Essays in Search for Match Quality

by

Kosti Takala

B.A. Economics, Aalto University (2011)
M.A. Economics, Aalto University (2013)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

Signature redacted

Author . . ..
Department of Economics
August 10, 2018

Signature redacted

Certified by . . ..
Glenn Ellison
Gregory K. Palm Professor of Economics
Thesis Supervisor

Signature redacted

Certified by ....
Michael Whinston
Sloan Fellows Professor of Management
Thesis Supervisor

Signature redacted

Accepted by . ...
Ricardo Caballero
Ford International Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Essays in Search for Match Quality

by

Kosti Takala

Submitted to the Department of Economics
on August 10, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

This thesis consists of three essays on search for reviews. In Chapter 1, I study how homogeneous consumers behave when they are considering to buy a new experience good, such as a pair of shoes, from a monopolist with an observable price. They have costly access to consumer reviews which are perfectly or imperfectly informative of match quality. Knowing how consumers behave, the monopolist sets an optimal price inducing certain behavior; sometimes the firm will find it optimal to set a high price to induce the consumers to search for earlier reviews, sometimes a low price to induce them to purchase the product. Consumer, producer and social surplus are non-monotone in search cost, and this result extends to settings with heterogeneous consumers.

In Chapter 2, we extend the model of Chapter 1 to allow for competition between ex-ante identical firms. Consumers are homogeneous and all prices observed at no cost. They can search for earlier reviews which perfectly reveal match quality. Consumers can keep searching as long as they have not found a match or exhausted all of their options. We learn that high search costs lead to relatively high prices but when costs decrease, surplus is first transferred from firms to consumers but further reductions may decrease consumer surplus. Additional effects are noted as reviews get even cheaper to access, and consumer surplus, profits, and social surplus are all non-monotone in the cost of reading reviews.

In Chapter 3, we analyze selection into consumption in the case of movies. People leaving reviews are assumed to be a representative sample of consumers, so that reviews perfectly reveal experiences. New consumers observe these reviews but do not know if their preferences are aligned with those of the reviewers. Examining two datasets with movie reviews and box office revenue, both in a cross-section of movies and within movies over time, we learn that selection decreases in the expected quality of a movie, the precision of the prior quality, and consumer homogeneity. Selection may increase or decrease over time and it tends to increase in the number of movies.

Thesis Supervisor: Glenn Ellison
Title: Gregory K. Palm Professor of Economics

Thesis Supervisor: Michael Whinston
Title: Sloan Fellows Professor of Management

Acknowledgments

I am grateful to my advisors Glenn Ellison, Michael Whinston and Alessandro Bonatti for their continuous guidance, encouragement and support in the past few years.

I would like to thank Bengt Holmström, Arda Gitmez, and Harry Pei for giving me their time and help whenever I needed it. My discussions with Bengt were always more philosophical than academic; Harry was an excellent officemate for two and half years and a big help in compiling this thesis; and Arda, also an officemate, accompanied me to dozens of movies that inspired and sparked me to write.

I also offer my thanks to Mark Hou, Ameaya Muley, and Javier Garcia-Arenas who were excellent housemates for two years, which included a collapsed roof and a burst pipe during the February of 2015. That was actually the only time it rained inside. Special thanks to Mark for being a reliable all-you-can-eat sushi companion.

Finally, I would like to thank my family for their unconditional support over the last five years. Special thanks to Florence Lee for having been an understanding partner throughout the years of research that culminated in this dissertation. The countless gallons of low-fat ice cream would have been far from cool, had you not been there to warm me up.

Contents

List of Figures

List of Tables

Chapter 1

Product Reviews and the Curse of Easy Information

When there is a new experience good of unknown quality and known price, consumers would like to obtain information on match quality before making their purchase decisions. One way to acquire this information is to search for product reviews which represent the experiences of earlier consumers. In this paper, I study product reviews as signals of match quality. By considering a Bayesian learning model with homogeneous consumers, one product and one period, I analyze the equilibrium behavior of the firm and the consumers, and show that consumer surplus is a non-monotone function of search cost – with an interior maximum. This is due to endogenous pricing, where price may be used to induce more search when search cost is low. The result is shown to extend to heterogeneous consumers, where even social surplus is maximized at a strictly positive search cost. Finally, I show that a highly informative signal technology may be as bad for the consumer as a low search cost – they are two sides of the same coin.

I am thankful to Glenn Ellison and Michael Whinston for their valuable feedback. I have also greatly benefited from my conversations with Alessandro Bonatti and Harry Di Pei.

1.1 Introduction

Product reviews from earlier consumers have become an increasingly important way of acquiring information on product quality, both on the vertical (absolute) and the horizontal (individual match) component, attracting a fair share of recent research (see Chevalier and Mayzlin (2006), among others). It is reasonable to think that technology has made obtaining product information cheaper. Understanding how consumers use reviews when making their decisions is important because it seems to affect observed prices, purchase decisions, and therefore also welfare. However, there has not been a satisfactory model of gradual learning about match quality to explain how and why the ease of search and quality of information matter in equilibrium – when price is endogenous. This paper fills that gap by studying how the mere existence of these reviews matters for the pricing strategy of the firm, resulting consumer surplus, and overall welfare – as a function of the search cost. More precise information may lead to a loss in surplus because searching is costly and sometimes unnecessary from a social point of view. The availability of information practically forces the consumer to use it because she cannot commit to not doing so.[1]

The availability of reviews on hotels, restaurants, movies, services, and all kinds of other experience goods suggests that they serve a purpose but the jury is out on whether or not they improve consumer and overall welfare. Chevalier and Mayzlin (2006) have shown empirically that book reviews matter for sales but not whether they improve consumer surplus – or even social surplus for that matter. They also obtain evidence that consumers not only consider the average rating but also read the content of a review. In everything that follows, I will assume that reviews are truthful but will discuss how promotional reviews could be incorporated in my model and what the predictions would be.[2]

Much of earlier search literature has focused on the search for the lowest price (e.g. Diamond (1971) and Stahl (1989)) and obfuscation (see Ellison and Ellison (2009)), but here price is known and search is for the unknown match quality of a product. In Stahl (1989), equilibrium prices decrease when search becomes cheaper but reviews more often provide information about product attributes or quality. The model of Kessler (1998) is among the examples in the literature where more information is not necessarily beneficial because of strategic pricing. This motivates the study of how product reviews matter for consumer surplus, profit, and social welfare in a model where firms set prices endogenously, knowing how prices impact consumer search.

My aim is not merely to show that the consumer-optimal search cost is strictly positive, but to derive it analytically and study the complementarity between search cost and signal precision. More importantly, this paper contributes to understanding how each price leads to certain consumer behavior, be it searching three times or not searching at all, and how this implies that the firm's

[1]This is slightly in contrast with some of the concurrent research that tries to improve the informational content of an overall rating in terms of vertical quality without considering the horizontal match component (see Dai, Jin, Lee, and Luca (2016)).

[2]There is a growing literature on promotional reviews, prominent examples of which include Mayzlin, Dover, and Chevalier (2014) and Luca and Zervas (2016). These papers argue empirically that promotional reviews exist and are used more when the competition between two firms is intense or when reputational concerns are small.

equilibrium price is a non-monotone function of the search cost. Under some conditions, the firm prices high because it wants the consumers to gather as much information as possible and then extract the resulting surplus. Social welfare is maximized by eliminating all search costs if consumers are identical, but this is not true if we allow for heterogeneity in the population. Finally, the consumer wants the reviews to be informative but not too much so; she wants the best possible search technology given that the firm's optimal price leads to a purchase with certainty.

The Bayesian learning model of this paper shares features with some previous papers on search for match quality (Wolinsky (1984), Anderson and Renault (1999)) and consumer information gathering (Branco, Sun, and Villas-Boas (2012), Roesler and Szentes (2017)). There is one period, one firm and one product with consumer-specific quality. The firm faces a continuum of identical, risk-neutral consumers (or a single representative consumer), who freely observe the price before making their decision. The specificity of quality means that any given consumer may find the product to be a good or a bad match, independently of one another. Consumption utility is assumed to be binary, with the prior probability of a good match being π_0. This prior corresponds to the share of consumers in the population who would find the product to be of high quality (utility of 1) should they consume.[3] However, no consumer knows whether they will like the product before trying it out. The purpose of searching for reviews is to figure out whether or not people with similar preferences have enjoyed the product.[4,5] When a useful review is observed, it is either good or bad: a good product always produces good reviews but a bad product gives good reviews with probability μ.[6] This implies that, when a signal is observed, the belief either moves up by an amount that depends on the imprecision μ, or falls down to 0. The expected cost of obtaining one useful review is s, which means that the total ex-ante expected search costs are fully determined by the firm's price and the model parameters.[7] Therefore, there exists a threshold belief, $\bar{\pi}$, such that the consumer will search until this threshold is reached, and only then will she buy. If she ever gets a negative signal, she will exit.

In Section 1.3, I solve for the equilibrium of the game, show why payoffs are non-monotone, and derive an analytical expression for the optimal search cost. First, I explain how a consumer sequentially decides between searching for reviews, purchasing the good, and exiting the market

[3]One can think of the prior belief (π_0) as the current average rating of a product, assuming binary ratings. It tells the consumer and the firm that a certain share of the reviews are positive but it does not reveal which types of consumers like the product and whether or not some people dislike it because of mere, idiosyncratic, bad luck. In this sense, the purpose of reading reviews is to learn whether types with beliefs similar to yours liked or disliked the product. It is not about learning the absolute quality but where the product is located horizontally.

[4]The consumer may need to search for a while before she gets to a review that matters to her because she is not interested in reviews that seem to come from people unlike her. She may not know how much longer she has to search but she knows the expected time it takes (which is identical for good and bad products).

[5]Note also that we can think of the prior belief, π_0, as the posterior expectation after observing the average rating of a product, but before studying the actual reviews in more detail and assessing whether they are useful for you.

[6]The assumption that a good product never has any bad reviews is extreme but greatly simplifies the analysis. If we assumed that both types of the product can produce both types of reviews, the main results would remain (as long as the signal is informative enough) but the analysis would be more involved because the belief would be able to move up and down. What matters for the consumer (and the firm) is how expensive and unlikely it is for her to reach a belief where purchasing is the optimal decision.

[7]One can define the search cost as $s = \tilde{s}/\lambda$, where $\tilde{s}$ is the search cost per unit of time, and $1/\lambda$ is the expected time between signals.

when price is fixed and known. Second, having pinned down the consumer's optimal behavior for a given price, I solve for the firm's optimal price, assuming that it is no better informed than the consumer. Both players' behavior depends crucially on how costly it is for the consumer to search and how informative the reviews are. Throughout, I assume that only the consumer can observe the reviews and will abstract away from what happens when the firm knows which reviews are relevant to the consumer and how.[8,9] Finally, I show the existence of a consumer-optimal, strictly positive search cost due to a discontinuity in the firm's pricing policy. This optimal search cost makes the firm indifferent between two prices, one of which gives the consumer strictly positive surplus. However, it is socially optimal to have no search cost. The intuition for non-monotone equilibrium utility is that, when the cost of search is low, the firm has an incentive to increase the price because it needs to choose between selling the good at a lower price and a higher probability, or selling it at a higher price and a lower probability. This implies that a low search cost can lead to low consumer surplus because the firm raises its price.

In Section 1.4, I introduce a continuum of consumers with uniformly distributed valuations for a good match (baseline model assumed valuation was 1 for everyone), so that a consumer of type $v \in [0,1]$ gets a consumption utility of v with probability π_0 and nothing otherwise. If the cost of production is low, both consumer and social surplus will be maximized at one strictly positive value of search cost. The intuition is that low production costs imply the good should always be sold to many people but the firm will exclude low-valuation consumers in equilibrium in order to extract more surplus from the ones that end up purchasing. Equilibrium is characterized by two thresholds, one for the search cost and one for the valuation, so that for search costs lower than the threshold, the firm prefers to set a high price and make high-value consumers search. These consumers are left with strictly positive utility. For search costs higher than the threshold, the firm prices so that participating consumers buy the good. One artifact of the model is that the consumers, whose valuation is higher than the threshold value, always make the same decision; they all search or they all buy, depending on the search cost, while the low-valuation consumers exit immediately.

In Section 1.5, I generalize the models of Sections 1.3 and 1.4 by getting rid of good and bad matches and assuming that each consumer's value follows some distribution on $[0,1]$ (uniform being the familiar case). This means that "good" matches are products for which value exceeds the price. The main results do not change, and the consumers still prefer relatively high search costs. I also study how the informativeness of the signal matters in equilibrium, showing that the consumer will

[8] When the firm is in the know, there is a signaling game where the price is a signal of quality, in the spirit of Wolinsky (1983) and Riordan (1986), among others, with one twist: the consumer can (imperfectly) verify the firm's type by paying the search cost. This game involves multiple equilibria, and is extremely difficult to solve. However, as I will justify in the next section, even if the firm knows the content of the reviews, it may not understand what aspects of those reviews are useful to which type of consumer. Therefore, it is a reasonable assumption to make that the price itself contains no information.

[9] Although not shown here, if the firm perfectly understands how the reviews matter for the consumer's decision process, then it will be able to extract all ex ante surplus in a semi-separating signaling equilibrium. Basically, higher-type firms price higher on average but low types sometimes imitate them. Full pooling does not work, at least for small values of search cost, because the consumers can then verify the firm's type by searching.

not always want the perfect signal because it may lead to high prices. This adds to the work of Roesler and Szentes (2017).

Section 1.6 explores a set of binary signal structures and analytically derives the optimal (binary) signal technology as a function of search cost, which complements Section 1.5 where we cannot get analytical results but show that imperfect signals are often optimal unless search cost is high enough. The consumer is often hurt by a better search technology (be it a low search cost or a highly informative signal) because she cannot commit to not using it should the firm set an intermediate price. Hence, price will be set high when the consumer has access to a good technology (low s and μ).

Another important result of Section 1.6 is that, even though search cost and signal informativeness can be seen as two sides of the same coin, the consumer will always choose a perfectly informative signal if she is also allowed to pick her search cost. This is because both of them are tools for controlling the firm's pricing decision but the consumer never actually ends up searching in equilibrium. This section complements the more general results obtained by Roesler and Szentes (2017) who show that a certain unbiased signal distribution maximizes consumer surplus (and minimizes the firm's profit), without any distributional assumptions. However, they assume that the firm has no production cost and that there is no cost for the consumer to use the search technology.[10]

In Appendix A.5, I consider a bad news model. This means that no news is considered good news because every signal we observe is perfect bad news. This may be seen as a realistic assumption because, while shopping online, we are usually looking for bad news as good signals are more often fabricated.[11] The results obtained in this section mirror those of the main model.

The paper is structured as follows. Section 1.1.1 offers a brief overview of the literature. In Section 1.2, I introduce and justify the main model. Section 1.3 is devoted to using the model to show how the consumer responds to the firm's price and what this means for the pricing decision. We also learn why the consumer's equilibrium utility is non-monotone in the search cost. Section 1.4 allows for heterogeneous consumers with different valuations, and the results are similar to the model where consumers are identical. Section 1.5 considers a general model and shows that imperfect signals may be consumer-optimal unless search costs are high, while Section 1.6 derives the optimal binary signal analytically. Section 3.5 concludes, giving directions for future work.

1.1.1 Prior Literature

This paper is related to the literatures on information design, optimal search under endogenous pricing, obfuscation, and word-of-mouth communication (WOM). The paper most closely related to mine is Branco, Sun, and Villas-Boas (2012). They use a continuous-time model of gradual learning to explore how much review information consumers decide to gather in a model where prices are endogenous and affect how much information will be collected. The results are very

[10]The issue of no production cost is just a matter of normalization.

[11]This is not to say bad news cannot be fabricated as well.

similar to mine: lower search costs may hurt consumer and social surplus because the firm will change its price in response, but this is not true for all beliefs or prior expected utilities. Rather than modeling consumers as getting signals, they assume that consumer values drift according to a Brownian motion for as long as they are paying a cost of s per unit of time. Their model leads to demand being linear in the price. A special feature is that, whenever the consumer searches in equilibrium, her utility is directly proportional to the firm's profit, which leads to the result that social surplus is maximized at the same search cost as consumer surplus. In my model, changes in search cost can affect both social welfare and division of surplus between the firm and consumers. They do not incorporate heterogeneity but mention it as a possible extension, while I study uniformly distributed valuations. While utility in their model is a random walk, meaning that every signal has the same informativeness, in my model signals become less (or more) informative as beliefs increase. I also study signal precision and its effects in greater detail than they do.

In Lewis and Sappington (1994), the consumers do not know their type but they observe a binary signal, the precision of which is controlled by the seller, and all this is costless. The result is that the seller will either provide no information to sell under the prior beliefs, or he will provide full information and only sell to the consumers who obtain a good signal. The intuition is that for the signal to be profit improving, the seller has to be able to raise his price from the equilibrium with no information (since all consumers are ex-ante identical), but this necessarily excludes all those who obtain a bad signal (expected posterior has to equal the prior). Therefore, if any information is provided, it should be perfect. The most important difference between my model and theirs is that the consumers in my model have to pay for the signals, and they make decisions sequentially. Also, my firm cannot choose the signal structure.

As mentioned already, my paper is not concerned with the dynamics of price setting since the firm makes a one-time decision which then influences how much the consumer will search. However, most of the quality search literature is concerned with either how search matters for price competition (see Bergemann and Välimäki (1996) for the case where quality is unknown to all parties but can be learned over time), or what the dynamically optimal price is.

More generally, my paper is separated from the bandit literature in a few ways: while that literature usually assumes a number risky arms with payoffs varying over time and arms, emphasizing the exploration-exploitation trade-off, I only have one risky arm (the firm) with a constant reward that differs from consumer type to consumer type (match quality) and can only be obtained once. The most important difference is that the reward the consumer may obtain at the end is endogenously chosen by the firm, which determines how long the consumer will search. For a slightly dated survey on this literature, see Bergemann and Valimaki (2006).

From the persuasion literature, the most closely related is Kamenica and Gentzkow (2011) where a sender is free to design an experiment to persuade the receiver to take the action the sender prefers. My paper is not directly related but has a similar flavor: the firm (sender) chooses a price to induce an experiment (review search) conducted by the consumer (receiver). If the result of this experiment is bad, there will be no exchange. If the result is good, the consumer will end up with enough consumption utility to at least cover her search cost.

It is instructive to make a comparison to the obfuscation literature. Ellison and Ellison (2009) study how sellers use obfuscation to capture a larger slice of the pie. One of the main conclusions is that a better search technology may lead to firms investing in more obfuscation, thereby reducing consumer surplus. In my paper, on the other hand, the seller is not really trying to fool anyone but to give the consumer the incentive to search even though the resulting surplus will largely end up in the seller's pocket. This is because my seller has only one decision variable, the price, which is perfectly observed by the consumer before she makes any decisions. This is in stark contrast to Ellison and Ellison (2009) because, in their paper (and in most of the literature), the consumers are searching for the lowest price and the sellers know the quality of their products.

1.2 Model: Costly Binary Signals on Quality

There is a firm selling a product of unknown quality by setting a single, fixed price, and this is the only decision the firm makes. The quality is binary and consumer specific, which means that each consumer will obtain a consumption utility of 1 or 0, independent from one another. In the model, we consider one consumer (representing a continuum of ex-ante identical consumers) who needs to decide whether or not to buy the product. To aid her decision, she can search for existing consumer reviews which help her update her beliefs. Not all of the reviews are relevant for the consumer because some may come from people whose opinions differ from hers by too much.[12] The expected cost of obtaining a useful review is s (called the search cost), which depends on how widespread her preferences are in the population.[13] There is no discounting. Timing is as follows:

1. Nature draws consumer-specific qualities, $\theta \in \{L, H\}$, according to their true distribution, and earlier consumer reviews are realized.

    - Consumption utility is binary: $u(L) = 0$, $u(H) = 1$, and an outside option yields $\bar{u} = 0$.
    - The prior on the product being of high quality is $\pi_0 := \mathbb{P}(\theta = H) \in (0, 1)$, and it is shared between the firm and the consumer.

2. The firm sets its price p without observing the quality or the reviews.[14] Production cost is $c \geq 0$

3. The consumer observes the price and sequentially decides whether or not to search for more reviews. If she stops searching after t reviews, she either buys the product or exits without buying, depending on whether her belief exceeds the price or not.

[12] Think about searching for restaurant reviews. If your main criterion is the quality of their spicy food, you should not consider reviews from people who hate spicy food.

[13] Note: one can think of the expected search cost as a function of a unit search cost and the arrival rate of signals: $s = \tilde{s}/\lambda$. However, this adds unnecessary notation, so it is dropped in the analysis.

[14] If the firm did observe the reviews, the price would act as a signal of quality. This would unnecessarily complicate the model. However, one can also argue that observing the reviews gives the firm no competitive edge because each consumer will find different reviews helpful and will obtain the reviews in a random order.

- If the consumer searches, she has to pay an expected cost of $s \geq 0$ to obtain a binary signal $\sigma_t \in \{G, B\}$, where t is the number of reviews read so far. That is, the signal can be good or bad, with the following conditional probabilities:

$$\mathbb{P}(\sigma_t = G|\theta) = \begin{cases} 1, & \text{if } \theta = H \\ \mu, & \text{if } \theta = L \end{cases}$$

4. Payoffs are realized.

The assumed signal structure means that there will be a jump in the consumer's beliefs only when a signal is observed. Given belief π_t, the posteriors can be written as:

$$\pi_{t+1} := \mathbb{P}(\theta = H|\sigma_0 = \ldots = \sigma_t = G) = \frac{\pi_t}{\pi_t + (1 - \pi_t)\mu}$$
$$\pi_b :- \mathbb{P}(\theta = H|B \in \{\sigma_0, \ldots, \sigma_t\}) = 0.$$

In particular, this structure leads to a belief that either moves up (but never reaches 1), or falls down to 0. That is, $0 < \pi_t < \pi_{t+1} < 1$ for all t, as long as the consumer has observed only good signals. Note also that the arrival rate of the signals is the same for both good and bad quality, meaning that there is no learning without signals – there is no drift in the beliefs.

Given the way the beliefs evolve, the consumer's value function after t good reviews is:

$$V(\pi_t) = \max\Big\{0, \;\; \pi_t - p, \;\; \Big(\pi_t + (1 - \pi_t)\mu\Big)V(\pi_{t+1}) - s\Big\},$$

because the consumer can exit (first term), buy (second term), or search for a review (last term), where the expected cost of obtaining a useful review is s and the probability of this review being good is $\pi_t + (1 - \pi_t)\mu$. If the review is bad, continuation value is zero.

The firm's profit function at π_0 can be written as:

$$\Pi(p, \pi_0) = \mathbb{P}(\text{ consumer purchases } | \; p, \pi_0) \; (p - c),$$

where the probability of purchase (conditional on p) will be derived in the next section.

1.3 Search, Pricing, and Welfare

In this section, I will use the above model to explore the interaction between the firm's price and the consumer's search behavior, and how the interaction depends on the search cost in equilibrium. This will lead to the observation that equilibrium utility is not monotone in the consumer's search cost. As a by-product, we will get an analytic solution for the consumer-optimal search cost. I will also discuss what forces are driving the equilibrium behavior. In a nutshell, the consumer is always hurt by a higher search cost, conditional on behaving the same way because her bargaining power is reduced. However, if increasing the search cost leads to a change in the consumer's equilibrium behavior, her utility will jump up, after which it will decrease again. This is due to the interaction between the consumer's behavior and the firm's pricing policy where price is used to control search.

1.3.1 Consumer Behavior Given Price

The consumer's strategy can be defined by a threshold belief, $\bar{\pi}$, such that, for $\pi \geq \bar{\pi}$, the consumer will buy instantly, and, for beliefs between her prior and $\bar{\pi}$, the consumer will search until she either crosses the threshold and buys, or obtains a bad review and exits (one of which will surely happen).[15] The firm will never choose a price that makes the consumer exit at the prior belief, as long as production costs are not too high, which we will assume by requiring that $c \leq \pi_0$. The threshold $\bar{\pi}$ will depend on the price (p) and the search cost (s).[16,17]

The probability of observing $k \in \mathbb{N}$ good reviews in a row is:

$$\mathbb{P}(\text{k good reviews}) = \pi_0 + (1 - \pi_0)\mu^k,$$

which is easy to understand because, if the product is good (which happens with probability π_0), the consumer will surely observe k good reviews in a row, but if the product is bad, then the probability of getting these reviews is μ^k.

Now, given the prior π_0, the purchase threshold $\bar{\pi}$, and the price, the consumer knows the number of good reviews she needs to obtain in a row for her to purchase. Thus, if we assume she needs n good reviews to purchase, her expected value (not taking the search cost into account) is:

$$\begin{aligned}\mathbb{E}[\text{value}] &= (\pi_0 + (1 - \pi_0)\mu^n)\left(\frac{\pi_0}{\pi_0 + (1 - \pi_0)\mu^n} - p\right) \\ &= \pi_0 - (\pi_0 + (1 - \pi_0)\mu^n)p \\ &= \pi_0(1 - p) - (1 - \pi_0)\mu^n p.\end{aligned}$$

Moreover, her expected cost of trying to obtain n good reviews (and exiting when this fails) is:

$$\begin{aligned}\mathbb{E}[\text{cost}] &= s(\pi_0 n + (1 - \pi_0)(1 + \mu + \mu^2 + \ldots + \mu^{n-1})) \\ &= s\left(\pi_0 n + (1 - \pi_0)\frac{1 - \mu^n}{1 - \mu}\right).\end{aligned}$$

Here, s is the expected cost of obtaining one signal, so the expected cost given that the product is good is ns, and the expected cost given that the product is bad is $\frac{1-\mu^n}{1-\mu}s$. This means that the

[15]One way to show this is as follows: The consumer prefers searching-and-buying over buying instantly if and only if $(\pi_t + (1 - \pi_t)\mu)(\frac{\pi_t}{\pi_t + (1-\pi_t)\mu} - p) - s \geq \pi_t - p \Leftrightarrow \pi_t \leq \bar{\pi} \equiv 1 - \frac{s}{(1-\mu)p}$. Since π_t is increasing in t, we will reach this threshold at some point and the consumer will prefer searching to buying as long as we are below the threshold. Note also that if she finds searching optimal at π_0, she will surely do so at $\pi_t \in (\pi_0, \bar{\pi})$ because the expected cost of reaching $\bar{\pi}$ is smaller. In equilibrium, exiting is therefore not an option at any point unless a bad signal is observed.

[16]Note, however, that the price itself depends on the model parameters (s, c). Everything also depends on the distributional assumptions, here characterized by (π_0, μ).

[17]One might think that we should also consider a lower threshold $\underline{\pi}$, below which the consumer will stop searching and choose to exit. This is true for a more general model where bad news are not conclusive but not for this model because beliefs can only go up or crash down.

consumer's ex ante expected utility from searching $n \geq 0$ times and then buying is:

$$V_n(\pi_0, p) = \pi_0 - (\pi_0 + (1-\pi_0)\mu^n)p - s\left(n\pi_0 + \frac{(1-\pi_0)(1-\mu^n)}{1-\mu}\right). \tag{1.1}$$

For the consumer to perform exactly n searches, this value function needs to be non-negative and maximized at n (where $n = 0, 1, 2, \ldots$). The following proposition tells us how many times the consumer will search for a given price.

Proposition 1.1. *Given a price, p, let n^* be the number of searches that satisfies:*

$$\frac{s}{(1-\mu)(1-\pi_{n^*-1})} < p \leq \frac{s}{(1-\mu)(1-\pi_{n^*})}. \tag{1.2}$$

If $V_{n^*}(\pi_0, p) \geq 0$, the consumer will search exactly n^* times and buy. Otherwise she will exit without search.

Condition (1.2) guarantees that the value function is maximized at n^* searches but it does not guarantee that this value is non-negative. That is why we also need the non-negativity condition, $V_{n^*}(\pi_0, p) \geq 0$. Note that the consumer will not search at all if $p \leq \frac{s}{(1-\mu)(1-\pi_0)}$.

Proof. Assume that n^* satisfies (1.2) given some price, p. For any $n \in \mathbb{N}$, it is easy to see that $V_{n+1}(\pi_0, p) \leq V_n(\pi_0, p) \Leftrightarrow p \leq \frac{s}{(1-\mu)(1-\pi_n)}$.[18] The right-hand side of this inequality is strictly increasing in n (and it grows without limit as n grows). Thus, if $V_{n^*} \geq V_{n^*+1}$ for some n^*, then $V_{n^*+1} > V_{n^*+2} > V_{n^*+3} > \ldots$ (inductively).[19] On the other hand, because $p > \frac{s}{(1-\mu)(1-\pi_{n^*-1})}$, we know that $V_{n^*} > V_{n^*-1} > \ldots > V_0$.[20] Therefore, n^* is the number of searches that maximizes the consumer's search utility. However, if this utility were negative, the consumer should exit immediately. This is why we need $V_{n^*}(\pi_0, p) \geq 0$. ☐

1.3.2 Pricing Policy

While the previous section derived the consumer's optimal search strategy for a given price (and parameters), in this section I show that the choice the firm faces is relatively simple. For Proposition 1.2 and what will follow, let us define the following (all of which are functions of s):

- $p_k := \frac{s}{(1-\mu)(1-\pi_k)}$, which is the highest price so that the consumer's optimal number of searches is **exactly** k as can be seen in (1.2), but not requiring that $V_k(\pi_0, p_k) \geq 0$.[21]
- $\Pi_k := (\pi_0 + (1-\pi_0)\mu^k)\left(\frac{s}{(1-\mu)(1-\pi_k)} - c\right)$, which is the profit at p_k.

[18] This is because, assuming the consumer has already seen n reviews, the gain from doing one more search and then buying relative to just buying is $(1-\pi_n)(1-\mu)p - s$. The reason is that your expected consumption value stays the same (π_n) but you end up saving the price in case the product is bad ($1-\pi_n$) and you get a bad signal ($1-\mu$).

[19] Because $p \leq \frac{s}{(1-\mu)(1-\pi_{n^*})}$ implies that $p \leq \frac{s}{(1-\mu)(1-\pi_{n^*+k})}$ for any $k \geq 1$.

[20] Since $p > \frac{s}{(1-\mu)(1-\pi_{n^*-1})}$ implies that $p > \frac{s}{(1-\mu)(1-\pi_{n^*-k})}$ for any $k \geq 1$.

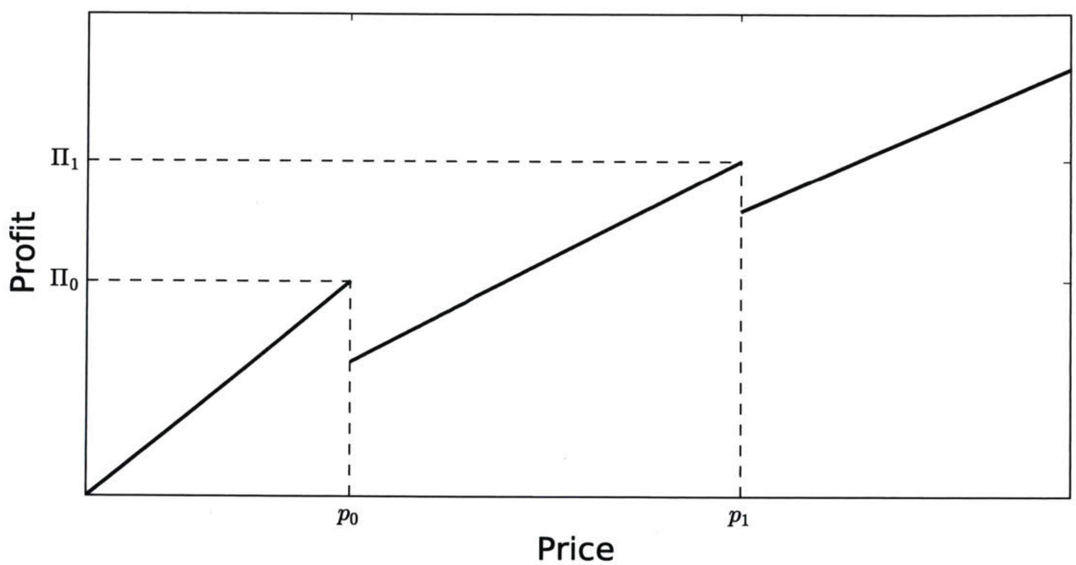
[21] This is saying that the consumer may not be willing to participate at p_k but if she did, she would search exactly k times.

- $k^* := \max\{k \mid V_k(\pi_0, p_k) \geq 0\}$, which is the maximum number of searches such that the consumer is willing to participate at p_k,

- $\hat{p}_{k+1}$, which is the price that satisfies $V_{k+1}(\pi_0, \hat{p}_{k+1}) = 0$.[22]

We see the firm's profit at each price in Figure 1-1 below. Each line segment ends at (p_k, Π_k) for $k = 0, 1, 2, \ldots$ but profit drops discretely at p_k because the consumer searches one more time. Note that each of the lines is less steep than the previous.

Figure 1-1: Profit as a function of price when $c = 0$, $s = 0.01$, and $\pi_0 = \mu = \frac{1}{2}$



Note how increasing the price makes the consumer search more as long as she is willing to participate, as we saw in (1.2). p_k is the best the firm can do, given that it wants the consumer to search exactly k times, while p_{k^*} is the highest such price so that the consumer is willing to participate. $\hat{p}_{k+1}$ is the highest possible price for $k+1$ searches that makes the consumer search $k+1$ times and get no utility.

The following Lemma shows that the firm's profit function Π_k is either always increasing or first decreasing and then increasing. In terms of Figure 1-1, this means that the peaks are either always increasing or first decreasing and then increasing.

Lemma 1.3.1. *Π_k is either always increasing or first decreasing and then increasing in k.*

Proof. Note first that $\Pi_{k+1} \geq \Pi_k \Leftrightarrow \frac{c}{s} \geq \frac{\mu - L_k^2}{(1-\mu)\mu}$. $L_k := \frac{\pi_k}{1-\pi_k}$ is strictly increasing in π_k which is strictly increasing in k, so that $\frac{\mu - L_k^2}{(1-\mu)\mu}$ is strictly decreasing in k. In fact, it is decreasing without

[22] We can write $\hat{p}_{k+1} := \frac{\pi_0 - \left[\pi_0(k+1) + (1-\pi_0)\frac{1-\mu^{k+1}}{1-\mu}\right]s}{\pi_0 + (1-\pi_0)\mu^{k+1}}$. We always require that $p_k < \hat{p}_{k+1} \leq p_{k+1}$, because otherwise the firm would not want the consumer to search exactly $k+1$ times, nor would the consumer be willing to. However, such $\hat{p}_{k+1}$ only exists for $k \geq k^*$ because the consumer gets a strictly positive utility for all prices below p_{k^*}.

bound. Therefore, if $\frac{c}{s} \geq \frac{\mu - L_{\tilde{k}}^2}{(1-\mu)\mu}$ for some $\tilde{k}$, it also holds for all $k > \tilde{k}$. However, there may be a finite number of periods at the beginning where the profit function is decreasing. □

The lemma means that Π_k is always increasing if and only if $\frac{c}{s} \geq \frac{\mu - L_0^2}{(1-\mu)\mu}$ (strictly so if the inequality is strict). It helps us prove the following proposition which characterizes the firm's optimal pricing policy as a function of the model parameters.

Proposition 1.2. *Assume that $\pi_0 > c$. If $s \geq (1-\mu)(1-\pi_0)\pi_0$, the firm's optimal price is π_0. If $s < (1-\mu)(1-\pi_0)\pi_0$, we have two cases:*

1. *If $\Pi_{k^*} \geq \Pi_0$, the firm's optimal price is either p_{k^*} or $\hat{p}_{k^*+1}$, and the firm prefers p_{k^*} if and only if $s \geq \bar{s}_k$, where:*

$$\bar{s}_k := \frac{\pi_0 + (1-\pi_0)(1-\mu)\mu^k c}{\pi_0(k+1) + \frac{(1+\pi_0)}{1-\mu} + (1-\pi_0)\mu^k + \frac{\pi_0^2}{(1-\mu)(1-\pi_0)\mu^k}}.$$

2. *If $\Pi_{k^*} < \Pi_0$, firm will set either p_0 or $\hat{p}_{k^*+1}$, depending on which one gives higher profits:*

   (a) $\Pi(\pi_0, \hat{p}_{k^*+1}) > \Pi_0 \Rightarrow \hat{p}_{k^*+1}$ is optimal

   (b) $\Pi(\pi_0, \hat{p}_{k^*+1}) \leq \Pi_0 \Rightarrow p_0$ is optimal.

Proof. See Appendix A.1. □

What Proposition 1.2 tells us is that the firm wants to induce either as much search as the consumer is willing to perform (when it is comparing p_{k^*} and $\hat{p}_{k^*+1}$) or no search at all (when setting p_0). If the firm makes higher profits at p_{k^*} than at p_0, it wants to induce $k \geq k^*$ searches, but if $\Pi_{k^*} < \Pi_0$, not inducing any search ($k = 0$) may be in order. We need the search cost to be lower than $(1-\mu)(1-\pi_0)\pi_0$ because otherwise, due to the consumer's search condition (1.2), she would be willing to purchase even for $p = \pi_0$, and she would never search even once, so the game would be trivial. If we make a parametric assumption, we can simplify the firm's pricing policy significantly:

Corollary 1.1. *Assume that $\mu \leq L_0^2$. The firm will always set either p_{k^*} or $\hat{p}_{k^*+1}$, where p_{k^*} will be firm-optimal if and only if $s \geq \bar{s}_{k^*}$.*

Proof. If the condition holds, $\frac{c}{s} \geq \frac{\mu - L_0^2}{(1-\mu)\mu}$ for all c and s, so the firm's profit Π_k is always increasing in k. Thus, the firm will want to set the highest p_k accepted by the consumer, p_{k^*}, or the price that brings the consumer's utility down to zero, $\hat{p}_{k^*+1}$. For $\Pi_{k^*} \geq \Pi(\pi_0, \hat{p}_{k^*+1})$, we need $s \geq \bar{s}_{k^*}$. □

1.3.3 Consumer Optimal Search Costs

In this subsection, I will characterize the consumer-optimal search cost and derive it analytically in some cases. Corollary 1.3 will show how the consumer's equilibrium utility depends on the search cost.

Before going into the main theorem of the section, let us first introduce one more Lemma which makes the proof of the theorem easier, and also helps us understand the problem better:

Lemma 1.3.2. *Let s_k solve $V_k(\pi_0, p_k) = 0$. Then $s_k > s_{k+1}$ $\forall k \in \mathbb{N}_0$, and $k^* = \max\{k \in \mathbb{N}_0 \mid s_k \geq s\}$.*

Proof. First of all, $V_k(\pi_0, p_k) = \pi_0 - \left[\frac{(\pi_0+(1-\pi_0)\mu^k)^2}{(1-\mu)\mu^k(1-\pi_0)} + k\pi_0 + (1-\pi_0)\frac{1-\mu^k}{1-\mu}\right] s$, which is strictly decreasing in k (since the ex-ante expected consumption value is always π_0 but the firm's price p_k is strictly increasing in k and the probability of success is strictly decreasing in k). Thus, for a given pair (π_0, μ), the consumer will never become worse off if the firm-optimal k decreases. Because $V_k(\pi_0, p_k)$ is also strictly decreasing in s, we can define a decreasing sequence of search cost thresholds, $s_0 > s_1 > s_2 > \ldots$, such that if search cost falls strictly between two thresholds, say s_k and s_{k+1}, then $V_k(\pi_0, p_k) > 0 > V_{k+1}(\pi_0, p_{k+1})$. This is due to the fact that $V_k(\pi_0, p_k) > V_{k+1}(\pi_0, p_{k+1})$, which means that we need a strictly higher search cost to bring $V_k(\pi_0, p_k)$ to zero than for $V_{k+1}(\pi_0, p_{k+1})$. If $s > s_0$, even $V_0(\pi_0, p_0) < 0$, in which case the firm will set $\hat{p}_0 \equiv \pi_0$, giving the consumer zero utility.

Now, remember that we defined k^* as the highest number of searches for which $V_k(\pi_0, p_k) \geq 0$. Therefore, we can also express k^* as the highest k such that $s_k \geq s$, which means that the consumer is (weakly) willing to search k times at p_k but she is not willing to search at all at p_{k+1}. □

Now we are ready to state the main theorem. It characterizes the consumer-optimal search cost, and provides an analytic solution in some cases.

Theorem 1.1. *Assume that $\pi_0 > c$. Let $L(s) = \frac{c}{s}$ and $R(\pi_0, \mu) = \frac{\mu - L_0^2}{\mu(1-\mu)}$. Define $\hat{s}$ as follows:*

$$\hat{s} := \max\{\bar{s}_0, s_1\} = \begin{cases} \bar{s}_0 \equiv \frac{(1-\mu)(1-\pi_0)}{(1-\mu)(1-\pi_0)+1}\left(\pi_0 + (1-\pi_0)(1-\mu)c\right), & \textit{if} \quad \pi_0 \geq \frac{\sqrt{\mu}-\mu}{1-\mu} \quad \textit{or} \quad c \textit{ is high}, \\ s_1 \equiv \frac{(1-\mu)(1-\pi_0)\mu\pi_0}{(1-\mu)(1-\pi_0)\mu + (\pi_0+(1-\pi_0)\mu)^2}, & \textit{if} \quad \pi_0 < \frac{\sqrt{\mu}-\mu}{1-\mu} \quad \textit{and} \quad c \textit{ is low}, \end{cases}$$

where $\bar{s}_0$ makes the firm indifferent between p_0 and $\hat{p}_1$, and s_1 is the lowest search cost so that the consumer will not participate at p_1. If $L(\hat{s}) \geq R(\pi_0, \mu)$, consumer-optimal search cost is $s^* = \hat{s}$. If $L(\hat{s}) < R(\pi_0, \mu)$, optimal search cost is $s^* \leq \hat{s}$.

Proof. See Appendix A.1. □

Note that $c \leq p_0(\hat{s}) \leq \pi_0$, so that both the consumer and the firm are willing to participate at p_0. This is because $p_0(\hat{s}) \geq p_0(\bar{s}_0) > c$ and $\pi_0 \geq \max\{p_0(\bar{s}_0), p_0(s_1)\}$.

The condition $L(s) \geq R(\pi_0, \mu)$ corresponds to the case where the firm's profit function Π_k is always increasing. If this holds, the consumer has to have a big enough search cost so that the firm is not willing to induce search. If, however, $L(\hat{s}) < R(\pi_0, \mu)$, the firm's profit function at the candidate for optimal search cost ($\hat{s}$) is decreasing at first. This means that the consumer's search cost can be lower than $\bar{s}_0$ and the firm will still find p_0 optimal.[23]

[23] This is due to the fact that the firm will under no circumstance consider setting $\hat{p}_1$ because that gives lower profits than p_1 which is worse than p_0 due to V-shaped profits.

Figure 1-2: Equilibrium utility and profit at consumer-optimal search cost ($\hat{s} = 0.1$) when $c = 0$, $\pi_0 = \mu = 0.5$

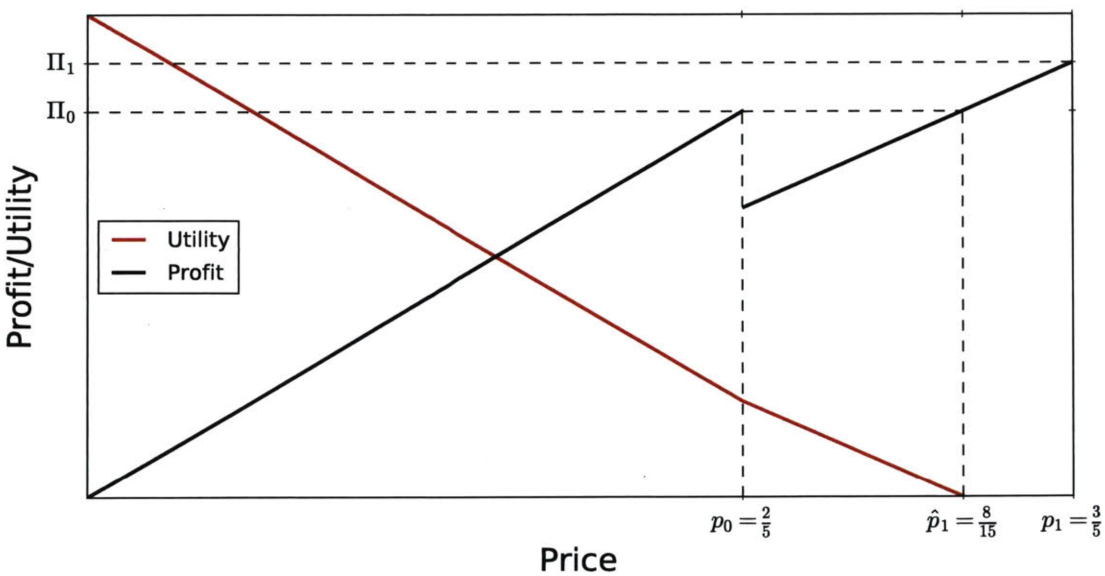


Figure 1-2 depicts utility and profit as functions of the price when search cost is at its consumer-optimal level ($\hat{s} = \bar{s}_0 = \frac{1}{10}$ when $\pi_0 = \mu = \frac{1}{2}$ and $c = 0$). One can see that the optimal search cost equalizes the firm's profit from $\hat{p}_1$ to Π_0 and that $V(\pi_0, \hat{p}_1) = 0$. If the search cost was any lower, utility would shift up, meaning that $\hat{p}_1$ would be higher and the firm's profit at $\hat{p}_1$ would exceed that at p_0 (especially because a lower s implies as lower p_0). On the other hand, if search cost was higher, the firm would still price at p_0 but this price would be higher than for a low s, meaning that the consumer would be hurt.

When $L(\hat{s}) \geq R(\pi_0, \mu)$, the Theorem provides an analytic solution to the problem of finding the optimal search cost for the consumer, which can be seen in the following Corollary.

Corollary 1.2. *Assume that* $\pi_0 \geq \frac{\sqrt{\mu}-\mu}{1-\mu}$. *Then the consumer-optimal search cost is* $\bar{s}_0$.

Proof. Obvious since the assumption guarantees that $R(\pi_0, \mu) \leq 0$, which in particular implies that $L(s) \geq R(\pi_0, \mu)$ for all s and c. Thus, due to the above Theorem, $\bar{s}_0$ is optimal. □

We can state another Corollary which makes it easy to compute the consumer's equilibrium utility when there is no production cost.

Corollary 1.3. *Assume that* $\mu < L_0^2$ *and* $c = 0$. *The consumer's equilibrium utility can be written as:*

$$U^* = \begin{cases} V_k(\pi_0, p_k), & \textit{if} \quad s \in [\bar{s}_k, s_k] \\ 0, & \textit{otherwise} \end{cases}$$

where $k \in \mathbb{N}_0$ is the number of searches in equilibrium. The consumer's utility jumps up at $s = \bar{s}_k$ and is strictly decreasing in s whenever $s \in [\bar{s}_k, s_k)$ for some $k \in \mathbb{N}_0$. Otherwise it is constant at zero.

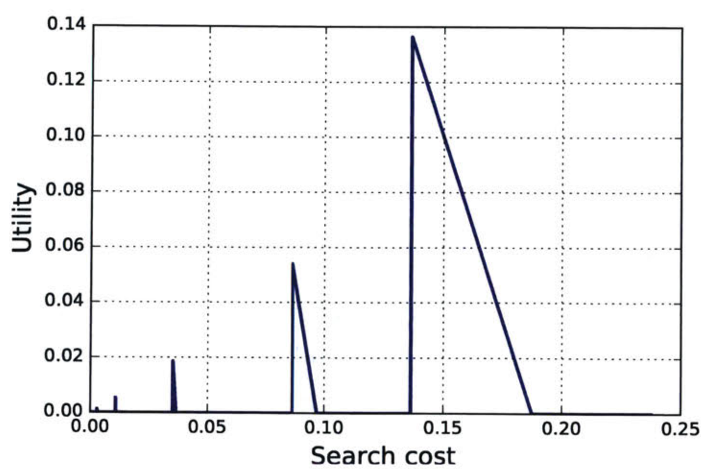
Proof. The assumption that $\mu < L_0^2$ implies that the firm's profit function Π_k is always increasing in k, so that the firm will always set either p_{k^*} or $\hat{p}_{k^*+1}$. The firm prefers p_{k^*} if and only if $s \geq \bar{s}_{k^*}$, where now:

$$\bar{s}_k = \frac{\pi_0}{\frac{\pi_0}{(1-\mu)(1-\pi_k)\pi_k} + \pi_0(k+1) + (1-\pi_0)\frac{1-\mu^{k+1}}{1-\mu}}.$$

With the two assumptions we made, for all k, $s_{k+1} < \bar{s}_k < s_k$, where s_k is the search cost for which $V_k(\pi_0, p_k) = 0$ (and s_k is decreasing in k). This means that if $s \in [\bar{s}_k, s_k]$, the firm's optimal price will be p_k, and if $s \in (s_{k+1}, \bar{s}_k)$, the optimal price will be $\hat{p}_{k+1}$. If $s > s_0$, the optimal price will naturally be equal to the prior since the consumer will never search. The consumer obtains strictly positive utility if and only if $s \in [\bar{s}_k, s_k)$, and this utility is decreasing in s within each interval because the firm is setting $p_k = \frac{s}{(1-\mu)(1-\pi_k)}$ which is increasing in s (and nothing else happens when s increases within the interval). Utility jumps up at $\bar{s}_k$ because there the firm changes its price from the one that makes the consumer search $k+1$ times to the one that makes her search only k times, which surely improves the consumer's utility. □

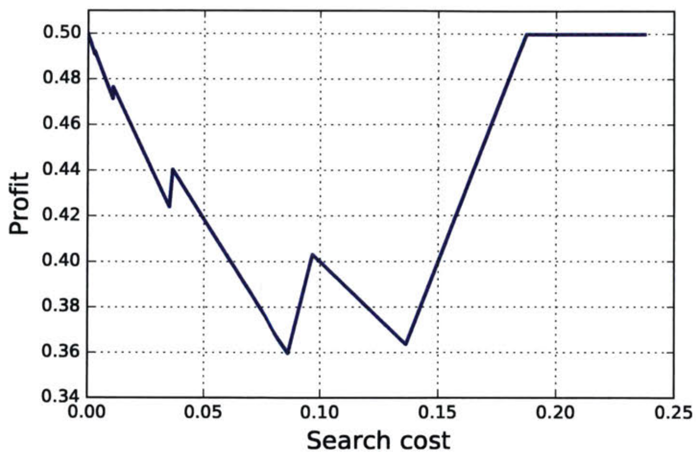
Note that when there is no production cost and $\mu < L_0^2$, maximum utility is $\bar{s}_0$. That is, utility equals the optimal search cost. Figures 1-3 and 1-4 illustrate what equilibrium payoffs look like, as a function of the search cost. We see that the the consumer's utility is either zero or strictly decreasing, other than at a few points where it jumps up discretely.

Figure 1-3: Equilibrium utility as a function of search cost when $c = 0$, $\pi_0 = \frac{1}{2}$, and $\mu = \frac{1}{4}$



Utility jumps up at these points because the firm changes its pricing policy to induce less search from the consumer, which can only be done by setting a lower price. On the other hand, given

Figure 1-4: Equilibrium profit as a function of search cost when $c = 0$, $\pi_0 = \frac{1}{2}$, and $\mu = \frac{1}{4}$



that the firm does not change its pricing policy, the consumer's utility is strictly decreasing in the search cost, while the firm's profit is strictly increasing as can be seen in Figure 1-4. Finally, when equilibrium utility is zero, the firm is getting all the social surplus, which means that its profits are decreasing in s.[24]

Profit is continuous and alternates between being increasing and decreasing in the search cost. The spikes for large k (small s_k) are so small that they do not show up on the scale. We can easily see that the consumer prefers a low search cost conditional on number of searches but a high search cost in that it brings the number of searches down. The optimal search cost is always $\bar{s}_0$ which leads to no search.

