Consumer Search with Observational Learning*

Daniel Garcia[†] Sandro Shelegia[‡]

December 1, 2014

Abstract

This paper studies social learning in a search environment. New consumers observe previous consumers' purchasing decisions before embarking on their own search for the best-fitting product. This form of social learning has two distinct effects on consumer search and firm pricing. First, consumers *emulate* others in the sense that they always make the first visit to the firm where their predecessor has purchased. Second, consumers *free-ride* on their predecessor's information and search less intensively than in the standard search model. Emulation encourages price competition because firms fight for consumer visits, while free-riding has the opposite effect due to reduced search. Both effects increase in the search cost, but emulation is shown to dominate for most commonly used distributions of consumer preferences, therefore prices (eventually) fall as search cost increases, and may even go down to the marginal cost. We show that the results derived in a static duopoly model remain valid with a large number of firms and discuss welfare implications.

JEL Classification: D11, D83, L13

Keywords: Consumer Search; Observational Learning, Emulation

---

*We thank Maarten Janssen, Marco Haan, Jose Luis Moraga-Gonzalez and Stephan Lauermann for their comments as well as the audience at the 6th Consumer Search and Switching Workshop at Indiana University and the MACCI Conference in Consumer Search at Bad Homburg. 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. Email: daniel.garcia@univie.ac.at

[‡]Department of Economics, University of Vienna. Email: sandro.shelegia@univie.ac.at

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 closely resembles the problem faced by consumers in many markets, where previous customers' 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, new consumers may *free ride* on the search effort of others and follow their advice. This induces a *social multiplier* of demand and may affect prices.

To the best of our knowledge, no paper has studied this issue. We attempt to bridge the gap between the observational learning literature and the consumer search literature by analyzing a simple oligopoly model of search with heterogenous products in the spirit of Wolinsky (1986) and Anderson and Renault (1999) (henceforth ARW). In this model, a large number of consumers derive utility from consuming a single unit of a given good, that comes in two varieties sold by two different firms. Consumers are initially uninformed about their valuation for each variety or the price charged by each firm, but may learn it 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 stores and buy from the highest-surplus offering firm. Taking this behavior as given, firms simultaneously choose prices to maximize expected profits.

We depart from this model by allowing each consumer to observe the purchasing decision of a single predecessor. Importantly, they do not observe whether their predecessor visited other stores nor the price he paid. We further assume that they visit first the firm offering the variety their predecessor bought. This behavioral modification, which we refer to as *emulation*, amounts to a change in the tie-breaking rule used in ARW. This is consistent with recent evidence in economics and marketing.[1] For instance, Mobius et al. (2005) present results of a field experiment designed to understand individual demand for different products for which individuals receive information. They show that those subjects who are connected through social links to others who are informed about a particular good, value the goods more. Importantly, they find that this effect is stronger for gadgets than for services, which are more likely to be subject to direct observation. 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)

[1] It is also implied by trembling-hand perfection.

analyzes the movie market where an unexpected increase in the first-week's box office has a persistent a significant effect in future attendance. Finally, Zhang (2010) shows that patients draw negative quality inferences from earlier refusals in the queue, thus becoming more inclined towards refusal themselves.

Strikingly, this change seemingly small modification drastically changes equilibrium outcomes. Since consumers buy from the store their predecessor bought, the market share of a firm in a given period equals the probability that a random consumer visits the store first in the following period. Since consumers are more likely to buy the variety they sample first, this induces a *social multiplier* of demand. In ARW model firms lower prices in order to increase the number of consumers they retain. In our model, in addition, firms lower prices in order to increase the number of consumers they attract. We show that in the resulting equilibrium, prices are lower than in ARW, and, contrary to the baseline model, for all commonly used distributions prices decrease in search cost once search cost is sufficiently large. The difference between ARW's and our model originates in those consumers whose utility realization in both stores exceeds their reservation utility, and, therefore, always buy in the store they visit first. In the baseline model, the firm who is visited by these consumers first is effectively a monopolist and, thus, as the proportion of such consumers increases, prices tend to increase. In our model, however, the likelihood that one such consumer visits a given store depends on this 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 store increases, firms engage in increasingly fiercer competition for searchers, leading to Bertrand-like competition and eventually to prices that can be as low as the marginal cost.

In light of this result, we then introduce proper learning in the model. To do so, we assume that consumer valuations for each variety are positively correlated, but valuations across stores remain independent. 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 has realized, we can focus on the effect of learning in consumer behavior and prices. 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. Hence, 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 effects of learning on equilibrium prices are, however, 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 beliefs that incoming consumers have. 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 is the price of the store 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 some distributions, such as uniform and normal, 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, provided that the likelihood ratio of the two distributions is unbounded, prices converge to marginal costs as search costs increase.

We also show that social welfare may be non-monotonic in search costs. A marginal change in search costs has three different effects on social welfare. First, it has a direct effect in utility since some consumers actively search in equilibrium. Second, it has an indirect effect on the search rule that consumers use, which crucially determines the allocation. While the direct effect is zero because consumers choose the reservation utility so as to maximize their own utility, there is an indirect effect on market shares that distorts learning. Intuitively, as search cost increases, consumers are increasingly diverted towards the firm with better utility distributions, which is is useful given that almost no consumer searches past the first firm. We show that this effect on welfare is always negative. We further show that for some distributions, such as uniform, it dominates the direct effect and so welfare may increase in search costs.

In the reminder of the paper we consider several extensions of the model. First we study the effects of competition on equilibrium prices in the model without learning. We show that prices decrease in search cost and converge to marginal costs as the proportion of searchers diminishes, for any number of firms. We further show that prices decrease in the number of firms for any search cost Of special interest is the limit price when the number of firms grows large. In this case, whether there is correlation across individuals or not, there is nothing to be learnt about rival firms' qualities from the realization of sampled varieties. Armstrong and Chen (2009) shows that, as long as it is exogenous, the order of visits does not change prices in the baseline model with a very large number of firms since the elasticity of demand is the same for all firms. In our model, however, prices are lower than in the baseline model and have a inverse U-shaped relationship with search costs. If search costs are very low and there are many firms, Bertrand competition for buyers obtains. On the other hand, if search costs are very high, we know that firms fiercely compete for first visits and prices also converge to marginal costs. For intermediate search costs, however, firms charge positive prices. In another extension we introduce a zero outside option, so that some consumers may choose not to buy.

Literature Review

Several recent papers have analyzed consumer search with observational learning, but all assume that prices are fixed exogenously. Kircher and Postlewaite (2008) study consumers who differ in their willingness to search and firms differ in quality. Although prices are fixed, firms may decide to offer a valuable service to any consumer who visit their store. They show that equilibria may arise where high-quality firms offer service to those consumers who search more actively and those who search less actively follow their advice. Hendricks et al. (2012) presents 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 quality of a single good, buying it right-away or not buying. The model is cast in a more traditional herding framework where consumer receive a signal before deciding whether to engage in costly search. They focus on the long-run dynamics of sales for high and low quality products and the possibility of bad herds arising (see also Ali (2014) and Mueller-Frank and Pai (2014)).

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 customers. 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 share of first visits depends on price. Armstrong et al. (2009) show that as the number of firms grows, (exogenous) prominence becomes irrelevant while in our model, competition becomes even fiercer and prices decrease in the number of firms. 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 cost leads to (stochastically) lower prices. One may view observability of prices as an extreme example of observational learning where consumers observe market shares, and thus prices. We show that imprecise information about prices results in a pure strategy equilibrium featuring inverse relationship between search cost and prices. Another closely related model can be found in 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).

Our paper is also related to the literature on monopoly pricing in the presence of social learning. Campbell (2013) and Chuhay (2010) analyze the impact of word-of-mouth communication on the monopoly price and product design. Perhaps closest to our work are two contributions by Bose et al. (2006) and Bose et al. (2008) who study dynamic interaction between a monopolist and a sequence of consumers with common valuation who observe each other's purchasing decisions. While most of this literature has studied monopoly, an exception is Kovac and Schmidt (2014) who 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[2], we view our work as complementary to theirs.

Finally, a number of papers have studied the relation between current market shares and future demand. Becker (1991) directly introduces aggregate demand into the individual utility function. 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 to 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 remaining of the paper is organized as follows. Section 3 introduces the basic model and characterizes the equilibrium. Section 4 introduces correlation in consumers' preferences. Section 5...

2 The Baseline Model

Consider a market populated by a large number of consumers $i = 1, 2, .., N$, 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 the first visit is free, but the second 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 $G(u)$ (with density $g(u)$) on a closed and finite support $[\underline{u}, \overline{u}]$. Realizations are independent between consumers and firms. In order to guarantee that pricing equilibrium is well behaved, we assume that $G(u)$ is Log-concave on $[\underline{u}, \overline{u}]$.[3] The outside option is assumed to be sufficiently bad so that each consumer buys one unit regardless of prices. We consider finite outside option in Section 5.1.

We assume that the market unfolds over time and consumers arrive at the market sequentially. In particular, we assume that consumer i arrives to the market in period i and leaves the market before the next consumer arrives. All consumers with the exception of the first one do not know their arrival times and hold a common prior $\nu(i)$ over it.[4]

[2]See Section 4

[3]Anderson and Renault (1999) show that this condition is sufficient for the equilibrium price to be increasing in c. It is satisfied for most common distribution functions such as uniform and normal.

