social spillover, η, is the increase in the individual demand of consumer i in a given firm resulting from consumer $i-1$ buying there. Therefore, $\gamma = \frac{\eta}{1-\eta}$ is the social multiplier of demand and captures the total additional demand obtained from one extra purchase.*

In this baseline model, $\eta = M_1 - M_2$ represents the change in the probability that consumer i purchases at firm 1, given that she visited it first rather than second. Because each consumer's increased purchasing probability at firm 1 leads to an increase in the probability of a first visit (and thus purchase) by her successor, the total additional demand, or the social multiplier of demand, is $\gamma = \sum_{i=1}^{\infty} \eta^i = \frac{\eta}{1-\eta}$. Since M_1 and M_2 are probabilities and $M_1 \geq M_2$, η ranges from 0 to 1, whereas γ ranges from 0 to infinity.

Before we provide the equilibrium characterization, it is convenient to define $\hat{p}$ as the symmetric equilibrium price in the ARW model ($\delta \to 0$ in our model) which can be obtained by taking the derivative of $Q_1(p_1)$ given by (10) with respect to p_1, imposing full symmetry ($p_1 = p^*$ and $Q_1(p^*) = 1/2$) and noting that $p^* = -Q_1(p^*)/Q_1'(p^*)$. This gives the familiar

$$\hat{p} = \frac{1}{2\int_{\underline{u}}^{\hat{w}} g(u)^2\, du + (1 - G(\hat{w}))g(\hat{w})}. \tag{12}$$

As shown in Anderson and Renault (1999), $\hat{p}$ is increasing in c for all $c < \bar{c}$.

Now we are ready to characterize the symmetric equilibrium price for any δ. Recall that $\hat{p}$ is the price in the ARW model (i.e. the price for $\delta = 0$).

Proposition 1. *In equilibrium, both firms charge*

$$\tilde{p} = (1 - \delta\eta^*) \cdot \hat{p},$$

where $\eta^* = (1 - G(\hat{w}))^2 \leq 1$ *is the social spillover in equilibrium. Thus, for any* $\delta > 0$ *and* $c > 0$ *we have* $\tilde{p} < \hat{p}$.

Proof. All proofs are provided in the Appendix. □

As δ increases, the weight of the initial batch of individuals in the discounted expected demand decreases and demand is increasingly dependent on the stationary market shares. As already mentioned, the latter depends only on the behavior of searchers. Since searchers are more price-elastic, prices decrease in δ. Intuitively, a price reduction leads to a sequence of demand increases through emulation that, when discounted to present, result in $\frac{1}{1-\delta\eta^*}$ extra demand, thus the price in the model with emulation is $(1-\delta\eta^*)$ times the price in ARW.

The comparative statics of the equilibrium price with respect to search costs is, perhaps, more surprising.

Proposition 2 (Comparative Statics)**.** *In the symmetric equilibrium, the following holds:*

(i) $\tilde{p}$ *is increasing in* c *when* c *is is sufficiently low.*

(ii) There exists a threshold $\bar{\delta} < 1$ *such that for all* $\delta \in (\bar{\delta}, 1)$*,* $\tilde{p}$ *decreases in* c *when* c *is sufficiently high.*

(iii) As $\delta \to 1$ *and* $c \to \bar{c}$*,* $\tilde{p}$ *converges to zero.*

As the search cost increases, first visits become more attractive, but the mass of searchers, consumers who compare both prices and thus react to price differences, decreases. Moreover, searchers have very low valuations relative to prices, so small price reductions attract their large share. The net effect on equilibrium prices depends on the weight that searchers have on total profits. As δ increases, searchers become more and more important and price elasticity increases, inducing lower prices. Indeed, as $c \to \bar{c}$ and $\delta \to 1$ prices approach zero (i.e. the marginal cost). Intuitively, when almost no consumers search, the fact that consumers follow each other induces a Bertrand-like competition for the few consumers who do search and react to prices. These consumers are extremely price sensitive (they have very low utility realizations in both firms), and their successors are unlikely to search giving further advantage to the firm that retains them. Indeed, recall that the (total) social multiplier of demand, γ, converges to infinity as $c \to \bar{c}$, leading to the marginal cost pricing for $\delta \to 1$.

This is a rather surprising result. In the ARW model without observational learning, a higher search cost relaxes price competition and results in a higher price, whereas in the model with observational learning (δ sufficiently high), exactly the opposite is true. The result is related to Armstrong and Zhou (2011) where the average equilibrium price is decreasing in search cost. The main difference is that in their model consumers observe prices, whereas here they observe each-other's purchasing behavior. Thus our contribution here is to show that observational learning is qualitatively similar to price advertising.

For illustration, consider Figure 1 that depicts equilibrium prices as a function of the search cost for uniformly distributed valuations in the baseline model with $\delta \to 1$ and $\delta \to 0$. As per Proposition 2, the price is increasing in the search cost starting at zero for both levels of δ. For $\delta \to 0$ (ARW model) prices continue to increase for the whole range of the search cost. In the model with emulation, however, as $\delta \to 1$ the price eventually decreases in c, and in the limit goes to zero as $c \to \bar{c}$.[11]

It remains to be shown that $\tilde{p}$ is indeed an equilibrium. In the Proof of Proposition 1 we establish that the profit function is locally concave for any log-concave distribution and globally concave for the uniform distribution.[12] We also show that, in any symmetric equilibrium, and for any distribution function, the lower bound of the price distribution converges to zero as $c \to \bar{c}$. This implies that expected profits vanish as search costs grow large.

Before we proceed, two points should be noted. First, we have assumed that firms commit to serve all consumers at the same price. In models of social learning, firms may use dynamic or stochastic pricing schemes in order to maximize inter-temporal profits. This assumption, however, can be justified if price-adjustment is *slow* when compared to

[11] For $G(\cdot) \sim U[0,1]$, $\bar{c} = 0.5$.

[12] The recent work on consumer search with horizontally differentiated products has focused mostly on this case. See, e.g., Armstrong and Zhou (2011)

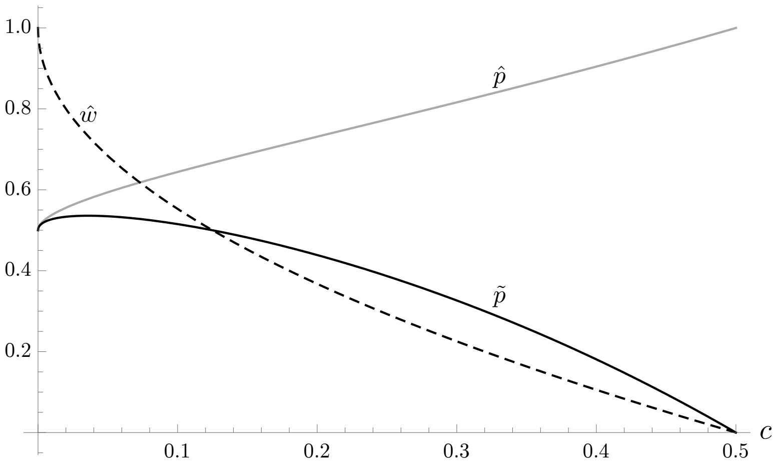

Figure 1: The equilibrium price in the baseline model for $\delta \to 1$ ($\tilde{p}$, black) and for $\delta \to 0$ ($\hat{p}$, gray). The dashed curve depicts the reservation utility $\hat{w}$.

the speed of learning among consumers. Further, as we show in Section 6, if firms are able to increase their price in the future, they have additional incentives to increase their market share today, resulting in lower equilibrium profits (see also Campbell (2013)).

Second, in this simple model, total welfare is unaffected by the introduction of emulation. In a symmetric pure strategy equilibrium, consumers extract no useful information from the purchasing decision of their predecessor and prices are irrelevant since the market is fully covered.[13] Nevertheless, consumers are better off (and firms are worse off) in the presence of emulation, the higher the discount factor. Therefore, emulation may emerge as a beneficial social norm through a process of social experimentation.

4 The Effect of Competition

We now extend the model of the previous section to the case of an oligopoly with n firms. The aim is to understand whether the social multiplier of demand is still relevant in more competitive markets. For the sake of comparison, the reader should keep in mind that in the model where the order of first visits is exogenously determined (e.g. Armstrong et al. (2009)), the effect of prominence completely disappears as the number of firms in the market grows.

For ease of exposition, we focus in the remainder of the paper on the two polar cases where $\delta \to 0$ and $\delta \to 1$. In the latter case, profits depend on the stationary market shares, while the former case is identical to ARW and serves as a useful benchmark.

We now proceed in a similar fashion as before. We first define M_1 and M_2 for this more general model. Recall that M_1 and M_2 are the probabilities that a consumer buys from firm 1 if she makes her first visit to firm 1 and any other firm, respectively. These

[13]This is not true in the model of Appendix C, where we consider an extension of our model where consumers have a zero outside option.

are given by

$$M_1(p_1, p^*) = (1 - G(\hat{w} + p_1 - p^*)) + \int_{\underline{u}}^{\hat{w}+p_1-p^*} G(u - p_1 + p^*)^{n-1} g(u)du \tag{13}$$

$$M_2(p_1, p^*) = \tfrac{(1-G(\hat{w}+p_1-p^*))}{n-1} \sum_{j=1}^{n-1} G(\hat{w})^j + \int_{\underline{u}}^{\hat{w}+p-p^*} G(u - p_1 + p^*)^{n-1} g(u)du. \tag{14}$$

M_1 is similar to the one from duopoly. M_2 accounts for the fact that a consumer may arrive at firm 1 after having visited 1 to $n-1$ other firms.

Consider $\delta \to 1$. If firm 1 sets price p_1 and if all of its rivals set p^*, the firm's stationary market share $x(p_1, p^*)$ can be computed as

$$x(p_1, p^*) = \frac{M_2(p_1, p^*)}{1 - (M_1(p_1, p^*) - M_2(p_1, p^*))},$$

which is analogous to (11).

For $\delta \to 0$, as before,

$$x(p_1, p^*) = \frac{M_1(p_1, p^*) + M_2(p_1, p^*)}{2},$$

and the equilibrium price can be computed by equating marginal revenue to zero and imposing symmetry. This yields the familiar ARW price in oligopoly

$$\hat{p}_n = \frac{1}{n\big(\int_{\underline{u}}^{\hat{w}} (n-1)G(u)^{n-2} g(u)^2\, du - G(\hat{w})^{n-1} g(\hat{w})\big) + \frac{(1-G(\hat{w})^n)g(\hat{w})}{1-G(\hat{w})}}. \tag{15}$$

Before we solve for the price for $\delta \to 1$, we compute the social spillover of demand for oligopoly as follows

$$\eta_n^* = M_1(p^*, p^*) - M_2(p^*, p^*) = 1 - \frac{nG(\hat{w}) - G(\hat{w})^n}{n-1}. \tag{16}$$

That is, the social spillover is now the difference between the probability that a consumer purchases in store 1 if she visits it first rather than later. We can therefore extend Proposition 1 to the oligopolistic competition for the special case where $\delta \to 1$.

Proposition 3. *In the symmetric equilibrium, all firms charge*

$$\tilde{p}_n = (1 - \eta_n^*) \cdot \hat{p}_n,$$

where $\eta_n^* = 1 - \frac{nG(\hat{w}) - G(\hat{w})^n}{n-1} \leq 1$, *thus for any* $c > 0$ *we have* $\tilde{p}_n < \hat{p}_n$. *Further, both* $\tilde{p}_n$ *and* $\hat{p}_n$, *but also* $\tilde{p}_n/\hat{p}_n$ *are decreasing in* n.

Notice, therefore, that, as the number of firms grows, the relative gap between prices in the model with and without emulation also grows, so that competition strengthens its effect.

It is instructive to look at both models in the limit case where $n \to \infty$. One immediate implication of large n is that the share of returning consumers vanishes. This greatly simplifies pricing in both models. As shown in Anderson and Renault (1999), the ARW model with an infinite number of firms is effectively a model of price competition with differentiated goods and so the price can be computed using the following formula

$$\hat{p}_\infty = \frac{1 - G(\hat{w})}{g(\hat{w})}. \tag{17}$$

Here a firm optimally trades off the exploitation of the captives (represented by $1 - G(\hat{w})$) and the retention of searchers ($g(\hat{w})$). In our model with emulation, however, for $n \to \infty$ the price is

$$\tilde{p}_\infty = (1 - \eta^*_\infty) \cdot \hat{p}_\infty = G(\hat{w}) \cdot \hat{p}_\infty, \tag{18}$$

and is strictly lower for any $c > 0$ since $G(\hat{w}) < 1$.