Example

Consider a simple example with $\pi_0 = \mu = \frac{1}{2}$, meaning that the prior is very uncertain and the signal is not very informative because even a bad product gives a good signal half of the time. Assume, as always, that $c < \pi_0 = \frac{1}{2}$. Now, $s_0 = \frac{1}{8}$, $s_1 = \frac{1}{11}$, and $\bar{s}_0 = \frac{1}{10} + \frac{c}{20}$. With these parameter values, $L(\hat{s}) \geq R(\pi_0, \mu) = -2$, so Theorem 1.1 tells us that the optimal search cost should be $\hat{s}$ (for any c). We also have that $\pi_0 \geq \frac{\sqrt{\mu}-\mu}{1-\mu}$ (since $\frac{1}{2} > \sqrt{2} - 1$). Thus, the consumer's optimal search cost is $\bar{s}_0 = \frac{1}{10} + \frac{c}{20}$, which is significant relative to the prior of $\frac{1}{2}$, even for zero production costs. The optimal search cost is increasing in the production cost because a higher c makes the firm more likely to induce search. However, it increases very slowly. When search cost is s^*, the firm will set $p^* = p_0 = \frac{s^*}{(1-\mu)(1-\pi_0)} = 4s^* = \frac{2}{5} + \frac{c}{5}$, which will give the consumer an expected utility of $V_0 = \frac{1}{2} - \frac{2}{5} - \frac{c}{5} = \frac{1}{10} - \frac{c}{5}$. The firm, on the other hand, will make a profit of $\Pi_0 = p_0 - c = \frac{2}{5} - \frac{4}{5}c$. Therefore, the product is always sold and the firm and the consumer share the production costs (the consumer pays 20%). The firm will not pass an increase in the costs directly to the consumer because she would then search. The assumption $c < \pi_0$ guarantees that $p^* > c$ for all c. Note that all of the above assumes that the consumer can always choose her search cost for any c so that she

[24]This is implied by the fact that when there are no production costs, it is never socially optimal to search, as the only reason to search is to avoid incurring the production cost in case the product is bad.

will be made to purchase at a low price. Because $\bar{s}_0 > s_1$, it is the firm's behavior that determines the optimal search cost, not the consumer's ($s \geq \bar{s}_0$ guarantees that the firm prefers to sell the good without search, whereas $s \geq s_1$ guarantees that the consumer is not willing to search at p_1).[25]

However, if $\pi_0 = \mu = \frac{1}{4}$ (so that the prior is worse but the signal is more informative), the theorem still tells us that $\hat{s}$ is optimal, and the important values of the search cost are $s_0 = \frac{9}{64}$, $s_1 = \frac{9}{85}$, and $\bar{s}_0 = \frac{9}{100} + \frac{81}{400}c$. We now have that $\pi_0 < \frac{\sqrt{u}-\mu}{1-\mu}$, which means that the consumer-optimal search cost is

$$s^* = \begin{cases} s_1 = \frac{9}{85}, & \text{if} \quad c \leq \frac{4}{51} \\ \bar{s}_0 = \frac{9}{100} + \frac{81}{400}c, & \text{if} \quad c > \frac{4}{51}. \end{cases}$$

The price is now:

$$p^* = \begin{cases} \frac{16}{85}, & \text{if} \quad c \leq \frac{4}{51} \\ \frac{4}{25} + \frac{9}{25}c, & \text{if} \quad c > \frac{4}{51} \end{cases}$$

Note how the firm bears the whole production cost for low values of c but shares the burden with the consumer for higher values. This is due to the equilibrium nature of the example: we are assuming that the consumer can choose her search cost to be exactly equal to s^*. In this case, for low values of c, everything stays constant as c increases because it is not the firm's behavior that matters but the consumer's (the optimal search cost is high because that makes sure the consumer is not willing to search at p_1). However, for higher c, it is again the firm's behavior that determines the optimal search cost, and, therefore, the value of the production cost matters for price and profits.

1.3.4 Discussion

The most important points of this section are that the consumer's search behavior is fully determined by the firm's price, and that the firm is essentially choosing between two pricing policies: a purchasing and a searching policy. Implementing the purchasing policy means setting the highest price that leads to purchase with certainty, while the searching policy involves a high price which incentivizes the consumer to search, extracting (almost) all consumer surplus. Under a simple condition ($\frac{c}{s} \geq \frac{\mu - L_0^2}{1-\mu}$) the firm always wants to extract all (or almost all) consumer surplus. If the consumer had the ability to commit to not searching, both the firm and the consumer would benefit in some cases, which means that the problem here is one of commitment. If the search cost is not too high, the consumer will search for $p = \pi_0 - \epsilon$ (where $\epsilon > 0$ is small) which makes the firm set a higher price (since the consumer is searching anyway). Being able to write a contract that specifies trade at $p = \pi_0 - \epsilon$ might increase both players' payoff.

More interestingly, as shown in Appendix A.5, the fact that the consumer can get a strictly positive surplus in some cases even when the firm's profit is strictly increasing in the price is due to

[25]If the consumer was willing to search at p_1, the firm would make her do that because the firm's profit is increasing in k, the number of searches.

the discreteness of the model. If the amount of search the consumer performs was continuous, as in Appendix A.5, the firm would always extract the full consumer surplus. Even though the model of the Appendix is slightly different (a bad news model), it still has the feature that the firm always extracts the full consumer surplus if prior belief is at least one half ($\pi_0 \geq \frac{1}{2}$). Therefore, the only way the consumer can get something strictly positive is if the common prior is lower than $\frac{1}{2}$ (the exact threshold depends on the production cost). Otherwise the firm will want the consumer to search because there is a probability of more than a half that the consumer will get a good signal and be willing to pay more (as long as the signal is not useless).

1.3.5 Welfare Considerations

So far we have only studied why a consumer almost always prefers a strictly positive search cost. However, the following proposition establishes that social surplus (consumer surplus plus firm profit) is always maximized at $s = 0$, but what the firm does is often not socially optimal.

Proposition 1.3. *Social surplus (consumer surplus + profit) is maximized at $s = 0$. For $s > 0$, there may be too much search in equilibrium relative to what would be socially optimal.*

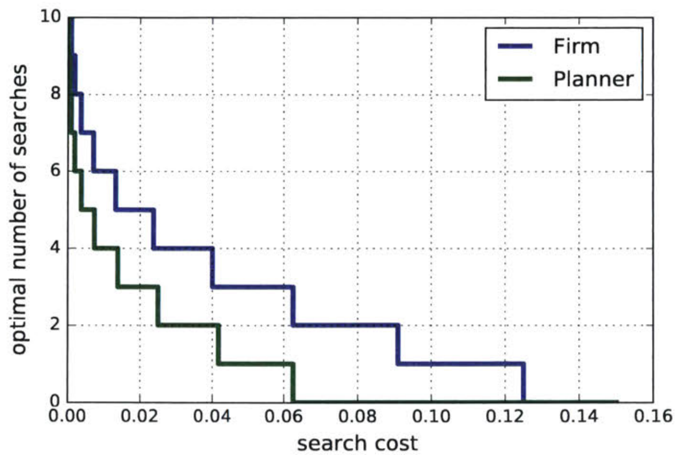
Proof. For the first part ($s = 0$), note that when there is no search cost, it is optimal to search forever (or as long as one possibly can), which leads to $k^* = \infty$ and $p^* = 1$ (or slightly less if there are only a finite number of signals). Therefore, the probability of success is π_0, and we will perfectly learn whether the good is of high or low quality. Therefore, social welfare is $W_\infty = V_\infty + \Pi_\infty = \pi_0(1 - c)$, which is the theoretical maximum of what we can get in ex-ante terms.

For the second part ($s > 0$), a benevolent planner, who controls everything, would search only when $W_{k+1} - W_k \geq 0$, which is the social gain from one more search. This can be written as $c \geq \frac{s}{(1-\mu)(1-\pi_k)}$. If the search cost is high relative to the production cost, searching will stop early (or not commence at all). This also means that welfare is either always decreasing in the number of searches k, or inverse V-shaped: first increasing, then decreasing (see Lemma A.1.1 in Appendix A.1).

On the other hand, the private gain of the firm can be written as $\Pi_{k+1} - \Pi_k = (\pi_k + (1 - \pi_k)\mu)(p_{k+1} - c) - (p_k - c) = (1 - \mu)(1 - \pi_k)c + \frac{(\pi_k + (1-\pi_k)\mu)^2 - \mu}{(1-\mu)(1-\pi_k)\mu} s$. Now, the private gain is bigger than the social gain if and only if $s > 0$ and $(1 - \mu)\pi_k > 0$, both of which hold. Therefore, the firm may induce too much search from a planner's point of view. □

To see this concretely, consider the following example. Take $\mu = \pi_0 = \frac{1}{2}$, $c = \frac{1}{4}$ and $s = \frac{1}{100}$. We can compute that the social gain is positive if and only if $k < \frac{\ln(2/23)}{\ln(1/2)} \approx 3.5$. Thus, the planner would search at $k = 3$ but not at $k = 4$. However, the firm's profit Π_k would be everywhere increasing, so it would find k^* such that $V_{k^*+1} < 0 \leq V_{k^*}$. In this case, $k^* = 5$. Using Proposition 1.2, we can now check that the firm prefers $\hat{p}_{k^*+1}$ over p_{k^*}, meaning that it will induce search at $k = 5$ but not at $k = 6$. The firm is not internalizing the loss it imposes on the consumer. We can use the same parameter values but let search cost vary to obtain Figure 1-5 below. In the figure, we see that the firm often induces strictly more search than would be socially optimal. The two lines overlap only for zero or high search costs. Note that we have plotted starting from a positive s because otherwise the number of searches would be infinite.

Figure 1-5: Optimal number of searches for the firm and the planner ($\mu = \pi_0 = \frac{1}{2}$, $c = \frac{1}{4}$)



1.3.6 Comment on Promotional Reviews

It is in order to comment on the assumption of reviews as trustworthy signals. What is the effect of promotional/fake reviews? Because we are assuming that bad products are thought to be good with probability μ (the receiver of the reviews does not fully understand what the reviewer is saying or what the reviewer's type is), it is reasonable to think that positive promotional reviews will, depending on the consumer, have two effects: 1) for some consumers, μ will increase because good reviews are now more probable and reviewers understand this in equilibrium (slower learning) but cannot observe which reviews are trustworthy, and 2) for other consumers, μ will stay the same because they can spot the fakes but s will increase because it now costs more to get useful reviews as the proverbial haystack is taller.

This means that if the firm is already setting $p_0 = \frac{s}{(1-\mu)(1-\pi_0)}$, increasing μ will help it increase the price (since increasing μ also leads to lower optimal s^*) until $p_0 = \pi_0$. If, however, $s_1 < s < \bar{s}_0$, then increasing μ may lead to the firm having to switch from $\hat{p}_1$ to p_0, which will surely improve consumer surplus. On the other hand, if the consumers are sophisticated enough so that they can spot the fakes, there is no effect on μ but the increased s will work the same way: if $s > \bar{s}_0$ to start with, the consumers will be worse off but if $s_1 < s < \bar{s}_0$, the consumers cannot be worse off but may benefit.

1.4 Heterogeneous Consumers

Previously, we assumed that there was a single consumer (or a continuum of identical consumers), which meant that the firm was able to extract (nearly) all consumer surplus if it found it optimal to make the consumer search. It also meant that there was no deadweight loss because zero-value consumers dropped out.

In this section, I will extend the previous model by allowing each consumer's valuation to be either v (match) or 0 (no match), with v uniform on $[0,1]$ in the population.[26] This means that the firm will never find it optimal to extract all the consumer surplus because this would exclude all of the consumers and give zero profit. Moreover, there may be a real deadweight loss due to the exclusion of low-value consumers whose value is higher than the production cost. Adding heterogeneity to the previous model does not change the result that consumer-optimal search cost is strictly positive, showing that the results are not due to full rent extraction. However, now a strictly positive search cost may be even socially optimal. This is due to an interaction between the search cost and deadweight loss, with that the firm is excluding too many consumers when search cost is low.

We will still maintain the assumption that consumption utility is binary: it is always either 0 or v, independently across consumers. This corresponds to saying that a consumer of type v has consumption utility θv, where $\theta \in \{0,1\}$, and $\mathbb{P}(\theta = 1) = \pi_0$. This prior is again shared among the consumers. In the rest of the section, we will see what happens in two cases: 1) no information, and 2) one available signal.[27]

1.4.1 No information

Assume first that no additional information is available at any cost. Now, consumer (of type) v will buy if and only if her utility from buying exceeds that of exiting. Buying gives $V_B(v) = \pi_0 v - p$, where v is the consumer's type and p the price. Exiting always gives nothing. Therefore, the consumer will buy if and only if $v \geq \frac{p}{\pi_0}$. The firm will then maximize $\Pi_B(p) = (1 - \frac{p}{\pi_0})(p - c)$, where the first term is the mass of consumers who prefer to buy. This yields an optimal price of $p_0^* = \frac{\pi_0 + c}{2}$. The maximized profit is $\Pi_0^* \equiv \Pi(p_0^*) = \frac{(\pi_0 - c)^2}{4\pi_0}$. Consumer v gets an expected utility of $V_0^*(v) = \pi_0 v - \frac{\pi_0 + c}{2}$ (given that $v \geq \frac{p_0^*}{\pi_0}$; otherwise nothing). As a whole, the consumers get an average utility of $V_0^* = \left(1 - \frac{p_0^*}{\pi_0}\right) \frac{V_0^*(1)}{2} = \frac{(\pi_0 - c)^2}{8\pi_0}$. This is because the lowest-value buyer gets 0, the highest type ($v = 1$) gets $V_0^*(1)$, and utility is increasing linearly in v between the two. The first term represents the mass of the buyers.

1.4.2 One available signal

To make things easier, I will assume that, at cost s, each consumer has access to one binary signal σ (the outcome of which is independent across consumers) with the conditional probabilities (as in

[26]The distributional assumption is for convenience; the results hold for more general distributions.

[27]Adding more signals would just needlessly complicate the analysis because we could always use arguments like those in Section 1.3 to show that, all else equal, a consumer prefers not to search.

Section 3):

$$\mathbb{P}(\sigma = G|\theta) = \begin{cases} 1, & \text{if} \quad \theta = H \\ \mu, & \text{if} \quad \theta = L. \end{cases}$$

Every consumer can choose to search once or not at all. Thus, each consumer has three decisions: Search, Buy, or Exit. The expected values of these decisions for consumer v can be expressed as: $V_S(v) = \pi_0 v - (\pi_0 + (1-\pi_0)\mu)\,p - s$, $V_B(v) = \pi_0 v - p$, and $V_E(v) = 0$, respectively.

Proposition 1.4. *Assume that $c < \pi_0$. Then there exist two thresholds, $\bar{s}$ and $\underline{v}^*(s)$, such that only consumers with $v \geq \underline{v}^*(s)$ participate and others exit. All consumers with $v \geq \underline{v}^*(s)$ make the same decision, buy or search, depending on the search cost as follows:*

1. *For $s < \bar{s}$, the firm will set a high price p_S^* to make the consumers with $v \geq \underline{v}^*(s)$ search, which maximizes the expected monopoly profits in this smaller population.*
2. *For $s \geq \bar{s}$, the consumers with $v \geq \underline{v}^*(s)$ will purchase the good at a price (p_B^*) that makes them indifferent between buying and searching.*

Note that the threshold value $\underline{v}^*(s)$ depends on the value of search cost. This is because this is the type of consumer who is exactly indifferent between exiting and participating in the market (be it buying or searching). It is increasing in s but experiences a discrete drop at $\bar{s}$ because the firm reduces its price discretely at that point (from p_S^* to p_B^*).

Proof. See Appendix A.2. □

The intuition for the proof is as follows. If one consumer prefers searching over buying, then all consumers do so, and vice versa if someone prefers buying to searching. This is because the expected posterior is equal to the prior and type v drops out when comparing buying to searching. Thus, the valuation v matters only for exiting. Therefore, there is a price, p_B^*, which is the highest price so that the consumers prefer buying over searching. If this price is below the no-signal monopoly price, it is the firm-optimal purchase price. If it is higher than the no-signal price, the firm prefers the no-signal price (since profit is decreasing in p in this range).

On the other hand, if the firm wants the consumers to search, it will essentially act like a monopolist maximizing a demand-weighted expected margin. This leads to an optimal search price p_S^*. Now the firm has two prices that can be optimal and it needs to choose one. Therefore, the firm will choose p_S^* if and only if $\Pi(p_S^*) > \Pi(p_B^*) \Leftrightarrow s < \bar{s}$. The mass of participating consumers, $1 - \underline{v}^*(s)$, is then defined through the optimal price and search cost.

The search cost threshold $\bar{s}$ is a rather complicated function of the parameters (π_0, μ, c), so we cannot get far analytically unless we make specific parametric assumptions.[28] This is exactly what

[28]We can express the threshold as $\bar{s} \equiv \frac{B-\sqrt{B^2-4AC}}{2A}$, where $A = 4+(1-\mu)^2(1-\pi_0)^2$, $B = 4(1-\mu)(1-\pi_0)(\pi_0+c)+2(1-\mu)^2(1-\pi_0)^2(\pi_0-(\pi_0+(1-\pi_0)\mu)c)$, and $C = (1-\mu)^2(1-\pi_0)^2(\pi_0-(\pi_0+(1-\pi_0)\mu)c)^2+4(1-\mu)^2(1-\pi_0)^2\pi_0 c$. See Appendix A.2.

we will do in the example below.

Example

Let $\pi_0 = \mu = \frac{1}{2}$. We have a flat prior and the signal technology is reasonably good, giving a posterior of $\pi_1 = \frac{2}{3}$ after a good signal. Plugging these values into the formula for $\bar{s}$ in the previous footnote gives us: $\Pi_S > \Pi_B \Leftrightarrow s < \bar{s} = \frac{1}{26} + \frac{9}{52}c$.

To compute the threshold $\underline{v}^*(s)$, we need to know the firm's equilibrium price as a function of the parameters. To calculate the firm's optimal pricing rule, we know that it will set p_S^* for search costs lower than $\bar{s}$, and p_B^* for values higher than $\bar{s}$. The purchase price is the minimum of $\frac{s}{(1-\mu)(1-\pi_0)}$ and the no-information monopoly price p_0^* because the former is the maximum price the consumer is willing to accept without search and the latter is the optimal price if there is no search in any case. The search price maximizes the expected search profit as seen in Appendix A.2. These give:

$$p^* = \begin{cases} \frac{1}{3} + \frac{1}{2}c - \frac{2}{3}s, & \text{if} \quad s < \frac{1}{26} + \frac{9}{52}c \\ 4s, & \text{if} \quad \frac{1}{26} + \frac{9}{52}c \leq s \leq \frac{1}{16} + \frac{1}{8}c, \end{cases}$$

because at $\bar{s}$ the firm switches from search pricing to purchase pricing. This allows us to write a single consumer's utility (knowing that she will search for low s and purchase for high s, as long as she is willing to participate) as:

$$V^*(v) = \begin{cases} \frac{1}{2}v - \frac{1}{4} - \frac{3}{8}c - \frac{1}{2}s, & \text{if} \quad s < \frac{1}{26} + \frac{9}{52}c \\ \frac{1}{2}v - 4s, & \text{if} \quad \frac{1}{26} + \frac{9}{52}c \leq s \leq \frac{1}{16} + \frac{1}{8}c. \end{cases}$$

However, this works only if $V^*(v)$ is non-negative, which finally gives us the threshold values for participation:

$$\underline{v}^*(s) \equiv \begin{cases} \frac{1}{2} + \frac{3}{4}c + s, & \text{if} \quad s < \frac{1}{26} + \frac{9}{52}c \\ 8s, & \text{if} \quad \frac{1}{26} + \frac{9}{52}c \leq s \leq \frac{1}{16} + \frac{1}{8}c. \end{cases}$$

Therefore, the consumers as a whole get V^* and the firm gets Π^*, where:

$$V^* = \begin{cases} \frac{1}{4}(\frac{1}{2} - \frac{3}{4}c - s)^2, & \text{if} \quad s < \frac{1}{26} + \frac{9}{52}c \\ \frac{1}{4}(1 - 8s)^2, & \text{if} \quad \frac{1}{26} + \frac{9}{52}c \leq s \leq \frac{1}{16} + \frac{1}{8}c \end{cases}$$

$$\Pi^* = \begin{cases} \frac{1}{2}(\frac{1}{2} - \frac{3}{4}c - s)^2, & \text{if} \quad s < \frac{1}{26} + \frac{9}{52}c \\ (1 - 8s)(4s - c), & \text{if} \quad \frac{1}{26} + \frac{9}{52}c \leq s \leq \frac{1}{16} + \frac{1}{8}c \end{cases}$$

As you may have noticed, I have omitted what happens when s is high (higher than $\frac{1}{16} + \frac{1}{8}c$, to be exact). However, in this case the equilibrium is exactly like the one with no available signals. The intuition is that the search costs are so high the firm is able to charge the no-search monopoly price p_0^*, and no one is willing to search.

The following proposition shows that the mere existence of a signal can reduce consumer and social surplus compared to the case with no information. This means that for some intermediate values of search cost, it would be socially optimal to prevent consumers from obtaining a signal.

Proposition 1.5. *Let $\pi_0 = \mu = \frac{1}{2}$ and $c < \pi_0$. The consumers are (strictly) hurt by their having access to the signal if and only if $s \in \left(\frac{1}{4}c, \frac{1}{26} + \frac{9}{52}c\right)$. In fact, the exact same result applies to social surplus. That is, it may be socially inefficient for the consumers to have access to a signal.*

Proof. The proof involves simply comparing $V^*(v)$ to $V_0^*(v)$ and W^* to W_0^*, and noting that if $V^*(v) > V_0^*(v)$, then $V^* > V_0^*$ (for the consumers as a whole). We can note that for $s < \frac{1}{4}c$, we always have $s < \frac{1}{26} + \frac{9}{52}c$ and $V^*(v) > V_0^*(v)$. Similarly, if $s \geq \frac{1}{26} + \frac{9}{52}c$, then $V^*(v) \geq V_0^*(v)$. The same reasoning applies to social surplus. □

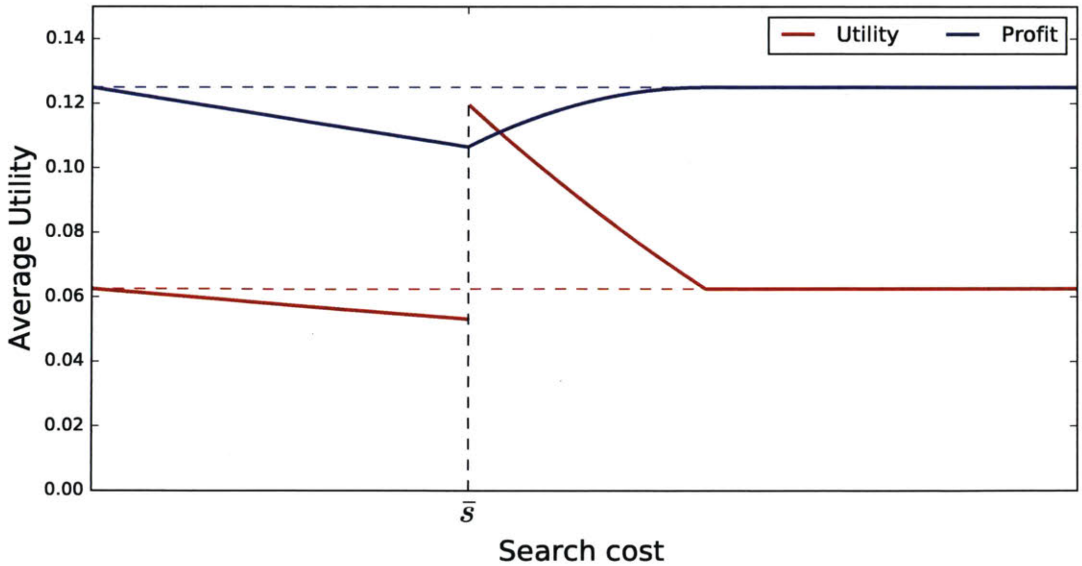
Corollary 1.4. *Assume that $\pi_0 = \mu = \frac{1}{2}$ and $c = 0$. The consumers' having access to a signal strictly hurts them and the society as a whole if and only if $s \in (0, \frac{1}{26})$.*

To understand the proposition (and the corollary), consider Figures 1-6 and 1-7 below. In these figures, we are plotting the firm's profit and average utility across all consumers in equilibrium. We are fixing the prior and the imprecision of the signal at $\frac{1}{2}$. Figure 1-6 fixes the production cost at zero and Figure 1-7 at 0.1. The dotted lines represent what profit and utility would be if there was no possibility of obtaining a signal (as in the no-information case of the previous subsection). Note that profit is a continuous function but utility jumps when the firm switches from the higher search price to the lower purchase price. Thus, we see that there are values of s where both profit and utility are lower with than without a signal. The society as a whole loses under these intermediate search costs.

Increasing c from 0 to 0.1 means that both the consumers and the firm prefer having a signal to no signal if search cost is low (less than $\frac{1}{4}c$, to be exact). This is why the dotted lines in Figure 1-7 are below the solid ones for low s.

The intuition behind the previous two figures is that, for low search costs, the consumers benefit from the signal because searching and utilizing the signal is cheap but the firm cannot set a high price because it does not want to exclude too many consumers (which is in stark contrast to what happens with identical consumers in Section 1.3). For high search costs, it is too expensive for the firm to make the consumers search but the consumers may still have some bargaining power (depending on how high s is) so that they may gain relative to the no-signal world (for very high search costs everything is the same in both worlds). However, for intermediate search costs, as defined in the proposition, the firm has the incentive to make the consumers search but they are hurt because they have to use the technology, which is not cheap. In this case, both the consumers and the firm are hurt because consumers cannot commit to buying at an intermediate price. In short, in some cases having access to a signal may hurt the society because there will be too much

Figure 1-6: Average utility and profit in equilibrium when $c = 0$ and $\pi_0 = \mu = \frac{1}{2}$.



search in equilibrium (which in turn is due to a lack of commitment power). Being able to write ex-ante contracts would solve this problem.

To understand why the lower bound in the proposition is increasing in c, we just need to remember that a high production cost leads to a high price and in those cases it is socially optimal to be able to utilize a signal technology (in fact, it is even consumer optimal to do so).

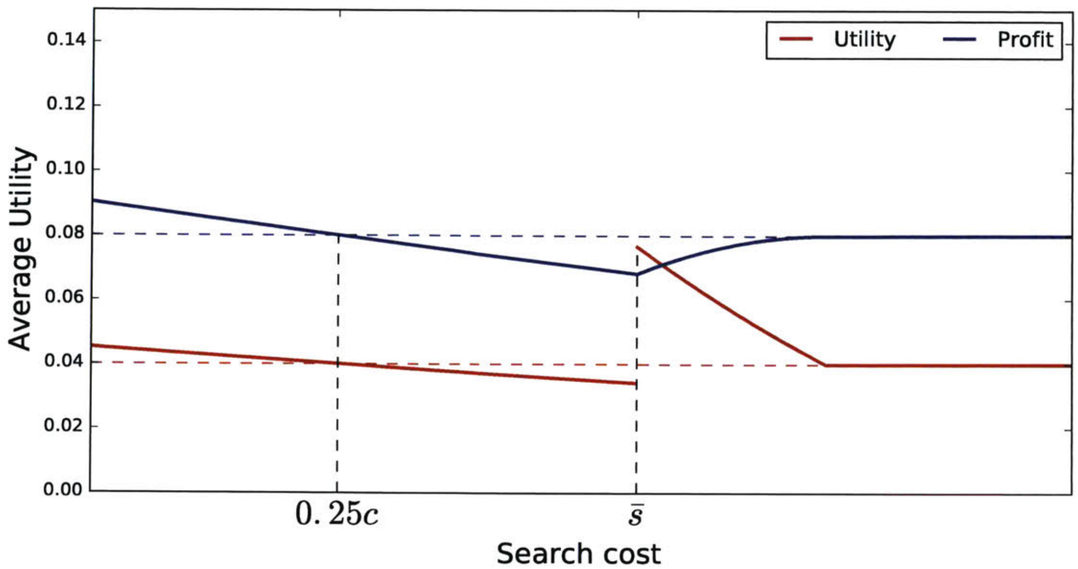
The following lemma establishes that a low search cost may not imply high participation in equilibrium.

Lemma 1.4.1. *Continue to assume that $\pi_0 = \mu = \frac{1}{2}$. If $c < \frac{10}{33}$, the firm's policy is more exclusive at $s = 0$ than at $s = \bar{s} := \frac{1}{26} + \frac{9}{52}c$. That is, the mass of participating consumers is bigger at $\bar{s}$.*

Proof. We can simply check what condition has to be satisfied for $\underline{v}^*(0) > \underline{v}^*(\bar{s})$. It turns out to be $c < \frac{10}{33}$. □

To get some intuition for the lemma, consider what happens when there is no production cost. If search cost is zero, the firm has to induce search because all of the consumers have perfect bargaining power in case the firm wants to set a low price, which would imply a zero price. Thus, all consumers who participate will search, and a fraction $\pi_0 + (1 - \pi_0)\mu = \frac{3}{4}$ of them will receive good signals. This means that the participating mass of consumers is $1 - \frac{p}{\pi_1} = 1 - \frac{3}{2}p$, where π_1 is the posterior expectation. Therefore the optimal monopoly price is $p^* = \frac{1}{3}$ and mass of participants $\frac{1}{2}$. However, when search cost is $\bar{s}$, the firm prefers selling the good without search because the consumers do not have that much bargaining power. This implies that the optimal price is the highest one that makes the consumers willing to buy, which is $p^* = 4\bar{s}$. Thus, the mass of participants is $1 - 8\bar{s} = \frac{9}{13} > \frac{1}{2}$. When the production cost increases, the mass of participants at

Figure 1-7: Average utility and profit in equilibrium when $c = 0.1$ and $\pi_0 = \mu = \frac{1}{2}$.



$s = 0$ stays constant but the mass at $s = \bar{s}$ decreases because $\bar{s}$ increases due to the firm's increased desire to induce search. The simplest way to think about the result is that every consumer's utility is always decreasing in the price (strictly if they participate), so that the mass of participating consumers is always strictly decreasing in the price. Because the firm charges a lower price at $s = \bar{s}$ than at $s = 0$ for low c, the mass of consumers is bigger under the former than the latter.

We can now state the main result of the section:

Proposition 1.6. *Let $\pi_0 = \mu = \frac{1}{2}$. Assume each consumer can obtain at most one signal at cost s. Then:*

1. *For $c \leq \frac{10}{33}$, each individual consumer's expected utility is maximized at $\bar{s} := \frac{1}{26} + \frac{9}{52}c$ (given that they purchase). Moreover, total consumer surplus is also maximized at $\bar{s}$.*

2. *For $c \leq \bar{c} \equiv \frac{\sqrt{51}/13-1/2}{2\sqrt{51}/13-3/4} \approx 0.142$, total social surplus is maximized at $\bar{s}$.*

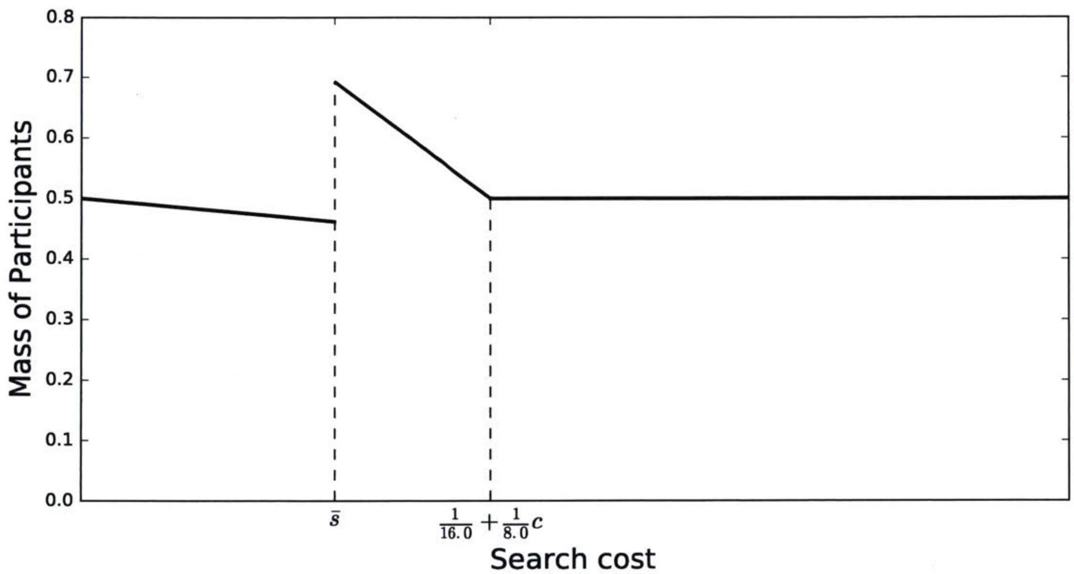
Note the contrast between this and Theorem 1.1: social surplus may not be maximized at $s = 0$. The intuition is that the firm is excluding too many consumers, whose valuation exceeds the production cost, in order to maximize its (monopoly) profits. This leads to a deadweight loss. Looking at Figures 1-6 and 1-7, we see how the firm changes its pricing policy at $s = \bar{s}$ so that the consumers get a discrete benefit.

Proof. To show that each individual consumer benefits from having $s = \bar{s} > 0$, note that their utility is piecewise linear in the search cost with one jump and one kink. When $s < \frac{1}{16} + \frac{1}{8}c$, this utility is decreasing in the search cost everywhere but at $\bar{s}$ where it jumps up (for low enough c).

Therefore, the consumer's maximum is either at $s = 0$ or at $s = \bar{s}$. When $c \leq \frac{10}{33}$, the utility is maximized at $\bar{s}$ – independent of the type, v. Because every consumer's utility is maximized at $\bar{s}$, one can expect the total consumer surplus to be maximized there as well. However, it might be that some consumer who was willing to participate at $s = 0$ would not do so at $s = \bar{s}$, but this will not happen with $c \leq \frac{10}{33}$ as shown in Lemma 1.4.1. This proves the first part of the Proposition.

For the second part, we can see that, just like consumer surplus, total surplus W^* is always decreasing in s but experiences a discrete jump at $\bar{s}$. Therefore, we can perform the same steps as above and compare total surplus at two points: $s = 0$ and $s = \bar{s}$. It turns out that total surplus is maximized at $\bar{s}$ whenever $c \leq \bar{c}$. □

Figure 1-8: Mass of participating consumers as a function of search cost when $c = 0$ and $\pi_0 = \mu = \frac{1}{2}$.



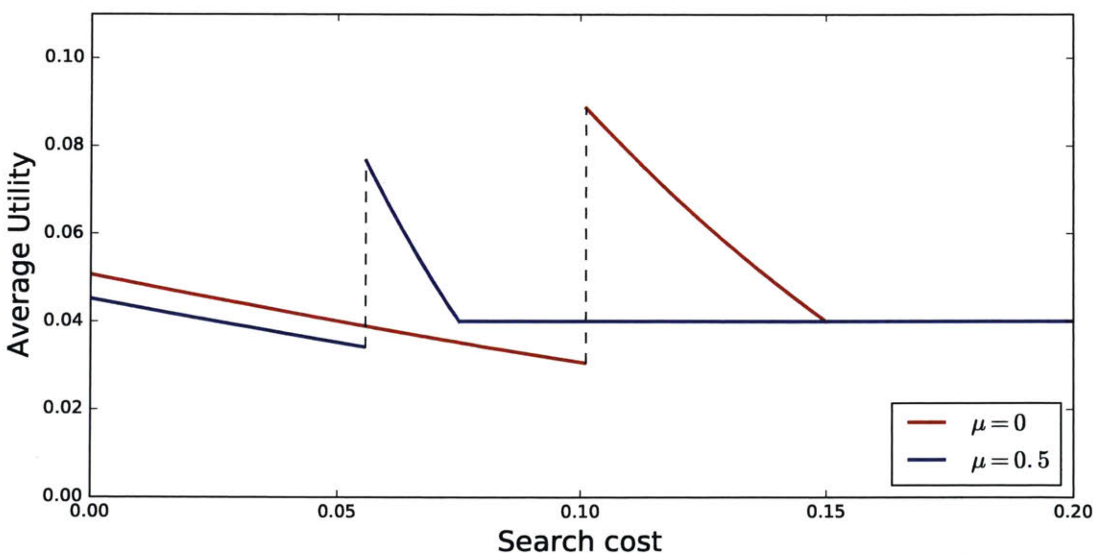
The intuition for the above Proposition is that, for low s, the firm is pricing too high, which hurts the consumers in two ways: it forces them to search (extracting more consumer surplus), and it increases $\underline{v}$ (the marginal consumer's valuation) – the firm excludes too many consumers even though all of them have a low search cost. When $c \leq \bar{c}$, both consumer and social surplus are maximized at a strictly positive search cost because that enables the firm to set a slightly lower price. This can be seen in Figure 1-8 above, where the mass of participating consumers first decreases in s because fewer of them are willing to search but then jumps up at $\bar{s}$ when the firm finds it optimal to induce no search.[29] The firm is not alone at fault in setting a high price to induce search because a medium price would still make the consumers search. Only when the search cost is high, can the firm make the consumers buy because they cannot search anymore. The consumers'

[29] We also see that social surplus can be lower or higher than in a model without any signal, depending on the search cost.

ability to search for information hurts them because they cannot commit to not searching if the price is intermediate. In fact, the firm is pricing too high even for high values of search cost, s, which leads to sub-optimal social surplus (from the point of view of a social planner).

To see how the precision of the signal (μ) matters for the equilibrium, consider Figure 1-9 below where we plot the average equilibrium utility for two values of μ. Note how the two curves start off almost identical.[30] However, the firm changes its pricing policy sooner (lower s) when the signal is imprecise (high μ) because the consumer's beliefs do not increase much after search. On the other hand, the consumer with $\mu = 0$ has to be compensated more for her not to search (she has higher bargaining power in case the firm wants to make her buy). Note, however, that for intermediate s, a less precise signal is consumer-optimal. This is due to the fact that the low-μ consumer cannot commit to not searching which forces the firm to make her search.

Figure 1-9: Utility as a function of search cost and the precision of the signal when $c = 0.1$ and $\pi_0 = \frac{1}{2}$.



1.5 Optimal Search Under General Distributions of Values

In the models of Sections 1.3 and 1.4 we assumed that, with probability π_0, the consumer's valuation for a match was 1 and $v \in [0, 1]$, respectively, and the consumer knew what her value for a match was. In this section we investigate how the results differ if we assume that the consumer only knows that her value is $v \sim F[0, 1]$, which is unknown to everyone, so that "bad matches" correspond to values lower than the price and "good matches" to values higher than the price. I will also analyze optimal choices of information structures, which I will continue in Section 1.6. For simplicity, I will

[30]They start off exactly the same if there is no production cost and the difference grows in c.

assume that the consumer can search only once at cost s and that the firm has no production cost. Search perfectly reveals the consumer's value. The main takeaway is that the results of Section 1.3 are preserved, as the consumer prefers a strictly positive search cost, the value of which depends on the prior distribution of values.

Assuming a uniform distribution, we get to compare the results to the no-search-cost results of Roesler and Szentes (2017). They ask what a consumer should optimally do if she was allowed to choose a signal structure leading to a distribution of posteriors (for a given prior), and if the firm knew the structure but not the realization when setting its price.[31] The optimal signal would induce a unit-elastic distribution of posteriors so that the firm would be indifferent between a bunch of prices, all leading to purchase with certainty.[32] I find that their signal structure is optimal for no search costs but a perfectly informative signal technology is optimal for a high enough s. In between, the optimal distribution of posteriors may be more informative than Roesler-Szentes but less so than the perfect signal I analyze. Appendix A.4 takes a brief look at what the optimal posterior distribution should be, without solving for it.

The motivation for studying optimal signal structures here and henceforth is that they make platforms more appealing to consumers. If there is intense competition between different platforms, they want to make their services as good as possible for the consumers – which means designing the signals optimally. This is a motive similar to the earlier analysis of the optimal search costs, and it turns out that the two are connected.

Now, the consumer's value of searching is $U_S(p) = \int_p^1 (v-p)dF(v) - s = \int_p^1 vdF(v) - [1 - F(p)]p - s$, while she gets $U_B(p) = \int_0^1 (v-p)dF(v) = \int_0^1 vdF(v) - p$ from buying. This implies that the consumer will prefer buying over searching if and only if:

$$\begin{aligned} & U_S(p) \leq U_B(p) \\ \Leftrightarrow\ & pF(p) - s \leq \int_0^p vdF(v) = [vF(v)]_0^p - \int_0^p F(v)dv = pF(p) - \int_0^p F(v)dv \\ \Leftrightarrow\ & \int_0^p F(v)dv \leq s. \end{aligned} \tag{1.3}$$

We can now define the highest price, $p_B(F,s)$, which makes the consumer prefer purchasing to searching, as the solution to $\int_0^{p_B} F(v)dv = s$. Implicitly differentiating this equation with respect to the search cost, we get $\frac{\partial p_B}{\partial s} = \frac{1}{F(p_B)} > 0$ for all $p_B > 0$. Thus, the higher the search cost is, the higher is the price the firm can charge while still preventing the consumer from searching (as long as $p_B(F,s) \leq \mathbb{E}_F[v]$). The intuition for the purchase condition (1.3) is that the consumer prefers buying to searching whenever the cost of search is higher than its benefit. The benefit is the higher, the more often the consumer's true value is below the price and the lower the true value is relative to the price because in these cases the consumer does not need to pay the price if she searches first (saving $p - v$ utils). The following proposition states that the highest price the consumer is willing

[31]Each of their signals refers to a posterior expectation, so that the signals are a reduced form representation of an experiment. This is why the average of the signals has to equal the prior expectation.

[32]And the firm would then be willing to choose the lowest price.

to accept without search is decreasing in the distribution when the distributions are ordered in terms of second-order stochastic dominance.

Definition 1.1. *For any two distributions, F and G, with supports on the unit interval $[0,1]$, we say that $G \overset{SO}{\succeq} F$ if and only if $\int_0^x F(v)dv \geq \int_0^x G(v)dv$ for all $x \in [0,1]$.*

Proposition 1.7. *Take any two distributions F and G. Then $p_B(G,s) \geq p_B(F,s) \forall s \Leftrightarrow G \overset{SO}{\succeq} F$.*

Proof. If $G \overset{SO}{\succeq} F$, we have that $\int_0^{p_B(F,s)} F(v)dv = s = \int_0^{p_B(G,s)} G(v)dv \leq \int_0^{p_B(G,s)} F(v)dv$, where the equalities are by definition of p_B and the inequality is due to second-order stochastic dominance. Therefore, due to $\int_0^x F(v)dv$ being increasing in x, we know that $p_B(G,s) \geq p_B(F,s)$. It is equally simple to show the converse. □

What the proposition says is that, conditional on the firm setting p_B to prevent search, the consumer wants to have the "worst" distribution (given an expected value) when it comes to second-order stochastic dominance.

We can next derive the maximum price, $p_{max}(F,s)$, at which the consumer is still willing to search:

$$U_S(p_{max}) = 0 \Leftrightarrow p_{max} + \int_{p_{max}}^1 F(v)dv = 1 - s. \tag{1.4}$$

This p_{max} exists whenever $0 \leq s < \mathbb{E}_F[v]$ because then $U_S(0) = \mathbb{E}_F[v] - s > 0$, $dU_S/dp = F(p) - 1 < 0$ for all $p < 1$, and $U_S(1) \leq 0$. Again, implicitly differentiating with respect to s gives: $\frac{\partial p_{max}}{\partial s} < 0$ because the left-hand side of this equation is strictly increasing in p, so that if s increases, p has to decrease.

Even though the consumer will accept p_{max} ex ante (but not necessarily after she is done with searching), it may not be the optimal price for the firm. We can write expected search profit as $\Pi_S(p) = [1 - F(p)]p$, which leads to the first-order condition:

$$p^* = \frac{1 - F(p^*)}{f(p^*)}.$$

The FOC characterizes a unique unconditional optimum under the usual condition of increasing hazard rate:

Assumption 1.1. *The distribution of valuations for a good match, F, satisfies the increasing hazard rate condition: $\frac{f(p)}{1-F(p)}$ is increasing in p for all p.*

This assumption allows us to write the firm-optimal search price as $p_S = \min\{p^*, p_{max}\}$, where p^* is the firm-optimal price without any conditions on the consumer surplus, while p_{max} is the highest price the consumer will accept.

Proposition 1.8. *If it exists, the firm's optimal search price, p_S, can be written as:*

$$p_S = \begin{cases} p^*, & \text{if} \quad s \leq \bar{s}(F) := \int_{p^*}^1 v dF(v) - [1 - F(p^*)]p^* \\ p_{max}, & \text{if} \quad s > \bar{s}(F). \end{cases}$$

Proof. It is obvious that the firm prefers p^* as long as the consumer is willing to accept it, which is the case when $s \leq \bar{s}(F)$. If $s > \bar{s}(F)$, the firm sets the highest price the consumer will accept because that is below the price that maximizes monopoly profits. Note that $p_{max}(F, 0) = 1$, so the only way $p^* = p_{max}(F, 0)$ is if $p^* = 1$. Otherwise, $\bar{s}(F) > 0$. □

However, next we will see that, for high-enough search costs, the consumer will never search.

Proposition 1.9. *Let s_μ be the search cost that solves $s_\mu = \int_0^\mu F(v)dv$, where μ is the expected value under F.[33] Then we have $p_B(F, s_\mu) = \mu = p_{max}(F, s_\mu)$. The consumer buys the good at price μ for all $s \geq s_\mu$.*

Proof. Now, by definition, $p_B(F, s_\mu) = \mu$. Then we also have that $p_{max} + \int_{p_{max}}^1 F(v)dv = 1 - s_\mu \Leftrightarrow p_{max} + \int_0^\mu F(v)dv + \int_{p_{max}}^1 F(v)dv = 1$. We can check that $p_{max} = \mu$ solves this equation because $\mu + \int_0^1 F(v)dv = \mu + 1 - \mu = 1$. It is also the only solution because there is a unique p_{max} for every s. Therefore, $p_{max} = \mu = p_B$.

Now, if $s \geq s_\mu$, the consumer will purchase the product at a price equal to the prior because she will prefer buying to searching at $p = \mu$, and this is also the highest price the firm can charge. □

Now, if we assume that $p^* \leq p_{max}$ (which always holds for low s due to the previous propositions), then the firm will set the purchase price, p_B, if and only if $\Pi_B \geq \Pi_S \Leftrightarrow p_B \geq [1 - F(p^*)]p^*$. This holds as an equality for some search cost, s^*, because the RHS is strictly greater for $s = 0$ but $\partial p_B / \partial s > 0 > \partial \Pi_S / \partial s$.

This allows us to write the consumer's equilibrium utility:

$$U^* = \begin{cases} \max\{0, 1 - p^* - \int_{p^*}^1 F(v)dv - s\}, & \text{if} \quad s < s^* \\ \mathbb{E}_F[v] - p_B, & \text{if} \quad s^* \leq s < s_\mu \\ 0, & \text{if} \quad s_\mu \leq s, \end{cases}$$

where s_μ satisfies: $p_B(F, s_\mu) = \mu := \mathbb{E}_F[v]$, so that the consumer buys the good at her prior expectation. When $s < s^*$, the firm either extracts all consumer surplus (if $p_{max} \leq p^*$) or sets $p^* < p_{max}$ which leaves the consumer with some utility. Remember that s^* is the search cost (that depends on F) which makes the firm indifferent between p_S and p_B.

If we take a closer look at the equilibrium utility, we see that it always jumps up at s^* as long as $p^* > p_B$. This is because $\partial U_S / \partial p < 0$, so that $U_S(p^*) < U_S(p_B) = 1 - p_B - \int_{p_B}^1 F(v)dv - s = 1 - p_B - \int_0^1 F(v)dv = \mathbb{E}_F[v] - p_B = U_B(p_B)$, where the second equality is due to the definition of p_B. It is also obvious that social surplus is maximized at s^* because then the good is always sold

[33] $\mu = \int_0^1 v dF(v)$.

(which is efficient because there is no production cost), while if $s < s^*$, there is a non-zero chance that the good will remain unsold, and the consumer will incur a search cost. This leads us to a theorem about the consumer-optimal distribution F:

Theorem 1.2. *Consider distributions with a fixed mean μ. For a given search cost, s, the consumer's optimal prior distribution of values, F^*, satisfies $\Pi_S = p_B(F^*, s)$.*

Proof. Since we know that p_B is the lower, the "worse" F is in the sense of SOSD, and because the consumer's utility is maximized at the point where $\Pi_S = p_B(F^*, s)$ (for a fixed μ), she wants the most uninformative distribution that still makes the firm willing to set p_B instead of a search price. □

Example: Consider the case of uniformly distributed valuation, $F(v) = v$. Now, the optimal search price is always $p^* = \frac{1}{2}$ because $p_{max} = 1 - \sqrt{s} > \frac{1}{2} \Leftrightarrow s < \frac{1}{4}$, and for s higher than $\frac{1}{4}$ the firm prefers selling the good without search. We can derive $p_B = \sqrt{2s}$ by solving $s = \int_0^{p_B} v dv$. Now, the prior expectation is $\mathbb{E}_F[v] = \frac{1}{2}$, so that if $s \geq \frac{1}{8}$, the firm can set its price equal to the prior (because $p_B \geq \frac{1}{2}$ for those s). The firm prefers p_B over p^* if and only if $\sqrt{2s} \geq (1 - \frac{1}{2})\frac{1}{2} = \frac{1}{4} \Leftrightarrow s \geq \frac{1}{32}$. This allows us to write the firm's optimal pricing policy:

$$p_{opt} = \begin{cases} \frac{1}{2}, & \text{if} \quad s < \frac{1}{32} \\ \sqrt{2s}, & \text{if} \quad \frac{1}{32} \leq s \leq \frac{1}{8} \\ \frac{1}{2}, & \text{if} \quad \frac{1}{8} < s. \end{cases}$$

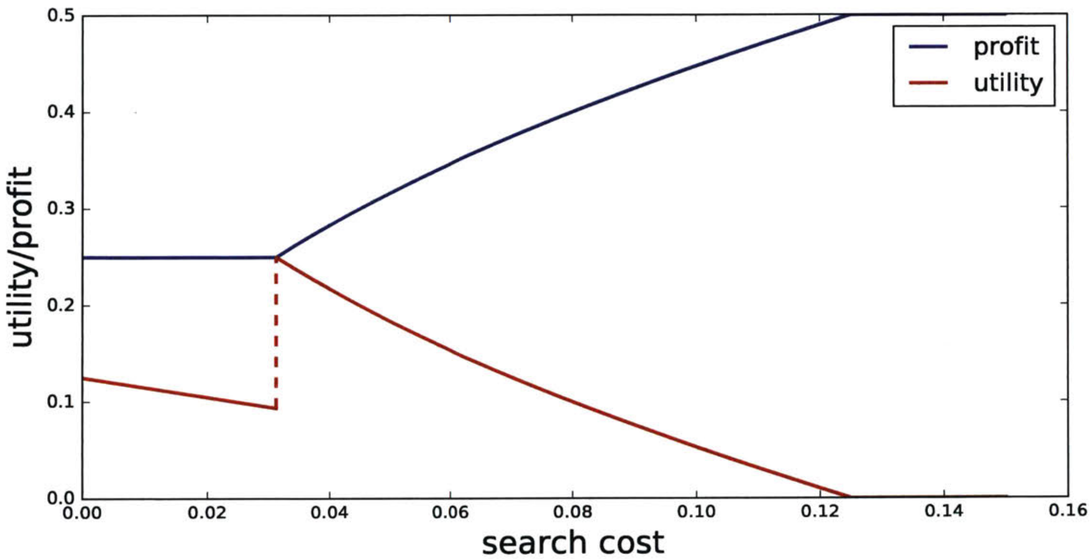
This leads to an equilibrium utility of

$$U^* = \begin{cases} \int_{\frac{1}{2}}^{1}(v - \frac{1}{2})dv - s = \frac{1}{8} - s, & \text{if} \quad s < \frac{1}{32} \\ \frac{1}{2} - \sqrt{2s}, & \text{if} \quad \frac{1}{32} \leq s \leq \frac{1}{8} \\ 0, & \text{if} \quad \frac{1}{8} < s. \end{cases}$$

We can see the equilibrium payoffs in Figure 1-10 below. The firm sets a constant price for low and high s but when $s \in (\frac{1}{32}, \frac{1}{8})$, it is bound by the highest price the consumer is willing to accept now instead of searching ($p_B = \sqrt{2s}$). The consumer's utility is first decreasing because she has to search, which becomes costlier in s, even though the benefit remains constant. At $s = \frac{1}{32}$, the firm changes its price from $\frac{1}{2}$ to $\frac{1}{4}$, and the consumer becomes willing to buy at that price. This improves her utility by a discrete amount. However, as the search cost increases further, the consumer will accept a higher price, which extracts more of her ex-ante surplus.