[4]A similar model has been studied in Monzón and Rapp (2014). They show that the classical results

We introduce observational learning by allowing each individual consumer $i > 1$ of the purchasing decision of her predecessor $i-1$. Given this information, the consumer decides which store to visit first and finally decides where to buy. Consumer $i = 1$ has no predecessor and, thus, has no additional information. Following ARW we assume that she chooses which store to visit by flipping a 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 in firm 1 given prices (p_1, p_2).

$$\Pi_1(p_1, p_2) = p_1 \frac{(1-\delta)}{1-\delta^N} \sum_{t=1}^{N} \delta^N x^i(p_1, p_2). \tag{1}$$

Notice that if $\delta = 0$, the model is essentially ARW. Our main focus will be in the case where the number of consumers is large and the firm is sufficiently patient, so that the firm maximizes

$$\lim_{\delta \to 1} \lim_{N \to \infty} \Pi_1(p_1, p_2). \tag{2}$$

Each of the limits has different implications. We consider N large so as to guarantee that market shares converge to their stationary distribution. None of our results depend on this and proofs are provided for the case where N is finite but sufficiently large. On the other hand, the discount factor captures the difference between our model and ARW, and taking it to the limit makes the comparison easier.

Following the literature, we shall devote our attention to symmetric Pure Strategy Equilibrium.

2.1 Consumer Behavior

In this model a consumer is indifferent between following its predecessor for the first visit or not. Wolinsky model can be considered as a special case of this model where consumers ignore actions of their predecessor. This however, is behavior would not be trembling-hand-perfect because if firms may charge prices other than the equilibrium price, then on average, the firm where predecessor bought charges a lower price. Thus following the predecessor can be justified with trembling-hand-perfection. Moreover, in due course we shall study a model where consumers' utilities are positively correlated, thus giving consumers a strict incentive to follow their predecessors. In this section we shall assume that consumers will always follow their predecessor for the first visit.

Let $\hat{p}$ denote the symmetric equilibrium price. If firm i charges $p_i \neq \hat{p}$ while consumers expect the other firm to charge $\hat{p}$, then a consumer who visits it first will search further

in herding models survive even if consumers cannot condition their strategies on their position if and only if her utility realization u_i at firm i is below $\hat{w} - \hat{p} + p_i$ where $\hat{w}$ solves

$$\int_{\hat{w}}^{\bar{u}} (u - \hat{w})g(u)du = c. \tag{3}$$

Notice that as c increases $\hat{w}$ decreases. Define $\bar{c}$ as

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

From now on, we shall focus on the case where $c \leq \bar{c}$. The consumer who arrives in period 1 has no a-priori information to discriminate firms. Hence, she makes her first visit randomly and buys at the first store if and only if $u_{i1} - p_i > \hat{w} - p^*$ and searches the other store otherwise. All remaining consumers observe a predecessor buying from firm $i \in \{1, 2\}$ and expect the same price p^* in both stores before embarking on their first search. Since predecessors choices are not informative about unobserved prices or utilities, once at the first store, consumers use the same reservation utility $\hat{w} - p + p^*$. As explained above, we assume consumers follow their predecessor. We can write firm 1's market share in period i as follows

$$\begin{aligned} x^i(p_1, p^*) &= (1 - G(\hat{w} + p - p^*))x^{i-1}(p_1, p^*) + (1 - x^{i-1}(p_1, p^*))G(\hat{w})(1 - G(\hat{w} + p - p^*)) + \\ &+ \int_{\underline{u}}^{\hat{w}+p-p^*} G(u - p + p^*)g(u)du, \end{aligned} \tag{5}$$

with a convention that $x^0 = 1/2$. The intuition for this equation is simple. Consumer i first visits firm 1 with probability $x^{i-1}(p_1, p^*)$ and buys there without searching further with the probability $1 - G(\hat{w} - p + p^*)$. She may also buy at firm 1 after having visited firm 2 first, provided that she draws a utility below $\hat{w}$ at firm 2, and then buys at firm 1 if her utility draw is above $\hat{w} - p + p^*$. In all other cases she searches both stores and compares prices, in which case she buys at firm 1 if her utility there exceeds her utility at firm 2 plus the price difference. This is captured by the last term in(5). Using this equation recursively, we obtain the following formula for profits per consumer:

$$\Pi_1(p, p^*) = p(1 - \delta)\frac{(1-G(\hat{w}+p-p^*))((1-\delta)(1-G(\hat{w}))x^0+G(\hat{w}))+\int_{\underline{u}}^{\hat{w}+p-p^*} G(u-p+p^*)g(u)du}{1-\delta(1-G(\hat{w}+p-p^*))(1-G(\hat{w}))}$$

That is, expected discounted demand is a combination of the demand of the first individual (who behaves just like in the standard Wolinsky model) and the demand from the *mature* market. The difference between these two demand functions stems from those consumers whose utility realizations are such that, regardless of the store 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 in each store with the probability that a searcher buys in that store (since they are following them, or following others who follow them, etc.) Notice also that this profit function includes the profit function of Anderson and Renault (1999) as a special case. In particular, when $\delta = 0$, this becomes

$$\Pi_1(p, p^*; \delta = 0) = \frac{p}{2}(1 - G(\hat{w} + p - p^*))(1 + G(\hat{w})) + p \int_{\underline{u}}^{\hat{w}+p-p^*} G(u - p + p^*)g(u)du.$$

On the other hand, if $\delta \to 1$, initial individuals whose visiting probabilities are not equal to the stationary ones are no longer relevant and we get

$$\Pi_1(p, p^*) = p\frac{(1 - G(\hat{w} + p - p^*))G(\hat{w})) + \int_{\underline{u}}^{\hat{w}+p-p^*} G(u - p + p^*)g(u)du}{1 - (1 - G(\hat{w} + p - p^*))(1 - G(\hat{w}))}.$$

In a Symmetric Pure Strategy Equilibrium, firms maximize profits choosing prices simultaneously, taking other firm's prices and consumer behavior as given.

Proposition 1. *The symmetric equilibrium price is*

$$p^*(\delta) = \frac{1 - \delta(1 - G(\hat{w}))^2}{2\int_{\underline{u}}^{\hat{w}} g^2(u)du + (1 - G(\hat{w}))g(\hat{w})} \tag{6}$$

which is decreasing in δ. In particular, $p^*(1) = G(\hat{w})(2 - G(\hat{w}))p^*(0) < p^*(0)$.

Proof. All proofs are relegated to the Appendix. □

As δ increases, the weight of the initial batch of individuals decreases and the expected demand is increasingly dependent on the long-run market shares. As already mentioned, the latter depend only on the behavior of searchers. Since searchers are more price-elastic, prices are lower the higher the discount factor. More surprisingly, however, we have the following comparative statics with respect to search costs.

Corollary 1. *For any δ, as $c \to \bar{c}$, $p^*(\delta) = \frac{1-\delta}{g(\underline{u})}$. Thus, the price converges to 0 if $\delta \to 1$.*

As search costs increase, first visits become more attractive but the mass of 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 those searchers have on total profits. As δ increases, searchers become more and more important and price elasticity increases, inducing lower prices. Indeed, as $c \to \bar{c}$ prices approach zero (i.e. marginal cost).

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 sensitive to prices (in the world with high search costs, they were induced to search and compare prices due to very low utility realizations for both products), and each one of them brings a very large number of other consumers in, who do not search.

This is an astonishing result. In a model without observational learning, high search costs relax price competition and result in high prices, whereas in the model with observational learning, exactly the opposite is true. It is noteworthy that while consumers are indifferent between emulating or not, collectively they are always better off when emulating, so one may expect such a social norm to emerge due to experimentation.

2.2 The Effects of Competition

We now extend this results to the case of an oligopoly with n firms. In order to save notation, and for the remaining of the paper, we focus on the two extreme cases where $\delta \to 1$ and $T \to \infty$ or $\delta = 0$. In this case, profits depend on the long-run, or *stationary*, market share. If a firm chooses a p if all its rivals choose p^*, its stationary market share is

$$
\begin{aligned}
x(p,p^*) &= x(p,p^*)(1-G(\hat{w}+p-p^*)) + \frac{1-x(p,p^*)}{n-1}\sum_{j=1}^{n-1} G(\hat{w})^j(1-G(\hat{w}+p-p^*)) \\
&+ \int_{\underline{u}}^{\hat{w}+p-p^*} G(u-p+p^*)^{n-1} g(u-p+p^*)du.
\end{aligned}
$$

Notice that market shares only determine the number of first visits to each firm, but not the subsequent searches. Let p_n^* be the price in a symmetric equilibrium with n firms. For an arbitrary distribution G, we have

$$
p_n^* = \frac{\frac{nG(\hat{w})-G(\hat{w})^n}{n-1}}{g(\hat{w})\left(\sum_{i=0}^{n-2} G(\hat{w})^i+(n-1)G(\hat{w})^{n-1}\right)-(n-1)n\int_{\underline{u}}^{\hat{w}} G(u)^{n-2}g(u)^2\,du}. \tag{7}
$$

Recall that $\hat{w}$ is still defined by (13). This can be contrasted with the equilibrium price in the baseline model given by

$$
\hat{p}_n = \frac{1}{g(\hat{w})\left(\sum_{i=0}^{n-2} G(\hat{w})^i-(n-1)G(\hat{w})^{n-1}\right)+(n-1)n\int_{\underline{u}}^{\hat{w}} G(u)^{n-2}g(u)^2\,du}. \tag{8}
$$

It is easy to show that $\frac{nG(\hat{w})-G(\hat{w})^n}{n-1} \leq 1$, therefore prices are always lower with observational learning.