To better understand this pricing rule, note that in the ARW model with a large number of firms, a unit price increase leads to a unit increase in the reservation utility, resulting in $g(\hat{w})$ consumers leaving (and never returning). This trade-off is resolved using the price in (17). In our model, a unit increase in price induces a loss of $g(\hat{w})$ searchers, each of whom diverts an incoming consumer away from the firm, never to return. Therefore, $\eta_\infty = 1 - G(\hat{w})$. As search cost increases, by log-concavity of $1 - G(w)$, $\frac{1-G(\hat{w})}{g(\hat{w})}$ grows initially but eventually falls, leading to a bell-shaped pattern in prices.

Figure 2 illustrates equilibrium prices for $n = 2, 4$, and ∞. As shown in Proposition 3, prices decrease in n for a given search cost. Interestingly, in the limit case where $n \to \infty$, the market is perfectly competitive both when the search cost is very low and when it is very high. In the former case, this reflects the fact that consumers are willing to visit a large number of firms to find a suitable product and, therefore, each of them has negligible market power. In the latter case, however, very few consumers visit more than one firm, and thus, firms compete fiercely to attract first visits. In other words, the equilibrium price in our model is the same whether consumers make fully informed decisions (buying from the firm that offers the highest surplus) or fully uninformed decisions (buying from the first firm they visit).

To close this section, we note that the oligopoly model highlights how observational learning differs from prominence, where search order is exogenously given. In prominence models, prices differ across firms because they face different shares of returning and incoming consumers and, therefore, different demand elasticities. As shown in Armstrong et al. (2009), this difference vanishes as the number of firms grows because the share of returning consumers goes to zero.[14] As a result, the (exogenously) determined order or prominence is irrelevant for prices (although it is important for profits) with sufficiently

[14]In this class of models, consumer go back to a firm they have visited only after they visit all firms, which becomes very unlikely as n grows.

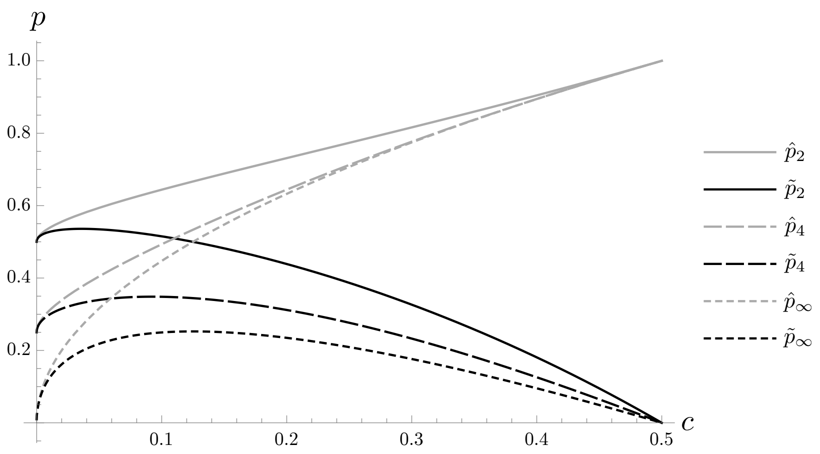

Figure 2: Equilibrium prices as a function of search cost for our model (black) and the ARW model (gray) for n equal to two (solid), four (long dash) and infinity (short dash). $G(\cdot) \sim U[0, 1]$.

many firms. In contrast, in our setting, even when the number of firms grows large, the model with emulation features a lower equilibrium price. This is because while firms do not price to returning consumers, they do internalize the social multiplier of demand that arises through emulation.

5 The Strength of Social Interaction

We now relax the assumption that every consumer has access to the purchasing decision of a single predecessor. In order to do so, we start with a reduced-form model of social interaction. We simply assume that firm 1's market share x determines its share of first visits, denoted by $z(x)$. For example, in the baseline model, each consumer observes a single predecessor, so the firm's market share and its share of first visits are equal, hence $z(x) = x$.

Using $z(x)$, firm 1's demand can be written as

$$x(p_1, p^*) = z(x(p_1, p^*))M_1(p_1, p^*) + (1 - z(x(p_1, p^*)))M_2(p_1, p^*). \tag{19}$$

Assuming that $z(x)$ is a well-behaved and differentiable function of x, we can use implicit differentiation to solve for a candidate symmetric equilibrium. First, notice that the social spillover is now defined as

$$\eta = z'(x)(M_1 - M_2).$$

The social spillover may be decomposed in two distinct terms. First, $z'(x)$ captures the marginal increase in the probability that a random consumer visits a given firm as a function of cumulative demand. This term depends only on the structure of the social interaction. Second, $M_1 - M_2$ captures the increase in demand accruing to a firm with an additional first visit. This term depends on the distribution of valuations and the search cost but it is independent of the structure of social interactions. Notice finally, that in the baseline model, $z'(x) = 1$ so that $\eta = M_1 - M_2$.

To find the putative equilibrium price, we equate firm 1's marginal revenue to zero and impose symmetry to write

$$\tilde{p} = (1 - \left(M_1^* - M_2^*\right) z'(1/2))\hat{p} = (1 - \eta^*)\hat{p}. \tag{20}$$

Again, the equilibrium price is a fraction of the ARW price. This fraction is smaller, the larger the measure of non-shoppers, i.e. those who do not search regardless of the firm they visit first, as represented by $M_1^* - M_2^* = (1 - G(\hat{w}))^2$, and the larger the marginal change in the probability of attracting one of them by increasing the market share above half, $z'(1/2)$. Given this, it is easy to see that the equilibrium price derived in Section 2 for $\delta \to 1$ obtains for $z'(1/2) = 1$. The ARW price obtains when $z'(1/2) = 0$, i.e. when the social multiplier of demand is zero.

The strength of social interactions as measured by $z'(1/2)$ has a clear threshold. If $z'(1/2)$ exceeds unity, i.e. a share of first visits increases at least one for one with the market share, then there exists sufficiently high search cost c, and thus sufficiently low $\hat{w}$, such that the putative symmetric equilibrium price is equal to the marginal cost. For learning models where $z'(1/2) < 1$, even for c so hight that $G(\hat{w}) = 0$, the marginal cost pricing does not obtain. This is because the stationary market share is never fully determined by the share of first visits, so that price elasticities are always positive and, therefore, prices exceed marginal cost.

The flexibility of the $z(x)$ formulation also allows for consumers to behave as contrarians - i.e. they may go to the store where the predecessor has not purchased. For instance, assume that each consumer observes one predecessor, and that consumers always visit first the other firm. Then $z(x) = 1 - x$, so that $z'(x) = -1$. This means that $\tilde{p} = \left(1 + (1 - G(w))^2\right) \hat{p}$, which is higher than $\hat{p}$, and $\tilde{p}$, but also $\tilde{p}/\hat{p}$, is increasing in c. In the limit where $c \to \bar{c}$, $\tilde{p} = 2\hat{p}$. As expected, contrarian behavior increases prices by discouraging price reductions with a negative multiplier of demand. This is more so when search cost is high, because losing first visits is particularly damaging with high search cost, so firms are deterred from reducing prices. This suggests that in a generalization of our model where consumer preferences are negatively correlated, prices will be higher with than without observational learning.[15]

We now proceed to rationalize $z(x)$ as the outcome of a micro-founded model of social

[15]Negative correlation results in free-riding in search, as in Section 5 where we study positive correlation of preferences. Both the contrarian search behavior and the fact that consumers search less due to free-riding lead to higher prices. Although we do not study a model with negative correlation preferences, it can be developed in the same spirit as the model of Section 7.

interaction. We assume that every consumer (except the first few) observe an odd number $(2k+1)$ of randomly drawn predecessors, where k is a nonnegative integer.[16] We interpret k as a measure of the strength of social interactions in the market. The baseline model where each consumer observes a single predecessor obtains when $k = 0$. We still assume that consumers emulate their predecessors, so that they start their search in the firm where $k + 1$ of their predecessors bought. Given that x is firm 1's market share,[17]

$$z(x) = \sum_{j=k+1}^{2k+1} \binom{2k+1}{j} x^j (1-x)^{2k+1-j}$$

Taking the derivative with respect to x and evaluating it at $x = 1/2$ yields

$$z'(1/2) = \frac{(1+2k)!}{2^{2k}(k!)^2} \equiv \zeta(k),$$

which is equal to 1 for $k = 0$ and increasing in k since

$$\zeta(k+1) = \frac{3+2k}{2(1+k)}\zeta(k) > \zeta(k).$$

As k grows, the sample of previous observations becomes more precise and, therefore, the probability that one of the firms is more popular in a given sample becomes increasingly large the larger is its market share (as captured by $\zeta(k)$.) Therefore, the price elasticity in a symmetric equilibrium explodes for a given $c > 0$. Intuitively, as consumers observe many others, the firm's share of first visits becomes increasingly sensitive to its market share, leading to an ever higher social multiplier of demand and ever lower price.

For every k there exists a measure of captive consumers $(1 - G(\hat{w}))^2$ corresponding to a search cost $\bar{c}_k$ such that the only candidate for a symmetric equilibrium price is zero, i.e. $\bar{c}_k$ solves $1 = \zeta(k)(1 - G(\hat{w}))^2$. Because $\zeta(k)$ is increasing in k and goes to infinity, $\bar{c}_k$ is decreasing in k and converges to zero as k goes to infinity. This putative outcome, however, cannot be supported as an equilibrium for any $k \geq 1$ because, in that case $\bar{c}_k < \bar{c}$, and thus even though the putative symmetric price is zero, a positive measure of consumers search. Because of the product differentiation, a deviation to a higher price yields positive profits for any search cost $c < \bar{c}$. Hence, we have the following result.

[16] Similar results obtain when the number of observed predecessors is even. In that case there are possibilities of a tie, that result in random search as in ARW. For instance, the model where consumers observe two predecessors has exactly the same equilibrium as our baseline model where they observe one predecessor.

[17] This probability is computed assuming that all draws are independent. The probability that a randomly drawn predecessor shares a predecessor with one of her successors equals the clustering coefficient of the resulting random network. This coefficient converges to zero as the number of consumers grows.

Proposition 4. *If a symmetric pure strategy equilibrium exists, both firms charge*

$$\tilde{p}_k = (1 - \eta_k^*) \cdot \hat{p},$$

where $\eta_k^* = \zeta(k)(1 - G(\hat{w}))^2 \leq 1$*, which is decreasing in* k *for any* $c > 0$*. For all* $c \in [\bar{c}_k, \bar{c})$ *no pure strategy equilibrium exists.*

Notice then that the model presented in Section 2 exhibits the weakest social interactions conducive to marginal cost pricing. Models with weaker interactions result in positive profits for any level of search cost, while models with stronger interactions result in lower prices and possibly mixed strategy equilibria. Indeed, we have the following corollary of Proposition 4 using the fact that $\lim_{k \to \infty} \bar{c}_k = 0$.

Corollary 1. *For any* $c > 0$*, there exists* K_c *sufficiently large such that for any* $k \geq K_c$*, no pure strategy equilibrium exists.*

Whenever $z'(1/2) > 1$, the putative equilibrium price becomes zero for a search cost well below the level where no consumers search. This immediately implies no pure strategy equilibrium. We are unable to characterize the mixed strategy equilibrium, but it is clear that as $c \to \bar{c}$, the equilibrium distribution of prices will converge to the marginal cost because,then, a price deviation above the marginal cost leads to zero demand, and thus zero profit.

Figure 3 illustrates our results. It shows $\tilde{p}$ for the uniform distribution and various levels of k. Solid black lines depict the equilibrium price, whereas dashed gray lines show the putative equilibrium price that is not an equilibrium because an upward price deviation exists. For $k = 0$, which corresponds to our baseline model, the putative equilibrium price actually constitutes an equilibrium for all $c < \bar{c}$. This is not the case for sufficiently high c and any $k \geq 1$. As per Proposition 4, for a given c, whenever a symmetric pure strategy equilibrium exists, $\tilde{p}$ is decreasing in k.