Under the assumption of uniformly distributed values, it is easy to compare the consumer's equilibrium utility to the one obtained in Roesler and Szentes (2017). In their paper, if the consumer has a zero search cost and is allowed to choose any Bayes-plausible posterior distribution, the optimal choice leads to an equilibrium price of $p^* \approx 0.2$ and utility of $U^* \approx 0.3$, because the consumer always searches and the optimal posterior distribution leads to purchase with certainty

Figure 1-10: Equilibrium utility and profit as a function of search cost when $v \sim U[0,1]$.



at the lowest possible price. However, if the consumer had even the slightest of search costs, she would not search at p^* because she could get the same utility by purchasing and not incurring the search cost. This would imply that the firm should raise its price to the maximum value the consumer is willing to accept without search. In fact, we can calculate numerically what the firm's optimal price should be in the Roesler-Szentes case when the consumer chooses the technology that leads to the following distribution of posterior expected values [34]

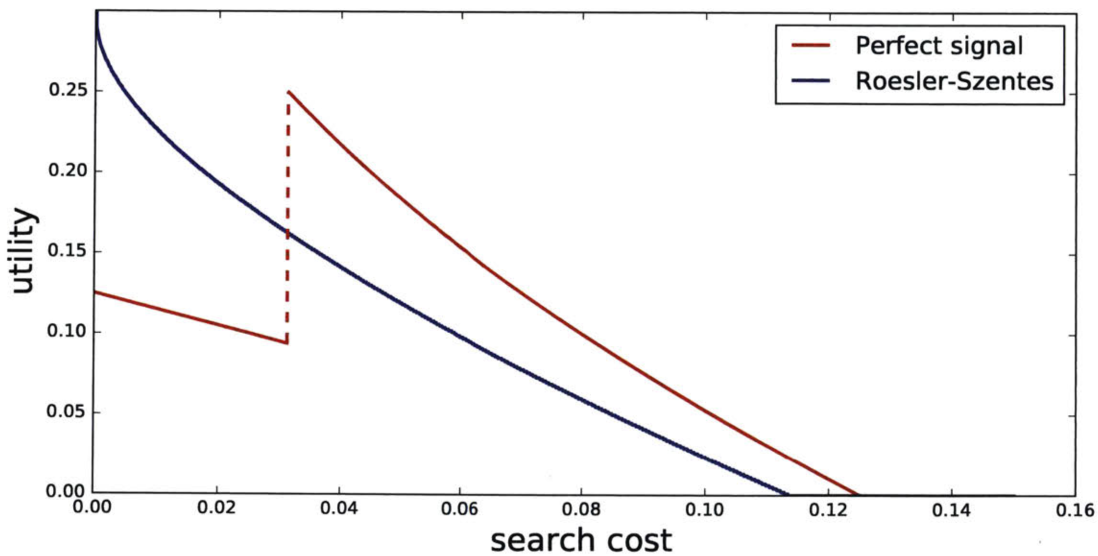
$$G_{0.2}^{0.87}(v) = \begin{cases} 0, & \text{if} \quad v < 0.2 \\ 1 - \frac{0.2}{v}, & \text{if} \quad 0.2 \leq v < 0.87 \\ 1, & \text{if} \quad v \geq 0.87, \end{cases}$$

but cannot commit to using the technology. Roesler and Szentes show that $G_{0.2}^{0.87}(v)$ is the consumer-optimal posterior distribution when she gets to choose and use it under no costs. Under the above distribution of posteriors, expected value of searching is $V_S(p) = 0.2\ln(\frac{0.87}{p}) - s$, while that of buying is $V_B(p) = 0.5 - p$. Now, $V_B(p) \geq V_S(p) \Leftrightarrow p - 0.2\ln(p) \leq 0.5 - 0.2\ln(0.87) + s$. The highest price the firm can charge while still making the consumer buy is $p_B(s) = \min\{\hat{p}(s), 0.5\}$ where $\hat{p}(s)$ satisfies:

$$\hat{p}(s) - 0.2\ln(\hat{p}(s)) = 0.5 - 0.2\ln(0.87) + s.$$

[34]This is only approximately true since their optimal price cannot be expressed analytically. This is the distribution of posterior expectations, meaning that each v corresponds to the expectation of some posterior distribution.

Figure 1-11: Equilibrium utility as a function of search cost when $v \sim U[0,1]$ under two signal structures.



Even though $\hat{p}(s)$ is not unique, we can choose the larger of the two candidates because that is optimal for the firm. In Figure 1-11 we plot the consumer's utility in two cases: 1) when she chooses the perfect signal (my setup), and 2) when she chooses the Roesler-Szentes signal. We see that the Roesler-Szentes signal is clearly better for the consumer when search costs are very low but, when the cost of searching is higher than $\frac{1}{32}$ (which, under the perfectly informative signal, makes the firm indifferent between searching and purchasing but leaves the consumer with the highest bargaining power), the consumer benefits from the perfect signal relative to the one that makes demand unit elastic (Roesler-Szentes).

1.6 Consumer Optimal Information Frictions

In this section, it is assumed that the consumer can costlessly choose from a menu of experiments each leading to two possible posterior expectations. However, conducting any one of these experiments will cost the consumer s. Note that all the experiments the consumer can choose will cost the same, no matter how informative they are in the Blackwell sense, which will only make the results stronger.[35] This setup is less general than the continuous setup of Section 1.5 but we get clean results.

We will consider two different setups: 1) one in which the consumer gets to choose which posterior she receives after a good signal but the bad posterior is fixed at zero, and 2) another

[35]Since we are studying how the informativeness of the consumer's signal matters in equilibrium, the result will be stronger if the consumer's cost of acquiring any signal technology is the same.

where the consumer gets to choose both the good and the bad posterior. In the first case, the consumer prefers an imperfect good posterior ($\pi_1(s) < 1$) for low values of search cost because otherwise it will be profitable for the firm to induce search. However, if the technology is not good enough (low π_1), the firm will want the consumer to buy and it has to set a low price to do this because the experiment still gives the consumer some bargaining power. In the second case, we learn that the consumer always prefers a perfectly informative good posterior $\pi_1 = 1$, but she may want a strictly positive bad posterior so as to incentivize the firm to set a low price. The consumer gets commitment power from the fact that she cannot improve her technology after the firm sets its price. If the consumer's search cost is high enough, she will always want the perfect experiment because it can only give her more bargaining power but will not change the firm's decision to sell the good without search.

These results complement the ones obtained in Roesler and Szentes (2017) by adding a search cost and focusing on the relationship between signal precision and search cost.

1.6.1 First Model: Choosing a Good Posterior

As in the previous section, there is one consumer (or a continuum of identical consumers) and one firm with a product whose quality can be high or low, yielding a consumption utility of 1 or 0, respectively. Prior probability on the quality being high is π_0. The firm has a cost of production, $c \geq 0$, and the consumer a search cost, $s \geq 0$. However, there is one twist compared to the model of Section 1.3: the consumer begins the game by **costlessly** choosing a search technology which we call π_1. This technology allows the consumer to conduct an experiment that leads to a posterior of $\pi_1 \geq \pi_0$ with probability $\frac{\pi_0}{\pi_1}$ and a posterior of 0 with the probability $1 - \frac{\pi_0}{\pi_1}$. Conducting the experiment costs s but the consumer can choose any experiment – and the firm observes this choice prior to setting its price. In other words, the timing is as follows: 1) the consumer chooses $\pi_1 \in [\pi_0, 1]$, 2) the firm observes this choice and sets $p \geq 0$, 3) the consumer makes a **one-time decision** to search or buy, and 4) payoffs are realized.[36]

One way to think about the search technology π_1 is as a search algorithm for reviews. If there was an option to enter your preferences on Amazon and automatically get a summary of all the reviews relevant to **your** preferences, your posterior would either be high or low. A better technology (higher π_1) would correspond to a larger range of posterior beliefs for a given prior. Products would be priced higher because firms would know the only ones to buy their products would be those who obtain high posteriors.

Note that this model is still of the same flavor as the baseline model because the consumer's posterior will always be 0 after a bad signal. In a sense, she can now choose to reach any good posterior π_1 at cost s, while in the baseline model she had to pay more to reach a higher posterior.

The following proposition computes the equilibrium of the game and shows that the consumer

[36]Note that there are no dynamics here. The consumer chooses the posterior she wishes to get by choosing her technology, and by using the technology at cost s, she has the chance of reaching that posterior. The fact that there are no dynamics means that the consumer's posterior does not depend on the firm's price directly, while in the previous model a higher price implied a higher purchase threshold $\overline{\pi}$.

may choose a less-than-perfect search technology. She will also prefer an interior search cost (as before).

Proposition 1.10. *Assume* $s \leq \frac{(\pi_1-\pi_0)\pi_0}{\pi_1}$, *and* $c \leq \pi_0$. *Let* $A(\pi_1, c) \equiv \frac{\pi_0(\pi_1-\pi_0)}{2\pi_1-\pi_0} + \frac{(\pi_1-\pi_0)^2}{\pi_1(2\pi_1-\pi_0)}c$. *Then, if search cost is fixed, the consumer's optimal choice of search technology is* $\pi_1^*(s) = \min\{1, \hat{\pi}_1(s)\}$, *where* $\hat{\pi}_1(s)$ *solves* $s = A(\hat{\pi}_1(s), c)$. *If both* s *and* π_1 *are for the consumer to choose, she will choose* $\pi_1^* = 1$ *and* $s^* = A(1, c)$.

What this proposition is saying is that the consumer wants the firm to set the low purchase price by choosing a bad enough search technology because, if she ends up searching, the firm can extract all of her surplus. The condition $s = A(\hat{\pi}_1(s), c)$ determines the search technology that just makes the firm indifferent between its searching and purchasing policies. If the search cost was any lower, the firm would prefer the search equilibrium and the consumer would not get anything.

Proof. See Appendix A.3. □

In particular, if the search cost is fixed, the proposition is saying that the consumer will often be better off not obtaining the best possible search technology even though that would be socially optimal. This is because the firm would then exploit this technology and make her search (it would know the technology exists when setting the price). When the consumer gets to choose her technology as well as the cost of using it (s), she always prefers the socially optimal technology. This is the most important insight of this section: even though s and π_1 are complementary tools in that both a high s and a low π_1 make the firm set the purchase price (p_B), the consumer prefers using s instead of π_1. Why? Because in equilibrium she will not end up searching, it does not directly matter what the value of s is. We know that in any equilibrium where the condition binds ($s = A(\pi_1, c)$) and $\pi_1 < 1$, the consumer prefers to raise both π_1 and s so that the condition still binds because that leads to a lower purchase price p_B.
So far we have been assuming that the consumer chooses her search technology before the firm sets its price. However, in some settings the opposite might be true. Therefore, the following proposition briefly explores what happens when the firm gets to commit to a price before the consumer chooses her technology.

Proposition 1.11. *Assume that* $s \leq \pi_0(1-\pi_0)$ *is fixed. Assume that the timing is slightly different from the previous propositions: 1) the firm sets p, 2) the consumer chooses* π_1, *3) the consumer buys, searches, or exits. Then the consumer's optimal policy is to always use the best search technology* ($\pi_1^* = 1$). *The consumer's using the optimal search technology will be reflected in the price, giving her the utility:*

$$V^* = \begin{cases} 0, & \text{if} \quad s < A(1, c) \\ \pi_0 - \frac{1}{1-\pi_0}s, & \text{if} \quad s \geq A(1, c), \end{cases} \tag{1.5}$$

where $A(1, c) = \frac{\pi_0(1-\pi_0)}{2-\pi_0} + \frac{(1-\pi_0)^2}{2-\pi_0}c$ *is the threshold where the firm changes its price.*

Note that this result is essentially the same as the results in Section 1.3, the only difference being that now the consumer can choose her technology. She would again want to have an intermediate search cost.

Proof. The consumer now chooses π_1^* to maximize $\frac{\pi_0}{\pi_1}(\pi_1 - p) - s$, after which she chooses to search (giving her a payoff of $\frac{\pi_0}{\pi_1^*}(\pi_1^* - p) - s$), purchase $(\pi_0 - p)$, or exit (0). We see immediately that the optimal choice of search technology is $\pi_1^* = 1$ because it does not impact the price in any way. Then we see that the consumer's search region is defined by $p \in \left(\frac{s}{1-\pi_0}, 1 - \frac{s}{\pi_0}\right]$ This means that the firm will be choosing between two prices: $p_B \equiv \frac{s}{1-\pi_0}$ and $p_S \equiv 1 - \frac{s}{\pi_0}$, leading to the cut-off search cost $A(1,c) = \frac{\pi_0(1-\pi_0)}{2-\pi_0} + \frac{(1-\pi_0)^2}{2-\pi_0}c$ (where the firm sets $p = p_S$ if and only if $s < A(1,c)$). Thus, the consumer's utility function can be written as:

$$V^* = \begin{cases} 0, & \text{if} \quad s < A(1,c) \\ \pi_0 - \frac{1}{1-\pi_0}s, & \text{if} \quad s \geq A(1,c). \end{cases}$$

□

This shows that the firm's issue in the previous model is one of commitment. There, the firm cannot commit to a price, which means that the consumer can affect the firm's decision by choosing an intermediate search technology. Here, however, the firm is able to commit, meaning that it can force the consumer to search and get zero utility when the search cost is low (which is essentially just Theorem 1.1 in disguise).

1.6.2 More General Model: Choosing Both Posteriors

So far, we have assumed that the consumer can choose her technology such that she will either perfectly learn that the product is of low quality, or she will obtain a posterior expectation of π_1 with probability $\frac{\pi_0}{\pi_1}$. Now, to be more general, and to allow the quality to be any real number between 0 and 1, assume that the consumer has the freedom to choose any experiment with two possible signals: a bad and a good signal. The only restriction is that the expected posterior has to be equal to the prior.

To fix notation, assume that the prior distribution of quality, θ, on $[0,1]$ is F, and that there are two signals which lead to posteriors F_B and F_G, respectively. The only requirement here is that the expected posterior distribution be equal to the prior. In other words, $F = qF_G + (1-q)F_B$, where q is the probability of a good signal. Assume also that $E_G \equiv \mathbb{E}_G[\theta] \geq \mathbb{E}[\theta] \equiv E \geq \mathbb{E}_B[\theta] \equiv E_B$, so that the good signal always (weakly) improves the expected quality relative to the prior (and similarly, the bad signal lowers the expected quality). The actual distributions do not matter, and we only need to know their expected values (as long as they satisfy the requirement that expected posterior be equal to the prior).[37]

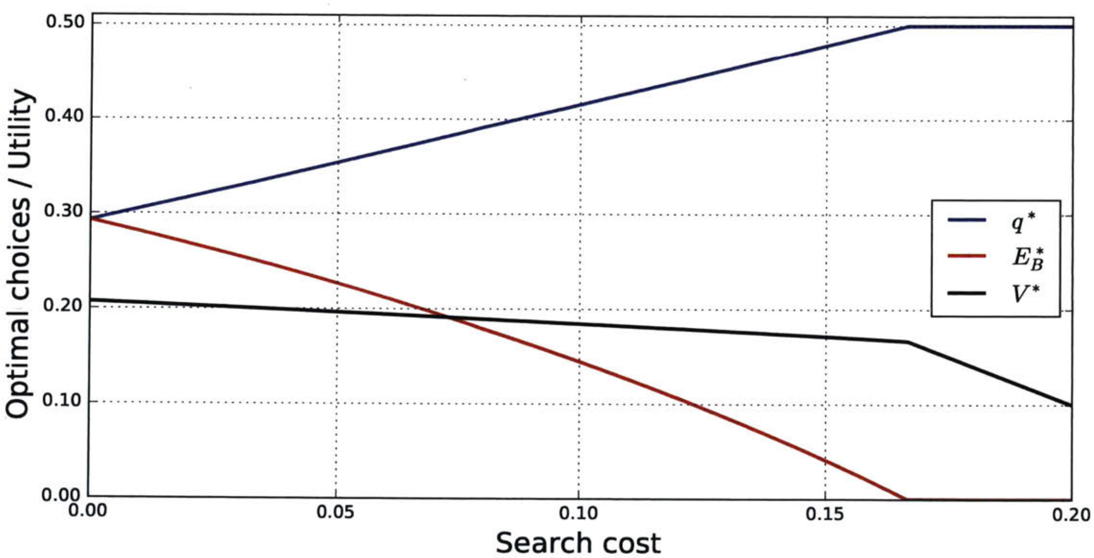
[37]Note that assuming only two possible signals is without loss of generality when the consumer can choose the posterior beliefs because the only two things that matter are 1) the probability that the posterior exceeds the price, and 2) expected posterior given that it exceeds the price. The first one is the only thing that matters for the firm

Proposition 1.12. *Assume that $s < 1 - E$ and $c < E$. The consumer's optimal search technology characterized by the good posterior, $E_G^* = 1$, and the probability of a good signal, $q^* = \min\{E, \hat{q}\}$, where $\hat{q}$ solves $s = A(\hat{q}, 1, E, c) \equiv q - \frac{E}{2-q} + \frac{(1-q)^2}{2-q}c$.*

Here, $A(q, E_G, E, c)$ is a search cost threshold such that the firm switches from making the consumer search to making her buy at the threshold. Thus, the consumer wants to be exactly at the threshold.

Proof. See Appendix A.3. □

Figure 1-12: Optimal q^*, bad posterior (E_B^*) and maximized utility (V^*), assuming $\pi_0 = \frac{1}{2}$ and $c = 0$.



In Figure 1-12, we see how the optimal probability of a high posterior (q^*) and the resulting bad posterior (E_B^*) depend on the search cost. Note that if the consumer was forced to choose $q = E$ (perfect experiment because $E_G^* = 1$), her utility would remain at zero until it jumped up at the point where $\hat{q}(s) = E$.

The intuition for the proposition is that the consumer wants to get the highest possible utility when a good signal is observed ($E_G^* = 1$) but she may want to make sure not to observe the good signal too often (if $q^* < E$). One can verify that, if it is easy for the consumer to search, she wants to have $q^* < E$, which also implies that $E_B > 0$.[38] This is because the firm does not care about the consumer's posterior conditional on it being above p, while the consumer wants the highest

when it is setting its price while the second matters for the consumer's utility. However, conditional on the firm setting a given price, the actual number of signals is inconsequential because the consumer can choose which price the firm will set.

[38]Simply because $E_G^* = 1$, $E_B = \frac{E-q}{1-q} > 0 \Leftrightarrow q < E$.

posterior because, combined with a low enough q, it gives her more bargaining power and forces the firm to charge a lower price. However, when searching is expensive (high s), the consumer wants the best possible technology because it improves her bargaining power and forces the firm to charge a lower price ($q^* = E$, $E_B = 0$, and $E_G = 1$). We can also state this as a corollary:

Corollary 1.5. *For $s \geq \frac{1-E}{2-E}[1-(1-c)(1-E)]$, the consumer will choose the perfect technology: $E_G^* = 1$ and $q^* = E$.*

Proof. $\hat{q} \geq E \Leftrightarrow 2(1-E)(1-c)+s \geq \sqrt{s^2+4(1-c)(1-E-s)} \Leftrightarrow s \geq \frac{1-E}{2-E}[1-(1-c)(1-E)]$. □

Corollary 1.6. *Assume that $c < E$ and that s is small enough for $q^* < E$. Then q^* is increasing in s and decreasing in c. Furthermore, this means that E_B^* is decreasing in s and increasing in c.*

Note how E_B^* is positive whenever $q^* < E$ (which happens when s is low), which makes sense because the consumer wants to limit the informativeness of her experiment when search cost is low. Similarly, a high production cost will make the firm more willing to induce search, which means that the consumer needs a worse technology – this manifests in the form of a lower q^* and a higher E_B^*. However, the posterior after a good signal always remains at 1 because that way the consumer can be sure that she is buying a good product after a good signal.

Proof. That the inequalities $\partial q^*/\partial s > 0$ and $\partial q^*/\partial c < 0$ hold can be seen by simply differentiating q^* (we need $q^* < E$, because otherwise there is no change). To obtain the results regarding E_B^*, we can just note that $E_B^* = \frac{E-q^*E_G^*}{1-q^*} = \frac{E-q^*}{1-q^*}$, which is decreasing in q. Thus, E_B^* moves in the opposite direction from q^*. □

Now, as an example, consider the case with $s = c = 0$. In this case, $q^* = 1-\sqrt{1-E} < E$, and $E_B^* = \frac{E-q^*}{1-q^*} = 1-\sqrt{1-E}$. Thus, the consumer will set a high E_B^* to get a better position vis-à-vis the firm. This is socially inefficient because we are not using the most effective technology but it is optimal for the consumer.

1.7 Concluding Remarks

In this paper, we have seen how the consumer may benefit from having an imperfect search technology and/or a high search cost because then the firm will find it optimal to set a lower price. However, the consumer always prefers to have a perfect technology if she can choose a high enough search cost so as to make the firm willing to sell without search. If the consumer cannot choose her search cost but she has more flexibility in choosing her technology, she may control the firm's behavior by choosing a technology that is not socially optimal (an experiment that is not perfectly informative).

The reason for non-monotone utility (and profits) is that the consumer has two actions that the firm can induce by pricing high or low. For low search costs, the consumer has a lot of bargaining power when she buys because the firm has to offer a low price to prevent her from searching. However, a low search cost means that the firm can set a very high price and still make the consumer search. This is why the firm prefers to induce search when search is cheap. On the other hand, when search costs are high, the consumer is willing to purchase the good at a higher price because she does not have as much bargaining power in case the firm wants to sell. The maximum price the firm can set and make the consumer search is low when search costs are high, so the firm prefers selling the good outright.

Social surplus in equilibrium is generally lower than what would be optimal if one party had all the bargaining power. If search cost was zero or so high that the consumer had no bargaining power, social surplus would be at the planner-optimal level. If there is an ex-ante contracting stage, we can always reach the socially efficient course of action by giving one party the full bargaining power and requiring that this party make a lump-sum transfer to the other party. The size of this transfer can be anything as long as both players are weakly better off than in the original equilibrium.

When we introduce a continuum of consumers with different valuations for a good match, we see that the main result holds: a non-zero search cost is still optimal for low production costs – both for the consumers and the society as a whole. However, the equilibrium of the game is never efficient in the first-best sense; a social planner would like to include more consumers but the firm is excluding them to extract more from those who participate.

It is left for future research to determine what happens in this game when there is competition and (ex-ante) product differentiation. It would also be interesting to learn what the effect of promotional reviews is: how do they matter for profits and utility?

Chapter 2

Optimal Search Frictions Under Competition: Searching for Match Quality

This paper examines the effects of consumers' ability to gather match-quality information on markets for differentiated products. Consumers, who costlessly observe all firms' prices, have access (at a cost) to product reviews which perfectly reveal whether a given product will meet their needs. Directed consumer search leads to a mixed-strategy pricing equilibrium which takes different forms at different levels of search cost. Initial reductions in search costs making such information gathering possible transfer surplus from firms to consumers. Additional reductions, however, can decrease consumer surplus (and increase profits) as firms find it less attractive to pay consumers to forego search. Additional effects are noted as reviews get even cheaper to access and consumer surplus, profits, and social surplus are all non-monotonic in the cost of reading reviews.

I am thankful to Glenn Ellison and Michael Whinston for their valuable feedback.

2.1 Introduction

Ubiquitous but understudied, crowd-sourced product reviews seem to matter for consumer behavior. Many of the markets, where reviews have become common (e.g. restaurants, movies, products of Amazon), are markets in which consumers would typically be choosing among a number of differentiated offerings. Characteristically in these markets, it is often difficult to tell two (or more) products apart prior to consuming or researching further. However, ex post, many of these products turn out to be of different match quality.

Consumer reviews can help both in matching of consumers to goods that fit their needs and in finding high-quality products. The former refers to matching each consumer to the horizontally differentiated product that fits her the best, while the latter refers to vertically differentiated products where consumers benefit from learning which products are the best on an absolute scale. I assume away any considerations for vertical quality, and focus solely on the horizontal aspect where each consumer is equally likely to find any given product to be a match. These matches are also assumed independent across consumers and products. Consumers' obtaining reviews sequentially, one product at a time, leads to asymmetries in competition, which may provide firms with incentives to price in a way that deters consumers from investing in further information gathering.

In this paper, I develop a model to examine both equilibrium pricing and search patterns, and the effects on welfare. I obtain interesting non-monotonicities in payoffs, as I let search costs vary.[1] In every equilibrium, the firms are playing symmetric mixed pricing strategies which depend on the value of the search cost so that expected prices are high for both high and low search costs but lower for intermediate values. The consumer is assumed to freely observe the prices of all the relevant, ex-ante identical but ex-post differentiated goods.

If anything, the products' being ex-ante symmetric should intensify competition between firms. To get away from pure Bertrand competition, however, it is assumed that the consumer has a particular need to fulfill and that there are no budget constraints so that if she purchases a good that turns out to be useless, she can still go and buy from another firm – as long as she has not found a match or exhausted all of her options.

To make matters simple, consumption utility is binary, unknown to all parties, and independent across firms and consumers. This means that any single firm's product is a good match for any consumer with an ex-ante probability of π_0, but it is uncertain which consumers like which firms. Therefore, prices contain no quality signals and no firm or consumer has an advantage over another. Each consumer has access to a signal technology (e.g. reading reviews) which perfectly reveals a given firm's type (so that the firm turns out to be a good match for this consumer with probability π_0). The consumer can use the technology at cost s per firm. Because she only has one need, she only needs to find one good satisfying that need, after which she will leave the market forever.

Under these assumptions, the consumer will always deal with the lowest-price firm first, and she will consider each firm separately, without having to take into account continuation values. When

[1] Search cost refers to how expensive it is for the consumer to search for information on match quality (be it through reading reviews or some kind of introspection).

dealing with a firm, the consumer can either purchase, search, or exit (which never happens in equilibrium). We will see that there are two intuitive threshold prices for the consumer's decision problem, so that the consumer will purchase for low prices, search for intermediate prices, and exit if prices are really high (not in equilibrium). The fact that the consumer may fail to find a match at a firm that charges a low price means that other firms have less of an incentive to price low. However, as the number of firms increases, prices decrease.

Knowing the consumer's behavior, each firm will use a symmetric mixed pricing strategy over a range of prices that depends on how expensive searching is. This is, as usual, due to the fact that the highest-price firm will essentially be a monopoly for those consumers who fail to find a match in the lower-priced firms. There are no pure-strategy equilibria. Thus, each price in the range gives the same expected profit (a firm may still get some demand even if it does not have the lowest price because the consumer may fail to find a match at the cheaper firms). One way to explain mixing is to assume that the consumer randomly faces a set of N firms selling an ex-ante homogeneous good so that the firms cannot know which competitors they are facing when a consumer is considering to buy the good. This means that the firm can only take the equilibrium distribution of prices as given without knowing what the realizations are.[2]

The effects of varying the search cost are interesting: When search costs are really high, searching is never an option, so all firms price below the prior and the consumer always purchases, starting from the cheapest firm, until she finds a match or exhausts her options. As the search cost decreases, some surplus will be transfered from the firms to the consumer in the form of lower prices due to her increased bargaining power (searching is a threat the consumer can use to keep prices low), but there is never any search because it is still too expensive.

As the search cost keeps going down, the firms will find it optimal to sometimes charge high prices which incentivize the consumer to search. This transfers surplus back to the firms from the consumer because the consumer may now be held up at times. The simple reason for this is that the consumer now lacks commitment power; her search cost is so low that the firms can charge high prices and she is willing to search. Social surplus decreases in this range as the search cost decreases because search is socially inefficient. When the search cost decreases even further toward zero, the firms will always price high and the consumer will always search before buying. In this range the consumer benefits from a lower search cost because it makes searching cheaper but the firms' competing against each other prevents prices from increasing much. Similarly, social surplus grows as the search cost decreases since searching becomes cheaper while the benefit remains constant.

As an example, consider a consumer who needs a new pair of sneakers. For simplicity, let there be two, ex-ante identical pairs that she is comparing. Now, she can just order the cheaper pair without any further trouble – or she can go to a brick-and-mortar store to try that pair on, which involves paying the search cost, s.[3] Because every consumer has a unique set of feet, the firms or

[2]One way to achieve this online is for a platform to present the consumer with a random set of results when she searches for a product. In some ways this is sensible because people use different search words and may have slightly different consideration sets for exogenous reasons – even if the products are ex-ante identical. The platform could also do the mixing for the firms, presenting every consumer-firm pair with random prices.

[3]We could also assume that, instead of trying the shoes on, the consumer can go on Amazon to read reviews from

the consumers cannot tell whether a particular shoe will be a good fit for a particular consumer. However, they share some prior expectation. Allowing the consumer to make multiple purchases until she finds a good that fits her needs is reasonable because, when the first pair she purchases turns out to be a bad fit, she can still go and buy the other firm's pair. The consumer's equilibrium utility first decreases in her search cost, obtaining a global minimum at a strictly positive search cost, after which it is increasing and obtains a global maximum at an even higher search cost, after which it is decreasing again until it levels off at the utility she would get if there were no quality signals – which incidentally equals her utility when there is no search cost.[4]

The intuition is that there is an interplay of two distinct effects: 1) the usual competitive effect that pushes prices down and makes the consumer better off compared to facing a single monopolist firm, and 2) a search effect which may either hurt or benefit the consumer. When search is cheap, the consumer is always made to search which is the worse for social surplus the higher is the search cost. Because, in this case, the consumer gets a constant share of the social surplus, she is hurt by an increase in the search cost when it is low. This mimics what happens with a monopolist. However, when the search cost crosses a certain threshold, the consumer starts to benefit from a higher search cost because there is an increasing probability that the prices the consumer faces are low enough for her not to search. Once search cost crosses another threshold, the consumer will always buy the good(s) without search but, because she has the option to search, she has some bargaining power which brings prices down. Therefore, a higher search cost in this range leads to lower consumer surplus because of reduced bargaining power. This also mirrors the case with a monopolist.[5] When search cost is even higher, it makes no difference because the consumer is never willing to search – eliminating all of her bargaining power.

In Section 2.5, we introduce strictly positive production costs to the above setup, but the main intuitions do not change, although the algebra gets messier. However, if production costs are high enough (e.g. $c > \frac{2}{7}$ when prior is $\pi_0 = \frac{1}{2}$), the consumer is best off not having any search cost. The most interesting observation is that equilibrium utility is decreasing in the production cost for any value of the search cost, and that the equilibrium where the consumer always searches is played for a larger range of s when production costs are higher.

This paper is organized as follows: Section 2.1.1 provides a quick overview of related literature, while Section 2.2 introduces the basic model of two firms which is analyzed in Section 2.3 where I show that every equilibrium involves the firms playing symmetric mixed strategies which depend on the search cost. I also show that having access to a match-quality signal may benefit or hurt the consumer, depending on her search cost. Section 2.4 generalizes the two-firm model to N firms, showing that the results remain the same, and that consumer surplus is strictly increasing in the

consumers with similar feet (which is up to her to judge).

[4]Think of the no-search-cost case as one where the consumer can fully and simply determine the match quality of a given shoe online, without having to consult the brick-and-mortars.

[5]We see that the case with two oligopolistic firms is the same as the one with a single firm for high and low search costs in that consumer surplus is moving in the same way. However, even though the consumer's utility jumps up discretely at some point when she is interacting with a single firm, with two firms it increases continuously and has no jumps (due to randomization).

number of firms for all values of the search cost. Section 2.5 shows that adding production costs does not change much. Section 3.5 concludes.

2.1.1 Earlier Literature

As shown in earlier research, internet has not lowered markups or prices (see, e.g., Ellison and Ellison (2009)), which suggests that search frictions may be good for consumers. A recent paper by Choi, Dai, and Kim (2016) shows that, in markets with horizontal differentiation and freely observable prices, increasing the value of search leads to higher prices.[6] If sellers are symmetric, prices increase whenever search costs decrease or the distribution of match values becomes more disperse. I obtain very similar results but the effect of decreasing search costs is more nuanced in my model because it is non-monotone. One of their contributions is that there is a pure-strategy equilibrium while I get a mixed-strategy one. I achieve this with a very simple model, while the model of Choi, Dai, and Kim (2016) is more complicated (although it also yields more results).

In Chapter 1, I showed analytically why and how easier access to relevant product information may actually hurt consumers facing a monopolist – the main reasons being endogenous pricing and lack of commitment. The same intuition carries over to the oligopoly case but now competition prevents the firms from extracting all of consumer surplus when search costs are low. Similar results for the monopoly case are obtained by Branco, Sun, and Villas-Boas (2012).

Ke, Shen, and Villas-Boas (2016) solve for the optimal way to search for information on multiple products in a continuous-time model, showing that a product is considered for search or purchase only if it gives high-enough expected utility. However, they ignore strategic pricing and only allow the consumer to buy once and for all, whereas most of my results stem from endogenous pricing and the fact that the consumer has the option to purchase multiple times.

Chen and Yao (2016) study consumers' sequential search using a structural model and find that low search costs do not necessarily mean higher consumer surplus because they may engage in excessive search. This is in line with my results although my explanation involves endogenously higher prices while theirs is bounded rationality (consumers not aware of website ranking rules). Similarly, Dukes and Liu (2015) find that a low search cost makes the consumers evaluate too many sellers, so that it is optimal for a platform to have a "sufficiently high" search cost when searching for match quality. However, their focus is on optimizing from the platform's point of view, balancing the needs of the firms and the consumers, and they separate depth (how deeply to search one firm) from width (how many firms to search) while I only allow for binary depth and focus more on the interplay of search costs and equilibria.

Wolinsky (1986) studies what conditions give rise to true monopolistic competition and whether imperfect competition can turn an oligopolistic market into a monopolistically competitive one. His model differs from mine in that his consumers do not observe prices for free – they have to pay a search cost to learn both the price and their valuation for the product. In my model, the consumer

[6]Traditionally, lower value of search would mean that a consumer who visits a given seller would be more likely to stay – leading to higher prices. However, here prices direct consumer search, which means that it is valuable to be the lowest-price seller. When search becomes more difficult, price competition becomes more severe.

knows the price, which allows her to purchase without searching first. I am also assuming, contrary to Wolinsky, that the consumer can, and often will, purchase multiple times until she finds a product that suits her. If the first product turns out to be of low quality, she can purchase again because the other product has the potential of giving her an excess utility of one. This is due to the assumption that the consumer's utility is always either 1 or 0 but nothing in between while Wolinsky allows for a continuum of valuations.

2.2 Basic Model and Consumer Behavior

Assume that there are two ex-ante identical sellers and one consumer. The product the consumer seeks to purchase can be either of low or high match quality, denoted as $\theta \in \{0,1\}$, where 1 corresponds to high quality. However, ex ante, the consumer merely has a prior, π_0, on the quality being high – and this prior is the same for each firm. The qualities are independent of each other. The product costs nothing to produce, and the firms themselves have no private information (they share the prior with the consumer). If the consumer purchases both products, her ex-post utility is $U = \max\{\theta_i, \theta_j\} - p_i - p_j$, where p_i and p_j are the prices charged by the two firms (this naturally generalizes to N firms). That is, she only needs to find one match. Therefore, her expected utility of purchasing from firm i alone is $U_i = \mathbb{E}[\theta_i] - p_i = \pi_0 - p_i$, where p_i is the price the firm charges. If she has already purchased from firm i, her utility of purchasing from firm j is $-p_j$ if firm i was a match, and $\pi_0 - p_j$ if it was not. The consumer can also pay a search cost, s, to **perfectly** learn her match quality with a firm. That is, the consumer can find out whether or not a firm's product is a match by either paying the search cost or buying the product. The following assumption states that the consumer can sample both firms if the first draw she gets is bad.

Assumption 2.1. *If the consumer chooses to buy from one firm, and gets a bad quality draw, she can still go and buy from the other firm. That is, she has no budget constraint.*

Intuitively, this assumption implies that the consumer will always deal with the lower-priced firm first. If we only allowed the consumer to buy once and for all, the option value of searching would be higher – sometimes it might even be better to search the higher-price firm first (see Lemma B.1.1 in Appendix B.1). However, when the consumer can buy multiple times until she gets a good draw, she should always choose the lower-price firm first.[7] Assumption 2.1 makes the game separable in that the consumer can consider each firm separately. The below lemma proves these statements.

Lemma 2.2.1. *Under Assumption 2.1, the consumer will always search or buy from the lower-price firm first, and consider the other firm only if the first product turns out to be of low match quality. The consumer's optimal strategy for each firm is the same as when facing a monopolist.*

Proof. See Appendix B.1. □

[7]Essentially, this is due to the fact that if the consumer wants to buy without search, the benefit is higher for the low-price firm, and if she were to search the high-price firm first, she would either end up buying that product or going back to the first firm and buying its good, which she could have done in the first place.

Now, under Assumption 2.1, we know that the consumer will always choose the low-price firm first. She will purchase from a given firm if and only if $p \le \min\{\pi_0, p_B\}$, and she will search that firm if and only if $p_B < p \le p_{max}$, where we have written $p_B \equiv \frac{s}{1-\pi_0}$ and $p_{max} \equiv 1 - \frac{s}{\pi_0}$ for the highest purchase and search price, respectively. Note that if $s \ge \pi_0(1-\pi_0)$, the consumer will never search (since $p_{max} \le p_B$) and she will buy if and only if $p \le \pi_0$ (since $\pi_0 \le p_B$). As in the proof of Lemma 2.2.1, assuming that firm i has the lower price, we can write the consumer's expected utility as $U_{ij} = V(p_i) + (1-\pi_0)V(p_j)$, where $V(p) = \max\{0, \pi_0 - p, \pi_0(1-p) - s\}$. This utility can take three forms depending on the prices:

$$U_{ij} = \begin{cases} \pi_0 - p_i + (1-\pi_0)(\pi_0 - p_j), & \text{if } \ p_i \le p_j \le p_B \\ \pi_0 - p_i + (1-\pi_0)[\pi_0(1-p_j) - s], & \text{if } \ p_i \le p_B < p_j \le p_{max} \\ \pi_0(1-p_i) - s + (1-\pi_0)[\pi_0(1-p_j) - s], & \text{if } \ p_B < p_i \le p_j \le p_{max}, \end{cases}$$

where we have assumed that $s < \pi_0(1-\pi_0)$ because otherwise the consumer always purchases the product (as long as the firms set $p \le \pi_0$, which is true in equilibrium).[8] Note that the consumer will never search firm i but purchase from firm j without search because $p_i \le p_j$ and the search decision is increasing in the price.

2.3 Pricing Game and Non-Monotone Utility

Next, I will show that the type of equilibrium the firms will play depends strongly on the value of the search cost, s, and that the consumer's equilibrium utility is non-monotone in s. In fact, it achieves both its unique global maximum and minimum at separate, strictly positive values of search cost.

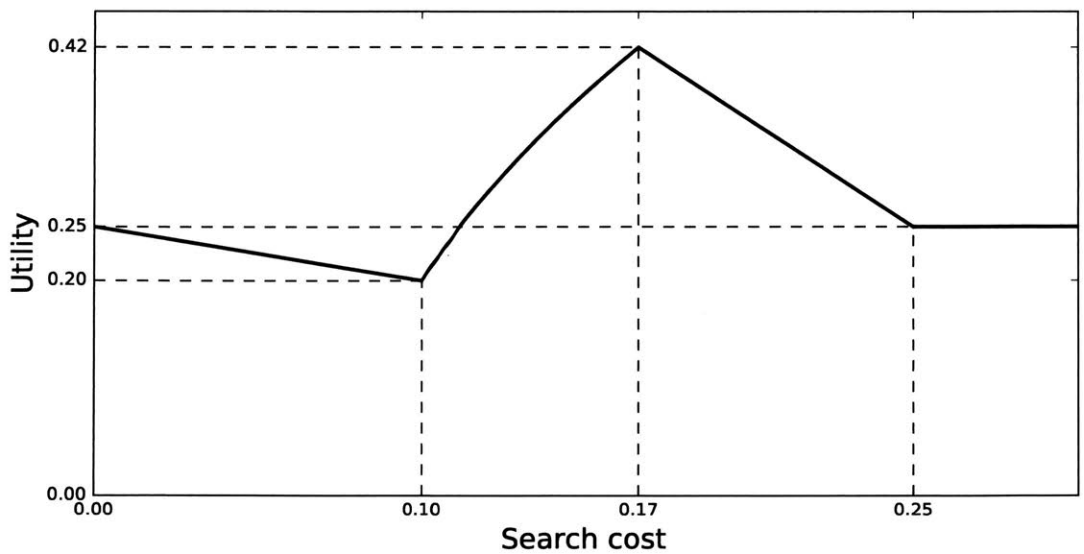
Theorem 2.1. *Assume that $\pi_0 \in (0,1)$. If Assumption 2.1 holds, the game has a unique mixed-strategy equilibrium for each value of the search cost, s. The form of this equilibrium depends on the exact value of s: For a low s, both firms always price high and the consumer always searches. For intermediate s, the firms sometimes price low and sometimes high, so that the consumer sometimes purchases without search and sometimes searches first. For high s, both firms price low so that the consumer never searches.*

The consumer's equilibrium utility is continuous but non-monotone in the search cost: first decreasing, then increasing, and decreasing again before leveling off, so that utility for a high s always equals that of $s = 0$. Utility attains its unique global minimum at $\hat{s} := \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}$, while the unique global maximum is achieved at $s^* := \frac{\pi_0(1-\pi_0)}{2-\pi_0}$, where $0 < \hat{s} < s^*$.

Assuming a prior of $\pi_0 = \frac{1}{2}$, we can see the Theorem in action in Figure 2-1. Here, the global minimizer is $\hat{s} = \frac{1}{10}$, while the maximizer is $s^* = \frac{1}{6}$. Note how the utilities for no search cost and

[8]Note that we have assumed the consumer always chooses firm i if the prices are equal but this will not matter in equilibrium because the probability of a tie is zero (due to there being only a mixed equilibrium).

Figure 2-1: Equilibrium utility as a function of s, fixing $\pi_0 = \frac{1}{2}$



very high search costs are the same. After the following proof, we will analyze the behaviors that lead to this graph.

Proof. Now, assuming that there is a lower-priced firm, the consumer should always choose this firm because the firms are ex-ante identical and have no informational advantage over the consumer.[9] Her optimal strategy at a given firm is the same as if she was facing a monopolist, as was shown in Lemma 2.2.1. Once the consumer has searched or purchased (both of which allow her to learn the type of the firm she deals with), she may search/purchase from the second firm – but this only happens if the first firm turns out to be a mismatch.

Assuming $s < \pi_0(1-\pi_0)$, the lower-price firm will then get:

$$\Pi_L(p) = \begin{cases} p, & \text{if } \ p \leq p_B \\ \pi_0 p, & \text{if } \ p_B < p \leq p_{max}, \end{cases}$$

while the higher-price firm will get:

$$\Pi_H(p) = (1-\pi_0)\Pi_L(p) = \begin{cases} (1-\pi_0)p, & \text{if } \ p \leq p_B \\ (1-\pi_0)\pi_0 p, & \text{if } \ p_B < p \leq p_{max}. \end{cases}$$

This is because the higher-price firm gets a customer only if the one with the lower price turns out to be of low match quality. Thus, if the high-price firm knew it had the highest price, it

[9]If both firms set the same price, the consumer can randomize.

should always charge p_B or p_{max}, whichever maximizes profits. However, there is no pure-strategy equilibrium as we will see next.

Given the consumer's strategy, the firms are going to play a pricing game (à la Bertrand). However, this game will not lead to marginal cost pricing, and the form the equilibrium takes depends on the search cost.

Lemma 2.3.1. *There is no equilibrium where both firms charge a fixed price. This holds for all s.*

Proof. See Appendix B.2. □

Lemma 2.3.2. *There is no equilibrium where one firm charges a fixed price and the other randomizes.*

Proof. See Appendix B.2. □

The above two Lemmas show that the equilibrium has to have both firms mixing over some range. The range has to be the same for both firms because otherwise one firm would have a profitable deviation. The range of prices depends on the value of the search cost as we will see next.

Lemma 2.3.3 (Low search cost). *Assume that $s \leq \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}$. The **unique** equilibrium involves both firms randomizing over $[p_S, p_{max}]$, where $p_{max} := 1 - \frac{s}{\pi_0}$ and $p_S := (1-\pi_0)p_{max}$. Probability of choosing a price less than p is given by $F(p) = \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0 - s)}{\pi_0^2 p}$.*

Lemma 2.3.3 says that when cost of search is low enough, both firms find it optimal to always set prices that induce search from the consumer. This is exactly what the condition $s \leq \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}$ does; it guarantees that a firm prefers setting p_S over p_B.[10] Note that in this range the search cost is always less than $\pi_0(1-\pi_0)$ so that the game is interesting.

Proof. Now, assuming a symmetric equilibrium (shown at the end of the proof), setting p_{max} gives a profit of $(1-\pi_0)\pi_0 p_{max} = (1-\pi_0)(\pi_0 - s)$ (because the consumer will only search your firm if the other one turns out to be bad), while setting p_S gives $\pi_0 p_S$ (because the consumer will search your firm first). These have to be equal, which implies that $p_S = (1-\pi_0)p_{max}$. Because we have assumed $s \leq \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}$, the firm setting p_S will never consider setting any lower price. The last thing to do is to find $F(p)$, which can be done by equalizing the profits from p_{max} to those from any $p \in [p_S, p_{max}]$:

$$[F(p)(1-\pi_0) + 1 - F(p)]\,\pi_0 p = (1-\pi_0)(\pi_0 - s)$$
$$\Leftrightarrow F(p) = \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0 - s)}{\pi_0^2 p}.$$

[10]Both of these prices would make the consumer choose the firm first (because the other one will be setting a higher price with probability one in equilibrium). Profit from setting p_S is $\pi_0(1-\pi_0)(1-\frac{s}{\pi_0})$ because the consumer's search will be successful with probability π_0, while p_B gives $\frac{s}{1-\pi_0}$, where I have simply used the definitions of these prices. The former profit exceeds the latter if and only if $s \leq \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}$. Note that $\pi_0 p_S > p_B \Rightarrow p_S > p_B$.

This shows that we have an equilibrium. Is it unique? The answer is yes because the firm setting the highest price will always choose p_{max} in any equilibrium (it gives the highest profits given not being the consumer's first choice). This means that the lower bound of the mixing range cannot be lower than p_S (any price between p_B and p_S leads to lower profits than p_S because the consumer will search either way and you will be the lower-price firm for sure, while any price below p_B leads to lower profits than p_S as seen in footnote 10). If the firms were mixing over $[p_S - x, p_{max}]$ for $x > 0$, the firm setting p_S would get a lower profit than when setting p_{max} because there would be a positive probability of the other firm undercutting, which would break the equilibrium.

As for the symmetric equilibrium being the only one, note first that no firm will ever charge $p > p_{max}$ because the consumer would never purchase at that price. On the other hand, for low s, one firm's pricing range has to include p_{max} because whichever firm is charging the highest price, would want to choose p_{max} instead of any lower price. Now, assume that firm i randomizes over a range containing p_{max}, and also assume that the other firm randomizes using some distribution $F(p)$ on some range. Then, firm i's profit when choosing price p will be $\Pi(p) = [1-F(p)+(1-\pi_0)F(p)]\pi_0 p$ (the first term is the probability of getting a consumer and π_0 is the probability that this consumer purchases). This profit has to be equal to $(1-\pi_0)\pi_0 p_{max}$ for all p in firm i's range because of the usual indifference condition for randomization. Solving this equation allows us to find $F(p)$ for the other firm, and we see that this $F(p)$ is continuously increasing between p_S and p_{max}, as defined in the Lemma. This also means that both firms will have the same upper bound, p_{max}, the same lower bound, p_S, and the same distribution, $F(p)$. □

Lemma 2.3.4 (High search cost). *Assume that* $\frac{\pi_0(1-\pi_0)}{2-\pi_0} \leq s \leq \pi_0(1-\pi_0)$. *In the unique equilibrium of the pricing game, both firms randomize over* $p \in [s, p_B]$, *where* $p_B = \frac{s}{1-\pi_0}$, *with* $F(p) = \frac{1}{\pi_0} - \frac{s}{\pi_0 p}$.

Lemma 2.3.4 says that when the cost of search is high (but not too high so that the game is interesting), both firms always set prices that induce the consumer to purchase without searching first. $s < \pi_0(1-\pi_0)$ is required for search to be possible under some prices. The lower bound for the search cost guarantees that the firm setting the highest price (p_B) prefers doing that over setting p_{max} – leading to a purchase with certainty if the consumer ever chooses that firm. It is also easy to show that the lower bound for the search cost always exceeds the upper bound from Lemma 2.3.3.[11]

Proof. Note that the firms always make the consumer purchase. The reason is that for $s \geq \frac{\pi_0(1-\pi_0)}{2-\pi_0}$, a firm prefers p_B over p_{max} given that the probability of the consumer choosing the firm is $(1-\pi_0)$. Thus, the firm that sets the higher price will always set p_B rather than any search price, which means that both firms have to be setting prices that are at most p_B. However, the upper end of the pricing range has to be exactly $p_B = \frac{s}{1-\pi_0}$ because otherwise the higher-price firm would have forgone profits. This implies that the lower bound for the pricing range has to be $(1-\pi_0)p_B = s$

[11] $\frac{\pi_0(1-\pi_0)}{2-\pi_0} \geq \frac{\pi_0(1-\pi_0)^2}{(1-\pi_0)^2+1} \Leftrightarrow \pi_0 \geq 0$, which holds.

so that the profits will be the same. Finally, $F(p)$ can be obtained by equating profits from the different prices.

The symmetry and uniqueness of the equilibrium is due to the same argument as in the previous Lemma. □

Lemma 2.3.5 (Very high search cost). *Assume that $s > (1-\pi_0)\pi_0$. The unique equilibrium has both firms mixing over prices $p \in [(1-\pi_0)\pi_0, \pi_0]$ with $F(p) = \frac{1}{\pi_0} - \frac{1-\pi_0}{p}$.*

In the search cost range of Lemma 2.3.5, the consumer will never search, and will be willing to purchase for any price $p \leq \pi_0$, so that both firms should always price below (or at) π_0.

Proof. Because s is too high, the consumer will never search for any price $p \geq \pi_0$ so there will be no search. This means that the firm that prices at the upper end of the price range has to set π_0. This implies that the lower bound has to be $(1-\pi_0)\pi_0$ so that profits will be the same at both prices. Finally, $F(p)$ can be solved by equating the profits from any p in the range to $(1-\pi_0)\pi_0$. The uniqueness is obvious based on the previous two Lemmas. □

The equilibria for the above values of search cost have been simple in the sense that the firms have always charged prices that lead to the consumer taking the same action for both firms. However, as the following Lemma will establish, this is not the case for some intermediate search costs.

Lemma 2.3.6 (Intermediate search cost). *Assume that $\frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1} < s < \frac{\pi_0(1-\pi_0)}{2-\pi_0}$. Now, the game has a unique equilibrium where the firms randomize over $p \in [(1-\pi_0)(\pi_0 - s), p_B] \cup [p_H, p_{max}]$, where $p_B := \frac{s}{1-\pi_0}$, $p_H := \frac{s}{(1-\pi_0)\pi_0}$, and $p_{max} := 1 - \frac{s}{\pi_0}$. The distribution function is given by:*

$$F(p) = \begin{cases} \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0-s)}{\pi_0 p}, & \text{if} \quad p \in [(1-\pi_0)(\pi_0-s), p_B] \\ \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0-s)}{\pi_0^2 p}, & \text{if} \quad p \in [p_H, p_{max}]. \end{cases}$$

Lemma 2.3.6 handles the last remaining case of intermediate search costs. We have already explained the thresholds: under these values of s, the search equilibrium is not possible because the firm setting p_S would prefer p_B, and, on the other hand, the purchase equilibrium is not possible because the firm setting p_B would prefer p_{max}. Therefore there will be a gap in the equilibrium distribution of prices, as we will show next.

Proof. The consumer will purchase if and only if $p \leq p_B = \frac{s}{1-\pi_0}$, which means that the firms should never price between p_B and $p_H = \frac{s}{(1-\pi_0)\pi_0}$ (where p_H is constructed by equating the profits from p_B to those from p_H, assuming the probability of consumer choosing the firm does not change). This is why there is a gap in the support of equilibrium prices. Note also that because s is in the range assumed in the Lemma, all else equal, the firms prefer p_{max} over p_B and they prefer p_B over p_S (where p_S was defined earlier as the lower bound when both firms make the consumer search).

Now, the highest price has to be p_{max} because that is preferred to p_B (and all other $p < p_B$). This means that the lowest price has to give the same expected profit. Thus, it has to be $(1-\pi_0)(\pi_0 - s)$.

We have now defined the highest and the lowest price. We also know that p_B has to be in the support because if the upper bound of the lower range was $p < p_B$, the firm setting p would like to raise its price to p_B and lose nothing. On the other hand, p_H, which we defined in the previous paragraph, gives the same probability as p_B of the consumer choosing that firm (since there is nothing between the two prices), so they have to give the same expected profits given the consumer choosing them. Now, we can just construct the cdf for prices as follows:

$$[F(p)(1-\pi_0)+1-F(p)]\,p = (1-\pi_0)(\pi_0 - s), \quad \text{if} \quad p \leq p_B$$
$$[F(p)(1-\pi_0)+1-F(p)]\,\pi_0 p = (1-\pi_0)(\pi_0 - s), \quad \text{if} \quad p \geq p_H.$$

Solving this gives the $F(p)$ introduced in the Lemma. □

Note that this equilibrium will converge to the equilibria mentioned in the previous Lemmata when s approaches the limits. To understand the mixing probabilities, we can take a look at Figures 2-2, 2-3, and 2-4 below. Fixing $\pi_0 = \frac{1}{2}$, these figures depict the probability density functions of prices when the firms are playing their (symmetric) equilibrium strategies, and we let the search cost vary. In Figure 2-2, we compute the densities for three different values of search cost (0, 0.05, and 0.1), all of which lead to an equilibrium where both firms always price high to incentivize the consumer to search. Note, however, that both the lower and upper bound of the mixing range go down as search costs increase. In Figure 2-3, we choose three search costs (0.105, 0.13, and 0.16) such that the firms sometimes set prices that make the consumer buy, and sometimes prices that make her search. When s is just above 0.1 (which is the knife-edge case that still makes both firms always price high), there is only a small probability that a given price will be low and a large range of higher prices that make the consumer search. However, as the cost of search increases, the range of purchase prices (low prices) widens and that of search prices shrinks (continuously).