Proposition 2. *Assume that $G(u)$ has bounded support. Then, for every n and every $c < \bar{c}$, the price with emulating consumers is lower than the price in the baseline model and converges to marginal cost as $c \to \bar{c}$. Further, $\hat{p}_n$ but also $p_n^*/\hat{p}_n$ decreases in n.*

As with the baseline model prices go to marginal costs (zero) with n in our model. An important distinction is that as the number of firms grows, the relative gap between prices in our and baseline model also grows, thus competition compounds the emulation effect.

As the number of firms grows, the share of returning consumers vanishes. In this case, as shown in Anderson and Renault (1999), ARW with an infinite number of firms becomes a model of price competition with differentiated goods and so the price can be computed using the following formula

$$\hat{p}_\infty = \frac{1 - G(\hat{w})}{g(\hat{w})} \tag{9}$$

where a firm optimally trades-off exploitation of *captive* consumers (represented by $1 - G(\hat{w})$ and the retention of marginal ones ($g(\hat{w})$). In our model, however, the price becomes

$$p^*_\infty = \frac{(1 - G(\hat{w}))G(\hat{w})}{g(\hat{w})}. \tag{10}$$

In our model, an additional consumer retained attracts a queue of new consumers of expected length $\frac{1}{G(\hat{w})}$. Thus, the trade-off between exploitation and retention becomes tighter and prices fall. In the basline model, a unit price increase lead to a unit increase in the reservation utility, which results in the demand loss of the order $g(\hat{w})$, which given that demand per firm is equal to $1 - G(\hat{w})$ leads to the pricing rule in (9). In our model the loss of $g(\hat{w})$ is exasperated by all the lost first visits that result from a decrease of "primary" demand by $g(\hat{w})$. The magnitude of this effect is $1/G(\hat{w})$, which leads to the pricing rule in (10). Prices are lower with emulation by the factor of $G(\hat{w})$, the expected stream of consumers who would have arrived for the first visit and stayed had the price were one unit lower. As search cost increases, $G(\hat{w})$ falls, leading to even lower prices with emulation becuase firms can attract an infinite stream of staying consumers by attracting on extra consumer.

In our model firms set competitive prices when search costs are very small but also when they are very large, and the same outcome is achieved for the opposite reasons. When search cost is small, almost all consumers search almost all firms. Because there are infinitely many firms, for each consumer any firm has a almost perfect substitute, and thus firms compete a la Bertrand. The opposite is true when search cost is high - very few consumers search more than once, and almost all their followers stay with the firm they visit first. Because searching consumers are so valuable to attract, firms are forced to charge very low prices.

It is also worth noticing that our model differs substantially from those models in which the order of search is exogenously given. In these 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, and the (exogenously) determined order, the prominence, becomes irrelevant for prices. Thus prices are the same in the standard Wolinsky model and the model with exogenous firm prominence. In contrast, in our setting, even when the number of firms grows large, the model with emulation features lower equilibrium prices. This is because while firms do not care about returning consumers,[5] they do internalize the demand multiplier through emulation.

The next figure illustrates equilibrium prices for different n. As shown in Proposition 2, prices decrease in n for a given search cost. Interestingly, in the limit case where $n \to \infty$, prices approach zero both when search costs are very small and when search costs are very large. In the former case, this reflects the fact that consumers are willing to visit a large number of stores to find a suitable product and, therefore, each of them has very small 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 highest surplus) or (almost) fully uninformed decisions (buying from the first firm they visit almost surely).

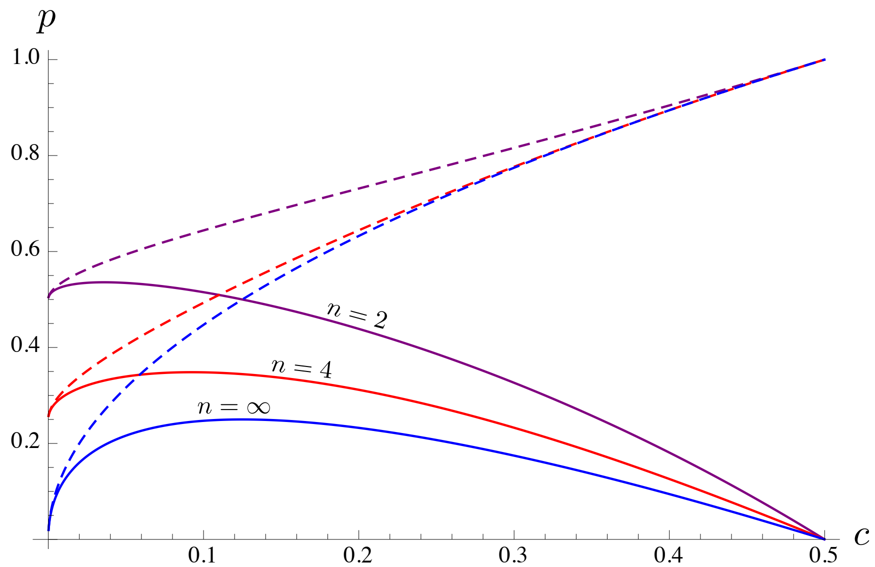

Figure 1: Equilibrium prices as function of search cost for our model (solid) and the Wolinsky model (dashed) and n equal to $2, 4$ and ∞. $G(\cdot) \sim U[0, 1]$.

3 Correlated Preferences

As suggested by the empirical evidence cited in the Introduction, consumers may follow others because they believe it is likely that, were they to have the same information, they would decide similarly. That is, their preferences are similar. A positive correlation in product matches introduces a new channel for price setting in our model. Whereas with

[5]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.

simple emulation, observational learning leads to higher price elasticity, with positive correlation of preference, this need not be true. In particular, consumers will free-ride on each others efforts and search less than in the model where their preferences are independent. If free-riding leads to less search with than without observational learning search, then prices could be higher because firms take advantage of the free-riding by increasing prices.

We introduce this into the model by adding correlation in the realization of the utility draws of each consumer in a given store. In particular, 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 ($G_L(u)$). Both distributions are equally likely to realize at each store, 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 densities. As is 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.[6] Thus consumers' preferences are correlated because they are drawn from the same distribution (G_H or G_L), and utility realizations are informative about the realized distribution. Finally let $(S_1, S_2)_k \in \{H, L\}^2$ be the realized state. Even though consumers do not know the utility that the consumer they observed derived from her chosen variety, the fact that that consumer buys from a firm is informative about which state $(S_1, S_2)_k$ has realized.

In order to simplify the analysis we focus on the case where firms are arbitrarily patient (i.e. $\delta \to 1$) and the number of individuals is arbitrarily large (but finite). .

3.1 Consumer Behavior

As in the Baseline Model, the consumer who arrives in period 1, lacking any information to discriminate firms, makes her first visit randomly and buy there if and only if $u_i 1 - p_i > \hat{w} - p^*$, where $\hat{w}$ is computed as in the baseline model without learning. In order to simplify analysis, we assume that all remaining consumers hold a common prior $\nu(j)$ over the distribution of their arrival time. This allows us to specify the same search rule for all consumer other than the first one. As shown by Monzón and Rapp (2014), the standard herding model does not change qualitatively if the same type of position uncertainty is introduced there.

All consumers after the first one observe their predecessor buying from firm $i \in \{1, 2\}$ and expect the same price p^* in both stores before embarking on their first search. Since valuations are positively correlated, the expected surplus from firm visiting firm i is larger

[6]Notice that Log-concavity of G_H and G_L does not guarantee Log-concavity of G.

than that of firm $-i$ and so the consumer should visit that store. Upon visit consumer learns his utility realization for good i, u_{ij} and the price that the firm i charges. Consumer will search if and only if $u_{ij} < w(p_i)$, where $w(p_i)$ solves

$$\int_{w-p_j+p^*}^{\bar{u}} (u - w + p_i - p^*)(q(w;p_i)g_H(u) + (1 - q(w;p_i))g_L(u))du = c \tag{11}$$

where $q(u;p)$ is the (posterior) probability that at the other store the valuations are drawn from the high distribution. This probability will, in general, depend on the valuation drawn from firm j (because preferences are correlated) and the price of firm i because it affects the probability consumers buy at that firm, and thus conditional distribution of the other store. In principle, $q(u;p)$ is a very complicated object, since each consumer may infer from u not only how likely it is that a given distribution realized but also his cohort (which is potentially informative about the informational content of the purchasing decision of the predecessor).