This result is closely related to Armstrong and Zhou (2011) who show that in a search model if consumers observe prices prior to search, there is no pure strategy equilibrium in prices. The intuition stems from the fact that first visits are valuable ($M_1 > M_2$) and, as a result, firms are willing to infinitesimally undercut each other to get all of the first visits. But since some consumers search beyond the first firm, firms have some monopoly power and thus a Bertrand outcome cannot constitute an equilibrium. Corollary 1 describes a qualitatively similar result. As k grows, each consumer observes an increasingly large sample of purchases, thus even an infinitesimal price reduction leads to a large increase in first visits because a very large sample of purchasing decisions serves as a precise statistic for the lowest price. As a a result, for a sufficiently high search cost, the only candidate for a symmetric pure strategy equilibrium requires prices to be close to zero, but due to the monopoly power firms cannot converge on a Bertrand outcome. This is not the

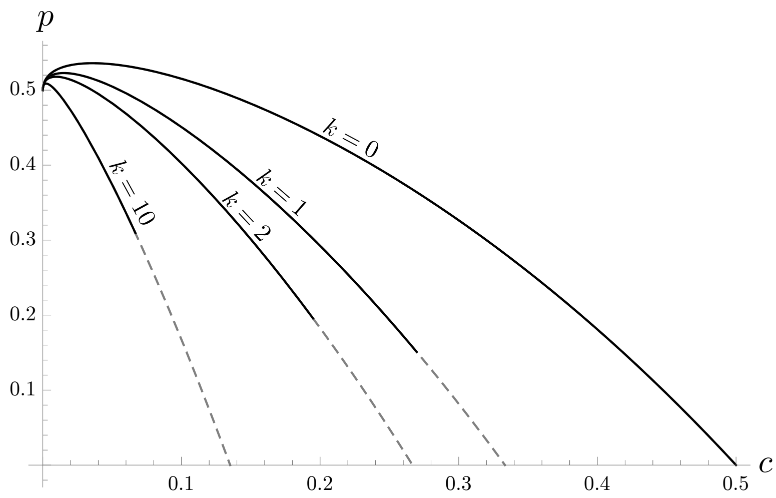

Figure 3: The symmetric equilibrium price, putative (dashed gray) and actual (solid black), for various sample sizes.

case with $k = 0$ because, in that case, zero prices obtain for $c \to \bar{c}$ where the measure of searchers vanishes and so firms' monopoly power is fully erode

We conclude this section with a brief remark over the empirical content of these results. In our model, prices depend on the size of the social multiplier which is composed of two different terms. The first term, $z'(x)$, measures the response of the share of first visits to a marginal increase in current demand. As we have shown, this effect can be directly linked to the strength of social interactions as captured by k, and can be approximated by the average degree of a social network. The second term, $M_1 - M_2$, measures the "conversion rate" of first visits into sales. In a search model, its magnitude is inversely related to the search cost. As a result, prices are decreasing in the average degree of the network and hump-shaped in search cost. This yields a novel testable implication on the effects of the Internet on prices in horizontally-differentiated markets. Namely, the expansion of online social networks (which facilitates observational learning) should result in lower prices while the increased access to product information online (which reduces the search cost) should have an ambiguous effect on prices. Moreover, we predict that the effect of more product information on prices is higher the higher is the degree of social learning.

6 Correlated Preferences

The empirical literature suggests that consumers follow others because they believe that, had they had the same information, they would have decided similarly. That is, their preferences are similar. A positive correlation in valuations across consumers introduces a new channel through which social learning may influence prices. Whereas in the model with "pure" emulation considered so far, observational learning leads to a higher price elasticity, with positively correlated valuations, this need not be true. In particular, consumers will free-ride on each other's efforts and search less than in the model where their preferences are independent. If free-riding leads to less search, firms may increase their market power and, as a result, prices may be higher.

In order to introduce these considerations, we shall return to the general model with state-dependent distributions $G(u; s)$. We assume that all consumers draw their valuations for each firm's product from one of two potential distributions, a High distribution (denoted by $G_H(u)$) and a Low distribution ($G_L(u)$).[18] Let $(s_1, s_2) \in \{H, L\}^2$ be the aggregate state. Both distributions are equally likely to realize at each firm, and these realizations are independent across firms. We assume that both distributions have the same (finite) support $[\underline{u}, \bar{u}]$. Let g_H and g_L be the corresponding continuously differentiable densities. As standard in the economics literature, we assume that the Monotone Likelihood Ratio property holds so that $\frac{g_H(x)}{g_L(x)}$ is an increasing function of x. Let $\lambda(u) = \frac{g_H(u)}{g_H(u)+g_L(u)}$ be the conditional probability of H given u. Let $G(u) = \frac{1}{2}G_H(u) + \frac{1}{2}G_L(u)$ be the unconditional distribution of valuations of a random consumer. As in the baseline model, we shall assume that $G(u)$ is log-concave. Preferences are correlated across consumers because they are drawn from the same distribution (G_H or G_L).

6.1 Consumer Behavior

The consumer who arrives in period 1 lacks any information to discriminate between firms, and, therefore, makes her first visit randomly. She buys outright if and only if $u_i - p_i > \hat{w} - p^*$, where $\hat{w}$ is computed as in the baseline model without learning. All remaining consumers hold a common prior $\nu(j)$ over the distribution of their arrival time. Therefore, all consumers use the same search rule, except the first one. They observe their predecessor buying from firm $j \in \{1, 2\}$ and expect the same price p^* at both firms before embarking on their first search. Since valuations are positively correlated, the expected surplus from visiting firm j is larger than that from visiting firm $-j$ and so the consumer should visit that firm. Upon visit, a consumer learns her utility realization for good j, u_{ij}, and the price set by firm j. The consumer will search further if and only if $u_{ij} < \tilde{w}(p_j)$, where $\tilde{w}(p_j)$ solves

$$\int_{\tilde{w}-p_j+p^*}^{\bar{u}} (u - \tilde{w} + p_j - p^*)(q(\tilde{w}; p_j)g_H(u) + (1 - q(\tilde{w}; p_j))g_L(u))du = c. \tag{21}$$

Here, $q(u; p)$ is the (posterior) probability that the distribution is High at the other firm. This probability will, in general, depend on the valuation drawn from firm j (because

[18]This simple structure ensures that the model remains tractable. Often used "truth-or-noise" cannot be used because the optimal search rule does not satisfy the reservation property. This is because when the posterior belief that a consumer holds about the probability that her predecessor searched is discontinuous in her own realization, search rules become exceedingly complex. If one looks at the extreme case of perfectly correlated preferences, then the model is the same as ARW because only the first consumer searches, and all others follow him.

preferences are correlated) and firm j's price, because the price affects firm j's market share, and thus the probability of an arrival there. In principle, $q(u;p)$ is a complicated object, since each consumer may infer from u not only how likely it is that a given distribution is realized but also her cohort, which is potentially informative about the purchasing decision of the predecessor.

In order to understand $q(u;p)$, we first look at the probability that consumer i buys from firm 1 if the state is $(ss') \in \{H,L\}^2$. Since the probability she starts her search in that firm is equal to the probability that her predecessor bought there, $x_{ss'}^{i-1}$, we have that consumer i buys with the probability

$$x_{ss'}^{i} = x_{ss'}^{i-1} M_1^{ss'} + (1 - x_{ss'}^{i-1}) M_2^{ss'}, \tag{22}$$

where

$$\begin{aligned} M_1^{ss'} &= (1 - G_s(\tilde{w}(p))) + \int_{\underline{u}}^{\tilde{w}(p)} u\ G_{s'}(u - p + p^*) g_s(u) du \\ M_2^{ss'} &= G_{s'}(w^*)(1 - G_s(w^* + p - p^*)) + \int_{\underline{u}}^{w^*+p-p^*} u\ G_{s'}(u - p + p^*) g_s(u) du, \end{aligned}$$

where $w^* = \tilde{w}(p^*)$. It is straightforward to see that if $\tilde{w}(p) > \underline{u}$, this mapping has a fixed point where market shares are stationary. In Appendix 2 we show that, provided that the number of consumers is sufficiently large, consumers' optimal search strategy is arbitrarily close to the one computed for a stationary market share distribution. In this case, dropping the time subscript, we can write a firm's market share in state ss' when it charges p while the other firm charges the equilibrium price p^* as the solution to (22) where we impose $x_{ss'}^{i} = x_{ss'}^{i-1} = x_{ss'}$. This gives

$$x_{ss'} = \frac{M_2^{ss'}}{1 - (M_1^{ss'} - M_2^{ss'})}.$$

Once we have market shares in every state (which depend on p and p^*, but also on $w(p)$ that is yet to be determined) we find $q(u;p)$ as

$$q(u;p) = \frac{x_{HH}(p)\lambda(u) + x_{LH}(p)(1-\lambda(u))}{(x_{HH}(p) + x_{HL}(p))\lambda(u) + (x_{LH}(p) + x_{LL}(p))(1-\lambda(u))} \tag{23}$$

The above formula uses market shares in every state and weights the conditional probability of H given u by these market shares using Bayes rule.

This completes the characterization of the consumer search rule. In particular, $w(p)$ is implicitly defined in (21) where $q(u;p)$ is given by (23) where $x_{ss'}$ is given in (22).

6.2 Equilibrium Conditions

In equilibrium, both firms charge the same price. Therefore, consumers' search rule in (21) can be rewritten as

$$\int_{\tilde{w}}^{\bar{u}} (u - \tilde{w})(q(\tilde{w})g_H(u) + (1 - q(\tilde{w}))g_L(u))du = c \tag{24}$$

where $q(\tilde{w}) = q(\tilde{w}; p^*)$ is the equilibrium probability that the rival firm has a High distribution given the reservation utility. Notice that $q(u) \leq q(\bar{u}) \leq \frac{1}{2}$ since observing the predecessor buying at a firm is bad news about the prospects at the rival. Hence, we have the following observation.

Proposition 5. *In a symmetric equilibrium, $\tilde{w} < \hat{w}$, and, hence, consumers free-ride on others' effort relative to the model without correlations.*

Firms' profits depend on the search rule used by consumers both on and off the equilibrium path. In particular, let $\tilde{w}'(p)$ be the implicit derivative of the reservation utility with respect to p. In the baseline model without learning, as well as in the ARW model, $\tilde{w}'(p) = 1$ so that a unit increase in the price is compensated with a unit increase in required utility. This is no longer the case with correlated preferences and observational learning. Consumers who visit a firm with a different price adjust their beliefs about the distribution of valuations in the rival firm, while keeping their beliefs about its price constant.[19] In particular, the higher the price, the less likely it is that a consumer ends up in that firm if its rival has a High realization of valuations. Thus, in general, $\tilde{w}'(p) < 1$.

Firm 1's expected demand when it charges p_1 while its competitor charges p^* and consumers use reservation utility function $\tilde{w}(p)$ is the average of demands over all possible states. Because states between firms are independent, and equally likely, the demand is given by:

$$x(p_1, p^*) = \frac{1}{4} \sum_{ss'} x_{ss'}(p_1, p^*).$$

Here $x(p_1, p^*)$ implicitly depends on $w(p)$ and thus on consumer search rule. In the symmetric equilibrium, the price satisfies

$$\sum_{ss'} x_{ss'}(p^*, p^*) + p^* \sum_{ss'} \frac{\partial x_{ss'}(p^*, p^*)}{\partial p_1} = 0.$$

Once we impose $\sum_{ss'} x_{ss'}(p^*, p^*) = \frac{1}{2}$ because of symmetry, the equilibrium pricing rule simplifies to

$$\frac{1}{2} + p^* \sum_{ss'} \frac{\partial x_{ss'}(p, p^*)}{\partial p_1} = 0. \tag{25}$$

[19] The assumption of passive beliefs is common in the literature. For a discussion see Janssen and Shelegia (2014).

In order to solve the model, then, we need to characterize the derivative of demand with respect to the price in every state. This derivative crucially depends on the response of consumers to a marginal change in price by one firm. In turn, this response depends on the poster estimate of state H at the other firm, which in turn depends on the demand in each state. No closed form solutions exist, but one can obtain a numerical solution by finding $\tilde{w}'(p)$ using the implicit function theorem, and imposing symmetry everywhere. In what follows we provide the numerical solution for the Uniform distribution and analytically characterize the limit when the mass of searchers vanishes for any distribution.

6.3 Equilibrium Characterization

If $\lambda(\underline{u}) = 0$, upon drawing $\underline{u}$ at firm j, the probability that the rival firm has a High distribution is arbitrarily small and so $x_{L,H}(p^*, p^*) = 0$. Thus, $q(\underline{u}) = 0$ and we define

$$\bar{c}_L = \int_{\underline{u}}^{\bar{u}} u g_L(u) du$$

to be the largest search cost such that there is search in equilibrium. For every $c < \bar{c}_L$, a firm offering its product for a price $p^* - (\bar{c}_L - c)$ retains all its incoming consumers and, therefore, obtains a stationary market share equal to one. This puts an upper bound on the equilibrium price. The following proposition establishes that this upper bound converges to zero as $c \to \bar{c}_L$.

Proposition 6. *As $c \to \bar{c}_L$, $p^* \to 0$ for every (G_H, G_L) such that $\lambda u = 0$.*

As the search cost increases, retaining consumers becomes increasingly cheap. Market shares are fully determined by the relative proportion of searchers who buy in each store, so that a deviating firm can attract all consumers by offering an increasingly small discount, leading to Bertrand competition and marginal-cost pricing.

6.4 Comparative Statics

We discuss now the effects of learning on prices. In order to do so, we shall compare the equilibrium prices if consumers' preferences are correlated ($G_H \neq G_L$) with the case where consumer preferences are independent ($G_H = G_L = G$). Let, as before, $\tilde{p}$ be the price with uncorrelated preferences and $\tilde{p}_{cor}$ be the equilibrium price if preferences are correlated. The following corollary shows that learning may put even further pressure on prices.