Figure 2-2: Equilibrium distributions of prices, as a function of search cost, fixing $\pi_0 = \frac{1}{2}$

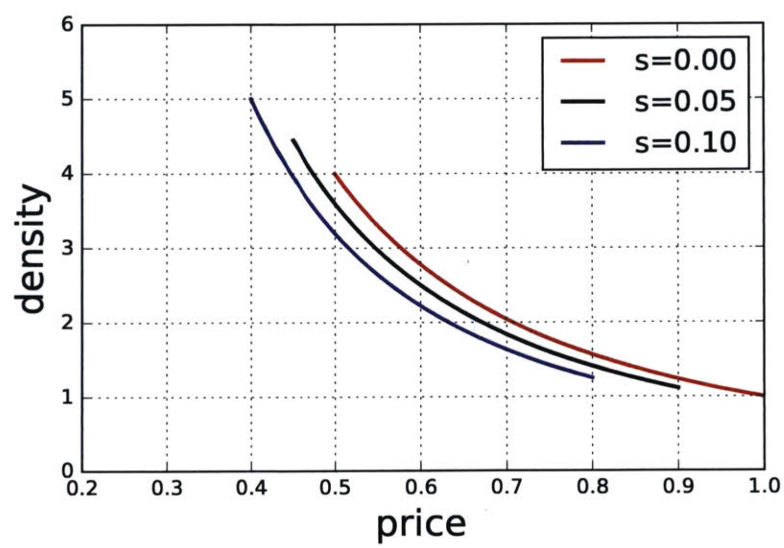
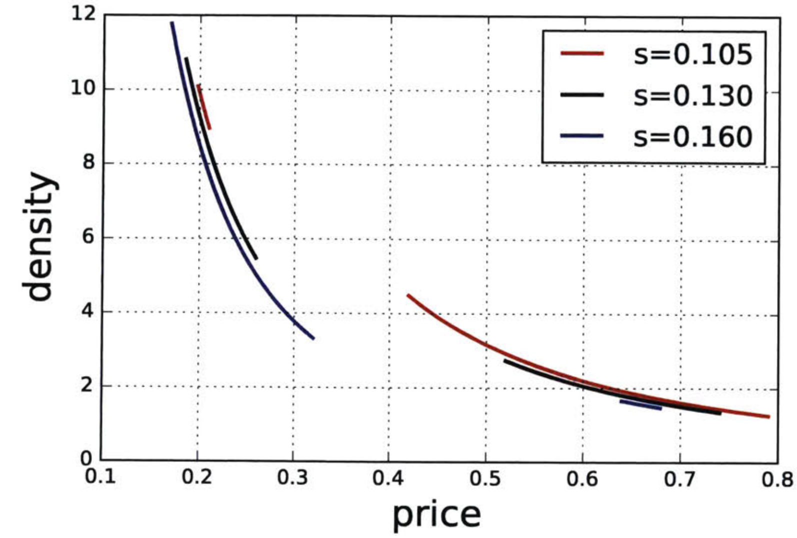
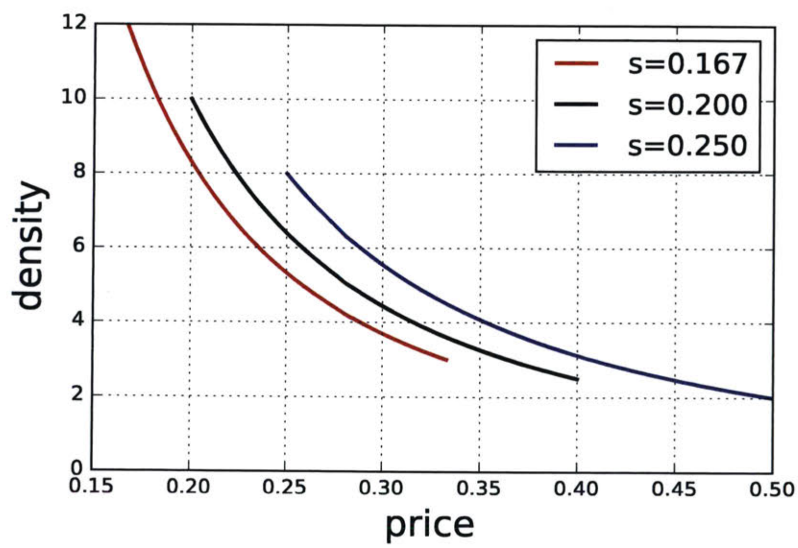


Figure 2-3: Equilibrium distributions of prices, as a function of search cost, fixing $\pi_0 = \frac{1}{2}$



Finally, in Figure 2-4, we consider some high values of search cost ($\frac{1}{6}$, $\frac{1}{5}$, and $\frac{1}{4}$), all of which lead to prices that always make the consumer buy (and never search). Now, as the search cost increases, the distribution of prices moves right because the consumer's bargaining power decreases, which allows the firms to price higher. As s increases beyond $\frac{1}{4}$, there is no change because this is the search cost where the consumer loses all of her bargaining power. Note, however, that her utility will not go to zero because the firms are still competing in prices.

Figure 2-4: Equilibrium distributions of prices, as a function of search cost, fixing $\pi_0 = \frac{1}{2}$



We can use all of the previous to compute the consumer's utility and the firms' profits at each search cost:

Lemma 2.3.7. *In equilibrium, the consumer's utility and a firm's profits are:*

$$U^* = \begin{cases} \pi_0(\pi_0 - s), & \textit{if} \quad s \leq \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1} \\ \pi_0^2 + (1-\pi_0)^4(2 - \frac{\pi_0}{s}) + (1-\pi_0)[(1-\pi_0)^2 + 3 - \pi_0]s, & \textit{if} \quad s \in \left(\frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}, \frac{\pi_0(1-\pi_0)}{2-\pi_0}\right) \\ \pi_0(2-\pi_0) - 2s, & \textit{if} \quad s \in \left[\frac{\pi_0(1-\pi_0)}{2-\pi_0}, \pi_0(1-\pi_0)\right] \\ \pi_0^2, & \textit{if} \quad s > \pi_0(1-\pi_0) \end{cases}$$

$$\Pi_i^* = \begin{cases} (1-\pi_0)(\pi_0 - s), & \textit{if} \quad s \leq \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1} \\ (1-\pi_0)(\pi_0 - s), & \textit{if} \quad s \in \left(\frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}, \frac{\pi_0(1-\pi_0)}{2-\pi_0}\right) \\ s, & \textit{if} \quad s \in \left[\frac{\pi_0(1-\pi_0)}{2-\pi_0}, \pi_0(1-\pi_0)\right] \\ \pi_0(1-\pi_0), & \textit{if} \quad s > \pi_0(1-\pi_0). \end{cases}$$

Proof. Profits of firm i are simply the same as the expected profits of the firm setting the highest price, which in the first two cases gives $(1-\pi_0)\pi_0 p_{max} = (1-\pi_0)(\pi_0 - s)$. On the other hand, in the third case, the highest-price firm gets $(1-\pi_0)p_B = s$, and in the fourth case, it gets $(1-\pi_0)\pi_0$.

The consumer's utility is slightly more difficult to compute but once we understand that it is the same as social surplus minus the two firms' combined profits, the problem becomes easy. In the first case, social surplus is $S = \pi_0 - s + (1-\pi_0)(\pi_0 - s) = (2-\pi_0)(\pi_0 - s)$ because the consumer searches once for sure and searches again if she gets a bad draw. Subtracting the firms' profits from this number gives the consumer's equilibrium utility. Similarly, in the last two cases, social surplus is $S = (2-\pi_0)\pi_0$, so subtracting two times firm i's profits will give the consumer's utility.

However, case number two is more difficult. If both firms set $p \leq p_B$, social surplus is $(2-\pi_0)\pi_0$. If both firms set $p > p_B$, social surplus is $(2-\pi_0)(\pi_0 - s)$. If the firms price on different sides of p_B, social surplus is $(2-\pi_0)\pi_0 - (1-\pi_0)s$ (because there will be search if the first product turns out to be of low quality after purchase). Multiplying these surpluses with the probabilities of them happening gives the social surplus:

$$S = F(p_B)^2(2-\pi_0)\pi_0 + 2F(p_B)[1-F(p_B)][(2-\pi_0) - (1-\pi_0)s] + [1-F(p_B)]^2(2-\pi_0)(\pi_0 - s),$$

where $F(\cdot)$ is the distribution from Lemma 2.3.6 because we are considering those values of s. Subtracting the firms' combined profits from this gives the expression for the consumer's utility given in the Lemma. □

Now, we can see that the consumer's equilibrium utility starts off at π_0^2, after which it decreases in s (slope of $-\pi_0$), reaching its unique global minimum of $U_{min} = \frac{\pi_0^2}{(1-\pi_0)^2+1}$ at $s = \hat{s} := \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}$. Then it starts to increase in s and reaches its **unique global maximum** of $U_{max} := \frac{2\pi_0 - (2-\pi_0)\pi_0^2}{2-\pi_0}$ at $s = s^* := \frac{\pi_0(1-\pi_0)}{2-\pi_0}$. This is the unique maximum because we assumed that the prior belief is strictly between 0 and 1.[12] After the maximum, equilibrium utility starts to decrease in s (slope of -2) until $s = (1-\pi_0)\pi_0$, after which there can never be any search and the equilibrium is the same as it would be if there was no option to search in the first place. This gives a utility of π_0^2, just like the case with $s = 0$. □

The theorem we have just shown offers some interesting insights that can be inferred from Figure 2-1. When search cost is low ($s \leq \frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1}$), the consumer will always search unless the firms set really low prices, which means that the benefit of setting a low price is low, and leads to the only equilibrium being one of high prices that induce search. These prices can be the higher the lower the consumer's search cost is because, in essence, the firms are compensating the consumer for searching. However, different from the case of a unique firm, there is price competition, which means that the consumer will get strictly positive utility.

If we denote the social surplus from searching exactly once by $S_1 := \pi_0 - s$, we see that each firm always gets an expected profit of $(1-\pi_0)S_1$ and the total social surplus from two potential searches is $(2-\pi_0)S_1$. Intuitively, the firms get a constant share of total surplus because the highest-price firm always gets the full social surplus, given that the consumer fails to find a match in the lower-priced firm. Thus, because all of the prices give each firm the same expected profit, the firms always get a constant share of social surplus. Naturally, the consumer then gets a constant share of the surplus as well. That is, she gets $\pi_0 S_1$. In effect, the consumer is splitting the social gains from the first search with the first firm, while the second firm gets the expected surplus from the second search.[13] This explains why the consumer's utility is decreasing in her search cost – she is getting a constant share of the social surplus which is decreasing in the search cost because search is always socially inefficient due to there being no production costs.

Another way to see this is to notice that the consumer gets the full social surplus but has she has to pay a transfer to the firms. The consumer's expected utility is $U = (2-\pi_0)(\pi_0 - s) - \pi_0(\mathbb{E}[p_{low}] + (1-\pi_0)\mathbb{E}[p_{high}]) = (2-\pi_0)(\pi_0 - s) - 2(1-\pi_0)(\pi_0 - s)$, where p_{low} and p_{high} are the expected lowest and highest price, respectively (order statistics). Now, increasing the search cost has two effects: 1) it will decrease social surplus at a rate of $(2-\pi_0)$ but it will also shift the firms' price distributions so that the consumer's expected transfers decrease at a rate of $2(1-\pi_0)$ (she is searching with probability one so a higher search cost just means lower prices). The overall effect is negative but the social loss is greater than the consumer's.

However, when search cost increases further ($\frac{(1-\pi_0)^2\pi_0}{(1-\pi_0)^2+1} < s < \frac{\pi_0(1-\pi_0)}{2-\pi_0}$), the firms will find it optimal to sometimes set a lower purchase price, sometimes a higher search price. Here, a higher search cost implies a larger range of purchase prices and smaller range of search prices, which benefits the consumer because she does not have to search as often, and she has some bargaining power versus the firms (her search cost is not too high, giving her the opportunity to search). The consumer benefits from a higher search cost in this range because, even though her bargaining power is decreasing in s, she will not have to search as often. One way to understand this is to notice that

[12] $U_{max} > \pi_0^2 \Leftrightarrow \pi_0 \in (0,1)$.

[13] The second firm is always essentially a monopolist although it has randomized its price before the realization of the first firm's type.

the consumer is the residual claimant to the full social surplus (the transaction between her and the firms is just a transfer). In this range, social surplus is increasing in the search cost (which can be verified numerically or differentiating the expression for social surplus), which is due to search being socially wasteful (and higher search cost implying less search). Furthermore, the expected prices of the firms are decreasing in the search cost, which in turn follows from the observation that a shrinking search range ($[p_H, p_{max}]$) benefits the consumer more than the widening purchase range hurts her.[14]

Once search cost enters the next range $\left(\frac{\pi_0(1-\pi_0)}{2-\pi_0} \leq s \leq \pi_0(1-\pi_0)\right)$, the consumer will always purchase the good in equilibrium because the firms find it too expensive to induce search (since the consumer has to be compensated more for search and she can be charged higher purchase prices). However, the consumer still has some bargaining power because her search cost is not prohibitively high – this leads to lower prices. Naturally, her equilibrium utility is decreasing in the search cost because the firms are already making her purchase and a high search cost means lower bargaining power. Now, social surplus is constant but the consumer is hurt by the increasing prices as search cost increases.

Finally, when search cost is high ($s > (1-\pi_0)\pi_0$), there is no chance the consumer will ever search, which eliminates all the bargaining power the consumer may have had so far (she buys even when price equals the prior). However, the firms will still be competing against each other à la Bertrand, giving the consumer strictly positive utility. This equilibrium does not depend on the exact value of the search cost, as long as its high enough. Moreover, the consumer (and the firms) will get a constant share of the social surplus, for the same reason as in the first case.

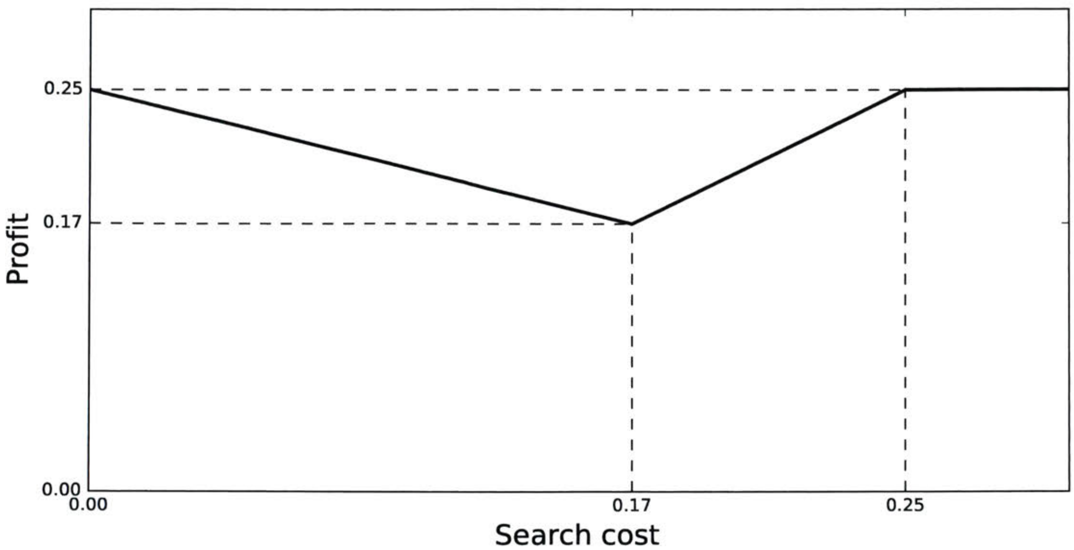
Note: The assumption that the consumer can purchase twice, in case the first purchase turns out to be of low quality, is important because it mitigates the effects of Bertrand competition. It means that the firms do not have to set their prices equal to marginal cost (assumed to be zero here) because even the higher-price firm may get some demand.

Note also: The consumer will get a strictly lower equilibrium utility for low but nonzero values of s than she would get if there was no signal at all (or if the cost of obtaining that signal was high). Essentially, in this case her search cost is too low for the firms to set low prices which ends up hurting her. She would like to commit not to search but cannot do that. We could get around this problem if there were ex-ante contracts specifying prices and a purchase decision because then the consumer could commit to buying even with low search costs, which would allow the firms to charge intermediate prices – benefiting all parties.

Final note: As can be seen in Figure 2-7, it is more important for the consumer to have a high-enough search cost when her prior is around 0.6 because there she is at her most willing to search

[14]The derivative of the expected price charged by a single firm is decreasing in the search cost, which can be verified by computing $\mathbb{E}[p] = \frac{(1-\pi_0)(\pi_0-s)}{\pi_0}\left[\log(\frac{s}{(1-\pi_0)^2(\pi_0-s)}) + \frac{1}{\pi_0}\log(\frac{(1-\pi_0)(\pi_0-s)}{s})\right]$, and differentiating it with respect to s.

Figure 2-5: Equilibrium profit for one firm as a function of s, fixing $\pi_0 = \frac{1}{2}$



for information, and the only way to commit not to search is to have a high search cost. On the other hand, the range of search costs for which the consumer's utility is increasing (values of search cost between the minimizer and maximizer) is at its widest when the prior is uncertain. We can also see that the minimizer and maximizer of utility are very close when the prior is small, which is due to the fact that the utility function is essentially flat for a low prior.[15]

Corollary 2.1. *With no production cost, the firms are worst off when the consumer is best off and their profits are maximized whenever there is no search cost or the search cost is high enough.*

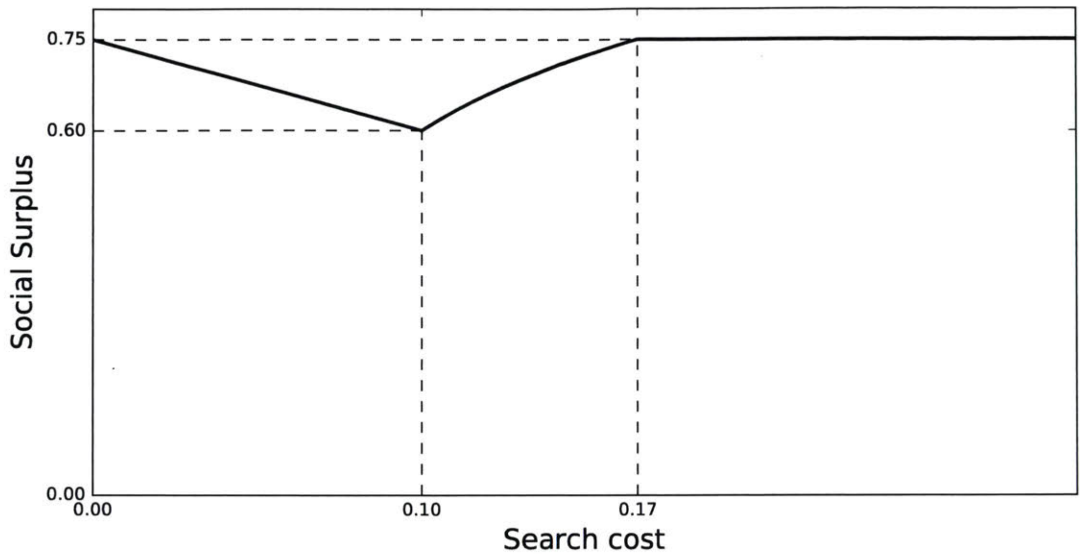
We can see Corollary 2.1 in Figure 2-5 where the firms' profits are minimized at the search cost the consumer prefers. On the other hand, Figure 2-6 reveals that social surplus is at its maximum for no search costs or for high enough search costs while it is lower for intermediate costs, achieving a minimum at the same point where the consumer is worst off.

2.3.1 Comparative Statics for the Prior

In this section we will consider what happens in the model when the common prior changes. We can think of the prior as the expectation a consumer has for the match quality of the products she is shown on a platform. This setup allows different people to have differing priors for goods as long as the goods that the platform shows them have an equal prior. In other words, the consumers and the firms know that the platform's search algorithm produces a set of products for which the prior

[15]The difference between U_{min} and $U(s=0)$ is $\frac{(1-\pi_0)^2}{(1-\pi_0)^2+1}\pi_0^2 < \pi_0^2$, which is really small when π_0 is small. On the other hand, the difference between U_{max} and $U(s=0)$ is $\frac{2(1-\pi_0)\pi_0}{2-\pi_0} < \pi_0$, which is also small.

Figure 2-6: Equilibrium social surplus as a function of s, fixing $\pi_0 = \frac{1}{2}$



quality is π_0.[16] This also introduces randomness into which firms and customers a given firm will face (in which case the firm can only assume a distribution of competitor prices but cannot observe them).

The rest of this section explores what happens to each of 1) firm profits, 2) consumer surplus, and 3) social surplus as we change the prior, and how this depends on the search cost. Proposition 2.1 below shows that increasing the prior may or may not be good for the firms, depending on both the prior and the search cost.

Proposition 2.1. *Given search cost s, the profit-maximizing prior is*

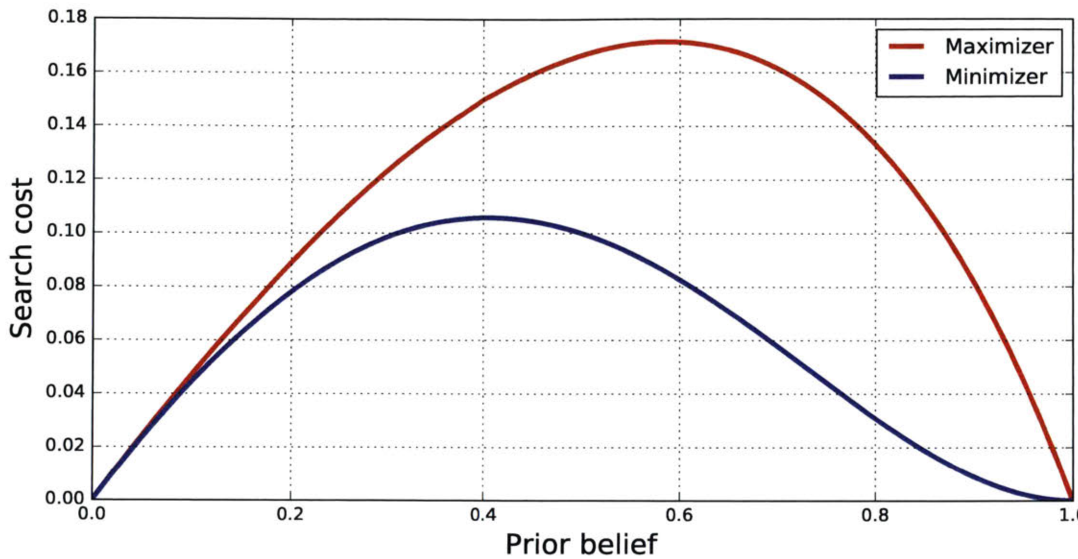
$$\pi_0^* = \begin{cases} \frac{1+s}{2}, & \text{if} \quad s \leq 3 - 2\sqrt{2} \\ \frac{1}{2}, & \text{if} \quad s > 3 - 2\sqrt{2}. \end{cases} \tag{2.1}$$

Figure 2-8 shows how Proposition 2.1 works. Consider first the case of $s = 0$. When we increase the prior from 0 to $\frac{1}{2}$, profit increases from 0 to $\frac{1}{4}$, after which it decreases again. Profits at priors π_0 and $(1 - \pi_0)$ are always equal with no search cost (see the case of 0.2 and 0.8 in the figure). However, when we increase the search cost, profit at any prior decreases linearly until we reach the (global) minimum profit for that prior.[17] As we show in the proof of Proposition 2.1, the global minimizer $s^*(\pi_0)$ increases in the prior until $\pi_0 = 2 - \sqrt{2} \approx 0.59$, at which point it starts to decrease again. We see that each of the priors between $\frac{1}{2}$ and $2 - \sqrt{2}$ is firm-optimal for some search cost,

[16] The algorithm may be based on prior purchases and reviews, or other kinds of consumer-specific variables that are observable to the platform.

[17] Achieved at $s^*(\pi_0)$, where $s^*(\pi_0)$ is the consumer-optimal search cost for prior π_0.

Figure 2-7: Search costs that maximize and minimize consumer utility as a function of the prior belief



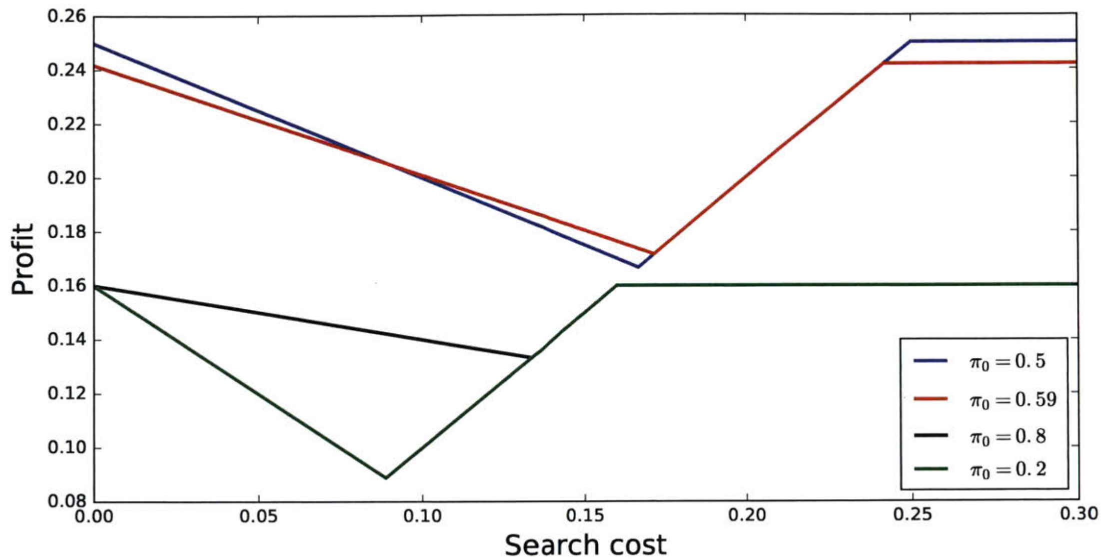
even though profits for these priors are lower than those for $\pi_0 = \frac{1}{2}$ at $s = 0$ (in the figure this manifests in that profits at prior 0.59 exceed those at 0.5 whenever s is in the intermediate range). This is due to the fact that profits decline at a slower rate for higher priors (see Lemma 2.3.7).

Proof. First, notice that $s^*(\pi_0) \equiv \frac{(1-\pi_0)\pi_0}{2-\pi_0}$ is maximized at $\pi_0 = 2-\sqrt{2}$. Note also that s^* is strictly increasing below the maximum and decreasing above it. Plugging this maximum value in, we obtain $s^*(2-\sqrt{2}) = 3-2\sqrt{2}$, which is the highest possible search cost (for any prior) that minimizes firm profits (and maximizes consumer utility). For priors higher than $2-\sqrt{2}$, the minimizer of profits, s^*, is lower than $3-2\sqrt{2}$, meaning that these priors will give profits lower than $\pi_0 = 2-\sqrt{2}$ at any search cost. On the other hand, any prior $\pi_0' < \frac{1}{2}$ is dominated by $\pi_0 = \frac{1}{2}$ because profits for π_0 are higher at $s = 0$, and decline at a slower rate.

Finally, notice that the prior that maximizes profits when $s \leq 3-2\sqrt{2}$ can be obtained by differentiating the profit function in the decreasing range. The maximizer is the solution to $1+s-2\pi_0^* = 0 \Leftrightarrow \pi_0^* = \frac{1+s}{2}$. □

Proposition 2.1 implies that, if a platform's main source of revenue is getting a cut of firm revenues (equal to profit with no production cost), then it may not optimal for the platform to improve its matching algorithm. That is, it is not in the platform's best interest to present the consumers with products that they have higher expectations for. In fact, the platform never wants a prior that is higher than $2-\sqrt{2} \approx 0.59$ for any value of search cost. The only priors that can be optimal for any search cost are $\pi_0 \in [1/2, 2-\sqrt{2}]$. For higher priors, the firms are hurt because the consumers are very likely to find a match at the first firm they visit, meaning that it is desirable

Figure 2-8: Equilibrium profit as a function of s for different priors.



to have the lower of the two prices – leading to more intense price competition and lower profits. Profits for low priors are low simply due to the fact that the consumers are not willing to pay much ex ante.

The following Proposition shows that consumers may be hurt by an increasing prior, too.

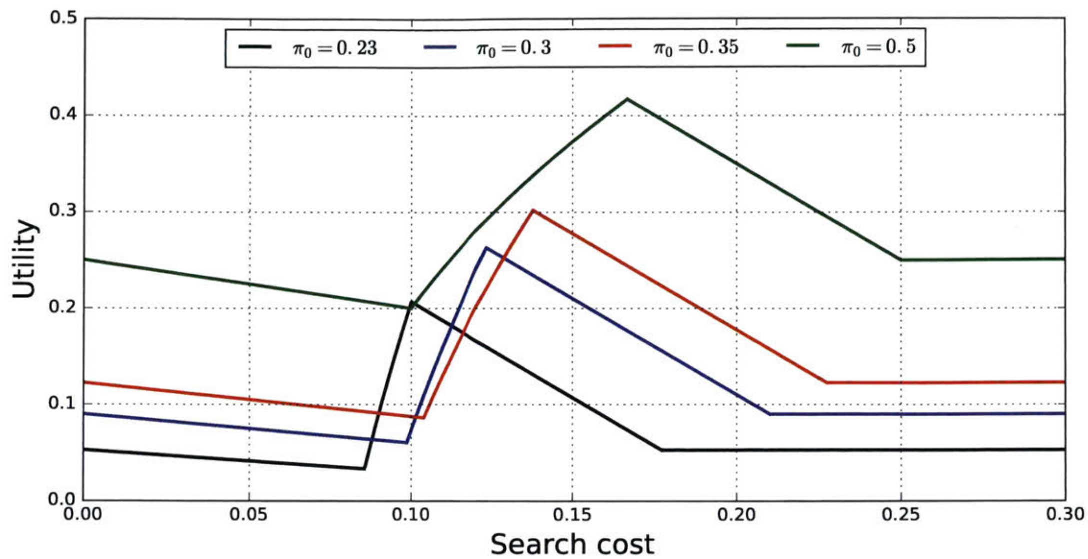
Proposition 2.2. *The consumers' utility is not everywhere monotone in the prior.*

What the Proposition is saying is that there are values of search cost where equilibrium utility is not monotonically increasing in the common prior. This is due to the fact that a higher prior may lead to an increase in $\hat{s}(\pi_0)$, the minimizer of utility, which means that the consumer may have a lower utility for a higher prior since the firms price higher (search equilibrium is played for a wider range of s).[18] A higher prior may also make the increase in utility **slower** in the range where utility is increasing.

Proof. We are not going to show it formally, but the optimal prior for any search cost is $\pi_0 = 1$ because in that case the firms will compete perfectly à la Bertrand, leading to marginal-cost (0) pricing for any search cost. We can show the claim of utility not being monotone in the prior by considering two different priors at a single search cost. To this end, let us take priors $\pi_0 = 0.23$ and $\hat{\pi}_0 = 0.5$, and the search cost $s = \frac{1}{10}$. With these choices, s is in fact the minimizer of utility when the prior is 0.5, while it is the approximate maximizer for the prior 0.23. We find

[18]The minimizer $\hat{s}$ will increase in the prior if and only if the prior is less than $1 - \frac{1}{\sqrt[3]{\sqrt{2}-1}} + \sqrt[3]{\sqrt{2}-1} \approx 0.404$, which is obtained just by differentiating the expression for $\hat{s}$ with respect to the prior, and noticing that it has a unique maximum.

Figure 2-9: Equilibrium utility as a function of s for different priors.



that $U(\hat{\pi}_0) = 0.20 < 0.2071 = U(\pi_0)$. Almost everywhere else, the higher prior produces a higher equilibrium utility. □

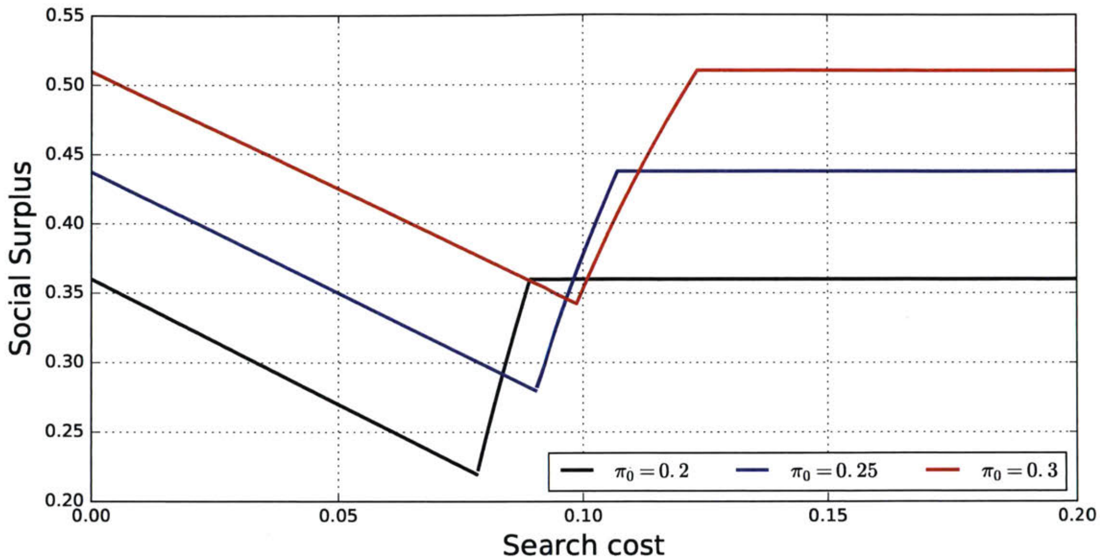
To see the Proposition in action, let us take a look at Figure 2-9 where we plot the consumer's equilibrium utility for different priors. As shown in the above proof, the consumer may indeed prefer a prior of 0.23 over 0.5 if search cost happens to be equal to 0.10. By comparing the utility curves for priors 0.23, 0.3, and 0.35, we also see that increasing the prior will not increase utility monotonically for every search cost, and that the increasing part of utility is steeper for lower priors.

Finally, the third main Proposition of this section shows that even social surplus may be decreasing in the prior expectation.

Proposition 2.3. *Social surplus is not everywhere monotone increasing in the prior.*

This is a stark result. It tells us that it may not be optimal for a social planner to maximize the prior match quality if he cannot guarantee the algorithm will produce a high enough prior (or if search cost cannot be lowered all the way to zero). In other words, improvements in the prior match quality are valuable only if the prior is high enough or the search cost low or high enough. I will skip the proof of the Proposition and refer the reader to Figure 2-10, where we see that intermediate search cost may lead to non-monotonicities in social surplus as a function of the prior.

Figure 2-10: Equilibrium social surplus as a function of s for different priors.



2.4 General Setup: More than two firms

Previously we assumed that there were only two firms but this setting naturally extends to $N+1$ firms for any $N \in \mathbb{N}/\{0\}$, as we will see in this section. We need to make the same assumption as in Section 2.2 but for $N+1$ firms.

Assumption 2.2. *If the consumer chooses to search/buy from firm i, and gets a bad signal (in case of search) or a bad quality draw (in case of purchase), she can still go and search/buy from any other firm. She can do this for as long as she wishes.*

With this assumption, we can show that the consumer-optimal search cost is strictly positive but she prefers no search cost to a low one.

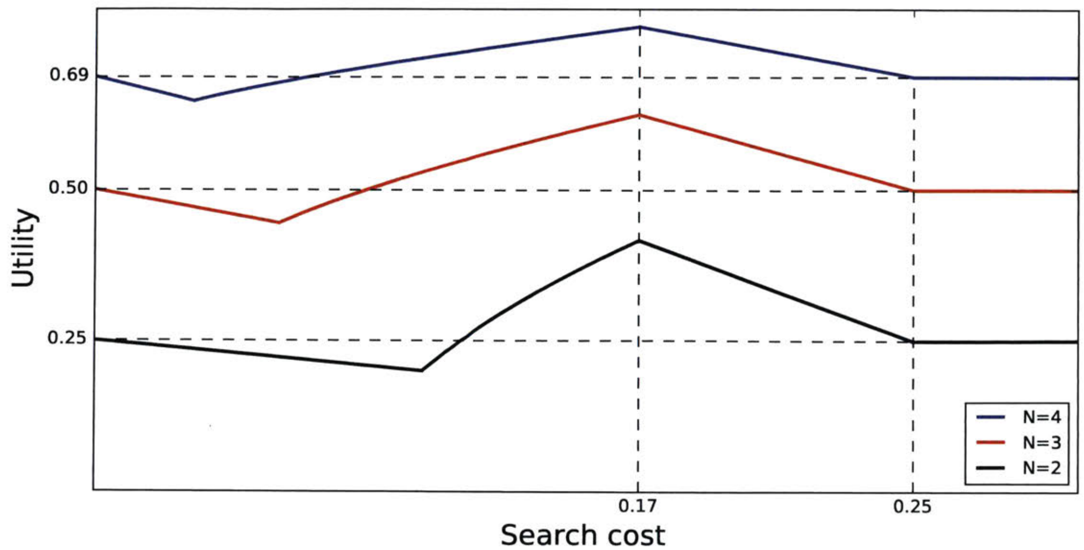
Theorem 2.2. *Assume that $\pi_0 \in (0,1)$ and that there are $N+1$ ex-ante-identical firms Bertrand competing. If Assumption 2.2 holds, the consumer's equilibrium utility is non-monotone in the search cost and achieves its unique global maximum at $s^* := \frac{\pi_0(1-\pi_0)}{2-\pi_0}$. It attains a unique global minimum at $\hat{s} := \frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1} < s^*$.*

Proof. See Appendix B.3. □

In particular, as in Section 2.2, the consumer is first hurt by increasing search cost. However, the following corollary will tell us that the "range of hurt" is shrinking in the number of firms and vanishes in the limit.

Corollary 2.2. *The range of s, for which the consumer always searches in equilibrium, shrinks as the number of firms grows. When N grows without bound, the search equilibrium can be played*

Figure 2-11: Equilibrium utility as a function of s and the number of firms, fixing $\pi_0 = \frac{1}{2}$



only for $s = 0$.[19] *On the other hand, as N grows, the equilibrium where some firms make the consumer search and others do not is played for a larger range of s. Equilibria for other values of s are independent of N.*

Corollary 2.3. *The consumer's utility is always maximized at $s^* := \frac{\pi_0(1-\pi_0)}{2-\pi_0}$, regardless of the number of firms. However, the minimizer, $\hat{s}$, is strictly decreasing in N.*

Corollary 2.4. *Equilibrium utility is strictly increasing in the number of firms for all s. As the number of firms grows without bound, utility approaches 1 for all s.*

This last corollary is in contrast with the result of Stahl (1996) where monopoly pricing remains as an equilibrium even when the number of firms increases without bound, unless there is an atom of shoppers. The difference, of course, is that there search is for prices, and the goods are even ex-post homogeneous, while here we merely have homogeneity ex-ante.[20]

Figure 2-11 above depicts the consumer's utility as a function of her search cost for different numbers of firms. It can be immediately said that utility is increasing in the number of firms and that utility becomes flatter as N increases. This is due to the fact that the minimizer decreases and and both the minimized and maximized utility get closer to the no-search utility.

[19] Although it does not matter which equilibrium the firms play in this case since the consumer always gets a surplus of 1.

[20] If goods in my model were homogeneous ex post, and the consumers knew this, they would only need to get one quality draw to learn every firm's quality.

2.5 Effect of Production Costs

So far, we have assumed that there is no cost of production. However, the production of physical goods such as sneakers often involves significant costs. The question then is, whether or not the main results change if we introduce a non-negative production cost, c, into the model. The short answer is 'no', as long as the cost of production is low enough relative to the prior. If c is high, it is always even consumer optimal to get rid of all search costs.

Proposition 2.4. *Assume that there are two firms with production cost $c \geq 0$. If $c \leq \pi_0$, there exist two thresholds, $0 \leq \underline{s} \leq \bar{s} \leq \pi_0(1-\pi_0)$, so that, for each s, the unique equilibrium of the pricing/search game is:*

1. *The consumer chooses the lower-price firm first. For each firm, when facing price p, the consumer purchases without search if $p \leq \min\{\pi_0, p_B\}$, she searches if $p_B < p \leq p_{max}$, and she exits otherwise, where $p_B \equiv \frac{s}{1-\pi_0}$ and $p_{max} \equiv 1 - \frac{s}{\pi_0}$.*

2. *Each firm uses a symmetric mixed strategy which depends on the value of the search cost as follows:*

   (a) *If $s \leq \underline{s}$, the firm mixes over $p \in [p_S, p_{max}]$, using $F(p) = \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0 - s - \pi_0 c)}{\pi_0^2 (p-c)}$, where $p_S \equiv (1-\pi_0)p_{max} + \pi_0 c$.*

   (b) *If $\underline{s} < s < \bar{s}$, the firm mixes over $[p_L, p_B] \cup [p_H, p_{max}]$, using*

   $$F(p) = \begin{cases} \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0 - s - \pi_0 c)}{\pi_0 (p-c)}, & \text{if} \quad p \in [p_L, p_B] \\ \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0 - s - \pi_0 c)}{\pi_0^2 (p-c)}, & \text{if} \quad p \in [p_H, p_{max}], \end{cases}$$

   where $p_L \equiv (1-\pi_0)\pi_0(p_{max} - c)$ and $p_H \equiv \frac{p_B}{\pi_0} - \frac{1-\pi_0}{\pi_0} c$.

   (c) *If $\bar{s} \leq s \leq \pi_0(1-\pi_0)$, the firm mixes over $p \in [s + \pi_0 c, p_B]$, using $F(p) = \frac{1}{\pi_0} - \frac{s - (1-\pi_0)c}{\pi_0 (p-c)}$.*

   (d) *If $s > \pi_0(1-\pi_0)$, the firm mixes over $p \in [\pi_0(1-\pi_0) + \pi_0 c, \pi_0]$, using $F(p) = \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0 - c)}{\pi_0 (p-c)}$.*

If $c > \pi_0$, the consumer's behavior is the same, while for the firms there are two cases, depending on the value of the search cost: 1) if $s < \pi_0(1-c)$, each firm mixes over $p \in [p_S, p_{max}]$, using $F(p) = \frac{1}{\pi_0} - \frac{(1-\pi_0)(\pi_0 - s - \pi_0 c)}{\pi_0^2 (p-c)}$, and 2) if $s \geq \pi_0(1-c)$, each firm will price at c (which will not be accepted by the consumer).

Proof. Let first $c \leq \pi_0$. The form of the equilibrium is exactly the same as in the case with no production cost, and the proof is also identical, which is why it is omitted. The only difference is that each firm's profit function now includes the production cost. Therefore the formulas for the mixing ranges and probabilities differ from the ones obtained earlier. However, setting $c = 0$, one obtains the simpler formulas we derived in the main section. The thresholds for search costs are:

$$\underline{s} = \frac{(1-\pi_0)^2\pi_0 + (1-\pi_0)[1-\pi_0(1-\pi_0)]c}{(1-\pi_0)^2+1}$$
$$\bar{s} = \frac{(1-\pi_0)[\pi_0 + (1-\pi_0)c]}{2-\pi_0}.$$

These two equations can be obtained as before by making sure that the firms setting extreme prices want to do so instead of deviating somewhere else. When $s \leq \underline{s}$, the firm setting p_S will not want to deviate to p_B, and when $s \geq \bar{s}$, the firm setting p_B will not want to deviate to p_{max}.

However, if $c > \pi_0$, the thresholds, $\underline{s}$ and $\bar{s}$, do not exist. It turns out that search is the only option because the consumer is not willing to buy the good at a price that exceeds π_0 unless she searches first. On the other hand, the firms are not willing to sell at a price lower than the production cost. Therefore, every equilibrium has to be a pure search equilibrium. The form of this search equilibrium is the same as when $c \leq \pi_0$ and $s \leq \underline{s}$, but now we need $s < \pi_0(1-c)$ instead, because otherwise $c \geq p_S \geq p_{max}$, which does not work (the firms would make no profit and the consumer would get no utility, so that there would be no exchange). □

We can see that the two thresholds for search costs ($\underline{s}$ and $\bar{s}$) are increasing in the production cost and they get closer to each other and π_0 as c increases. This is formalized in the following corollary and can be seen in Figure 2-12 below.

Corollary 2.5. $0 < \underline{s} < \bar{s} < \pi_0(1-\pi_0)$ *for all* $c < \pi_0$. *When* $c = \pi_0$, $\underline{s} = \bar{s} = \pi_0(1-\pi_0)$. *In fact, both* $\underline{s}$ *and* $\bar{s}$ *are strictly increasing in* c *but* $\underline{s}$ *increases faster.*

Proposition 2.5. *There is a threshold,* $\bar{c}$, *such that having* $s = 0$ *is consumer-optimal if and only if* $c \geq \bar{c}$.

Proof. As in the main sections, we can find the consumer's equilibrium utility by computing the social surplus and deducting the firms' combined profits. For $s \leq \underline{s}$, social surplus is $W_{low} = (2-\pi_0)[\pi_0(1-c)-s]$ and combined profits $\Pi_{low} = 2(1-\pi_0)[\pi_0(1-c)-s]$, so that the consumer's utility is $U_{low} = \pi_0^2(1-c) - \pi_0 s$. For $s \in (\underline{s}, \bar{s})$, social surplus is $W_{med} = F(p_B)^2[(2-\pi_0)(\pi_0 - c)] + 2F(p_B)[1-F(p_B)][\pi_0 - c + (1-\pi_0)[\pi_0(1-c)-s]] + [1-F(p_B)]^2[(2-\pi_0)(\pi_0(1-c)-s)]$, where $F(p_B)$ is the probability that a given firm charges a price low enough for the consumer not to search. This gives three cases depending on whether zero, one, or two firms set a low-enough price. The actual price does not matter (other than directing consumer behavior) as it is just a transfer from a social point of view. The firms get the same profit as for low search costs, $\Pi_{med} = \Pi_{low}$, because the mixing range includes p_{max} in both cases.

When $s \in [\bar{s}, \pi_0(1-\pi_0)]$, the consumer always purchases, so social surplus is $W_{high} = (2-\pi_0)(\pi_0 - c)$, while the firms get a total of $\Pi_{high} = 2[s-(1-\pi_0)c]$, which leaves the consumer with $U_{high} = (2-\pi_0)\pi_0 - \pi_0 c - 2s$. Finally, if $s > \pi_0(1-\pi_0)$, social surplus is the same as above, $W_{vh} = W_{high}$, the firms get $\Pi_{vh} = 2(1-\pi_0)(\pi_0 - c)$, and the consumer then gets $U_{vh} = \pi_0(\pi_0 - c)$.

To summarize, the consumer's utility is:

$$U(s,c) = \begin{cases} \pi_0^2(1-c) - \pi_0 s, & \text{if } \ s \leq \underline{s} \\ W_{med} - \Pi_{med}, & \text{if } \ \underline{s} < s < \bar{s} \\ (2-\pi_0)\pi_0 - \pi_0 c - 2s, & \text{if } \ \bar{s} \leq s \leq \pi_0(1-\pi_0) \\ \pi_0(\pi_0 - c), & \text{if } \ s > \pi_0(1-\pi_0). \end{cases}$$

As in the case with $c = 0$, the consumer's equilibrium utility is first decreasing, then increasing, and then decreasing again.[21] This means that the global maximum can be found either at $s = 0$ or at $s = \bar{s}$. When we compare $U(0, c)$ to $U(\bar{s}, c)$, we get:

$$U(0,c) \geq U(\bar{s}(c), c) \Leftrightarrow \pi_0^2(1-c) \geq \frac{(1-\pi_0)^2 + 1}{2 - \pi_0}(\pi_0 - c).$$

Noting that this inequality is easier to satisfy for a larger c, we can define $\bar{c}$ as the unique solution to $U(0, \bar{c}) = U(\bar{s}(\bar{c}), \bar{c})$. For $c \geq \bar{c}$, the consumer is better off having no search cost than having any strictly positive cost of search. For $c < \bar{c}$, the consumer's utility is maximized at $\bar{s}$. □

Assuming $\pi_0 = \frac{1}{2}$, the above inequality allows us to solve for $\bar{c} = \frac{2}{7}$. This says that, if the cost of production is greater than $\frac{2}{7}$ (when the prior is $\frac{1}{2}$), the consumer would be best off having no search cost.

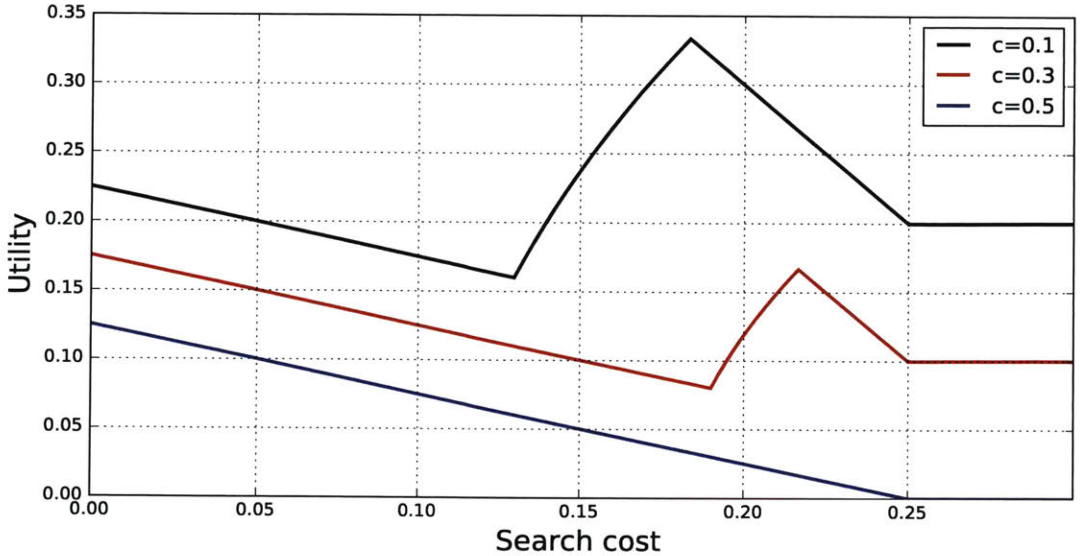
We can see how Propositions 2.4 and 2.5 work in Figure 2-12, where we plot the consumer's equilibrium utility as a function of her search cost for different values of production cost. As is to be expected, the consumer's utility is decreasing in the production cost for all search costs and the "peak" moves right and becomes smaller as the cost of production goes up. At some point ($c = \bar{c} = \frac{2}{7}$), the peak becomes lower than where utility starts off at, so that the consumer is better off not having any search cost at all. When $c \geq \pi_0$, the peak disappears altogether, and the only type of equilibrium (for any s) that remains is the one where the consumer always ends up searching. This is simply due to the fact that purchasing without search is not good for anyone when production costs are high.

2.6 Conclusion

We have shown that an N-firm pricing game with differentiated products and freely observable prices, where a single consumer can search for match quality information and purchase multiple times, leads to interesting equilibrium behavior, which depends non-trivially on the consumer's search cost. When production costs are zero, searching is never socially valuable. However, because the consumer cannot commit to not searching, the firms will set high prices that incentivize search when it is cheap. The consumer's optimal search cost is strictly positive because this guarantees

[21] The intermediate range, $s \in (\underline{s}, \bar{s})$ is algebraically messy, but utility is increasing there since searching is not socially valuable ($c \leq \pi_0$) and increasing s makes it less likely that the consumer ends up searching. The firms' profits are also decreasing in s in this range, so that utility has to be increasing.

Figure 2-12: Equilibrium utility as a function of s for different values of c, fixing $\pi_0 = \frac{1}{2}$



that the she prefers buying to searching while still giving her bargaining power so that the firms cannot set high prices. The consumer's utility is strictly increasing in the number of firms and strictly decreasing in the production cost, for all values of search cost. These results can be used as a first step in designing optimal online platforms, where optimality depends on the revenue-sharing agreements between the platform, the firms, and the consumers.

Chapter 3

Selection and Importance of Consumer Reviews Over Time: Evidence from the Movies

This paper analyzes how big of a factor selection into consumption is in the case of movies. For this purpose, I build two models where utilities are normally distributed and consumers obtain independent signals of their own utility. Those who chose to consume leave a review that perfectly reveals their realized utility. Examining two datasets with movie reviews and box office revenue both in a cross section of movies and within movies over time, we learn that selection is decreasing in the expectation of population mean utility, its precision, and consumer homogeneity. Selection may increase or decrease over time and it tends to increase in the number of movies, although having a close competitor may actually decrease selection. These observations match the predictions of the theoretical model.

I want to thank my advisors Glenn Ellison and Mike Whinston for their feedback and help. I also had the pleasure to discuss the models with Arda Gitmez.

3.1 Introduction

This paper aims to establish a link between underlying qualities of movies and the selection of consumers into consumption, where selection is defined as the difference between the mean utility of the reviewers and the mean population utility. Utilities have two components: a common vertical, and a consumer-specific horizontal component. Reviews perfectly mirror realized utilities but come from consumers who chose to watch the movie, implying that these consumers were more horizontally inclined too see it than an average person. Thus, the reader of reviews has to understand how selected the reviewers are, in order to make the correct inference about the vertical component.

My paper acts as a step toward understanding how consumer reviews evolve over time and how selection depends on certain parameters of the model – most prominent of which include the underlying quality and its distribution, and the distribution of consumer preferences. I will first present two models and the theoretical predictions they offer, after which I will take these predictions to data, verifying that they – for the most part – hold. The data, which come from IMDb and Rotten Tomatoes, include both consumer and critic reviews, and box office revenue. Most importantly, the data are on a daily level, offering insights on the dynamics of reviews. I will abstract away from manipulation and issues regarding which consumers take the time to write reviews. This assumption is bound to introduce some bias into the analysis, but I think it is a reasonable first proxy to assume no manipulation in a large population. It is also not too painful to leave a numerical score (which is the case on movie rating platforms such as IMDb and Rottentomatoes) compared to writing a full review.

The theoretical Section 3.3 builds two models. The first one is a simple model where a continuum of consumers knows the population mean utility of consumption but not where their own preferences are. These utilities are assumed normal (so that everyone knows the mean and variance) and every consumer is assumed to observe an independent and identically distributed signal on their own utility. Having observed the signal, a consumer will consume if and only if her expected utility is positive, which introduces selection into the model. Everyone who has consumed, will leave an honest review on a platform. Selection is defined as the difference between the mean of these reviews and the population mean. There may also be truncation in the reviews due to the limitations of the platform (only allowing reviews to be within a certain range, say), which may make selection seem a problem less severe than it actually is.