In order to understand $q(w;p)$, we first look at the probability that consumer j buys from firm 1 if the state is $(SS') \in \{H, L\}^2$. Since the probability with which she visits this firm first is equal to the probability that her predecessor bought there, $x_{SS'}^{i-1}$, we have that a consumer arriving in cohort i buys with probability

$$x_{SS'}^{i}(p,p^*) = x_{SS'}^{i-1}(p,p^*)(1 - G^S(w(p))) + (1 - x_{SS'}^{i-1}(p,p^*))G_{S'}(w^*)(1 - G_S(w^* + p - p^*)) +$$
$$+ x_{SS'}^{i-1}(p,p^*) \int_{\underline{u}}^{w(p)} G_{S'}(u - p + p^*)g_S(u)du + (1 - x_{SS'}^{i-1}(p,p^*)) \int_{\underline{u}}^{w^*+p-p^*} G_{S'}(u - p + p^*)g_S(u)du.$$

which increases in $x_{SS'}^{i-1}$, leading to a link between past and current market shares. It is straightforward to see that this mapping has a fixed point where market shares are stationary ($x_{SS'}(p_1,p_2)$). In Appendix 1 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 (12) where we impose $x_{SS'}^{i} = x_{SS'}^{i-1} = x_{SS'}$.

Once we have market shares in every state, that depend on p, p^*, but alto on $w(p)$ that yet needs tone computed, we find $q(u;p)$ as

$$q^*(u;p) = \frac{x_{HH}(p)\lambda(u) + x_{LH}(p)(1 - \lambda(u))}{(x_{HH}(p) + x_{HL}(p))\lambda(u) + (x_{LH}(p) + x_{LL}(p))(1 - \lambda(u))} \tag{12}$$

The above formula uses market shares in every state and weights the conditional probability of H given u by these market shares. As expected, if in all states both firms have equal market share ($x_{SS'} = 1/2$), which given our assumption about G_H and G_L imply $G_L = G_H$, the first visit is not informative and $q^*(u;p) = 1/2$. Notice that $q^*(u;p)$ is increasing in u because λ is an increasing function and the market shares satisfy $x_{HH} \geq x_{LH}$ and $x_{HL} \geq x_{LL}$. Intuitively, a higher u leads the consumer to update upwards the probability he attaches to the current firm having a high distribution ($S = H$). Importantly, this also increases the belief he holds about the other firm's distribution, since it reduces the informativeness of the predecessor's purchase at this firm regarding the other firm's distribution.

This completes the characterization of the consumer search rule. In particular, $w(p)$ is implicitly defined in (11) where $q(u; p)$ is given by (12) where $x_{SS'}$ is defined in (12).

3.2 Equilibrium Conditions

In a symmetric equilibrium, prices are equal and, thus, do not affect search behavior. The consumers' search rule can be rewritten as

$$\int_{\hat{w}}^{\bar{u}} (u - \hat{w})(q(\hat{w})g_H(u) + (1 - q(\hat{w}))g_L(u))du = c \tag{13}$$

where $q(\hat{w}) = q(\hat{w}; p^*)$ is the equilibrium probability that the rival firm has the High distribution given the reservation utility. Notice that $q(u) \leq q(\bar{u}) \leq \frac{1}{2}$ since observing a previous consumer buying in a store is bad news about the prospects in the rival store. This effect mitigated if the consumer learns that the current firm offers high valuations, in which case the probability that the other firm has a high valuation remains high. Hence, we have the following trivial observation.

Proposition 3. *Fix $c > 0$, in the unique symmetric equilibrium, $\hat{w} < w^*$, and, hence, consumers free-ride on others' effort.*

Firms' profits depend on the search rule used by consumers both on and off-the-equilibrium path. In particular, let $w'(p)$ be the implicit derivative of the search cutoff with respect to p. In the baseline model without learning, $w'(p) = 1$ so that an increase in price is compensated with a higher required utility. This is no longer the case with observational learning. Consumers visiting 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.[7] 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, $w'(p) < 1$. Given this search behavior, consumers are split into four groups in the (u_1, u_2) space. As is standard in Wolinsky-type models, those whose valuation for both varieties is lower than the reservation utility $\hat{w}$ search independently of the variety they sample first and then buy from the highest-surplus offering firm. Those whose utility profile satisfies $u_j < \hat{w} < u_{-}j$ only search if they visit firm j and always buy from $-j$. Finally, those

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

whose valuations are higher than their reservation utility buy from the firm they visit first. This group's demand is relatively inelastic and in most consumer search models is assigned randomly across firms. In our model, these consumers are assigned to each firm with probabilities equal to market shares and, thus, become quite elastic. We term this the *emulation* effect of observational learning.

Expected demand for a firm that charges p while its competitor charges p^* and consumers use reservation utility function $w(p)$ is the average of demands over all possible states. Because states between firms are independent, and equally likely, demand is given by:

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

Here $x(p, p^*)$ implicitly depends on $w(p)$ and thus on consumer search rule.

In equilibrium, symmetric equilibrium price satisfies the following condition

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

and further $\sum_{SS'} x_{SS'}(p^*, p^*) = \frac{1}{2}$ by symmetry. So the equilibrium pricing rule simplifies to

$$\frac{1}{2} + p^* \sum_{SS'} \frac{\partial x_{SS'}(p_1, p^*)}{\partial p_1} = 0 \tag{14}$$

In order to solve the model, then, we need to characterize the elasticity of demand in each state with respect to the price. This elasticity crucially depends on the response of consumers to a marginal change in prices by one firm. On its turn, this response depends on the beliefs they hold about the quality of the other store which depends on the demand in each state. In what follows we provide numerical solutions for the Uniform distribution and analytically characterize the limit when the mass of searchers vanishes for any distribution.

4 Equilibrium with Learning

The effects of learning on prices is ex-ante ambiguous because learning adds two competing effects. First, as highlighted by Proposition 1, for a given c and a symmetric price p^*, $\hat{w}$ is lower the bigger the difference between distributions. In equilibrium, a consumer arriving at a store offering low surplus correctly infers that her predecessor was also likely to obtain a low surplus which induces her to update her beliefs about the competing store's distribution of valuations. This induces the *level* of search to decrease. Additionally, firms are tempted to raise prices in order to distort learning. If a firm increases its price in the Baseline model, the consumer adjusts her cutoff valuation so that $\hat{w} - p + p^*$

is independent of p. If learning is present, however, higher prices induce a different distribution of expected market shares and thus affects the posterior beliefs that a given consumer holds. In particular, if a firm deviated to a higher price, its relative demand on those states where the competitor has a High distribution of valuations decrease with respect to that in which the competitor offers a Low distribution. Thus, higher prices lead consumers to adjust their beliefs about the other store's valuations downwards and the consumer would willingly stop her search with lower surplus ($\hat{w}(p) - p + p^*$ decreases in p), so that $\hat{w}'(p) < 1$). This *marginal* change in the search intensity pushes prices upwards.

When search costs are low, prices are increasing in search costs and learning induces prices to raise. As the cost of search increases, however, the level effect of free-riding joins forces with emulation and prices decrease faster and converge to marginal cost for lower search costs than without learning. In particular, if the likelihood ratio at the lower bound converges to zero, the maximum search cost at which a pure-strategy equilibrium exists equals the expected valuation of G_L. To see this, notice that a consumer who arrives at a store offering sufficiently low utility is almost sure that the current firm's distribution of valuations is Low. But because the probability that her predecessor bought in this store if the rival firm offers High valuations converges to

$$\lim_{c \to \bar{c}} x_{SS'}(p, p^*) = \frac{G_{S'}(\hat{w})}{G_{S'}(\hat{w}) + G_S(w(p))} \tag{15}$$

If $\lambda(\underline{u}) = 0$, the probability that the rival firm offers a High distribution of valuations is arbitrarily small and $x_{L,H}(p^*, p^*) = 0$.Thus, $q(\underline{u}) = 0$. Define $\bar{c}_L = \int_{\underline{u}}^{\bar{u}} u g_L(u) du$ to be the largest search cost that there is search in equilibrium. Since prices approach marginal costs as the share of movers vanishes, we have the following counterpart of Proposition 3:

Proposition 4. *Suppose G has a finite support and $\lambda(\underline{u}) = 0$. As $c \to \bar{c}_L$, we have $p^* \to 0$.*

Prices converge to zero as the search cost converges to the expected valuation of G_L, since in such a case total demand depends on the behavior of those individuals with lower valuations at each store and, for search cost high enough, their demand becomes infinitely elastic.

4.1 Comparative Statics

In order to disentangle changes in information from changes in the distribution of valuations, it is convenient to define two auxiliary distributions $G_1(u)$ and $G_2(u)$ such that $G_H(u) = (1-r)G(u) + rG_1(u)$ while $G_L(u) = (1-r)G(u) + rG_L(u)$ and $G(u) = \frac{1}{2}G_1(u) + \frac{1}{2}G_1(u)$, where $r \in [0, 1]$ measures the degree of correlation across consumers. This specification guarantees that for any r the unconditional distribution (independent of the state S) for a firm is $G(u)$, while r controls how close G_L is to G_1 and G_H is to G_2. If $r = 0$ valuations are independent, and drawn from $G(u)$, while for $r > 0$ valuations are positively correlated. Conveniently, for every r the unconditional distribution of valuations remains constant but the amount of information contained in the purchasing decision of a predecessor may vary significantly (as allowed by the disparity between G_1 and G_2). Thus r proxies correlation of valuations among consumers.

The following Corollary suggests that correlation and learning may lead to lower prices for high enough search costs. Let $\hat{p}(r)$ be the price if the distribution is governed by $r \in [0,1]$ and recall that higher r correspond to higher correlation and, therefore, higher informativeness of the purchasing decision of a predecessor.