Corollary 2. *Take an (unconditional) distribution G with associated distributions G_H and G_L. There exists $\epsilon > 0$, such that for $c \in (\bar{c}_L - \epsilon, \bar{c}_L)$, $\tilde{p}_{cor} < \tilde{p}$.*

That is, if the purchasing decision of a predecessor is informative about the distribution of utilities of her successor, the equilibrium price elasticity increases and, thus, prices decrease. This is because consumers whose valuation for the variety they sample first is low become increasingly pessimistic about their prospects at the other firm, the higher is the correlation across consumers, and, therefore, they are less inclined to search. Since the elasticity of demand decreases in the proportion of searchers when the latter is sufficiently high, the result follows.

The following figures illustrate our results for $G(\cdot) \sim U[0,1]$. In this example G_L is a triangular distribution on $[0,1]$ with mode 0 while G_H is a triangular distribution on $[0,1]$ with mode 1. As required, the mixture of the two is uniform on $[0,1]$ and G_H FOSD G_L. Also, $\lambda(\underline{u}) = 0$.

Figures 1 and 2 show prices and reservation utilities in various models. As expected,

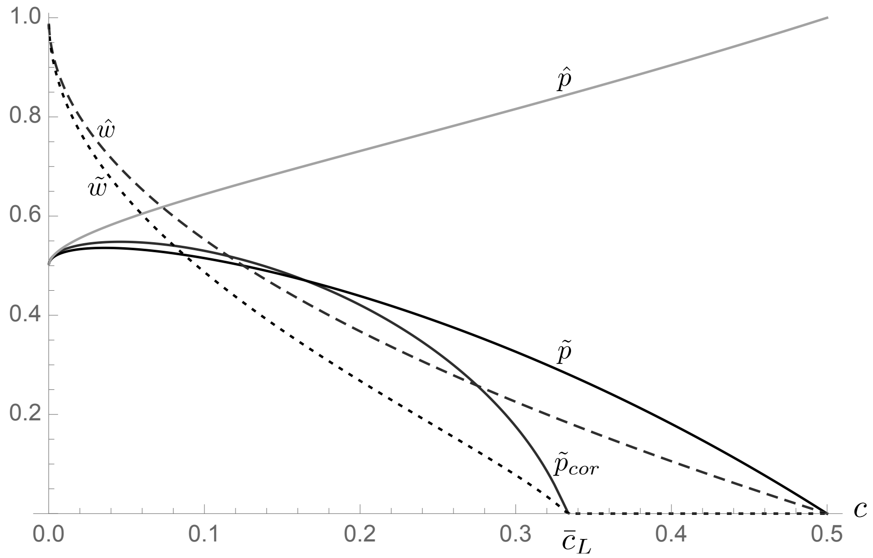

Figure 4: Prices in our model with correlated ($\tilde{p}_{cor}$), uncorrelated ($\tilde{p}$) and the ARW ($\hat{p}$, gray) for triangular distributions.

the price is lower in the model with pure emulation ($\tilde{p}$) than in the ARW model ($\hat{p}$). The equilibrium price in the ARW model is increasing in c and reaches 1 at $c = 1/2$. This is where even a consumer who draws 0 at the first firm refuses to search further. In our model with correlation of preferences, as predicted by Proposition 6, because $\lambda(0) = 0$, the price converges to zero when $c \to \bar{c}_L = 1/3$. In our model with pure emulation and the same unconditional distribution, the price converges to zero at $c = 1/2$, as implied by Proposition 2. Intuition for both results is simple. When $c = 1/3$ in the model with learning, a consumer who is sure that utilities from the other firm are drawn from G_L will not search. But this is what happens in equilibrium when a consumer draws utility close to 0 - she reasons that because the distribution at the current firm is almost surely G_L, the fact that she came here indicates that the distribution is also G_L at the other firm, or otherwise she would have almost surely ended up at the other firm for the first visit.

Thus even though u close to $\underline{u}$ is bad news about the current firm, it is also bad news about the next firm. Because the search cost at the current firm is sunk, no consumer searches beyond the first firm.

Figure 5 shows $\tilde{w}'(\tilde{p}_{cor})$ as a function of c. For $c \to 0$, the derivative of reservation utility approaches 1. This is because when almost all consumers search, the price contains almost no information and so the reservation utility matches it one for one. For all $c \in (0, \underline{c}_L)$, $\tilde{w}'(\tilde{p}_{cor})$ is less than 1, reflecting the informational content of price deviations.

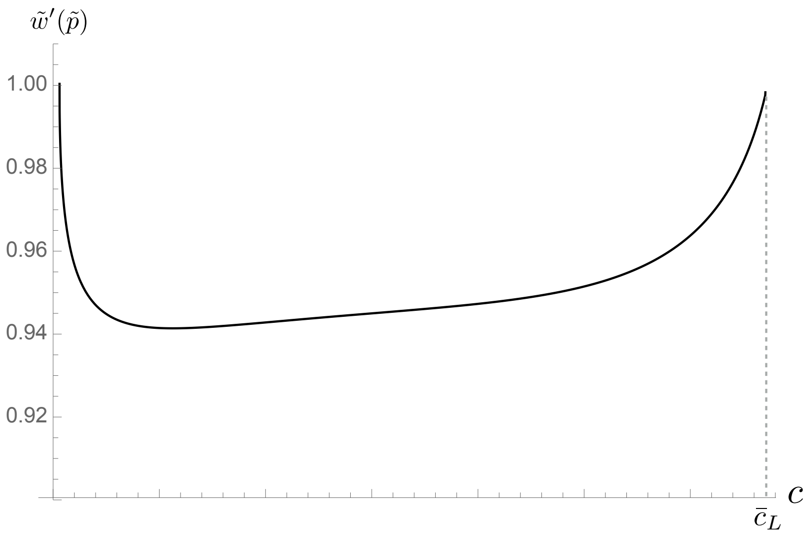

Figure 5: The derivative of reservation utility with respect to p in our model for triangular distributions.

Figure 6 illustrates equilibrium demands in various states as a function of c. In symmetric states (LL and HH), demand for both firms is equal to 0.5 and is independent of the search cost. Matters are more interesting in asymmetric states. As the search cost increases, demand for a firm with low distribution facing a firm with high distribution shrinks (the opposite is true for high distribution vs low distribution). Naturally, when search cost is low some consumers buy at the firm where utility is drawn from G_L, and so the same proportion of consumers visits this firm first. As $c \to 1/3$, search vanishes and almost all consumers are steered toward and stay at the firm with the high distribution. Thus, a *herd* forms whenever (almost) all consumers make the first visit to a given firm and buy outright.[20]

These results can be linked to those obtained in Bar-Isaac et al. (2011) who show that a reduction in search costs leads to an increase in demand for "superstars" (accompanied by a simultaneous increase in the demand for niche products.) In contrast, with observational learning the superstar effect increases in search costs because individual experimentation becomes increasingly costly.

[20]In the classical definition, a herd occurs when individuals disregard their private information and follow the crowd. In our simple model, individuals do not possess prior information.

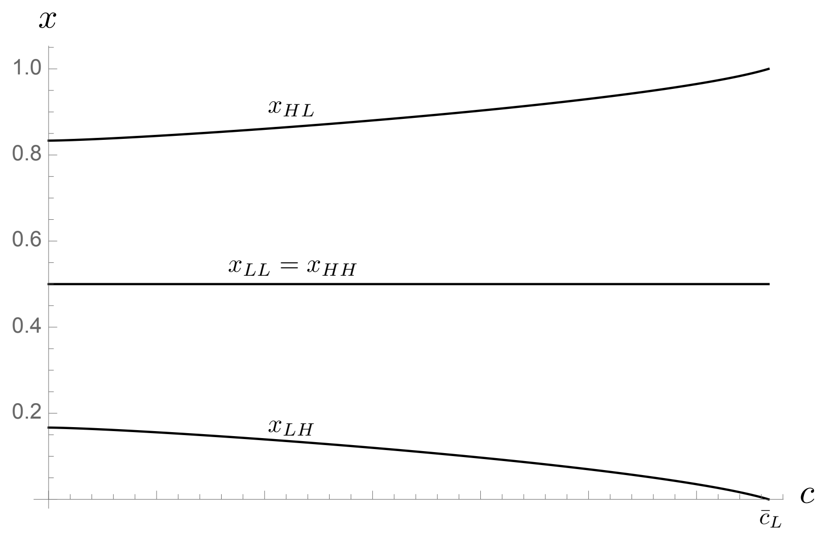

Figure 6: Equilibrium demand in various states for triangular distributions.

Finally, a brief remark about the welfare properties of our model is in order. In Appendix D we show that observational learning results in higher aggregate welfare and, more surprisingly, total welfare may not be monotone in search costs. In particular, if search cost is sufficiently close to $\bar{c}_L$, an increase in search cost may induce higher total welfare. The reason is as follows: whenever search costs are sufficiently high, market shares are arbitrarily informative and searchers are more likely to move from a high-distribution firm to a low-distribution firm than viceversa. Therefore, they impose a negative externality on non-searchers. As search costs increase, the measure of searchers decreases and welfare of non-searchers increase. This indirect effect dominates the direct effect of higher search costs for some distribution functions (e.g. uniform).

7 Dynamic Pricing

We have assumed throughout that firms choose prices once and for all. While this assumption is rather crude, it captures well the strategic interaction as long as price adjustment is slow when compared with learning. Nevertheless, if firms are able to adjust price frequently (relative to learning), then one may wonder if the mechanism that we have described is still at work. As we show next, not only do prices remain low even if firms adjust prices frequently, but actually prices *decrease even further* in the case of frequent adjustment. If firms adjust prices frequently, they have an extra incentive to undercut because they can harvest their market-share gains at higher prices in the future. This logic is also present in recent models of dynamic pricing with switching costs.[21] The differences are, however, stark. First, our baseline model does not provide opportunities for harvesting since prices are fixed once. In such a case, switching costs would be irrelevant

[21]See Cabral (2013); Doganoglu (2010) and Rhodes (2014)

for prices while search costs are not. Second, in our model prices converge to marginal costs as search costs increase while in the switching cost model prices are always bounded away from the competitive outcome. Third, switching costs have been shown to reduce prices only if firms can price discriminate across consumers (Cabral (2013)) or when firms are infinitely lived and sufficiently patient (Rhodes (2014)). Neither of these assumptions is required in our model.

To formalize the argument, we shall assume that utilities are uniformly distributed on some closed interval $[a, b]$ and that preferences are independent across consumers. We will use this assumption to invoke Armstrong and Zhou (2011) result that firms with higher share of first visits charge lower prices. This puts a lower bound on the continuation payoff of a firm who has deviated in the past to a low price that resulted in a higher market share.

Consider a modification of our model in which two firms can adjust prices every period, everything else is as in the baseline model from before. Then firms' strategies are a sequence of pricing functions as a function of the history of price realizations, $\{p_{j,t}(h_t)\}$.[22] A Symmetric Stationary Markov Perfect Equilibrium is a sequence of price functions such that if a history up to period t involves a market share x_t, then $p_{j,t}(h_t) = p(x)$. In this simple class of equilibria, the problem of the consumer is unchanged.[23] Let p^* be the stationary price so that $p(1/2) = p^*$. Hence, let $\hat{w}$ be the reservation utility and let $x^t(p_{1,t}, p_{2,t}, x^{t-1})$ be the market share of firm j in period t given prices and the market-share in the previous period. It follows that

$$x^t = M_2(p_{1,t}, p_{2,t}) + x^{t-1}(M_1(p_{1,t}, p_{2,t}) - M_2(p_{1,t}, p_{2,t}))$$

Notice that, following a downward price deviation by firm 1 in period t, $x^t > 1/2$. Since market share in period t equals first visits in period $t+1$, the share of first-visits that firm 1 receives is relatively large (as a function of her total demand) when compared to firm 2. Since valuations are uniformly distributed, we can use the result in Armstrong and Zhou (2011) to establish the following Lemma.

Lemma 1. *Suppose that $h_t = \{(p^*, p^*), .., (p, p^*)\}$ with $p < p^*$. Then, in every period $t' > t$, $p_{i,t} \leq p^* \leq p_{j,t}$.*

Following a marginal downward price deviation, a firm makes her rival less competitive because it faces a lower price elasticity. This means that the worst possible path of future prices for the rival firm following such a deviation is p^* in every period. A lower bound

[22] Firms cannot condition on demand realizations. One could simply assume that there is a large number of consumers per cohort, in which case this assumption is without loss of generality.

[23] Formally, we assume that if the consumer observes a price which is different from the equilibrium price, she assumes that this is the first period in which a deviation occurs and, therefore, expects the rival firm to price at the stationary price.

in the continuation value is, thus, $V(p^*, p^*; x^t) = p^* Q(p^*, p^*; x^t)$. We can now state the following result.