The second model takes a slightly different but related approach and assumes the population mean utility to be unknown but normally distributed (with known mean and variance). Here, the simplification is that the consumers are assumed to know their own preferences around the population mean. That is, each consumer knows whether or not they tend to like certain types of movies but they do not know if a particular movie is of high enough mean quality for consumption to be worthwhile to them. These preferences will be referred to as horizontal preferences, which are normally distributed around 0 in the population. A more general model allows the variance of these preferences to depend on the genre or type of movie. Here, selection is the mean horizontal preference of those who consume because the population mean is zero.

The main results of the two models are that selection is decreasing in the population mean quality (or its expectation) and increasing in the variance of consumer utilities. This relates to the oft-cited wisdom that it is better to be loved by some than to be mediocre for everyone. Better signals increase selection. Selection into consumption is also increasing in the variance of the prior quality distribution, or how imprecise the expectation of population mean is. When we consider the dynamics of reviews, sophisticated consumers will understand that reviews are selected and will be able to back out the true mean quality, so that there is perfect learning from period two onwards. However, if some consumers are irrational, or have some behavioral biases, then consumers may make mistakes and there may be learning even after many periods. Selection is increasing in sophistication because rational consumers are better at choosing whether or not to see the movie while unsophisticated movie watchers often make the wrong decision. A final interesting result is that when we consider two or more movies (that are assumed similar in ex-ante terms), selection may counter-intuitively decrease as the number of movies increases. This stems from the fact that if consumers have identical horizontal preferences for two movies and the signals are independent, then it is more likely for consumers with low preferences to obtain high signals (two independent draws), so that the population of consumers will become less selected as a whole. However, if the movies are sufficiently horizontally differentiated (negatively correlated preferences), then selection can increase as there are more competing films.

When we take the model(s) to the data in Section 3.4, we see that many of the predictions hold in the data but that others are not testable. We use both a simple OLS and a fixed-effects model to show that consumer reviews decrease over time but more slowly for movies of high "cinematographic quality" as measured by critic reviews. The most interesting observation is that, having movie fixed effects take care of unobserved qualities, higher box office revenue implies lower consumer reviews (and less selection) because high box office is associated with a less selective population of consumers (if everyone consumed, there would be no selection). Constructing a variable for the variance of utilities for each genre of movies, we learn that consumer reviews are increasing in the variance, which is what was predicted by the theoretical model. However, we do not find evidence that a larger number of movies would mean less selection – in fact, the opposite seems to be the case. This is a more intuitive result because if there are more movies to choose from, every consumer will find the one they love the most, so that selection should increase. The theoretical result is different because it assumes that everything about the different movies is identical to start with and that consumers have the same horizontal preferences all films.

The paper is organized as follows: Section 3.1.1 offers a brief look into the literatures on selection and movies. Section 3.2 introduces the two datasets used in the paper and makes some simple observations, while Section 3.3 introduces and analyzes the two main models. Section 3.4 takes the models to the data to see which predictions hold, and Section 3.5 concludes.

3.1.1 Literature

There is a growing literature on online ratings and selection, and a slightly older literature on movies in general. In this section, I will go over some of the most-closely related papers. Contrary to the earlier work, I focus solely on quantifying selection and how it responds to changes in the underlying quality and preference distributions. However, as a side product, I also obtain evidence for the conclusions made in Li and Hitt (2008) (reviews decrease over time) and Moretti (2011) (word of mouth has an impact of box office revenue and its decline).

Ratings and Selection

Dai, Jin, Lee, and Luca (2016) argue why it is not always optimal to aggregate consumer ratings into a single, simple average because the underlying quality may be changing and consumers may have different horizontal preferences. Therefore, they suggest a different, systematic way of aggregating reviews so that they give more weight to informative and more recent reviews. They do this by estimating a structural model of restaurant reviews in Seattle (taking into account what makes a consumer decide to leave a review and how socially conforming they are), focusing on the decision to review and what kind of reviews to write, while I abstract away from these concerns by assuming that everyone reviews (or that the distribution of reviewers is the same as the distribution of consumers). However, I allow people to select into consumption and analyze what reviews tell us about selection.

Li and Hitt (2008) document how Amazon reviews tend to decrease over time (probably due to the most enthusiastic consumers consuming first), which is similar to what I find during the theatrical run of a movie. They, however, do not estimate how the rate of decline depends on the distributions of quality or consumer preferences. Their setting is more similar to mine than that of Dai, Jin, Lee, and Luca (2016) because they do not allow the underlying quality to change. However, different from Li and Hitt (2008), my setting has identical and fixed prices for all products, so that there is no extraction of consumer surplus – at least in the short run. Contrary to Li and Hitt, Godes and Silva (2012) find evidence that the decrease in consumer ratings is due to mistakes, where high ratings make the wrong consumers want to try the product. This may partly be the story in my model, as well, but I find it more believable that most of the decline in reviews comes from a subpopulation of consumers who make the right decision to consume but are just less enthusiastic about the product than early consumers. However, this is almost impossible to test without experimental data. My theoretical results, however, agree with theirs in that the more reviews there are, and the more heterogeneous the consumers, the less informative the signals (reviews) become, so that consumers make more mistakes and selection decreases.

Movies

Moretti (2011) constructs a Bayesian model of social learning, where consumers observe signals of movie quality and may learn through social interactions (which are unobserved). His main idea is to use the number of opening weekend screens as a proxy for expected popularity/quality because it is a reduced form statistic of how the profit-maximizing sellers see the expected appeal. Using this knowledge, he then constructs a measure of surprise compared to the expected appeal and shows that movies that underperform expectations experience a fast decline in box office revenues, while positive surprise leads to a slower decline. He does multiple robustness checks (different measures of revenue, conditioning for critic reviews and advertising, and using a different measure of surprise). I find the same relationship in my data, although I am not trying to show there is word of mouth. He argues that a movie's competition should not matter for the surprise because the allocated number of screens should already reflect the expected intensity of competition.

Moretti (2011) also tests whether more or less diffuse priors make box office decline faster or slower (where a prior is more diffuse given a genre if the first-weekend surprise has a wider distribution). I borrow his measure of diffuse priors, and show that genres with more variance in critic reviews tend to have higher consumer ratings, which is evidence for selection. Moretti then estimates a model with a quadratic term for decline, obtaining that sales for positive-surprise movies are concave (declining faster over time), and they are convex for negative surprises (declining slower over time) – this is also what I find. The most interesting part of his paper is using national weather shocks (he only has aggregate data) to test whether weather-induced surprise in first-weekend box office generates the same decline patterns in sales, and finds that this is not the case. He then draws the conclusion that network externalities (consumers caring about whether their peers have seen a movie) cannot fully explain the sales patterns, so that word-of-mouth (only caring about peers' seeing a movie because it conveys information) may play a role. I control for word-of-mouth directly using consumer and critic reviews, and find that positive surprise in first-day box office has a larger effect the higher the consumer ratings on IMDb are. Assuming that network externalities do not affect reviews directly (other than through deciding to see a movie), we can see my finding as evidence for people using reviews to decide what movies to see, because positive surprise should already be measuring the unobserved box-office appeal of a movie.

Duan, Gu, and Whinston (2008) investigate how word of mouth impacts movie sales, concluding that ratings seem to make no difference other than through the number of reviews written. I am not focusing on how ratings affect box office but almost the opposite; I offer evidence that higher box office is associated with lower ratings, conditional on quality, because it means that a wider audience, with more heterogeneous preferences, saw the movie. Naturally, in reduced form, good reviews correlate with high box office, but this is in large part due to unobserved quality. Related to Moretti (2011), I show that IMDb reviews have a higher impact for movies with positive surprise, meaning that consumers respond not only to quality but also to higher reviews and word of mouth.

Einav (2007) uses a long panel data set on movies and box office (1985-1999) to estimate a nested logit model and separate the underlying demand seasonality from the quality/appeal of movies. This is a common problem when timing decisions are endogenous, because we cannot know whether high box office was due to high underlying demand or better-quality movies. It is assumed that the decay rate in movie appeal is independent of the release date (conditional on observables, of course). Einav identifies a market expansion effect (substitution between movies and the outside good in the logit model) by comparing the quality of movies released in the same week in different years, while the quality measure is based on movie fixed effects. Einav finds that about one third of the sales variation is due to movie quality/quantity while the rest is due to seasonality. I do not seek to separate seasonality from quality but to show that there is selection in who goes to see what.

Finally, Chiou (2008) builds a model very similar to that of Einav (2007) to show that the seasonality in the home video market implies that Memorial Day and 4th of July are more favorable for theatrical releases than Labor Day (because then the video release date coincides with Christmas season). She uses the box office revenues of similar movies as instruments for a given movie's within-nest share. Similar to Einav, she finds that there seems to be no segmentation in the home video or theatrical market.

3.2 Data and Observations

This section introduces the data used for the paper and makes some reduced-form observations. The third subsection runs some regressions as in Moretti (2011) to show that his results can be replicated, although I also comment on what is different. The last subsection takes a brief look at how variance of consumer reviews changes over time and what this means.

3.2.1 Two Datasets

In this section I will describe the sets of data used for this paper. The first one is an almost nine-month panel dataset of daily observations on movies, their box office, and consumer and critic reviews. The daily national box office figures and statistics were scraped from the Box Office Mojo website for movies that were roughly in the top-30 for each day in the US, and consumer and critic reviews were scraped daily at noon (ET) from IMDb and Rotten Tomatoes from October 9, 2016, through June 27, 2017. This data includes average user ratings both on IMDb and Rotten Tomatoes, critic ratings on Rottentomatoes (both the Tomatometer and numerical ratings), numbers of votes on both sites (for each day), and the actual distribution of IMDb ratings for each day in the observation period. I also have the number of movie theaters playing each movie on any given day and the genres of movies (a single movie may have multiple genres). There are 190 unique titles in the cleaned data, with an average of 38 observations each (I drop titles with no ratings or box office information). This is the dataset I use for most of the analysis because it contains a decent number of titles and observations. For parts of the analysis, I also constrain the movies to be less than 40 days old, because many "bad" movies drop out at that point (if not earlier), which makes it look like ratings are increasing late in the theatrical run. 178 unique titles are observed at some point in these constrained data.

The second dataset extends the first one to cover the years 2010 through 2017 (eight full years), and it includes one observation (total US box office revenue and consumer and critic reviews from Rotten Tomatoes, as of April 2018) per movie (2137 movies). This dataset also includes information on the genres. The obvious limitation is that there is only one, final observation for each film, so that

Table 3.1: Summary statistics for the (short) panel dataset (40 or fewer days since release)

Across movies	Mean	Std	# obs	max	min
IMDb average (1-10)	6.70	1.12	5380	9.4	1.6
-First day	6.74	1.23	123	9.3	1.6
-22nd day	6.74	1.09	118	8.9	3.6
IMDb daily average	6.86	1.23	5162	10.0	0.0
-First day	6.73	1.23	123	9.3	1.6
-22nd day	6.88	1.15	116	10.0	3.0
Tomatometer (0-100)	58.3	27.5	5160	100	0
RT User Average (0.5-5)	3.66	0.55	4963	4.9	1.9
RT Critic Average (1-10)	5.90	1.41	5015	9.0	2.0
Daily Box Office	$1,312,670	$3,771,428	5380	$71,094,394	$5
-First day	$6,289,281	$12,302,150	123	$71,094,394	$436
-22nd day	$938,939	$1,477,032	118	$7,396,325	$722
Theaters	1,467	1,391	5380	4,347	1

Within movies (mean)	Mean	Std	# obs	max	min
IMDb daily average	6.80	0.58	29	8.00	5.53
IMDb daily variance	4.38	2.46	29	9.55	0.00
Daily Box Office	$1,242,053	$1,386,763	30	$5,753,993	$155,689
Theaters	1,352	549	30	1,963	599

we cannot measure the movie-specific variance of reviews or the dynamics of variables. However, the long nature of the data allows us to see far.

The IMDb ratings are consumer reviews on a scale from 1 to 10, while user ratings on Rotten Tomatoes go from 0.5 to 5 (although I often normalize them by multiplying by two). The Tomatometer measures how many percent of the critics have given the movie a "fresh" rating, where fresh means that they thought the movie was at least decent, and worthy of a rating of 3/5 or better.[1] This means that even if a movie has a Tomatometer score of 100 (percent), it may only mean that its quality is 60% of maximum if every critic agrees that it is barely decent. On the other hand, a movie which divides opinions may have a Tomatometer score of 50 if half the critics think it does not deserve to be called decent but the other half think it is the best film in history. Rotten Tomatoes also has numerical critic ratings on a scale from 1 to 10 but it is my impression that people do not really look at these. Table 3.1 above gives some summary statistics both across and within movies. To understand the within-movies table, take the standard deviation of *IMDb daily average* as an example.[2] We first calculate, for each movie, the standard deviation of daily IMDb average over all days it was observed, and then take the mean of these standard deviations.

[1]The website collects critic reviews from different sources and assigns a "fresh" or "rotten" stamp to them, in effect making the reviews binary.

[2]Where daily IMDb average corresponds to the average of reviews on a given day, not the cumulative average. Similarly, daily IMDb variance is the within-day variance of a movie's IMDb reviews.

Its value means that, on average, a single movie's IMDb average has a standard deviation of 0.58 over its theatrical run.

3.2.2 Reduced-form Observations

This subsection takes a look at the datasets, making a few reduced-form observations. First, we see that selection and variance of reviews seem to decrease when quality increases – if we trust critic reviews to be an adequate representation of quality. Second, both variance and selection also decrease over the theatrical run of a movie. Finally, there is a non-monotonic relationship between box office and reviews. All the figures have been relegated to Appendix C.3.

Selection decreases in movie quality

In Figure C.1, we plot the difference between consumer and critic reviews on Rotten Tomatoes using the long (cross-sectional) dataset. The figure shows that selection seems to decrease as we increase the critical quality of a movie, and it does so fairly linearly, which is due to the fact that consumer reviews tend to increase slowly as we increase quality. This is intuitive since critics are somewhat forced to see all kinds of movies but consumers can choose to go or not. Therefore, consumers should be a more selected group and leave higher reviews, on average. However, as we can see, selection is in fact negative for movies with critical averages of 8 and above. Some reasons for negative selection include critics not being a random sample from the population, manipulation (consumer reviews manipulated down, or up), truncation (reviews have to be between 1 and 10), and irrational/mistaken consumers. The critics explanation is the most intuitive because there are some movies that the critics find uniformly bad but many consumers enjoy, meaning that the critics do not accurately represent the general population.

To see how likely it is to observe negative selection, let us take a look at Figure C.2 which runs a Logistic Regression on the probability of observing positive selection as a function of quality (critic average). We see that negative selection is estimated to be more and more likely as quality increases, and we should observe it more than half the time for qualities of 8 and above. The dots represent conditional averages.

Variance of reviews decreases in quality

Figure C.3 depicts how the variance of consumer reviews decreases as the quality of a movie increases. This means that a high-quality movie will get reviews that are closer to their mean than their low-quality counterparts, where quality is again defined as the average critic review. On the other hand, Figure C.4 plots the daily IMDb variance (how variable one single day's consumer reviews are) as a function of that day's IMDb average (more precisely, the decile of all averages in the data). Again, we see a downward-sloping pattern.

Selection decreases over time

Figure C.5 shows that consumer reviews on Rotten Tomatoes tend to decrease slightly as time goes by relative to critic reviews (which stay more or less constant). This means that the consumers, who arrive later in a movie's run, like it less, on average.

Variance of reviews decreases over time

Figure C.6 reveals that IMDb variance decreases during a movie's theatrical run (up to 60 days). This means that the later consumers agree more on the quality of the movie than the first-period consumers (which might be due to manipulation although this is not considered in the present paper).

Selection and Reviews Not Monotone in Box Office

Lastly, Figures C.7 and C.8 show how reviews and selection vary by box office decile in the longer dataset. In a way, this is the final picture after the movies are done with their theatrical run (and, for the most part, the home video run as well). We see that reviews are relatively high for movies that do badly in the box office, which might be due to the fact that these movies get a very selected audience. This is true even for the critics, meaning that they may also be selective in some sense (maybe they do not see all the movies). However, we see that critics are more likely to leave lower reviews. For some reason, reviews are at their lowest for the middle deciles of the box office distribution. One reason for this could be that these movies are the ones that are not of high quality but they nonetheless have decent box office appeal, drawing a large audience (who just barely like the movie). Selection is maximized for movies in the eight decile of box office revenue and it is single peaked.

3.2.3 Reviews, Surprise and Selection

In this subsection we will see how the Moretti-type surprise affects box office in the presence of consumer reviews and, more importantly, how it matters for selection over the theatrical run. Surprise, as defined by Moretti (2011), measures how well a movie opens relative to producer/public expectations, and he argues that this measure is not sensitive to advertising spending, genre, production costs, or critic reviews. However, even if the measure of surprise is conditioning on advertising and such, it may not capture the effectiveness of ads; you can pour millions of dollars down the drain if consumers do not follow the ads. One way to see if his measure of surprise is good for detecting word of mouth is to regress log of box office on days since release, surprise, IMDb rating, and interactions of the three. This is done in Table 3.2 (with mixed effects in intercept).

We see that the effect of surprise on box office revenue is positive only when IMDb ratings are high enough (higher than 5.0, to be exact). Since surprise measures the excess demand on top of what could be expected based on screens and critical reviews, high opening-day revenue is bad if ratings are low because ratings measure quality and high surprise means more people are going to

Table 3.2: Log(box office) on surprise and consumer reviews (136 movies, 5626 obs).

Variable	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	6.770**	0.682	9.931	0.000	5.434	8.107
Surprise	-1.805*	0.741	-2.436	0.015	-3.257	-0.352
IMDb	1.030**	0.098	10.558	0.000	0.839	1.221
Days	-0.221**	0.006	-37.471	0.000	-0.233	-0.209
Days * Surprise	-0.067**	0.008	-8.919	0.000	-0.082	-0.052
Days * Surprise^2	-0.015**	0.001	-14.197	0.000	-0.017	-0.013
Days * IMDb	0.021	0.001	24.371	0.000	0.019	0.023
Days * Surprise * IMDb	0.013**	0.001	11.830	0.000	0.011	0.015
Surprise * IMDb	0.360**	0.104	3.481	0.001	0.157	0.563
Intercept RE	4.757	0.790				

** significant at 5%, ** is at 1%. IMDb reviews are the reviews posted around noon each day (day of revenue).*

hear about the low quality (through official or unofficial channels). The effect of *Days* is negative, as should be expected, and it is even more negative for movies with positive surprise, as long as the user ratings are low. In other words, this means that low-quality movies decline faster if their opening box office is unexpectedly high, which could be due to word of mouth. More specifically, the higher is the opening-day revenue, the more people will hear about the movie's quality, and the more the user rating matters. High quality (IMDb ratings) leads to slower decline in revenues, an effect amplified by surprise because high ratings only matter if people are talking about the film.

Table 3.3 shows the regression of box office revenue on days since release (*Days*) and an indicator for whether the surprise was positive (*Positive Surprise*). It also includes dummies for the weekend and different months. We see that revenue decreases as the movie gets older but the shape of decline depends on whether or not the movie experienced positive surprise on opening day. Movies with positive surprise decline slower but the pattern is concave while the decline for negative surprise is convex.[3] Thus, the effect of a positive surprise dies out over time.

When we add IMDb reviews in Table 3.4, we see that the coefficient on *Days* becomes more negative (from -0.141 to -0.225). This is due to the fact that the rate of decline is negatively correlated with ratings – better-quality films decline more slowly (coefficient on *Days * IMDb* is 0.014). We also see that positive surprise makes IMDb ratings more relevant; that is, the effect of high user ratings is bigger when surprise is positive. In fact, the effect of positive surprise is negative at first unless IMDb ratings are high enough, supporting the theory that an unexpectedly good opening creates more buzz which leads more people to consider the movie.

Now we know that "surprisingly" good first-day box office leads to a slower decline in revenue but what does this mean for selection? If regress selection (average user review minus average critic review of RT) on days since release, indicator of positive surprise, IMDb reviews, and movie and

[3] A t-test for the sum of *Days*^{2} and *Days*^{2} ** Positive Surprise* reveals that the sum is negative (-0.0001), with a p-value of 0.034.

Table 3.3: Effect of surprise on box office without reviews (136 movies, 5626 obs).

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept (January)	13.626	0.216	62.959	0.000	13.202	14.050
Days	-0.141	0.005	-28.400	0.000	-0.151	-0.132
Days^2	0.001	0.000	9.525	0.000	0.001	0.001
Friday	0.666	0.038	17.661	0.000	0.592	0.740
Saturday	1.124	0.035	32.266	0.000	1.056	1.192
Sunday	0.783	0.035	22.552	0.000	0.715	0.851
Days * Positive Surprise	0.082	0.006	13.094	0.000	0.069	0.094
Days^2 * Positive Surprise	-0.001	0.000	-8.919	0.000	-0.001	-0.001
Positive Surprise	1.364	0.286	4.766	0.000	0.803	1.925

Table 3.4: Effect of surprise on box office with reviews (136 movies, 5626 obs).

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	8.260	0.853	9.682	0.000	6.588	9.932
Days	-0.225	0.009	-26.316	0.000	-0.242	-0.208
Days^2	0.001	0.000	8.521	0.000	0.000	0.001
Friday	0.652	0.033	19.591	0.000	0.587	0.717
Saturday	1.126	0.031	36.617	0.000	1.066	1.187
Sunday	0.794	0.031	25.907	0.000	0.734	0.854
Days * Positive Surprise	-0.028	0.011	-2.520	0.012	-0.049	-0.006
Days^2 * Positive Surprise	-0.001	0.000	-9.483	0.000	-0.001	-0.001
Positive Surprise	-6.020	1.123	-5.359	0.000	-8.221	-3.818
Days * IMDb	0.014	0.001	11.927	0.000	0.012	0.017
IMDb	0.754	0.122	6.170	0.000	0.515	0.994
Days * Positive Surprise * IMDb	0.015	0.001	9.978	0.000	0.012	0.018
Positive Surprise * IMDb	1.034	0.156	6.630	0.000	0.728	1.339

Table 3.5: Effect of surprise on selection (102 movies, 3377 obs). Movie and month FE.

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	-1.043	0.220	-4.735	0.000	-1.475	-0.611
Days	-0.013	0.002	-7.195	0.000	-0.017	-0.010
Positive Surprise	-0.027	0.256	-0.107	0.915	-0.529	0.474
Days * Positive Surprise	-0.006	0.001	-10.055	0.000	-0.007	-0.004
Days * IMDb	0.001	0.000	5.028	0.000	0.001	0.002
IMDb	0.356	0.018	19.468	0.000	0.320	0.391

month fixed effects, we get the results shown in Table 3.5. It is not surprising that selection is decreasing over time as, intuitively, later consumers are less and less enthusiastic about a movie. Movie fixed effects should control for the unobserved difference between consumers and critics that stays constant over the theatrical run. Therefore, the coefficient on *IMDb* can be understood as the effect of making movies that better fit consumer preferences, holding critic preferences constant. We also see that a higher IMDb user rating implies a slower decline in selection because it means that the movie is a better fit to the consumers, in general. Finally, and most interestingly, movies with positive first-day surprise experience a faster decline in selection as time goes by. This is because movies that do well (and better than expected) in the opening weekend will get hyped more, creating more word of mouth, and therefore attracting a wider audience. A wider audience, however, is also associated with lower selection because it includes people that were not excited about the movie to start with but became interested due to the "buzz".

3.2.4 Variance of Reviews

Let us now take a look at Figure C.6. In the figure, we plot the within-movie variance of IMDb reviews conditional on how many days the movie has been playing at the theaters. Cumulative variance is the variance of all ratings so far, while daily variance is the variance of today's (new) ratings. We see that first-weekend (especially first-day) variance is much higher than that of following days and weeks. Because we also get most reviews during the opening weekend, the cumulative variance decreases slower and is less volatile than its daily counterpart. This behavior suggests that the distribution of people (in a horizontal-preference space) is wider during the opening weekend than in the following weeks, producing a more dispersed set of initial reviews. The people who see a movie after the opening weekend seem to have a more precise prior on their match quality with the movie, producing a more uniform set of reviews. This is already different from Moretti (2011) because he (essentially) assumes away learning about the horizontal match quality.[4]

[4]People get unbiased signals of their own utility, which naturally contains the horizontal component, but this is all. There is no match learning based on other people's experiences since they are just indicative of the common quality component.

3.3 Models for Selection

In this section, I will show how to model selection, particularly in the context of movies. We will first construct a simple model which highlights the main features of selection into consumption, assuming that everyone knows the mean utility in the population. Using the simple model as a springboard, we will then construct a more complicated model to add dynamics and explain how consumers learn from earlier reviews. This extended model allows the population mean to be unknown. However, even if the population mean is unknown in the first period, reviews perfectly reveal the underlying mean quality in the second period – assuming that consumers are sophisticated. This offers a testable prediction on whether consumers are rational.

3.3.1 Basic Model: Signals of Utility with Known Mean

Consider the following model to understand how selection works when consumption utility is normally distributed and unknown to consumers, so that everyone knows the population distribution but not their own realization. Let $u \sim N(q, \frac{1}{\tau_u})$ be the normally distributed consumption utility in the consumer population for a given movie, where q is the **known** expected quality of the movie and τ_u is the inverse variance of the consumption utilities. Thus, a high τ_u leads to utilities that are concentrated around q (more homogeneous consumers). Assuming that movie critics are a reasonably random sample from the underlying population, we can see q as the mean critic review (observed by the consumers), but more generally we can think of q as a function of the mean critic review (in case critics have preferences different from the general population). Similarly, we can understand τ_u as a function of the variance of these critic reviews, the genre, and other factors. The consumers do not observe (their personal) u but have access to a signal $s = u + \epsilon$, where $\epsilon \sim N(0, \frac{1}{\tau_\epsilon})$ is the error. τ_ϵ can be understood as the precision of the signal.[5]

Let us assume there is only one movie with the above prior and signal distributions. The distribution of signals is therefore:

$$s \sim N\left(q, \frac{\tau_u + \tau_\epsilon}{\tau_u \tau_\epsilon}\right).$$

If a consumer observes a signal s, her posterior will follow the following distribution:

$$u|s \sim N\left(\frac{\tau_u q + \tau_\epsilon s}{\tau_u + \tau_\epsilon}, \frac{1}{\tau_u + \tau_\epsilon}\right).$$

Assume further that any given consumer needs to get positive expected utility for them to consume, so that we need $\mathbb{E}[u|s] \geq 0$.[6] Using the posterior expectation, this is equivalent to saying that the consumer's signal has to be at least $\hat{s} \equiv -\frac{\tau_u}{\tau_\epsilon} q$ for her to consume. Therefore, the ex-ante probability

[5]The precision of the signal depends on the consumers' ability to read critic reviews/earlier consumer reviews correctly. If a critic dislikes all horror films, this does not mean you should follow his recommendation if you have different preferences – although there may still be some value in the review.

[6]This is not a restrictive assumption since we can always shift the distribution of utility so that the threshold is exactly at zero. However, the threshold is assumed to be the same for each consumer.

of seeing the movie can be written as:

$$\mathbb{P}(s \geq \hat{s}) = 1 - \Phi\left((\hat{s} - q)\sqrt{\frac{\tau_u \tau_\epsilon}{\tau_u + \tau_\epsilon}}\right) = 1 - \Phi\left(-q\sqrt{\frac{\tau_u(\tau_u + \tau_\epsilon)}{\tau_\epsilon}}\right) = \Phi\left(q\sqrt{\frac{\tau_u(\tau_u + \tau_\epsilon)}{\tau_\epsilon}}\right),$$

where Φ is the standard normal cdf, and we have used the fact that $s \sim N(q, \frac{\tau_u + \tau_\epsilon}{\tau_u \tau_\epsilon})$. Knowing this, we can also derive the distribution of $\mathbb{E}[u|s]$ (which is a random variable due to the randomness of the signal):

$$\mathbb{E}[u|s] \sim N\left(q, \left(\frac{\tau_\epsilon}{\tau_u + \tau_\epsilon}\right)^2 \frac{\tau_u + \tau_\epsilon}{\tau_u \tau_\epsilon}\right) = N\Bigg(q, \underbrace{\frac{\tau_\epsilon}{\tau_u(\tau_u + \tau_\epsilon)}}_{\equiv \sigma^2}\Bigg).$$

Now, using the law of total expectation (law of iterated expectations) and the fact that $\mathbb{E}[X|X \geq 0] = \mu + \sigma\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}$ for $X \sim N(\mu, \sigma^2)$, where ϕ and Φ are the pdf and cdf of the standard normal distribution, respectively, we can obtain our first result:

Proposition 3.1. *The expected consumer review can be written as:*

$$\mu = \mathbb{E}\Big[u \;\Big|\; \mathbb{E}[u|s] \geq 0\Big] = \mathbb{E}\Big[\mathbb{E}[u|s] \;\Big|\; \mathbb{E}[u|s] \geq 0\Big] = q + \sigma\frac{\phi(q/\sigma)}{\Phi(q/\sigma)},$$

where $\sigma^2 \equiv \frac{\tau_\epsilon}{\tau_u(\tau_u + \tau_\epsilon)}$ *is the variance of* $\mathbb{E}[u|s]$.

The result is simply saying that consumer reviews will always be more positive than the population average (q), and that selection depends on that average and the variance of the conditional expectation (σ^2)

Proof. The proof is simply due to iterated expectations. One can check the result also by writing $\mathbb{E}\Big[u \;\Big|\; \mathbb{E}[u|s] \geq 0\Big] = \frac{1}{\Phi(q/\sigma)} \int_{\hat{s}}^{\infty} \mathbb{E}\Big[\mathbb{E}[u|s] \;\Big|\; \mathbb{E}[u|s] \geq 0\Big] \frac{1}{\sigma}\phi(\frac{s-q}{\sigma}) ds$, using the formula for truncated normals and simplifying. ☐

Now that we have obtained an expression for the expected review, we can do comparative statics to see how reviews depend on the model parameters.

Proposition 3.2. *The average consumer rating, μ, is strictly increasing in the ex-ante expectation (q) and the variance (σ^2) of the ex-post expectation. Selection (the difference between reviews and population expectation) is strictly decreasing in q and strictly increasing in σ^2.*

Proof. First of all, note that for any function $f(x)$, $\frac{\partial \phi(f(x))/\Phi(f(x))}{\partial x} = -f'(x)\frac{\phi(f(x))}{\Phi(f(x))}\left(f(x) + \frac{\phi(f(x))}{\Phi(f(x))}\right)$, because $\phi'(x) = -x\phi(x)$ and $\Phi'(x) = \phi(x)$. Using this, we can write:

$$\frac{\partial \mu}{\partial q} = 1 - \sigma\frac{1}{\sigma}\frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\left[\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right] = 1 - \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\left[\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right] > 0.$$

The inequality follows because, defining $h(x) = \frac{\phi(x)}{\Phi(x)}\left(x + \frac{\phi(x)}{\Phi(x)}\right)$, one can show that $h(x) \in (0,1)$ and $h'(x) < 0$ for all x. Thus, consumer ratings are increasing in the population mean. However, for selection, we have:

$$\frac{\partial(\mu - q)}{\partial q} = \frac{\partial \mu}{\partial q} - 1 < 0,$$

meaning that reviews increase more slowly than q. It is equally easy to show that

$$\frac{\partial(\mu - q)}{\partial \sigma} = \frac{\partial \mu}{\partial \sigma} = \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\left[1 + \frac{q}{\sigma}\left(\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right)\right] > 0,$$

which proves that both selection and consumer ratings are increasing in the variance of $\mathbb{E}[u|s]$. □

Corollary 3.1. *Selection and consumer ratings are strictly increasing in τ_ϵ and strictly decreasing in τ_u.*

This Corollary tells us that if the population is more homogeneous (a larger τ_u implies utilities are more concentrated around q), then reviews are lower and there is less selection for any given q. On the other hand, if the signal is of better quality (a higher τ_ϵ), then consumer reviews are better and there is more selection (holding q fixed). What is the intuition? High signal quality means that only the people, who will enjoy the movie ($u > 0$), go see it, while low signal quality means that there are more people with $u < 0$ who obtain high signals and choose to watch the movie. Conversely, if population utilities are more homogeneous (a high τ_u), signals are going to be more homogeneous, so that decisions are also going to be more similar, implying that consumer ratings are going to be lower (more concentrated around q) and less selective.

Proof. We have that $\frac{\partial \sigma^2}{\partial \tau_\epsilon} = \frac{1}{(\tau_u + \tau_\epsilon)^2} > 0$, and that $\frac{\partial \sigma^2}{\partial \tau_u} = -\frac{2\tau_u + \tau_\epsilon}{\tau_u^2(\tau_u + \tau_\epsilon)^2} < 0$. Therefore, using the previous Proposition, we have that $\frac{\partial(\mu - q)}{\partial \tau_\epsilon} = \frac{\partial \mu}{\partial \tau_\epsilon} > 0$ and $\frac{\partial(\mu - q)}{\partial \tau_u} = \frac{\partial \mu}{\partial \tau_u} < 0$. □

Proposition 3.3. *Consumer surplus increases in q and τ_ϵ, and it decreases in τ_u.*

Understanding that here reviews are the (ex-post) realized consumer utilities, for any given q, increasing τ_ϵ and decreasing τ_u will increase consumer surplus. In other words, whatever the mean quality of a movie, the more it divides opinions and the better the consumers understand their preferences for the movie, the higher is the realized consumer surplus.[7] This is the usual result that it is better to be loved by some and hated by others than to be decent for everyone, as long as the signal is informative ($\tau_\epsilon > 0$), because then the consumers with extremely low utilities will not watch the movie but there are many who have extremely high utilities and end up watching. Naturally, better signal quality implies fewer wrong decisions (false positives or negatives), and therefore higher surplus.

[7] Assuming away endogenous pricing. This assumption is realistic in the short run because all movies have the same price – although that price can vary between theaters and times of day and week.

Proof. The ex-ante expected utility is $U = \mathbb{P}(s \geq \hat{s})\mu = \Phi(q/\sigma)\left(q + \sigma\frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right) = q\Phi(q/\sigma) + \sigma\phi(q/\sigma)$, because the consumers obtain good signals with probability $\mathbb{P}(s \geq \hat{s})$ and the expected utility having obtained a good signal is μ. Therefore, $\partial U/\partial q = \Phi(q/\sigma) > 0$ (two terms cancel out) and $\partial U/\partial \sigma = \phi(q/\sigma) > 0$. Because σ^2 is increasing in τ_ϵ and decreasing in τ_u, we obtain the Proposition. □

One obvious concern with this approach is that reviews are restricted to be between 1 and 10, while I have assumed the underlying distribution of utility to be normal. This may not be as restrictive as it sounds if we assume that all really bad reviews are labeled 1, and good ones 10 (truncation), which would simultaneously be one explanation for why there are often more ones and tens compared to other categories. Naturally, another explanation is that there is manipulation for either behavioral or monetary reasons. The previous point also highlights a key problem with reviews: there is truncation both ex ante and ex post. The former means that only consumers with high signals consume, while the latter forces all reviews to be within some, possibly narrow, range.

Proposition 3.4. $s^2 \equiv var(reviews) = \frac{1}{\tau_u} - \sigma^2\frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\left(\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right) = \frac{1}{\tau_u} - \mu(\mu - q)$, *which is strictly increasing in* q, *with* $\lim_{q\to-\infty} s^2 = \frac{1}{\tau_u+\tau_\epsilon}$, *and* $\lim_{q\to\infty} s^2 = \frac{1}{\tau_u}$. *Moreover, the variance of reviews is strictly decreasing in* τ_ϵ, *with* $\lim_{\tau_\epsilon\to 0} s^2 = \frac{1}{\tau_u}$, *and* $\lim_{\tau_\epsilon\to\infty} s^2 = \frac{1}{\tau_u} - \frac{1}{\tau_u}\frac{\phi(q\sqrt{\tau_u})}{\Phi(q\sqrt{\tau_u})}\left(q\sqrt{\tau_u} + \frac{\phi(q\sqrt{\tau_u})}{\Phi(q\sqrt{\tau_u})}\right)$. *For* τ_u *we have* $\lim_{\tau_u\to 0} s^2 = \infty$ *and* $\lim_{\tau_u\to\infty} s^2 = 0$.

Proof. See Appendix C.1. □

Remember that $\sigma^2 = \frac{\tau_\epsilon}{\tau_u(\tau_u+\tau_\epsilon)}$ is the variance of $\mathbb{E}[u|s]$. That is, it is the variance of the conditional expectation due to the randomness of the signal. The first part of Proposition 3.4 gives us the variance of reviews, taking into account the fact that these reviews come from a selected subpopulation (those consumers who get good enough signals). This variance is always less than the unconditional variance of population utility ($\frac{1}{\tau_u}$), and it is increasing in the expected population utility. It is decreasing in the precision of the population utility (τ_u, "uniformity of preferences") and also decreasing in the precision of the signal (τ_ϵ).

The following Corollary makes a more general mathematical observation that can be obtained as a consequence of the proof of Proposition 3.4.

Corollary 3.2. *More generally, we have that* $var\left(u\middle|\mathbb{E}[u|s] \geq 0\right) = var\left(\mathbb{E}[u|s]\middle|\mathbb{E}[u|s] \geq 0\right) + var\,(u|s)$.

3.3.2 Explaining Negative Selection

With the assumptions we made at the beginning of the previous section, we showed that consumer reviews will always be more selected than critic reviews. In this section, we will briefly consider two cases that may in fact lead to lower consumer than critic reviews.

Truncation

When consumer reviews are truncated, the observed mean review may in fact be lower than the mean utility in the whole population, even though the non-truncated mean of the selected population is always higher than critic reviews (which are a proxy for the population mean).

Proposition 3.5. *If consumer reviews are truncated in some way, the observed average review may be lower than average utility in the population. That is, there may appear to be negative selection due to truncation.*

Proof. To show this, consider a setting where consumer reviews are restricted to be in the interval $[0, 10]$. Thus, if your experienced utility was less than 0, your review would be 0, and correspondingly for high experienced utilities.

Now, we cannot write the observed mean review in closed form but, using simulations, we see that if $q = 9$, $\tau_u = \frac{1}{81}$, and $\tau_\epsilon = \frac{1}{25}$, the observed mean review will be 7.5.

Thus, the observed mean rating is less than the mean utility in the whole population, even though consumers are selected to watch the movie (the true, non-truncated mean being about 10.9). This completes the proof. □

One thing we have yet failed to mention is that we are implicitly assuming that critic reviews are not truncated even though consumer reviews are. This may be a strong assumption but we could obtain the same result if critics were a sample of people more concentrated around the population mean than regular consumers. In this case, truncation would affect the critics less and lead to a smaller bias than for the consumers. On the other hand, as long as consumers know the mean population utility, q, their reviews will be selected but may be lower than q if there is truncation.

Biased Critics

Another way for the consumer reviews to be lower than critic reviews – even with selection – is that the populations are different. What this means is that the average critic review may not be a good proxy for the mean population utility. In fact, the critics may be inclined to review some types of movies more negatively than the consumers and vice versa for other genres. To better understand this notion, consider a model where the population mean is $q = \alpha_0 + \alpha_1 c$, with c being the critic average. In this model, it would be reasonable to expect that $\alpha_0 > 0$ and $\alpha_1 < 1$ so that $q < c$ for high c. In other words, consumers may like "bad" movies more than the critics but fail to see the appeal of some critically acclaimed films. In this setting, even if consumers are selected, they may still leave reviews that are lower than the critics'.

3.3.3 Model: Learning about mean utility

In the previous subsection we assumed that the mean utility in the population was known to be q prior to observing an individual signal. We will now relax that assumption and think of the population mean as a random variable, $x \sim N(q, \frac{1}{\tau_x})$. We will assume that every individual i

observes a private signal $s_i = x + \epsilon_i$, where $\epsilon_i \sim N(0, \frac{1}{\tau_\epsilon})$, meaning that $s_i \sim N(q, \frac{\tau_x+\tau_\epsilon}{\tau_x\tau_\epsilon})$. This signal can be thought of as a result of reading critic reviews which reveal the true mean utility with some precision. Individual i's consumption utility is $u_i = x + v_i$, where $v_i \sim N(0, \frac{1}{\tau_v})$ is what we will call the horizontal component. This component is known to the individual but not to anyone else. Thus, population utilities are normally distributed around a common mean which is unknown and itself normally distributed.

Proposition 3.6. *Under the above assumptions, the expected first-period selection can be written as:*

$$\bar{v}_1(q) = \frac{a}{\tau_v}\frac{\phi(aq)}{\Phi(aq)}, \tag{3.1}$$

where $a = \sqrt{\frac{\tau_v}{1+\sigma^2\tau_v}}$ *and* $\sigma^2 = \frac{\tau_\epsilon}{\tau_x(\tau_x+\tau_\epsilon)}$. *Therefore, the mean review is* $\mu_1 = x^* + \bar{v}_1(q)$, *where* x^* *is the true mean quality.*

Proposition 3.6 offers us a closed-form solution to the question "How much selection is there?", as a function of the four parameters. If there was no selection, the mean horizontal preference would be exactly zero, but here it is always strictly positive. This is regardless of the true mean quality in the population, x^*, because first-period consumers make their consumption decisions acting on the prior alone. Here, σ^2 is the variance of $\mathbb{E}[x|s_i]$, and it corresponds to the variance of the ex-post population mean.[8]

Proof. The assumptions imply that, observing a signal s_i, individual i's posterior belief on the population mean is distributed as $x|s_i \sim N(\frac{\tau_x q+\tau_\epsilon s_i}{\tau_x+\tau_\epsilon}, \frac{1}{\tau_x+\tau_\epsilon})$. Now, knowing her horizontal preference for this particular movie, individual i will watch it if and only if $v_i + \frac{\tau_x q+\tau_\epsilon s_i}{\tau_x+\tau_\epsilon} \geq 0 \Leftrightarrow s_i \geq -\frac{\tau_x}{\tau_\epsilon}q - \frac{\tau_x+\tau_\epsilon}{\tau_\epsilon}v_i$ (due to the same assumption of positive expected utility as before). Thus, the ex-ante probability that someone with v_i will consume is $P(v_i) = 1 - \Phi\left(-\frac{v_i+q}{\sigma}\right) = \Phi(\frac{v_i+q}{\sigma})$, where $\sigma^2 \equiv \frac{\tau_\epsilon}{\tau_x(\tau_x+\tau_\epsilon)}$ is the variance of $\mathbb{E}[x|s_i]$ and Φ is the standard normal cdf.[9] Once the consumer has watched the movie, she will leave a review equal to the realized utility: $r_i = x^* + v_i$, where x^* is the true mean utility. For future reference, this means that we can back out the mean utility if we know the horizontal preferences of the reviewers.[10]

Now, because we know the (ex-interim) probability of consuming conditional on v_i, we can

[8]Note that we this is **almost** the same parameter as the σ^2 we defined in the previous subsection; only that here τ_u is replaced by τ_x.

[9]This is due to the fact that s_i is normally distributed.

[10]Reviews accurately depicting consumer utilities is still a strong assumption due to the possible existence of a myriad of behavioral biases.

calculate the ex-ante expected probability of consumption (without knowing v_i).

$$\begin{aligned} P &= \int_{-\infty}^{\infty} P(v_i) f(v_i) dv_i \\ &= \int_{-\infty}^{\infty} \Phi(\frac{q+v_i}{\sigma}) \sqrt{\tau_v} \phi(\sqrt{\tau_v} v_i) dv_i \\ &= \int_{-\infty}^{\infty} \Phi(\alpha + \beta Y) \phi(Y) dY \\ &= \Phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right) \\ &= \Phi\left(\frac{\sqrt{\tau_v} q}{\sqrt{1+\tau_v \sigma^2}}\right) \equiv \Phi(aq), \end{aligned} \tag{3.2}$$

with $a = \sqrt{\frac{\tau_v}{1+\sigma^2 \tau_v}}$ as already defined in the Proposition. Here we have first used the fact that v_i is normally distributed (with zero mean and precision τ_v), then changed variables using $Y = \sqrt{\tau_v} v_i$ so that $dv_i = dY/\sqrt{\tau_v}$, and $\alpha = q/\sigma$ and $\beta = 1/(\sqrt{\tau_v}\sigma)$. Note that $Y \sim N(0,1)$, which allows us to use a known result for standard normals on the second-to-last line. The last line just follows by writing out α and β and simplifying.

Knowing the probability of consumption, we can calculate the expected value of v_i for those who consume (either due to high signals or high horizontal preferences, or both).

$$\begin{aligned} \bar{v}_1(q) &= \frac{1}{P} \int_{-\infty}^{\infty} v_i \Phi(\frac{q+v_i}{\sigma}) f(v_i) dv_i \\ &= \frac{1}{P} \frac{1}{\sqrt{\tau_v}} \int_{-\infty}^{\infty} Y \Phi(\alpha + \beta Y) \phi(Y) dY \\ &= \frac{1}{P} \frac{1}{\sqrt{\tau_v}} \frac{\beta}{\sqrt{1+\beta^2}} \phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right) \\ &= \frac{a}{\tau_v} \frac{\phi(aq)}{\Phi(aq)}, \end{aligned}$$

where the second line follows from a change of variables and the third from using a known formula for standard normals. The last line involves plugging in P from Equation (3.2), writing out α and β, and simplifying. □

Note that we can understand $a^2 = (\frac{1}{\tau_v} + \sigma^2)^{-1}$ as the ex-ante precision of the expected post-signal consumption utility. That is, a^2 is the inverse variance of $v_i + \frac{\tau_x q + \tau_\epsilon s_i}{\tau_x + \tau_\epsilon}$. Therefore, a high a means that every consumer will make the same decision, depending on the sign of q.

With the expression for selection in hand, we can do comparative statics. Proposition 3.7 below summarizes these results.

Proposition 3.7. *Selection is strictly decreasing in the expected mean of population utilities, q. If $q \geq 0$, it is also decreasing in the precision of horizontal preferences, τ_v. We also have, approxi-*

mately,

$$\frac{\partial \bar{v}_1}{\partial a} < 0 \Leftrightarrow aq > 0.84,$$

so that selection is strictly decreasing in τ_x and strictly increasing in τ_ϵ if and only if $aq > 0.84$.

Proof. Looking at the expression for selection, and remembering that $\phi(y)/\Phi(y)$ is everywhere strictly decreasing in y, it is easy to see that selection is strictly decreasing in q (due to a being strictly positive). Furthermore, $\frac{a}{\tau_v} = \frac{1}{\sqrt{\tau_v(1+\sigma^2\tau_v)}}$ is strictly decreasing in τ_v, while a is strictly increasing in τ_v, which means that $\frac{\phi(aq)}{\Phi(aq)}$ is decreasing in τ_v whenever $q \geq 0$. Combining the two, we see that selection is decreasing in τ_v for non-negative values of q.

We can numerically evaluate the derivative of $\bar{v}_1$ with respect to a at different points to see that, approximately, $\bar{v}_1$ is strictly decreasing in a for $aq > 0.84$ and strictly increasing otherwise. It is easy to see that a is strictly decreasing in σ^2, and that σ^2 is strictly increasing in τ_ϵ and strictly decreasing in τ_x. Combining these observations, we see that a is strictly increasing in τ_x and strictly decreasing in τ_ϵ, so that $\bar{v}_1$ is strictly decreasing in τ_x and strictly increasing in τ_ϵ whenever $aq > 0.84$. If $aq < 0.84$, the opposite is the case. □

Proposition 3.7 says that selection is always decreasing in the expected mean quality, q. This is as intuitive as in the simple model we saw earlier because a higher q implies a larger number of consumers with low v. The result that selection is decreasing in τ_v for positive values of q is also intuitive because a higher τ_v means that the consumer population has more homogeneous preferences for the movie, and that the population utilities are more concentrated around the expected mean, $q \geq 0$. Thus, a higher τ_v leads to more and more people consuming the movie (although some consumers will still withdraw due to bad signals).

When it comes to the other two parameters of the model, τ_x and τ_ϵ, the intuition becomes slightly less clear. To make matters simpler, let us just consider what happens when a increases because a is strictly increasing in τ_x and strictly decreasing in τ_ϵ. In a way, we can think of a as a measure for how much consumers will trust the signal.[11] First of all, note that when $q < 0$, a higher a leads to a lower probability of consumption ($\Phi(aq)$), and it also means that consumers will trust the prior more than the signal, so that fewer consumers get false positives, and therefore the only ones to consume are those with high v, leading to more selection.[12] Similarly, if $q = 0$, increasing a decreases the rate of false positives but has no effect on probability of consumption, leading to more selection.[13] Now, if $q > 0$ but $aq < 0.84$, we have the most interesting case where a higher a leads to both a higher probability of consumption and higher selection. The reason is

[11]We know that, as τ_x goes from 0 to ∞, σ^2 goes from ∞ to 0 so that a goes from 0 to $\sqrt{\tau_v}$. Similarly, τ_ϵ going from 0 to ∞ implies that σ^2 goes from 0 to $1/\tau_u$, so that a goes from $\sqrt{\tau_v}$ to $\sqrt{\frac{\tau_v}{(1+\tau_v/\tau_u)}}$. Thus, a high a is a sign of either a high τ_x or a low τ_ϵ, both of which are associated with more weight on the prior and less on the signal.

[12]Here, *false positives* mean signals that induce consumption even though realized utility turns out to be negative.

[13]The rate of false positives decreases because, holding τ_v fixed, a higher a is equivalent to a lower σ^2. This implies that consumers are more likely to make the right decision, or at least not be wrong by much if they make the wrong decision.

that, even though there are more consumers (implying a larger range of horizontal values and less selection), the consumers get fewer false positives (and fewer false negatives). The latter effect dominates when $aq < 0.84$, leading to the result. Finally, when $aq > 0.84$, increasing a has a larger (relative) effect on the probability of consumption than the false positive/negative rate, so that selection decreases.

The way to understand the above is that, whenever something increases the ex-ante probability of consumption, it will decrease selection because it makes horizontal preferences wider in the subpopulation that consumes. Furthermore, anything that leads to the consumers making fewer wrong decisions, increases selection. These effects go into the same direction if $q \leq 0$ but when q is positive, we need an extra condition (comparing aq to 0.84) to say what the overall effect is.

Dynamics of Selection

This subsection studies what happens to selection when we consider more than one period, assuming rational consumers. As we saw before, Equation (3.1) shows that people who end up consuming the movie are on average more positively inclined toward it than your average Joe. However, if a second-period consumer is sophisticated, she will understand that reviews are selected and will be able to back out the true quality:

$$x^* = \mu_1 - \bar{v}_1,$$

where μ_1 is the mean first-period review and $\bar{v}_1$ the expected selection. Because we have a continuum of consumers, who write reviews honestly and without error, these values are exact. This means that the second-period sophisticated consumers purchase if and only if $v_i \geq -x^*$, so that the mass of second-period consumers is $\Phi(\sqrt{\tau_v}x^*)$ (knowing x^* but not v_i, only that $v_i \sim N(0, 1/\tau_v)$). Finally, this means that the expected horizontal preference of those who end up consuming in the second period is:

$$\bar{v}_2(x^*) = \mathbb{E}[v_i | v_i \geq -x^*] = \frac{1}{\sqrt{\tau_v}} \frac{\phi(x^*\sqrt{\tau_v})}{\Phi(x^*\sqrt{\tau_v})}. \tag{3.3}$$

This leads us to the following Proposition:

Proposition 3.8. *For given parameter values τ_v, σ^2, and q, there exists a unique threshold quality $\hat{x}^*$ such that $\bar{v}_1(q) \geq \bar{v}_2(x^*)$ for all $x^* \geq \hat{x}^*$. That is, when the realized quality is high and consumers sophisticated, there is more selection in the first period than the second.*

Proof. We note that $\frac{\phi(x^*\sqrt{\tau_v})}{\Phi(x^*\sqrt{\tau_v})}$ is strictly decreasing in x^* for all x^*, and it approaches 0 as x^* increases without bound. Thus, $\bar{v}_2$ is everywhere strictly decreasing (and it decreases from $\frac{1}{\sqrt{\tau_v}}$ to 0), while $\bar{v}_1$ is constant in $\left(0, \frac{1}{\sqrt{\tau_v(1+\sigma^2\tau_v)}}\right)$. Therefore, $\bar{v}_2 > \bar{v}_1$ for low x^*, but the inequality is reversed for higher values. □

The intuition for this result is that, in the basic model, selection is decreasing in the (known) mean utility. Thus, when the expected mean (q) is sufficiently low compared to the realized mean utility (x^*), there is less selection in the second period after all uncertainty has been resolved. In the data, this would mean that movies, which turn out to be bad relative to expectations, should get more selective reviews after the first week than those that have high underlying quality, even if both types of films have the same ex-ante expectation.

One remark to make is that, in this model, there are no dynamics after the second period because everyone knows x^* and all the new consumers have the same distribution of tastes as old ones. One possible remedy is assuming that the consumers do not fully learn x^* from the first-period reviews (maybe because reviews do not fully correspond to people's utilities or because something gets lost in translation).[14] A way to model this would be to assume that, in period t, $x \sim \left(x_t, \frac{1}{\tau_x(B)}\right)$, where x_t corresponds to the updated mean ($\mu_t - \bar{v}_t$) and $\tau_x(B)$ is the precision of the prior as an increasing function of B, the box office revenue. In this case, holding x_t fixed, increasing B increases τ_x, which will lead to a decline in σ^2 and an increase in a, implying lower selection if and only if $ax_t > 0.84$. Thus, higher box office revenue could lead to lower selection, although it could also increase selection if x_t or a was small.