Corollary 2. *Take a distribution G. There exists $\epsilon > 0$, such that for $c \in (\bar{c}_L - \epsilon, \bar{c}_L)$, $p^*(1) = 0 < p^*(0)$.*

That is, as the informativeness of the purchasing decision increases, the equilibrium price elasticity increases and, thus, prices decrease. This is because those consumers whose valuation for the variety they sample first is low become increasingly pessimistic about their prospects in the other store the higher is the correlation across consumers, and, therefore, they are less inclined to search for the same cost. As the elasticity of demand decreases in the proportion of searchers when the later is sufficiently high, the result follows.

The following figures illustrate our results for a uniform distribution G on $[0,1]$. In this example G_1 is a triangular distribution on $[0,1]$ with mode 0 while G_2 is a triangular distribution on $[0,1]$ with model 1. As required, the mixture of the two is uniform on $[0,1]$, G_H FOSD G_L and the likelihood ration is monotone for any $r > 0$.

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

As expected, prices are always lower in the model with pure emulation (black) than in the baseline model. As predicted by the theory, the equilibrium price in the baseline model is increasing in s and reaches 1 at $c = 1/2$. This is where even a consumer who draws 0 at the first store refuses to search further (after this, prices are infinite because outside option is $-\infty$). In our model with learning ($r = 1$), as predicted by Proposition 4, because $\underline{u} = 0 > -\infty$, price is zero at $c = 1/3$, which is the average of $G_L = G_1$. In our model without learning ($r = 0$), price converges to zero at $c = 1/2$, as implied by Corollary 2. Intuition for both is simple. When $c = 1/3$ in the model with learning, a consumer who is sure that utilities from the other store are drawn from G_L will not search. But this is what happens in equilibrium when a consumer draws utility close to $\underline{u}$ - she reasons that because distribution in the current firm is almost surely G_L, the fact that she came here indicates that in the other store the distribution is also G_L, or otherwise almost all queues would end up at the other store. Thus even though u close to $\underline{u}$ is bad news about the current store, it is also bad news about the next store.

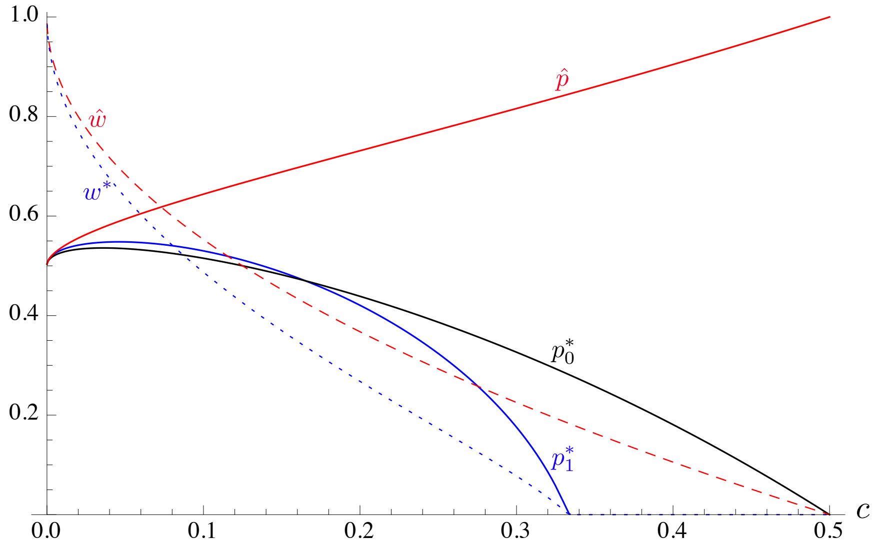

Figure 2: Prices in our model with $r = 1$ (blue), $r = 0$ (black) and baseline model (red) for triangular distributions.

Figure 3 shows $w'(p)$ as a function of c. For $c \to 0$, the derivative of reservation utility approaches 1. This is because when almost all consumers search, price contains almost no information and so reservation utility matches it one for one. For all $c > 0$, $w'(p)$ is less than 1, reflecting the informational content of price deviations. Moreover, as the second part of Figure 4 shows, $w'(p)$ can be as low as zero. This means that a price increase results in (almost) no change in reservation utility. Would would imagine that prices should be very high in this case, but in fact they are nearly zero. To understand this,

Finally, Figure 4 illustrates equilibrium demands in various states as functions of c. In symmetric states (LL and SS), demand for both firms is equal to 0.5 and is independent of 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 small some consumers buy at the store where utility is drawn from G_L, and so the same proportion of consumers visits this store first. As $c \to 1/3$, search vanishes and almost all consumers are steered first to the firm with high distribution, and because $\hat{w}$ approaches 0, almost all stay there. Thus, a *herd* forms where all consumers make the first visit to the store with high distribution, and then correctly avoid searching further.[8] The interesting aspect of such optimal collective behavior is that it leads with zero prices. Note though that prices are zero not because of optimal collective behavior in equilibrium, but rather for the opposite reason out of equilibrium. Namely, if a firm were to deviate and charge a higher price, even though with probability $1/4$ the state HL, almost no

[8]The meaning of "herd" in our model differs slightly from the way that it is interpreted in the bulk of the economics literature. In our model consumers do not possess pre-search information, and have to obtain it actively, thus herds in traditional sense cannot arise.

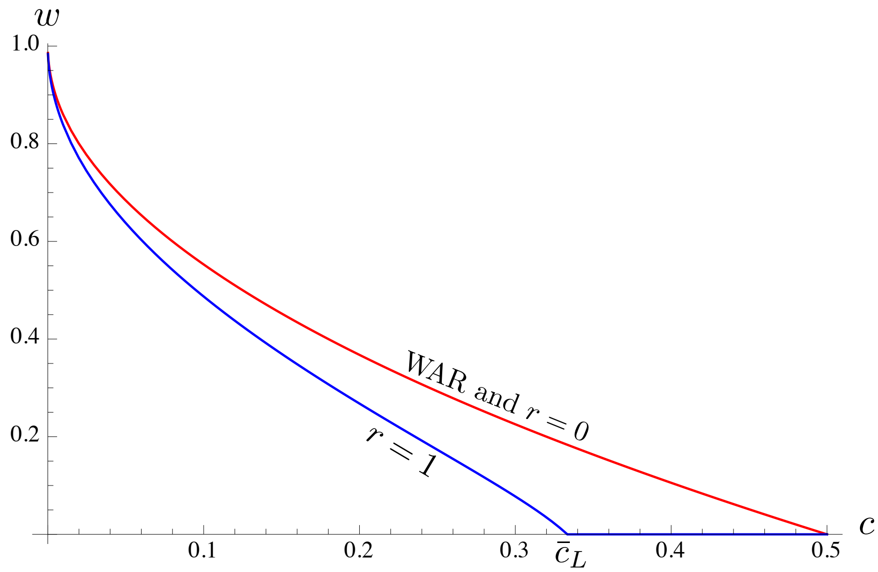

Figure 3: Reservation utilities in our model with $r = 1$ (blue) and our model with $r = 0$ and also baseline model (red) for triangular distributions.

consumers would visit it. This means that upon price deviations "bad herds" form where small price differences result in quarter of consumer incorrectly visiting a firm with a significantly worse distribution and negligibly better price.

4.2 Welfare

Which are the Welfare consequences of introducing Learning? 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:

$$W = \sum \{ x_{SS'} (\int_{\hat{w}} u dG_S(u) + G_{S'}(\hat{w}) \int_{\hat{w}} u dG_{S'}(u)) +$$
$$+ (1 - x_{SS'}) (\int_{\hat{w}} u dG_{S'}(u) + G_S(\hat{w}) \int_{\hat{w}} u dG_S(u)) + \int^{\hat{w}} (u G_S(u) dG_{S'}(u) + u G_{S'}(u) dG_S(u)) \}$$

The first result is a rather obvious one. Namely,

Proposition 5. *Welfare is higher if consumers observe others.*

A more interesting result concerns the comparative statics of welfare with respect to search costs. Information about predecessors and search act as substitute sources of information. As search cost vanish, search is a more efficient signal and so the prior information becomes irrelevant. On the other hand, as search costs increase, information about predecessors becomes increasingly important.

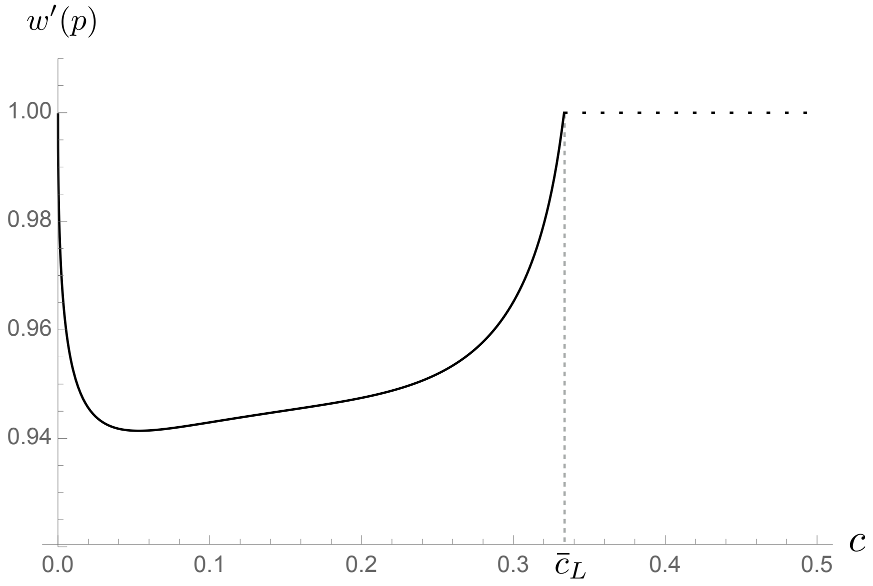

Figure 4: Derivative of reservation utility with respect to p in our model for $r = 1$ and triangular distributions.