Proposition 7. *If a SSMPE exists, the equilibrium price* $p^* \leq \tilde{p}$

Recall that the multiplier of demand captured the effect of an additional purchase on the stream of future demand. That is precisely the margin that a firm can affect by changing her price in period t and pricing at p^* from there on. Therefore, if $V(p, p^*, x^t) = V(p^*, p^*, x^t)$, then $p^* = \tilde{p}$. In general, however, $V(p, p^*, x^t) > V(p^*, p^*, x^t)$ for $p < p^*$ since the rival firm would optimally charge higher prices and so equilibrium prices and profits decrease.

As we can see, then, even when prices can be adjusted very frequently, the price reducing effect of observational learning is still present, and in fact is further strengthened by the possibility of profit harvesting. This seems to be in contrast to the intuition that a firm would eventually increase its price once it has built up a sufficiently large market share. This does not happen in our model because learning does not "slow down" in the sense that current consumers do not become less important for future learning than earlier consumers were for current learning. As a result, pricing incentives do not change over time, and are always affected by the ongoing learning.

In an alternative model where learning slows down over time, one can obtain different pricing results that are more closely aligned with intuitions from switching cost models. To this end, consider a simple extension of our baseline model involving two phases. Consumers in the introductory phase do not see any of their predecessors and, thus, behave just as in ARW. Consumers in the mature phase observe exactly one consumer from the introductory phase. This means that learning is only relevant for the first period price. In a symmetric equilibrium, the first period market share in the first phase will be equal to one half and, in the second period, the resulting equilibrium outcome will be the ARW (static) price. Out of equilibrium, however, as shown in Armstrong and Zhou (2011), profits in the second phase will be increasing in the first phase market share. Thus, the incentive to undercut in the first phase at any price is higher than in the static model, and lower prices result, leading to an increasing schedule of prices.

In light of these results, we may conclude that social learning puts downward pressure on prices. As explained, if social learning becomes less important over time, we should expect prices to start low and increase as learning slows down. In our baseline model, however, social interactions do not decay over time, and, therefore, dynamic pricing does not necessarily result in higher prices.

8 Conclusion

In this paper, our aim was to introduce observational learning into a standard consumer search model with horizontally differentiated products. We have developed a simple and tractable extension of the ARW framework, whereby consumers observe others and visit first the firm where their predecessor has purchased. In the simplest version of this model, this amounts to a behavioral modification of the standard framework that, nonetheless, has strong implications for the equilibrium prices. We further developed a rationale for such behavior through correlated preferences and showed that the qualitative implications of the original model remain unchanged.

In order to highlight the core mechanism through the social multiplier of demand, we have abstracted from a number of important issues. First, firms are restricted to serve all incoming consumers at the same price. This assumption, however, can be justified if price-adjustment is *slow* when compared to the speed of learning among consumers. Second, in our model firms choose only prices, so we have abstracted from such important choice variables as product design or advertising.[24] For instance, niche products induce a different learning process than mass products and this is likely to have implications for prices (see Chen et al. (2011)). Similarly, advertising spurs consumer awareness for new products and, therefore, may soften the price-elasticity of the residual demand. On the other hand, many advertising decisions are intrinsically linked to consumers' interactions. For instance, most products have a recognizable logo stamped on them so that other consumers may identify it. Thus, firms may actively affect the speed of learning about new products by advertising and branding. We believe that these features can be incorporated within our basic framework.

Finally, our model provides a simple framework to evaluate the effect of the Internet on consumer search markets. On the one hand, online platforms facilitate the comparison of different products and, thus, reduce search costs. On the other hand, online social networks foster social interactions. As it turns out, these two features have opposite effects on the size of the social spillover and, so, provide different implications for equilibrium prices. Therefore, a potential avenue for future research is to empirically separate the effects of these two sources of information on prices and demand elasticities in consumer search markets.

[24]See Bar-Isaac et al. (2011) and Haan and Moraga-Gonzalez (2011) who build on the ARW model to endogenize product design and to study informative advertising, respectively.

References

Ali, S. Nageeb, "Social Learning with Endogenous Information," Technical Report, Mimeo 2014.

Anderson, Simon P. and Regis Renault, "Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model," *RAND Journal of Economics*, Winter 1999, *30* (4), 719–735.

Armstrong, Mark and Jidong Zhou, "Paying for Prominence," *The Economic Journal*, 2011, *121* (556), F368–F395.

_ , John Vickers, and Jidong Zhou, "Prominence and Consumer Search," *The RAND Journal of Economics*, 2009, *40* (2), 209–233.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, August 1992, *107* (3), 797–817.

Bar-Isaac, Heski, Guillermo Caruana, and Vicente Cuñat, "Search, Design and Market Structure," *American Economic Review*, 2011, *102* (2), 1140–60.

Becker, Gary S, "A note on restaurant pricing and other examples of social influences on price," *Journal of Political Economy*, 1991, *99* (5), 1109.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change in Informational Cascades," *Journal of Political Economy*, October 1992, *100* (5), 992–1026.

Bose, Subir, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund, "Dynamic monopoly pricing and herding," *The RAND Journal of Economics*, 2006, *37* (4), 910–928.

Cabral, Luis, "Dynamic pricing in customer markets with switching costs," *New York University*, 2013.

Cai, Hongbin, Yuyu Chen, and Hanming Fang, "Observational Learning: Evidence from a Randomized Natural Field Experiment," *American Economic Review*, 2009, *99* (3), 864–882.

Caminal, Ramon and Xavier Vives, "Why Market Shares Matter: An Information-Based Theory," *RAND Journal of Economics*, Summer 1996, *27* (2), 221–239.

Campbell, Arthur, "Word-of-mouth communication and percolation in social networks," *The American Economic Review*, 2013, *103* (6), 2466–2498.

Chen, Yubo, Qi Wang, and Jinhong Xie, "Online social interactions: A natural experiment on word of mouth versus observational learning," *Journal of Marketing Research*, 2011, *48* (2), 238–254.

Chuhay, Roman, "Marketing via Friends: Strategic Diffusion of Information in Social Networks with Homophily," Working Papers 2010.118, Fondazione Eni Enrico Mattei September 2010.

Diamond, Peter A., "A Model of Price Adjustment," *Journal of Economic Theory*, June 1971, *3* (2), 156–168.

Doganoglu, Toker, "Switching costs, experience goods and dynamic price competition," *Quantitative Marketing and Economics*, 2010, *8* (2), 167–205.

Galeotti, Andrea, "Talking, Searching, and Pricing," *International Economic Review*, November 2010, *51* (4), 1159–1174.

Glaeser, Edward L, Bruce I Sacerdote, and Jose A Scheinkman, "The social multiplier," *Journal of the European Economic Association*, 2003, *1* (2-3), 345–353.

Haan, Marco A. and Jose L. Moraga-Gonzalez, "Advertising for Attention in a Consumer Search Model," *Economic Journal*, 05 2011, *121* (552), 552–579.

Hendricks, Kenneth, Alan Sorensen, and Thomas Wiseman, "Observational learning and demand for search goods," *American Economic Journal: Microeconomics*, 2012, *4* (1), 1–31.

Janssen, Maarten C.W. and Sandro Shelegia, "Beliefs, Market Size and Consumer Search," 2014.

Kircher, Philipp and Andrew Postlewaite, "Strategic firms and endogenous consumer emulation," *The Quarterly Journal of Economics*, 2008, *123* (2), 621–661.

Kovac, Eugen and Robert C. Schmidt, "Market share dynamics in a duopoly model with word-of-mouth communication," *Games and Economic Behavior*, 2014, *83*, 178–206.

Manski, Charles F, "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

Miller, Grant and A. Mushfiq Mobarak, "Learning About New Technologies Through Social Networks: Experimental Evidence on Nontraditional Stoves in Bangladesh," *Marketing Science*, 2015.

Möbius, Markus and Tanya Rosenblat, "Social Learning in Economics," *Annual Review of Economics*, 2014, *6*, 827–847.

Monzón, Ignacio and Michael Rapp, "Observational learning with position uncertainty," *Journal of Economic Theory*, 2014, *154* (0), 375 – 402.

Moretti, Enrico, "Social learning and peer effects in consumption: Evidence from movie sales," *The Review of Economic Studies*, 2011, *78* (1), 356–393.

Mueller-Frank, Manuel and Mallesh M. Pai, "Social Learning with Costly Search," *American Economic Journal: Microeconomics.* Forthcoming.

Rhodes, Andrew, "Re-examining the effects of switching costs," *Economic Theory*, September 2014, *57* (1), 161–194.

Smith, Lones and Peter Sørensen, "Pathological outcomes of observational learning," *Econometrica*, 2000, *68* (2), 371–398.

Wolinsky, Asher, "True Monopolistic Competition as a Result of Imperfect Information," *Quarterly Journal of Economics*, August 1986, *101* (3), 493–511.

Zhang, Juanjuan, "The sound of silence: Observational learning in the US kidney market," *Marketing Science*, 2010, *29* (2), 315–335.

Appendix A: Proofs

Proof of Proposition 1

Using the demand in (9), the FOC for firm 1's profit maximization after imposing that in equilibrium $M_1^* + M_2^* = 1$ is

$$\frac{1}{2} + p\frac{(\partial M_1/\partial p_1 + \partial M_2/\partial p_1)}{2(1-\delta(M_1 - M_2))} = 0.$$

In the symmetric equilibrium $M_1 + M_2 = 1$, $Q_1 = 1/2$, $\partial M_1/\partial p_1 = -\int_{\underline{u}}^{\hat{w}} g(u)^2\,du - (1 - G(\hat{w}))g(\hat{w})$ and $\partial M_2/\partial p_1 = -\int_{\underline{u}}^{\hat{w}} g(u)^2\,du$. So we can solve for $\tilde{p}$ from the above to obtain

$$\tilde{p} = \frac{1-\delta\eta^*}{2\int_{\underline{u}}^{\hat{w}} g(u)^2\,du + (1-G(\hat{w}))g(\hat{w})}.$$

It is well known that in the ARW model the price is

$$\hat{p} = \frac{1}{2\int_{\underline{u}}^{\hat{w}} g(u)^2\,du + (1-G(\hat{w}))g(\hat{w})},$$

thus $\tilde{p} = (1-\delta\eta^*)\hat{p}$.

Even though the candidate equilibrium price is clearly unique, it still remains to be shown that no firm wants to deviate from the candidate equilibrium. We shall first show that the SOC is satisfied locally for any $G(\cdot)$ that satisfies log-concavity of $1-G(u)$. Then we show that the profit function is globally concave for the special case where $G(u)$ is uniform.

We take the SOC for firm 1 and evaluate it at $p_1 = \tilde{p} = (1-\delta(1-G(\hat{w})))\hat{p}$.

$$\begin{aligned}
\left.\frac{\partial^2 \Pi_1(p_1,\tilde{p})}{\partial p_1^2}\right|_{p_1=\tilde{p}} &= \Big[(1-G(\hat{w}))(\delta + \delta(G(\hat{w})-2)G(\hat{w}) - 1)g(\hat{w}) \\
&- 4(\delta-2)(G(\hat{w})-1)g(\hat{w})\int_{\hat{w}}^{\hat{w}} g(u)^2\,du - 8\left(\int_{\hat{w}}^{\hat{w}} g(u)^2\,du\right)^2 \\
&+ g(\hat{w})^2(\delta + \delta(G(\hat{w})-2)G(\hat{w}) - 1) + (3(\delta-1) \\
&+ (3\delta-2)(G(\hat{w})-2)G(\hat{w}))g(\hat{w})^2\Big] \Big/ \Big[2(1-\delta+\delta(2-G(\hat{w}))G(\hat{w})) \times \\
&\times \left(2\int_{\underline{u}}^{\hat{w}} g(u)^2\,du + (1-G(\hat{w}))g(\hat{w})\right)\Big].
\end{aligned}$$

The denominator is positive. The derivative of the numerator with respect to δ is

$$(1-G(\hat{w}))\left(4g(\hat{w})\left(\int_{\underline{u}}^{\hat{w}} g(u)^2\,du\right) + (1-G(\hat{w}))g(\underline{u})^2 + (1-G(\hat{w}))\left((1-G(\hat{w}))g(\hat{w}) + 3g(\hat{w})^2\right)\right),$$

which is positive because $(1-G(\hat{w}))g(\hat{w})+3g(\hat{w})^2>0$ by the log-concavity of $1-G(u)$. Thus the SOC is increasing in δ. We then consider the numerator of SOC for $\delta=1$ to obtain

$$-8\left(\int_{\underline{u}}^{\hat{w}} g(u)^2\,du\right)^2-4(1-G(\hat{w}))g(\hat{w})\int_{\underline{u}}^{\hat{w}} g(u)^2\,du-(2-G(\hat{w}))G(\hat{w})\times$$
$$\times\left[g(\underline{u})^2+g(\hat{w})^2+(1-G(\hat{w}))g'(\hat{w})\right].$$

This is negative by the log-concavity of $1-G(u)$ which implies that the expression in square brackets is positive. This proves that $\frac{\partial^2\Pi_1(p_1,\tilde{p})}{\partial p_1^2}\Big|_{p_1=\tilde{p}}<0$, thus $p_1=\tilde{p}$ is the local maximizer of firm 1's profits.