The observation that selection often declines throughout a movie's theatrical run can thus be explained either by a more accurate prior (τ_x) or a timing of consumption where the consumers with the highest horizontal preferences see the movie first, then the group with the next-highest preferences, and so on. The model with only one movie (and a fixed outside option) cannot capture this phenomenon but adding more movies would imply that not everyone, who would enjoy a movie, goes to see it in the opening week because they have other options. Then, later weeks in the theatrical run have consumers with lower horizontal preferences seeing any given movie. Similarly, if people with high preferences (v) found more time to consider movies, they would be more likely to end up seeing something in the opening week.

As a final note, observe that even if there are no dynamics in selection after the second period, this does not mean that box office revenue will stay constant. It simply means that the probability of consumption conditional on "considering to see a movie" is constant but we can assume that the mass of people who consider to see a movie is exponentially decreasing as a function of time since release, everything else equal. Figure C.9 provides evidence for this claim. In the figure, we plot the seven-day rolling sum of daily (mean) revenues across movies conditional on having been out the same number of days. On the other hand, Figure C.10 reveals that the within-movie variance of IMDb reviews decreases during a movie's theatrical run and that the rate of decline is faster in the early days. This suggests that the audiences of a movie become more homogeneous as the movie gets older. It also suggests that the consumers that have the highest preferences for a movie see it early, so that the pool of consumers gets worse as time goes by.

[14]If there were other entertainment options available at $t-1$, and they are not observable to the period-t consumers, then these consumers may not be able to back out the true mean quality.

Sophistication Matters for Selection

Next we will analyze how the previous result depends on the level of sophistication of the consumers. To proceed, we make the following assumption which clarifies the meaning of sophistication.

Assumption 3.1. *$\theta \in [0,1]$ is a measure of consumer sophistication, such that second-period consumers expect population mean quality to be $x^{**} \equiv \bar{r}_1 - \theta \bar{v}_1$, and they completely ignore their personal signals, putting probability one on x^{**}.*

With this Assumption in hand, we know that completely ignorant consumers will blindly trust the mean review, $\bar{r}_1$, while fully sophisticated consumers will correct for selection (as was assumed above). This leads to the following Proposition.

Proposition 3.9. *Second-period selection, $\bar{v}_2$, is strictly increasing in θ, the level of sophistication.*

Proof. It is easy to see that $\bar{v}_1$ is always positive because the only way for it to be negative is for consumers with negative v_i to be more likely to consume in period one than consumers with positive preferences, which is clearly not true. This implies that x^{**} is strictly decreasing in θ. Thus, due to $\bar{v}_2(x^{**}) = \frac{1}{\sqrt{\tau_v}} \frac{\phi(x^{**}\sqrt{\tau_v})}{\Phi(x^{**}\sqrt{\tau_v})}$ being strictly decreasing in x^{**}, we have that $\bar{v}_2(x^{**})$ is strictly increasing in θ. □

The Proposition says that we should expect to see more selection in reviews the more sophisticated consumers are. This is simply due to the fact that more sophisticated consumers make fewer mistakes in deciding to see a movie, so that they are almost sure to enjoy it in period two, conditional on decision to watch.[15] Everything else equal (holding x^* and the population of arriving consumers fixed), lower selection in the weeks subsequent to the opening week implies that consumers are not fully rational. If they were rational, selection would always decrease after the opening week as consumers make few or no mistakes.

One thing left unmentioned is that we are implicitly assuming the consumers are smart enough to incorporate their own first-period signals in a Bayesian way but may be imperfect at understanding other people's preferences.

3.3.4 Two movies

In all of the above, we have assumed that there is a single movie the consumers can choose to watch. Thus, as long as the signal is higher than a threshold (normalized at zero), the consumer will choose to watch. However, in reality, there may be multiple films that satisfy the "high-signal" condition. How does this affect selection? The simple answer is that there will be **less** selection if consumers can choose from two very similar films instead of one, but the magnitude of this selection depends on how similar those movies are. Selection may in fact increase as the number of movies goes up if the movies are sufficiently different, attracting heterogeneous audiences. To show this,

[15]Fully sophisticated consumers make no mistakes because they know the true population mean and their own horizontal preference v. However, as consumers become less sophisticated, they trust the reviews too much, failing to take selection into account, which means that there will be more and more consumers with $v < -x^*$.

let us use the previous model where the population mean is unknown but consumers know how their preferences differ from that mean.

To be more specific, let us assume a population of consumers with utilities $u_{ij} = x_i + v_j$, where x_i is the unknown population mean utility for movie i, and $v_j \sim N(0, 1/\tau_v)$ is the individual preference of consumer j. Let us further assume that there are two movies with unknown but identically distributed population means x_1 and x_2, $x_i \sim N(q, 1/\tau)$, and that every consumer j has the same individual preference v_j for these two movies (because the movies belong to the same genre, say) – and she knows what it is.[16] Let $s_i = x_i + \epsilon_i$, where $\epsilon_i \sim N(0, 1/\tau_\epsilon)$ is an individual signal on the mean quality of movie i, such that signals are independent across consumers and movies.

Proposition 3.10. *Under the above assumptions, the amount of selection for each of the two movies is:*

$$\bar{v}_{(2)} = \frac{2a\phi(aq)\left[1 - \Phi(abq)\right]}{\tau_v\left[\Phi(aq) + 2T(aq, b)\right]}, \tag{3.4}$$

where $a = \sqrt{\frac{\tau_v}{1+\sigma^2\tau_v}}$, $b = \sqrt{\frac{\sigma^2\tau_v}{2+\sigma^2\tau_v}}$, $\sigma^2 = \frac{\tau_\epsilon}{\tau(\tau+\tau_\epsilon)}$, *and* $T(h,c) = \phi(h)\int_0^c \frac{\phi(hx)}{(1+x^2)}dx$ *is Owen's T function.*

Proof. See Appendix C.1. □

Proposition 3.10 has a complicated expression for expected selection but the main idea is simple: selection depends non-trivially on the parameters of the model and it is decreasing in q, the mean quality.

Corollary 3.3. *Selection* $\bar{v}_{(2)}$ *is strictly decreasing in* q *for* $q \geq 0$.

Proof. Note that, for positive q, the numerator is always strictly decreasing in q because $\phi(y)$ is strictly decreasing for positive y and $\Phi(y)$ is strictly increasing for all y. This relies on the fact that a, b are positive constants. Therefore, we only need to show that the denominator is strictly increasing in q. To this end, let us differentiate $\Phi(y) + 2T(y, b)$ with respect to y. This yields:

$$\begin{aligned}\frac{d(\Phi(y) + 2T(y,b))}{dy} &= \phi(y) - 2y\phi(y)\int_0^b \frac{\phi(yx)}{1+x^2}dx - 2\phi(y)\int_0^b \frac{yx^2\phi(yx)}{1+x^2}dx \\ &= \phi(y) - 2y\phi(y)\int_0^b \frac{1+x^2}{1+x^2}\phi(yx)dx \\ &= \phi(y) - 2\phi(y)\left[\Phi(yb) - \frac{1}{2}\right] \\ &= 2\phi(y)\left[1 - \Phi(yb)\right] > 0.\end{aligned}$$

This proves that selection is decreasing in q. □

[16] This is the crucial assumption which we will discuss after the next proposition and its corollaries.

The following Corollary offers an interesting observation on how adding a second movie decreases selection compared to just one film.

Corollary 3.4. $\bar{v}_{(2)} = 2\frac{[1-\Phi(abq)]\Phi(aq)}{\Phi(aq)+2T(aq,b)}\bar{v}_1 < \bar{v}_1$, *for* $q \geq 0$. *For all* $q \in \mathbb{R}$, *we have the lower bound:* $\bar{v}_{(2)} > \frac{2[1-\Phi(abq)]}{2-\Phi(aq)}\bar{v}_1$.

Here, $\bar{v}_1$ is the expected (first-period) selection for a movie that has no competition. It was defined in Equation (3.1). The Corollary says that selection for a movie that is playing alone is **higher** than for a movie that has a competitor. Numerical calculations seem to reveal that the Corollary also holds for $q < 0$, which will not be shown here.

Proof. The fact that we can express $\bar{v}_{(2)}$ as a multiple of $\bar{v}_1$ is simply due to rewriting Equation (3.4) and multiplying and dividing by $\Phi(aq)$. Now, if $q \geq 0$, the proof for the inequality is simple: $2[1-\Phi(abq)] \leq 1$ because a and b are always positive and $\Phi(0) = \frac{1}{2}$. On the other hand, $\frac{\Phi(aq)}{\Phi(aq)+2T(aq,b)} < 1$ because $T > 0$.

If $q < 0$, the proof is not as simple because $2[1-\Phi(abq)] > 1$, which is why it is not provided here, even though it seems to hold. However, we can always obtain a lower bound for the expected selection with two movies using two facts about Owen's T function: (1) $T(aq,1) > T(aq,b)$ because $b < 1$ and T is strictly increasing in b, and (2) $T(aq,1) = \frac{1}{2}\Phi(aq)[1-\Phi(aq)]$. These allow us to evaluate the factor in front of $\bar{v}_1$:

$$\begin{aligned}\frac{2\left[1-\Phi\left(abq\right)\right]\Phi\left(aq\right)}{\Phi\left(aq\right)+2T(aq,b)} &> \frac{2\left[1-\Phi\left(abq\right)\right]\Phi\left(aq\right)}{\Phi\left(aq\right)+\Phi(aq)[1-\Phi(aq)]} \\ &= \frac{2\left[1-\Phi\left(abq\right)\right]}{2-\Phi(aq)},\end{aligned}$$

where the inequality follows using the two facts. This expression is always less than 1. □

It is in order to discuss why selection is decreasing in the number of movies and when this would not be the case. Here, we have crucially assumed that consumer j's horizontal valuation, v_j, is the same for both films (implying perfect positive correlation), so that the movies can be thought to be in the same genre and people's preferences one dimensional. Consider what happens if consumers are allowed to have different horizontal preferences, v_{j1} and v_{j2}, for the two films. Intuitively, when the correlation between these horizontal preferences decreases from 1 toward -1, the consumers will be more and more divided into two camps. Some of them will be (much) more likely to consume 1 than 2, and vice versa for others.[17]

Let us now, through numerical analysis, check what happens when there is perfect negative correlation, so that $v_{j1} = -v_{j2} \equiv v_j \sim N(0, 1/\tau_v)$. Conditional on v_j, the probability of choosing

[17]The intuition is that a high v_{j1} makes it very unlikely that v_{j2} is high as well, so that the probability of choosing movie 1 is much higher than that of choosing 2, conditional on having a high v_{j1}.

movie 1 is then:

$$
\begin{aligned}
P(v_j) &= \mathbb{P}(1 \text{ acceptable}|v_j)\mathbb{P}(1 \text{ better than } 2 \mid 1 \text{ acceptable}, v_j) \\
&= \int_{-\frac{\tau}{\tau_\epsilon}q - \frac{\tau+\tau_\epsilon}{\tau_\epsilon}v_j}^{\infty} \sqrt{\frac{\tau\tau_\epsilon}{\tau+\tau_\epsilon}}\phi\left(\sqrt{\frac{\tau\tau_\epsilon}{\tau+\tau_\epsilon}}(s_1 - q)\right)\Phi\left(\sqrt{\frac{\tau\tau_\epsilon}{\tau+\tau_\epsilon}}(s_1 + 2\frac{\tau+\tau_\epsilon}{\tau_\epsilon}v_j - q)\right)ds_1 \\
&= \int_{-\frac{q+v_j}{\sigma}}^{\infty} \phi(Y)\Phi\left(Y + \frac{2v_j}{\sigma}\right)dY,
\end{aligned}
$$

where again $\sigma^2 = \frac{\tau_\epsilon}{\tau(\tau+\tau_\epsilon)}$ and Y is a standard normal variable, using the fact that s_1 is a normally distributed signal. The second line follows because the consumer will prefer movie 1 if and only if $v_j + \frac{\tau q + \tau_\epsilon s_1}{\tau+\tau_\epsilon} \geq -v_j + \frac{\tau q + \tau_\epsilon s_2}{\tau+\tau_\epsilon} \Leftrightarrow s_2 \leq s_1 + 2\frac{\tau+\tau_\epsilon}{\tau_\epsilon}v_j$. Now, we can calculate the expected share of consumers choosing movie 1 by integrating over the distribution of values, v_j.

$$
\begin{aligned}
Q \equiv \mathbb{P}(\text{choose movie } 1) &= \int_{-\infty}^{\infty} \sqrt{\tau_v}\phi(\sqrt{\tau_v}v_j)P(v_j)dv_j \\
&= \int_{-\infty}^{\infty} \phi(X)\int_{-\frac{X/\sqrt{\tau_v}+q}{\sigma}}^{\infty} \phi(Y)\Phi\left(Y + \frac{2X}{\sqrt{\tau_v}\sigma}\right)dYdX,
\end{aligned}
$$

where $X = \sqrt{\tau_v}v_j$ is a standard normal variable, and it is independent of Y. Finally, we can calculate the expected horizontal preference of those who choose to watch movie 1 as follows:

$$
\mathbb{E}[v \mid \text{consume } 1] = \frac{1}{Q}\int_{-\infty}^{\infty} \frac{X}{\sqrt{\tau_v}}\phi(X)\int_{-\frac{X/\sqrt{\tau_v}+q}{\sigma}}^{\infty} \phi(Y)\Phi\left(Y + \frac{2X}{\sqrt{\tau_v}\sigma}\right)dYdX.
$$

We can now numerically estimate the expected selection for movie 1. Simultaneously, we will also be estimating the expected selection for movie 2, because the distribution of v_j is symmetric around zero. Figure C.11 plots the amount of selection as a function of q and σ, for two values of τ_v. We learn that selection decreases when q, σ, or τ_v increases. However, this does not tell us anything about the effect of having two movies instead of one. Therefore, in Figure C.12, I plot the difference in selection $(\bar{v}_{(2)} - \bar{v}_1)$ as a function of q and σ, for two values of τ_v. Interestingly, selection seems to always be higher with two movies than one, as long as they are perfectly negatively correlated. Moreover, the difference $(\bar{v}_{(2)} - \bar{v}_1)$ seems to be increasing in q, and decreasing in σ and τ_v. Thus, we would expect that increasing the number of movies has a larger impact on selection if the movies are of high expected quality or if the qualities are well-known in advance (high τ, which implies low σ), or if people have heterogeneous horizontal preferences (low τ_v).

When expected qualities increase, having two movies makes almost every consumer watch something, so that selection decreases relative to low expected qualities but increases relative to having just one movie because everyone has a positive preference for something ($v_j < 0 \Rightarrow v_{j2} > 0$). On the other hand, when τ_v is low, there are many people with a high or low v_j, so that there will be more selection both relative to a high τ_v and relative to just one movie. We can summarize our findings in the following Proposition:

Proposition 3.11. *Selection with two movies is higher than with a single movie if correlation* *between the horizontal preferences for the two movies is below some threshold.*

I offer no formal proof of the Proposition but we have shown that selection decreases when correlation is perfectly positive and increases when it is perfectly negative. We would also expect there to be a point where selection with one film equals that with two movies.

The main intuition for lower selection under competition when correlation is perfect is because people have a higher chance of obtaining a "good" signal (there are two independent signals, so the probability of at least one crossing a threshold is higher). Thus, the more movies there are, the more likely it is for people with lower horizontal preferences to watch something. Since movies are assumed ex-ante identical, each of them will get the same share of these low-preference consumers, pushing mean selection down. Conversely, if correlation is perfectly negative, the intuition for higher selection is that everyone will be very likely to consume the movie they have a positive preference for (they always have a positive preference for one movie and negative for the other), driving selection up as compared to the case of only one movie (where people are likely to consume something even if their preference is negative because they get two shots for a high signal).

How about the assumption that the movies have the exact same ex-ante quality (both in terms of expectation and variance)? If we keep assuming that the variance of quality is still the same for both films and t here is perfect correlation, but let the expectation for movie 1 be higher than that for movie 2, anyone will have a higher chance of watching the first movie instead of the second.[18] This leads to lower selection for the first movie and higher selection for the second because the movies become islands in the sense that most consumers get signals that induce them to choose 1, and people for whom the difference in signals is high enough for them to choose 2 follow almost the same distribution as people who only have one movie to choose from (because whenever someone prefers 2 to 1, they are more likely to also find 2 alone acceptable, while preferring 1 to 2 is less of a guarantee that 1 is acceptable alone).

3.3.5 (Testable) Predictions

The models of this section have offered us a number of predictions on what selection will look like and how it depends on the parameters. To summarize, we would expect the following:

1. Selection decreases in (expected) population mean quality (q).

2. Reviews decrease if τ_x (precision of population mean), or τ_v (consumer homogeneity) increases.

3. Selection increases in signal precision (τ_ϵ) and the variance of population utilities ($1/\tau_u$).

4. Selection in period 2 is lower than period 1 if realized quality (x^*) is high compared to expectation (q).

5. Selection decreases in the number of (similar) movies.

[18]The condition for choosing movie 2 instead of 1 becomes $s_2 \geq s_1 + \frac{\tau}{\tau_\epsilon}(q_1 - q_2) > s_1$. That is, signal for movie 2 has to be the higher, the higher is the quality difference, but we also know that the distribution of signals is higher for movie 1.

3.4 Putting the Predictions to Test

In this section we will go over the predictions produced by the previous models. To properly test all of them would require more comprehensive data but we will see how far our data takes us.

3.4.1 Selection Decreases in Population Mean Quality

To get at selection and how consumer reviews/realized utilities depend on quality and horizontal preferences, consider the following regression:

$$R_{jt} = \alpha_j + \beta_C C_{jt} + \beta_B \ln(B_{jt}) + \lambda(t - d_j) + \epsilon_{jt},$$

where R_{jt} is the (cumulative) period-t user rating of movie j, α_j is a movie-specific constant (intercept), C_{jt} is the critical average, B_{jt} the total box office revenue up and until t, and d_j the release date of film j. Everything else has been lumped into the error term ϵ_{jt}. We have included the critic average to see if the relationship between realized user ratings and critical reviews is positive, while box office has been included to better understand selection. This is because, as we will see, we can think about box office as a measure of selection: higher box office means less selection, everything else equal (for example, within a movie). The more people go watch a film, the less selected this population is, on average. Thus, we would expect ratings to decrease in box office. We would also expect ratings to decrease in time since release, implying a negative λ in the above regression, because of the most hardcore fans seeing a movie first. λ differs from β_B in that it measures how the appeal of a movie decreases over time regardless of how many people have seen it already (measured by box office). People who see a movie later are less excited to see it, which is captured by $\lambda < 0$. The constant α_j and the coefficient on critic reviews, β_C, measure the movie-specific, opening-day review potential, which can be understood as a signal of quality (although it will overestimate the true population mean because it corresponds to the mean utility of first-period consumers as a function of the critics).

Why am I using cumulative variables instead of daily observations? First, for critic reviews, this makes sense because most of them are published before or during the opening weekend and, if consumers care about critic averages when choosing to watch a movie, they care about the current, cumulative average instead of just yesterday's mean review (which may not even exist). Second, using cumulative box office statistics is better when we are trying to understand how reviews react to the number of people watching a movie. The cumulative revenue is much less noisy than its daily counterpart, not least because people may write reviews with a lag. Finally, I use cumulative user reviews because, again, there is more random noise on the daily level.[19] The daily average may jump up and down merely because there are only a handful of people watching/reviewing a movie on a given day, and using averages gives a day with thousands reviews the same weight as a day with 10 reviews.

[19]It turns out that some people change their mind or remove their review altogether so that the daily average can be negative or higher than the highest possible review, in the case of less popular movies.

Table 3.6: Regression of RT user average on some predictors, using the panel dataset.

	(1)	(2)	(3)	(4)	(5)
intercept	4.7451**	3.1556**	6.701**	7.755**	3.815**
	(0.105)	(0.153)	(0.177)	(0.037)	(1.216)
critic avg	0.4504**	0.4787**	0.116**	0.079**	0.078**
	(0.017)	(0.017)	(0.027)	(0.027)	(0.027)
days	-0.0111*	-0.0103*	-0.009**	-0.010**	-0.010**
	(0.005)	(0.005)	(0.001)	(0.001)	(0.001)
days*critic avg	0.0011	0.0002	0.000*	0.001**	0.001**
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
log(box office)		0.0931**		-0.055**	-0.055**
		(0.007)		(0.005)	(0.005)
genre var					1.870**
					(0.574)
# obs	4963	4963	4963	4963	4963
# titles	178	178	166	166	166
R^2	0.358	0.383	-	-	-
Movie FE			Y	Y	Y

***Notes**: Standard errors in parentheses (*=95%, **=99%). (1) and (2) are simple OLS, while (3) and (4) fit a movie-specific (random) intercept. The dataset includes the first 40 days of each movie's theatrical run, if available.*

Table 3.6 presents the results of the above regression with four modifications. The first two are overly simplistic OLS regressions, with and without box office revenue, ignoring the fact that we observe the same movies multiple times. The next two, however, take movie fixed effects into account, but are differentiated in terms of whether box office is included or not. Regression (5) is identical to (4) but it includes an extra variable (*genre var*) which will be defined later. All of the regressions reveal that consumer reviews are increasing in mean critic reviews and decreasing over the theatrical run although the rate of decline is slower for critically acclaimed movies. Including box office revenue does not change how the length of the theatrical run matters ($\hat{\lambda} \approx -0.01$ in all specifications), even though it is true that deeper theatrical runs mechanically imply more revenue. We can interpret $\hat{\lambda}$ as as a measure of how selection decreases over the theatrical run, everything else equal. That is, regardless of how many people have seen a movie, its audiences will be less horizontally selected over time – the most enthusiastic viewers, if any, see it early. The interaction between *days* and *critic avg* tells us that movies with higher "quality", as measured by critic reviews, tend to receive more homogeneous audiences over time, leading to slower decline in reviews.[20]

Let us get back to the interpretation of *critic avg*. The fact that the coefficient is much smaller than 1 tells us that consumers are less sensitive to "cinematographic quality" than critics, at least conditional on seeing the movie. Thus, higher critic reviews are associated with higher consumer

[20]This means that there are more people who like the movie but see it late, the higher the "quality" is.

reviews (μ), which implies that the underlying population mean utility (q) is also increasing in critic reviews (if this were not the case, consumer reviews would not be increasing either). But now, we can write selection as $\mu - q = \sigma\phi(q/\sigma)/\Phi(q/\sigma)$, which is decreasing in q, and q in turn is increasing in critic reviews, so that higher critic reviews seem to lead to less selection.

3.4.2 Reviews Decrease in τ_x and τ_v.

Still looking at Table 3.6, let us discuss the effect of box office revenue on user ratings, β_B. First, why does this effect change signs when we include movie-specific intercepts α_j? Without these fixed effects, box office acts as an indicator of quality, however imperfect, because it varies from movie to movie (as we can see, the inclusion of box office in the simple OLS shrinks the intercept but has no other significant effects) and it adds another dimension on top of *critic avg.* It is important to understand that the model without box office assumes that consumer utilities and critic average always move in the same direction, while including box office allows the consumers sometimes to enjoy movies the critics find horrible.[21] Finally, and possibly most importantly, let us look at specification (4) and why the effect of higher box office is negative. In this model, we have already controlled for the movie-specific (period-0) mean quality by including fixed effects, so that what box office measures are the horizontal preferences. That is, higher box office is now associated with an audience with more heterogeneous preferences, so that reviews conditional on quality are going to be lower, or less selected in our language. In terms of our model, high box office corresponds to high $\Phi(aq)$, and therefore high a conditional on q. Knowing that a is increasing in τ_v and τ_x, higher box office can be seen to indicate a more homogeneous population of consumers and/or a movie with a precise prior quality. The way to interpret the estimate $\hat{\beta}_B = -0.055$ is that a doubling of box office is associated with a decline of 0.055 in consumer reviews, holding everything else constant.

3.4.3 Selection Increases in Signal Precision and the Variance of Utilities

To understand this prediction, let us assume that the variances of critic and consumer utilities are positively correlated, so that whenever a movie has a high variance of critic reviews, it has an underlying distribution of consumer utilities with high variance. With this assumption, even though we do not observe the variance of within-movie critic reviews, we can lump movies into genres (highly overlapping), calculate variances of mean critic reviews for each genre, and use these variances as a measure of the underlying variance of utilities. Thus, if a genre has a high variance of critic reviews, then we would expect $1/\tau_u$ to be higher, but cannot say much about signal precision.

Looking at column (5) in Table 3.6, the only difference to column (4) is that we have added the variable *genre var* which takes into account the mean variance of critic reviews for that genre.[22] As

[21]Including box office as a predictor allows consumers to like movies that critics do not, so that two movies may have the same consumer average even though they have different critic averages.

[22]This has been calculated as follows: 1) take the variance of critical reviews for each genre in the eight-year dataset of Rottentomatoes ratings, 2) convert these genres into the (slightly) different genres in the shorter panel data, 3) calculate mean variances for each movie taking into account that it may belong to multiple genres simultaneously (if

Table 3.7: OLS of second-period selection on underlying quality, using the panel dataset ($R^2 = 0.765$, 99 obs).

	coef	std err	t	P>|t|	[95.0% Conf. Int.]
intercept	1.4120	0.227	6.226	0.000	0.962 1.862
critic hype	-0.0014	0.002	-0.706	0.482	-0.005 0.002
$\hat{q}_0$	-0.6002	0.046	-12.949	0.000	-0.692 -0.508

we can see, this has practically no effect on any of the variables of interest but it affects the (mean) intercept because we are now grouping movies into genres. The positive sign on *genre var* suggests that genres with a higher variance of critic reviews (and therefore a higher variance of consumer utilities, as assumed) have higher consumer reviews. This is evidence in favor of the hypothesis that selection is increasing in the variance of utilities because, conditional on underlying quality, higher variance is associated with higher reviews, and therefore higher selection. This result is not sensitive to the definition of selection; we could have as well computed the estimated underlying quality (assuming our model is approximately correct) and shown that reviews increase in variance. See Appendix C.2 for the details.

3.4.4 Selection in Period 2 is Lower Than Period 1 if Realized Quality is High

Now, if we want to show that selection decreases from period 1 to period 2 if the underlying quality turns out to be high compared to expectations, we need to have a proxy for what the expectations were. If we use the pre-release Tomatometer (labeled *critic hype* in the regression) score as a proxy for these expectations, and let $\hat{q}_0$ be the realized quality, as estimated from first-period reviews (see Appendix C.2 for the details), we can regress the difference between user ratings and realized quality in the second period on *critic hype* and $\hat{q}_0$. The results are shown in Table 3.7. We see that the ex-ante expectation, if measured by critical hype, has no impact on what the selection in period 2 is going to be, but that the estimated first-period underlying quality has a strong negative impact on selection. However, this impact is less than one-to-one, so that even though selection decreases, consumer reviews are going to increase if realized quality is higher.

3.4.5 Selection Decreases in Number of Movies

In this final subsection, we will consider how the number of movies affects selection. The simple prediction says that user ratings (and selection) should decrease in the number of movies if there is positive enough correlation in horizontal preferences for the movies but decrease if correlation is negative enough. As the results of a simple regression – presented in Table 3.8 – suggest, selection seems to increase in the number of movies. In the regression, we are using the total number of movies playing in theaters on any given day (*movies*) as an explanatory variable for user ratings.

a movie is a drama-comedy, its variance is the mean of the variances of the drama and comedy genres).

Table 3.8: Mixed LM regression of RT user average on the number of movies

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	7.450	0.200	37.216	0.000	7.058	7.843
critic avg	0.095	0.027	3.536	0.000	0.042	0.147
log(box office)	-0.045	0.005	-9.125	0.000	-0.054	-0.035
movies	0.002	0.001	2.316	0.021	0.000	0.004
days	-0.004	0.000	-14.574	0.000	-0.005	-0.004

If anything, the relationship is positive, which is sensible because it means that a larger choice set makes people go see movies they have the highest preferences for. In terms of the model, this would also mean that the movies are horizontally differentiated, which makes sense when we understand that the release dates have been chosen optimally by the producers (even though we have abstracted away from this endogeneity issue). For robustness, I checked if conditioning on the genre makes a difference but there is no clear relationship (drama movies have less selective reviews when there is competition from other drama movies, while the opposite is the case for action films).

3.5 Conclusion

In this paper, we have seen how to model selection in two similar models: one where population mean is known but horizontal preferences are not, and another where population mean is unknown but horizontal preferences known. Both of these models produce interesting hypotheses regarding the amount of selection and how it depends on the distributions of preferences and movie qualities. When we put these hypotheses to the test, we see that they seem to hold. Probably the most interesting result is that, conditional on movie quality, consumer reviews tend to decrease as a function of box office revenue because higher revenue means that more people with low preferences see the movie – leading to less selection. My estimate is that when box office is doubled, consumer reviews decrease by 0.055 units (on a scale of 1 to 10), conditional on quality. This means that, if one merely uses box office revenue to measure movie quality (sometimes called appeal), one may fail to capture the horizontal component of utility. Namely, movies that are highly popular among certain groups of consumers may have low box office but high quality, while some films with high box office may be of lower quality but wider potential audience. If prices were allowed to adjust, niche movies of high quality would then command a high price while box-office hits of mediocre quality would sell for less.

Furthermore, higher critic reviews are associated with better consumer reviews but lower selection. I also find evidence that selection decreases over time (by approximately 0.07 units per week, conditional on box office), and it does so faster for movies with high box office, as mentioned above. On the other hand, surprisingly high first-day box office makes consumer reviews more important in predicting future box office, which is evidence for the impact of word of mouth. Moreover, selection is higher for genres with divided critical opinion because then only the (very) positively inclined consumers will see the movie. Finally, there is weak evidence that selection increases in the number of competing movies, which supports the theoretical finding that horizontally differentiated movies get more selective reviews when there is competition than when they are playing alone. If there was only one film to choose from, it would not get a selected audience because anyone with positive expected utility would see it.

It is left for future research to determine how consumers read reviews and how sophisticated they are in making their decisions to see a movie. If consumers are not perfectly rational or sophisticated, they may mistakenly go see movies with high reviews, even though those reviews come from a very selected subpopulation, meaning that the simple average rating is not good for the consumers but may be optimal for the producers. This would also explain why some movie trailers seem to be so bad they would drive 80% of the consumers away; they are catering to the 20% that react well to the trailer, go see the movie, and leave a very positive review, which in the end leads to other, less excited, consumers seeing the movie. This is aligned with what Li and Hitt (2008) conclude.

As a final remark, the theoretical result that better signals lead to higher reviews and more selection (in most cases) suggests that the birth of platforms that make reviews personal (think Netflix) would increase selection by better allocating consumers to the right movies but it might also lead to less selection if people were given wrong signals and they followed them blindly. This is where sophistication enters the picture.

Appendix A

Appendix to Chapter 1

A.1 General Results for Section 1.3

Let $k_{fb} := \min\left\{k \;\middle|\; c < \frac{s}{(1-\mu)(1-\pi_k)}\right\}$, which is the socially optimal (first best) number of searches. It is derived in the Lemma below:

Lemma A.1.1. *k_{fb} is the socially optimal number of searches.*

Proof. Note that social surplus can be written as $W_k = (\pi_0 + (1-\pi_0)\mu^k)(\pi_k - c) - \mathbb{E}[\text{cost}]_k = \pi_0 - (\pi_0 + (1-\pi_0)\mu^k)c - \mathbb{E}[\text{cost}]_k$. Therefore, $W_{k+1} - W_k = (1-\mu)(1-\pi_0)c - s \geq 0 \Leftrightarrow c \geq \frac{s}{(1-\mu)(1-\pi_k)}$. Because the RHS of this inequality is strictly increasing in k and without bound, there has to be some k for which the inequality is reversed, meaning that there is some k_{fb} such that $\frac{s}{(1-\mu)(1-\pi_{k_{fb}-1})} \leq c < \frac{s}{(1-\mu)(1-\pi_{k_{fb}})}$. This means that $W_0 < W_1 < \ldots W_{k_{fb}-1} \leq W_{k_{fb}}$ and $W_{k_{fb}} > W_{k_{fb}+1} > \ldots$. □

The following Lemma shows that the firm will not set the socially optimal price $p_{k_{fb}}$ because it will always have an option that is at least as good.

Lemma A.1.2. *Assume that $\pi_0 > c$. Then $k_{fb} \leq k^* + 1$, which means that the firm will never set $p > \hat{p}_{k^*+1}$.*

Proof. If, contrary to the claim, $k_{fb} > k^*+1$, we would know that $c \geq \frac{s}{(1-\mu)(1-\pi_{k^*+1})} > \hat{p}_{k^*+1} > p_{k^*}$. However, this would mean that the firm makes negative profits at $\hat{p}_{k^*+1}$ and the consumer gets zero at that price. Thus, social surplus would be negative at $\hat{p}_{k^*+1}$. However, the only way social surplus can be negative is if $\pi_0 < c$ (since social surplus is increasing in this range and $W_0 = \pi_0 - c$). But this is a contradiction because we assumed that $\pi_0 > c$ □

Proposition 1.2 Assume that $\pi_0 > c$. If $s \geq (1-\mu)(1-\pi_0)\pi_0$, the firm's optimal price is π_0. If $s < (1-\mu)(1-\pi_0)\pi_0$, we have two cases:

1. If $\Pi_{k^*} \geq \Pi_0$, the firm's optimal price is either p_{k^*} or $\hat{p}_{k^*+1}$, and the firm prefers p_{k^*} if and only if $s \geq \bar{s}_k$, where:

$$\bar{s}_k := \frac{\pi_0 + (1-\pi_0)(1-\mu)\mu^k c}{\pi_0(k+1) + \frac{(1+\pi_0)}{1-\mu} + (1-\pi_0)\mu^k + \frac{\pi_0^2}{(1-\mu)(1-\pi_0)\mu^k}}.$$

2. If $\Pi_{k^*} < \Pi_0$, firm will set either p_0 or $\hat{p}_{k^*+1}$, depending on which one gives higher profits:
   (a) $\Pi(\pi_0, \hat{p}_{k^*+1}) > \Pi_0 \Rightarrow \hat{p}_{k^*+1}$ is optimal
   (b) $\Pi(\pi_0, \hat{p}_{k^*+1}) \leq \Pi_0 \Rightarrow p_0$ is optimal.

Proof. $\pi_0 > c$ implies that the firm will not consider $\hat{p}_{k_{fb}}$ (by Lemma A.1.2). If search cost is higher than $(1-\mu)(1-\pi_0)\pi_0$, the equilibrium will feature $p^* = \pi_0$ because the consumer will just buy at that price and, on the other hand, she will never get positive utility for $p > \pi_0$ (since her utility is strictly decreasing in p, given optimal behavior). The firm is now choosing its price knowing that it will induce a number of searches from the consumer. If $\Pi_{k^*} \geq \Pi_0$, the firm should set either p_{k^*} or p_{k^*+1} because the consumer is willing to search at these prices and the firm's profit is at least as high as at p_0. Because Π_k is either always increasing or V-shaped (Lemma 1.1), the firm should not consider any price between p_0 and p_{k^*}. The firm will choose p_{k^*} over p_{k^*+1} if and only if $\Pi_{k^*} \geq \Pi(\pi_0, \hat{p}_{k^*+1}) \Leftrightarrow s \geq \hat{s}_{k^*}$. Note that the firm does not want to set a price higher than $\hat{p}_{k^*+1}$ because it will be the residual claimant to the full social surplus but social surplus is decreasing at that point due to Lemma A.1.2.

If, however, $\Pi_{k^*} < \Pi_0$ so that Π_k is V-shaped and the firm's profit at p_{k^*} is less than that at p_0, equilibrium price will be either p_0 or $\hat{p}_{k^*+1}$. It cannot be anything higher than $\hat{p}_{k^*+1}$ because the firm again gets the full social surplus which is decreasing by Lemma A.1.2. It also cannot be anything between p_0 and p_{k^*} because those prices give profits that are lower than Π_0. Which one of the two prices is optimal depends on the profits as given in the Proposition. □

Let us remind ourselves of the Lemma introduced before Theorem 1.1:

Lemma 1.2 Let s_k solve $V_k(\pi_0, p_k) = 0$. Then $s_k > s_{k+1}$ $\forall k \in \mathbb{N}_0$, and $k^* = \max\{k \in \mathbb{N}_0 \mid s_k \geq s\}$.

Now we can restate and prove the Theorem.

Theorem 1.1 Assume that $\pi_0 > c$. Let $L(s) = \frac{c}{s}$ and $R(\pi_0, \mu) = \frac{\mu - L_0^2}{\mu(1-\mu)}$. Define $\hat{s}$ as follows:

$$\hat{s} := \max\{\bar{s}_0, s_1\} = \begin{cases} \bar{s}_0 \equiv \frac{(1-\mu)(1-\pi_0)}{(1-\mu)(1-\pi_0)+1}\left(\pi_0 + (1-\pi_0)(1-\mu)c\right), & \text{if} \quad \pi_0 \geq \frac{\sqrt{\mu}-\mu}{1-\mu} \quad \text{or} \quad \text{c is high,} \\ s_1 \equiv \frac{(1-\mu)(1-\pi_0)\mu\pi_0}{(1-\mu)(1-\pi_0)\mu + (\pi_0 + (1-\pi_0)\mu)^2}, & \text{if} \quad \pi_0 < \frac{\sqrt{\mu}-\mu}{1-\mu} \quad \text{and} \quad \text{c is low.} \end{cases}$$

If $L(\hat{s}) \geq R(\pi_0, \mu)$, consumer-optimal search cost is $s^* = \hat{s}$. If $L(\hat{s}) < R(\pi_0, \mu)$, optimal search cost is $s^* \leq \hat{s}$.

Proof. When $L(s) \geq R(\pi_0, \mu))$, the firm's profit function Π_k is strictly increasing in k, meaning that the firm always wants to induce the highest possible number of searches k. The lower s is, the more the consumer is willing to search, which implies a higher equilibrium number of searches, k^*. Using (1.1), we can compute the consumer's value function at the relevant prices:

$$V_k(\pi_0, p_k) = \pi_0 - \left[\frac{(\pi_0 + (1-\pi_0)\mu^k)^2}{(1-\mu)\mu^k(1-\pi_0)} + k\pi_0 + (1-\pi_0)\frac{1-\mu^k}{1-\mu}\right] s$$

$$V_{k+1}(\pi_0, \hat{p}_{k+1}) = 0.$$

Due to Lemma 1.2, we can construct a decreasing sequence of threshold values for the search cost, $s_0 > s_1 > s_2 > \ldots$, such that if the search cost falls between two thresholds, say s_k and s_{k+1}, then $V_k(\pi_0, p_k) > 0 > V_{k+1}(\pi_0, p_{k+1})$. If $s > s_0$, even $V_0(\pi_0, p_0) < 0$, in which case the firm will set $\hat{p}_0 \equiv \pi_0$, giving the consumer zero utility.

The consumer will (typically) get a strictly positive utility if and only if her search cost $s \geq \bar{s}_k$, as long as $s < s_0 = (1-\mu)(1-\pi_0)\pi_0$ because otherwise the consumer will always just buy at a price that makes her indifferent between exiting and buying.[1] $\bar{s}_k$ is a threshold such that the firm will choose p_k over $\hat{p}_{k+1}$ for $s \geq \bar{s}_k$, and this was introduced in Proposition 1.2.

To obtain the value of the search cost that maximizes the value function, one needs to find the lowest value of the search cost that still makes the firm want to set $p = p_0$. The consumer's equilibrium value when the firm prices at p_0 is $V_0(\pi_0, p_0) = \pi_0 - \frac{s}{(1-\mu)(1-\pi_0)}$, and the value at p_1 is $V_1(\pi_0, p_1) = \pi_0 - \frac{(1-\mu)\mu(1-\pi_0)+(\pi_0+(1-\pi_0)\mu)}{(1-\mu)\mu(1-\pi_0)} s$, because these prices make the consumer search 0 and 1 times, respectively. Thus, $s_0 = (1-\mu)(1-\pi_0)\pi_0$ and $s_1 = \frac{(1-\mu)(1-\pi_0)\mu\pi_0}{(1-\mu)(1-\pi_0)\mu+(\pi_0+(1-\pi_0)\mu)^2}$, where $s_0 > s_1$. Furthermore, $\bar{s}_0 = \frac{(1-\mu)(1-\pi_0)}{(1-\mu)(1-\pi_0)+1}(\pi_0 + (1-\pi_0)(1-\mu)c)$, which makes the firm indifferent between p_0 and $\hat{p}_1$. Note also that $\bar{s}_0 > (1-\mu)(1-\pi_0)c$ because $\pi_0 > c$, which means that if $s \geq \bar{s}_0$, the firm will never have to consider $\hat{p}_{k_{fb}}$.[2]

Now, $L(\hat{s}) \geq R(\pi_0, \mu)$ guarantees that the firm will set $p = p_0$. We can verify that $s_0 > \bar{s}_0$ as long as $\pi_0 > c$ (this is assumed), which means that $\hat{s} := \max\{\bar{s}_0, s_1\} < s_0$. One can check that $\bar{s}_0 \geq s_1$ for all $c \geq 0$ if and only if $\pi_0 \geq \frac{\sqrt{\mu}-\mu}{1-\mu}$.[3] However, if $\pi_0 < \frac{\sqrt{\mu}-\mu}{1-\mu}$, $s_1 > \bar{s}_0$ for low values of production cost.

If $L(\hat{s}) < R(\pi_0, \mu)$, we can do better than $\hat{s}$ because the firm's profit function is first decreasing, meaning that the consumer does not have to constrain the firm as much by choosing a high search cost. Since $\Pi_1 < \Pi_0$, there exists $s^* \leq s_1 \leq \hat{s}$ such that, using Proposition 1.2, the firm will set $p_0 = \frac{s}{(1-\mu)(1-\pi_0)}$ for any $s \geq s^*$. Thus, the consumer-optimal choice is the smallest such search cost, s^*. □

[1]Here, we have broken the tie in the firm's decision so that it will make the consumer buy if it is indifferent between the two strategies.

[2]The explanation is that when $s > (1-\mu)(1-\pi_0)c$, it is not socially optimal to search at $k = 0$, which means that it is also not optimal for the firm to induce search more than once (since it is the residual claimant to the full social surplus at $\hat{p}_{k+1}$ for any k).

[3]Note that this condition guarantees that $\bar{s}_0 \geq s_1$ for **all** c.

A.2 Proof of Proposition 1.4

Proposition 1.4. Assume that $c < \pi_0$. Then there exist two thresholds, $\bar{s}$ and $\underline{v}^*(s)$, such that, for $s < \bar{s}$, the firm will make the consumers with $v \geq \underline{v}^*(s)$ search. For $s \geq \bar{s}$, the consumers with $v \geq \underline{v}^*(s)$ will purchase the good at a price that makes them indifferent between buying and searching. All consumers with $v \geq \underline{v}^*(s)$ make the same decision: buy or search. Similarly, everyone with $v < \underline{v}^*(s)$ exits.[4]

Proof. Now, given a price p, a consumer with valuation v will search if and only if $V_S(v) \geq \max\{V_B(v), V_E(v)\}$. The value of search is higher than the value of buying if and only if $p > \hat{p}_B \equiv \frac{s}{(1-\mu)(1-\pi_0)}$, where $\hat{p}_B$ is the highest price that still makes the consumer buy the good without search. Note how this does not depend on the consumer's type as the expected consumption value is the same whether or not you search or buy. Consumer v prefers searching to exiting if and only if $v \geq \underline{v}(p) \equiv \frac{(\pi_0+(1-\pi_0)\mu)p+s}{\pi_0}$. In other words, if one consumer finds searching optimal, then every consumer whose value is high enough finds searching optimal and no one finds buying optimal.

The firm cannot just maximize the profits assuming that everyone will buy the good because the consumers have the option to search. Therefore, the firm has to choose between making the consumers search and making them buy. Above we found that, as long as $p \leq \hat{p}_B$, the consumers will indeed buy the good without search. Therefore, the firm's optimal purchase price is:

$$p_B^* = \begin{cases} \hat{p}_B \equiv \frac{s}{(1-\mu)(1-\pi_0)}, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \leq \frac{\pi_0+c}{2} \\ p_0^* \equiv \frac{\pi_0+c}{2}, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \geq \frac{\pi_0+c}{2}. \end{cases}$$

In other words, should the firm decide to sell the good without search, it will always set the maximum price that makes the consumer buy instead of searching ($\hat{p}_B$) unless this price is higher than the optimal price with no information ($p_0^* = \frac{\pi_0+c}{2}$ from the case with no signals). This leads to the purchase profit function:

$$\Pi_B^* = \begin{cases} \frac{\pi_0+c}{(1-\mu)(1-\pi_0)\pi_0}s - \frac{s^2}{(1-\mu)^2(1-\pi_0)^2\pi_0} - c, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \leq \frac{\pi_0+c}{2} \\ \frac{(\pi_0-c)^2}{4\pi_0}, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \geq \frac{\pi_0+c}{2}. \end{cases}$$

Therefore, type v consumer will get (given that she is willing to buy):

$$V_B^*(v) = \begin{cases} \pi_0 v - \frac{s}{(1-\mu)(1-\pi_0)}, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \leq \frac{\pi_0+c}{2} \\ \pi_0 v - \frac{\pi_0+c}{2}, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \geq \frac{\pi_0+c}{2}. \end{cases}$$

[4] We can express the threshold as $\bar{s} \equiv \frac{B-\sqrt{B^2-4AC}}{2A}$, where $A = 4+(1-\mu)^2(1-\pi_0)^2$, $B = 4(1-\mu)(1-\pi_0)(\pi_0+c)+2(1-\mu)^2(1-\pi_0)^2(\pi_0-(\pi_0+(1-\pi_0)\mu)c)$, and $C = (1-\mu)^2(1-\pi_0)^2(\pi_0-(\pi_0+(1-\pi_0)\mu)c)^2+4(1-\mu)^2(1-\pi_0)^2\pi_0 c$.

This translates into an average utility of:

$$V_B^* = \begin{cases} \frac{\pi_0}{2} + \frac{s^2}{2(1-\mu)^2(1-\pi_0)^2\pi_0} - \frac{s}{(1-\mu)(1-\pi_0)}, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \leq \frac{\pi_0+c}{2} \\ \frac{(\pi_0-c)^2}{8\pi_0}, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \geq \frac{\pi_0+c}{2}. \end{cases}$$

And, finally, we can compute the social surplus:

$$W_B^* = \begin{cases} \frac{\pi_0}{2} - \frac{s^2}{2(1-\mu)^2(1-\pi_0)^2\pi_0} - \left(1 - \frac{s}{(1-\mu)(1-\pi_0)\pi_0}\right) c, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \leq \frac{\pi_0+c}{2} \\ \frac{3(\pi_0-c)^2}{8\pi_0}, & \text{if } \frac{s}{(1-\mu)(1-\pi_0)} \geq \frac{\pi_0+c}{2}. \end{cases}$$

Now, if the firm decides to make the consumers search, its profit function will be: $\Pi_S = (\pi_0 + (1-\pi_0)\mu)(1 - \underline{v}(p))(p - c)$, because only the consumers who expect a positive utility will search (those above $\underline{v}(p) = \frac{(\pi_0+(1-\pi_0)\mu)p+s}{\pi_0}$) and only a fraction of them will obtain a good signal $(\pi_0 + (1-\pi_0)\mu)$. Maximizing this yields the optimal search price[5]:

$$p_S^* = \frac{\pi_0 - s + (\pi_0 + (1-\pi_0)\mu)c}{2(\pi_0 + (1-\pi_0)\mu)}.$$

This means that the equilibrium threshold value for those who search is:

$$\underline{v}_S^* = \frac{\pi_0 + s + (\pi_0 + (1-\pi_0)\mu)c}{2\pi_0},$$

and those with $v < \underline{v}_S^*$ will exit. The firm's maximized search profit is:

$$\Pi_S^* = \frac{(\pi_0 - s - (\pi_0 + (1-\pi_0)\mu)c)^2}{4\pi_0}.$$

Finally, this means that consumer v ($v \geq \underline{v}_S^*$) gets:

$$V_S^*(v) = \pi_0 v - \frac{\pi_0 + s + (\pi_0 + (1-\pi_0)\mu)c}{2},$$

and the consumers as a whole get:

$$V_S^* = \frac{(\pi_0 - s - (\pi_0 + (1-\pi_0)\mu)c)^2}{8\pi_0},$$

which means that the social surplus is:

$$W_S^* = \frac{3(\pi_0 - s - (\pi_0 + (1-\pi_0)\mu)c)^2}{8\pi_0}.$$

Whether or not the consumers will be made to search depends on the firm's profits in the two cases (Π_S^* versus Π_B^*). However, because we have assumed that $\pi_0 \geq c$, $\Pi_B \geq \Pi_S$ always holds when

[5]As long as $\pi_0 \geq s + (\pi_0 + (1-\pi_0)\mu)c$, because otherwise optimal price is less than production cost.

$\frac{s}{(1-\mu)(1-\pi_0)} > \frac{\pi_0+c}{2}$ – that is, for high values of the search cost (and not too high production costs), the firm prefers selling the good without search. So, if we then assume that $\frac{s}{(1-\mu)(1-\pi_0)} \leq \frac{\pi_0+c}{2}$, this implies $\Pi_S^* > \Pi_B$ if and only if the condition of the Proposition holds (if and only if $s < \bar{s}$, where $\bar{s}$ is a function of A, B, and C). □

A.3 Results and Proofs for Section 1.6

In this section, I will prove some of the results whose proofs were relegated into the Appendix, and will also introduce a few extra corollaries.

Proposition 1.10 Assume $s \leq \frac{(\pi_1-\pi_0)\pi_0}{\pi_1}$, and $c \leq \pi_0$. Let $A(\pi_1, c) \equiv \frac{\pi_0(\pi_1-\pi_0)}{2\pi_1-\pi_0} + \frac{(\pi_1-\pi_0)^2}{\pi_1(2\pi_1-\pi_0)}c$. Then, if search cost is fixed, the consumer's optimal choice of search technology is $\pi_1^*(s) = \min\{1, \hat{\pi}_1(s)\}$, where $\hat{\pi}_1(s)$ solves $s = A(\hat{\pi}_1(s), c)$. If both s and π_1 are for the consumer to choose, she will choose $\pi_1^* = 1$ and $s^* = A(1, c)$.

Proof. To solve this game, let us start from the end. Given a technology π_1 and a price p, the consumer will search if and only if $\frac{\pi_0}{\pi_1}(\pi_1 - p) - s \geq \max\{0, \pi_0 - p\} \Leftrightarrow p \in \left(\frac{\pi_1}{\pi_1-\pi_0}s, \frac{\pi_0-s}{\pi_0}\pi_1\right]$.[6] The assumption that $s \leq \frac{(\pi_1-\pi_0)\pi_0}{\pi_1}$ guarantees that there are some prices that satisfy the above inequalities.[7]

Thus, given a search technology π_1, the firm will essentially have the choice between two prices: $p_S \equiv \frac{\pi_0-s}{\pi_0}\pi_1$ ("search price" which makes the consumer search) and $p_B \equiv \frac{\pi_1}{\pi_1-\pi_0}s$ ("purchase price" which makes the consumer buy without search). The trade-off from the firm's perspective is that the search price gives a higher profit conditional on selling but the probability of a sale is less than one, whereas the purchase price is lower but leads to a sale with certainty. Therefore, the firm prefers the search price if and only if: $\frac{\pi_0}{\pi_1}(\frac{\pi_0-s}{\pi_0}\pi_1 - c) \geq \frac{\pi_1}{\pi_1-\pi_0}s - c \Leftrightarrow s \leq A(\pi_1, c) \equiv \frac{\pi_0(\pi_1-\pi_0)}{2\pi_1-\pi_0} + \frac{(\pi_1-\pi_0)^2}{\pi_1(2\pi_1-\pi_0)}c$. This leads to the pricing scheme:

$$p^* = \begin{cases} \frac{\pi_0-s}{\pi_0}\pi_1, & \text{if} \quad s < A(\pi_1, c) \\ \frac{\pi_1}{\pi_1-\pi_0}s, & \text{if} \quad s \geq A(\pi_1, c). \end{cases}$$

One thing to note is that the assumption $c \leq \pi_0$ guarantees that $A(\pi_1, c) \leq \frac{\pi_0}{\pi_1}(\pi_1 - \pi_0)$, meaning that there are parameter values such that it is best to sell the product without search. Similarly, the assumption that $s \leq \frac{(\pi_1-\pi_0)\pi_0}{\pi_1}$ makes it possible for the consumer to search. Now, the consumer's

[6]There is a slight abuse of notation here because the inequalities are weak but I assume that the lower bound on price is strict. This is just a matter of convention because we can always make the consumer strictly prefer one action over another by changing the price infinitesimally. Here I have assumed that if the consumer is indifferent between buying and searching, she will buy; and if she is indifferent between searching and exiting, she will search.