Proposition 6. *Increasing search costs has a non monotone effect on welfare. For arbitrarily low search costs, it decreases welfare. On the other hand, for high enough search costs there are distributions G_H and G_L such that welfare increases in search costs.*

The first part is rather trivial. If search costs are low enough, almost every consumer searches and lowering search costs further reduces their expenditure. Since consumers do not stick to their first option, information is not valuable and Welfare is enhanced. The second part, however, may be less intuitive but is definitely more interesting. When almost every consumer buys in the first store she visits, decreasing search costs has a negligible effect on total "expenditure" since very few consumers search. Marginal consumers decide to search but those do not matter for welfare because their gain is negligible (being $\hat{w}$ an optimal strategy). 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 predecessors. Importantly, however, their predecessors 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 predecessor. Since predecessors deem it unlikely that their predecessor searched, they follow them to the store with a Low distribution. Therefore, the net effect of a decrease in search cost depends on the relative magnitude of the direct effect in cost savings 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 search costs.

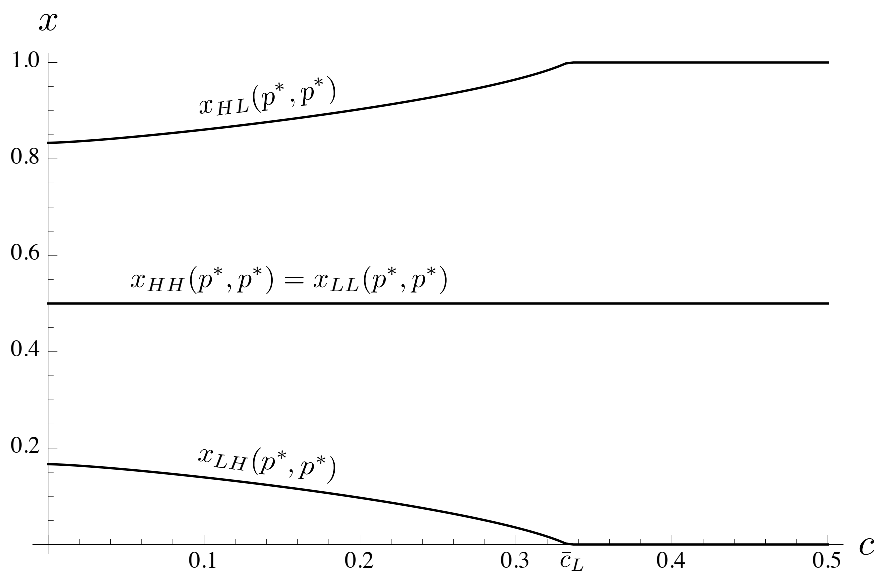

Figure 5: Equilibrium demand in various states for triangular distributions.

5 Extensions

5.1 Finite Outside Option

In the baseline model we 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.[9] This means that there is a limit to how low prices can be when search costs are high, because eventually consumer stop searching and prices jump up.

For illustration purposes, we revert back to the baseline model with two firms and no correlation of preferences. We also assume that $\delta = 1$. There is one caveat, however. Some consumers leave the market with buying anything, and so we have to specify what their predecessors observe. Since we are considering the case where firms maximize stationary profits ($\delta = 1$), it is not particularly important how the queue unfold. Therefore, we shall assume that each consumer sees a randomly drawn previous buyer, and look for fixed point where the first visits are proportional to market shares. As before, let x stand for the demand of firm 1 chat charges p_1 while the other firm charges p^*. Then x solves

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

[9]This is akin to the Diamond Paradox (Diamond (1971)), except that it occurs only when s is sufficiently high. Diamond's result holds for any $s > 0$, but is derived in an environment with homogeneous goods, while here goods are heterogenous.

where X denotes total demand, and thus x/X is firm 1's share of first visits. This expression is similar to (5) 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 $\underline{u}$.)

The total demand is easily obtained using the fact that consumers who do not purchase end up searching both firms, therefore the probability that a consumer does not purchase (regardless of where she makes the first visit) 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 $p_1 = p^*$.

There resulting pricing equation is

$$p = \frac{\big(1-G(p)^2\big)\big((2-G(w))G(w)-G(p)^2\big)}{2(1-G(p)^2)\int_p^w g(v)^2\,dv+G(p)(1-2G(p)^2+(2-G(w))G(w))g(p)+(1-G(p)^2)(1-G(w))g(w)}. \tag{16}$$

For comparison, consider the standard Wolinsky pricing rule derived under the assumption that half of consumers make first visits to firm 1.

$$p = \frac{1-G(p)^2}{2\big(\int_p^w g(v)^2\,dv+G(p)g(p)+\frac{1}{2}(1-G(w))g(w)\big)}. \tag{17}$$

It is easily verified that the price with observational learning is lower than the Wolinsky price.

To close the model, recall that the above only holds for $p \leq w$, which is not the case for s sufficiently large. In both cases, there exists a threshold $\bar{s}$, such that for all s above it, the equilibrium involves monopoly prices, and no consumers search. In both cases, we need to verify that for $s > \bar{s}$, p^M is indeed larger than w. For the standard Wolinsky model, see Janssen and Shelegia (2014) for the proof. For our model, note that $\bar{s}$ is defined as s such that $p = w$. From (16), $\bar{s}$ is such that w solves

$$w = \frac{2G(w)\,(1 - G(w)^2)}{(2G(w)^2 + G(w) + 1)\,g(w)} \tag{18}$$

Given that $G(\bar{u}) = 1$, which makes the RHS equal to 0, such a solution always exits, but might not be unique.[10] If it is indeed unique, it is trivial to show that $p^M > w$, so that indeed once $s > \bar{s}$, firms charge $p^M > w$ and no consumer searches.[11]

In figure 6 below we depict equilibrium prices for $G(\cdot) \sim U[0, 1]$. The red curve depicts standard Wolinsky price as a function of s. For s relatively small, the price is below the reservation utility and increasing in s. Once s reaches the threshold $\bar{s}_1$, price is the monopoly level, and stays there for all s even larger. As shown above, the price with

[10]It is unique for uniformly distributed valuations

[11]In principle, it is possible that at the threshold, $w > p^M$, in which case no pure strategy equilibrium exists for s immediately above $\bar{s}$ becuase conditional on consumers no searching, firms charge p^M and induce search, but conditional on consumers searching, they charge $p > w$, precluding search. See Janssen and Shelegia (2014) for an example of such an equilibrium in the context of vertical relations.

emulation is lower (black curve). It is also non-monotone, but never reaches zero because once $s > \bar{s}_2$, search stops and prices jump to the monopoly level (1/2 in the example). Note that in both model when s exceeds the respective threshold, a version of the Diamond Paradox prevails - consumers do not search, and firms charge the monopoly price. Because, $\bar{s}_2 > \bar{s}_1$, with emulation the Diamond Paradox appears for higher search cost than in the baseline model. Thus emulation allows consumers to collectively escape the paradox by putting sufficient downward pressure on prices. Moreover, once the search cost exceeds $\bar{s}_2$, prices jump up discontinuously, leading to a start contrast between markets with small and large search costs. Markets with $s \leq \bar{s}_2$ are characterized by low prices, whereas markets with $s > \bar{s}_2$ are characterized with monopoly prices, In contrast, in the baseline model the transition is smooth.

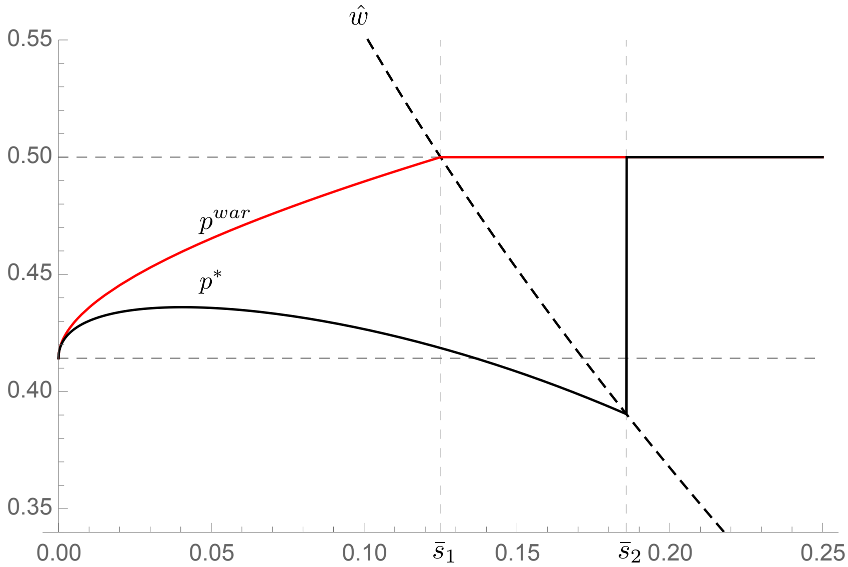

Figure 6: Equilibrium prices in the model with zero outside option for $G \sim U[0, 1]$.