Now we show that $\tilde{p}$ is the global maximizer for the special case where $G(\cdot)\sim U(0,1)$. In that case $\tilde{p}=\frac{1-\delta(1-\hat{w})^2}{1+\hat{w}}$.

There are three cases depending on p_1. For $p_1\in[\tilde{p},1+\tilde{p}-\hat{w})$, after some manipulations, the second derivative of profits can be written as

$$\frac{\partial^2\Pi_1(p_1,\tilde{p})}{\partial p_1^2}=-\frac{(\hat{w}+1)^3(\delta(\hat{w}-1)+\hat{w}+1)\left(\delta(\hat{w}-1)^2-1\right)^2}{\left(\delta(1-\hat{w})\left(p\hat{w}+p+\hat{w}^2-2\right)+\delta^2(1-\hat{w})^3+\hat{w}+1\right)^3}.$$

The denominator is clearly positive since $1+\hat{w}-\delta(1-\hat{w})>0$. The numerator is increasing in p_1, thus it reaches its minimum at $p_1=\tilde{p}=\frac{1-\delta(1-\hat{w})^2}{1+\hat{w}}$. Plugging this value into the numerator, and simplifying yields $(1+\hat{w})^3\left(1-\delta(1-\hat{w})^2\right)^3$, which is positive, therefore the numerator is positive for all p_1 in the interval. Thus, the second derivative of firm 1's profit is negative for all p_1 in this interval.

Now consider $p_1\geq 1+\tilde{p}-\hat{w}$, which means that p_1 is so high that all consumers who visit firm 1 first go on to search firm 2. In this case,

$$\frac{\partial^2\Pi_1(p_1,\tilde{p})}{\partial p_1^2}=-2.$$

Now consider $\tilde{p}-\hat{w}<p_1<\tilde{p}$. It can be shown that in this case the second derivative of profits has the opposite sign of

$$f=\delta^2p_1^3(\hat{w}-1)^2(\hat{w}+1)^3+3\delta p_1^2(\hat{w}-1)(\hat{w}+1)^2(\delta(\hat{w}-1)+\hat{w}+1)\left(\delta(\hat{w}-1)^2-1\right)$$
$$+3p_1(\hat{w}+1)(\delta(\hat{w}-1)+\hat{w}+1)^2\left(\delta(\hat{w}-1)^2-1\right)^2+\left(\delta(\hat{w}-1)^2-1\right)^2$$
$$\times\left(4\delta+\delta^2\left(3\hat{w}^2-4\right)(\hat{w}-1)^2+\hat{w}^2(\hat{w}+2)^2+\delta^3(\hat{w}-1)^4+\delta\hat{w}(\hat{w}(\hat{w}(3\hat{w}+2)-7)-2)-1\right),$$

which is increasing in p_1, thus its minimum is reached at the lower bound of p_1. There are two cases depending on δ. For $\delta<\frac{1-\hat{w}(1+\hat{w})}{(1-\hat{w})^2}$ (in which case the $\tilde{p}-\hat{w}>0$) so that the lower bound on p_1 is $\tilde{p} - \hat{w}$. Plugging it in simplifies f as

$$\begin{aligned} f &= \delta^2(\hat{w}-2)\left(\hat{w}^2-1\right)^2(\hat{w}(2\hat{w}+3)-3)-(\hat{w}+1)^2\left(2\hat{w}^2+\hat{w}-2\right)+2\delta^3(\hat{w}+1)^2(\hat{w}-1)^5 \\ &+ \delta(\hat{w}+1)^2((\hat{w}-6)\hat{w}(\hat{w}+1)+6)(\hat{w}-1). \end{aligned}$$

The above is minimized at $\hat{w} = 0$ and $\delta = 1$ where it is equal to 0, thus it is negative for all other cases.

The second case is where $\delta \geq \frac{1-\hat{w}(1+\hat{w})}{(1-\hat{w})^2}$, so that the lower bound on p_1 is 0. Plugging 0 into the expression for f gives

$$\begin{aligned} f &= \left(4\delta+\delta^2\left(3\hat{w}^2-4\right)(\hat{w}-1)^2+\hat{w}^2(\hat{w}+2)^2+\delta^3(\hat{w}-1)^4+\delta\hat{w}(\hat{w}(\hat{w}(3\hat{w}+2)-7)-2)-1\right) \\ &\times \left(1-\delta(1-\hat{w})^2\right)^2. \end{aligned}$$

The term on the second line is positive. The term on the first line is minimized at $\hat{w} = 0$ and $\delta = 0$, where it equals zero, thus otherwise it is positive. This proves that f is positive in both cases, thus the second derivative of profits is negative for $\tilde{p} - \hat{w} < p_1 < \tilde{p}$

Finally, consider $p_1 < \tilde{p} - \hat{w}$, in which case p_1 is so low that all consumer who arrives to firm 1 buy from it. In this case

$$\frac{\partial^2 \Pi_1(p_1, \tilde{p})}{\partial p_1^2} = 0,$$

but the profit function is linear and increasing in p_1, thus the maximizer can never be in this interval.

This establishes that for the uniform distribution of valuations, firm 1's profit is concave in p_1 and so $\tilde{p}$ is its global maximizer. Thus $p_i = \tilde{p}$ for $i = 1, 2$ constitutes the unique symmetric pure strategy equilibrium.

We finally prove that for any distribution, prices converge to zero as $\delta \to 1$ and $c \to \bar{c}$ in any symmetric equilibrium. Let $\underline{p}$ be the lowest price in the support. A firm charging $\underline{p}$ obtains an expected market share $x(\underline{p})$. First, if the lowest upper bound of $x(\underline{p})$ is 1, then, it must be that $x(\underline{p}) \to 1$ for $c \to \bar{c}$, since firms retain market power for smaller search costs. If indeed $x(\underline{p}) \to 1$ as $c \to \bar{c}$, randomization implies that either $\underline{p} \to 0$ proving the result or the price distribution becomes degenerate. But the preceding argument shows that if a pure-strategy equilibrium exists it must have zero prices. Thus, $\underline{p} \to 0$ also in that case. Finally, assume that $x(\underline{p}) \leq \bar{x} < 1$ and consider a deviation to a price $\underline{p} - (\bar{c} - c)$. Clearly $w(p) = \underline{u}$ and, therefore, $x(p) = 1$. Hence, if $x(\underline{p})\underline{p} < \underline{p} - (\bar{c} - c)$, this deviation is profitable. Hence, in any equilibrium $\underline{p}(1 - \bar{x}) \leq \bar{c} - c$. Since $\bar{x}$ is independent of c, it must be that $\underline{p}$ converges to zero, and randomization implies that the expected price also converges to zero.

Proof of Proposition 2

The last part follows directly from $\eta^* = (1 - G(\hat{w}))^*$ which is positive and less than 1 for all $c > 0$. For $c \to \bar{c}$, $G(\hat{w}) \to 1$ and so $\eta^* \to 0$. Since the price in Wolinsky is bounded, $\tilde{p} \to 0$ as cost increases. To see that the price increases in the search cost for low enough c, notice that the derivative of $\tilde{p}$ with respect to $\hat{w}$ (which is decreasing in c) at $\hat{w} = \bar{u}$ is

$$\frac{\partial p}{\partial \hat{\omega}}\Big|_{\hat{\omega}=\bar{u}} = -\left(\frac{g(\bar{u})}{2\int_{\underline{u}}^{\bar{u}} g(u)du}\right)^2$$

which is negative for any distribution and any δ. On the other hand, at the lower bound, the derivative is

$$\frac{\partial p}{\partial \hat{\omega}}\Big|_{\hat{\omega}=\underline{u}} = 3\delta - 1 - \frac{(1-\delta)g(\underline{u})}{g(\underline{u})^2},$$

which is equal to 2 at $\delta = 1$. It follows directly that there exists $\bar{\delta} < 1$ such that for all $\delta > \bar{\delta}$ the derivative is positive, thus the price is decreasing in c.

Proof of Proposition 3

In order to show that the price is lower with emulation we need that $\frac{nG(\hat{w})-G(\hat{w})^n}{(n-1)} < 1$. To see this notice that $\frac{nG(\hat{w})-G(\hat{w})^n}{n-1} < 1$ is equivalent to

$$n > \frac{1 - G^n(\hat{w})}{1 - G(\hat{w})}.$$

This always holds for $c > 0$ because $\frac{1-G^n(\hat{w})}{1-G(\hat{w})} = \sum_{j=0}^{n-1} G^j(\hat{w}) < \sum_{j=0}^{n-1} 1 = n$.

To prove that the ratio $\tilde{p}_n/\hat{p}_n$ is decreasing in n note that $\tilde{p}_n/\hat{p}_n = \frac{nG(\hat{w})-G(\hat{w})^n}{n-1}$, which is decreasing in n if for any $n \geq 2$ we have

$$\frac{nG(\hat{w}) - G(\hat{w})^n}{n-1} < \frac{(n-1)G(\hat{w}) - G(\hat{w})^{n-1}}{n-2},$$

which can be rewritten as

$$G(\hat{w})^{n-2}(n - 1 - (n-2)G(\hat{w})) < 1.$$

The LHS of the above is strictly increasing in $G(\hat{w})$ for $G(\hat{w}) < 1$, thus the LHS reaches its maximum at $G(\hat{w}) = 1$, where it is equal to 1. It follows that $\frac{nG(\hat{w})-G(\hat{w})^n}{n-1}$ is decreasing in n for $c > 0 (G(\hat{w}) < 1)$.

To prove that $\tilde{p}_n$ is decreasing in n we note that Anderson and Renault (1999) show that $\hat{p}_n$ is decreasing in n, which given that $\tilde{p}_n/\hat{p}_n$ is also decreasing immediately implies that $\tilde{p}_n$ is also decreasing in n.

Finally, we show that the limit of $\tilde{p}_n$ when $c \to \bar{c}$ is 0. This follows from $\lim_{c\to\bar{c}} \eta_n^* = 1 - \frac{n-1^n}{n-1} = 0$. A simple modification of the argument presented in the Proof of Proposition 2 shows that this must hold in any equilibrium.

Proof of Proposition 4

The implicit derivative of x with respect to p_1 can be found from (19) as

$$\frac{\partial x}{\partial p_1} = \frac{z(x)(\partial M_1/\partial p_1 - \partial M_2/\partial p_1) + \partial M_2/\partial p_1}{1 - z'(x)(M_1 - M_2)}.$$

If we now impose symmetry, then $x = 1/2$ and $z(1/2) = 1/2$ so that the necessary condition for the symmetric pure strategy equilibrium price becomes

$$\tilde{p} = \frac{1 - (M_1^* - M_2^*)z'(1/2)}{\partial M_1/\partial p_1 + \partial M_2/\partial p_1} = \frac{(1 - G(\hat{w})^2\zeta(k))}{2\int_{\underline{u}}^{\hat{w}} g(u)^2 du + (1 - G(\hat{w}))g(\hat{w})} = (1 - \eta_k^*)\hat{p},$$

where in the last transformation we used the well-known expression for the ARW price

$$\hat{p} = \frac{1}{2\int_{\underline{u}}^{\hat{w}} g(u)^2 du + (1 - G(\hat{w}))g(\hat{w})}.$$

As shown in the text, $\zeta(k)$ is increasing in k, therefore, if the equilibrium exists, $\tilde{p}$ is decreasing in k.

As for the non-existence of the pure strategy equilibrium, note that for $c = \bar{c}_k$, $\tilde{p} = 0$ because, by definition, $1 - \eta_k^* = 1$. At the same time, $G(\hat{w}) > 0$, thus a positive fraction of consumers searches, so that firm 1 can earn strictly positive profit by a deviation to $p_1 > 0 = p^*$. In fact, there exists $c' < \bar{c}_k$, such that no pure strategy equilibrium exists for any $c > c'$.

Finally, notice that the same argument used in the Proof of Proposition 2 extends for any $z'(1/2) \leq 1$ and, therefore, the mixed strategy equilibrium becomes degenerate at $c \to \bar{c}$.