[7]If this was not the case, even for $\pi_1 = 1$, the firm would always set $p = \pi_0$ and the consumer would buy without using her search technology, giving her no expected utility.

utility, given π_1, can be written as follows:

$$V^*(\pi_1) = \begin{cases} 0, & \text{if} \quad s < A(\pi_1, c) \\ \pi_0 - \frac{\pi_1}{\pi_1 - \pi_0}s, & \text{if} \quad s \geq A(\pi_1, c). \end{cases} \tag{A.1}$$

We see that the consumer always gets zero utility when she is made to search. Note that her best-case utility is $V_{max}(\pi_1) = \pi_0 - \frac{\pi_1}{\pi_1 - \pi_0}A(\pi_1, c) = (\pi_0 - c)\frac{\pi_1 - \pi_0}{2\pi_1 - \pi_0}$. This shows that, if the consumer was free to choose both the technology, π_1, and her search cost, s, she would maximize her utility by choosing $\pi_1^* = 1$, and $s^* = A(1, c) = \frac{1-\pi_0}{2-\pi_0}(\pi_0 + (1 - \pi_0)c)$.[8] This is the lowest possible search cost that makes the firm choose the purchase price (making the consumer search is too expensive because the firm has to pay for it), given that the consumer is using the best possible technology, $\pi_1^* = 1$.

What if the consumer cannot choose her search cost but only the technology? Observing that, for $s > A(\pi_1, c)$, $V^*(\pi_1)$ is increasing in π_1, and that $\frac{\partial A(\pi_1, c)}{\partial \pi_1} > 0$, we obtain the result that the consumer's optimal search technology is $\pi_1^* = \min\{\hat{\pi}_1(s), 1\}$, where $\hat{\pi}_1(s)$ solves: $s = A(\hat{\pi}_1(s), c)$. The consumer benefits from increasing π_1 as long as the firm is still willing to sell the good without search, leaving the consumer with positive utility. □

To see how this works, consider the following corollary:

Corollary A.1. *Assume $c = 0$. Then the consumer-optimal posterior as a function of search cost is*

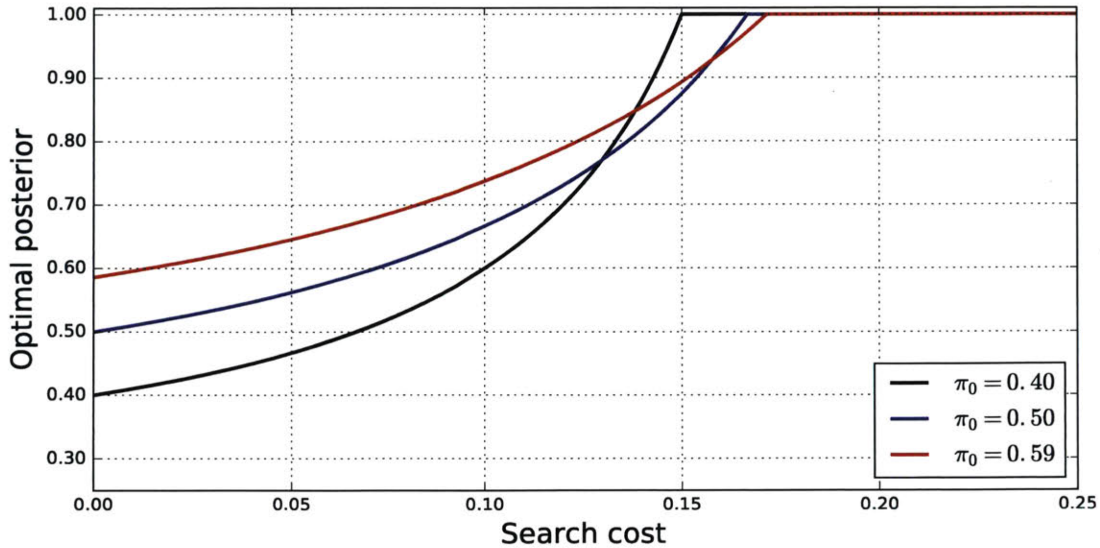
$$\pi_1^*(s) = \begin{cases} \frac{\pi_0^2 - \pi_0 s}{\pi_0 - 2s}, & \textit{if} \quad 0 < s \leq \frac{\pi_0(1-\pi_0)}{2-\pi_0} \\ 1, & \textit{if} \quad \frac{\pi_0(1-\pi_0)}{2-\pi_0} \leq s < \pi_0(1 - \pi_0). \end{cases}$$

Proof. We need $s = A(\pi_1^*, 0) \Leftrightarrow \pi_1 = \frac{\pi_0^2 - \pi_0 s}{\pi_0 - 2s}$ but this is less than one only if $s \leq \frac{\pi_0(1-\pi_0)}{2-\pi_0}$. Otherwise the consumer should obtain the perfectly precise signal. □

To understand the corollary, consider Figure A-1 below. We plot the consumer's optimal (good) posterior as a function of the search cost for three priors when there are no production costs. We see that the consumer will not choose the perfectly informative experiment unless her search cost is high enough. The threshold search cost for choosing the perfect experiment is increasing in the prior until $\pi_0 = 2 - \sqrt{2}$ (the red line), after which it starts to decrease (not seen in the figure). An interesting observation is that there are values of s for which the optimal posterior is higher under $\pi_0 = 0.4$ than under $\pi_0 = 0.59$. Thus, it is not the case that the optimal posterior increases monotonically in the prior for all s. If the consumer becomes more convinced of the product's high quality ex ante, she may need to use a less informative experiment (meaning a lower π_1) to keep the firm from setting a high price.

[8] It can be easily checked that $s^* \leq \frac{(\pi_1^* - \pi_0)\pi_0}{\pi_1^*}$ as long as $c \leq \pi_0$.

Figure A-1: Optimal posterior as a function of search cost for three priors, assuming $c = 0$.



The corollary means that the consumer's equilibrium utility is

$$V^*(\pi_1^*(s)) = \begin{cases} s, & \text{if} \quad s \leq \frac{\pi_0(1-\pi_0)}{2-\pi_0} \\ \pi_0 - \frac{s}{1-\pi_0}, & \text{if} \quad \frac{\pi_0(1-\pi_0)}{2-\pi_0} \leq s \leq \pi_0(1-\pi_0). \end{cases}$$

If she could choose both the cost of search and the technology, she would go with $s = \frac{\pi_0(1-\pi_0)}{2-\pi_0}$ and $\pi_1 = 1$. One way to understand the consumer's optimal policy and her value function, is to remember that she is the one who chooses the search technology, and the firm observes this choice when setting its price. The consumer will not choose the socially optimal search technology ($\pi_1 = 1$) when her search cost is low because if she did, the firm would make her use it and get no surplus. By choosing an intermediate search technology, the consumer forces the firm to set a low price that induces no search. In effect, the consumer is shooting herself in the leg to avoid having to conduct expensive search. Note also that the distortion is decreasing in the search cost, which is intuitive, because the higher the search cost is, the less of an incentive the firm has to make the consumer search. Thus, the consumer does not have to tie her hands as much when her search cost is high. All this leads to a value function that is increasing in the search cost at first but then decreases after the search cost crosses the critical threshold, $\hat{s} = \frac{\pi_0(1-\pi_0)}{2-\pi_0}$. After the threshold, increasing the search cost does not allow the consumer to improve her search technology because it is already set at its maximum value of 1. Therefore, the good will always be sold and higher search costs are detrimental to the consumer (because they reduce her bargaining power, as her ability to search is reduced).

Because the good is always sold without search, social surplus is always exactly π_0 (assuming no production cost, of course). However, the firm's profit, just like the consumer's value, is not monotone in the search cost:

$$\Pi^*(\pi_1(s)) = \begin{cases} \pi_0 - s, & \text{if} \quad s < \frac{\pi_0(1-\pi_0)}{2-\pi_0} \\ \frac{s}{1-\pi_0}, & \text{if} \quad s \geq \frac{\pi_0(1-\pi_0)}{2-\pi_0}. \end{cases} \tag{A.2}$$

Thus, the firm's profit is V-shaped, whereas the consumer's utility is the inverse of that. The consumer's optimum is the firm's minimum, and vice versa.

Proposition 1.12 Assume that $s < 1-E$ and $c < E$. The consumer's optimal search technology characterized by the good posterior, $E_G^* = 1$, and the probability of a good signal, $q^* = \min\{E, \hat{q}\}$, where $\hat{q}$ solves $s = A(\hat{q}, 1, E, c) \equiv q - \frac{E}{2-q} + \frac{(1-q)^2}{2-q}c$.

Proof. Now, given the search technology (q, E_G), the consumer will search if:

$$q(E_G - p) - s > \max\{0, E - p\} \quad \Leftrightarrow \quad p \in \left(\frac{E - qE_G}{1-q} + \frac{s}{1-q},\ E_G - \frac{s}{q}\right).$$

Thus, the firm has to choose between two prices: $p_S = E_G - \frac{s}{q}$, and $p_B = \frac{E-qE_G}{1-q} + \frac{s}{1-q}$, where the former makes the consumer search and the latter makes her buy. The firm will choose p_S if and only if:

$$\begin{aligned} \frac{E - qE_G}{1-q} + \frac{s}{1-q} - c &< q(E_G - \frac{s}{q} - c) \\ \Leftrightarrow s &< \underbrace{qE_G - \frac{E}{2-q} + \frac{(1-q)^2}{2-q}c}_{\equiv A(q, E_G, E, c)}. \end{aligned} \tag{A.3}$$

This allows us to write the consumer's utility, given the parameters (E, c) and the chosen technology (q, E_G):

$$V(q, E_G, E, c) = \begin{cases} 0, & \text{if} \quad s < A(q, E_G, E, c) \\ \frac{q}{1-q}(E_G - E) - \frac{s}{1-q}, & \text{if} \quad s \geq A(q, E_G, E, c), \end{cases}$$

where the function $A(\cdot)$ is as defined in (A.3). This shows, again, how the firm gets all the surplus when the consumer is made to search, and how the firm sometimes finds it optimal to give the consumer some surplus in exchange for her not searching. We see that the consumer's utility is strictly positive if and only if $A(q, E_G, E, c) \leq s < q(E_G - E)$, which is possible only when $c < E$.

Because both $A(q, E_G, E, c)$ and $V(q, E, G, E, c)$ are strictly increasing in E_G, we want to set it as high as possible without violating $s \geq A$ (for a given q, that is). Thus, we want to set E_G so that $s = A(q, E_G, E, c)$, which yields: $E_G^*(q, s) = \min\left\{\frac{E-(1-q)^2 c}{q(2-q)} + \frac{s}{q}, 1\right\}$.

Take then any $q < E$.[9] If the optimal posterior $E_G^*(q, s) < 1$, then we can show that

[9]Because if $q > E$, we need either $E_B < 0$ or $E_G > 1$, which is impossible. Furthermore, if $q = E$, then the only

$V(q, E_G^*(q,s), E, c) = \frac{1-q}{2-q}(E-c)$, which is decreasing in q. In particular, this implies that we should reduce q, which increases both $E_G^*(q,s)$ and $V(q, E_G^*(q,s), E, c)$.[10] If, on the other hand, $E_G^*(q,s) = 1$, we have $s \geq A(q,1,E,c)$. However, now $\partial V(q,1,E,c)/\partial q > 0$ and $\partial A(q,1,E,c)/\partial q > 0$ (these follow from the assumptions $s < 1-E$ and $c < E$, respectively), which imply that we should increase q as much as we can, as long as $s \geq A(q,1,E,c)$. Putting everything together implies that the consumer maximizes her utility by choosing $E_G^* = 1$ and $q^* = \min\{E, \hat{q}\}$ such that $s = A(\hat{q}, 1, E, c)$. We can even write $\hat{q}$ analytically:

$$\hat{q} = 1 - \frac{\sqrt{s^2 + 4(1-c)(1-E-s)} - s}{2(1-c)}$$

This concludes the proof. □

Corollary A.2. *If the consumer is not free to choose q, and she can only pick E_G, then the optimal technology may have $E_G^* < 1$.*

Proof. This is obvious because $E_G^*(q,s) = \min\left\{\frac{E-(1-q)^2c}{q(2-q)} + \frac{s}{q}, 1\right\}$ but we cannot adjust q, so it may be that $s = A(q, E_G^*, E, c)$ at $E_G^* < 1$. □

For example, if we assume that $E = \frac{1}{2}$ and restrict the consumer to consider only those experiments that have $q = \frac{1}{2}$, the condition for $E_G^* < 1$ becomes $s < \frac{1}{2} - \frac{2}{3}E$. This means that for low values of the search cost ($s < \frac{1}{6}$), the consumer's optimal experiment has $E_G^* < 1$ and $E_B^* > 0$. More specifically, $E_G^* = \frac{1/2}{3/4} + 2s = \frac{2}{3} + 2s$, and $E_B^* = \frac{1}{3} - 2s$. The consumer could easily (and for free) obtain a better technology but this would end up hurting her through the pricing of the firm.

A.4 Optimal Distribution of Posteriors With Search Costs

Complementing Section 1.5, this subsection characterizes the optimal posterior distribution, $G_s(v)$ as a function of the search cost when prior is $F(v)$. We need the two to have the same mean, μ. Note that we are focusing, without loss of generality, on the case where each v leads to a distribution of signals $\sigma \in [0,1]$ and each σ is thought as the posterior expectation of v. Thus, $G_s(v)$ is the distribution of posterior expectations. As in Section 1.5, we know that the consumer will prefer buying to searching as long as $U_B(p) \geq U_S(p)$, where $U_S(p) = \int_p^1 (v-p)dG(v) - s = \int_p^1 v dG(v) - [1-G(p)]p - s$, and $U_B(p) = \int_0^1 (v-p)dF(v) = \int_0^1 v dF(v) - p$. This implies that the consumer will prefer buying over searching if and only if:

$$\begin{aligned} & U_S(p) \leq U_B(p) \\ \Leftrightarrow & \int_p^1 G(v)dv \geq 1 - \mu - s \\ \Leftrightarrow & \int_0^p G(v)dv \leq s, \end{aligned}$$

option is $E_G = 1$ and $E_B = 0$.

[10] $\partial E_G^*/\partial q = \frac{2(1-q)(c-E)}{q^2(2-q)^2} - \frac{s}{q^2} < 0$ because $c < E$.

where μ is the prior expectation, and we get the equivalence by integrating by parts. This inequality implies the highest price $p_B(s,G)$ the consumer is willing to accept for given $G(v)$ and s. Now, the firm can also set a higher price and collect the search profit $\Pi_S(p) = [1-G(p)]p$, which is maximized at some p^* and gives $\Pi_S^*(G) := \Pi_S(p^*)$. This means that the lowest $p_B(s,G)$ the firm is willing to set is $p_B(s,G) = \Pi_S^*(G)$. Therefore, the consumer-optimal posterior distribution is characterized by two equations:

$$\min_{G(\cdot)} \Pi_S^*(G) \quad \text{s.t.} \quad (i) \quad \int_0^{\Pi_S^*(G)} G(v)dv = s \tag{A.4}$$

$$(ii) \quad \int_0^1 vdG(v) = \mu. \tag{A.5}$$

A.5 A Bad News Model

In this section, I will briefly go over a slightly different model and obtain results similar to the models covered earlier. In particular, the consumer's utility is shown to be non-monotone in her search cost. We will also see that the firm always wants the consumer to either search extensively or not at all.

Let the setup be similar to the main model in that news arrive exponentially. To be more specific, the product can be of high or low quality (yielding utilities 1 and 0, respectively). The consumer has a prior π_0 on the quality being high, but she can improve this belief by searching for information at a unit cost of $\tilde{s} \geq 0$. If the consumer searches for information, bad news arrive at rate $\lambda > 0$ only when the quality is low.[11] If the product is good, nothing is observed.

Proposition A.1. *Given a price p, search cost $s \equiv \frac{\tilde{s}}{\lambda}$, and an initial belief π_0, the consumer will search until her belief exceeds $\hat{\pi} \equiv 1 - \frac{s}{p}$. Her initial value of search can be written as:*

$$V_S(p) = \pi_0(1-p) - \left(1 + \pi_0 \frac{p-2s}{p-s} \ln\left(\frac{p-s}{L_0 s}\right)\right) s.$$

There exists a threshold $\hat{s}$ such that the firm will make the consumer buy the product for $s \geq \hat{s}$, giving her weakly positive utility, and the firm will make her search for information for $s < \hat{s}$, giving her no utility. The maximal utility is obtained at $s = \hat{s} > 0$. If $\hat{s} \geq (1-\pi_0)\pi_0$, the maximizer is not unique and the consumer gets zero surplus everywhere.

Proof. Assuming that the probability the consumer assigns on the product being of high quality at time $t \geq 0$ is $\pi_t \equiv \mathbb{P}(H)$, the belief can be shown to evolve as follows: $\pi_t + d\pi_t = \frac{\pi_t}{\pi_t + (1-\pi_t)(1-\lambda dt)}$. Taking the limit ($dt \to 0$), one obtains:

$$\dot{\pi}_t \equiv d\pi_t/dt = \lambda \pi_t (1-\pi_t) \tag{A.6}$$

$$\Rightarrow \pi_t = \frac{L_0 e^{\lambda t}}{1 + L_0 e^{\lambda t}}, \tag{A.7}$$

[11]Here it is informative to use not only the expected search cost $s = \tilde{s}/\lambda$, but also the arrival rate, λ.

where $L_0 = \frac{\pi_0}{1-\pi_0}$ is the (given) prior likelihood ratio. Note that, in a time interval of length dt, the searching consumer will obtain a bad signal and her belief will fall all the way down to 0 with probability $(1-\pi_t)\lambda dt$.

This allows us to determine the threshold belief, $\hat{\pi}$, such that for beliefs higher than the threshold, the consumer prefers to buy the good (instead of searching for more information). This threshold is obtained simply by equating the expected utility from buying now to the expected utility from searching for a "small amount of time" (dt) and then buying:

$$\begin{aligned} \hat{\pi} - p &= -\tilde{s}dt + (1-\hat{\pi})\lambda dt \cdot 0 + (1-(1-\hat{\pi})\lambda dt)(\hat{\pi} + d\hat{\pi} - p) \\ \Leftrightarrow \hat{\pi} &= 1 - \frac{s}{p}, \end{aligned} \tag{A.8}$$

where $s \equiv \tilde{s}/\lambda$ is the "normalized search cost". This buying threshold works as long as $\hat{\pi} > p$, which requires that $p \in (\frac{1}{2} - \sqrt{\frac{1}{4} - s}, \frac{1}{2} + \sqrt{\frac{1}{4} - s})$ and $s \le \frac{1}{4}$. Otherwise, exiting will yield a higher utility. Setting $\pi_0 = \hat{\pi}$ and solving for the price, gives us $\boxed{p_0 = \frac{s}{1-\pi_0}}$. It is easy to see that for prices $p \le p_0$, the purchase threshold is lower than the consumer's prior, implying that the consumer should either purchase or exit instantaneously (depending on whether $\pi_0 \ge p_0$ or not).

If we let $\hat{t}(\hat{\pi}) \equiv \min\{t \mid \pi_t \ge \hat{\pi}, t \ge 0\}$, this will be the first time the consumer's belief crosses the threshold (it is unique, and exists for all $\hat{\pi} \in (0,1)$). Using (A.7), we can, in fact, write $\hat{t} \equiv \hat{t}(\hat{\pi}) = \frac{1}{\lambda}\ln(\frac{p-s}{L_0 s})$, or $e^{\lambda\hat{t}} = \frac{p-s}{L_0 s}$. Thus, the time spent searching is increasing in the price and decreasing in the search cost (holding the price constant).

We are now ready to compute the consumer's value function at π_0, knowing that she will buy at the prior if $p \le \min\{p_0, \pi_0\} = \min\{\frac{s}{1-\pi_0}, \pi_0\}$. This also shows why we need $s < (1-\pi_0)\pi_0$ for the consumer to get a strictly positive utility because otherwise she would accept $p = \pi_0$ and get nothing.[12]

To compute the value of searching until the belief hits $\hat{\pi}$ or until there is a bad signal, one needs to compute the costs and benefits of search. To this end, expected costs can be written as:

$$\begin{aligned} \mathbb{E}[\text{costs}] &= \left(\pi_0\hat{t} + (1-\pi_0)\int_0^{\hat{t}} t\lambda e^{-\lambda t}dt\right)\tilde{s} && \Big| \text{ integrate by parts} \\ &= \left(\pi_0\lambda\hat{t} + (1-\pi_0) - (1-\pi_0)(1+\lambda\hat{t})e^{-\lambda\hat{t}}\right)s && \Big| \; e^{\lambda\hat{t}} = \frac{p-s}{L_0 s} \\ &= \left(1 + \pi_0\frac{p-2s}{p-s}\ln(\frac{p-s}{L_0 s}) - \pi_0\frac{p}{p-s}\right)s, \end{aligned}$$

because the consumer will surely search until $\hat{t}$ if the product is good, but, if it is bad, she may

[12] In this case, the firm would either set a price equal to the prior or such a high price that the consumer would just be willing to search but get zero in expectation.

drop out at any point if bad news arrive. Expected benefit can be written as:

$$\begin{aligned}\mathbb{E}[\text{benefit}] &= \mathbb{P}(\text{purchase})(\hat{\pi} - p) \qquad \Big| \quad \hat{\pi} = 1 - s/p \\ &= \left(\pi_0 + (1-\pi_0)e^{-\lambda\hat{t}}\right)\left(1 - \frac{s}{p} - p\right) \\ &= \pi_0\left(1 - \frac{p^2}{p-s}\right).\end{aligned}$$

Combining the benefits and the costs, and simplifying, we can write the consumer's search value as:

$$V_S(p) = \pi_0(1-p) - \left(1 + \pi_0\frac{p-2s}{p-s}\ln\left(\frac{p-s}{L_0 s}\right)\right)s, \quad \text{for} \quad p \geq p_0 \equiv \frac{s}{1-\pi_0}. \tag{A.9}$$

Remember that the above assumes that the price is such that the consumer will find it beneficial to start searching (and keep doing so until her belief hits $\hat{\pi}$ or 0).[13]

The firm will then choose its price, knowing that the consumer will search until her belief hits the threshold, $\hat{\pi}$. If the firm does not want the consumer to search at all, it can set its price at $p = \min\{\frac{s}{1-\pi_0}, \pi_0\}$. If the firm makes the consumer search, note first that the value of searching until $\hat{\pi} > \pi_0$ is strictly decreasing in p, and that it is positive at p_0 only if $s \leq (1-\pi_0)\pi_0$. If search cost is smaller than this threshold, there exists a price, $p^* \geq p_0$, such that $V_S(p^*) = 0$. If search cost is greater than the threshold, the firm will always set $p = \pi_0$. Let us then find the firm's optimal policy, given $p \geq p_0$:

$$\max_p \Pi(p) = \max_p\,(\pi_0 + (1-\pi_0)e^{-\lambda\hat{t}})(p-c) = \max_p\,\pi_0\frac{p(p-c)}{p-s} \tag{A.10}$$

$$\Rightarrow \Pi'(p) = \pi_0\frac{p(p-2s)+cs}{(p-s)^2} \geq 0 \quad \forall p \geq 2s. \tag{A.11}$$

This means that the firm should choose between two prices, p_0 and p^*, because its profit is either always increasing in p or V-shaped, so that the intermediate values cannot be optimal. In fact, remembering that $p_0 = \frac{s}{1-\pi_0} \geq 2s$ whenever $\pi_0 \geq \frac{1}{2}$, this shows that the firm's profit is increasing in p whenever $\pi_0 \geq \frac{1}{2}$. Thus, the firm will always extract all consumer surplus when $\pi_0 \geq \frac{1}{2}$, even if production cost is zero. If there is a strictly positive cost of production, the firm's profit function will be increasing in p even for $p < 2s$, so that the consumer's best-case utility may be zero even for some $\pi_0 < \frac{1}{2}$. This is formalized in Proposition A.2.

Purchase price, p_0, will be optimal if and only if:

$$\frac{s}{1-\pi_0} - c \geq \pi_0\frac{p^*(p^*-c)}{p^*-s}. \tag{A.12}$$

[13]Note that if the consumer decides to search until $\hat{\pi}$ at π_0, then she will choose to continue searching until $\hat{\pi}$ at every $\pi \in (\pi_0, \hat{\pi})$, as long as she obtains no negative signals.

Otherwise the search price, p^*, will maximize profits. One can show that p^* is decreasing in s (essentially because $V_S(p)$ is decreasing in both p and s, meaning that p^* has to decrease if s increases to keep $V_S(p^*) = 0$). Furthermore, the firm's search profits $\Pi(p^*)$ are decreasing in s because a higher s forces the firm to charge a lower price even though it would like to set $p = 1$.[14] On the other hand, when p^* is optimal, the firm gets the full social surplus, which is naturally decreasing in the search cost. Therefore, there exists a value of the search cost, $\hat{s}$, such that for $s \geq \hat{s}$, the left-hand side of (A.12) is at least as high as the right-hand side (because LHS is increasing in s, and RHS is decreasing in s). However, it can be that $\hat{s} \geq \pi_0(1-\pi_0)$ which leads the firm to set $p = \pi_0$, giving the consumer no utility.

The consumer obtains her maximal, strictly positive utility at $s = \hat{s}$ whenever $\hat{s} < \pi_0(1-\pi_0)$. Her utility is zero for $s < \hat{s}$ and positive but decreasing for $s \in [\hat{s}, \pi_0(1-\pi_0))$. We can write this out:

$$V^* = \begin{cases} 0, & \text{if} \quad s < \hat{s} \\ \pi_0 - \frac{s}{1-\pi_0}, & \text{if} \quad \hat{s} \leq s < \pi_0(1-\pi_0) \\ 0, & \text{if} \quad \pi_0(1-\pi_0) \leq s. \end{cases} \tag{A.13}$$

This concludes the proof. □

Proposition A.2. *Assume that $c = 0$. The consumer's utility can be strictly positive only when $\pi_0 < \frac{1}{2}$. If $\pi_0 \geq \frac{1}{2}$, the consumer always gets 0, for all values of search cost.*

The proposition makes sense because, when $c = 0$, $\pi_0 \geq \frac{1}{2}$ implies that $p_0 \geq 2s$, meaning that the firm will always prefer p^* when $\pi_0 \geq \frac{1}{2}$. If $c > 0$, even a lower π_0 leads the firm to prefer p^* over p_0 for all s, always leaving the consumer with no surplus.

Proof. Let $\hat{s}$ be the search cost at which the firm is indifferent between p_0 and p^* (that is, $\hat{s}$ satisfies (A.12) with equality). This is the consumer-optimal search cost, as shown above. Now, we can write:

$$\begin{aligned} \hat{s} &= \frac{p^* - \sqrt{(p^*)^2 - 4(1-\pi_0)\pi_0(p^*)^2}}{2} \\ &= \frac{p^*}{2}[1 - |1 - 2\pi_0|] \\ &= \begin{cases} \pi_0 p^*, & \text{if} \quad \pi_0 \leq \frac{1}{2} \\ (1-\pi_0)p^*, & \text{if} \quad \pi_0 \geq \frac{1}{2}, \end{cases} \end{aligned} \tag{A.14}$$

where p^* itself is a function of $\hat{s}$.

If $\pi_0 < \frac{1}{2}$, then $\hat{s} = \pi_0 p^*$, and we can plug it into the consumer's value function $V_S(p)$ at p^* to

[14]Technically, $p = 1$ never leads to a sale because the consumer's posterior never reaches 1, but we can assume that $p = 1$ corresponds to $p = 1 - \epsilon$ for small $\epsilon > 0$.

get:

$$V_S(p^*) = \pi_0(1-p^*) - \pi_0 p^* + 2\pi_0^2 p^* \frac{1-2\pi_0}{1-\pi_0}\ln(L_0) = 0$$
$$\Leftrightarrow p^* = \frac{1-\pi_0}{2[1-\pi_0-\pi_0(1-2\pi_0)\ln(L_0)]} < \frac{1}{2} \leq 1-\pi_0,$$

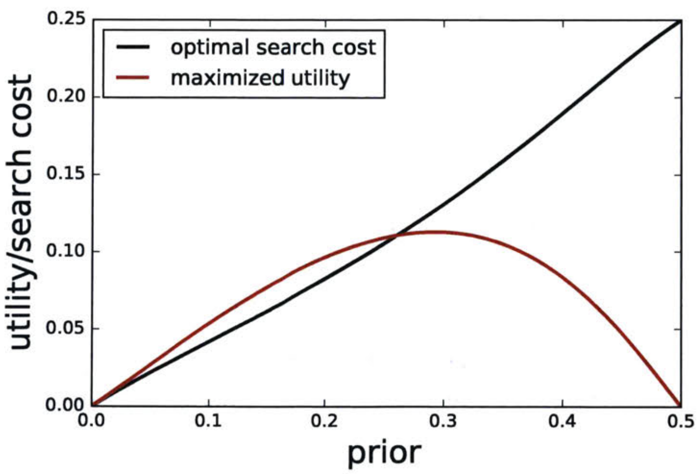
where the value function is zero at p^* by definition and the first inequality follows from the fact that $(1-2\pi_0)\ln(L_0) < 0$. But this implies that $\hat{s} = \pi_0 p^* < \pi_0(1-\pi_0)$, so that the consumer gets strictly positive utility at $\hat{s}$.

On the other hand, if $\pi_0 \geq \frac{1}{2}$, then $\hat{s} = (1-\pi_0)p^*$, and plugging it into the value function at p^* gives $V_S(p^*) = \pi_0 - p^* = 0 \Leftrightarrow p^* = \pi_0$. This means that $\hat{s} = \pi_0(1-\pi_0)$ when $\pi_0 \geq \frac{1}{2}$, so that the consumer always gets zero utility everywhere, no matter what the search cost is. □

The proposition tells us that for no production costs, if $\pi_0 < \frac{1}{2}$, then we can write the optimal search cost as $\hat{s} = \frac{(1-\pi_0)\pi_0}{2[1-\pi_0-\pi_0(1-2\pi_0)\ln(L_0)]}$. This also gives us the consumer's best-case utility $U(\hat{s}) = \pi_0 - \frac{\hat{s}}{1-\pi_0} = \pi_0\left(\frac{(1-2\pi_0)(1-2\pi_0\ln(L_0))}{2[1-\pi_0-\pi_0(1-2\pi_0)\ln(L_0)]}\right)$.

Figure A-2 plots the consumer's optimal search cost $\hat{s}$ and the resulting best-case utility as a function of the prior. The consumer's utility is always zero until $s = \hat{s}$ where it jumps up and then starts decreasing linearly in s (with a slope of $\frac{1}{1-\pi_0}$) until it hits zero again at $s = \pi_0(1-\pi_0)$. The consumer can get strictly positive utility only if $\pi_0 < \frac{1}{2}$, and nothing otherwise (because then the firm will be choosing between prices π_0 and p^*, both of which give zero expected utility).

Figure A-2: Equilibrium utilities as a function of search cost for different priors ($c = 0$).



Corollary A.3. *If $\pi_0 < \frac{1}{2}$, social surplus is a non-monotone function of the search cost. It is maximized at $s = 0$ but it is not monotonically decreasing because it jumps up at $s = \hat{s}$. If $\pi_0 \geq \frac{1}{2}$, social surplus is strictly decreasing in s.*

Proof. The proof of the corollary is simple. Social welfare is maximized at $s = 0$ because then the firm will make the consumer search for so long that we can be absolutely sure that the product is of high quality, giving an expected welfare of $S = \pi_0(1-c)$. Welfare is decreasing in s for $s < \hat{s}$ because we cannot make absolutely sure that the product is of high quality and we also incur a positive search cost. However, if $\pi_0 < \frac{1}{2}$, at $s = \hat{s}$, the firm is indifferent between making the consumer search and not but the consumer experiences a discrete jump in her utility, implying that social surplus jumps up. If $\pi_0 \geq \frac{1}{2}$, there is no jump in the consumer's utility since it is always zero. □

A.5.1 Heterogeneous Consumers

As in the main model, we can consider the continuous-time setup with heterogeneous consumers. Let the consumer types be indexed by v, where $v \sim U[0,1]$. Assume that there is no production cost ($c = 0$). Then, as in Section 1.4, it is easy to show that all consumers, who do not exit immediately, will make the same decision to buy or search:

Lemma A.5.1. *For all values of p, s, and π_0, there exists a threshold $\bar{v}(p, s, \pi_0)$ such that consumers with $v \leq \bar{v}$ exit immediately and consumers with $v > \bar{v}$ participate (search or purchase). The threshold can be expressed as follows:*

$$\bar{v}(p,s,\pi_0) = \begin{cases} \frac{p}{\pi_0}, & \textit{if} \quad p \leq p_0 := \frac{s}{1-\pi_0} \\ p + \left(\frac{1}{\pi_0} + \frac{p-2s}{p-s} \ln\left(\frac{p-s}{L_0 s}\right) \right) s, & \textit{if} \quad p > p_0. \end{cases}$$

If $p \leq p_0$, all participants purchase immediately, while if $p > p_0$, types $v > \bar{v}$ will search until their belief hits $\hat{\pi} := 1 - \frac{s}{p}$ and buy (or exit if this fails).

Proof. All consumers use the same decision rule because their beliefs evolve the same way and threshold belief $\hat{\pi}$ is the same for everyone.[15] However, some consumers want to exit because their maximized value of participation is negative (type 0 should never participate, implying that there exists a threshold $\bar{v}$ which may be greater than one). We can derive the threshold easily when $p \leq p_0$ because then everyone either buys or exits, and value of purchase is $\pi_0 v - p \geq 0 \Leftrightarrow v \geq \bar{v} := \frac{p}{\pi_0}$.

When $p > p_0$, the value function is a more complicated beast. However, we can notice that v does not matter for search costs and it only enters the benefit as $\pi_0 v$, meaning that we can write the value function of consumer v as:

$$V_v(p) = \pi_0(v-p) - \left(1 + \pi_0 \frac{p-2s}{p-s} \ln\left(\frac{p-s}{L_0 s}\right)\right) s, \quad \text{when} \quad p \geq p_0.$$

Requiring that $V_{\bar{v}}(p) = 0$ gives us $\bar{v}$. This completes the proof. □

[15] If we replace $\hat{\pi}$ with $\hat{\pi} v$ in equation A.8, v drops out.

We can now write the firm's profit function as:

$$\Pi(p) = \begin{cases} (1 - \frac{p}{\pi_0})p, & \text{if} \quad p \leq \min\{p_0, \pi_0\} \\ (1 - \bar{v})\frac{\pi_0 p^2}{p-s}, & \text{if} \quad p > p_0 \\ 0, & \text{otherwise,} \end{cases}$$

because when p is low enough, everyone who participates $(1 - \bar{v})$ will buy at that price with probability one. When p is higher, the probability that a searching consumer purchases in the end is $\frac{\pi_0 p}{p-s}$.

Proposition A.3. *For high enough search costs, the firm's optimal price is* $p^*(s) = \min\{p_0, \frac{\pi_0}{2}\}$.

Proof. For high enough search costs, the firm has to set a low search price for the consumers to search, so it will find it better to charge a price that is below or at the purchasing threshold $p_0 := \frac{s}{1-\pi_0}$.[16] Because the firm's profit in this range is $\frac{1}{\pi_0}(\pi_0 - p)p$, it will maximize its profits by setting $p^* = \frac{\pi_0}{2}$, unless this price is greater than p_0, in which case $p^* = p_0$ is optimal. □

Proposition A.4. *Assume that* $\pi_0 = \frac{1}{2}$*. Then the consumer-optimal search cost is* $s^* \approx \frac{1}{16}$*. This is also the socially optimal search cost.*

Proof. I provide no analytical proof but refer the reader to Figure A-3 below where we see the resulting equilibrium payoffs when the game is solved numerically. It is easy to see that social surplus is maximized at s^* because at that point equilibrium price p^* is increasing in s and social surplus can be written as $\pi_0 - \frac{(p^*)^2}{\pi_0}$ which is decreasing in p^*. □

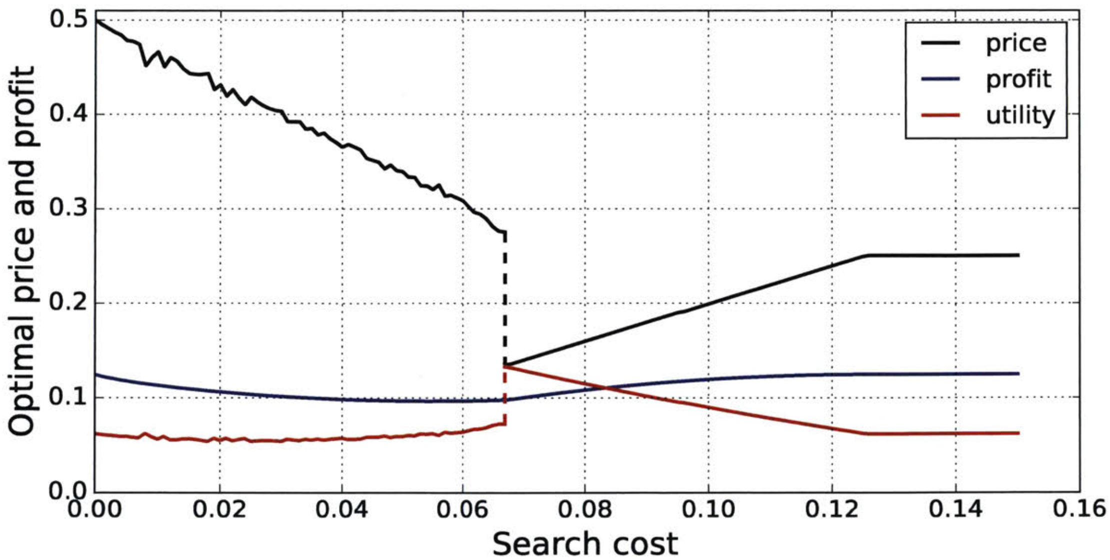
You can see the proposition in Figure A-3 where we plot the equilibrium price, profit and utility for different values of search cost (assuming that the prior is $\pi_0 = \frac{1}{2}$). For low values of s, the firm can charge a high price because all the consumers with $v > \bar{v}$ will search. However, the firm will not set a price higher than $\frac{\pi_0}{2}$ because it will essentially act like a monopolist once the consumers have done their costly search – and the monopolist wants to balance the mass of participating consumers and its price.

A.5.2 Heterogeneity in Search Costs

One may wonder how the equilibrium changes if we allow the consumers to have different search costs. If we assume that search costs are uniform on $[0, 1]$, then the firm will always just a set price $p = \pi_0$ because there will be a big mass of consumers who will not search and will just buy the product at the prior. If the firm decided to set a higher price, it would lose all the high-search-cost consumers but only gain a little bit from the others. This shows that one of the "correct" ways of introducing heterogeneity in search costs is to assume that $s \sim U[0, x]$ for some $x \geq 0$. Then a

[16] Similar to the single consumer case, optimal search price $p_S > p_0$ is decreasing in s

Figure A-3: Equilibrium as a function of search cost with heterogeneous consumers ($\pi_0 = \frac{1}{2}$).



high x means that there are fewer consumers with low s (in terms of their mass) and more people who have high search costs.[17]

Now, as long as $p \leq \pi_0$, all consumers with $s \geq (1-\pi_0)p$ will purchase because the purchasing threshold for consumer s is $\hat{\pi}(s) = 1 - \frac{s}{p}$ and this is less than the prior for high s. Everyone else will search until their belief hits $\hat{\pi}(s) := 1 - \frac{s}{p}$. What is the firm's optimal price, assuming that $p \leq \pi_0$ and $x > (1-\pi_0)\pi_0$? We know that a fraction $1-(1-\pi_0)p$ of the consumers will buy product and the others will search. Consumer s who searches will end up buying with probability $P(s) = \frac{\pi_0 p}{p-s}$, so that the firm's expected profit from the searching segment is, assuming no production cost:

$$\frac{1}{x}\int_0^{(1-\pi_0)p} \frac{\pi_0 p^2}{p-s} ds = -\frac{1}{x}\pi_0 p^2 \ln \pi_0,$$

which is strictly increasing in p. Therefore, the firm's profit from the two segments is

$$\Pi(p) = \frac{1}{x}[x-(1-\pi_0)p]p - \frac{1}{x}\pi_0 p^2 \ln \pi_0,$$

which is maximized at $p_B^* = \min\{\frac{x}{2(1-\pi_0+\pi_0 \ln \pi_0)}, \pi_0\}$. This is the optimal price if the firm serves everyone (ex ante). If the firm wants to charge a high price, it will exclude the high-search-cost consumers altogether and maximize:

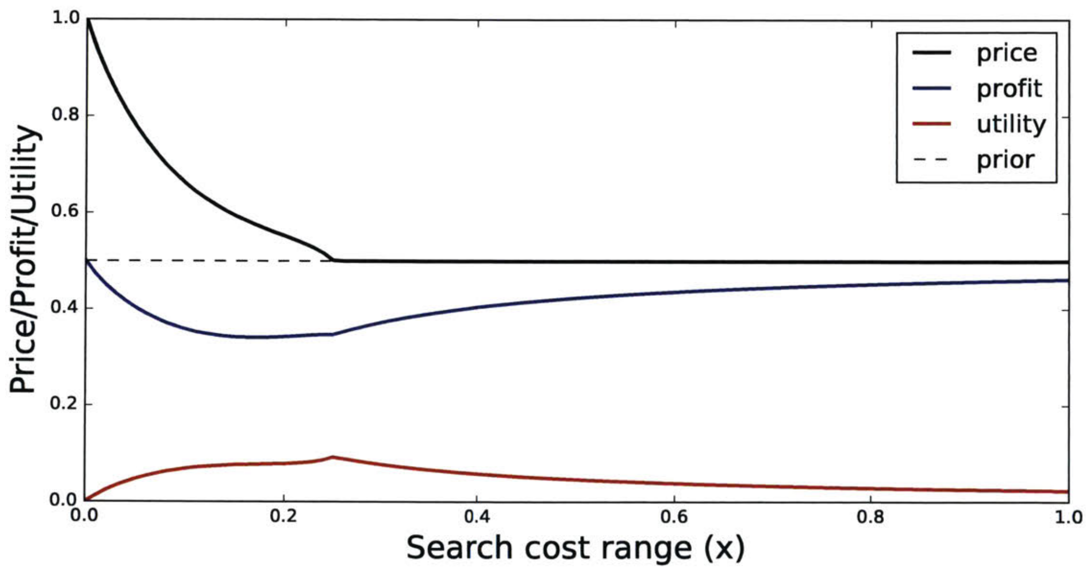
$$\Pi_S(p) = \frac{1}{x}\int_0^{\bar{s}(p)} \frac{\pi_0 p^2}{p-s} ds = \frac{1}{x}\pi_0 p^2 \ln\left(\frac{p}{p-\bar{s}(p)}\right),$$

[17] Another way to do this is to assume that $s \sim F([0,1])$ and see how the shape of F matters.

where $\bar{s}(p) \leq x$ is the search cost that sets $V_S(p,s) = 0$. That is, it is the mass of consumers who search, while consumers with $s > \bar{s}(p)$ exit. Note that $\bar{s}$ is decreasing in p as was obtained previously. Maximizing this profit will give an optimal search price p_S^*. Note also that if $\bar{s}(p) > x$, then the firm could easily raise p and lose no consumers.

The firm has to then choose p^* such that either $\bar{s}(p^*) = x$ or $p^* = p_S^*$ or $p^* = p_B^*$. If we fix $\pi_0 = \frac{1}{2}$, we can draw the figure below where we plot the equilibrium of this game as a function of x. In the figure, the firm is setting its price so that $\bar{s}(p^*) = x$ for $x \leq \frac{1}{4}$, where p^* is decreasing in x, but after that it finds it optimal to sell to all the high-search-cost consumers at $p^* = \pi_0$. The average consumer utility is maximized at $x = \frac{1}{4}$ where the firm starts serving the whole market. For x higher than this, the consumers are hurt because relatively fewer of them will have low search costs, and only the ones with $s < (1 - \pi_0)\pi_0$ will get a positive surplus.

Figure A-4: Equilibrium as a function of x when $s \sim U[0, x]$ and $\pi_0 = \frac{1}{2}$.



Appendix B

Appendix to Chapter 2

B.1 Proofs for Section 2.2

Lemma 2.1 Under Assumption 2.1, the consumer will always search or buy from the lower-price firm first, and consider the other firm only if the first product turns out to be of low match quality. The consumer's optimal strategy for each firm is the same as when facing a monopolist.

Proof. When the consumer faces a monopolist, buying is optimal if and only if $\pi_0 - p \geq \max\{0, \pi_0(1-p) - s\} \Leftrightarrow p \leq \min\{\pi_0, p_B\}$, where $p_B \equiv \frac{s}{1-\pi_0}$ is the price that makes the consumer indifferent between buying and searching (and it is called the purchase price). On the other hand, searching dominates exiting if and only if $\pi_0(1-p) - s \geq 0 \Leftrightarrow p \leq p_{max}$, where $p_{max} \equiv 1 - \frac{s}{\pi_0}$ (called the search price). We see that $p_B < \pi_0 \Leftrightarrow s < \pi_0(1-\pi_0)$, and that $p_{max} > \pi_0 \Leftrightarrow s < \pi_0(1-\pi_0)$. If $s \geq \pi_0(1-\pi_0)$, the consumer will never search and is willing to purchase the product at price π_0, which is the uninteresting case with a monopolist. When the cost of search is not high ($s < \pi_0(1-\pi_0)$), the firm has to choose between p_B and p_{max}, and it will choose the one that brings higher expected profits, depending on the value of s so that p_{max} is chosen for low search costs and p_B for higher ones.

What if the consumer faces two firms instead of just one? Assume first that Assumption 2.1 holds. If she has already searched or purchased from one firm (and obtained a product of low match quality), she will deal with the second firm as if it was the only firm in the world. Therefore, she will use exactly the same reasoning as above. This implies that, when the consumer is deciding whether to search or to purchase from the first firm, her continuation value in case of failure is independent of what she chooses to do with the first firm. Thus, the consumer deals with the first firm exactly as in the case of a monopolist. The consumer's utility of choosing firm i first can be written as $U_{ij} = V(p_i) + (1-\pi_0)V(p_j)$, where $V(p) = \max\{0, \pi_0 - p, \pi_0(1-p) - s\}$ is the value of facing a single firm. It is easy to see that the consumer should choose the lower-price firm first because $U_{ij} \geq U_{ji} \Leftrightarrow V(p_i) \geq V(p_j) \Leftrightarrow p_i \leq p_j$.[1] □

[1]The tie only occurs if $p_i = p_j$ or if $V(p) = 0$ for both prices.

Lemma B.1.1. *If we only allow the consumer to buy once, she may search the high-price firm first.*

Proof. If we do not make Assumption 2.1, then searching has a higher option value because buying from one firm does not allow the consumer to deal with the second firm at all, but it is possible to search one firm first and still go to the other one later. Now, last-period value is still $V(p_j)$ as in proof of Lemma 2.1 before, in case the consumer decides to search firm i first, but there is no last-period value if the consumer buys from firm i first and gets a bad draw. Thus, the consumer's value of buying from firm i is $U_{Bi} = \pi_0 - p_i$ and her value of searching firm i is $U_{Si} = \pi_0(1 - p_i) - s + (1 - \pi_0)V(p_j)$. The continuation value depends on whether the consumer searches or purchases from the first firm.

Assuming that $s < \pi_0(1 - \pi_0)$ and that the consumer chooses firm i first, she will purchase if and only if:

$$U_{Bi} \geq U_{Si} \Leftrightarrow p_i + V(p_j) \leq p_B, \tag{B.1}$$

where $p_B = \frac{s}{1-\pi_0}$ is the purchase price we defined earlier. This leaves two cases: 1) $p_j \leq p_B$, and 2) $p_B < p_j \leq p_{max}$, depending on what the consumer does with the second firm. In case 1), inequality (B.1) becomes $p_i \leq p_B - \pi_0 + p_j$, while in case 2) it becomes $p_i \leq \frac{2-\pi_0}{1-\pi_0}s - \pi_0(1 - p_j)$. We can now show that even if $p_i \leq p_j$ and $p_i \leq p_B$, the consumer may search firm j first. This happens if $U_{Sj} \geq \max\{U_{Si}, U_{Bi}, U_{Bj}\} \Leftrightarrow \pi_0(1-p_j) - s + (1-\pi_0)(\pi_0 - p_i) \geq \max\{\pi_0(1-p_i) - s + (1-\pi_0)V(p_j), \pi_0 - p_i\}$, where we know that $U_{Bi} \geq U_{Bj}$ because $p_i \leq p_j$, but we do not know what $V(p_j)$ is because we have not specified the range of p_j. If we assume that $p_j \leq p_B$, then $V(p_j) = \pi_0 - p_j$, and we can write the last (double)inequality as $p_j - p_i \leq p_{max} - \pi_0$ and $(1 - 2\pi_0)(p_j - p_i) \geq 0$. Thus, for low enough s, the consumer should search firm j first even when $p_i < p_j \leq p_B$, as long as $\pi_0 \leq \frac{1}{2}$. □

B.2 Some proofs for Section 2.3

Lemma 2.2. There is no equilibrium where both firms charge a fixed price. This holds for all s.

Proof. To the contrary, assume that there was an equilibrium where the firms were charging prices p_L and p_H, respectively, with $p_L \leq p_H$. If $p_L = p_H > 0$, then one firm could slightly undercut the other and gain a discrete jump in the probability of search/purchase (from $\frac{1}{2}$ to 1), which will always lead to higher expected profits. If $p_L = p_H = 0$, then either firm would have the incentive to raise its price because it would essentially become a monopoly if the lower-price firm turned out bad.

Assume then that $p_L < p_H \leq p_B$ or $p_B < p_L < p_H$. Now, firm L would have an incentive to increase its price which would not affect anything else. If $p_L < p_B < p_H$, both firms should increase their price until p_B and p_{max}, respectively. If $p_L = p_B < p_H = p_{max}$, one of the firms would be making profits weakly higher than the other. This would mean that the other firm should price just below the higher-profit firm and get strictly more than before. □

Lemma 2.3. There is no equilibrium where one firm charges a fixed price and the other randomizes.

Proof. Assume that one firm was charging p and the other was randomizing over $[p_L, p_H]$. Then if $p < p_L$ or $p_H \leq p < p_{max}$, the fixed-price firm should increase its price. On the other hand, the firm that was mixing should not mix because it would be winning or losing the price competition for all the prices, meaning that it could just charge the maximum winning or losing price.

If $p_L \leq p < p_H$, the higher-priced firm should not be mixing because it would make higher profits by setting either $p - \epsilon$ or p_H. This proves the claim. □

B.3 Proof of Theorem 2.2

Theorem 2.2. Assume that $\pi_0 \in (0, 1)$ and that there are $N + 1$ ex-ante-identical firms Bertrand competing. If Assumption 2.2 holds, the consumer's equilibrium utility is non-monotone in the search cost and achieves its unique global maximum at $s^* := \frac{\pi_0(1-\pi_0)}{2-\pi_0}$. It attains a unique global minimum at $\hat{s} := \frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1} < s^*$.

Lemma B.3.1. *If* $s < \frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1}$*, then the unique equilibrium involves all firms mixing over* $p \in [p_H, p_{max}]$*, where* $p_{max} := 1 - \frac{s}{\pi_0}$ *and* $p_H := (1 - \pi_0)^N p_{max}$*. The mixing probability is given by* $F(p) = \frac{1}{\pi_0} - \frac{1-\pi_0}{\pi_0}\left(\frac{1-\frac{s}{\pi_0}}{p}\right)^{1/N}$.

Proof. As with N=1 (two firms), the general case has a unique mixing equilibrium for every s because everyone has to be mixing over the same range (otherwise some firm would have a profitable deviation) and this range is fully determined by what the highest-price firm wants to do. For $s < \frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1}$, the firm setting the lowest price, p_H, will prefer this price to $p_B := \frac{s}{1-\pi_0}$ (highest price that makes the consumer buy instead of searching). p_H is defined by equating the profits from p_H and p_{max}: p_H with probability π_0 has to equal p_{max} with probability $(1 - \pi_0)^N \pi_0$. p_{max}, on the other hand, is the highest price for which the consumer is willing to search your firm instead of exiting (the highest-price firm should always set p_{max} for these search costs). Now, let us figure out what $F(p)$ is by equating the expected profits from any $p \in [p_H, p_{max}]$ to the profits from p_{max} (depending on how many firms price below p):

$$\left[\sum_{k=0}^{N} \binom{N}{k} F(p)^k (1 - \pi_0)^k [1 - F(p)]^{N-k}\right] \pi_0 p = (1 - \pi_0)^N \pi_0 (1 - \frac{s}{\pi_0})$$
$$\Leftrightarrow [1 - \pi_0 F(p)]^N p = (1 - \pi_0)^N (1 - \frac{s}{\pi_0})$$
$$\Leftrightarrow F(p) = \frac{1}{\pi_0} - \frac{1 - \pi_0}{\pi_0}\left(\frac{1 - \frac{s}{\pi_0}}{p}\right)^{1/N}$$

□

Lemma B.3.2. *Assume that* $s \in \left[\frac{\pi_0(1-\pi_0)}{2-\pi_0}, \pi_0(1-\pi_0)\right]$. *In the unique equilibrium of the pricing game, all firms randomize over* $p \in \left[(1-\pi_0)^{N-1}s, \frac{s}{1-\pi_0}\right]$ *with* $F(p) = \frac{1}{\pi_0} - \frac{(1-\pi_0)^{1-1/N}}{\pi_0}(\frac{s}{p})^{1/N}$.