To conclude this section, note that prices are never close to marginal cost here, so whether the price is decreasing in search cost depends on G. One approach is to find conditions under which price with emulation for $s = \bar{s}$ is lower than the price for $s = 0$. The condition for this is that the RHS of (18) should be below $\frac{1-G(w)^2}{2\left(\int_w^{\text{ubar}} g(v)^2\,dv + G(w)g(w)\right)}$, the RHS of the pricing equation for $s = 0$ (Perloff-Salop price). It is easily verified, that for uniform $G(\cdot)$ this always holds.

5.2 Richer Observational Learning

In our model consumers observe the purchasing decision of a single predecessor. While this assumption is rather crude, it is perhaps more plausible than assuming that consumers observe the whole sequence of purchasing decisions or the true market shares. In any case, this assumption greatly simplifies the analysis since the number of possible information sets a given consumer may end up in grows exponentially in the number of predecessors.

More importantly, in our model all consumers have the same reservation utility strategy. This would not be the case if they observe a larger set of consumers. On the other hand, if consumers observed market shares without noise, the equilibrium would converge to the mixed-strategy pricing equilibrium presented in Armstrong and Zhou (2011) for the case in which consumers observe prices. The reason for this is that even a small deviation in price would result in *all* consumers first visiting the deviating firm, thus profit would jump up discretely. This is the familiar Bertrand pressure on prices. But prices cannot be zero because for relatively small search cost a firm can still earn profit even if it is visited last. Therefore, there has to be a mixed strategy equilibrium.[12]

This shows that when consumers posses a lot of information about their predecessors' purchases, equilibrium results in mixed strategies. We analyze the other extreme where consumers have very limited information. What happens between these two extremes is extremely hard to analyze, but we believe that for relatively limited observational learning equilibrium is qualitatively similar to ours, until when a lot of learning results in mixed strategies.

Proposition 7. *Assume that every consumer observes at least one predecessor and that* $\delta \to 1$*. Then, as* $c \to \bar{c}$*, there exists an equilibrium with price* $\hat{p} \to 0$*.*

6 Discussion

Our aim in this paper was to introduce simple observational learning into a standard consumer search model with horizontally differentiated products. We did so and showed that consumer search models change qualitatively with such learning.

To achieve this goal, we have made several important assumptions that can be relaxed and or modified in future work. First, we followed Anderson and Renault (1999) in assuming that all consumers buy. This assumption greatly increases the tractability of the search model and simplifies learning across individuals. In its defense, it should be noted that an extensive margin of demand should push prices downwards. Since our main results concern surprisingly low prices, an active extensive margin would reinforce these results. Moreover, it allows us to concentrate on the *business-stealing* effect of information, since the amount of purchases is kept constant.

Second, we have introduced correlation across consumers' valuations in the simplest feasible way. If we went one step further in the direction of simplicity, and assumed that *all* consumers have the same utility, the equilibrium search rule would not have standard reservation utility property. This is because, in a putative cutoff-strategy equilibrium, the purchasing decision of a predecessor is more informative for utilities just below the

[12]This is in contrast to our results (Propositions 3 and 4) that say that prices go down to zero.There search costs are so high that almost no one searches beyond the first firm, thus not matching the other firm's low price results in zero profits.

cutoff and so search is less valuable there. Therefore, consumers would want to stop for utilities below the cutoff. As a result, the equilibrium would have to feature an interval of utilities where probability of searching further transitions smoothly from one to zero. Although very interesting, this search rule renders the full model less tractable. Moving in the direction of more general correlation of utilities is also hard to model since consumers' learning about the distribution in other stores is highly non-linear and their beliefs are formed by mixtures of truncated distributions. Hence, our model is a simple, yet appealing, compromise. We believe that pushing in either direction would be important for our further understanding of observational learning in search markets.

Finally, we have for the most part abstracted away from dynamic pricing considerations. It is highly intuitive that firms would take advantage of consumers' learning and change their prices accordingly. While we have addressed this issue in a somewhat rudimentary fashion by introducing a period where learning has been completed, price changes during the learning process may be interesting, but are fairly complicated to handle.

A Appendix 1

Proof of Proposition 1

Proof. Using the expression of demand recursively, we get

$$\Pi = p_1 \sum_{i=1}^{N} \delta^{i-1} x^i(p_1, p_2). \tag{19}$$

For simplicity, let $N = \infty$. Let $M_2(p_1, p_2)$ be the probability that a consumer purchases in store 1 if she visits it in her second visit and $M_1(p_1, p_2)$ be the probability that a consumer purchases in store 1 in her first visit. Notice that

$$x^i(p_1, p_2) = M_2(p_1, p_2) + x^{i-1}(p_1, p_2)(M_1(p_1, p_2) - M_2(p_1, p_2) \tag{20}$$

so that

$$\Pi = p_1\{(M_2(p_1, p_2) + x^0(p_1, p_2)(M_1(p_1, p_2) - M_2(p_1, p_2)) + \sum_{i=2}^{\infty} \delta^{i-1} x^i(p_1, p_2)\}$$

where $x^0 = \frac{1}{2}$ is the probability that the first consumer visits store 1. Applying this recursion we get

$$\Pi \;=\; p_1 \sum_{i=0}^{\infty} \delta^i x^0 (M_1(p_1, p_2) - M_2(p_1, p_2)) + p_1 \sum_{i=1}^{\infty} M_2(p_1, p_2)\delta^i \sum_{j=0}^{\infty} (M_1(p_1, p_2) - M_2(p_1, p_2))^j$$

$$\Pi = p_1 \frac{M_1(p_1, p_2) - M_2(p_1, p_2)}{1 - \delta(M_1(p_1, p_2) - M_2(p_1, p_2))} x^0 + p_1 M_2(p_1, p_2)\{\frac{1 - M_1(p_1, p_2) + M_2(p_1, p_2)}{(1-\delta)(1 - M_1(p_1, p_2) + M_2(p_1, p_2))}\}$$

So that

$$\Pi = \frac{p_1}{1 - \delta(M_1(p_1, p_2) - M_2(p_1, p_2))}((1-\delta)(M_1(p_1, p_2) - M_2(p_1, p_2))x^0 + M_2(p_1, p_2))$$

Taking First Order Condition with respect to p_1 and equating $p_1 = p_2$ so that demand equals 1/2 yields

$$p^* = -\frac{1 - \delta(M_1(p_1, p_2) - M_2(p_1, p_2))}{M_1'(p_1, p_2)} \tag{21}$$

which is clearly decreasing in δ since $M_1 > M_2$. Notice finally, that if $\delta = 0$, this is the standard ARW price

$$\hat{p} = -\frac{1}{M_1'(p_1, p_2)} = \frac{1}{2\int_{\underline{u}}^{\hat{w}} g^2(u)du + (1 - G(\hat{w}))g(\hat{w})} \tag{22}$$

on the other hand, for $\delta = 1$ we have

$$p^* = -\frac{1 - M_1(p_1, p_2) + M_2(p_1, p_2)}{M_1'(p_1, p_2)} = \frac{(2 - G(\hat{w}))G(\hat{w})}{2\int_{\underline{u}}^{\hat{w}} g^2(u)du + (1 - G(\hat{w}))g(\hat{w})} \tag{23}$$

□

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 1. *Suppose there is a unique cutoff $\hat{w}(p) > \underline{u}$ solving equation (11) and let $w(p)$ be the optimal cutoff rule. Then, for every $\epsilon > 0$, there exists a $T^* < \infty$ such that for all (non-negative) prices $\|w(p) - w(p)\| < \epsilon$.*

Proof. The idea for the argument follows the ideas of Lemma 2 in Thomas and Cripps (2014), although our model is much simpler. Let $M_1(p_1, p^*; \tilde{w})$ and $M_2(p_1, p^*; \tilde{w})$ represent the probabilities that a consumer who uses a reservation-utility strategy $\tilde{w}$ buys from firm i if he visits this firm first and second respectively. These probabilities are independent of t. The evolution of market shares can be readily computed as

$$|x^t - x^{t-1}| = |x^{t-1} - x^{t-2}|(M_1(p_1, p^*) - M_2(p_1, p^*)) \tag{24}$$

Notice that that $M_1 - M_2 \leq \max_S(1 - G_S(\tilde{w}(p_2))$.[13] If p_2 is such that $\tilde{w}(p_j) = \bar{u}$, then market shares are independent of time because all consumers go through firm $-j$. If $\tilde{w}(p_j) = \underline{u}$, firm j is an absorbing state of the process independently of p_{-j} and S'. Hence, assume that p_j is such that $G_S(\tilde{w}(p_j)) \in (0, 1)$[14]

$$|x^t - x^{t-1}| \leq [\max_S(1 - G_S(\tilde{w}(p_2)))]^t |x^1 - x^0| \tag{25}$$

or

$$|x^t - x^{t-1}| \leq [\max_S(1 - G_S(\tilde{w}(p_2)))]^t |x^W - \frac{1}{2}| \tag{26}$$

where x^W are 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 $\{\max_S(1 - G_S(\hat{w}(p_2))\}$ 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 $\phi > 0$. This is true, in particular for $\tilde{w} = \hat{w}$. Now, since $\lambda(u)$ and $x_{S,S'}(p)$ are continuous functions, the conditional probabilities of the different states

[13]Here M_1 is at most 1, and M_2 is at least $G_S(\tilde{w}(p_2))$.