Proof of Lemma 1 Consider any history in which firm 1 has a market share $x^{t-1} > \frac{1}{2}$. Let the continuation value of firm 2 be $V(x)$. In such a case, the optimal price for firm 2 in period t maximizes

$$p_2^t \left[M_2(p_2^t, p_1^t) + x^{t-1}(M_1(p_2, p_1) - M_2(p_2, p_1))\right] + \delta V(x^t) \tag{26}$$

Let $p_2^*(x^{t-1})$ be the profit-maximizing price. From Armstrong and Zhou (2011) we know that if $G(u)$ is uniform, fresh demand is more elastic than returning demand so that $p_2^*(x^{t-1}) > \hat{p}_2(x^{t-1})$ defined as the argmax of

$$p_2^t \left[M_2(p_2^t, p_1^t) + \frac{1}{2}(M_1(p_2^t, p_1^t) - M_2(p_2^t, p_1^t))\right] + \delta V(x^t) \tag{27}$$

where continuation values are as before. Now notice that the derivative of the value function with respect to x, V_x is the value of attracting an additional consumer to firm 2 tomorrow. This is exactly the social multiplier evaluated at tomorrow's prices.

$$\eta^{t+1} = M_1(p_2^{t+1}, p_1^{t+1}) - M_2(p_2^{t+1}, p_1^{t+1}) = (1-G(w^* - p_1^{t+1} + p_2^{t+1}))(1-G(w^* - p_2^{t+1} + p_1^{t+1})) \tag{28}$$

Which is independent of x for given continuation prices. For uniform distributions, the derivative of this function with respect to the difference in prices $p_1^{t+1} - p_2^{t+1}$ is $G(w^* - p_1^{t+1} + p_2^{t+1}) - G(w^* - p_2^{t+1} + p_1^{t+1})$, which is symmetric with respect to $p_1^{t+1} - p_2^{t+1}$ and it is increasing if $p_1^{t+1} < p_2^{t+1}$ and decreasing otherwise. Thus, $V_x(x) = V_x(1-x)$ and the continuation motive is the same. Hence, $p_1^t \leq p_2^t$. But then, $x^{t+1} \leq \frac{1}{2}$ as required. **Proof of Lemma 7**

Fix p^*. The profit from a deviation to some $p_t < p^*$ is at least

$$\begin{aligned}\Pi(p_t, p^*) &= (1-\delta)p_t Q(p_t, p^*; 1/2) + \delta^s p^* Q(p^*, p^*; x^t) \\ &= (1-\delta)x^t p_t + \delta 1 - \delta(M_1^* - M_2^*)[\delta M_2^* + (M_1^* - M_2^*)(1-\delta)x^t]\end{aligned}$$

Taking the First Order Condition and solving for p^* we get

$$p^* \leq \frac{1-\delta\eta^*}{\frac{\partial M_1}{\partial p} + \frac{\partial M_2}{\partial p}} = \tilde{p}$$

Appendix B: Correlated Preferences

The following lemma shows that the stationary equilibrium we study can be reached as the limit of the dynamic economy of the model for T large enough provided that $G(\hat{w}(p^*)) > 0$.

Lemma 2. *Suppose there is a unique cutoff $\tilde{w}(p_1) > \underline{u}$ solving equation (21) and let $w(p_1)$ be the optimal cutoff rule. Then, for every $\epsilon > 0$, there exists a $T^* < \infty$ such that for all (non-negative) prices $\|\tilde{w}(p_1) - w(p_1)\| < \epsilon$.*

Proof of Lemma 2

We first establish that market shares converge for any pair of non-negative prices. To see this, notice first that $M_1(p_1, p^*)$ and $M_2(p_1, p^*)$ are independent of t. The evolution of market shares can be readily computed as

$$|x^t - x^{t-1}| = |x^{t-1} - x^{t-2}|(M_1(p_1, p^*) - M_2(p_1, p^*)) \tag{29}$$

Notice that that $M_1 - M_2 \leq \max_s(1 - G_s(\tilde{w}(p^*))$.[25] If p^* is such that $\tilde{w}(p_1) = \bar{u}$, then market shares are independent of time because all consumers go through firm 2. If $\tilde{w}(p_1) = \underline{u}$, firm 1 is an absorbing state of the process independently of p_2 and the state

[25] Here M_1 is at most 1, and M_2 is at least $G_S(\tilde{w}(p^*))$.

S'. Hence, assume that p_1 is such that $G_S(\tilde{w}(p_1)) \in (0,1)$

$$|x^t - x^{t-1}| \leq [\max_s (1 - G_s(\tilde{w}(p^*)))]^t |x^1 - x^0| \tag{30}$$

or

$$|x^t - x^{t-1}| \leq [\max_s (1 - G_s(\tilde{w}(p^*)))]^t |x^W - \frac{1}{2}| \tag{31}$$

where x^W is the share of firm 1 in ARW (i.e. the first consumer). Clearly, this Fixed Point converges to the stationary market shares by Blackwell's Theorem and the parameter of convergence is $\{\max_s(1 - G_s(\hat{w}(p^*))\}$ which is strictly less than one. For every $\epsilon > 0$ and for every $p \geq 0$, there exists $T_1 < \infty$ such that $T_1 = \frac{\ln(\phi * \epsilon)}{\ln((1 - G_L(\tilde{w}(0))))}$ and

$$|x^t - x^{t-1}| \leq \phi\epsilon$$

for every $t > T_1$ and some $\phi > 0$. Now we can compute the optimal cutoff rule for a consumer who knew his own arrival time. Indeed, since $\lambda(u)$ and $x_{S,S'}(p_1)$ are continuous functions, the conditional probabilities of the different states given some (u, p_1) realization can be written as

$$q_{T_1}(u; p_1) = \frac{x_{HH}^{T_1}(p_1)\lambda(u) + x_{LH}^{T_1}(p_1)(1 - \lambda(u))}{(x_{HH}^{T_1}(p_1) + x_{HL}^{T_1}(p_1))\lambda(u) + (x_{LH}^{T_1}(p) + x_{LL}^{T_1}(p_1))(1 - \lambda(u))}. \tag{32}$$

which differs from q^* as computed from Equation (23) because it depends on t. The remainder of the proof establishes that this difference is arbitrarily small. Notice that $x_{S,S'}^{T_1} \in [x_{S,S'} - \phi\epsilon, x_{S,S'} + \phi\epsilon]$, we can write

$$sup \parallel q_{T_1}, q^* \parallel \leq \frac{\phi\epsilon}{\lambda(x_{HH} + x_{HL}) + (1 - \lambda)(x_{LH} + x_{LL})} < \frac{\phi\epsilon}{G_L(\hat{w}(0))(1 - G_L(\hat{w}(p_1))} \tag{33}$$

Let $T^* = RT_1$, for R large enough we have that

$$\begin{aligned} \sup \parallel \sum_{t=1}^{T^*} \frac{1}{T^*} q_t(u, p_1), \sum_{t=1}^{T^*} \frac{1}{T^*} q_t^*(u, p) \parallel &< \frac{(R-1)\frac{\phi\epsilon}{G_L(\hat{w}(0))(1 - G_L(\hat{w}(p_1))} + 1}{R} \\ &< \frac{2\phi\epsilon}{G_L(\hat{w}(0))(1 - G_L(\hat{w}(p_1))} \end{aligned}$$

and now let $\phi = \sup_{p:\bar{u} > \hat{w}(p_!) > \underline{u}} \frac{1}{2} G_L(\hat{w}(0))(1 - G_L(\hat{w}(p_1)))\underline{g}(u)$, so that the difference between the stationary probability measure and the actual probability measure that almost everyone observes is smaller than $\epsilon\underline{g}(u)$. Using Equation (21), it is easy to verify that $G(\hat{w}(p))$ is a continuous contraction in the probability measure q. Since G is strictly increasing, G^{-1} is uniformly continuous with parameter $\frac{1}{\underline{g}(u)} \geq 0$, and this implies that $\|\tilde{w}(p_1) - w(p_1)\| < \epsilon$ as required.

Proof of Proposition 5

In equilibrium $x_{HH} = x_{LL} = 1/2$ and $x_{HL} > 1/2 > x_{LH}$, thus $\lambda(u) < 1/2$ for all u. From (21) it follows that its LHS is higher than the LHS of (4) for the same w. This immediately implies that $\tilde{w} < \hat{w}$.

Proof of Proposition 6

Fix c and let p^* be the conjectured price of the rival firm. We prove that for any price $p^* > \bar{p}(c)$ there is a profitable deviation to $p^* - (\bar{c}_L - c)$ that captures the full demand.[26] Indeed, notice that

$$\int_{\underline{u}}^{\bar{u}} (u - \underline{u} - (\bar{c}_L - c))dG_L(u) = c$$

by definition of c_L. Since $\lambda(\underline{u}) = 0$ a firm charging $p^* - (\bar{c}_L - c)$ retains all its incoming consumers. Because of the social multiplier of demand, this implies that its long-run market share is 1. This yields a profit of $p^* - (\bar{c}_L - c)$. Therefore, every price p^* such that

$$\frac{1}{2}p^* \leq p^* - (\bar{c}_L - c)$$

is inconsistent with an equilibrium. Thus, $p^* \leq p(c) = 2(\bar{c}_L - c)$. Clearly, as $c \to \bar{c}_L$, $p(c) \to 0$ and the result follows.

Proof of Corollary 2

Proof. By Proposition 4, it holds that $\lim_{\delta\to 1} \lim_{c\to\bar{c}_L} \tilde{p} = 0$. On the other hand $\lim_{\delta\to 0} \lim_{c\to\bar{c}_L} \tilde{p} = 0$. For all $c \leq \bar{c}_L < \bar{c}$, $\hat{p} > 0$. Hence, the result follows. $\square$

Appendix C : Finite Outside Option

We have followed Anderson and Renault (1999) and assumed that consumers have to buy from one of the firms because their outside option is arbitrarily bad. This assumption greatly simplifies the analysis but it is not without loss. The main consideration is that when there is a sufficiently good outside option, for high enough search costs consumer do not search, and firms charge monopoly prices. This means that there is a limit to how low prices can be when search costs are high, because eventually consumers stop searching and prices jump up.

In this Appendix we extend the model to incorporate a zero outside option using the baseline model presented in Section 4. This is the Wolinsky's version of the ARW model. In order to deal with the issue that some consumers do not buy, we assume that every consumer observes the close predecessors who *purchased* one of the two goods. As before we focus on the case where $\delta \to 1$. Firm 1's share of first visits is now $z(x) = x/X$, where

[26] The argument can be extended to mixed strategies by choosing p^* to be the lower bound of the support of the price distribution.

x is the demand of firm 1 and X is total market demand. We have

$$\begin{aligned} x(p_1, p^*) &= \frac{x(p_1, p^*)}{X}(1 - G(\hat{w} + p_1 - p^*)) + \Big(1 - \frac{x(p_1, p^*)}{X}\Big)G(\hat{w})(1 - G(\hat{w} + p_1 - p^*)) \\ &+ \int_{p_1}^{\hat{w}+p_1-p^*} G(u - p_1 + p^*)g(u)du, \end{aligned}$$

where p_1 is the price charged by firm 1 and p^* the equilibrium price. This expression is similar to (7) except for the usage of stationary market shares x/X and the fact that only consumers with valuations above p_1 purchase the product (notice that the integral runs from p_1 instead of $\underline{u}$).

Total demand can be readily computed since those consumers who do not purchase in any of the two firms is $G(p_1)G(p^*)$, which gives the total demand $X = 1 - G(p_1)G(p^*)$. One can now solve for $x(p_1, p^*)$, maximize firm 1's profit given firm 1's demand, and impose that in equilibrium $p_1 = p^*$.

Before we fully characterize the above equilibrium, we should note that with a zero outside option consumers may stop searching, which is the case when $p^* > \hat{w}$. Thus when search cost is sufficiently large, consumers do not search and firms charge monopoly prices.

Proposition 8. *Under suitable conditions on $G(w)$, there exist two thresholds, c_1 and c_2 with $\bar{c}_1 < \bar{c}_2$, such that*
(i) with observational learning equilibrium price is $\tilde{p}$ for $c \leq \bar{c}_1$ and the market breaks down otherwise;
(ii) without observational learning equilibrium price is $\hat{p}$ for $c \leq \bar{c}_2$ and the market breaks down otherwise.
(iii) for $c \leq \bar{c}_1$ the price in the model with emulation is lower than in the model without it.

Proof. Solving for x from (34), taking derivative with respect to p_1 and imposing $p_1 = p^*$ gives the following pricing rule

$$\tilde{p} = \frac{\big(1-G(\tilde{p})^2\big)\big((2-G(\hat{w}))G(\hat{w})-G(\tilde{p})^2\big)}{2(1-G(\tilde{p})^2)\int_p^{\hat{w}} g(v)^2\,dv+G(\tilde{p})(1-2G(\tilde{p})^2+(2-G(\hat{w}))G(\hat{w}))g(\tilde{p})+(1-G(\tilde{p})^2)(1-G(\hat{w}))g(\hat{w})}. \tag{34}$$

The above only holds for $p \leq \hat{w}$, which gives the threshold on c, denoted by $\bar{c}_2$ From (34), $\bar{c}_2$ is such that the corresponding w solves

$$w = \frac{2G(w)\,(1 - G(w)^2)}{(2G(w)^2 + G(w) + 1)\,g(w)}. \tag{35}$$

Now consider the Wolinsky pricing rule derived under the assumption that half of consumers make first visits to firm 1.

$$\hat{p} = \frac{1 - G(\hat{p})^2}{2\left(\int_{\hat{p}}^{\hat{w}} g(v)^2\, dv + G(\hat{p})g(\hat{p}) + \frac{1}{2}(1 - G(\hat{w}))g(\hat{w})\right)}. \tag{36}$$

As before, the above only holds for c below a threshold, denoted by $\bar{c}_1$. $\bar{c}_1$ is such that $p = w$, and corresponds to w that solves

$$w = \frac{1 - G(w)}{g(w)}.$$

It is easy to show that $\frac{2G(w)\left(1 - G(w)^2\right)}{(2G(w)^2 + G(w) + 1)g(w)} < \frac{1 - G(w)^2}{2g(w)G(w)}$ in the relevant range of $\hat{w}$, therefore $\bar{c}_1 < \bar{c}_2$.