Proof. Note that the firms always make the consumer purchase. The reason is that for $s \geq \frac{\pi_0(1-\pi_0)}{2-\pi_0}$, any firm prefers p_B over p_{max} given that the probability of purchase is the same for both. Thus, the firm that sets the highest price will always set p_B rather than any search price, which means that all firms have to be setting prices that are at most p_B. However, the upper end of the pricing range has to be exactly $p_B = \frac{s}{1-\pi_0}$ because otherwise the highest-price firm would have forgone profits. This already implies that the lower bound for the pricing range has to be $(1-\pi_0)^N p_B = (1-\pi_0)^{N-1}s$ so that the expected profits will be the same. Finally, $F(p)$ can be obtained by equating profits from the different prices just like in the previous lemma. □

Lemma B.3.3. *Assume that* $s > (1-\pi_0)\pi_0$. *The unique equilibrium has all firms mixing over prices* $p \in [(1-\pi_0)^N\pi_0, \pi_0]$ *with* $F(p) = \frac{1}{\pi_0} - \frac{1-\pi_0}{\pi_0}(\frac{\pi_0}{p})^{1/N}$.

Proof. Because s is too high, the consumer will never search for any price $p \geq \pi_0$ so there will be no search. This means that the firm that prices at the upper end of the price range has to set $p_B = \pi_0$. This implies that the lower bound has to be $(1-\pi_0)^N\pi_0$ so that profits will be the same at both prices. Finally, $F(p)$ can be solved by equating the profits from any p in the range to $(1-\pi_0)^N\pi_0$. □

Lemma B.3.4. *Assume that* $s \in \left(\frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1}, \frac{\pi_0(1-\pi_0)}{2-\pi_0}\right)$. *Now, the game has a unique equilibrium where all of the firms randomize over* $p \in \left[(1-\pi_0)^N(\pi_0 - s), \frac{s}{1-\pi_0}\right] \cup \left[\frac{s}{(1-\pi_0)\pi_0}, 1 - \frac{s}{\pi_0}\right]$. *The CDF is given by:*

$$F(p) = \begin{cases} \frac{1}{\pi_0} - \frac{1-\pi_0}{\pi_0}\left(\frac{\pi_0 - s}{p}\right)^{1/N}, & \text{if } \; p \in \left[(1-\pi_0)^N(\pi_0 - s), \frac{s}{1-\pi_0}\right] \\ \frac{1}{\pi_0} - \frac{1-\pi_0}{\pi_0}\left(\frac{\pi_0 - s}{\pi_0 p}\right)^{1/N}, & \text{if } \; p \in \left[\frac{s}{(1-\pi_0)\pi_0}, 1 - \frac{s}{\pi_0}\right]. \end{cases}$$

Proof. Now, the consumer will purchase if and only if $p \leq p_B := \frac{s}{1-\pi_0}$, which means that the firms should never price between p_B and $p_H := \frac{s}{(1-\pi_0)\pi_0}$ (where p_H is constructed by equating the profits from p_B to those from p_H, assuming the probability of consumer choosing the firm does not change). This is why there is a gap in the support of equilibrium prices. Note also that because s is in the range assumed in the Lemma, all else equal, the firms prefer p_{max} over p_B.

Now, the highest price has to be p_{max} because that is preferred to p_B (and all other $p < p_B$). This means that the lowest price has to give the same expected profit. Thus, it has to be $(1-\pi_0)^N(\pi_0 - s)$. We have now defined the highest and the lowest price. We also know that p_B has to be in the support because if the lower range ended at $p < p_B$, the firm setting p would like to raise its price to p_B and lose nothing. On the other hand, p_H, which we defined in the previous paragraph, gives the same probability as p_B of the consumer choosing that firm (since there is nothing between the two prices), so they have to give the same expected profits given the consumer choosing them. Now, we can just construct the CDF for prices as follows:

$$\begin{aligned}[1-\pi_0 F(p)]^N p &= (1-\pi_0)^N(\pi_0 - s), \quad \text{if} \quad p \le p_B \\ [1-\pi_0 F(p)]^N \pi_0 p &= (1-\pi_0)^N(\pi_0 - s), \quad \text{if} \quad p \ge p_H.\end{aligned}$$

Solving this gives the $F(p)$ introduced in the Lemma. □

Lemma B.3.5. *In equilibrium, the consumer's utility and a firm's profits are:*

$$U^* = \begin{cases} (1-\frac{s}{\pi_0})[1-(1-\pi_0)^N(1+\pi_0 N)], & \textit{if} \quad s \le \frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1} \\ 1-(1-\pi_0)^{N+1}\left(1-\frac{s}{\pi_0}\right) - \frac{s}{\pi_0}[1-\pi_0 F(p_B)]^{N+1} & \\ \quad -(N+1)(1-\pi_0)^N(\pi_0 - s), & \textit{if} \quad s \in \left(\frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1}, \frac{\pi_0(1-\pi_0)}{2-\pi_0}\right) \\ 1-(1-\pi_0)^{N-1}[(1-\pi_0)^2+(N+1)s], & \textit{if} \quad s \in \left[\frac{\pi_0(1-\pi_0)}{2-\pi_0}, \pi_0(1-\pi_0)\right] \\ 1-(1-\pi_0)^N(1+\pi_0 N), & \textit{if} \quad s > \pi_0(1-\pi_0) \end{cases}$$

$$\Pi_i^* = \begin{cases} (1-\pi_0)^N(\pi_0 - s), & \textit{if} \quad s \le \frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1} \\ (1-\pi_0)^N(\pi_0 - s), & \textit{if} \quad s \in \left(\frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1}, \frac{\pi_0(1-\pi_0)}{2-\pi_0}\right) \\ (1-\pi_0)^{N-1}s, & \textit{if} \quad s \in \left[\frac{\pi_0(1-\pi_0)}{2-\pi_0}, \pi_0(1-\pi_0)\right] \\ \pi_0(1-\pi_0)^N, & \textit{if} \quad s > \pi_0(1-\pi_0) \end{cases}$$

Proof. Profits of firm i are simply the same as the expected profits of the firm setting the highest price, which in the first two cases gives $(1-\pi_0)^N \pi_0 p_{max} = (1-\pi_0)^N(\pi_0 - s)$. On the other hand, in the third case, the highest-price firm gets $(1-\pi_0)^N p_B = (1-\pi_0)^{N-1}s$, and in the fourth case, it gets $(1-\pi_0)^N \pi_0$.

The consumer's utility is slightly more difficult to compute but once we understand that it is the same as social surplus minus the firms' combined profits, the problem becomes easy. In the first case, social surplus is $S = (\pi_0 - s)[1 + (1-\pi_0) + (1-\pi_0)^2 + \ldots + (1-\pi_0)^N] = (\pi_0 - s)\frac{1-(1-\pi_0)^{N+1}}{\pi_0}$ because the consumer searches once for sure and always searches again if she gets a bad draw until she finds a good product or exhausts all of her options. Subtracting the firms' combined profits $\sum \Pi_i = (N+1)(1-\pi_0)^N(\pi_0 - s)$ from this number gives the consumer's equilibrium utility.

Similarly, in the last two cases, social surplus is $S = 1-(1-\pi_0)^{N+1}$ (surplus is one unless all firms turn out to be low quality), so subtracting firms' profits will give the consumer's utility.

However, the second case where prices can be both higher and lower than the highest purchase price, p_B, is more complicated. If all firms set $p \le p_B$, social surplus is $1-(1-\pi_0)^{N+1}$. If all firms set $p > p_B$, social surplus is $(\pi_0 - s)\frac{1-(1-\pi_0)^{N+1}}{\pi_0}$. All other cases yield a social surplus that is between these two numbers. We can calculate the expected social surplus as follows:

$$
\begin{aligned}
S =& F(p_B)^{N+1}[\pi_0 + (1-\pi_0)\pi_0 + (1-\pi_0)^2\pi_0 + \ldots + (1-\pi_0)^N\pi_0] \\
&+ (N+1)F(p_B)^N(1-F(p_B))[\pi_0 + (1-\pi_0)\pi_0 + \ldots + (1-\pi_0)^{N-1}\pi_0 + (1-\pi_0)^N(\pi_0 - s)] \\
&+ \binom{N+1}{2}F(p_B)^{N-1}(1-F(p_B))^2[\pi_0 + (1-\pi_0)\pi_0 + \ldots + (1-\pi_0)^{N-2}\pi_0 \\
&\qquad\qquad + (1-\pi_0)^{N-1}(\pi_0 - s) + (1-\pi_0)^N(\pi_0 - s)] \\
&\vdots \\
&+ \binom{N+1}{N+1}(1-F(p_B))^{N+1}[\pi_0 - s + (1-\pi_0)(\pi_0 - s) + \ldots + (1-\pi_0)^N(\pi_0 - s)] \\
=& \sum_{k=0}^{N+1}\binom{N+1}{k}F(p_B)^k[1-F(p_B)]^{N+1-k}\left[1-(1-\pi_0)^{N+1} - \frac{s}{\pi_0}\left[(1-\pi_0)^k(1-(1-\pi_0)^{N+1-k})\right]\right]
\end{aligned}
$$

where the last line is the sum of expected social surpluses in different cases. k is the number of firms that set a low price ($p \leq p_B$), and $\pi_0\frac{1-(1-\pi_0)^{N+1}}{\pi_0}$ is the expected social value of searching/buying from the $N+1$ firms (it is the same for all the prices). However, the cost of searching depends on how many firms set a high price $p > p_B$, where the expected search cost when k firms set a low price is $\frac{s}{\pi_0}\left[(1-\pi_0)^k(1-(1-\pi_0)^{N+1-k})\right]$. Now, we can simplify the above expression for social surplus to get:

$$
\begin{aligned}
S &= 1-(1-\pi_0)^{N+1} - \sum_{k=0}^{N}\binom{N+1}{k}F(p_B)^k[1-F(p_B)]^{N+1-k}\cdot\frac{s}{\pi_0}(1-\pi_0)^k\left[1-(1-\pi_0)^{N+1-k}\right] \\
&= 1-(1-\pi_0)^{N+1}\left(1-\frac{s}{\pi_0}\right) - \frac{s}{\pi_0}\left[1-\pi_0F(p_B)\right]^{N+1}.
\end{aligned}
$$

Subtracting the $N+1$ firms' profits from this expression finally allows us to write the consumer's utility as given in the lemma. □

Having the above lemmas in hand, we can finally prove the Theorem:

Proof. Using Lemma B.3.5, we see the consumer's equilibrium utility at $s = 0$ always equals that at any $s \geq \pi_0(1-\pi_0)$. We also see that equilibrium utility is first decreasing in s and reaches its global minimum of $U_{min} = \frac{1-(1-\pi_0)^N(1+\pi_0 N)}{(1-\pi_0)^{N+1}+1}$ at $s = \frac{(1-\pi_0)^{N+1}\pi_0}{(1-\pi_0)^{N+1}+1}$. After this it is increasing until $s = s^* := \frac{\pi_0(1-\pi_0)}{2-\pi_0}$, after which it is again decreasing and levels off at $s = \pi_0(1-\pi_0)$. Therefore, s^* is the unique global maximizer of utility for all $N \in \mathbb{N}/\{0\}$. □

Appendix C

Appendix to Chapter 3

C.1 Omitted Proofs

Proposition 3.4. $s^2 \equiv var(\text{reviews}) = \frac{1}{\tau_u} - \sigma^2 \frac{\phi(q/\sigma)}{\Phi(q/\sigma)} \left(\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)} \right) = \frac{1}{\tau_u} - \mu(\mu - q)$, which is strictly increasing in q, with $\lim_{q \to -\infty} s^2 = \frac{1}{\tau_u + \tau_\epsilon}$, and $\lim_{q \to \infty} s^2 = \frac{1}{\tau_u}$. Moreover, the variance of reviews is strictly decreasing in τ_ϵ, with $\lim_{\tau_\epsilon \to 0} s^2 = \frac{1}{\tau_u}$, and $\lim_{\tau_\epsilon \to \infty} s^2 = \frac{1}{\tau_u} - \frac{1}{\tau_u} \frac{\phi(q\sqrt{\tau_u})}{\Phi(q\sqrt{\tau_u})} \left(q\sqrt{\tau_u} + \frac{\phi(q\sqrt{\tau_u})}{\Phi(q\sqrt{\tau_u})} \right)$. For τ_u we have $\lim_{\tau_u \to 0} s^2 = \infty$ and $\lim_{\tau_u \to \infty} s^2 = 0$.

Proof. To show that $var(r) = \frac{1}{\tau_u} - \sigma^2 \frac{\phi(q/\sigma)}{\Phi(q/\sigma)} \left(\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)} \right)$, remember that $var(r) = var(u \mid \mathbb{E}[u|s] \geq 0) = \mathbb{E}[u^2 \mid \mathbb{E}[u|s] \geq 0] - \mathbb{E}[u \mid \mathbb{E}[u|s] \geq 0]^2$. We have already shown that $\mathbb{E}[u \mid \mathbb{E}[u|s] \geq 0] = q + \sigma \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}$, so what remains is to obtain an expression for the first term. To do this, we can write:

$$
\begin{aligned}
\mathbb{E}[u^2|\mathbb{E}[u|s] \geq 0] &= \mathbb{E}[u^2|s \geq \hat{s}] \\
&= \int_{\hat{s}}^{\infty} \frac{\mathbb{E}[u^2|s]}{1 - F(\hat{s})} f(s) ds \\
&= \int_{\hat{s}}^{\infty} \frac{var(u|s) + \mathbb{E}[u|s]^2}{1 - F(\hat{s})} f(s) ds \\
&= \frac{1}{\tau_u + \tau_\epsilon} + \frac{1}{(\tau_u + \tau_\epsilon)^2} \int_{\hat{s}}^{\infty} \frac{\tau_u^2 q^2 + 2\tau_u \tau_\epsilon q s + \tau_\epsilon^2 s^2}{1 - F(\hat{s})} f(s) ds \\
&= \frac{1}{\tau_u + \tau_\epsilon} + \frac{\tau_u^2}{(\tau_u + \tau_\epsilon)^2} q^2 + \frac{2\tau_u \tau_\epsilon q}{(\tau_u + \tau_\epsilon)^2} \mathbb{E}[s|s \geq \hat{s}] + \frac{\tau_\epsilon^2}{(\tau_u + \tau_\epsilon)^2} \mathbb{E}[s^2|s \geq \hat{s}], \qquad \text{(C.1)}
\end{aligned}
$$

where F is the cdf of the signal (and f the corresponding pdf). The second line follows due to the usual decomposition of variance, while the third line is just writing out the (known) expression for the conditional variance and expectation. The fourth line uses the definition of conditional expectation. Next, let us compute $\mathbb{E}[s^2|s \geq \hat{s}]$, after which we are basically done (since we already know how to compute everything else). To this end, we can use the variance decomposition to write:

$$\begin{aligned}\mathbb{E}[s^2|s \geq \hat{s}] &= var(s|s \geq \hat{s}) + \mathbb{E}[s|s \geq \hat{s}]^2 \\ &= \sigma_s^2 \left[1 - \frac{q}{\sigma}\frac{\phi(q/\sigma)}{\Phi(q/\sigma)} - \frac{\phi(q/\sigma)^2}{\Phi(q/\sigma)^2}\right] + \left[q + \sigma_s \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right]^2 \\ &= \sigma_s^2 + q^2 + 2q\sigma_s \frac{\phi(q/\sigma)}{\Phi(q/\sigma)} - \sigma_s^2 \frac{q}{\sigma}\frac{\phi(q/\sigma)}{\Phi(q/\sigma)},\end{aligned} \tag{C.2}$$

where we have just written out the formulas for the variance and expectation of the truncated normal distribution (knowing that $s \sim N(q, \sigma_s^2)$ with $\sigma_s^2 = \frac{\tau_u + \tau_\epsilon}{\tau_u \tau_\epsilon}$). Finally, combining equations (C.1) and (C.2), and cleaning things up (noting that $\frac{1}{\tau_u + \tau_\epsilon} + \sigma^2 = \frac{1}{\tau_u}$), we get the desired expression.

Looking at the simple expression for the variance, $\frac{1}{\tau_u} - \mu(\mu - q)$, it is easy to see that it is decreasing in μ (selection implies $\mu > q$). Using the earlier fact that μ is increasing in τ_ϵ, we therefore have that the variance is decreasing in τ_ϵ. Computing the derivative with respect to q is slightly more complicated:

$$\begin{aligned}\frac{\partial s^2}{\partial q} = \mu - (2\mu - q)\frac{\partial \mu}{\partial q} &= q + \sigma\frac{\phi(q/\sigma)}{\Phi(q/\sigma)} - \left(q + 2\sigma\frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right)\left(1 - \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\left(\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right)\right) \\ &= \left(q + 2\sigma\frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right)\left(\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right)\frac{\phi(q/\sigma)}{\Phi(q/\sigma)} - \sigma\frac{\phi(q/\sigma)}{\Phi(q/\sigma)} \\ &= \sigma\frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\left[\left(\frac{q}{\sigma} + 2\frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right)\left(\frac{q}{\sigma} + \frac{\phi(q/\sigma)}{\Phi(q/\sigma)}\right) - 1\right] > 0.\end{aligned}$$

Here we have just simplified the derivative and gathered terms. The inequality follows because the function $h(x) = \left(x + 2\frac{\phi(x)}{\Phi(x)}\right)\left(x + \frac{\phi(x)}{\Phi(x)}\right) - 1 > 0$ for all x (this function is strictly increasing from 0 to ∞). This proves that the variance is increasing in q. There is no such result for τ_u because there are some knife-edge cases where reviews are increasing in τ_u. The limit proofs are omitted. □

Proposition 3.10 Under the above assumptions, the amount of selection for each of the two movies is:

$$\bar{v}_{(2)} = \frac{2a\phi\left(aq\right)\left[1 - \Phi\left(abq\right)\right]}{\tau_v\left[\Phi\left(aq\right) + 2T(aq, b)\right]},$$

where $a = \sqrt{\frac{\tau_v}{1+\sigma^2\tau_v}}$, $b = \sqrt{\frac{\sigma^2\tau_v}{2+\sigma^2\tau_v}}$, $\sigma^2 = \frac{\tau_\epsilon}{\tau(\tau+\tau_\epsilon)}$, and $T(h, c) = \phi(h)\int_0^c \frac{\phi(hx)}{(1+x^2)}dx$ is Owen's T function.

Proof. As in the case of a single movie, consumer j finds movie i acceptable if and only if $v_j + \mathbb{E}[x_i|s_i] \geq 0 \Leftrightarrow s_i \geq -\frac{\tau}{\tau_\epsilon}q - \frac{\tau+\tau_\epsilon}{\tau_\epsilon}v_j$. However, now she also has to prefer movie i over the other one. For simplicity, and without loss of generality, let us compute everything for movie 1 because of the full symmetry. Consumer j prefers movie 1 over movie 2 if and only if $v_j + \frac{\tau q + \tau_\epsilon s_1}{\tau + \tau_\epsilon} \geq v_j + \frac{\tau q + \tau_\epsilon s_2}{\tau+\tau_\epsilon} \Leftrightarrow s_1 \geq s_2$. Note that assuming ex-ante symmetry simplifies things a lot here.

Now, exactly as with one movie, the probability of movie 1 being acceptable is:

$$P := \mathbb{P}(1 \text{ acceptable}) = \Phi\left(\frac{v_j + q}{\sigma}\right), \tag{C.3}$$

where $\sigma^2 = \frac{\tau_\epsilon}{\tau(\tau+\tau_\epsilon)}$ is the conditional variance of the expected mean quality given a signal.

As a small detour, we can also compute the probability of consumer v_j choosing movie 1. Given that movie 1 is acceptable, the probability of it being preferred over 2 is:

$$\begin{aligned}
\mathbb{P}(1 \succ 2 \mid 1 \text{ acceptable}) &= \mathbb{P}\left(s_1 \geq s_2 \mid s_1 \geq -\frac{\tau}{\tau_\epsilon}q - \frac{\tau+\tau_\epsilon}{\tau_\epsilon}v_j\right) \\
&= \frac{1}{P}\int_{-\frac{\tau}{\tau_\epsilon}q-\frac{\tau+\tau_\epsilon}{\tau_\epsilon}v_j}^{\infty} \sqrt{\frac{\tau\tau_\epsilon}{\tau+\tau_\epsilon}}\phi\left(\sqrt{\frac{\tau\tau_\epsilon}{\tau+\tau_\epsilon}}(s_1-q)\right)\Phi\left(\sqrt{\frac{\tau\tau_\epsilon}{\tau+\tau_\epsilon}}(s_1-q)\right)ds_1 \\
&= \frac{1}{P}\int_{-\frac{v_j+q}{\sigma}}^{\infty} \Phi(Y)\phi(Y)dY \\
&= \frac{1}{P}\cdot\frac{1}{2}\left[1-\Phi\left(-\frac{v_j+q}{\sigma}\right)^2\right] \\
&= \frac{\left[2-\Phi(\frac{v_j+q}{\sigma})\right]\Phi(\frac{v_j+q}{\sigma})}{2\Phi(\frac{v_j+q}{\sigma})} \\
&= 1 - \frac{1}{2}\Phi(\frac{v_j+q}{\sigma}).
\end{aligned} \tag{C.4}$$

Here, the first and second line are just rewriting the conditional probability using Equation (C.3) and the distrFibutions of the two signals, when s_1 has to be acceptable (lower bound of the integral) and s_2 has to be lower than s_1 (the last term in the integral). The third line transforms s_1 into a standard normal variable Y, and the fourth line uses the following interesting result for standard normal variables:

$$\int_a^b \Phi(Y)\phi(Y)dY = \frac{1}{2}\left[\Phi(b)^2 - \Phi(a)^2\right].$$

The fifth line rewrites $\Phi(-\frac{v_j+q}{\sigma}) = 1 - \Phi(\frac{v_j+q}{\sigma})$ and plugs in P, and the last line simplifies. End of detour.

To get back to the proof, and to calculate the expected selection for each movie (they are symmetric), we just need to note that the ex-ante probability of consumer v_j not choosing either movie is $\mathbb{P}(\text{consume neither} \mid v_j) = \left(1-\Phi(\frac{v_j+q}{\sigma})\right)^2$ (product of the complements of Equation (C.3)). Observe that we can also get this using Equations (C.3) and (C.4).

Therefore, the probability of an unknown consumer (unknown v) consuming neither movie is:

$$
\begin{aligned}
\mathbb{P}(\text{consumer neither}) &= \int_{-\infty}^{\infty} \left[1 - \Phi\left(\frac{v_j + q}{\sigma}\right)\right]^2 \sqrt{\tau_v}\phi(\sqrt{\tau_v}v)dv \\
&= \int_{-\infty}^{\infty} [1 - \Phi(\alpha + \beta Y)]^2 \phi(Y)dY \\
&= \int_{-\infty}^{\infty} \phi(Y)dY - 2\int_{-\infty}^{\infty} \Phi(\alpha + \beta Y)\phi(Y)dY + \int_{-\infty}^{\infty} \Phi(\alpha + \beta Y)^2\phi(Y)dY \\
&= 1 - 2\Phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right) + \Phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right) - 2T\left(\frac{\alpha}{\sqrt{1+\beta^2}}, \frac{1}{\sqrt{1+2\beta^2}}\right),
\end{aligned}
$$

where line two follows from transforming v into Y which is a standard normal variable and defining $\alpha = q/\sigma$, and $\beta = 1/\sqrt{\sigma^2\tau_v}$. Line three opens the brackets up and line four uses known results for integrals of standard normals. Here, $T(\cdot,\cdot)$ is Owen's T function.

Thus, the probability of consuming one of the two movies is the complement of the above, so that:

$$
\begin{aligned}
Q := \mathbb{P}(\text{watch a movie}) &= \Phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right) + 2T\left(\frac{\alpha}{\sqrt{1+\beta^2}}, \frac{1}{\sqrt{1+2\beta^2}}\right) \\
&= \Phi(aq) + 2T(aq, b) \qquad \text{(C.5)}
\end{aligned}
$$

where the second line uses our definitions from the Proposition, $a = \sqrt{\frac{\tau_v}{1+\sigma^2\tau_v}}$ and $b = \sqrt{\frac{\sigma^2\tau_v}{2+\sigma^2\tau_v}}$, to rewrite the equation.

We are finally ready to compute the expected v conditional on choosing one movie:

$$
\begin{aligned}
\bar{v}_{(2)} &= \mathbb{E}[v \mid \text{consume 1 or 2}] \\
&= \frac{1}{Q}\int_{-\infty}^{\infty} \frac{Y}{\sqrt{\tau_v}} \left(1 - [1 - \Phi(\alpha + \beta Y)]^2\right) \phi(Y)dY \\
&= \frac{1}{Q}\int_{-\infty}^{\infty} \frac{Y}{\sqrt{\tau_v}} \left(2\Phi(\alpha + \beta Y) - \Phi(\alpha + \beta Y)^2\right) \phi(Y)dy \\
&= \frac{1}{\sqrt{\tau_v}Q}\left[\frac{2\beta}{\sqrt{1+\beta^2}}\phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right) - \frac{2\beta}{\sqrt{1+\beta^2}}\phi\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right)\Phi\left(\frac{\alpha}{\sqrt{1+\beta^2}\sqrt{1+2\beta^2}}\right)\right] \\
&= \frac{2}{\sqrt{\tau_v(1+\sigma^2\tau_v)}Q}\phi\left(q\sqrt{\frac{\tau_v}{1+\sigma^2\tau_v}}\right)\left[1 - \Phi\left(q\sqrt{\frac{\tau_v}{1+\sigma^2\tau_v}} \cdot \sqrt{\frac{\sigma^2\tau_v}{1+\sigma^2\tau_v}}\right)\right] \\
&= \frac{2a\phi(aq)[1 - \Phi(abq)]}{\tau_v[\Phi(aq) + 2T(aq, b)]}, \qquad \text{(C.6)}
\end{aligned}
$$

where the first line uses the definition of conditional expectation (and already performs a change of variables), the second one opens up the brackets, the third uses known results for integrals of standard normals, the fourth rewrites in terms of the original parameters, and the fifth plugs in Q from Equation (C.5) and also rewrites in terms of what we had in the Proposition.

Table C.1: Mixed Linear model. Regressing RT user reviews on underlying quality.

No. Observations:	4963	Method:	REML
No. Groups:	166	Scale:	0.0213
Min. group size:	1	Likelihood:	1931.6966
Max. group size:	59	Converged:	Yes
Mean group size:	29.9		

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
intercept	2.511	0.974	2.577	0.010	0.601	4.420
$\hat{q}$	0.530	0.016	33.455	0.000	0.499	0.561
log(box office)	-0.081	0.005	-17.237	0.000	-0.090	-0.072
days	-0.007	0.001	-8.569	0.000	-0.009	-0.006
days*critic avg	0.001	0.000	5.510	0.000	0.001	0.001
genre var	1.383	0.459	3.010	0.003	0.482	2.283

Now, this is the expected horizontal preference for those who consume one of the two movies but it is also the expected preference for those who consume, say, movie 1. This follows from the perfect symmetry of the movies – each movie is going to get the same consumers, in expectation. We are done. □

C.2 Using Underlying Quality Instead of Critic Reviews

In this section, we will run regression (5) from Table 3.6 but replacing *critic avg* with $\hat{q}$, an estimate for the underlying quality. However, to do this, we need to first estimate the underlying quality. The approach we will take is to trust the basic model and assume the (total) variance of reviews at any given time is $s_t^2 = \frac{1}{\tau_u} - \mu_t(\mu_t - q_t)$.[1] Let us then estimate the following regression:

$$s_{jt}^2 = \alpha_j + \beta\mu_{jt}^2 + \gamma\mu_{jt}C_{jt} + \lambda(t - r_j) + \epsilon_{jt},$$

where r_j is the release date of movie j. The estimate for underlying quality can then be obtained by noting that $-\mu(\mu - q) = \mu(\beta\mu + \gamma C) \Leftrightarrow q = (1+\beta)\mu + \gamma C$. The regression gives us $\hat{\beta} = -0.151$ and $\hat{\gamma} = 0.06$, so that $\hat{q} = 0.849\mu + 0.06C$ for given consumer and critic averages. Plugging this $\hat{q}$ into regression (5) from Table 3.6, we get the estimates shown in Table C.1. As we can see, these values are similar to those before but $\hat{q}$ is a much better predictor of reviews than *critic avg* because it better corresponds to the underlying quality.

C.3 Figures

This Appendix includes the graphs that were omitted from the main text.

[1]Here we are saying that μ_t is the cumulative IMDb average at t, while s_t^2 is the cumulative variance.

Figure C-1: Difference in consumer and critic average as a function of critic average, long data.

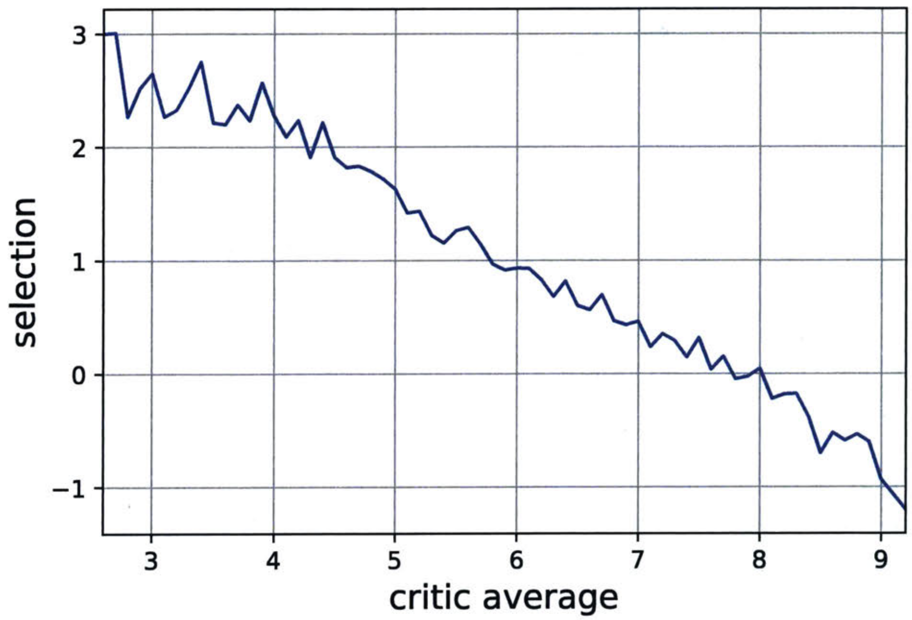


Figure C-2: Probability that consumer average is higher than critic average as a function of critic average, obtained using Logistic Regression. Reviews from Rotten Tomatoes, long data

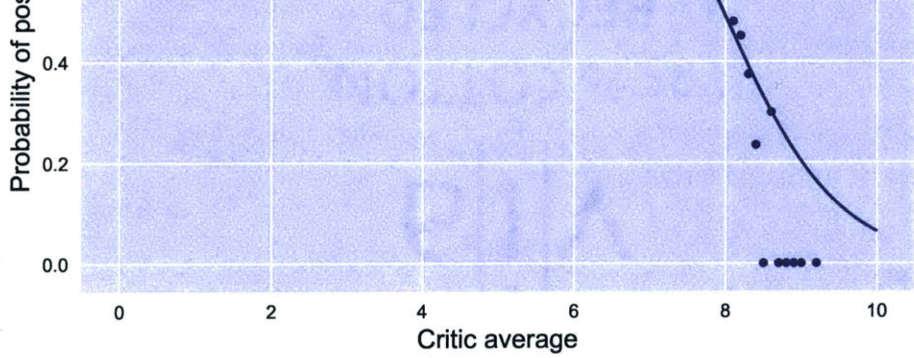


Figure C-3: IMDb variance (movie specific) for certain bins of critic averages, daily data.

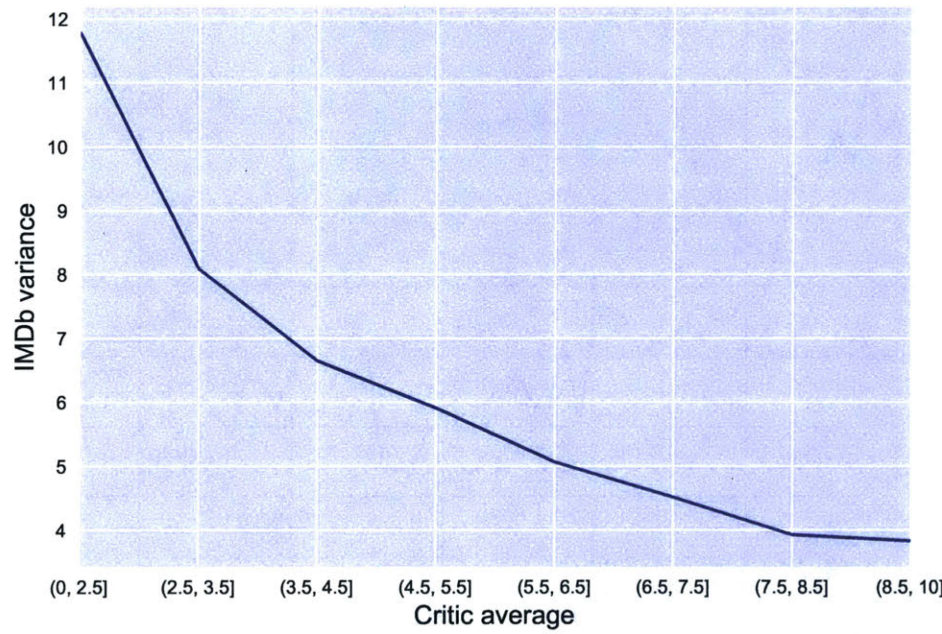


Figure C-4: Daily IMDb variance as a function of IMDb average. Averages binned into ten groups corresponding to the ten deciles, daily data.

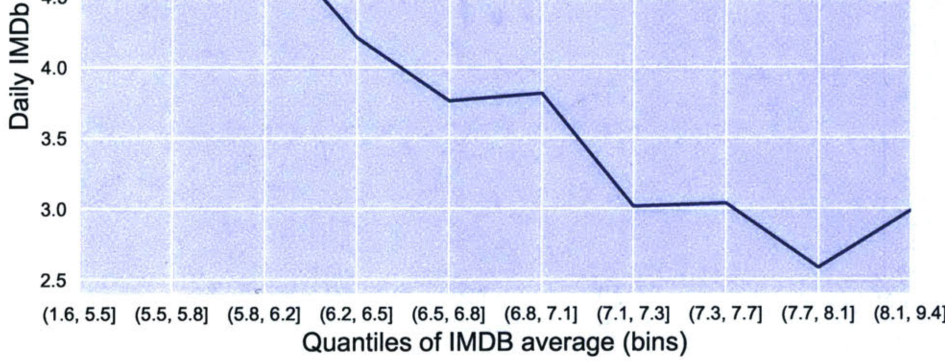


Figure C-5: Selection on Rotten Tomatoes as a function of time since release, daily data.

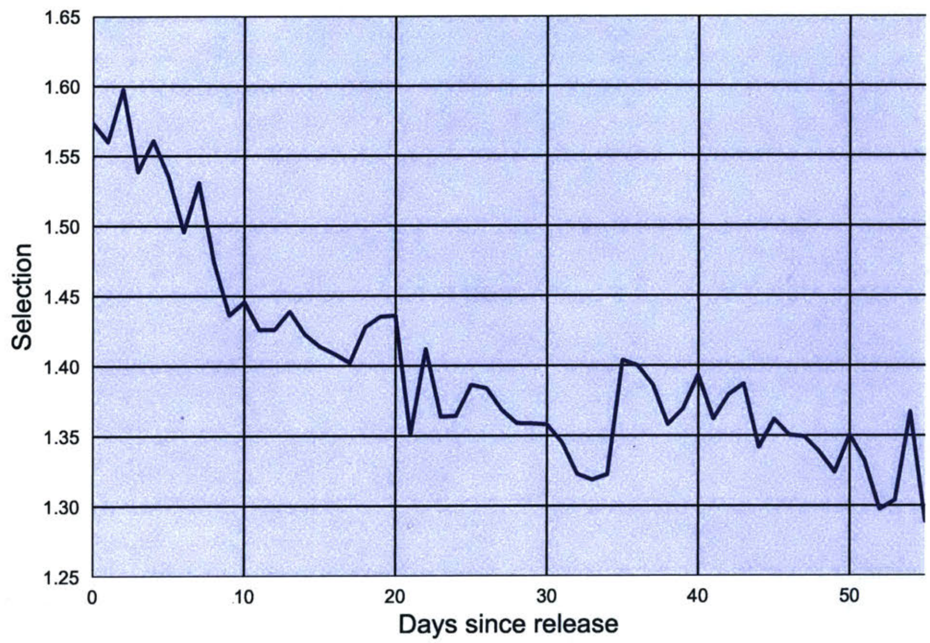


Figure C-6: Variance of IMDb user ratings (within movie) conditional on time since release, daily data.

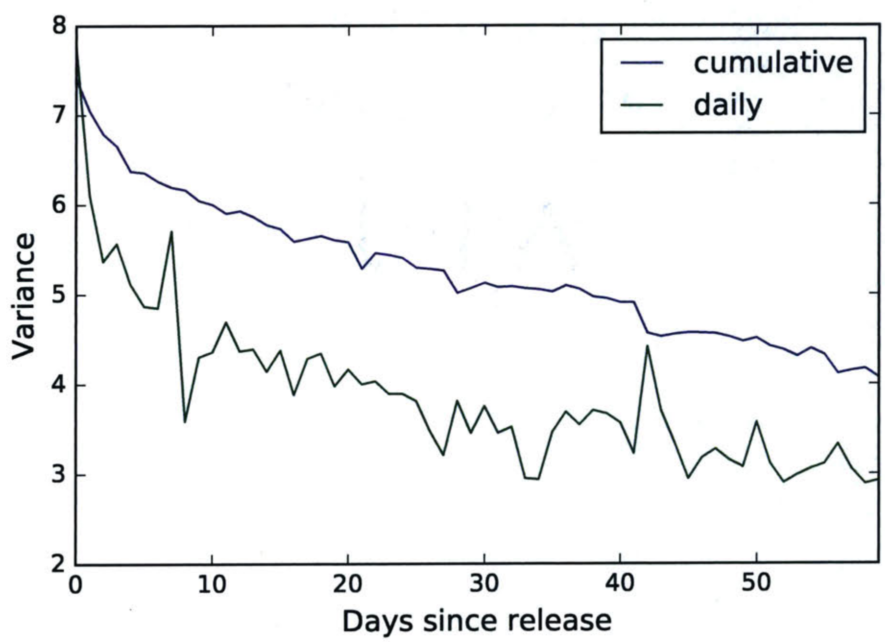


Figure C-7: Consumer and critic reviews by box office decile, long data.

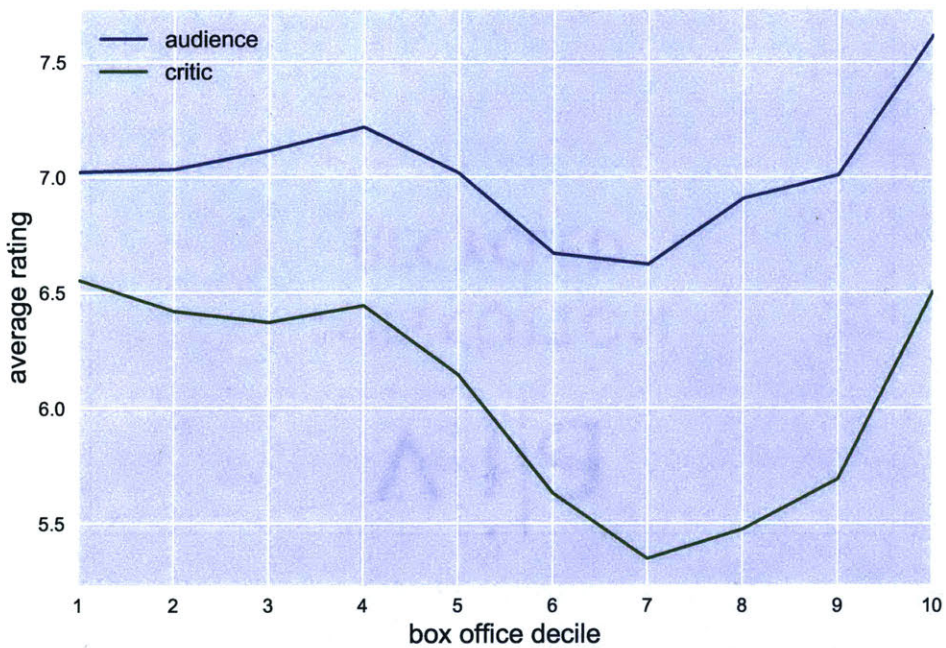


Figure C-8: Selection by box office decile, long data.

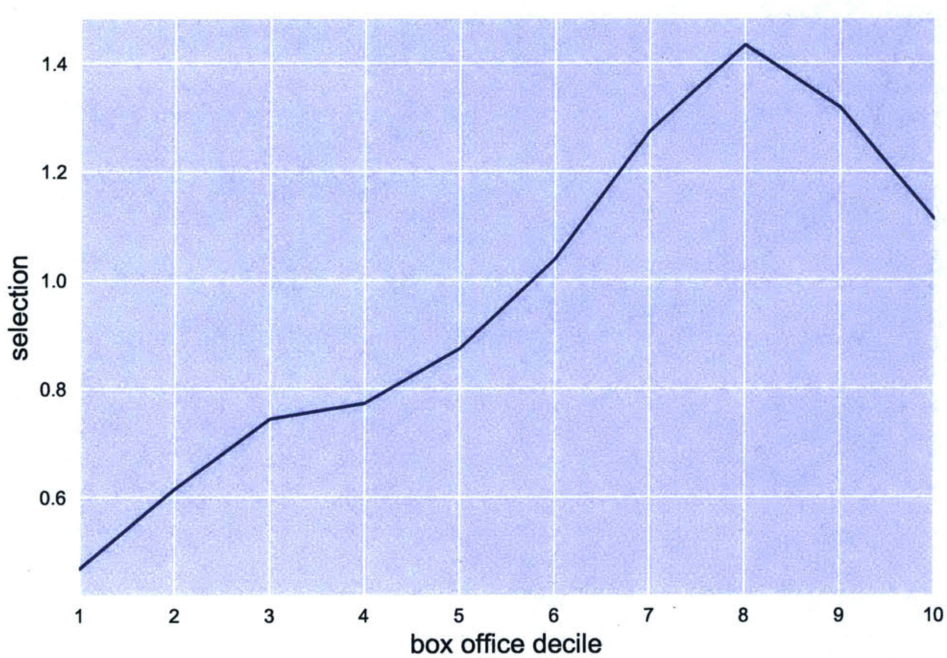


Figure C-9: Mean weekly box office revenue as a function of time (rolling seven-day sum), daily data.

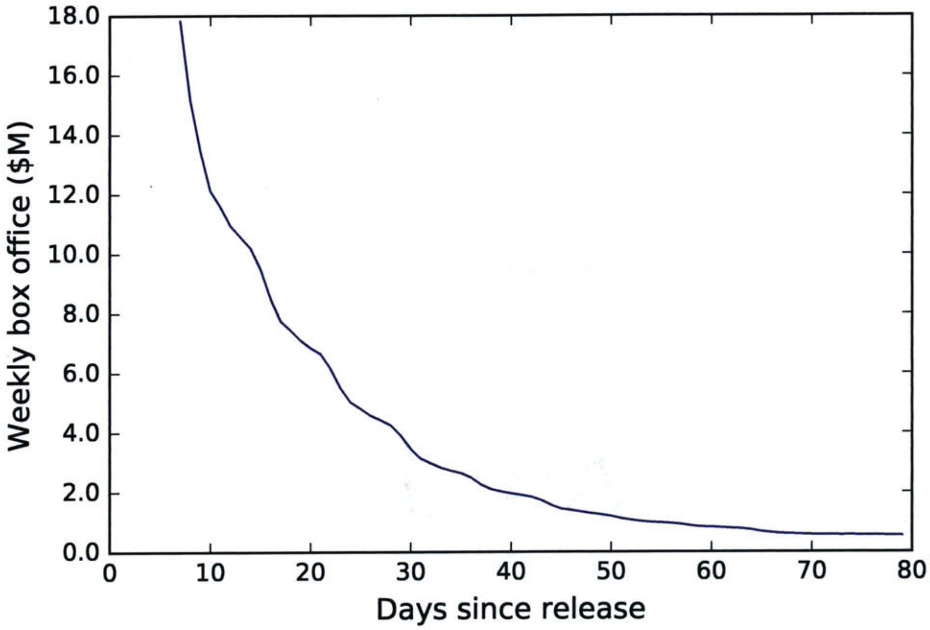


Figure C-10: Daily IMDb variance as a function of time (rolling seven-day average), daily data.

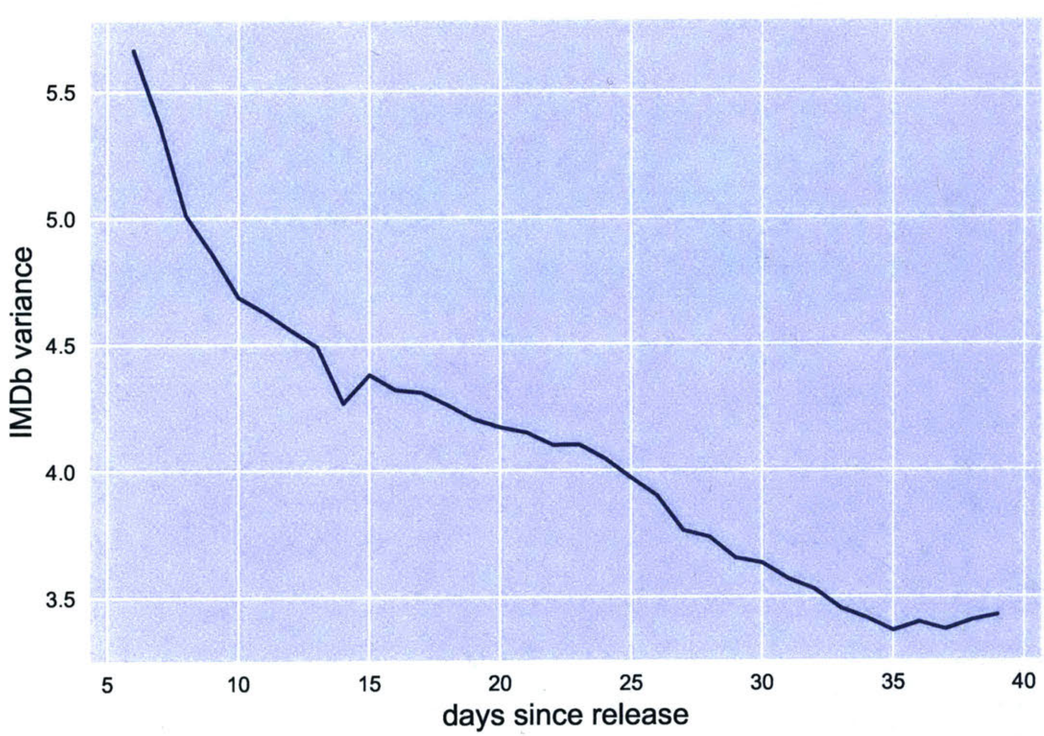
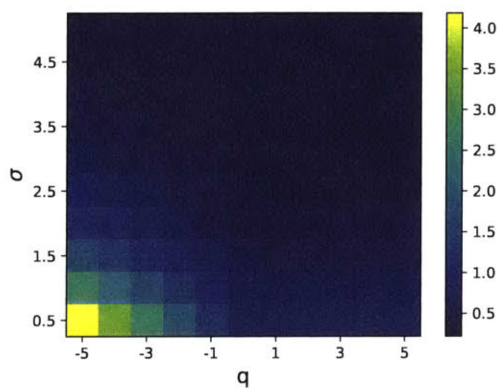


Figure C-11: Simulated selection with two movies and perfect negative correlation.

(a) $\tau_v = 1$



(b) $\tau_v = 10$

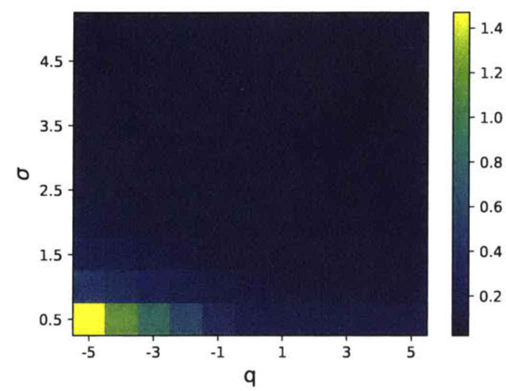
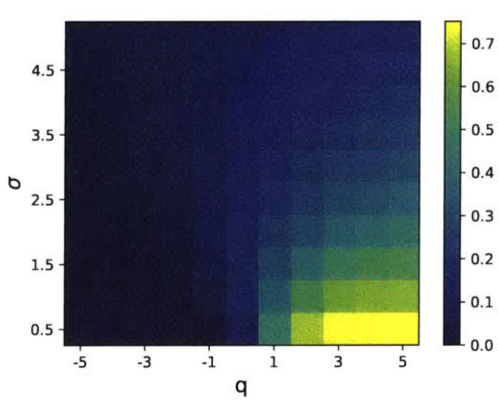
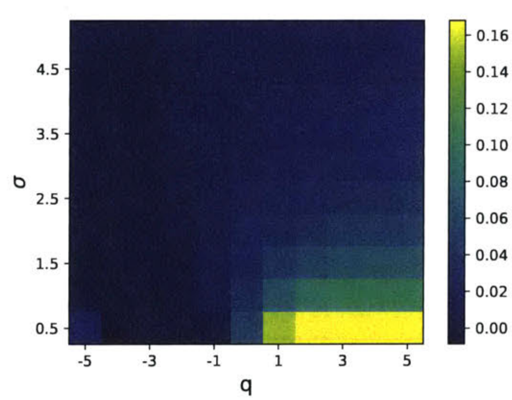


Figure C-12: How selection increases with two movies versus one (perfect negative correlation).

(a) $\tau_v = 1$



(b) $\tau_v = 10$



Bibliography

ANDERSON, S. P., AND R. RENAULT (1999): "Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model," *The RAND Journal of Economics*, pp. 719–735.

BERGEMANN, D., AND J. VÄLIMÄKI (1996): "Learning and strategic pricing," *Econometrica: Journal of the Econometric Society*, pp. 1125–1149.

BERGEMANN, D., AND J. VALIMAKI (2006): "Bandit problems," .

BRANCO, F., M. SUN, AND J. M. VILLAS-BOAS (2012): "Optimal search for product information," *Management Science*, 58(11), 2037–2056.

CHEN, Y., AND S. YAO (2016): "Sequential search with refinement: Model and application with click-stream data," *Management Science*.

CHEVALIER, J. A., AND D. MAYZLIN (2006): "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, 43(3), 345–354.

CHIOU, L. (2008): "The timing of movie releases: Evidence from the home video industry," *International Journal of Industrial Organization*, 26(5), 1059–1073.

CHOI, M., A. Y. DAI, AND K. KIM (2016): "Consumer search and price competition," .

DAI, W., G. JIN, J. LEE, AND M. LUCA (2016): "Optimal Aggregation of Consumer Ratings: An Application to Yelp.com," .

DIAMOND, P. A. (1971): "A model of price adjustment," *Journal of economic theory*, 3(2), 156–168.

DUAN, W., B. GU, AND A. B. WHINSTON (2008): "Do online reviews matter?An empirical investigation of panel data," *Decision support systems*, 45(4), 1007–1016.

DUKES, A., AND L. LIU (2015): "Online shopping intermediaries: The strategic design of search environments," *Management Science*, 62(4), 1064–1077.

EINAV, L. (2007): "Seasonality in the US motion picture industry," *The Rand journal of economics*, 38(1), 127–145.

ELLISON, G., AND S. F. ELLISON (2009): "Search, obfuscation, and price elasticities on the internet," *Econometrica*, 77(2), 427–452.

GODES, D., AND J. C. SILVA (2012): "Sequential and temporal dynamics of online opinion," *Marketing Science*, 31(3), 448–473.

Kamenica, E., and M. Gentzkow (2011): "Bayesian persuasion," *The American Economic Review*, 101(6), 2590–2615.

Ke, T. T., Z.-J. M. Shen, and J. M. Villas-Boas (2016): "Search for information on multiple products," *Management Science*, 62(12), 3576–3603.

Kessler, A. S. (1998): "The value of ignorance," *The Rand Journal of Economics*, pp. 339–354.

Lewis, T. R., and D. E. Sappington (1994): "Supplying information to facilitate price discrimination," *International Economic Review*, pp. 309–327.

Li, X., and L. M. Hitt (2008): "Self-selection and information role of online product reviews," *Information Systems Research*, 19(4), 456–474.

Luca, M., and G. Zervas (2016): "Fake it till you make it: Reputation, competition, and Yelp review fraud," *Management Science*, 62(12), 3412–3427.

Mayzlin, D., Y. Dover, and J. Chevalier (2014): "Promotional reviews: An empirical investigation of online review manipulation," *The American Economic Review*, 104(8), 2421–2455.

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

Riordan, M. H. (1986): "Monopolistic competition with experience goods," *The Quarterly Journal of Economics*, 101(2), 265–279.

Roesler, A.-K., and B. Szentes (2017): "Buyer-optimal learning and monopoly pricing," *American Economic Review*, 107(7), 2072–2080.

Stahl, D. O. (1989): "Oligopolistic pricing with sequential consumer search," *The American Economic Review*, pp. 700–712.

——— (1996): "Oligopolistic pricing with heterogeneous consumer search," *International Journal of Industrial Organization*, 14(2), 243–268.

Wolinsky, A. (1983): "Prices as signals of product quality," *The Review of Economic Studies*, 50(4), 647–658.

——— (1984): "Product differentiation with imperfect information," *The Review of Economic Studies*, 51(1), 53–61.

——— (1986): "True monopolistic competition as a result of imperfect information," *The Quarterly Journal of Economics*, 101(3), 493–511.