[14]In equilibrium, $G_S(\hat{w}(p^*)) > 0$ if $c < \bar{c}_L = \int u dG_L(u)$.

given some (u, p) realization can be written as

$$q_{T_1}(u; p) = \frac{x_{HH}^{T_1}(p)\lambda(u) + x_{LH}^{T_1}(p)(1-\lambda(u))}{(x_{HH}^{T_1}(p) + x_{HL}^{T_1}(p))\lambda(u) + (x_{LH}^{T_1}(p) + x_{LL}^{T_1}(p))(1-\lambda(u))}. \tag{27}$$

which, in principle, differs from q^* as computed from Equation (12). Since $x_{S,S'}^{T_1} \in [x_{S,S'} - \phi\epsilon, x_{S,S'} + \phi\epsilon]$, we can write

$$sup \parallel q_{T_1}, q^* \parallel \leq \frac{\phi\epsilon}{\lambda(x_{HH} + x_{HL}) + (1-\lambda)(x_{LH} + x_{LL})} < \frac{\phi\epsilon}{G_L(\hat{w}(0))(1 - G_L(\hat{w}(p))} \tag{28}$$

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

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

and now let $\phi = \sup_{p:\bar{u} > \hat{w}(p) > \underline{u}} \frac{1}{2} G_L(\hat{w}(0))(1 - G_L(\hat{w}(p)))\underline{g}(u)$, 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 (11), 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 $\frac{1}{\underline{g}(u)} \geq 0$, this implies that $\|w(p_i) - w(p_i)\| < \epsilon$ as required. □

Proof of Proposition 3

Proof. Form (11), it is obvious that w is continuous and decreasing in c. Hence, it suffices to show that $p^*(\underline{u}) < p^*(\bar{u})$. To see this, notice that

$$p^*(\bar{u}) = \bar{p}(\underline{u}) = \frac{1}{2\int_{\underline{u}}^{\bar{u}} g^2(u)du} \tag{29}$$

That is, the price is inversely proportional to the "mean density". Clearly, for all continuous distributions, $p^*(\bar{u}) > 0$. On the other hand $p^*(\underline{u})$ satisfies

$$p^*(\underline{u}) = \frac{2G(\underline{u})}{g(\underline{u})} \tag{30}$$

which is zero for all distributions with finite lower bound. □

Proof of Corollary 1

Proof. $G_L(u)$ is stochastically increasing in r, meaning that $\bar{c}_L$ is decreasing in r. Since $\hat{p}$ is strictly positive for $c < \bar{c}_L$ and $p(\bar{c}_L; r) = 0$, the result follows. □

Proof of Proposition 5

Proof. As shown earlier for $n = 2$, price is lower in our model with emulation then in the baseline model. This is also true for any n because $\frac{nG(\hat{w})-G(\hat{w})^n}{(n-1)} < 1$. To see this notice that $\frac{nG(\hat{w})-G(\hat{w})^n}{(n-1)} < 1$ is equivalent to

$$n > \frac{1 - G^n(\hat{w})}{1 - G(\hat{w})} \tag{31}$$

$$= \sum_{j=0}^{n-1} G^j(\hat{w}) \tag{32}$$

but $\sum_{j=0}^{n-1} G^j(\hat{w}) < \sum_{j=0}^{n-1} 1 = n$. To see the second part notice that

$$\frac{nG(\hat{w}) - G^n(\hat{w})}{n-1} \leq \frac{(n-1)G(\hat{w}) - G^{(n-1)}}{n-2} \tag{33}$$

since

$$G(\hat{w})\frac{(n-1)^2 - n(n-2)}{(n-1)(n-2)} \geq G^{(n-1)}(\hat{w})\frac{(n-1) - G(n-2)}{(n-1)(n-2)} \tag{34}$$

which can be rewritten as

$$n \geq G(\hat{w})^{n-2}(n(1 - G(\hat{w})) - (1 - 2G(\hat{w}))) \tag{35}$$

for $n \geq 2$. It is easy to see that a sufficient condition is

$$nG(\hat{w}) \geq 2G(\hat{w}) - 1 \tag{36}$$

which holds trivially for $n \geq 2$. Since the price in ARW decreases in n, the price in our model must be decreasing.

For the second part of the statement, we prove that for any number of firms, if $\underline{u} > -\infty$, prices must decrease in search cost. To see this notice that the price decreases in n and $\hat{w}$ is independent of n so that $p^*_\infty \leq p^*_n \leq p^*$. But since $p^*_\infty > 0$ for $c \in (0, \bar{c}_L)$ and $\lim_{c \to \bar{c}_L} p^* = 0$, $p^*_n > p^*_{n+1}$ for every n, $0 < c < \bar{c}_l$. □

Proof of Proposition 5

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 store 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. □

Proof of Proposition 6

Proof. By the Envelope Theorem, since $\hat{w}$ is chosen optimally given c, the only effects that a marginal change in c has on W come from the direct effect of savings in search expenditures and the indirect effect through 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 $\frac{\partial x_{HL}}{\partial \hat{w}}$.

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

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

$$\frac{\partial \hat{w}}{\partial c} = - \frac{1}{1 - \frac{1}{4} \sum x_{SS'} G_{S'}(\hat{w})} < - \frac{1}{1 - G_H(\hat{w})} \tag{41}$$

Thus, welfare increases in search costs if

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

which yields

$$(G_L(\hat{w}) g_H(\hat{w}) - G_H(\hat{w}) g_L(\hat{w}))(V_H(\hat{w}) - V_L(\hat{w})) \geq 2(G_H(\hat{w}) + G_L(\hat{w}))^2)(1 - G_H(\hat{w})) \sum \frac{G_j(\hat{w}) G_i(\hat{w})}{G_i(\hat{w}) + G_j(\hat{w})} \tag{42}$$

which is true for some distributions. For instance, for the triangular distribution we have that $0 < V_H - V_L \approx \frac{1}{3}$ when $\hat{w}$ is small enough and so

$$(\hat{w})^2 \frac{1}{3} > (\hat{w})^2 (1 - (\hat{w})^2)\{\hat{w} + (\hat{w})(2\hat{w} - (\hat{w})^2)\} = (\hat{w})^3 (1 - (\hat{w})^2)(1 + 2\hat{w} - (\hat{w})^2) \quad (43)$$

which holds for $\hat{w}$ small enough. □

Proof of Proposition 7

Proof. Given that (almost) every individual observes at least one predecessor, they sample first the firm they see more often. Denote by $\phi_k(x)$ the probability that a firm with market share x is the most popular among k random consumers.[15] Notice that, by definition $\phi_k(x) = 1 - \phi_k(1-x)$ and $\phi_k(0) = 0$. Further notice that for all $x > \frac{1}{2}$, $\phi_k(x) \geq \phi_1(x)$ and $\phi_k(x)$ is convex. The stationary market shares for a firm charging p given an equilibrium price p^* can then be computed as a Fixed Point of the following functional mapping

$$T(x) = \phi_k(x)(M_1(p_1, p^*) - M_2(p_1, p^*)) + M_2(p_1, p^*) \quad (44)$$

Suppose, without loss of generality, that $p \leq p^*$. Since $T(1) = M_1(p_1, p^*) < 1$ and $T(\frac{1}{2}) \geq \frac{1}{2}$, and $\phi_k(x)$ is convex, the Fixed Point is unique. Denote it by $x^k(p_1, p^*)$. Because $\phi_k(x) \geq x$, $x^k(p_1, p^*) \geq x^1(p_1, p^*)$. Thus, the demand for a firm undercutting its rival if individuals observe k predecessors is not lower than the demand she would obtain if they observe one. As the share of searchers vanish, we have that

$$x^k(p_1, p^*) \leq \frac{G(\hat{w})}{G(\hat{w}) + G(\hat{w} + p - p^*)} = x^1(p_1, p^*) \quad (45)$$

As shown in the Proof of Proposition ? for the case with one predecessor, in the limit when $c \to \bar{c}$, if the rival firm charges $p^* > \epsilon(c)$, it is profitable to undercut. Now, the demand increases and so, locally at least, the incentive to undercut is higher. Thus, the equilibrium must involve $p^* \leq \epsilon(c)$, with $\epsilon(\bar{c}) = 0$. □

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.

[15]In case of a tie, the consumer visits any firm at random

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

_ **and Yongmin Chen**, "Inattentive Consumers and Product Quality," *Journal of the European Economic Association*, 04-05 2009, *7* (2-3), 411–422.

_ **, 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.

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.

_ **,** _ **,** _ **, and** _ , "Monopoly pricing in the binary herding model," *Economic Theory*, 2008, *37* (2), 203–241.

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.

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.

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.

Mobius, Markus M., Paul Niehaus, and Tanya S. Rosenblat, "Social learning and consumer demand," *Harvard University, mimeograph. December*, 2005.

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," Technical Report 2014.

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

Thomas, Caroline D. and Martin W. Cripps, "Strategic Experimentation in Queues," Technical Report 2014.

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.