We still have to show the uniqueness and the existence of $\bar{c}_1$ and $\bar{c}_2$. For $\bar{c}_1$ see Janssen and Shelegia (2014) for the proof. For our model, note that $\bar{c}_2$ is defined as c such that $\tilde{p} = \hat{w}$. Given that $G(\bar{u}) = 1$, which makes the RHS of (35) equal to 0, such a solution always exits, but might not be unique. Thus we require that $G(w)$ is such that (35) has a unique solution.[27] If it is indeed unique, it is trivial to show that $p^M > w$, so that indeed once $c > \bar{c}_1$, firms charge $p^M > \hat{w}$ and no consumer searches.[28]

In both cases, for $c > \bar{c}_i$, the Diamond paradox (Diamond (1971)) prevails. That is, consumers do not search because they expect prices that are so high that they are not willing to pay the first search cost.

Finally, to prove (iii) we need to show that $\tilde{p} < \hat{p}$ for $c < \bar{c}_1$. To do so, we show that the ratio of the LHS of (34) to the LHS of (36) for the same p is less than 1. After some rearrangement one can write this condition to be the same as

$$2\int_p^{\hat{w}} g(v)^2\, dv + G(p)g(p) + (1 - G(\hat{w}))g(\hat{w}) > 0,$$

which holds for any $p < \hat{w}$. Thus for any $c \leq \bar{c}_1$, the solution to (34) is higher than the solution to (36), therefore $\tilde{p} < \hat{p}$. $\square$

In Figure 7 below we depict equilibrium prices for $G(\cdot) \sim U[0, 1]$. The gray curve depicts standard Wolinsky price as a function of c. For relatively low c, the price is below the reservation utility and increasing in c. Once c reaches the threshold $\bar{c}_1$, the market breaks down. As shown above, the price with emulation is lower (the black curve). It is also non-monotone, and never reaches zero because once $c > \bar{c}_2$, the market breaks down. Because, $\bar{c}_2 > \bar{c}_1$, with emulation the Diamond Paradox appears for higher search cost than in the Wolinsky model. Thus emulation allows consumers to collectively escape the

[27] It is unique for the uniform distribution.

[28] In principle, it is possible that at the threshold, $\hat{w} > p^M$, in which case no pure strategy equilibrium exists for c immediately above $\bar{c}_2$ because conditional on consumers not searching, firms charge p^M and induce search, but conditional on consumers searching, they charge $p > \hat{w}$, precluding search.

paradox by putting sufficient downward pressure on prices.

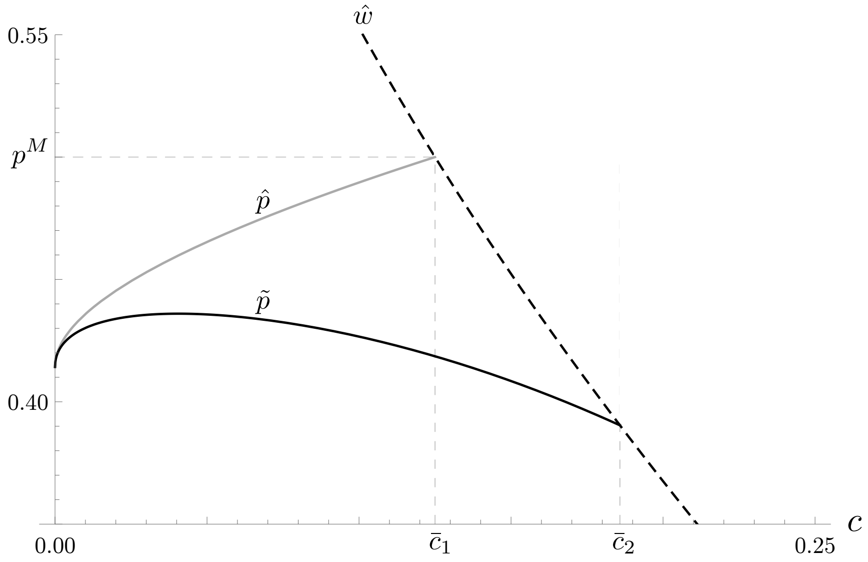

Figure 7: Equilibrium prices in the model with zero outside option for $G \sim U[0, 1]$.

To conclude, note that prices are never close to marginal cost because consumers stop searching when search cost is high, the reason for zero price in the model without outside option. As a result, whether the price is decreasing in search cost depends on the shape of the distribution G. This occurs, for instance, if the price with $c = 0$ is higher than the price when $c = \bar{c}_2$. This is easily verified for the uniform case.

Appendix D: Welfare

What are the welfare consequences of introducing correlation of preferences into the model? Consumers observe the purchasing decisions of their predecessors and thus visit the store offering higher utility realizations with higher probability. Since the search rule is a cutoff rule, we have that total Surplus can be written as:

$$
\begin{aligned}
W &= \sum_{ss'} \int_{\underline{u}}^{\hat{w}} \big(uG_s(u)g_{s'}(u) + uG_{s'}(u)g_s(u)\big)du \\
&+ \sum_{ss'} \Bigg[x_{ss'} \left(\int_{\hat{w}}^{\bar{u}} ug_s(u)du + G_s(\hat{w}) \int_{\hat{w}}^{\bar{u}} ug_{s'}(u)du\right) \\
&+ (1 - x_{ss'}) \left(\int_{\hat{w}}^{\bar{u}} ug_{s'}(u)du + G_{s'}(\hat{w}) \int_{\hat{w}}^{\bar{u}} ug_s(u)du\right)\Bigg]\Bigg) \\
&- c\sum_{ss'} \big(x_{ss'}G_s(\tilde{w}) + (1 - x_{ss'})G_{s'}(\tilde{w})\big)
\end{aligned}
$$

The first term represents the expected utility of a consumer whose utility draws in both stores are below the cutoff and, therefore, is independent of market shares because such a consumer searches regardless of where she makes her first visit. The remaining terms increase in the difference between the market share of a firm offering a High and a Low distribution ($x_{HL} - x_{LH}$), and, therefore, it is higher in our model as compared to ARW. This means that

Proposition 9. *Welfare is higher if consumers observe others.*

Proof. Welfare can be rewritten as

$$W = \sum x_{SS'} V_S(\hat{w}) + (1 - x_{SS'}) V_{S'}(\hat{w}) \tag{37}$$

where $V_S(x)$ is the value of a consumer visiting a firm with distribution S first and using x as a threshold rule. Clearly, $\hat{w}$ is optimal given the information that each individual consumer has access to. In other words, for given x_{ij}

$$W = \max_{\hat{w}} \{ \sum x_{SS'} V_S(\hat{w}) + (1 - x_{SS'}) V_{S'}(\hat{w}) \} \tag{38}$$

Notice that in states HH and LL, W is independent of $x_{SS'}$. In state HL, however, W is increasing in x_{HL} since $V_H \geq V_L$. But then,

$$W = \max_{\hat{w}} \{ \sum x_{SS'} V_S(\hat{w}) + (1 - x_{SS'}) V_{S'}(\hat{w}) \} \geq \max_{\hat{w}} \{ \sum \frac{1}{2} V_S(\hat{w}) + \frac{1}{2} V_{S'}(\hat{w}) \} = W^* \tag{39}$$

the welfare in the ARW model. Thus, consumers benefit from having access to information. □

Perhaps the more interesting result concerns the comparative statics of welfare with respect to search costs. Information about predecessors and search constitute substitute sources of information. As the search cost vanishes, search is a more efficient signal and so the prior information becomes irrelevant. On the other hand, as search cost increases, information about predecessors becomes increasingly important. This substitution effect is at the core of the following result.

Proposition 10. *Increasing the search cost may have a non monotone effect on welfare. For arbitrarily low search costs, welfare decreases in the search cost. On the other hand, for high enough search cost there are distributions G_H and G_L such that welfare increases in the search cost.*

Proof. By the Envelope Theorem, since $\tilde{w}$ is chosen optimally given c, a marginal change in c affects W only through the direct effect of savings in search expenditures and the indirect effect of information aggregation via $x_{ss'}$. When $c \to 0$, there is no value of information since search is free, and, thus, the only value comes from savings. On the other hand, if $c \to \bar{c}$, there is a (small) direct effect on savings since the measure of searchers is $\frac{1}{4} \sum x_{ss'} G_S(u)$. Information is useful since search is costly so that the only question is whether more or less information is aggregated. Since $x_{HH} = x_{LL} = \frac{1}{2}$, the only question is the sign of $\frac{\partial x_{HL}}{\partial \tilde{w}}$.

$$\frac{\partial x_{HL}}{\partial \tilde{w}} = \frac{G_L(\tilde{w})g_H(\tilde{w}) - G_H(\tilde{w})g_L(\tilde{w})}{(G_H(\tilde{w}) + G_L(\tilde{w}))^2} > 0 \tag{40}$$

where the first equality comes from the definition of market shares at $\tilde{w} \to \underline{u}$ and the last inequality comes from the MLRP, but $\tilde{w}$ is decreasing in c. In particular,

$$\frac{\partial \tilde{w}}{\partial c} = -\frac{1}{1 - \frac{1}{4}\sum x_{SS'}G_{S'}(\tilde{w})} < -\frac{1}{1 - G_H(\tilde{w})}$$

Thus, welfare increases in search costs if

$$\begin{aligned} \frac{G_L(\tilde{w})g_H(\tilde{w}) - G_H(\tilde{w})g_L(\tilde{w})}{(G_H(\tilde{w}) + G_L(\tilde{w}))^2} \frac{1}{1 - G_H(\tilde{w})}(V_H(\tilde{w}) - V_L(\tilde{w}))\frac{1}{2} \\ \geq \frac{1}{4}\sum \frac{G_S(\tilde{w})G_{S'}(\tilde{w})}{G_S(\tilde{w}) + G_{S'}(\tilde{w})} \end{aligned}$$

which yields

$$\begin{aligned} (G_L(\tilde{w})g_H(\tilde{w}) - G_H(\tilde{w})g_L(\tilde{w}))(V_H(\tilde{w}) - V_L(\tilde{w})) \geq \\ 2(G_H(\tilde{w}) + G_L(\tilde{w}))^2)(1 - G_H(\tilde{w}))\sum \frac{G_j(\tilde{w})G_i(\tilde{w})}{G_i(\tilde{w}) + G_j(\tilde{w})}. \end{aligned}$$

For the case of the triangular distribution this becomes

$$(\tilde{w})^2\frac{1}{3} > (\tilde{w})^2(1 - (\tilde{w})^2)\{\tilde{w} + (\tilde{w})(2\tilde{w} - (\tilde{w})^2)\} = (\tilde{w})^3(1 - (\tilde{w})^2)(1 + 2\tilde{w} - (\tilde{w})^2)$$

which holds for $\tilde{w}$ small enough.

$\square$

The first part is rather trivial. If the search cost is sufficiently low, almost every consumer searches and an increase of the search cost induces a first order, direct and negative effect on welfare. Since consumers do not stick to their first option, prior information is not valuable and, thus, welfare is reduced. The second part, however, is less intuitive. When almost every consumer buys at the first store she visits, an increase of the search cost has a negligible effect on total search expenditure. Marginal consumers are discouraged from searching but those do not matter for welfare because their utility loss is negligible (their cutoff rule being optimal). Importantly, however, these marginal consumers fail to internalize the information externality they originate when searching. Naive intuition would suggest that this externality is positive, since consumers who search acquire valuable information and pass it on to their successors. Importantly, however, their successors obtain only a biased signal since they do not learn their search pattern, only their purchasing decision. As it turns out, when search costs are very high, a searcher is more likely to buy in a store with a Low distribution than in a store with a High distribution and thus passes on *bad* information to her successor. Since successors deem it unlikely that their predecessor searched, they follow them to the store with a Low distribution. Therefore, the net effect of an increase in search cost depends on the relative magnitude of the direct negative effect and the indirect effect on information aggregation. In the proof we show that for Triangular distributions it is indeed the case that social welfare is non-monotone in the search cost.